ANALYSIS OF THE MECHANISM OF MAINTENANCE OF THE H3K27 TRIMETHYLATION MARK USING A NOVEL CHROMATIN TARGETING SYSTEM

by

SARAH ISABELLE MARIE LEPAGE

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Medical Genetics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

April 2012

Abstract

Chromatin replication during cell division must be accurately orchestrated to ensure genetic and epigenetic information is transmitted to cell progeny. The basic unit of chromatin is the nucleosome, which consists of 147 base pairs of DNA wrapped around a histone octamer composed of an $(H3\text{-}H4)_2$ tetramer and two H2A-H2B dimers. The covalent addition of various chemical groups to amino acid residues on the N-terminus of these core histones is associated with both the repression and activation of genes, repetitive sequences, and non-coding RNAs. Trimethylation of lysine 27 (K27me3) on histone H3 is associated with transcriptional repression, and is found widespread throughout facultative heterochromatin in the genome. This mark is deposited by EZH2, a member of the PRC2 complex. The PRC2 complex is one of two Polycomb group complexes, and is responsible for trimethylating H3K27 at target regions. The second member of the Polycomb group, PRC1, is recruited by regions marked by H3K27me3, and can subsequently inhibit transcription by ubiquitinating H2A to prevent RNA polymerase II elongation.

In association with DNA replication, parental histones disassociate and are segregated to nascent DNA, thereby diluting the number of parental nucleosomes incorporated on the leading strand and lagging strand. Newly synthesized histones assemble onto the newly formed chromatin to replace the disassociated parental histones. Such newly synthesized histones do not contain the same tail modifications as their adjacent parental histones. Therefore, in order to maintain the chromatin and transcription states of specific loci, post-translational histone modifications must be recapitulated after each cell division. However, the molecular basis of such "memory" remains obscure.

Using a Gal4 DNA binding domain (Gal4DBD) fused to EZH2 coupled with FLP/FRT-based deletion of a gal4 binding site cassette, I provide evidence that once established, the maintenance of H3K27me3 through multiple cell divisions does not require the presence of the DNA binding sites necessary for the initial deposition of this mark. These results suggest that the presence of specific histone marks may be sufficient to promote reiterative deposition of the same mark on nascent histones in association with DNA replication, presumably via "reader" domains within the KMTase or associated proteins.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Ash	-	Absent, small or homeotic
Bmi1	-	B lymphoma Mo-MLV insertion region 1 homolog
BRD4	-	Bromodomain containing protein 4
BRWD1	-	Bromodomain and WD40-domain containing protein
CAF1	-	Chromatin assembly factor 1
CBX	-	Chromobox
CBX	-	Chromobox homolog
CDYL	-	Chromodomain protein, Y-like
CpG	-	cytosine guanine dinucleotide
DMEM	-	Dulbecco's Modified Eagle Medium
Dnmt	-	DNA methyltransferase
EED	-	Embryonic ectoderm development
EZH2	-	Enhancer of zeste 2
FRT	-	Flp recombinase target site
Gal4	-	Galactose-induced gene 4
GFP	-	Green fluorescent protein
HDAC	-	Histone deacetylase
Hox	-	Homeobox
HP1	-	Heterochromatin protein 1
Jarid2	-	Jumonji, AT rich interactive domain 2
K	-	Lysine
KMT	-	Lysine methyltransferase
MEL	-	Mouse erythroleukemia
MLL	-	Mixed-lineage leukemia
MPP8	-	M-phase phosphoprotein 8
NURF	-	Nucleosome remodeling factor

ORC	-	Origin replication complex
PBS	-	Phosphate buffered saline
PcG	-	Polycomb Group
PCNA	-	Proliferating cell nuclear antigen
PHC	-	Polyhomeotic homolog
PRC1	-	Polycomb repressive complex 1
PRC2	-	Polycomb repressive complex 2
PRE	-	Polycomb response element
qPCR	-	Quantitative polymerase chain reaction
RbAp46/48	-	Retinoblastoma binding protein
RING	-	Ring finger protein
RMCE	-	Recombinase mediated cassette exchange
Setdb1	-	Set domain, bifurcated 1
Su(var)3-9	-	Suppressor of variegation 3-9
Suz12	-	Suppressor of zeste 12 homolog
SWI/SNF	-	SWItch/Sucrose NonFermentable
TrxG	-	Trithorax Group
UAS	-	Upstream activation sequence

Acknowledgements

Above all, I would like to thank my research team for their unfailing guidance, support, and cooperation throughout the course of this thesis. To my advisor, Dr. Matthew Lorincz, for his infectious scientific enthusiasm and patient mentorship. To my committee members, Dr. Carolyn Brown and Dr. Hugh Brock, for their formal and informal advice on my project and science in general. To the members of the Lorincz Lab: Irina, Preeti, Mehdi, Danny, Peter, and Sheng; I could not have done this without you. I would also like to thank Dr. Jacob Hodgson, for always helping me see the bigger picture.

I would also like to thank my friends and family, without whom I would not be where I am today. Mom, I am forever grateful for your continued support and belief in me. Jordan, your patience and warmth are what kept me motivated to push on despite many hardships. To everyone in Southlands, you provided me with a healthy dose of well-needed horse therapy after long days in the lab. I would also like to thank my MedGen peers, some for getting me through first year classes and some for their inspirational bits of advice by the water cooler or in the office.

Finally, I would like to thank my external colleagues for their kind words of advice and their donations of materials for my project. To Dr. Margaret Rush, who provided the parental cell lines used in this thesis and got my feet wet in the lab. To the labs of Dr. Carolyn Brown and Dr. Louis Lefebvre, for extensive use of their equipment and reagents. I would also like to thank Dr. Ann Rose, whose generous expertise helped me narrow my focus and bring my thesis outline from the abstract to reality. To all, I am extraordinarily thankful.

Dedication

To June and Jordan.

1 Introduction

1.1 Overview of epigenetic modifications and transcription

"Epigenetics" refers to the heritable changes seen in transcription that do not involve alterations to the underlying DNA sequence. Epigenetics has become a highly active field of research over recent years, mainly due to the major role it plays in development and disease. Epigenetic modifications to DNA and to histones influence transcription by either directly interfering with the binding of RNA polymerase II, or by indirectly recruiting other repressive factors. This wide variety of modifications can be dynamic or very stable, and may work both in concert or exclusive of one another to maintain or alter transcription status within a given cell type at a given time in development. These modifications may also arise stochastically in mature mammals, either due to chance or environmental influence (Jaenisch and Bird, 2003).

1.2 Chromatin and histone tail modifications

In eukaryotes, DNA is packaged into chromatin, the basic building block of which is the nucleosome core particle that is comprised of 147 bp of DNA wrapped around an octamer containing four core histones, H2A, H2B, H3 and H4. Once thought of as mere packaging units for DNA, histones are now recognized as dynamic, integral components of gene transcription. Covalent modifications, including methylation, acetylation, phosphorylation, sumoylation, and ADP-ribosylation decorate residues of each of the four core histones, predominantly on their amino terminal tails (***Table 1***). These modifications are the basis for the dynamic regulation of chromatin and gene expression.

Table 1. Histone modifications and their associated transcription status.

Type of Modification	Histone and Residue	Association
Acetylation	H2AK5	Transcriptional activation
	H2BK5 H2BK12 H2BK15 H2BK20	Transcriptional activation
	H3K9 H3K14 H3K18 H3K23 H3K27	Transcriptional activation
	H4K5 H4K8 H4K12 H4K16	Transcriptional activation
Methylation	H3K4 H3R8 H3K9 H3R17 H3K27 H3K36 H3K79	Transcriptional activation Transcriptional silencing Transcriptional silencing Transcriptional activation Transcriptional silencing Transcriptional activation Transcriptional activation
	H4R3 H4K20 H4K59	Activation/silencing* Transcriptional silencing Transcriptional silencing
Phosphorylation	H2AS1	Transcriptional activation
	H2BS14	DNA repair
	H3T3 H3S10 H3T11 H3S28	Mitosis Transcriptional activation Mitosis Transcriptional activation
	H4S1	DNA repair
Ubiquitinylation	H2AK119	Transcriptional silencing
	H2BK120	Meiosis

*Depending on which enzyme modifies the residue. Adapted from Cell Signaling Technology, Inc. (retrieved June 2011) Signaling Pathways: Histone Modification Table, copyright 1999-2011.

There are two main theories to explain how these modifications influence transcription: the charge neutralization hypothesis and the histone code hypothesis.

1.2.1 Charge neutralization hypothesis

The charge neutralization hypothesis suggests that histone modifications, in particular acetylation, directly impact chromatin structure. Since acetylation on histone lysines neutralizes their charge, it results in the reduction of interaction of DNA with histones, thus making the DNA more accessible to the transcription machinery. Without acetylation, histones have a net positive charge, while DNA possesses a negative charge, thereby resulting in an electrostatic interaction between DNA and histones. Upon acetylation of the lysines, the DNA loosens from the histones, providing a more open chromatin structure (Hong *et al*., 1993; Wade *et al*., 1997).

1.2.2 The "histone code" hypothesis

Although the charge neutralization hypothesis may explain the transcriptional effect seen with histone acetylation, it does not explain the full myriad of effects observed by all histone modifications. The histone code hypothesis suggests that post-translational modifications work in a combinatorial and/or sequential fashion that can be read by proteins and interpreted into a variety of downstream effects (Strahl and Allis, 2000). This code is now thought of as more of a "histone language" in that any given modification does not necessarily result in a given function; it acts in concert with other modifications to dictate transitions between transcriptionally active and silent states. The histone code therefore adds several layers of complexity and nuance to the underlying genetic code (Jenuwein and Allis, 2001).

If the histone code hypothesis is true, there should be proteins that bind to histone modifications that effect downstream functions. One such protein is HP1 (heterochromatin protein 1), a conserved protein capable of binding specifically to methylated H3K9. This binding is required

to maintain heterochromatin in H3K9 methylated regions in mice (Bannister *et al*., 2001; Maison and Almouzni, 2001; Maison and Almouzni, 2004). Several other proteins have been shown to bind to the H3K9me mark in murine cells, such as CDYL, CDYL2, CBX2, CBX4, CBX7, and MPP8 (Kokura *et al*., 2010; Bua *et al*., 2009; Liu *et al*., 2010; Kim *et al*., 2006; Mulligan *et al*., 2008), however the significance of their downstream effects remain unknown (Maksakova *et al*., 2011). These proteins have in common a chromodomain, capable of recognizing and binding to methylated lysines (Bannister *et al*., 2000). Bromodomains, on the other hand, bind to acetylated lysines (Dhalluin *et al.*, 1999). In mice, the bromodomain and WD-repeat-containing protein BRWD1 is required for normal spermiogenesis and the oocyte to embryo transition (Phillips *et al*., 2007). A mutation in BRWD1 leads to phenotypically normal but infertile mice. In humans, BRD4's downstream effect is important for the immune system; activation of BRD4 has been implicated in the management and prediction of the reduction of tumour growth in mice (LeRoy *et al*., 2008; Crawford *et al*., 2008). Furthermore, LRWD1, a member of the origin replication complex (ORC), has been shown to interact with H3K27me3, H3K9me3, and H4K20me3 using mass spectrometry; the ORC is key in the initiation of replication, and was known previously to localize to heterochromatic regions (Vermeulen *et al*., 2010; Prasanth *et al*., 2004). Taken together, these studies indicate that there are proteins that can mediate downstream effects upon binding to histone modifications, thus supporting the existence of a histone code.

1.3 PcG and Trithorax Group proteins

Over the past 20 years, genetic screens for *su*ppressors of position effect *var*iegation [Su(*var*)] (the phenomenon by which a gene is regulated by its chromatin environment) in *Drosophila* revealed that over 100 genes play important roles in the establishment and maintenance of

heterochromatin (Weiler and Wakimoto, 1995; Sinclair *et al*., 1983). Many of these genes are conserved in mammals, such as HP1, Su(var)3-9, as well as the developmentally vital families of proteins, the Polycomb Group (PcG) and the Trithorax Group (TrxG) (Goldberg *et al*., 2007). PcG and Trithorax Group proteins are required for the correct spatial and temporal expression of the homeotic (*Hox*) genes during development. *Hox* genes are responsible for patterning the vertebrate body axis (Kmita and Duboule, 2003). PcG and Trx mutant phenotypes show characteristic homeotic transformation and misexpression of the *Hox* genes (Kennison, J.A., 1995) in flies, mice and humans and these proteins have been implicated in several important cellular processes, such as cell cycle regulation, X-inactivation, cancer progression, and stem cell lineage and differentiation pathways (Oktaba *et al*., 2008; Leung *et al.,* 2004; Sparmann and van Lohuizen, 2006; Wang *et al.,* 2001; Pasini *et al*., 2007). Due to their importance in development and the characterization of their catalytic activities and target genes, PcG and TrxG proteins are exemplary in the study of how epigenetic enzymes influence transcription.

The PcG family contains two main protein complexes: polycomb repressive complex 2 (PRC2) and PRC1, and is associated with repressed transcription. PRC2 is responsible for catalyzing trimethylation of lysine 27 on histone H3. Knockouts of PRC2 subunits in mammals cause early embryonic lethality (O'Carroll *et al*., 2001; Pasini *et al.*, 2004). The core subunits include: EZH2, EED, and Suz12, all of which must associate for PRC2 methyltransferase (KMTase) activity (Pasini *et al*., 2004; Cao *et al*., 2002). EZH2 contains a SET domain, which has H3K27me3 catalytic activity, while Suz12 and EED both act to stimulate EZH2 (Kuzmichev *et al*., 2002). The PRC2 complex also contains the histone binding protein(s) RbAp46/RbAp48, as well as the histone deacetylases HDAC1 and HDAC2, however these are not required for stimulation of

KMTase activity (Cao *et al*., 2002). PRC1 on the other hand, is thought to act downstream of PRC2 at polycomb target genes and is responsible for the monoubiquitylation of lysine 119 of histone H2A. Mammalian PRC1 is comprised of 4 subunits: CBX, which contains a chromodomain that can bind to H3K27me3, PHC, which is required for PRC1 mediated repression, Bmi1, which is required for ubiquitinase activity, and RING1, whose Ring domain can ubiquitinate H2AK119 (Bernstein *et al*., 2006; Kim *et al*., 2005; Cao *et al*., 2005; Wang *et al*., 2004) (***Table 2***). The H2AK119ub1 mark is thought to repress transcription by inducing chromatin compaction and/or inhibiting transcriptional elongation (Francis *et al*., 2004; Zhou *et al*., 2008).

Table 2. The subunits of PRC1 and PRC2.

Subunit (mouse & human)	Functional Domain
PRC1 • CBX • PHC • BMI1 • RING1A/B	PRC1 • Chromodomain • SAM domain and zinc-finger domain • RING-finger domain • RING-finger domain
PRC2 • EED • EZH2 • SUZ12 • RbAp46/48	PRC2 • WD40 domain • SET domain • Zinc-finger domain • WD40 domain

Trithorax group proteins maintain expression of Hox genes, a group of genes responsible for the structure and orientation of the body, in cells where they must stay active (Francis and Kingston, 2001). TrxG proteins are a somewhat heterogeneous group, but they are characterized by antagonistic mechanistic properties to the PcG proteins. The *Drosophila* Ash1 and vertebrate MLL TrxG complexes are composed of a SET domain and can methylate lysine 4 on histone H3, a mark associated with transcriptional activation. Another class of TrxG proteins comprises

protein components of ATP-dependent chromatin remodeling complexes like the SWI/SNF or the NURF complexes, and includes subunits that can "read" the methylation mark established by the SET domain containing proteins, and manipulate the chromatin structure accordingly (Shuettengruber *et al*., 2007).

1.4 Recruitment of PcG and TrxG

Recruitment of the PcG and TrxG proteins has been a subject of intense study within the past decade. Knowing how, when and where these complexes are recruited to chromatin is crucial in furthering our understanding of epigenetic processes and development. Recent genome-wide studies in flies, mice and humans reveal that PcG proteins regulate the transcription of thousands of genes, many of which are involved in cell fate decisions (Schuettengruber and Cavalli, 2009). PcG proteins therefore do not bind to the same target in each cell type; rather, a complex network of polycomb recruitment factors must act in concert to ensure the correct spatial and temporal expression of genes. In *Drosophila*, PcG proteins are recruited to specific DNA elements called Polycomb Response Elements (PRE), that stretch along several hundred base pairs (Ringrose and Paro, 2007). These elements contain DNA-binding consensus sites for several transcription factors, some of which are required as mediators for PcG/PRE binding; these factors include the GAGA factor, PHO, PHO-like, and DSP1 (Muller and Kassis, 2006). Since PHO and PHO-like are conserved in mammalian cells, the search began for the mammalian PRE and its involvement in polycomb recruitment. Sing and colleagues uncovered the first vertebrate PRE, named PRE-kr, which is required for the repression of the mouse *MafB* gene (also known as kreisler) (Sing *et al.*, 2009). PRE-kr can recruit PRC2 and PRC1, with different sequence requirements; interestingly, PRC1 has a stronger affinity for the PRE than PRC2. This provided evidence that PRC1 can

preclude PRC2 and may act independently at some sites (Schoeftner *et al*., 2006). YY1, the mammalian ortholog of PHO, has also been implicated in the recruitment of PcG proteins. YY1 recruits EZH2 to muscle-specific genes in myoblasts, which then disassociate upon differentiation (Caretti *et al*., 2004). More recently, another potential PRE was discovered between the *HOXD11* and *HOXD12* loci; this 1.8 kb element contains YY1 binding sites, and can recruit PRC1 and PRC2 in human embryonic cells (Woo *et al*., 2010). However, while PREs exist in *Drosophila* and work to recruit polycomb proteins, this mechanism does not appear to be the general mode of targeting in mammals.

A flurry of recent studies has implicated non-coding RNAs (ncRNAs) in the recruitment of PcG proteins. Rinn *et al*. demonstrated that PRC2 could be recruited by a 2.2 kb long ncRNA called HOTAIR, which is transcribed from the *HOXC* locus and is required for the repression of the *HOXD* locus, 40 kilobases away (Rinn *et al*., 2007). A ncRNA coded from the *Kcnq1* locus is involved in imprinting; it can associate with EZH2, SUZ12 and the H3K9 KMTase G9a in the mouse placenta and cause repression of genes at the *Kcnq1* domain. This ncRNA, *Kcnq1ot1*, has a tissue-specific interaction with PRC2 and G9a; the association is detectable in the placenta but not in the liver (Pandey *et al*., 2008). Similarly, a 1.6 kb ncRNA found within the *Xist* ncRNA, called *RepA*, has been shown to associate with PRC2 and recruit it to the X chromosome, initiating, but not maintaining, X chromosome inactivation (Zhao *et al*., 2008). More recently, a class of ncRNAs termed large intergenic non-coding RNAs (lincRNAs) has been implicated in a number of biological processes, including cell cycle regulation and maintenance of pluripotency. A subset of these lincRNAs associates with PRC2, and their subsequent knockdown results in the upregulation of PcG target genes (Guttman *et al*., 2009; Khalil *et al*., 2009). This proposed

mode of recruitment (in *cis* only) couples PRC2 recruitment with transcription, and therefore assumes that the transcriptional machinery can access the polycomb target genes.

A number of PRC2-interacting proteins have been postulated to play a role in recruiting polycomb proteins to their targets. In *Drosophila*, the polycomb-like (Pcl) protein forms a subcomplex with PRC2 and may play a role in modulating PRC2 activity; *Pcl-/-* flies demonstrate reduced H3K27me3 levels at a subset of PcG target genes (Nekrasov *et al*., 2007). The human Pcl homolog, PHF1, also interacts with PRC2 and influences its activity, but it remains unclear whether H3K27me3 levels increase or decrease upon PHF1 knockdown, or how it affects mammalian *HOX* gene expression (Sarma *et al*., 2008; Cao *et al*., 2008).
The transcription factor SNAIL1 can interact with PRC2, recruiting it to the *CHD1* promoter (Herranz *et al*., 2008). The downregulation of *CHD1*, a tumour suppressor gene, is thought to contribute to cancer progression and/or metastasis. The recruitment of PRC2 to *CHD1* and its subsequent repression is an example of how deregulation of polycomb components can play an important role in the progression of a myriad of cancers.

Several recent studies have reported the interaction between the Jumonji C-containing protein JARID2 and PRC2. JARID2 is localized to approximately 90% of all PcG target genes, and binding of PRC2 to its targets is greatly reduced upon downregulation of JARID2 (Peng *et al*., 2009; Pasini *et al*., 2010). JARID2 has also been shown to be responsible for the subsequent recruitment of PRC1 and the Ser-5 phosphorylated form of RNA polymerase II (Landeira *et al*., 2010). JARID2 has an ARID domain as well as a zinc finger domain, both of which have the capacity to bind DNA, providing evidence that this protein may link PRC2 to DNA. Interestingly,

H3K27me3 levels do not drastically change upon JARID2 knockdown, and the phenotype seen in *Jarid-/-* embryos is not as severe as what is seen in embryos lacking PRC2 components. Therefore, while JARID2 appears to be an important novel PRC2 subunit and may even work to recruit PRC2 to its targets, other factors are required to maintain H3K27me3 levels in the cell. Several proposed modes of recruitment have been investigated; however none are able to completely abolish PRC2 targeting upon the downregulation of the recruitment protein used. A general model of PRC2 recruitment involves a combination of factors working together or sequentially to repress genes. Catalyzing the addition of H3K27me3 on non-PcG targets or ectopically expressed genes, which is useful in examining the repressive effects of newly deposited H3K27me3, must therefore be done via artificial recruitment of PRC2.

1.5 Formation of chromatin following replication

Replication coupled chromatin duplication involves the process of chromatin disassembly ahead of the replication fork, and chromatin reorganization behind the fork. Upon formation of the fork and passage of DNA polymerase, the histones associated with the parental strand become disrupted, and then likely re-associate with the nascent daughter strands in a random manner (Krude and Knippers, 1991). At the same time, canonical histones are synthesized to fill in the gaps amongst the parental histones (Marzluff *et al*., 2008). Production of canonical histones is a tightly regulated process, ensuring the appropriate "dosage" of these structural proteins; failure to meet demand for new histones causes irreversible growth arrest in yeast, and excess assembly of histones results in impaired S phase progression in humans (Kim *et al*., 1988; Groth *et al.*, 2007). This suggests that proper restoration of chromatin after DNA replication is vital in maintaining both genetic and epigenetic stability.

1.6 Maintenance of histone modifications

Maintenance of epigenetic information is crucial in both developing embryos and differentiated tissues; as mentioned above, deletion of any of the core PRC2 components results in early embryonic lethality. DNA replication during S phase raises a significant question: once DNA is duplicated, how are epigenetic modifications re-established on nascent chromatin?

"Epigenetic memory" is described as the conservation of epigenetic modifications upon cell division. As an example, induced pluripotent stem cells retain epigenetic signatures of their somatic tissue of origin (Kim *et al*., 2010). Establishment of the pattern of epigenetic marks on nascent chromatin may be guided by the pre-existing modifications on the parental histones. This method of histone mark propagation would be ideal in the context of replication coupled chromatin assembly, as modifications present ahead of the replication fork could be swiftly re-established onto newly synthesized chromatin. Therefore, the mechanism by which the parental nucleosomes are reassembled onto nascent DNA likely has a major impact on how the cell propagates post-translational modifications on histones upon DNA replication. However, the actual reassembly mechanism and the subsequent recapitulation of histone modifications are unclear (Annunziato, 2005). There currently exist two models by which parental histones are segregated amongst daughter strands (***Figure 1***).

1.6.1 Random segregation model of nucleosome assembly

The random segregation model was first proposed by Jackson and Chalkley, who conducted a series of radiolabelled histone experiments which revealed that intact parental H3-H4 tetramers associate randomly with either daughter strand behind the replication fork (Jackson and Chalkley, 1985). Subsequent studies suggest that parental histone octamers split into H2A-H2B dimers and

H3-H4 tetramers upon passage of the replication fork (Jackson, 1990). These studies also suggested that recycling of the H3 and H4 histones was more efficient than the recycling of H2A and H2B; the former containing most of the transcriptionally significant epigenetic modifications (Annunziato, 2005). This model is consistent with the observation that H2A and H2B have a much higher turnover rate in the cell than H3 and H4 (Kimura and Cook, 2001). However, this model is apparently incompatible with the notion that epigenetic information at any given domain is inherited; random assembly of whole tetramers along nascent DNA implies that the histone modifications at a particular locus can change upon each cell division. Modifications that are present in only a few nucleosomes would likely be lost by dilution after a few rounds of DNA replication (Margueron and Reinberg, 2010).

1.6.2 Semi-conservative model of nucleosome formation

An alternative possibility is that the pre-existing marks on parental histones can influence the deposition of marks on newly incorporated histones within the same nucleosome. The semi-conservative model of histone segregation suggests that the H3-H4 histone tetramer splits, followed by association of the parental H3-H4 dimer with a newly synthesized H3-H4 dimer, thereby producing nucleosomes where half of the H3-H4 tetramer possesses parental post-translational modifications. Several lines of evidence support this model. Tagami *et al*. (2004) discovered through biochemical purification that H3 and H4 (in complex with histone chaperones) exist as heterodimers, suggesting that they may be incorporated into chromatin as such, rather than tetrameric units. Benson *et al.* (2006) showed that cytosolic predeposition complexes purified from FLAG-H4 expressing cells contain H3/H4 dimers, not tetramers. These studies raised questions as to whether this tetramer splitting did in fact result in the association of

old and new H3/H4 dimers after the replication fork, and whether this occurred at distinct genomic locations or if it was a global phenomenon. Assembly of nucleosomes containing parental and newly synthesized H3/H4 dimers is an ideal model to explain the inheritance of post-translational histone modifications; the histone marks on the parental H3/H4 serves as a template for the newly synthesized histones, which are devoid of the pattern of modifications present on the parental histone. However, Katan-Khaykovich and Struhl (2011) showed that in yeast, tetramer splitting is a relatively infrequent event, and appears to occur exclusively at actively transcribing genes. Intriguingly, in mammals, newly synthesized H3.1 (canonical H3) does not form mixed tetramers with preexisting H3 molecules (Xu *et al*, 2010). Rather the histone variant H3.3, which is incorporated into chromatin at active genes, associates with parental H3 in a subset of genomic locations. Histone exchange with H3.3 is independent of DNA replication, even at active genes. However, this recent evidence indicates that the tetramer split model is not the general mode of assembly of new histones; rather, parental H3/H4 tetramers remain intact, and are segregated randomly to each daughter strand, perhaps with the exception of chromatin at active genes.

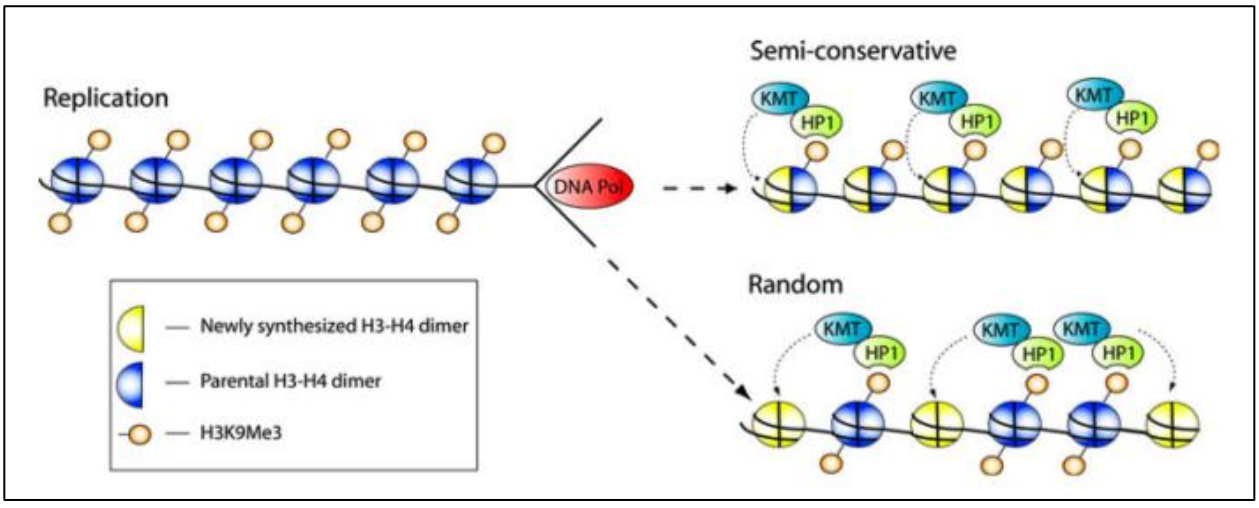

Figure 1. The semi-conservative model and the random model of histone segregation after DNA replication. KMT: lysine methyltransferase. (Figure 1, Blomen and Boonstra, 2011)

1.7 Mechanisms of "maintenance" of histone marks

Regardless of how old and new histones are organized into chromatin after the replication fork, there is likely to be a precise mechanism in place for the recapitulation of the histone marks present on the parent histones to maintain the transcription status of any given genomic locus. Newly synthesized histones are not completely devoid of marks prior to its association with chromatin. Prior to deposition, newly synthesized histones are globally diacetylated on H4 at the lysine 5 and lysine 12 residues and locally monomethylated at lysine 9 in H3 in humans, and are monoacetylated at lysine 56 in H3 and locally mono- and dimethylated at lysine 79 in H3 in yeast (Campos *et al*., 2010; Jasencakova and Groth, 2010). These newly synthesized H3/H4 histones are found in dimers, rather than tetramers, and are mostly associated with histone chaperones (Tagami *et al*., 2004; Benson *et al*., 2006; Chang *et al*., 1997). Upon deposition into chromatin, the majority of these marks are removed or further processed to resemble the histone modifications present before DNA replication. How precisely this is accomplished remains a subject of intense study; it is likely that the mechanism involved differs among types of modifications and their context. Five or more varieties of inheritance can be theorized: inheritance via genetic and DNA binding factors, replication-dependent marks, through cross-talk with other epigenetic modifications, through spatio-temporal regulation during replication, and copying marks from parental histones, using them as a blueprint (Jasencakova and Groth, 2010).

The replication-coupled method of histone modification establishment would ensure the fastest restoration of marks after the replication fork. PCNA interacts with a variety of factors involved in the maintenance of chromatin structure, such as nucleosome assembly factors,

histone deacetylases, DNA methyltransferases, nucleosome remodeling factors, and histone methyltransferases (Groth *et al*., 2007). This is an ideal model for maintenance of global epigenetic modifications, but since PCNA is found at all replication forks, more specificity is required for the majority of histone modifications. Studies done by Scharf *et al*. (2009) indicate that acetylation is very quickly processed during or after replication to resemble the parental pattern. Histone lysine methylation, on the other hand, appears to require a stepwise process to re-establish the mark, starting with an almost immediate deposition of monomethylation at a subset of lysine residues, which become further modified into di- and trimethylated lysines throughout the progression of the cell cycle. Since higher methylated states are thought to aid in condensing chromatin and also recruit structural proteins such as HP1 and Polycomb, the slow re-establishment of di- and trimethylated lysine residues may prevent premature condensation of chromatin.

Crosstalk between epigenetic marks is also an appealing model for the propagation of histone modifications. The presence of one mark may promote the deposition of another, or may prevent its establishment. The first examples of this complex phenomenon were discovered in *Saccharomyces cerevisiae*, where phosphorylation of serine 10 on H3 promotes the acetylation of lysine 14 during gene activation (Walter *et al*., 2008). Methylation of H3K36 in yeast results in the recruitment of the deacetylase Rpd3S, which deacetylates histones H3 and H4 after passage of RNA polymerase (Lee and Shilatifard, 2007). Histone crosstalk among different histones has also been documented; for example, H3K4 and H3K79 methylation are regulated by H2BK123 monoubiquitination (Weake and Workman, 2008). Finally, as mentioned above, the monoubiquitination of H2AK119 by PRC1 is usually a direct result of the presence of H3K27me3. This phenomenon can only explain modifications that occur as a result of the

presence of another modification, and does not explain how individual, independent marks are propagated.

Using the epigenetic marks present on parental histones as a template for reconstitution of the marks on newly synthesized histones seems like the ideal method of maintenance. Indeed, maintenance of DNA methylation (Jaenisch and Bird, 2003), the most well characterized epigenetic mark, involves "template copying", as described in detail below.

1.7.1 DNA methylation

DNA methylation inhibits gene expression specifically when found in promoter regions. Deposition of this covalent mark is catalyzed by the addition of a methyl group from an S-adenosyl-L-methionine substrate to the 5' carbon on a cytosine base. In mammalian cells, cytosines are methylated predominantly in the context of cytosine-guanine dinucleotides (CpG), giving rise to 5-methylcytosine, also known as the fifth base of DNA (Turek-Plewa and Jagozinski, 2005). This reaction is catalyzed by three DNA methyltransferases in mammalian cells: Dnmt1, Dnmt3a, and Dnmt3b. Dnmt3a and Dnmt3b, the *de novo* DNA methyltransferases, bind to unmethylated double stranded DNA targets, and catalyze the methylation of cytosines regardless of DNA strand. Dnmt1 and its isoforms are the only known maintenance DNA methyltransferases. Dnmt1 performs its function during DNA replication by recognizing hemimethylated DNA and using the methylated parent strand as a template for methylation of the CpG on the complementary nascent strand, thus propagating the DNA methylation pattern, as seen in ***Figure 2*** (Flynn *et al*., 1996).

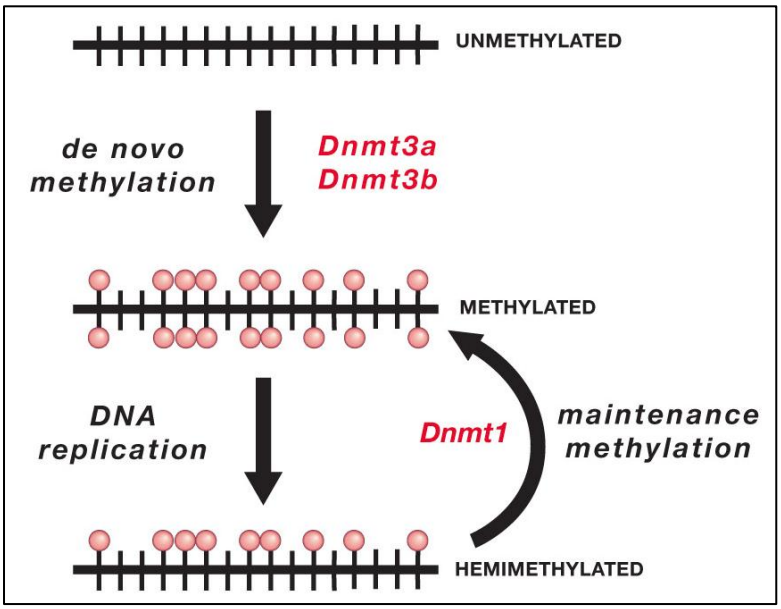

Figure 2. Schematic of *de novo* and maintenance DNA methylation. *De novo* DNA methylation is performed by Dnmt3a/b, and maintenance DNA methylation is performed by Dnmt1.

1.7.2 Maintenance by template copying

It is tempting to postulate that histone marks are maintained in a similar manner to DNA methylation, ie. the mechanism for maintaining the mark is separate from the mechanism for its *de novo* establishment. Indeed, there is convincing evidence that suggests that this mechanism of inheritance is used by a handful of histone marks. HP1 recognizes methylated H3K9 and subsequently induces the "spreading" of this mark at heterochromatic regions on newly synthesized histones in most eukaryotes (Probst *et al*., 2009). By binding to di- and trimethylated H3K9 via its chromodomain, HP1 is thought to indirectly propagate this mark by also interacting with H3K9 methyltransferases, SUV39h1/2 and G9a (Eissenberg and Elgin, 2000; Grewal and Jia, 2007; Melcher *et al*., 2000). In addition, HP1 also binds with CAF-1, a histone chaperone that associates with newly formed H3/H4 dimers and supplies them to newly replicated DNA; this interaction is required for heterochromatin formation and S-phase progression in mice

(Quivy *et al*., 2004; Quivy *et al*., 2008). CAF-1 can also bind SETDB1, another H3K9 KMTase (Loyola *et al*., 2009). In light of these interactions, one could speculate that HP1 is recruited to sites of DNA replication via CAF-1, then upon recognizing the methylation of H3K9 on parental histones redistributed after the replication fork, recruit H3K9 KMTases to methylate the newly synthesized H3 histones. However, experiments conducted in the Lorincz lab reveal that deletion of either HP1a or HP1b does not lead to loss of H3K9me3 in the flanks of endogenous retroviruses, or the activation of these elements in mouse ES cells (Maksakova *et al*., 2011).

Nevertheless, recent studies suggest that the H3K27me3 mark may be inherited in a template-copying manner. PRC2 localizes to sites of DNA replication, as demonstrated by Hansen *et al*., (2008) who used synchronized human fibroblasts to observe the co-localization of EZH2, PCNA and CAF-1. This group also demonstrated that the PRC2 complex can bind to the H3K27me3 peptide, and it was later shown in another study that the EED subunit binds specifically to trimethylated lysines on H3 (Margueron *et al*., 2009). Finally, it was shown that the H3K27me3 mark can be established on an integrated transgene through Gal4-mediated recruitment of EED, and that this mark persists after downregulation by tetracycline washout of the tetracycline-inducible Gal4DBD-EED fusion protein.

The challenge with using the template-copying model of histone mark transmission is that it provides no explanation for the recruitment of histone modifiers, nor does it explain how the spreading of post-translational modifications is halted at the loci's boundaries. The PRC2 studies also fail to describe whether any accessory PRC2 components, such as JARID2 (which has a DNA binding domain) associate with the version of PRC2 that seems to be involved in maintenance. Focusing on the Hansen article specifically, there are several technical issues that complicate interpretation of the results. Firstly, as the epigenetic modification status of the locus

in which the transgene was integrated is unknown, the basal level of H3K27me3 and/or the presence of other histone marks or DNA methylation could affect the persistence of H3K27me3 along the transgene. Secondly, leaky expression of the Gal4DBD-EED fusion protein after tetracycline washout could still lead to recruitment of other PRC2 components and result in maintenance of the H3K27me3 mark. In a paper published by our lab in 2009, a Gal4DBD-EZH2 fusion protein was used in a similar manner to recruit PRC2 to a transgene, and it was noted that even very low expression of Gal4DBD-EZH2 (undetectable by western blot) was sufficient for good enrichment of H3K27me3 on the cassette to which it was targeted (Rush *et al*., 2009). Therefore, although Hansen *et al*. provided evidence in favour of an H3K27me3 maintenance mechanism mediated by PRC2 binding, I set out to establish a system to determine whether the presence of H3K27me3 is indeed sufficient for its own inheritance in which the initial binding sites responsible for recruitment of EZH2 could be efficiently deleted.

1.8 Statement of thesis

Given the interesting dynamics of chromatin and the importance of maintaining transcription states during development and beyond, I sought to study the mechanisms of heritability of repressive chromatin states. While this is a fairly broad topic, garnering interest from diverse fields such as biochemistry, developmental biology and cancer research, I chose to focus on polycomb-mediated silencing of an introduced transgene. As aforementioned, PcG proteins are vital for the regulation of expression of the Hox genes responsible for the structure and orientation of the body, and play an important role in cancer progression. How these proteins are recruited to their target genes in mammals remains unclear, and thus any study on the maintenance of the post-translational histone modifications catalyzed by polycomb must be done

via artificial means. Our lab and several others have successfully utilized the Gal4-UAS system of targeting, which was first developed by Brand and Perrimon (Brand and Perrimon, 1993), as a means to study gene expression. Originating from yeast, the Gal4 transcriptional activation protein binds very specifically to a short upstream activation sequence (UAS), which results in gene activation when placed upstream of a reporter gene. Previously, we generated a translational fusion between the DNA binding domain of the GAL4 transcription factor (referred to as the Gal4DBD below) and EZH2. By introducing a transgene containing the UAS (referred to as "Gal4 sites") into cells expressing the Gal4DBD-EZH2 fusion, we have a means of specifically recruiting EZH2 to the transgene under study. The Gal4DBD-EZH2 fusion protein can form a complex with the other PRC2 subunits, resulting in the trimethylation of H3K27 along the transgene (Rush *et al*., 2009).

Given the previous data suggesting that PRC2 does indeed bind to H3K27me3 and its enzymatic activity is stimulated by this binding, I hypothesized that PRC2 does not require the DNA sequence responsible for its initial recruitment to propagate H3K27me3 onto newly synthesized histones. In order to test this hypothesis, I generated murine cell lines expressing GAL4DBD-EZH2 and a cassette containing Gal4 sites upstream of a *GFP* gene. These Gal4 sites were flanked by Flp recombinase target sites, providing an elegant means of deleting the Gal4 sites after establishment of H3K27me3, thus abolishing the initial recruitment mechanism, and analyzing the effect of removing these sites on the maintenance of the H3K27me3 mark.

2 Materials and Methods

2.1 Cell lines, MEL and ES cell culture

2.1.1 MEL cells

MEL cells were maintained on plates in MEL media (Dulbecco's Modified Eagle Media (DMEM) (Hyclone Cat. SH30022.01), supplemented with 10% fetal bovine serum, 100 units/ml penicillin, 2 mM glutamine, and 0.05 mM streptomycin). Passaging conditions were as follows: Every 48-72 hours, cells were pipetted up and down repeatedly with a P1000 pipette to render culture in a single cell suspension. $1/10^{th}$-$1/15^{th}$ of the single cell suspension was then transferred to a new plate containing warmed media. If the single cells suspension exceeded approximately $1/8^{th}$ of the volume of media in the new plate, additional steps were taken. The volume of cells needed were taken out of the single cell suspension and pipetted into a 15 mL Falcon tube. The cells were centrifuged at 1000rpm for 4 minutes in a Heraeus Labofuge 400 to pellet cells. The media was then aspirated off the pellet, and 500-1000 μl of media was pipetted from the new plate and used to resuspend the pellet. The resuspended cells were then gently pipetted into the new plate containing warmed media. This new plate was then rocked gently in a circular motion to disperse cells across the plate, then placed in a 37^{o}C incubator under 5% CO_2 to grow.

2.1.2 ES cells

ES cells were maintained on gelatinized plates in ES media (Dulbecco's Modified Eagle Media (DMEM) (Hyclone Cat. SH30022.01), supplemented with 15% fetal bovine serum, leukemia inhibitory factor, 0.1mM nonessential amino acids, sodium pyruvate, 20mM HEPES, 0.1mM 2-mercaptoethanol, 100units/ml penicillin, and 0.05mM streptomycin) in the absence of feeder cells. Passaging conditions were as follows: Every 48-72 hours cell media was aspirated from

cell monolayers and 1-5ml phosphate buffered saline (PBS) was added depending on the size of the culture plate. PBS was then aspirated off, and 500-1000 μl of trypsin (Hyclone Cat. SH30042.02) was added to disperse cells. Trypsinization was carried out at room temperature or in the 37^{o}C incubator under 5% CO_2 for 3-4 minutes. Trypsinization was stopped by addition of 1-3 ml of ES media to the plates. Cells were then pipetted repeatedly up and down using a P1000 pipette tip in order to disperse cell clumps. 1/5th-1/15th of this mix was then passaged to a pregelatinized tissue culture plate with ES media. This new plate was then rocked gently in a circular motion to disperse cells across the plate, then placed in a 37^{o}C incubator under 5% CO_2 to grow.

Plates were gelatinized by the addition of 1-5mls of sterile 0.1% porcine skin gelatin in water (Sigma Cat. G2500-100G) to a tissue culture plate, followed by incubation at RT for ~10 minutes. Excess gelatin was then aspirated from the plate before addition of ES media.

Table 3. Cell lines used in this thesis.

Line name	Source
RL5 MEL Gal4DBD	(Rush et al., 2008)
RL5 MEL Gal4DBD-EZH2	(Rush et al., 2008)
RL5 MEL Gal4DBD L1-CMV-GFP-1L	Sarah Lepage
RL5 MEL Gal4DBD-EZH2 L1-CMV-GFP-1L	Sarah Lepage
RL5 MEL Gal4DBD L1-FRTgal4FRT-GFP-1L	Sarah Lepage
RL5 MEL Gal4DBD-EZH2 L1-FRTgal4FRT-GFP-1L	Sarah Lepage
RL5 MEL Gal4DBD L1-FRTgal4FRT-p16-GFP-1L	Sarah Lepage
RL5 MEL Gal4DBD-EZH2 L1-FRTgal4FRT-p16-GFP-1L	Sarah Lepage
HA36	Dirk Schubeler
HA36 Gal4KAP1	Sarah Lepage

2.2 Plasmids and primers

2.2.1 Generation of cassettes

FRTgal4FRT was excised from pBS2FRT.gal49x (see **Table 4** for source) by restriction digest using EcoRV (Fermentas Cat. ER0301), and then cloned into the EcoRV site in L1-GAG-*GFP*-1L to generate L1-FRTgal4FRT-*GFP*-1L. To generate L1-FRTgal4FRT-p16-*GFP*-1L, the FRTgal4FRT fragment was cloned into the EcoRV site in pSL1180, and then cut from this plasmid using HindIII (Fermentas Cat. ER0501) and EcoRI (Fermentas Cat. ER0271). The newly digested FRTgal4FRT fragment was cloned into L1-SD-SA-p16-*GFP*-1L at the HindIII and EcoRI sites to generate L1-FRTgal4FRT-p16-*GFP*-1L.

Table 4. Plasmids used in this thesis.

Plasmid Name	Reference
L1-GAG-GFP-1L	Rush et al., 2008
pBS2FRT.gal49x	Margaret Rush
pSL1180	Amersham Biosciences
L1-p16-GFP-1L	Sarah Lepage
L1-FRTgal4FRT-GFP-1L	Sarah Lepage
L1-FRTgal4FRT-p16-GFP-1L	Sarah Lepage
pCAGGS+Gal4e-IRES-Puro	Louis Lefebvre
L1-CMV-GFP-1L	unknown
pC3-ERHBD-GAL4-KAP1	Sripathy et al., 2006

Table 5. Primers used in this thesis.

Primer Name	Sequence
Gag140R	CGGTCGGTCCAGTTGTTCTTGGTAGGC
243 lox-	CATTAATGCAGCTGGCACGACAGG
Gag45F	CCCTCTCTCCAAGCTCACTTACAGGCTCTC
Chr4_RL5_R	GGACAGGGGATAGTGCATCTGTTTCTTACTAACCTG
GFPdF	AGTCCGCCCTGAGCAAAGA
GFPcF	ACTACAACAGCCACAACGTCTATATCA
GFPcR	GGCGGATCTTGAAGTTCACC
+2783+RL5	TTATCTGCCCACTCACTCTGGAC
+2927-RL5	TGCCATAATGACTTCTGTGTTCG
R5GFP	CAAAGTAGACGGCATCGCAGC
roGFP4	GGTGGTGCAAATCAAAGAAC
CMV_1	CCATTGCATACGTTGTATC
-strCMV3	AGTTATGTAACGCGGAACTC
Gal4DBD+29	ACCGAAGTGCGCCAAGTGTCT
Gal4DBD-117	CTGTCAGATGTGCCCTAGTCAG

2.3 Generation of cell lines

2.3.1 RMCE

L1-FRTgal4FRT-*GFP*-1L, L1-FRTgal4FRT-p16-*GFP*-1L and L1-CMV-*GFP*-1L were introduced into RL5 MEL Gal4DBD and RL5 MEL Gal4DBD-EZH2 by recombinase-mediated cassette exchange (RMCE). RMCE was done as follows using the Lipofectamine™ LTX system of transfection (Invitrogen Cat. 15338-100). 250 ng of plasmid was mixed with 750 ng of CMV-

CRE, and Opti-MEM reduced serum media (Invitrogen Cat. 11058-021) was added to bring the volume up to 100 µl. A mock (no plasmid) transfection was included. 2 µl of Plus reagent was added to the mixture and incubated for 5 minutes at room temperature. 3 µl of LTX Lipofectamine reagent was then added to the mixture and incubated for 30 minutes at room temperature. RL5 MEL Gal4DBD and Gal4DBD-EZH2 cells were plated in a 24 well plate at a density of 100,000 cells per well in 500 µl of complete MEL media and incubated at 37^{o}C under 5% CO_2. After 30 minutes, approximately 100 µl of the DNA:lipid complexes were added dropwise to each well containing cells, and grown in a 37^{o}C incubator under 5% CO_2 for 72 hours. 10 µM ganciclovir was then added to the cells which were observed for cell death for 12 days.

2.3.2 Flow cytometry

After 12 days of ganciclovir selection, RL5 MEL Gal4DBD and Gal4DBD-EZH2 RMCE'd pools were assayed for *GFP* expression by flow cytometry. Approximately 500,000 cells were centrifuged as per the cell culture protocol and resuspended in 300-600 µl PBS, 3% FBS, and 1 µg/ml propidium iodide. Fluorescence was then analyzed on a BD LSR II flow cytometer, gating for live cells first using forward/sidescatter, then propidium iodide, then finally against *GFP* expression (green fluorescence). A mininum of 10,000 events were collected per sample.

2.3.3 Single-cell derived clones

Single-cell derived clones were generated from the RMCE'd pools by limiting dilution. Cells were diluted to 100 cells/10 ml of complete media and 100 µl of this was aliquoted into each well of a 96 well plate. After 7 days, each well containing one single colony was expanded in 500 µl in a 24-well plate. After 5-7 days, cells were further expanded in 1 ml of complete media

in 12-well plates for 5-7 days, then again expanded in 3 ml of complete media in 6-well plates to make enough cells to extract genomic DNA.

2.3.4 Genomic DNA extraction

Approximately 5 x 10^6 cells were centrifuged as described in the cell culture protocol, then resuspended in 200 µl TE buffer and transferred to 1.5 ml Eppendorf tubes. 200 µl of Bradley's reagent (20mM Tris pH 7.5, 4mM EDTA, 20mM NaCl, and 1% SDS) and 0.5 mg/ml Proteinase K was added to the tubes, then they were incubated overnight at 55°C. The next day, 800 µl of ice-cold 3M NaOAc was slowly added to the lysed cells and incubated on ice for 30 minutes. The tubes were then inverted several times to mix then incubated on ice again for 20 minutes. Tubes were vortexed then spun down in a tabletop centrifuge for 5 minutes at 6,000 rpm. The DNA pellet was then washed twice with 500 µl of 70% ethanol, and then allowed to dry for 15-20 minutes at room temperature. The pellet was then resuspended in 40 µl of TE buffer and incubated in a 37°C waterbath for 6 hours to overnight to bring DNA into solution. DNA concentration and purity was measured using a spectrophotometer.

2.4 Validation of cell lines

2.4.1 Orientation screen of RMCE'd cassettes

Single-cell derived clones were assayed for presence and orientation of the integrated cassette by PCR. The reaction was conducted using Taq polymerase and its associated reagents (Fermentas Cat. EP0401). GFPdF and chr4_RL5_R primers were used in this PCR (sequences in **Table 5**). Program was 3 min at 95°C initial denaturing, 30 seconds at 95°C subsequent denaturing, 30 seconds at 59°C annealing, 30 seconds at 72°C extension, and repeated from the second step 35 times. Amplicon size was around 500 bp and was visualized on a 1.5% agarose gel.

2.4.2 Copy number qPCR of GFP

Clones that contained the integrated cassette in the correct orientation were further screened for *GFP* copy number. The qPCR was conducted on a Bio-Rad Opticon 2 thermal cycler with each sample in triplicate using HotStart Taq polymerase (Fermentas Cat. EP0401), under EvaGreen chemistry. GFPcF and GFPcR primers were used in this PCR (sequences in **Table 3**). Program was 5 min at 95^{o}C initial denaturing, 30 seconds at 95^{o}C subsequent denaturing, 30 seconds at 59^{o}C annealing, 30 seconds at 72^{o}C extension, 1 second at 80^{o}C plate read, repeated from second step 40 times.

GFP copy number in RMCE'd clones were compared internally to *β-major* to normalize amount of DNA, and then compared to a control RMCE clone (RL5 L1-LTR-*GFP*-1L 6U, single *GFP* copy). Copy number was calculated by subtracting the C(t) values of the *GFP* amplicon from the C(t) values of the *β-major* amplicon within each clone (including the control), then subtracting the C(t) difference from the clones from the C(t) difference from the control. Calculation of the *GFP* fold enrichment of the samples over the 6U control gave approximate copy number.

2.4.3 Native chromatin immunoprecipitation (ChIP)

Protein A and Protein G agarose beads were prepared by mixing 300 µl of each and washing with IP buffer (10 mM Tris-HCl [pH 8.0], 1% Triton X-100, 0.1% deoxycholate, 0.1% SDS, 90 mM NaCl, 2 mM EDTA) containing 1X PIC (Protease Inhibitor Cocktail) then blocked for 3 hours with 1 ml of IP buffer + PIC, 300 µg of salmon sperm DNA, and 750 µg of BSA. After blocking, beads were washed again and resuspended in IP buffer + PIC.

10 x 10^{6} MEL cells were pelleted as described in the cell culture protocol, then washed once with 10 ml of PBS. Pellets were flash frozen in liquid N_2 then stored at -80°C until needed for ChIP.

Pellets were thawed and resuspended in 250 µl of douncing buffer (10 mM Tris-HCl [pH 7.5], 4 mM $MgCl_2$, 1 mM $CaCl_2$) + PIC, then homogenized using a 1 ml syringe (25 5/8 gauge) and passing through 20 times. Homogenized cells were incubated with 25 units of micrococcal nuclease (Worthington Cat. LS004797) for 5 mins in a 37°C waterbath. Digestion was stopped by adding 10mM EDTA and incubating on ice for 5 mins. 1 ml of hypotonic lysis buffer (0.2mM EDTA, 0.1mM benzamidine, 0.1mM PMSF, and 1.5mM DTT) + PIC was added and cells were incubated on ice for 1 hour. Chromatin was collected by spinning down tubes at 3,000 g for 5 mins at 4°C and transferring supernatant to a new tube.

100 µl of the blocked beads was added to each sample and rotated at 4°C for 2 hours. Beads were removed by centrifugation and 100 µl of chromatin was removed to check size of digested fractions. 2-2.5 x 10^6 cells were divided into 0.5 ml siliconized tubes, one for each IP. A 50% input was removed at this time and stored at -20°C. The volume of each sample was brought up to 325 µl with IP buffer + PIC. H3K27me3 and IgG antibody were added to the appropriate tubes and rotated at 4°C for 1 hour (amounts found in **Table 6**). 20 µl of beads were then added and the tubes were rotated at 4°C overnight.

Beads were pelleted at 4,000 rpm for 2 minutes at 4°C, then washed twice with ChIP wash buffer + PIC (20 mM Tris-HCl [pH 8.0], 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 150 mM NaCl), then once with ChIP final wash buffer + PIC (20 mM Tris-HCl [pH 8.0], 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 500 mM NaCl). DNA was eluted by adding 100 µl of elution buffer (100 mM $NaHCO_3$, 1% SDS) to the beads, and bringing the volume of the input DNA to 200 µl with elution buffer. Samples were incubated at 68°C for 2 hours (with gentle vortexing at half hour intervals) then beads were spun down as described above. A second elution was performed for 5 minutes at 68°C, and then both elutions were pooled into one tube.

DNA was purified using the QiaQuick PCR Purification Kit (Qiagen Cat. 28104) with the following modifications: the second spin after addition of PE buffer was performed for 2 mins. The Elution Buffer (EB) was heated to 42°C before use. The columns were air-dried for 1 minute prior to the addition of EB, and then 60 µl of EB was added and allowed to sit in the columns for 1 minute prior to centrifugation.

Enrichment was assayed by quantitative, realtime PCR in technical triplicates on a Bio-Rad Opticon 2 thermal cycler, with 2µl of sample being used for each reaction.

Bar graphs for ChIP experiments were generated by qPCR on a sample representing a portion of the total input chromatin in qPCR, and then calculating the total amount of chromatin (in nanograms) that was present in each sample prior to IP. This value was used as the denominator for each sample in the bar graphs. The numerator was calculated using the nanogram value of the immunoprecipitated fraction of each sample as determined by qPCR. This final fraction was then converted into a percentage and graphed.

2.4.4 Crosslink ChIP

Protein A and Protein G agarose beads were prepared by mixing 200 µl of each then washing with 1X TE buffer three times. The beads were blocked with 800 µg of salmon sperm DNA by rotating overnight at 4°C. After blocking, the beads were resuspended in TE buffer.

2 x 10^7 MEL cells were harvested and washed as described above, then resuspended in 10 ml of PBS. Cells were fixed using 1.5% formaldehyde and incubated at room temperature for 10 minutes, inverting the tubes intermittently. Glycine solution was added to a final concentration of 0.8M and incubated at room temperature for 5 mins with intermittent inversion, to stop the reaction. Fixed cells were washed twice with PBS.

1 ml of Collection Buffer + 1X PIC (100 mM Tris-HCl [pH 9.4], 100 mM DTT) was added to each sample and incubated for 10 mins on ice, then 10 mins in a 30°C waterbath. Cells were pelleted by centrifugation at 2,000 rpm for 5 minutes at 4˚C. 1 ml of Buffer A + PIC (10 mM EDTA, 0.5 mM EGTA, 10 mM HEPES [pH 6.5], 0.25% Triton X-100) was added, cells were pelleted, and then the supernatant was aspirated off. 1 ml of Buffer B + PIC (1 mM EDTA, 0.5 mM EGTA, 10 mM HEPES [pH 6.5], 200 mM NaCl) was added, cells were pelleted, then the supernatant was aspirated off. 1 ml of lysis buffer (MLB) + PIC (10 mM EDTA, 50 mM Tris-HCl [pH 8.0], 1% SDS) was added and incubated on ice for 20 minutes.

DNA fragments were generated by sonication for 18 minutes, 30 secs on/30 secs off, in a Diagenode Bioruptor. After sonication, DNA was collected by spinning tubes down at 13,000 rpm for 10 mins at 4˚C. 6 million cells per sample were transferred to siliconized tubes, and then diluted 2.5X in IP buffer + PIC (2 mM EDTA, 100 mM NaCl, 20 mM Tris-HCl [pH 8.1], 0.5% Triton X-100). Samples were pre-cleared using 3 µl of salmon sperm DNA, 3 µl IgG antibody, and 90 µl blocked beads, then rotated at 4˚C for 2 hours. Beads were pelleted as described above, and pre-cleared chromatin was divided into 3 0.5 ml tubes, one for each antibody. 10% input was removed at this time, and cross-links were reversed to check size of sonicated DNA fragments (reversal protocol described below). Gal4DBD and IgG antibodies were added (amounts found in **Table 6**). Samples were rotated at 4˚C overnight, then 1 µl of salmon sperm DNA and 40 µl of beads were added the next day, then rotated at 4˚C for 2 hours.

Beads were pelleted as described above, then washed in 1 ml of wash buffer 1 (WB1) + PIC (2 mM EDTA, 10 mM Tris-HCl [pH 8.0], 0.1% SDS, 1% Triton X-100, 150 mM NaCl), rotated at 4˚C for 15 mins, then pelleted. Beads were washed with WB2 + PIC (2 mM EDTA, 20 mM Tris-HCl [pH 8.0], 0.1% SDS, and 250 mM NaCl) and WB3 + PIC (1 mM EDTA, 10 mM Tris-HCl

[pH 8.0], 1% NP-40, 1% deoxycholate, 0.25 M LiCl) in the same way, and then washed twice in TE for 5 mins each.

100 µl of elution buffer (EB) (1% SDS, 0.1 M $NaHCO_3$) was added to the beads, which were then rotated at room temperature for 20 mins. Beads were spun down at 4,000 rpm for 2.5 mins at room temperature. Supernatant was transferred to a safelock tube and the elution was repeated. Input DNA was brought up to 200 µl with EB. 2 µg of RNase A was added to each sample and incubated at 37˚C for 30 mins.

To reverse crosslinks, 8 µl of 5M NaCl was added to each sample and incubated overnight at 65˚C. 0.05 mg/ml Proteinase K, 0.1M EDTA, and 25mM Tris-Cl pH 8.0 were added the next day, then incubated at 55˚C for 2 hours. DNA was purified as described above and eluted in 50 µl of elution buffer.

Table 6. Antibodies used for ChIP.

Antibody	Amount (µg)	Host	Catalog No.
IgG	11	Rabbit	Sigma I8140
H3K27me3	6	Rabbit	Upstate ABE44
Gal4DBD	5	Rabbit	Santa Cruz sc-577

2.4.5 Generation of HA36 Gal4KAP1

The pC3-ERHBD-GAL4-KAP1 plasmid was linearized with SalI (Fermentas ER0641), gel extracted and purified using the QiaQuick Gel Extraction Purification Kit (Qiagen 28704). HA36 cells were grown in 3 M hygromycin for ≥10 days prior to transfection. Transfection was done as follows using the Lipofectamine™ 2000 system of transfection (Invitrogen Cat. 11668-019). The day before transfection, 1 x 10^5 cells were plated in 500 µl of ES media without

antibiotics. Next day, 2 µg of plasmid was diluted in 50 µl of Opti-MEM and mixed. A mock (no plasmid) transfection was included. 3 µl of Lipofectamine was diluted in 50 µl of Opti-MEM and incubated for 5 mins at room temperature. DNA was then mixed with Lipofectamine and incubated for 20 mins at room temperature. After incubation, DNA:lipid complexes were added to the wells containing cells and incubated at 37˚C with 5% CO_2. Media was changed to complete ES media the following day, then 200 µg/ml G418 was added to the cells 48 hours after transfection. Cells were cultured in G418 for 5 days, then single cell-derived clones were generated and genomic DNA was extracted as described above. Clones containing pC3-ERHBD-GAL4-KAP1 were assayed by PCR using the Taq polymerase and its associated reagents. Gal4DBD+29 and Gal4DBD-117 primers were in this PCR (sequences in Table 5). Program was 2 min at 94°C initial denaturing, 30 seconds at 94°C subsequent denaturing, 30 seconds at 59°C annealing, 1 min at 72°C extension, and repeated from the second step 30 times. Amplicon size was around 100 bp and was visualized on a 2% agarose gel. RNA was extracted from clones containing pC3-ERHBD-GAL4-KAP1 using the GenElute™ Mammalian Total RNA Extraction Kit (Sigma Cat. RTN70), and then assayed for expression by qRT-PCR as described above, using the Gal4DBD+29 and Gal4DBD-117 primers.

3 Results

3.1 Experimental design

In order to effectively study the maintenance of the H3K27me3 mark, a few key components needed to be included in the experimental design. H3K27me3 maintenance was studied in MEL cells by integrating transgenes at a known genomic site called the "RL5" locus. This targeted integration was achieved by recombinase mediated cassette exchange (RMCE), a method of site-specific chromosomal integration that involves the exchange of a transgene with a pre-localized selectable marker via the action of Cre recombinase on inverted lox sites (Feng *et al.*, 1999). This provides a superior means of studying epigenetic marks, as the RL5 locus was previously shown to be devoid of H3K27me3 (Rush *et al*., 2009).

Recruitment of PRC2 is achieved by expressing a Gal4DBD-EZH2 fusion protein in the RL5 MEL cell line, which can then bind specifically to gal4 sites found in the integrated transgene. The Gal4DBD-EZH2 fusion protein can effectively recruit other members of the PRC2 complex and catalyze the trimethylation of H3K27 along the transgene (Rush *et al*., 2009).

In contrast to the work done by Hansen *et al*., I sought to eliminate the initial PRC2 recruitment event rather than downregulate the Gal4DBD fusion protein to study the heritability of H3K27me3. With the addition of Flp recombinase target (FRT) sites flanking the gal4 sites used to recruit PRC2 (see ***Figure 4***), I developed an efficient method of deleting the gal4 sites after H3K27me3 establishment. Expressed Flp recombinase acts on the FRT sites within the integrated transgene, excising the gal4 sites within the two FRT sites.

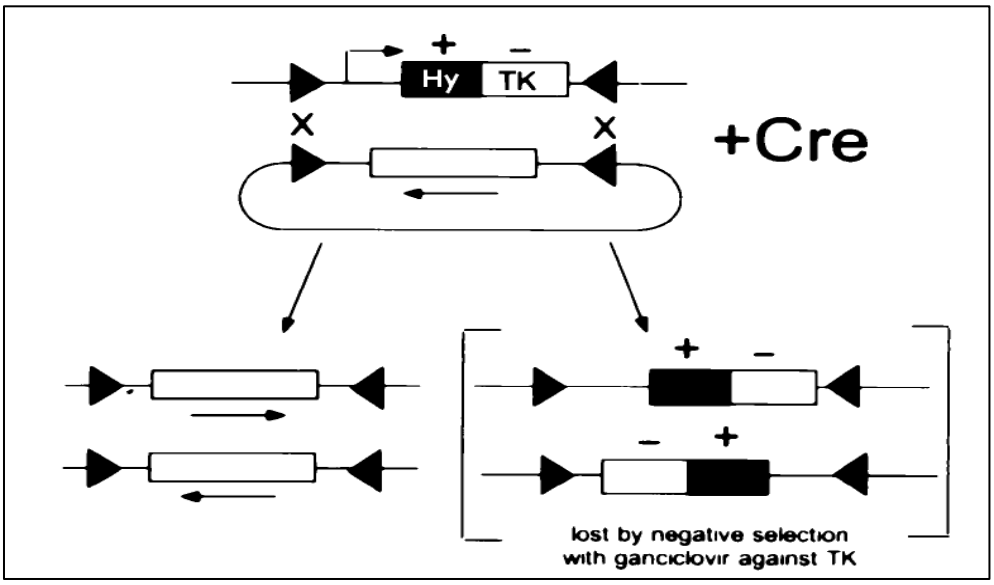

Figure 3. RMCE. The L1-HYTK-1L cassette is integrated in the RL5 locus in MEL cells. Co-transfection of a L1-1L plasmid and Cre recombinase results in exchange of the HYTK cassette with the transgene of interest, in either orientation, upon successful recombination.

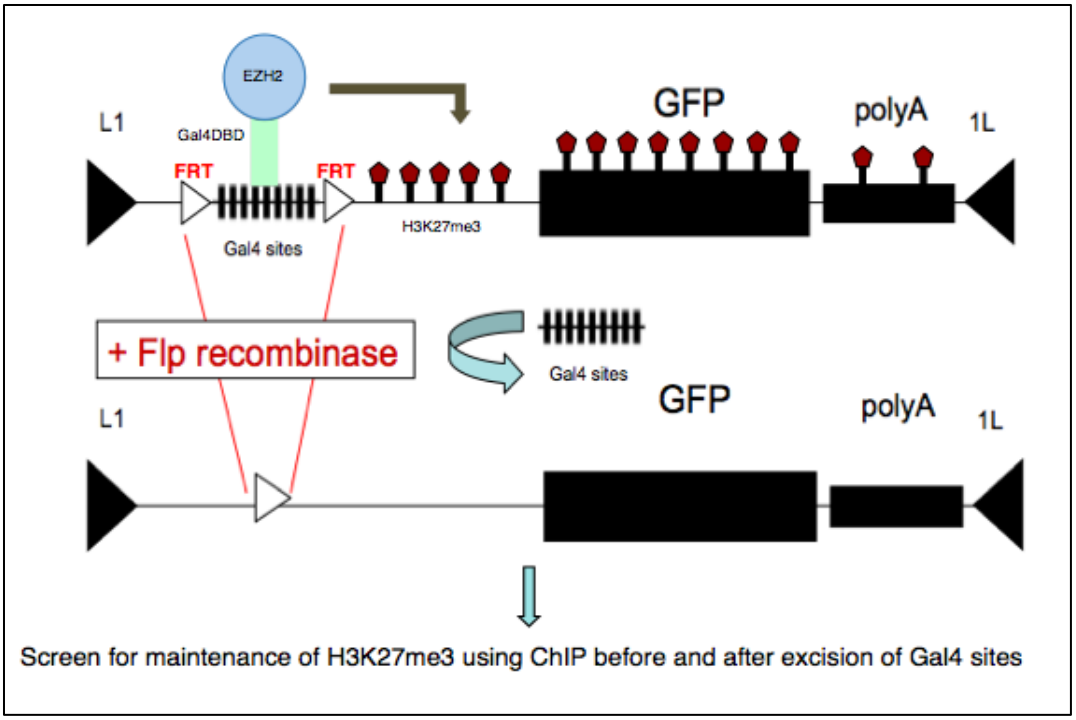

Figure 4. Outline of experimental design. H3K27me3 maintenance upon cell division was examined by firstly establishing the mark via Gal4DBD recruitment of PRC2, and then abolishing this recruitment mechanism by excising the gal4 sites. FRT = Flp recombinase target.

3.2 Generation of constructs

Constructs were generated with all the necessary components for successful RMCE, recruitment of PRC2 and deletion of the gal4 sites. Each cassette contained inverted *loxP* sites, the *GFP* gene with or without a p16 promoter, and gal4 sites flanked by FRT sites. L1-FRTgal4FRT-e*GFP*-1L was generated to screen for H3K27me3 maintenance alone, while L1-FRTgal4FRT-p16-*GFP*-1L was generated in hopes that the transcription of *GFP* would be affected by the recruitment of PRC2. The p16 promoter is a well-studied promoter of a tumour suppressor gene, downregulated in a number of cancers, and is known to be a target of polycomb proteins (Liggett and Sridansky, 1998, Bracken *et al*., 2007). Therefore, this promoter was incorporated into the cassette with the hopes that *GFP* expression would be reduced when the H3K27me3 mark was established at the cassette.

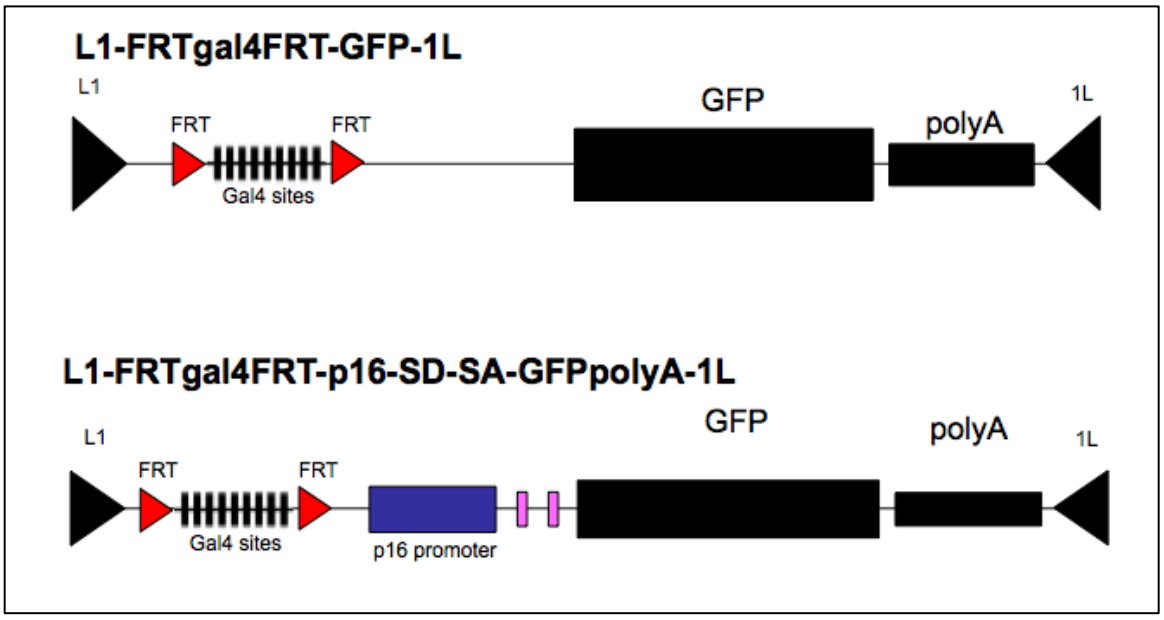

Figure 5. Experimental constructs used for RMCE.

3.3 RMCE

Each of the constructs shown in ***Figure 5***, along with the L1-CMV-*GFP*-1L and the L1-p16-*GFP*-1L control constructs, were integrated into RL5 MEL cells via RMCE. Two MEL lines were used: the Gal4DBD-EZH2 expressing line, and the Gal4DBD expressing control line. Plasmid and Cre recombinase transfection was performed in parallel across all cell lines. After transfection and 12 days of ganciclovir selection, RMCE success rate was analyzed by flow cytometry (***Figure 6***).

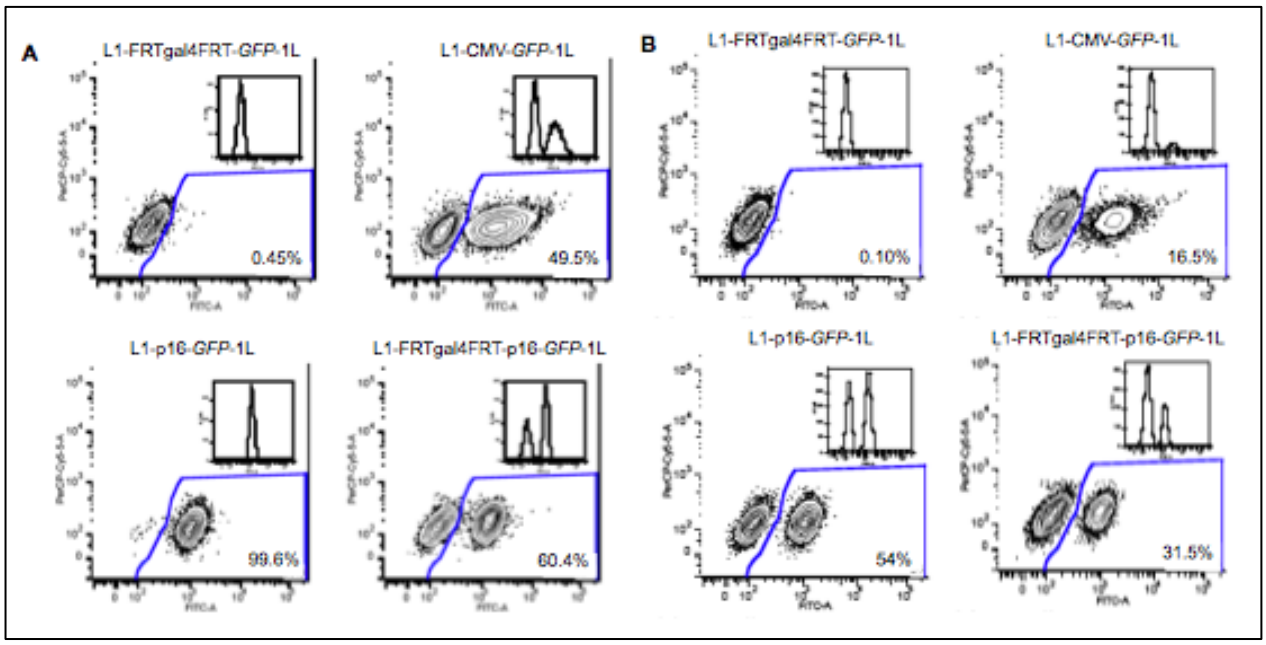

Figure 6. Flow cytometry on RMCE'd RL5 MEL pools. **A**: Gal4DBD-EZH2 expressing MEL cell lines. **B**: Gal4DBD expressing MEL cell lines. L1-FRTgal4FRT-*GFP*-1L RMCE used as a negative RMCE control for *GFP* expression, L1-CMV-*GFP*-1L and L1-p16-*GFP*-1L used as positive RMCE controls for *GFP* expression. Percentages indicate portion of *GFP* positive cells in each pool.

RMCE success rate ranged from 49.5%-99.6% *GFP* positive cells in the Gal4DBD-EZH2 MEL line, and 16.5%-54% *GFP* positive cells in the Gal4DBD MEL line. Importantly, the L1-FRTgal4FRT-p16-*GFP*-1L cassette is expressed in both the Gal4DBD and the Gal4DBD-EZH2

MEL lines, which was unexpected because PRC2 is presumably recruited to the cassette in the Gal4DBD-EZH2 line (Rush *et al*, 2009), resulting in the trimethylation of H3K27 along the p16 promoter. Median fluorescence was measured to be similar between the two lines, indicating that *GFP* expression is not reduced significantly in the Gal4DBD-EZH2 line (data not shown). It appears as though recruitment of the H3K27me3 mark in this context is not sufficient to downregulate transcription of *GFP*. Therefore, I continued my studies with the cell lines containing the promoterless *GFP* construct.

3.4 Validation of cell lines

After single cell-derived clones were generated from the L1-FRTgal4FRT-*GFP*-1L RMCE pools, a subset of clones were picked and assayed for cassette orientation, *GFP* copy number and H3K27me3 status. Cassette orientation was determined by PCR. As mentioned above, the process of RMCE can result in the cassette being integrated in either orientation, and therefore primers needed to be designed to distinguish one orientation from another. To screen for the correct orientation (in which all clones were matched), one primer was located within the *GFP* gene, while the other was located outside the cassette in the RL5 locus. Out of five clones screened, one clone from the RMCE'd Gal4DBD-EZH2 L1-FRTgal4FRT-*GFP*-1L clones and one clone from the RMCE'd Gal4DBD L1-FRTgal4FRT-*GFP*-1L clones showed the cassette integrated in the proper orientation (***Figure 7***).

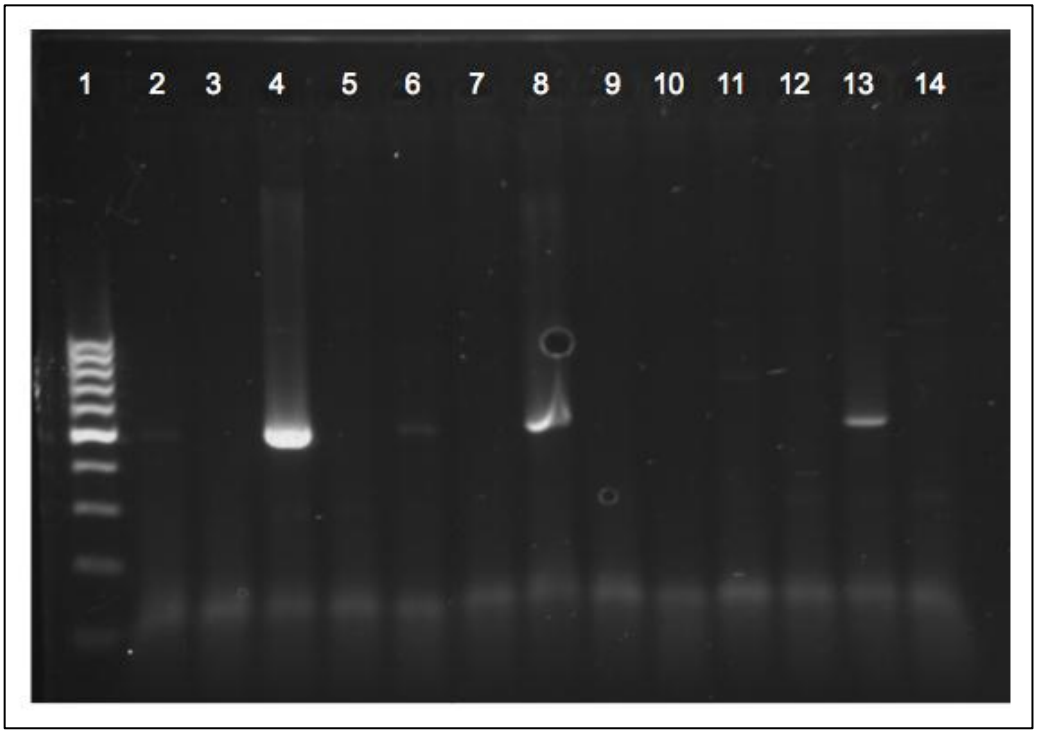

Figure 7. PCR orientation screen of L1-FRTgal4FRT-*GFP*-1L RMCE'd clones. Lane 1, 100 bp ladder. Lanes 2-6, Gal4DBD clones 1-5. Lanes 7-11, Gal4DBD-EZH2 clones 1-5. Lane 12, orientation "A" negative control. Lane 13, orientation "B" positive control. Lane 14, H_2O control.

Positive clones from each line were further screened by qPCR to determine *GFP* copy number. The process of RMCE results in only one integration event per cell in greater than 90% of isolated clones (Lorincz *et al*, 2004; Appanah *et al*., 2007), and indeed, when compared to a previously validated RMCE clone (Appanah *et al*., 2007), this is what was observed (***Figure 8***).

Both clones were then analyzed for H3K27me3 enrichment by native ChIP at two regions along the cassette, and one region outside the cassette in the RL5 locus. As expected, the Gal4DBD-EZH2 L1-FRTgal4FRT-*GFP*-1L clone 2 showed H3K27me3 enrichment of various levels across the cassette, while the Gal4DBDL1-FRTgal4FRT-*GFP*-1L clone 3 showed no significant enrichment above background (***Figure 9***). As expected, these data taken together confirm that the Gal4DBD-EZH2 clone is able to recruit PRC2 and establish H3K27me3 prior to deletion of the gal4 sites.

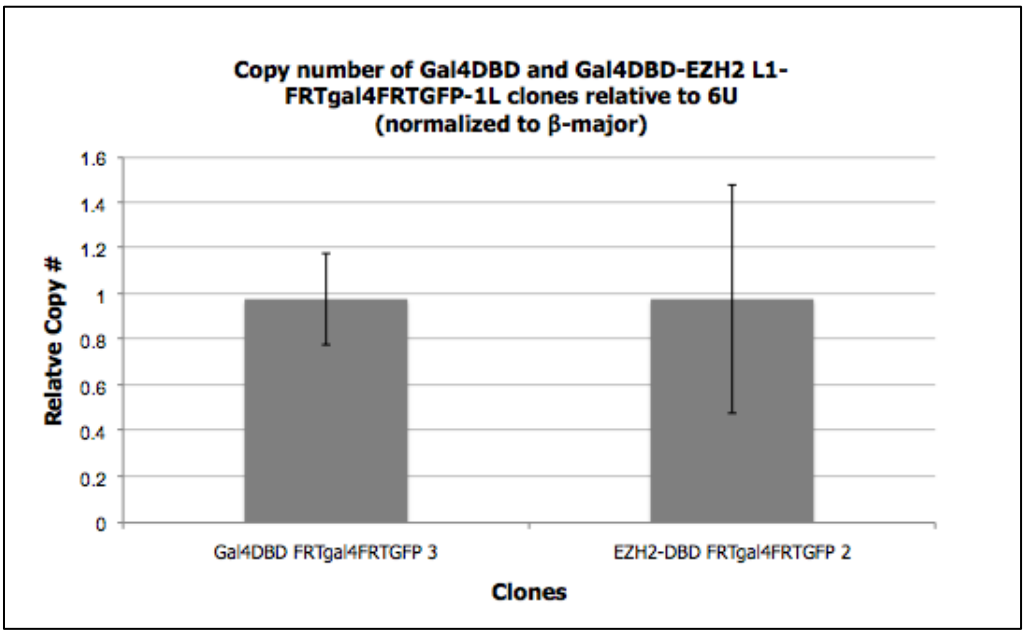

Figure 8. *GFP* copy number. GFP-containing Gal4DBD L1-FRTgal4FRT-GFP-1L clone 3 and Gal4DBD-EZH2 FRTgal4FRTGFP clone 2 were assayed by qPCR to determine GFP copy number per cell. *GFP* copy number was normalized to an internal control (*β-major*) and compared to the previously validated 6U RMCE clone. Error bars represent the standard deviation of PCR replicates performed in triplicate.

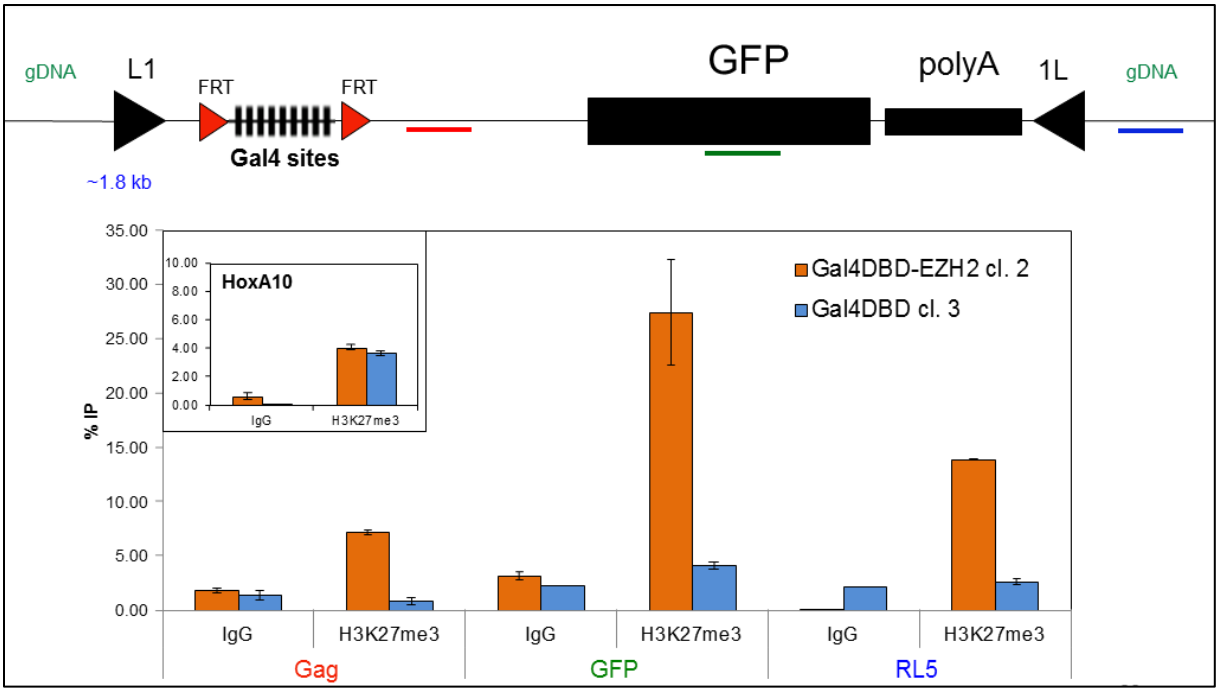

Figure 9. H3K27me3 enrichment over input at the L1-FRTgal4FRT-*GFP*-1L cassette in Gal4DBD-EZH2 clone 2 and Gal4DBD clone 3. H3K27me3 levels were measured at Gag, GFP, and RL5 regions (colours of text match coloured regions in schematic). Error bars represent the standard deviation of PCR replicates performed in triplicate.

3.5 Expression of Flp recombinase and Gal4 deletion

Transient expression of Flp recombinase was first attempted using a reporter that expressed dsRed only in cells where the co-transfected Flp recombinase was active. The dsRed reporter contained an antibiotic resistance gene flanked by FRT sites between the promoter and the *dsRed* gene, and thus the gene could only be successfully transcribed once Flp recombinase excised the antibiotic resistance gene. However, although this reporter system worked well in human Phoenix A cells, repeated co-transfection attempts in MEL cells were unsuccessful, and it was concluded that the dsRed protein may be toxic to MEL cells. Therefore, the Flp expression plasmid was co-transfected instead with a CMV-*GFP* reporter to achieve transient expression. 40 hours post transfection, *GFP* positive cells were sorted into 96 well plates to generate single cell-

derived clones (***Figure 10***). Clones were later screened by PCR to ensure that the CMV-*GFP* plasmid did not integrate (data not shown).

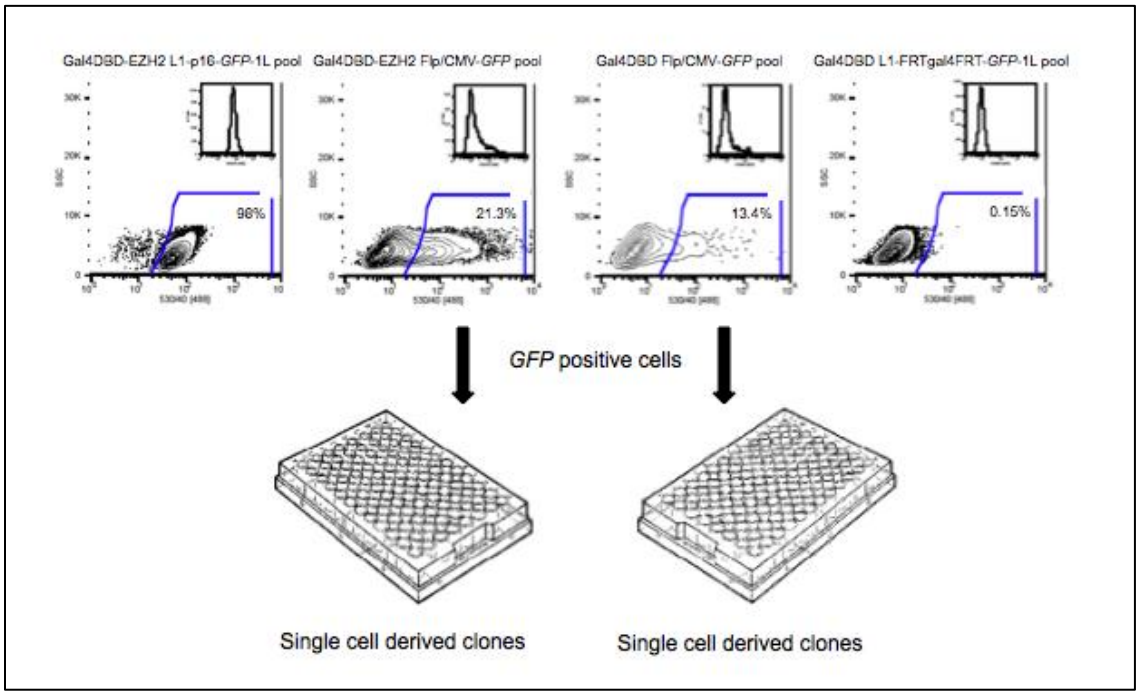

Figure 10. Transient co-expression of Flp recombinase and L1-CMV-*GFP*-1L. Transfected cells were underwent fluorescence-activated cell sorting (FACS) and *GFP* positive cells were sorted into 96 well plates to generate single cell-derived clones.

Cells were grown for approximately 40 cell divisions before subsequent Gal4 deletion and H3K27me3 maintenance screens were performed. This allowed ample time for the H3K27me3 mark to be lost passively after deletion of the Gal4 sites, if the deletion was indeed successful. The Gal4 site deletion screen was performed by PCR, on 10 *GFP* positive clones (***Figure 11***).

Clones harbouring the Gal4 deletion should generate a smaller amplicon (500 bp) than clones in which the Gal4 sites are still present (800 bp). Indeed, clones of each class were identified. Two clones from each line were chosen for comparative ChIP analysis; one that still contained the Gal4 sites (Gal4DBD/Gal4DBD-EZH2 L1-FRTgal4FRT-*GFP*-1L +Gal4) and one that had lost

the Gal4 sites (Gal4DBD/Gal4DBD-EZH2 L1-FRTgal4FRT-*GFP*-1L –Gal4). These clones were cultured and further analyzed in parallel.

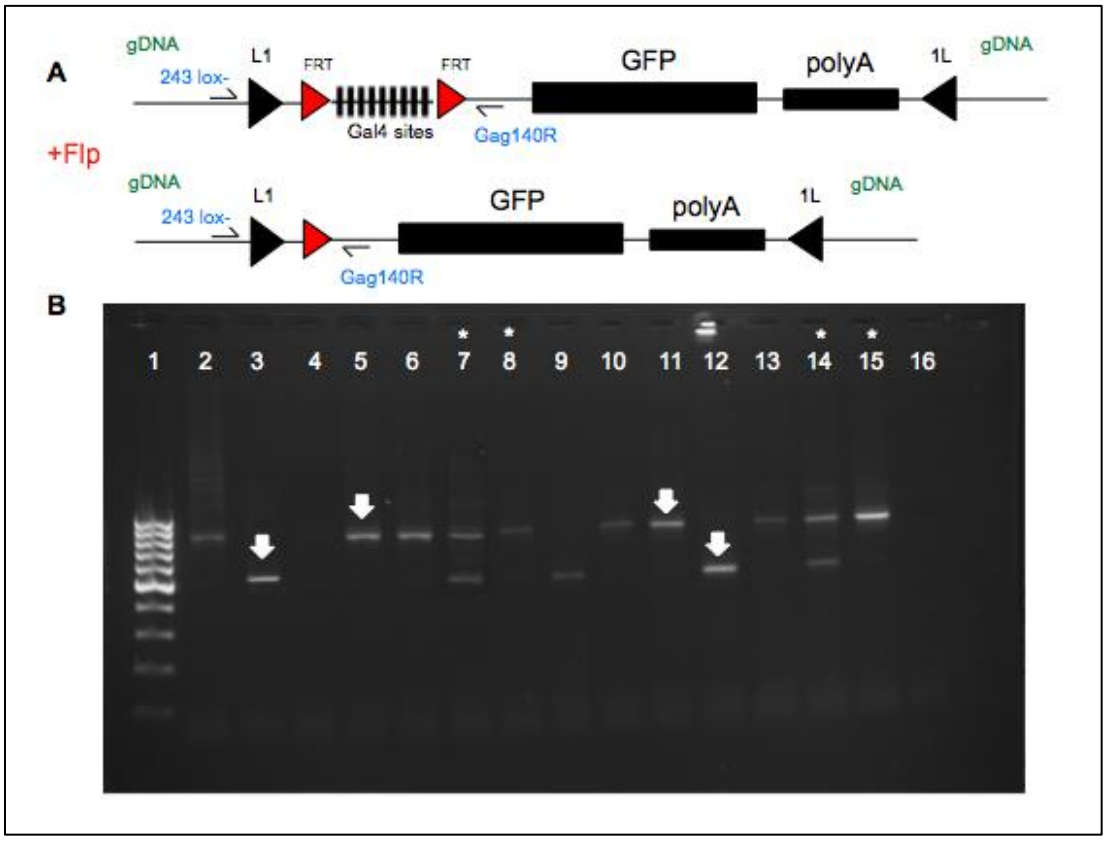

Figure 11. PCR screen of Gal4 site deletion. Lane 1, 100 bp ladder. Lanes 2-6, Gal4DBD-EZH2 L1-FRTgal4FRT-*GFP*-1L selected Flp-transfected clones. Lane 7, Gal4DBD-EZH2 L1-FRTgal4FRT-*GFP*-1L Flp-transfected pool. Lane 8, Gal4DBD-EZH2 L1-FRTgal4FRT-*GFP*-1L clone 2 pre+Gal4. Lanes 9-13, Gal4DBD L1-FRTgal4FRT-*GFP*-1L selected Flp-transfected clones. Lane 14, Gal4DBD L1-FRTgal4FRT-*GFP*-1L Flp-transfected pool. Lane 15, Gal4DBD L1-FRTgal4FRT-*GFP*-1L clone 3 pre+Gal4. Lane 16, H_2O control. Asterisks indicate PCR controls; arrows indicate clones chosen for ChIP.

3.6 ChIP analysis of +Gal4 and -Gal4 clones

Gal4DBD/Gal4DBD-EZH2 L1-FRTgal4FRT-*GFP*-1L +Gal4 and Gal4DBD/Gal4DBD-EZH2 L1-FRTgal4FRT-*GFP*-1L –Gal4 clones were assayed for H3K27me3 enrichment at 4 different regions across the cassette. As expected, the Gal4DBD control showed no H3K27me3 enrichment in either the +Gal4 or the -Gal4 clones (***Figure 12***). However, in the Gal4DBD-

EZH2 lines, the H3K27me3 mark appeared to persist in both the +Gal4 and the -Gal4 clones, in 3 out of 4 genomic locations screened. Significant enrichment of H3K27me3 in the Gal4DBD-EZH2 FRTgal4FRTGFP +Gal4 and -Gal4 clones was noted over the Gag, 5' GFP, and *GFP* locus (the latter by approximately 4.5 fold), while the Gal4DBD-EZH2 FRTgal4FRTGFP -Gal4 clone was not enriched over the RL5 locus.

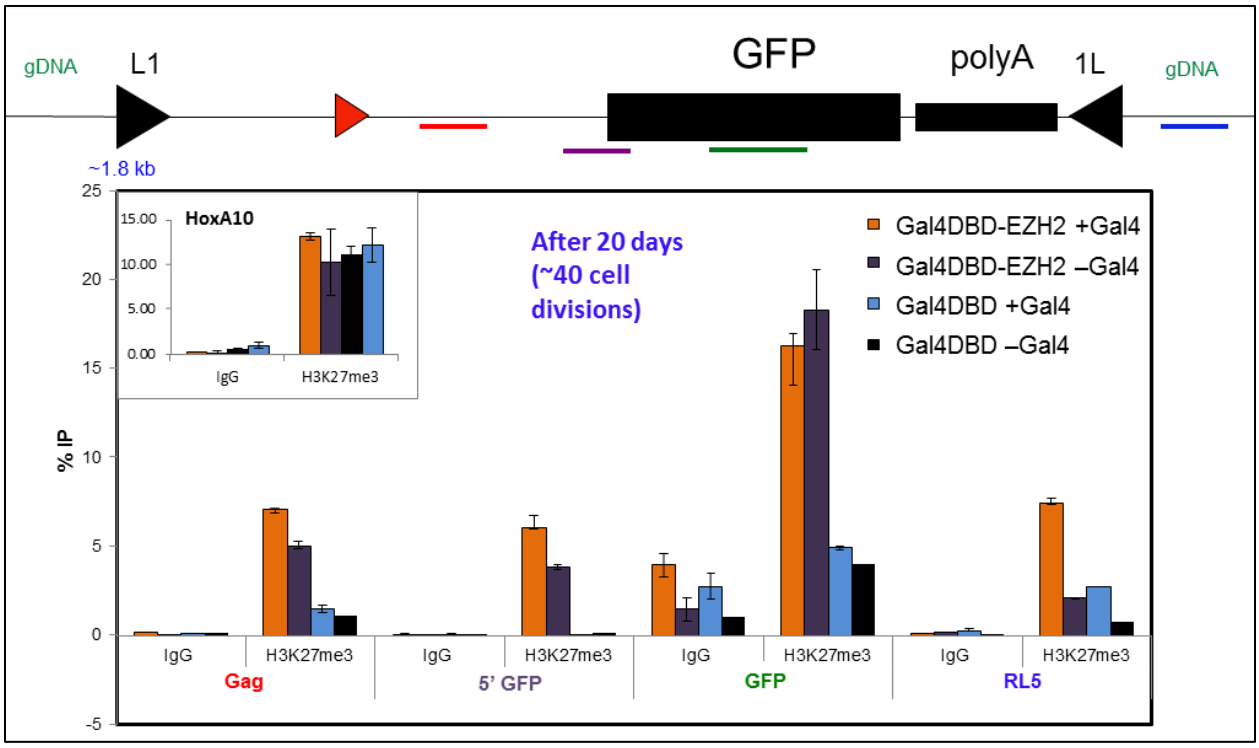

Figure 12. Comparison of H3K27me3 enrichment levels before and after Gal4 site deletion in the Gal4DBD-EZH2 and the Gal4DBD expressing cell lines. Colour coded names on the X axis correlate with the locations/amplicons screened in the overhead schematic of the cassette. IgG was used as a negative control. Error bars represent the standard deviation of PCR replicates performed in triplicate.

To further confirm that this enrichment was not a result of Gal4DBD-EZH2 being continuously recruited to the cassette, and as expected, presence of Gal4DBD-EZH2 decreased in the -Gal4 clone (***Figure 13*****).**

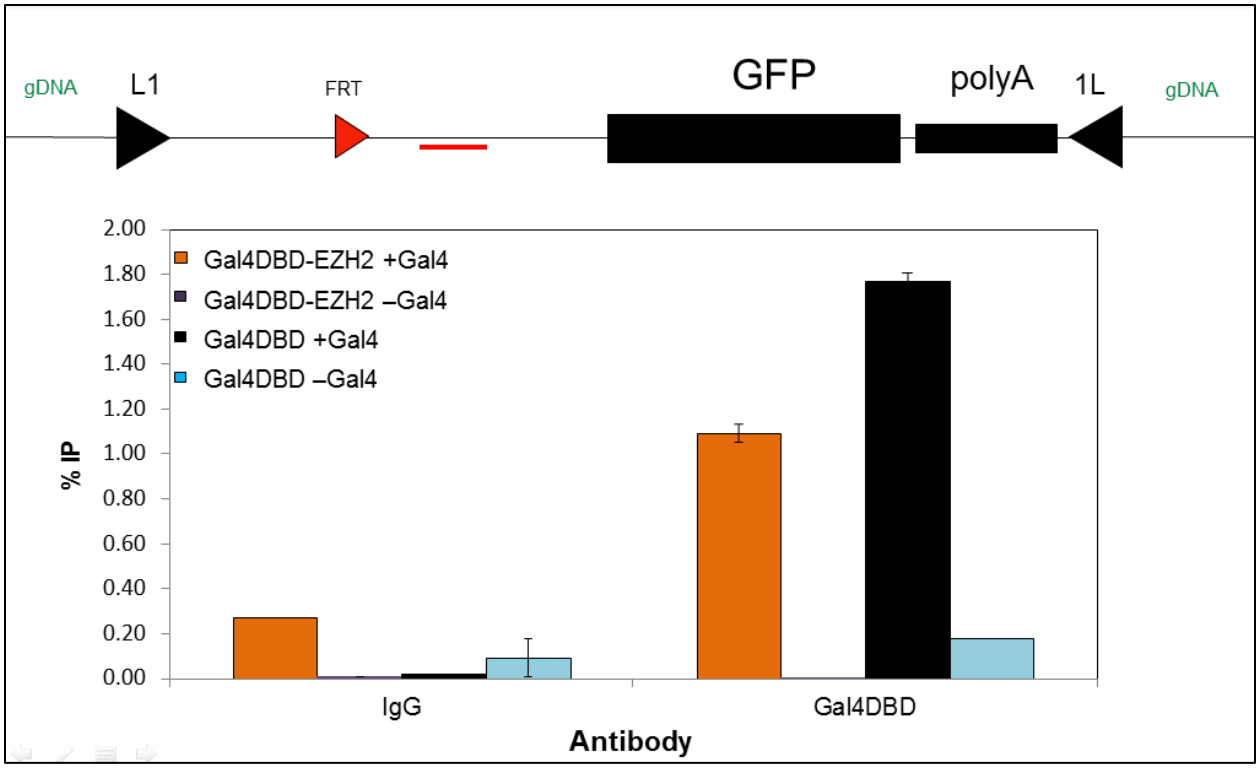

Figure 8. Comparison of Gal4DBD enrichment levels before and after Gal4 site deletion. An antibody specific for the Gal4DBD was used in crosslink ChIP to screen for Gal4DBD at the FRTgal4FRTGFP locus in pre- and post- +Gal4 clones. IgG was used as a negative control. Error bars represent the standard deviation of PCR replicates performed in triplicate.

Therefore, although the initial means of recruitment of PRC2/H3K27me3 has been abolished, the H3K27me3 mark appears to persist after many cell divisions, consistent with previous data (Hansen *et al*., 2008).

3.7 Generation and validation of Gal4KAP1 cell line

In addition to studying the maintenance of H3K27me3, I also sought to develop a cell line that would allow similar experiments to be performed to determine the inheritance behaviour of the H3K9me3 mark. A Gal4KAP1 fusion protein was chosen to target the H3K9me3 mark to the integrated cassettes; as the KAP1 protein associates with the H3K9 methyltransferase SETDB1 in mammalian ES cells, SETDB1 can thus be recruited to the target sites to deposit H3K9me3

via interaction with Gal4KAP1 (White *et al*., 2006). A previous study has shown that the Gal4KAP1 protein can repress transcription as a result of direct tethering to DNA via gal4 sites (Sripathy *et al*., 2006). As the KAP1/SETDB1 repression system has not been observed in MEL cells, the Gal4KAP1 fusion protein was transfected into the HA36 cell line. HA36 cells contain an RMCE site, so that the experiments performed in the MEL cells could be replicated.

HA36 cells were stably transfected with a linearized pC3-ERHBD-GAL4KAP1 neomycin-resistant plasmid and cultured in 200 µg/ml G418 for 5 days. Surviving cells were expanded and single cell derived clones were generated. Validation of integration of pC3-ERHBD-GAL4KAP1 was performed by PCR, and positive clones were chosen for future RMCE experiments (**Figure 14**).

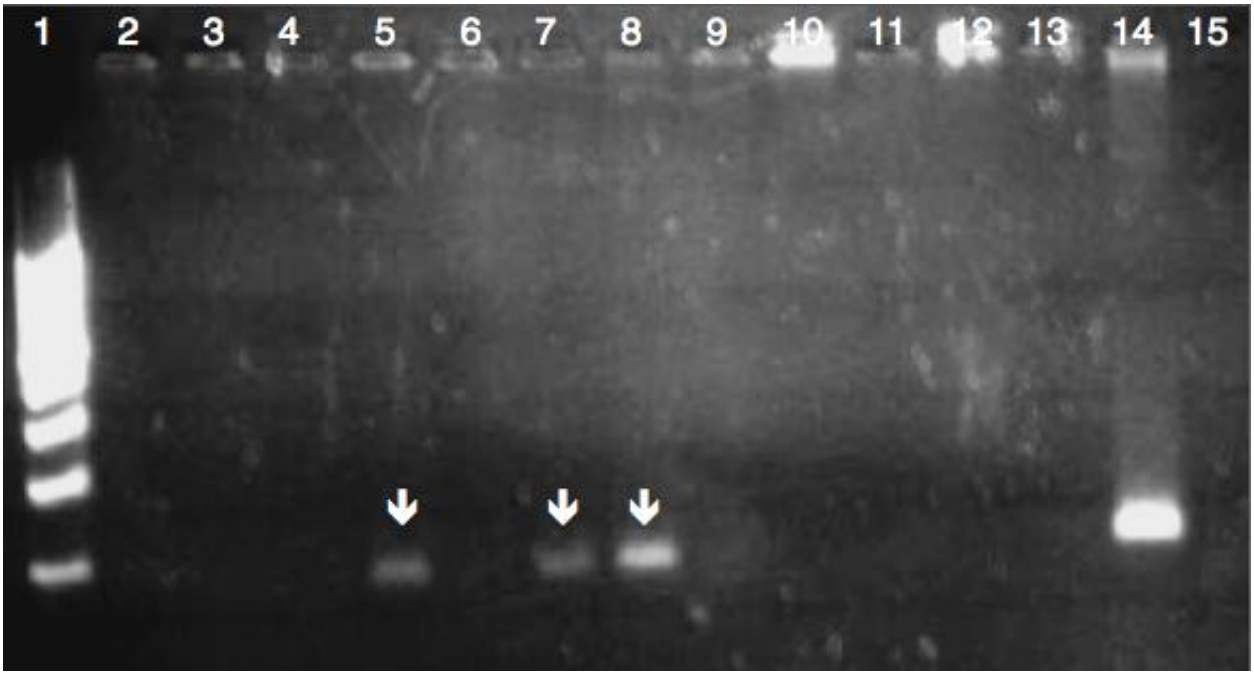

Figure 14. Validation of integration of pC3-ERHBD-GAL4KAP1 in HA36 cells. Lane 1, 100 bp ladder. Lane 2-13, pC3-ERHBD-GAL4KAP1 clones 1-12. Lane 14, pC3-ERHBD-GAL4KAP1 control plasmid. Lane 15, H_2O. Arrows indicate positive clones containing the pC3-ERHBD-GAL4KAP1 fusion gene.

4 Discussion

While previous evidence supported the notion that PRC2 could bind to and recognize its own H3K27me3 mark, the approach employed relied on tight regulation of expression of the Gal4DBD-EED fusion gene from a TET inducible promoter (Hansen *et al*., 2008). Due to the developmental importance of polycomb mediated silencing, it is important to thoroughly investigate all aspects of polycomb recruitment and H3K27me3 maintenance in mammalian cells. Therefore, I chose to establish an alternative methodology for determining whether the mechanism of H3K27me3 heritability is indeed distinct from the mechanism of its initial establishment, using the FRT/FLP recombinase-system to delete the gal4 binding sites from the transgene under study.

In this thesis, I posed the question of whether the H3K27me3 mark persists on an integrated transgene after the initial recruitment mechanism had been removed. Using a previously published cell line from our lab (Rush *et al*., 2009), I introduced a cassette capable of recruiting the Gal4DBD-EZH2 fusion protein into a targeted euchromatic integration site. This fusion protein formed a complex with the remainder of the PRC2 subunits and subsequently catalyzed the trimethylation of H3K27 >1.8 kb downstream along the transgene and in genomic DNA. Upon deletion of the Gal4 sites responsible for recruiting the Gal4DBD-EZH2, several rounds of cell division took place before screening for the persistence of H3K27me3. Indeed, I observed that the H3K27me3 mark appeared to be inherited upon DNA replication, as the enrichment levels remained similar to those of a cell line that retained its Gal4 sites. This maintenance is due at least in part to the endogenous PRC2, as Gal4DBD enrichment was negligible after Gal4 site deletion. Thus, it appears that the H3K27me3 mark can persist through multiple cell divisions at

a heterologous site, even after removal of the DNA binding sites required for the initial recruitment of the EZH2 fusion, and in turn the initial establishment of this mark.

The RMCE method used in our lab was first published by Feng *et al*. in 1999. It is a two-step system that uses both negative and positive selection to result in efficient and highly successful targeted cassette integration. A pre-localized cassette containing a *HyTK* gene initially provides the cell with hygromycin resistance. Exchange of this negative selection marker with an introduced plasmid occurs via Cre-mediated recombination between inverted Lox sites, which flank the endogenous cassette and the plasmid cassette. Cells containing a successful recombination event can now be positively selected for using ganciclovir. This method results in an integration efficiency of >90% in MEL cells, and 10%-50% in embryonic stem cells (Feng *et al*., 1999).

Initially, I monitored the efficiency of Cre-mediated recombination/RMCE targeting by including GFP reporter constructs RMCE'd in parallel with the promoterless *GFP* construct. Interestingly, although the Gal4DBD-EZH2 and the Gal4DBD pre- and post-RMCE lines were selected for the same amount of time and at the same concentration of antibiotic, the success rate (measured by *GFP* fluorescence) varied widely within and across parent MEL lines (***Figure 6***). One explanation for this obvious difference in RMCE efficiency between the two lines may be the observation that the parent Gal4DBD-EZH2 MEL line grows faster than the parent Gal4DBD MEL line, perhaps as a result of the location of integration of the Gal4DBD-EZH2 fusion gene or the expression of the EZH2 fusion. The difference in growth rate may impact the antibiotic's effectiveness, thus killing off more of non-RMCE'd cells in the Gal4DBD-EZH2 pool. The L1-p16-*GFP*-1L cassette appears to have a higher RMCE success rate than the L1-CMV-*GFP*-1L within both lines. It is known that the L1-CMV-*GFP*-1L cassette is resistant to silencing at the

RL5 locus in MELs (Lorincz lab, unpublished); therefore, I speculate that there may have been a contaminant, perhaps mycoplasma, that reduced the effectiveness of Cre recombinase and/or the ganciclovir selection. Nevertheless, since single-cell clones were derived from the MEL lines containing the promoterless *GFP* construct and further screened, a relatively low efficiency of RMCE was of little importance.

Surprisingly, the L1-FRTgal4FRT-p16-*GFP*-1L construct was not silenced in the Gal4DBD-EZH2 line. Since the L1-FRTgal4FRT-*GFP*-1L construct successfully recruited the Gal4DBD-EZH2 fusion protein and was enriched for H3K27me3 as a result (***Figure 9***), it was expected that the H3K27me3 would be sufficient to silence the expression of L1-FRTgal4FRT-p16-*GFP*-1L. The endogenous *INK4A* locus (driven by the p16 promoter) is blanketed by H3K27me3 *in vivo* in mouse embryonic stem cells, silencing the locus, yet is activated by oncogenic insults and also with aging (Bracken *et al*., 2007; Kotake *et al*., 2011). It is possible that since MEL cells are virally transformed, and maintained at the proerythroblast stage of maturation, their oncogenic nature may override the induced polycomb silencing of Gal4DBD-EZH2 (Elnitski and Hardison, 1999). It is also possible, though not tested in this study, that H3K4 methylation is also present at the cassette, thus preventing complete repression. The cassette used in the Rush *et al*. paper was enriched for H3K4 methylation as well as H3K27me3 methylation (Rush *et al*., 2009). Therefore, the promoterless L1-FRTgal4FRT-*GFP*-1L cassette was used for the remainder of the H3K27me3 maintenance study presented here.

Further validation of the clones derived from the post-selection RMCE pool of L1-FRTgal4FRT-*GFP*-1L cells produced results consistent with what has been observed in our lab. Matched orientation clones (***Figure 7***) were chosen to be compared with one another to accurately assess H3K27me3 levels at the same regions across the cassette. Since the integration event only

occurred at the RL5 locus and not in a second copy elsewhere in the genome (***Figure 8***), any H3K27me3 screens performed on the cassette must be a consequence of recruitment of the Gal4DBD-EZH2 fusion, rather than spreading of this mark from flanking genomic sequence. Indeed, a cassette introduced at the same integration site but lacking gal4 sites is not marked by H3K27me3 (Rush *et al*., 2009).

As expected, only the L1-FRTgal4FRT-*GFP*-1L containing clone expressing Gal4DBD-EZH2 was marked by H3K27me3 (***Figure 9***). As shown previously in our lab, the gal4 sites placed upstream of *GFP* can successfully recruit Gal4DBD-EZH2 and establish H3K27me3, and here I show that the FRT sites flanking the Gal4 sites do not interfere with PRC2 recruitment. Interestingly, instead of the H3K27me3 levels decreasing steadily away from the Gal4 sites, the GFP region appears to be marked more than 2-fold higher than the Gag region, and 1.5 fold higher than the RL5 region. This may be due to the fact that the *GFP* gene contains a CpG island. While the original *GFP* gene found in *Aequorea victoria* contains only 12 CpG sites, the gene was modified for optimal expression in mammalian cells and now contains 60 CpG sites (Prasher *et al*., 1992; Yang *et al*., 1996). CpG islands in the genome are generally devoid of DNA methylation (Deaton and Bird, 2011). As observed by several research groups, genomic regions devoid of DNA methylation tend to be marked by H3K4 and/or H3K27 methylation, the latter occurring at silent loci (Mathieu *et al*., 2005; Hawkins *et al*., 2010; Mendenhall *et al*., 2010; Kondo *et al*., 2008; Gal-Yam *et al*., 2008). While it is possible that the CpG island in the *GFP* region in the cassette is recruiting endogenous PRC2, the IgG enrichment of the same locus is also higher than the Gag and RL5 regions. However, the H3K27me3 enrichment increases ~2.5 fold at the *GFP* region compared to Gag and RL5, whilst the IgG enrichment only increases ~1.8 fold at the *GFP* region over input. This result is consistent with Rush *et al*.'s previous study,

which also demonstrates an increase of H3K4 methylation at the *GFP* site. Therefore, I conclude that the result observed in ***Figure 9*** is a result of H3K27 methylation being catalyzed by endogenous PRC2, albeit inefficiently. It does not significantly affect subsequent studies, because H3K27me3 levels are measured comparatively between the Gal4DBD and Gal4DBD-EZH2 lines.

In order to delete the Gal4 sites and further study its effect on the maintenance of H3K27me3, I employed FRT sites and transient expression of Flp recombinase. Because the construct containing Flp recombinase did not itself carry a reporter gene, it was necessary to co-transfect it with a second plasmid containing a reporter. A reporter containing dsRed was used in early attempts, but after several failed transfections, it was concluded that the dsRed was likely toxic to MEL cells. Thus, the Flp recombinase was co-transfected with a plasmid containing *GFP* that expresses higher levels of GFP and shows strong fluorescence when transiently expressed. The efficiency of transfection ranged from 13.4% to 21.3%, as determined by *GFP* positive cells.

Clones that were derived from the pool of *GFP* positive cells were screened for FLP-mediated deletion by PCR, using primers that generate a smaller amplicon and, as expected, the Gal4 sites were successfully deleted in a subset of clones post+Gal4 transfection (***Figure 11***). The non-deleted clones served as ideal controls, as they were transfected, sorted, and cultured in parallel with the deleted clones. Interestingly, out of the samples screened, the Gal4DBD line seemed to harbour more deleted clones than the Gal4DBD-EZH2 line. While the difference may be insignificant, one could speculate that the Flp recombinase has a more difficult time performing its action at the FRT sites when PRC2 components are recruited to nearby regions within the cassette.

Two clones from each line, one harbouring the Gal4 site deletion and one retaining the sites, were assayed by native ChIP to determine if the H3K27me3 mark persisted after abolishment of Gal4DBD-EZH2 recruitment. As these clones were derived from single cells, they had ample time to lose the H3K27me3 actively or passively as they were cultured and expanded for ChIP analysis (21-25 days, approximately 40 cell divisions). Despite the length of time these clones were cultured, the Gal4DBD-EZH2 clone with deleted Gal4 sites retained similar H3K27me3 levels across the cassette (***Figure 12***). As previously noted, the H3K27me3 enrichment at *GFP* exceeds the levels observed at the more proximal Gag region as well as the distal RL5 region; therefore, a fourth region was screened, containing the transcription start site of *GFP*. This region is enriched for H3K27me3 at a level similar to Gag and RL5, thus supporting the speculation that the CpG island within the gene body of *GFP* has some capacity to recruit endogenous PRC2 independent of the gal4 sites introduced. Interestingly, the H3K27me3 levels in the RL5 region, just outside of the cassette, appear to decrease to the level of the Gal4DBD control line upon deletion of the Gal4 sites. This unexpected result will be discussed in further detail later.

Persistence of H3K27me3 across the cassette following deletion of the gal4 sites is consistent with previous research indicating that the initial recruitment mechanism is not necessary to maintain the mark at a specific locus (Hansen *et al*., 2008). The data presented in this thesis contributes to the template-copying hypothesis of histone modification inheritance. While this evidently does not appear to be a highly precise mechanism, the template-copying model may indeed be the method of inheritance for histone modifications that undergo spreading. Certain histone modifications (including H3K27me3) can spread along chromatin in a blanket-like fashion over kilobases of DNA, such as in the Hox cluster (Pauler *et al*., 2009). Such spreading

is presumably regulated by proteins that can read the modification, stimulate an effector protein, and then catalyze the mark on an adjacent nucleosome in a positive feedback fashion. HP1 in S. pombe is a prime example, as well as Sir3 in S. cerevisiae, which is an effector protein that binds to acetylated H4K16 and associates with Sir2, an H4K16 deacetylase (Onishi *et al*., 2007). The PRC2 complex may act as both an effector and a chromatin-modifying enzyme in the case of H3K27me3. Since EED has been shown to bind specifically to H3K27me3 and stimulate the activity of EZH2, the simplest model would be that the PRC2 complex recognizes its own mark, likely H3K27me3, and promotes the catalytic activity of EZH2 towards newly synthesized histone H3, either within the same nucleosome and/or adjacent nucleosomes. While a few other proteins have been shown to have binding specificity for H3K27me3 and thus could serve as readers, specific polycomb proteins, such as EED (Margueron *et al*., 2009) have a high association with H3K27me3 (Vermeulen *et al*., 2010). In a follow up study on EED done by Xu *et al*., the structure of EED allows it to bind to methylated lysines, but the actual lysine bound plays a role in EED's subsequent action; EED can bind to H1K26me3, but this binding stimulates EZH2 to methylate EED, not H3K27me3 (Xu *et al*., 2010). Therefore, while EED can bind a handful of targets, only when bound to H3K27me3 can it stimulate EZH2 to catalyze the addition of a trimethyl group on an H3K27.

Based on our lab and others' analysis of the inheritance of the H3K27me3 mark, I wanted to determine whether other histone marks are maintained in the same fashion upon cell division. I chose to look at H3K9me3 (another repressive mark), and observe whether it persists after abolishment of its initial recruitment mechanism.

For this study, HA36 ES cells were used instead of MEL cells. HA36 cells harbour the L1-HyTK-1L cassette at a specific genomic site, rendering it available for RMCE in the same

fashion as MEL cells. In order to observe the effect of H3K9 methylation on the cassettes described previously, it was necessary to express a Gal4 protein fused with KAP1 (KRAB-ZFP associated protein 1), which interacts with SETDB1 (an H3K9 methyltransferase in HA36 cells) to promote H3K9me3 deposition at the target cassette. KAP1 (KRAB-ZFP associated protein 1) interacts with both SETDB1 (an H3K9 methyltransferase) and HP1 at heterochromatin; this interaction of these repression-associated proteins has been shown to be required for the differentiation of F9 cells into parietal endoderm-like cells *in vitro* (Cammas *et al*, 2004). KAP1 is believed to be recruited to chromatin by the KRAB zinc finger protein, via the DNA binding activity of its Zinc finger domain (Huntley *et al*, 2006). In coordination with these findings, the transcriptional repression of a chromatinized reporter gene by a heterologous KRAB repressor protein correlates with localized enrichment of KAP1, SETDB1, and HP1, as well as H3K9 methylation at promoter sequences of the transgene (Ayyananathan *et al*, 2000; Schultz *et al*, 2004). In addition, SETDB1 is important for stem cell maintenance and is required for the repression of endogenous retroviruses in ES cells (Bilodeau *et al*, 2009; Matsui *et al*, 2010); thus, HA36 cells were chosen over MEL cells for this experiment for their biological relevance. Upon induction with 4-OHT, the ERHBD-GAL4KAP1 fusion protein (a gift from David Schultz) can directly target an introduced transgene containing Gal4 sites, resulting in transcriptional repression of the targeted gene (Sripathy *et al*, 2006).

By utilizing the ERHBD-GAL4KAP1 fusion protein in RMCE'd HA36 ES cells, we sought to observe whether endogenous SETDB1 could be recruited by ERHBD-GAL4KAP1 (upon induction of the fusion protein with 4-OHT), and result in subsequent trimethylation of H3K9 across the cassette, similar to what was observed with H3K27me3. I chose to express

ERHBD-GAL4KAP1 in the HA36 cells prior to RMCE so that single cell clones could then be derived to ensure the consistency of ERHBD-Gal4KAP1 expression. After the linearized pC3-ERHBD-Gal4KAP1 plasmid was transfected into HA36 cells and subsequently selected for integration using G418, three single cell derived clones appear to have the fusion gene integrated (**Figure 14**). To test the efficacy of ERHBD-Gal4KAP1 targeting in these positive clones, RMCE with the gal4 cassettes used in the H3K27me3 study was attempted using RMCE conditions identical to those used previously in our lab for ES cells. Using both the test cassettes as well as the L1-CMV-*GFP*-1L control reporter construct, successful targeting was not achieved, despite repeated efforts using varying concentrations of antibiotics, transfection reagents, Cre recombinase, and/or DNA as well as early and late passage HA36 cells.

A subset of cells did survive ganciclovir selection in several experiments, but upon further analysis by flow cytometry to screen for *GFP* in the control line, no expression could be detected (result not shown). Due to time constraints, I did not continue with this project, though the experimental design remains an elegant and suitable method to screen for the persistence of the H3K9me3 mark.

4.1 Applications and future directions

The mechanism by which histone marks are inherited upon cell division is an interesting topic to many facets of biology and medicine. Researchers in epigenetics are interested in knowing how and why histone marks, which influence many cell processes, may persist under certain conditions. The prospect that an abundant repressive histone mark, H3K27me3, is heritable through a template copying model in a differentiated cell line could potentially have important

implications in the maintenance of cellular identity; after all the energy involved in undergoing DNA replication, chromatin assembly, and cell division, the cell must maintain its characteristic pattern of gene expression and silencing. As global changes in the epigenetic landscape are hallmarks of cancer (Sharma *et al*, 2010), knowledge of the heritability of histone modifications may open doors to new treatments and prevention of cancer progression. The data presented in this thesis suggests that H3K27me3 is inherited upon cell division by its recognition by PRC2, which subsequently catalyzes the mark anew on newly synthesized histones. Thus, researchers concerned with the maintenance of cellular identity, specifically the heritability of histone modifications, may find it productive to delve further into the prospect that H3K27me3 inheritance follows the template copying model. For example, instead of using an artificial means of recruitment (Gal4) to target H3K27me3 to an integrated transgene, it may be worthwhile to generate a similar FRT/Flp recombinase-based target cassette using endogenous DNA sequences that promote deposition of the H3K27me3 mark. There are several regions in the mammalian genome that are blanketed with H3K27me3; testing some of these regions to see if they can recruit EZH2 will not only provide insight into the mechanics of PRC2 recruitment, but also to test if a more natural targeting of EZH2 demonstrates the same persistence of H3K27me3 after the recruiting DNA has been removed by Flp recombinase.

In conclusion, I would like to highlight some potential future studies related to this project. Firstly, though Gal4DBD was detected only before Flp recombinase was expressed, it would be wise to test if endogenous PRC2 components persist at the transgene along with H3K27me3. Additionally, since this cassette does not have a transcription profile, it would be interesting to note whether PRC1 is recruited and/or persists with H3K27me3; this would reveal whether we would expect to see the cassette remain silent after several cell divisions. Indeed, since our L1-

FRTgal4FRT-p16-*GFP*-1L cassette was not silenced in the Gal4DBD-EZH2 cell line, it is possible that PRC1 is not recruited in this case. Secondly, validation of the protocol used to test for heritability of H3K9me3 would without a doubt provide useful information to epigeneticists. While H3K9me3 is recognized by HP1 and subsequently recruits H3K9 methyltransferases SUV39H1 and SUV39H2 (Margueron and Reinberg, 2010), this interaction may only play a modest role (if any) in the spreading of the H3K9me3 mark (Maksakova *et al*., 2011). I hope that this research provides insight and direction to epigeneticists who will continue to uncover details of how chromatin and gene transcription are regulated.

References

Annunziatio, A. T. (2005) Split Decision: What Happens to Nucleosomes during DNA Replication? *The Journal of Biological Chemistry*, **280**(13):12065-12068.

Ayyanathan, K., Lechner, M. S., Bell, P., Maul, G. G., Schultz, D. C., Yamada, Y., Tanaka, K., Torigue, K., and Rauscher, F. J. (2003) Regulated recruitment of HP1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: a mammalian cell culture model of gene variegation. *Genes Dev*, **17**:1855-1869.

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

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

Bastow, R., Mylne, J. S., Lister, C., Lippman, Z., Martienssen, R. A., and Dean, C. (2004) Vernalization requires epigenetic silencing of *FLC* by histone methylation. *Nature*, **427**:164-167.

Benson, L. J., Gu, Y., Yakovleva, T., Tong, K., Barrows, C., Strack, C. L., Cook, R. G., Mizzen, C. A., and Annunziato, A. T. (2006) Modifications of H3 and H4 during Chromatin Replication, Nucleosome Assembly, and Histone Exchange. *J Biol Chem*, **281**:9287-9296.

Bernstein, E., Duncan, E. M., Masui, O., Gil, J., Heard, E., and Allis, C. D. (2006) Mouse Polycomb Proteins Bind Differentially to Methylated Histone H3 and RNA and Are Enriched in Facultative Heterochromatin. *Mol and Cell Biol*, **26**(7):2560-2569.

Bilodeau, S., Kagey, M. H., Frampton, G. M., Rahl, P. B., and Young, R. A. (2009) SETDB1contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. *Genes Dev*, **23**:2484-2489.

Bracken, A. P., Kleine-Kohlbrecher, D., Dietrich, N., Pasini, D., Gargiulo, G., Beekman, C.,

Theilgaard-Monch, K., Minucci, S., Porse, B. T., Marine, J-C., Hansen, K. H., and Helin, K. (2007) The Polycomb group proteins bind throughout the INK4A-ARF locus and are disassociated in senescent cells. *Genes and Dev*, **21**(5):525-530.

Brand, A., and Perrimon, N. (1993) Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development*, **118**:401-415.

Bua, D. J., Kua, A. J., Cheung, P., Liu C. L., Migliori, V., Espejo, A., Casadio, F., Bassi, C., Amati, B., Bedford, M. T., Guccione, E., and Gozani, O. (2009) Epigenome microarray platform for proteome-wide dissection of chromatin-signaling networks. *PLoS One*, **4**(8):e6789.

Cammas, F., Herzog, M., Lerouge, T., Chambo, P. and Losson, R. (2004) Association of the transcriptional corepressor TIF1B with heterochromatin protein 1 (HP1): an essential role for progression through differentiation. *Genes Dev*, **18**:2147-2160.

Campos, E. I., Fillingham, J., Li, G., Zheng, H., Voigt, P., Kuo, W-H. W., Seepany, H., Gao, Z., Day, L. A., Greenblatt, J. F., and Reinberg, D. (2010) The program for processing newly synthesized histones H3.1 and H4. *Nature Struc and Mol Biol*, **17**(11):1343-1352.

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

Cao, R., Wang, H., He, J., Erdjument-Bromage, H., Tempst, P., and Zhang, Y. (2008) Role of hPHF1 in H3K27 Methylation and Hox Gene Silencing. *Mol and Cell Biol*, **28**(5):1862-1872.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S., and Zhang, Y. (2002) Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing. *Science*, **298**(5595):1039-1043.

Caretti, G., Di Padova, M., Micales, B., Lyons, G. E., and Sartorelli, V. (2004) The Polycomb Ezh2 methyltransferase regulates muscle gene expression and skeletal muscle differentiation. *Genes and Dev*, **18**:2627-2638.

Chang, L., Loranger, S. S., Mizzen, C., Ernst, S. G., Allis C. D., Annunziato, A .T. (1997) Histones in transit: cytosolic histone complexes and diacetylation of H4 during nucleosome

assembly in human cells. *Biochemistry*, **36**:469-480.

Crawford, N. P. S., Alsarraj, J., Lukes, L., Walker, R. C., Officewala, J. S., Yang, H. H., Lee, M. P., Ozato, K., and Hunter, K. W. (2008) *Bromodomain 4* activation predicts breast cancer survival. *PNAS*, **105**(17):6380-6385.

Deaton, A. M., and Bird, A. (2011) CpG islands and the regulation of transcription. *Genes and Dev*, **25**:1010-1022.

Dhalluin, C., Carlson, J. E., Zeng, L., He, C., Aggarwal, A. K., Zhou, M-M., and Zhou, M-M. (1999) Structure and ligand of a histone acetyltransferase bromodomain. *Nature*, **399**:491-496.

Eissenberg, J. C., and Elgin, S. C. (2000) The HP1 protein family: getting a grip on chromatin. *Curr Opin Genet Dev*, **10**:204-210.

Elnitski, L., and Hardison, R. (1999) Efficient and reliable transfection of Mouse Erythroleukemia cells using cationic lipids. *Blood Cells, Molecules, and Diseases*, **25**(19):299-304.

Feng, Y-Q., Seibler, J., Alami, R., Eisen, A., Westerman, K. A., Leboulch, P., Fiering, S., and Bouhassira, E. E. (1999) Site-specific chromosomal integration in mammalian cells: highly efficient CRE recombinase-mediated cassette exchange. *J Mol Biol*, **292**:779-785.

Flynn, J., Glickman, J. F., and Reich, N. O. (1996) Murine DNA Cytosine-C5 Methyltransferase: Pre-Steady- and Steady-State Kinetic Analysis with Regulatory DNA Sequences. *Biochemistry*, **35**:7308-7315.

Francis, N. J., Kingston, R. E., and Woodcock, C. L. (2004) Chromatin Compaction by a Polycomb Group Protein Complex. *Science*, **306**(5701):1574-1577.

Gal-Yam, E. N., Egger, G., Iniguez, L., Holster, H., Einarsson, S., Zhang, X., Lin, J. C., Liang, G., Jones, P. A., and Tanay, A. (2008) Frequent switching of Polycomb repressive marks and DNA hypermethylation in the PC3 prostate cancer cell line. *Proc Natl Acad Sci*, **105**(35):12979-12984.

Goldberg, A. D., Allis, C. D., and Bernstein, E. (2007) Epigenetics: A Landscape Takes Shape. *Cell*, **128**:635-638.

Grewal, S. I., and Jia, S. (2007) Heterochromatin revisited. *Nat Rev Genet*, **8**:35-46.

Groth, A., Corpet, A., Cook, A. J., Roche, D., Bartek, J., Lukas, J., and Almouzni, G. (2007) Regulation of Replication Fork Progression Through Histone Supply and Demand. **318**(5858):1928-1931.

Groth, A., Rocha, W., Verreault, A., and Almouzni, G. (date) (2007) Chromatin challenges during DNA replication and repair. *Cell*, **128**:721-733.

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**(11):1291-1300.

Hawkins, R. D., Hon, G. C., Lee, L. K., Ngo, Q., Lister, R., Pelizzola, M., Edsall, L. E., Kuan, S., Luu, Y., Klugman, S., Antosiewicz-Bourget, J., Ye, Z., Espinoza, C., Agarwahl, S., Shen, L., Ruotti, V., Wang, W., Stewart, R., Thomson, J. A., Ecker, J. R., and Ren, B. (2010) Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell*, **6**(5):479-491.

Herranz, N., Pasini, D., Diaz, V. M., Franci, C., Gutierrez, A., Dave, N., Escriva, M., Hernandez-Muñoz, I., Di Croce, L., Helin, K., Garcia de Herreros, A., and Peiro, S. (2008) Polycomb Complex 2 is Required for *E-cadherin* Repression by the Snail1 Transcription Factor. *Mol and Cell Biol*, **28**(15):4772-4781.

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. *J of Biol Chem*, **268**(1):305-314.

Huntley, S., Baggott, D. M., Hamilton, A. T., Tran-Gyamfi, M., Yang, S., Kim, J., Gordon, L., Branscomb, E., and Stubbs, L. (2006) A comprehensive catalog of human KRAB-associated zinc finger genes: insights into the evolutionary history of a large family of transcriptional repressors. *Genome Res*, **16**:669-677.

Jackson, V. (1990) In vivo studies on the dynamics of histone-DNA interaction: evidence for nucleosome dissolution during replication and transcription and a low level of dissolution independent of both. *Biochemistry*, **29**:719-731.

Jackson, V. and Chalkley, R. (1985) Histone Segregation on Replicating Chromatin. *Biochemistry*, **24**:6930-6938.

Jaenisch, R. and Bird, A. (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature Genetics, 33 Suppl,* 245-254.

Jasencakova, Z., and Groth, A. (2010) Restoring chromatin after replication: How new and old histone marks come together. *Cell and Dev Biol*, **21**:231-237.

Jennison, J. A. (1995) The Polycomb and trithorax group proteins in *Drosophila*: trans-regulators of homeotic gene function. *Annu Rev Genet*, **29**:289-303.

Jenuwein, T. and Allis, C. D. (2001) Translating the histone code. *Science*, **293**:1074-1079.

Kim, C. A., Sawaya, M. R., Cascio, D., Kim, W., and Bowie, J. U. (2005) Structural Organization of a Sex-comb-on-midleg/Polyhomeotic Copolymer. *The Journal of Bio Chem*, **280**(30):27769-27775.

Kim, J., Daniel, J., Espejo, A., Lake, A., Krishna, M., Xia, L., Zhang, Y., and Bedford, M. T. (2006) Tudor, MBT, and chromo domains gauge the degree of lysine methylation. *EMBO Rep*, **7**:397-403.

Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M. J., Ji, H., Ehrlich, L. I. R., Yabuuchi, A., Takeuchi, A., Cunniff, K. C., Hongguang, H., Mckinney-Freeman, S., Naveiras, O., Yoon, T. J., Irizarry, R. A., Jung, N., Seita, J., Hanna, J., Murakami, P., Jaenisch, R., Weissleder, R., Orkin, S. H., Weissman, I. L., Feinberg, A. P. and Daley, G. Q. (2010) Epigenetic memory in induced pluripotent stem cells. *Nature*, **467**:285-292.

Kim, U. J., Han, M., Kayne, P., and Grunstein, M. (1988) Effects of histone H4 depletion on the cell cycle and transcription of *Saccharomyces cerevisiae. EMBO*, **7**:2211-2219.

Kiruma, H. and Cook, P. R. (2001) Kinetics of core histones in living human cells: little exchange of H3 and H4 and some rapid exchange of H2B. *J Cell Biol*, **153**:1341-1353.

Kmita, M. and Duboule, D. (2003) Organizing axes in time and space; 25 years of colinear tinkering. *Science*, **301**:331-333.

Kokura, K., Sun, L., Bedford, M. T., and Fang, J. (2010) Methyl-H3K9-binding protein MPP8 mediates E-cadherin gene silencing and promotes tumour cell motility and invasion. *EMBO J*, **29**(21):3673-3683.

Kondo, Y., Shen, L., Cheng, A. S., Ahmed, S., Boumber, Y., Charo, C., Yamochi, T., Urano, T., Furukawa, K., Kwabi-Addo, B., Gold, D. L., Sekido, Y., Huang, T. H-M., and Issa, J-P, J. (2008) Gene silencing in cancer by histone H3 lysine 27 trimethylation independent of promoter DNA methylation. *Nat Genet*, **40**(6):741-750.

Kotake, Y., Nakagawa, T., Kitagawa, K., Suzuki, S., Liu, N., Kitagawa, M., and Xiong, Y. Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of $p15^{INK4B}$ tumor suppressor gene. *Oncogene*, **30**:1956-1962.

Krude, T., and Knippers, R. (1991) Transfer of Nucleosomes from Parental to Replicated Chromatin. *Mol and Cell Biol*, **11**(12):6257-6267.

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 and Dev*., **16:**2893-2906.

Landeira, D., Sauer, S., Poot, R., Dvorkina, M., Mazzarella, L., Jorgensen, H. F., Pereira, C. F., Leleu, M., Piccolo, F. M., Spivakov, M., Brookes, E., Pombo, A., Fisher, C., Skarnes, W. C., Snoek, T., Bezstarost, K., Demmers, J., Klose, R. J., Casanova, M., Tavares, L., Brockdorff, N., Merkenschlager, M., and Fisher, A. G. (2010) Jarid2 is a PRC2 component in embryonic stem cells required for multi-lineage differentiation and recruitment of PRC1 and RNA Polymerase II to developmental regulators. *Nature Cell Biology*, **12**(6):618-625.

Lee, J-S., and Shilatifard, A. (2007) A site to remember: H3K36 methylation a mark for histone deacetylation. *Mut Res*, **618**:130-134.

Leung, C., Lingbeek, M., Shakhova, O., Liu, J., Tanger, E., Saremaslani, P., van Lohuizen, M. and Marino, S. (2004) BmiI is essential for cerebellar development and is overexpressed in human medulloblastomas. *Nature*, **428**:337-341.

Lindroth, A. M., Cao, X., Jackson, J. P., Zilberman, D., McCallum, C. M., Henikoff, S., and Jacobsen, S. E. (2001) Requirement of *CHROMOMETHYLASE3* for maintenance of CpXpG methylation. *Science*, **292**(5524):2077-2080.

Lindroth, A. M., Shultis, D., Jasencakova, Z., Fuchs, J., Johnson, L., Schubert, D., Patnaik, D., Pradhan, S., Goodrich, J., Schubert, I., Jenuwein, T., Khorasanizadeh, S., and Jacobsen, S. E. (2004) Dual histone H3 methylation marks at lysines 9 and 27 required for interaction with *CHROMOMETHYLASE3. EMBO J*, **23**:4286-4296.

Liu, H., Galka, M., Iberg, A., Wang, Z., Li, L., Voss, C., Jiang, X., Lajoie, G., Huange, Z., Bedford, M. T., and Li, S. S. C. (2010) Systematic identification of methyllysine-driven interactions for histone and nonhistone targets. *J Proteome Res*, **9**:5827-5836.

Lorincz, M. C., Dickerson, D. R., Schmitt, M., and Groudine, M. (2004) Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. *Nat Struct Mol Biol*, **11**(11)1068-1075.

Loyola, A., Tagami, H., Bonaldi, T., Roche, D., Quivy, J. P., Imhof, A., Nakatani, Y., Dent S. Y. R., and Almouzni, G. The HP1-CAF1-SetDB1-containing complex provides H3K9me1 for Suv39-mediated K9me3 in pericentric heterochromatin. *EMBO Rep*, **10**:769-775.

Maison, C. and Almouzni, G. (2004) HP1 and the dynamics of heterochromatin inheritance. *Nature*, **5**:296-304.

Maksakova, I. A., Goyal, P., Bullwinkel, J., Brown, J. P., Bilenky, M., Mager, D. L., Singh, P. B. and Lorincz, M. C. (2011) H3K9me3 binding proteins are dispensable for SETDB1/H3K9me3-dependent retroviral silencing. *Epigenetics and Chromatin*, **4**(1):12.

Margueron, R. and Reinberg, D. (2010) Chromatin structure and the inheritance of epigenetic information. *Nat Rev Genet*, **11**:285-296.

Margueron, R., Justin, N., Ohno, K., Sharpe, M. L., Son, J., Drurylll, W. J., Voigt, P., Martin, S. R., Taylor, W. R., De Marco, V., Pirrotta, V., Reinberg, D., and Gamblin, S. J. (2009) Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature*, **461**:762-767.

Marzluff, W. F., Wagner, E. J., and Duronio, R. J. (2008). Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nat Rev Genet*, **9**:843-854.

Mathieu, O., Probst, A. V., and Paszkowski, J. (2005) Distinct regulation of histone H3 methylation at lysines 27 and 9 by CpG methylation in *Arabidopsis. EMBO J*., **24**(15):2783-2791.

Matsui, T., Leung, D., Miyashita, H., Maksakova, I. A., Miyachi, H., Kimura, H., Makoto, T., Lorincz, M. C., and Shinkai, Y. (2008) Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature*, **464**(7290):927-931.

Melcher, M., Schmid, M., Aagaard, L., Selenko, P., Laible, G., and Jenuwein, T. (2000) Structure-function analysis of SUV39H1 reveals a dominant role in heterochromatin organization, chromosome segregation, and mitotic progression. *Mol Cell Biol*, **20**:3728-3741.

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 cells. *PLoS Genet*, **6**(12):1-10.

Muller, J. and Kassis, J. A. (2006) Polycomb response elements and targeting of Polycomb group proteins in *Drosophila*. *Curr Opin Genet Dev*, **16**:476-484.

Mulligan, P., Westbrook, T. F., Ottinger, M., Pavlova, N., Chang, B., Macia, E., Shi, Y-J., Barretina, J., Liu, J., Howley, P. M., Elledge, S. J., and Shi, Y. (2008) CDYL bridges REST and histone methyltransferases for gene repression and suppression of cellular transformation. *Molecular Cell*, **32**(5):718-726.

Nekrasov, M., Klymenko, T., Fraterman, S., Papp, B., Oktaba, K., Köcher, T., Cohen, A., Stunnenberg, H. G., Wilm, M., and Müller, J. (2007) Pcl-PRC2 is needed to generate high levels of H3-K27 trimethylation at Polycomb target genes. *EMBO,* **26**:4078-4088.

O'Carroll, D. Erhardt, S. Pagani, M., Barton, S. C., Surani, M. A. and Junuwein, T. (2001) The

Polycomb-Group Gene *Ezh2* is Required for Early Mouse Development. *Mol and Cell Biol*., **21**(13):4330-4336.

Oktaba, K., Gutiérrez, L., Gagneur, J., Sengupta, A. K., Furlong, E. E. M., Müller, J. (2008) Dynamic Regulation by Polycomb Group Protein Complexes Controls Pattern Formation and the Cell Cycle in *Drosophila*. *Dev Cell*., **15**:877-889.

Onishi, M., Liou, G-G., Buchberger, J. R., Walz, T., and Moazed, D. (2007) Role of the conserved Sir3-BAH domain in nucleosome binding and silent chromatin assembly. *Molecular Cell*, **28**(6):1015-1028.

Pandey, R. R., Mondal, T., Mohammed, F., Enroth, S., Redrup, L., Komorowski, J., Nagano, T., Mancini-DiNardo, D., and Kanduri, C. (2008) *Kcnq1ot1* Antisense Noncoding RNA Mediates Lineage-Specific Transcriptional Silencing through Chromatin-Level Regulation. *Cell*, **32**:232-246.

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 and Cell Biol*, **27**(10):3769-3779.

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

Pasini, D., Cloos, P. A. C., Walfridsson, J., Olsson, L., Bukowski, J-P., Johansen, J. V., Bak, M., Tommerup, N., Rappsilber, J., and Helin, K. (2010) JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells. *Nature*, **464**:306-311.

Pauler, F. M., Sloane, M. A., Huang, R., Regha, K., Koerner, M. V., Tamir, I., Sommer, A., Aszodi, A., Jenuwein, T., and Barlow, D. P. (2009) H3K27me3 forms BLOCs over silent genes and intergenic regions and specifies a histone banding pattern on a mouse autosomal chromosome. *Genome Res*, **19**:221-233.

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.

Prasanth, S. G., Prasanth, K. V., Siddiqui, K., Spector, D. L., and Stillman, B. (2004) Human Orc2 localizes to centrosomes, centromeres and heterochromatin during chromosome inheritance. *EMBO J*, **23**(13):2651-2663.

Prasher, D. C., Eckenrode, V. K., Ward, W. W., Prendergast, F. G., and Cormier, M. J. (1992) Primary structure of the *Aequorea victoria* green-fluorescent protein. *Gene*, **111**(2): 229-233.

Probst, A. V., Dunleavy, E., and Almouzni, G. (2009) Epigenetic inheritance during the cell cycle. *Nat Rev Mol Cell Biol*, **10**:192-206.

Quivy, J. P., Gérard, A., Cook, A. J. L., Roche, D., and Almouzni, G. (2008) The HP1-p150/CAF-1 interaction is required for pericentric heterochromatin replication and S-phase progression in mouse cells. *Nat Struct Mol Biol*, **15**:792-797.

Quivy, J. P., Roche, D., Kirschner, D., Tagami, H., Nakatani, Y., and Almouzni, G. (2004) A CAF-1 dependent pool of HP1 during heterochromatin duplication. *EMBO J*., **23**:3516-3526.

Ringrose, L. and Paro, R. (2007) Polycomb/Trithorax response elements and epigenetic memory of cellular identity. *Development*, **134**:223-232.

Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Brugmann, S. A., Goodnough, L. H., Helms, J. A., Farnham, P. J., Segal, E., and Chang, H. Y. (2007) Functional Demarcation of Active and Silent Chromatin Domains in Human *HOX* Loci by Noncoding RNAs. *Cell*, **129**:1311-1323.

Sarma, K., Margueron, R., Ivanov, A., Pirrotta, V., and Reinberg, D. (2008) Ezh2 Requires PHF1 To Efficiently Catalyze H3 Lysine 27 Trimethylation In Vivo. *Mol and Cell Biol*, **28**(8):2718-2731.

Scharf, A. N. D., Barth, T. K., and Imhof, A. (2009) Establishment of Histone Modifications after Chromatin Assembly. *Nuc Acid Res*, **37**(15):5032-5040.

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*, **25**:3110-3122.

Schuettengruber, B., and Cavalli, G. (2009) Recruitment of the Polycomb group complexes and their role in the dynamic regulation of cell fate choice. *Development*, **136**:3531-3542.

Schultz, D. C., Ayyanathan, K., Negorev, D., Maul, G. G., and Rauscher, F. J. (2002) SETDB1: a novel KAP1-associated histone H3, lysine specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev*, **16**:919-932.

Sharma, S., Kelly, T. K., and Jones, P. A. (2010) Epigenetics in cancer. *Carcinogenesis*, **31**(1):27-36.

Sinclair, D. A. R., Mottus, R. C. and Grigliatti, T. A. (1983) Genes which suppress position-effect variegation in *Drosophila melanogaster* are clustered. *Mol Gen Genet*, **191**:326-333.

Sing, A., Pannell, D., Karaiskasis, A., Sturgeon, K., Kjabali, M., Ellis, J., Lipshitz, H. D., and Cordes, S. P. (2009) A vertebrate Polycomb response element governs segmentation of the posterior hindbrain. *Cell*, **138**(5):885-897.

Sparmann, A. and van Lohuizen, M. (2006) Polycomb silencers control cell fate, development and cancer. *Nat Rev Cancer*, **6:**646-656.

Sripathy, S. P., Stevens, J., and Schultz, D. C. (2006) The KAP1 corepressor functions to coordinate the assembly of *de novo* HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression. *Mol and Cell Biol*, **26**(22):8623-8638.

Strahl, B. D., and Allis, C. D. (2000) The language of covalent histone modifications. *Nature*, **403**:41-45.

Sung, S., and Amasino, R. M. (2004) Vernalization in *Arabidopsis thaliana* is mediated by the PHD finger protein VIN3. *Nature*, **427**:159-164.

Tagami, H., Ray-Gallet, D., Almouzni, G., and Nakatani, Y. (2004) Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell*, **116**:51-61.

Tagami, H., Ray-Gallet, D., Almouzni, G., and Nakatani, Y. (2004) Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell*, **116**:51-61.

Turek-Plewa, J. and Jagodzinski, P.P. (2005) The role of mammalian DNA methyltransferases in regulation of gene expression. *Cell and Mol Biol*, **10**:631-647.

Vermeulen, M., Eberl, H. C., Matarese, F., Marks, H., Denissov, S., Butter, F., Lee, K. K., Olsen, J. V., Hyman, A. A., Stunnenberg, H. G., and Mann, M. (2010) Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell*, **142**(6):967-980.

Wade, P. A., Pruss, D., and Wolffe, A. P. (1997) Histone acetylation: chromatin in action. *Trends Biochem Sci*, **22**(4):128-132.

Walter, W., Clynes, D., Tang, Y., Marmorstein, R., Mellor, J., and Berger, S. L. (2008) 14-3-3 Interaction with Histone H3 Involves a Dual Modification Pattern of Phosphoacetylation. *Mol and Cell Biol*, **28**(8):2840-2849.

Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R. S. and Zhang, Y. (2004) Role of histone H2A ubiquitination in Polycomb silencing. *Nature*, **431:**873-878.

Wang, J., Mager, J., Chen, Y., Schneider, E., Cross, J. C., Nagy, A. and Magnuson, T. (2001) Imprinted X-inactivation maintained by a mouse *Polycomb* group gene. *Nature*, **28:**371-375.

Weake, V. M., and Workman, J. L. (2008) Histone ubiquitination: triggering gene activity. *Mol Cell*, **29**:653-663.

Weiler, K. S. and Wakimoto, B. T. (1995) Heterochromatin and gene expression in *Drosophila*. *Annu. Rev. Genetics*, **29**:577-605.

White, D. E., Negorev, D., Peng, H., Ivanov, A. V., Maul, G. G. and Rauscher III F. J. (2006) KAP1, a novel substrate for PIKK family members, colocalizes with numerous damage response factors at DNA lesions. *Cancer Res*, **66**:11594-11599.

Woo, C. J., Kharchenko, P. V., Daheron, L., Park, P. J., and Kingston, R. E. (2010) A region of the human HOXD cluster that confers polycomb-group responsiveness. *Cell*, **140:**99-110.

Xu, M., Long, C., Chen, X., Huang, C., and Zhu, B. (2010) Partitioning of histone H3-H4 tetramers during DNA replication-dependent chromatin assembly. *Science*, **328**(5974):94-98.

Xu, M., Long, C., Chen, X., Huang, C., Chen, S., and Zhu, B. (2010) Partitioning of Histone H3-H4 Tetramers During DNA Replication-Dependent Chromatin Assembly. *Science*, **328**(5974):94-98.

Yang, T. T., Cheng, L., and Kain, S. R. (1996) Optimized codon usage and chromophore mutations provide enhanced sensitivity with the green fluorescent protein. *Nucleic Acids Res*, **24**(22): 4592-4593.

Zhou, W., Zhu, P., Wang, J., Pascual, G., Ohgi, K. A., Lozach, J., Glass, C. K. and Rosenfeld, M. G. (2008) Histone H2A Monoubiquitination Represses Transcription by Inhibiting RNA Polymerase II. *Cell*, **29**:60-80.