HARVARD UNIVERSITY
Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Division of Medical Sciences

in the subject of Biological and Biomedical Sciences

have examined a dissertation entitled

Locus-Dependent Epigenetic Inheritance of Polycomb-Mediated Gene Silencing

presented by Tiasha Ayumi Shafiq

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature: Carl D. Novina
Carl D. Novina (May 4, 2022 14:34 EDT)

Typed Name: Dr. Carl Novina

Signature: Ramesh Shivdasani

Typed Name: Dr. Ramesh Shivdasani

Signature: Stephen Buratowski

Typed Name: Dr. Stephen Buratowski

Signature:

Typed Name: Dr. Mo Khalil

Date: May 03, 2022

Locus-Dependent Epigenetic Inheritance of Polycomb-Mediated Gene Silencing

A dissertation presented

by

Tiasha Ayumi Shafiq

to

The Division of Medical Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biological and Biomedical Sciences

Harvard University

Cambridge, Massachusetts

May 2022

Dissertation Advisor: Dr. Danesh Moazed Tiasha Ayumi Shafiq

Locus-Dependent Epigenetic Inheritance of Polycomb-Mediated Gene Silencing

Abstract

During development, it is crucial that gene expression patterns, which define cell phenotypes and "epigenetic states", are stably inherited. The maintenance of epigenetic states involves changes in repressive histone posttranslational modifications (HPMs), such as H3K27me3 and H2AK119ub1, established and maintained by the Polycomb Group (PcG) of proteins, which assemble into Polycomb Repressive Complexes 1 and 2 (PRC1 and PRC2). One model of Polycomb epigenetic inheritance proposes a positive feedback read-write mechanism, in which the PRC2 histone-methyltransferase complex recognizes the modification on parentally inherited histones and catalyzes the same modification on adjacent newly deposited histones during replication. However, studies in fission yeast, *Drosophila* and mammalian cells suggest that locus-specific factors such as site-specific DNA-binding proteins also contribute to epigenetic inheritance. To investigate the mechanism of Polycomb-mediated epigenetic inheritance and the possible role of locus-specific factors in human cells, I inserted *5XtetO-CITRINE* reporters at two distinct types of loci, Polycomb target genes and housekeeping genes, together with expression and reversible binding of rTetR-CBX7 to establish ectopic Polycomb domains and silencing. I show that upon the release of the rTetR-CBX7 initiator, the Polycomb domain and the silent state are maintained at Polycomb target genes but not at housekeeping genes, suggesting that maintenance of the silent state is locus-dependent. Maintenance of the silenced state at Polycomb target genes is partially dependent on the hydrophobic cage domain of EED, the PRC2 subunit that recognizes H3K27me3, and requires the RING1A/B subunit of

PRC1, which catalyzes H2AK119ub. In addition, I show that maintenance is disrupted by point mutations in the DNA-binding domain of MTF2, a PRC2 accessory factor that binds to CpG-rich DNA sequences, supporting the contribution of DNA sequence to epigenetic inheritance. Furthermore, the loss of the silent state at housekeeping genes is partially reversed by DNA deletion that attenuate their transcription, suggesting a potential role for active transcription in counteracting epigenetic maintenance. These findings suggest that the heritability of Polycomb-mediated gene silencing requires DNA sequence-independent histone modification positive feedback but is also dependent on the ability of PRC2 accessory factors to bind to CG rich DNA and is opposed by proximal transcription at some loci.

Table of Contents

Acknowledgements

First and foremost I would like to thank my advisor, Dr. Danesh Moazed for his continued guidance and support during my PhD. Danesh is an exceptional scientist and mentor and he has helped me hone my scientific skills during my PhD. Throughout the challenging years of my PhD, he has been a great mentor and has taught me how to approach big picture and mechanistic questions in the epigenetics field. Beyond that, he has always been kind and supportive through the ups and downs of my project, and I am grateful to have had such a fantastic mentor who has challenged me to grow and foster analytical and critical thinking skills which I could apply to the next stage of my career.

Second, I would like to thank my Dissertation Advisory Committee members, Dr. Carl Novina, Dr. Bradley Bernstein and Dr. Yang Shi for their valued input on my project throughout the years.

Third, I would like to thank all the past and present members of the Moazed lab of whom I had the pleasure to overlap with. Every member of the lab has supported me by imparting their scientific knowledge and their willingness to help me on my journey as a scientist. Furthermore, they have fostered a fun and supportive work environment which elevated my PhD experience.

Finally, I would like to thank my parents who have always encouraged me to pursue my own journey in life, personally and professionally.

List of Figures

Chapter 1:

Chapter 2:

Chapter 3:

This Page is Left Intentionally Blank

Chapter 1: Introduction

1.1. Polycomb-Mediated Gene Silencing

The development of multicellular organisms relies on the establishment and maintenance of gene expression programs to achieve unique cellular identities. It is crucial that these gene expression patterns, which are referred to as the "epigenetic states" of the cell, are stably inherited. The maintenance of epigenetic states over numerous cell divisions involves a combination of mechanisms that can be broadly divided into trans-acting and cis-acting. In response to cues during embryogenesis, trans-acting mechanisms which include DNA-binding transcription factors, recruit the transcription machinery to shape gene expression profiles. Transcription factors also form positive feedback loops that maintain gene expression programs; however, epigenetic inheritance of gene expression and cell identity also involves changes in cis-chromatin and histone post-translational modifications (HPMs) that regulate transcription[1,2]. Some HPMs have been implicated in epigenetic inheritance of gene expression states by silencing key regulators outside of their proper domains of expression. Examples of HPMs that are involved in gene silencing are H3K9me (methylation on Histone 3 Lysine 9), H3K27me (methylation on Histone 3 Lysine 27) and H2AK119ub1 (mono-ubiquitination on Histone 2A Lysine 119). The latter two modifications are catalyzed by Polycomb Repressive Complexes, PRC1 and PRC2. Many of the genes that encode PRC1 and PRC2 subunits were first discovered in *Drosophila melanogaster* and were found to be necessary for maintenance of segmental identity by regulating the expression of the homeotic genes (Hox Genes)[3,4,5.] Since then, Polycomb Group (PcGs) genes have been shown to be conserved in plants, certain fungi, *C. elegans*, and mammals[6]. Studies over the past two decades have shown that PcGs regulate their target genes through modifying local chromatin structure and/or regulating higher-order

chromatin organization to direct epigenetic states. Mutations in PcGs in mammals result in developmental defects and embryonic lethality, have been implicated in human developmental diseases and are associated with many cancers[7–9].

1.2 The Polycomb Complexes: PRC2 and PRC1

The PcG proteins form two major complexes, PRC2 and PRC1. However, from *D. melanogaster* to mammals, PcGs have evolved and expanded their repertoire of subunits, which have functionally diverged.

1.2.1 PRC2

PRC2 is the sole histone-methyltransferase that mono, di and tri- methylates Histone H3 at Lysine 27[10]. H3K27me1 is poorly understood and is enriched at active genes, and is thought to facilitate the conversation to H3K27me2 and H3K27me3, with the latter associated with gene silencing[11]. PRC2 comprises of four core subunits, EZH1 or EZH2, SUZ12, EED, and RBBP4 or RBBP7. EZH1/2 is the catalytic subunit that transfers a methyl group from the cofactor S-adenosyl-L-methionine (SAM) to H3K27 through its SET domain. Enzymes such as EZH1/2 are commonly referred to as “writers”. EED, with its seven WD40 repeated domains can bind H3K27me3 (i.e. the “reader” of the complex)[12]. The binding of H3K27me3 to EED induces a conformational change in EZH2 by stabilizing its SET domain to activate its histone-methyltransferase (HMTase) function[13,14–16]. SUZ12 forms a scaffold for the complex that stabilizes EZH1/2, binds to EED and RBBP4/7, and interacts with accessory proteins that aid in the targeting and function of PRC2[17,18]. The composition of the auxiliary proteins defines two mutually exclusive subcomplexes of PRC2- PRC2.1 and PRC2.2 (Figure 1).

PRC2.1

PRC2.1 complex comprises of Polycomb like (PCL) proteins- 1, 2, or 3, also called PHF1, MTF2 and PHF19, respectively. The PCL proteins have an extended homology domain that can bind to unmethylated CG-rich DNA to help in PRC2 recruitment and can prolong the residency time of PRC2 bound to chromatin templates in vitro (See Chapter 3). In mouse Embryonic Stem Cells (mESCs), PRC2.1 is enriched at unmethylated CGIs (CpG islands).[19–22] Furthermore, PRC2.1 contains either PRC2-associated LCOR isoform 1 (PALI1), PALI2 or Elongin BC, and Polycomb repressive complex 2-associated protein (EPOP). EPOP and PALI1 enhance PRC2's HMTase function[23,24]. Further studies are needed to elucidate the role of EPOP and PALI1/2 in PcG's function.

PRC2.2

PRC2.2 is associated with accessory proteins JARID2 and AEBP2[25]. AEBP2 stabilizes PRC2 and through its binding to nucleosomes stimulates PRC2's HMTase activity. Additionally AEBP2 can bind methylated DNA, which is thought to be necessary for PRC2's role at imprinted genes, since methylated DNA generally repels PRC2/PCL binding[26,27]. AEBP2 also has the ability to bind to H2AK119ub1. JARID2 stimulates PRC2 activity by binding to nucleosomes through its JmjN domain. Additionally, methylation of JARID2 at Lysine 116 (K116me3) by EZH2 is recognized by EED to allosterically activate EZH2's HMTase activity[25,28]. Moreover, JARID2 also binds to H2AK119ub1 through its ubiquitin-interacting motif (UIM) to form a feedback loop between PRC1 and PRC2. Additionally, it exhibits DNA-binding activity through its AT-rich interaction domain (ARID). Therefore, JARID2 contributes to PRC2 function and targeting through multiple mechanisms[14,29,30,31].

PRC2.1 and PRC2.2's relationship:

PRC2.1 and PRC2.2 co-occupy at most target sites in mESCs, and the loss of one can compensate for the other to maintain H3K27me3 levels. This suggests that the subcomplexes can work via both synergistic and independent mechanisms[32,33]. Their independent and distinct roles become evident during mESC differentiation, whereby the loss of MTF2 results in enhanced and faster differentiation towards cell fates from all germ layers, while the loss of JARID2 predominantly leads to early differentiating precursors, with reduced efficiency towards mesendodermal lineage[34]. Additionally, recent evidence in human Embryonic Stem Cells (hESCs) suggests that the two complexes compete and are not functionally equivalent[35].

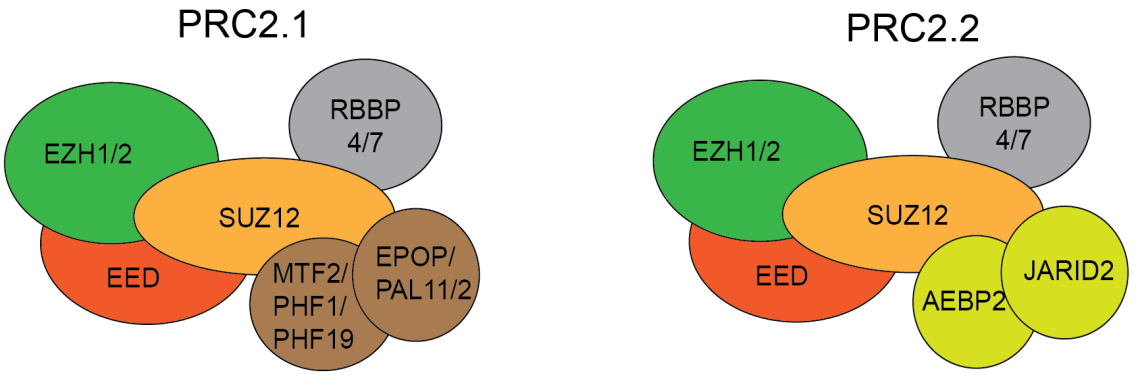


Figure 1. Components of PRC2.1 and PRC2.2

The core of Polycomb Repressive Complex 2 (PRC2) comprises of a SET (Su(var)3-9, Enhancer-of-zeste and Trithorax)-domain containing enhancer of zeste 1/2 (EZH1 or EZH2) protein, together with embryonic ectoderm development (EED) and the suppressor of zeste 12 (SUZ12). The N-terminal part of SUZ12 associates with retinoblastoma-binding protein 4 or 7 (RBBP4 or RBBP7). PRC2 is subdivided into PRC2.1 (left) and PRC2.2 (right). PRC2.1 complexes contain a Polycomb-like (PCL) subunit (MTF2/PHF1/PHF19) and Elongin BC and Polycomb repressive complex 2-associated protein (EPOP) or PRC2-associated LCOR isoform 1 (PALI1 or PALI2). PRC2.2 complexes contain adipocyte enhancer binding protein 2 (AEBP2) and Jumanji and AT-rich interaction domain containing 2 (JARID2).

Image adapted from Mierlo et al, 2019, Trends in Cell Biology[36].

1.2.2 PRC1

The core PRC1 is comprised of RING1A/RING1B, the catalytic subunit that catalyzes H2AK119ub1 and one of the six Polycomb group ring finger (PCGF) proteins (PCGF1, 2, 3, 4, 5, or 6)[37,38,39]. The RING1A/B and the PCGF proteins dimerize to facilitate their interaction with an E2 conjugating enzyme to enable mono-ubiquitylation on H2AK119. Depending on the PCGF subunit and its association with distinct proteins, PRC1 complexes are divided into two types- variant or non-canonical PRC1 (vPRC1) and canonical PRC1 (cPRC1). (Figure 2)

vPRC1:

vPRC1 comprises of RING1A/B, PCGF1-6, RYBP or its paralogue YAF2, and various auxiliary subunits depending on the PCGF present in the complex. This complex has been shown to have a higher catalytic activity in vitro due to RYBP enhancing the activity of E3 ligase relative to cPRC1[37,40]. RYBP also has the ability to bind to H2AK119ub1, potentially forming a read-write feedback mechanism of vPRC1-H2AK119ub1[41].

cPRC1:

cPRC1 is comprised of RING1A/B and CBX proteins (2/4/6/7/8), PCGF2 or 4 and other auxiliary subunits. cPRC1 might have a limited ability to mono-ubiquitinate H2AK119, but it has been shown to be involved in chromatin compaction through chromatin conformation capture-based approaches (i.e HiC) and has chromatin compaction activity in vitro[42–44]. The CBX2 subunit can also undergo liquid-liquid phase separation (LLPS) in vitro and forms nuclear

condensates[45,46]. This phenomenon is hypothesized to enhance Polycomb activity by concentrating PcGs at their target loci while restricting other proteins. Further studies are required to elucidate the functional consequences of compaction and LLPS in gene silencing.

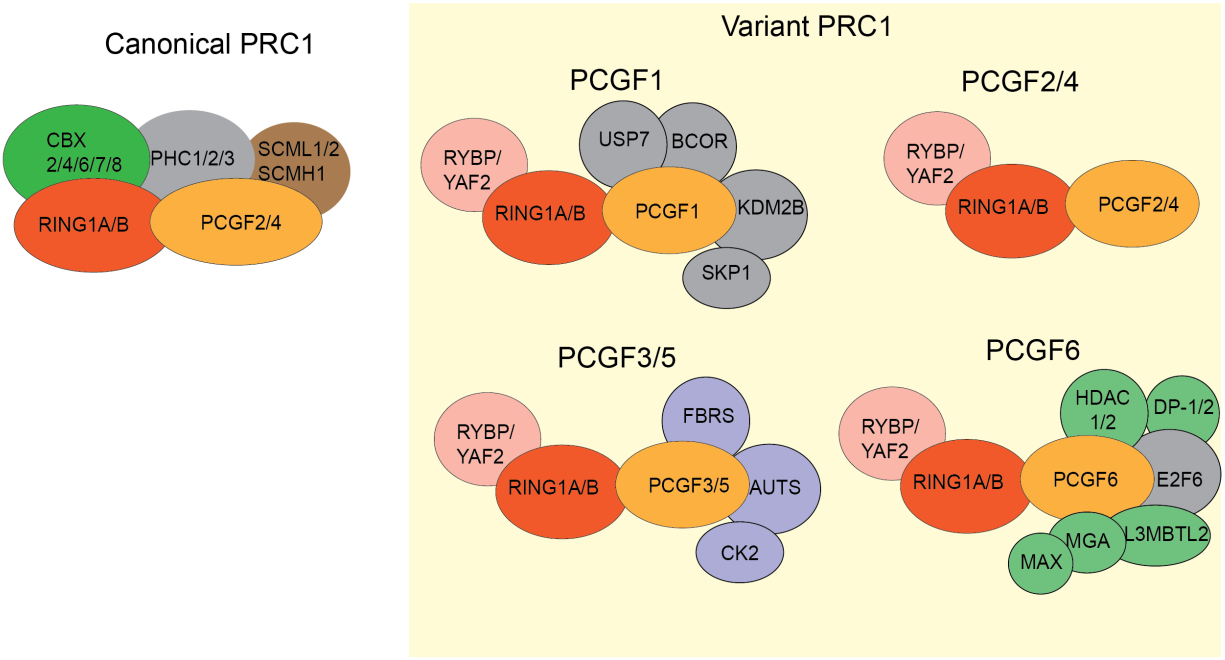


Figure 2. Components of Canonical and Variant PRC1.

The core of Polycomb repressive complex 1 (PRC1) contains RING finger protein 1 (RING1A/B) and one of six Polycomb group RING finger (PGCF) proteins (PCGF1–6). They interact with a range of accessory subunits, giving rise to biochemically distinct PRC1 complexes. Canonical PRC1 (cPRC1) complexes (left) comprises of PCGF2 or PCGF4, and include a chromobox protein (CBX2, 4, 6, 7 or 8) and Polyhomeotic (PHC) protein (PHC1, 2 or 3). In some cases, cPRC1 complexes also contain an SCM protein, SCM like 1 (SCML1 or 2) or sex combs on midleg homolog 1 (SCMH1). Variant PRC1 (vPRC1) complexes (right) comprise of all six PCGFs and contain RING and YY1 binding protein (RYBP) or YY1-associated factor 2 (YAF2). The identity of the PCGF protein dictates the assembly of other accessory subunits, resulting in a number of distinct vPRC1 complexes. AUTS2, autism susceptibility protein 2; BCOR, BCL6 corepressor; CK2, casein kinase 2; DP-1, dimerization partner 1 (also known as transcription factor Dp-1); E2F6, transcription factor E2F6; FBRS, fibrosin; HDAC1, histone deacetylase 1; KDM2B, lysine-specific demethylase 2B; L3MBTL2, lethal(3)malignant brain tumour-like protein 2; MAX, MYC-associated factor X; MGA, MAX gene-associated; SKP1, S-phase kinase-associated protein 1; USP7, ubiquitin carboxyl-terminal hydrolase 7.

Image adapted from Blackledge et al, 2021, Nat Rev Mol Cell Biol.[47]

1.2.3 Interplay of PRC1 and PRC2 and their Feedback Loops

PRC1 and PRC2 are largely associated with overlapping genomic regions, and the interplay between the two complexes play a role in gene silencing through the inhibition of the transcriptional machinery. PRC1 and PRC2 have independent functions in gene repression, but can synergistically act through linked feedback mechanisms to maintain robust Polycomb domains (Figure 3).

In ESCs, developmentally regulated genes (for example Hox genes) are bivalent[48]. They are marked by both H3K4me3 (an active HPM) and H3K27me3 and H2AK119ub1. vPRC1 and to some extent cPRC1 can deposit H2AK119ub1. vPRC1-RYBP can bind H2AK119ub1, which may help amplify H2AK119 ubiquitylation through its read-write positive feedback mechanism. Studies have shown that the catalytic activity of PRC1 and hence H2AK119ub1 is necessary for gene repression in mESCs.[49,50] The mechanism of H2AK119ub1-mediated transcriptional repression is still unknown. Current hypotheses include the recruitment of downstream proteins that repress transcription, the steric inhibition of the transcriptional machinery, and/or preventing transcription initiation[51–53].

H2AK119ub1 can be recognized and bound by PRC2.2's JARID2 and AEBP2. This provides a physical link between PRC2-H3K27me3 and PRC1-H2AK119ub and links positive feedback loops that may reinforce each other. This is suggested by experiments whereby loss of vPRC1 leads to reductions in H3K27me3 levels[49,50,54]. In *D. melanogaster*, The Scm protein can physically associate with PRC1 and PRC2, bridging the two complexes independently of HPMs[55], but this has not been demonstrated in mammalian cells yet.

It has been shown that cPRC1 can compensate for the loss of vPRC1 (abolishment of RYBP/YAF2) in mESCs and vice versa[41]. Even though in vitro, cPRC1 has been shown to have

low E3 ubiquitin ligase activity for H2AK119 relative to vPRC1, in vivo cPRC1 could possibly have the ability to catalyze H2AK119ub1 just as efficiently, as shown in one study[54,56]. Independently of H2AK119ub1, cPRC1 can compact chromatin and undergo LLPS. There is evidence that in *D. melanogaster* that PRC1 and RING1B mediate chromatin compaction and repression of the *Hox* genes independently of H2AK119ub1[57]. This suggests that under certain conditions cPRC1 might have important repressive activities that are H2AK119ub1-dependent or -independent in mESCs[41]. Furthermore, cPRC1 can recognize H3K27me2/3 through the chromodomain of CBX proteins, resulting in a link between PRC2-H3K27me3 and PRC1-H2AK119ub1[58].

Loss of PRC2 has minimal effect on gene repression in mESCs but leads to defects during their differentiation, emphasizing the role of PRC2 and H3K27me3 in maintaining gene expression states in differentiated cells[59,60]. In *D. melanogaster*, it has been shown that H3K27me3 conferred by PRC2 is necessary for proper development[61,62]. In mammals, reducing the levels of H3K27me3 also leads to developmental defects by de-repression of Polycomb target genes[6]. It is possible that the H3K27me3 mark recruits downstream histone deacetylases to counteract the transcriptional machinery or could be bound directly by cPRC1 leading to compaction and LLPS to limit the transcriptional machinery. Additionally, EZH1 has been shown in vitro to compact chromatin[63]. PRC2 also has non-histone substrates that could potentially contribute to gene repression[64]. Therefore, vPRC1, cPRC1 and PRC2 have distinct functions but can reinforce each other to promote gene silencing in mESCs[41].

The role of H2AK119ub1 and H3K27me3 in differentiated cell types has not been extensively studied. Therefore, further studies need to be conducted to understand the role of

PRC1's H2AK119ub1 dependent or independent properties and PRC2's H3K27me3 in differentiated cells.

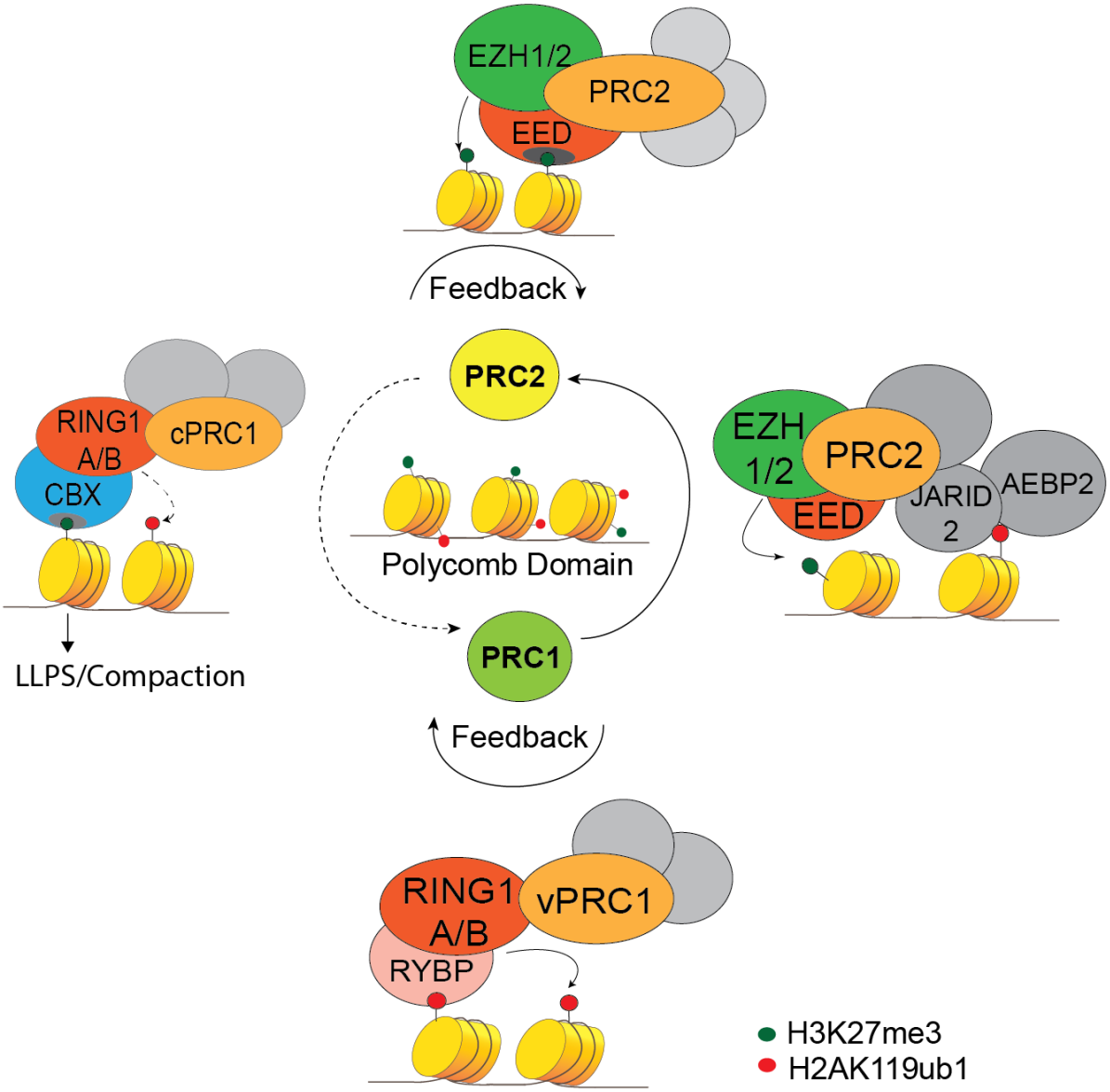


Figure 3. Feedback loops of PRC1 and PRC2.

H2AK119ub1 is recognized by the RING and YY1 binding protein (RYBP) subunit or by the YY1-associated factor 2 (YAF2) of variant PRC1 (vPRC1) complexes to allosterically activate RING1A/B to deposit H2AK119ub1. This creates a positive feedback mechanism that reinforces vPRC1 binding and amplifies H2AK119ub1. (Bottom)

H2AK119ub1 is recognized by Jumanji and AT-rich interaction domain containing 2 (JARID2) and adipocyte enhancer binding protein 2 (AEBP2) subunits of PRC2.2. The recruited PRC2 can

Figure 3 (Continued)

now catalyze H3K27me3. Therefore, H2AK119ub1 facilitates communication between PRC1 and PRC2 in Polycomb chromatin domains. (Right)

H3K27me3 is recognized by the embryonic ectoderm development (EED) subunit of PRC2, which allosterically activates HMTase activity of enhancer of zeste 2 (EZH2). This creates a feedback mechanism that reinforces PRC2 binding and amplifies H3K27me3. (Top)

H3K27me3 is also recognized by the chromobox (CBX) subunit (CBX2, 4, 6, 7 or 8) of canonical PRC1 (cPRC1) complexes. Although cPRC1 complexes are thought to be less catalytically active than vPRC1 complexes in vitro (as indicated by the dotted arrow), in some contexts cPRC1 can catalyze H2AK119ub1. Therefore, H3K27me3 can facilitate communication between PRC2 and PRC1 in Polycomb chromatin domains. (Left)

Image adapted from Blackledge et al, 2021, Nat Rev Mol Cell Biol.[47]

1.3 Models for Epigenetic Inheritance of Polycomb-Mediated Gene Silencing

Once ESCs differentiate, bivalent genes lose their Polycomb modifications and become active in some cell lineages while conversely other bivalent genes maintain their Polycomb modifications and lose their markers of active transcription in other lineages. Therefore, the Polycomb domains of H2AK119ub1 and H3K27me3 need to be maintained over multiple cellular divisions in actively dividing lineage committed cells. DNA replication will dilute the HPMs at each replication cycle but mechanisms have evolved to restore them on newly deposited unmodified nucleosomes to maintain the silent state. Several models for epigenetic inheritance of Polycomb-mediated gene silencing have been proposed. The earliest models are based on the “read-write” mechanism and do not take into account the possible contribution of underlying DNA sequences to inheritance. Subsequent models have been proposed whereby the necessity of DNA sequences and antagonism between opposing transcriptional states have been emphasized.

1.3.1. Model of Read-Write Positive Feedback Mechanism

The read-write positive feedback model proposes that the feedback loops that were described above are sufficient for epigenetic inheritance. In case of PRC2, the read-write ability is conferred through its EED ("reader" of H3K27me) and EZH2 ("writer") subunits while for PRC1 it is RYBP ("reader") and RING1A/B ("writer"). The reader subunits would read the parental HPMs during replication, resulting in allosterically activation of the writer subunits to catalyze modifications on newly deposited nucleosomes[1]. There is evidence that supports this model but shows that read-write is only sufficient for epigenetic inheritance of H3K9me under certain circumstances in fission yeast *(S. pombe*). In *S. pombe*, upon artificial establishment of a repressive H3K9me domain at an ectopic locus, it was observed that the read-write feedback mechanism of the H3K9me3 HMT was sufficient to maintain silencing of the domain upon release of the initial inducer only when a putative histone demethylase was deleted[65,66]. In mammalian cells, a similar approach was undertaken to artificially establish a domain of H3K27me3/H2AK119ub1 at an ectopic locus at a gene desert, that is devoid of active transcriptional input and CGIs. Under these conditions, sequence-independent and read-write dependent inheritance of Polycomb-mediated gene silencing was observed[56].

1.3.2. Models of Locus-Dependent Epigenetic Inheritance

1.3.2.1 DNA-Sequence Dependent Epigenetic Inheritance:

<u>Inheritance of H3K9me Mediated Heterochromatin in *S.pombe*</u>

In *S. pombe*, cis-DNA sequences have been identified that bind specific DNA binding factors at heterochromatic loci to maintain epigenetic inheritance of H3K9me. Deletion of these DNA sequences termed "maintainers" leads to loss of inheritance of H3K9me. The maintainer

has been proposed to function in epigenetic inheritance by working together with preexisting H3K9me3 to recruit the H3K9 methyltransferase and/or by recruiting downstream proteins that create a permissible chromatin environment for propagation of H3K9me[67,68].

Inheritance of Polycomb-Mediated Gene Silencing in *D. melanogaster*:

In *D. melanogaster,* it has been observed that the read-write positive feedback loops are insufficient for epigenetic inheritance of Polycomb-mediated gene silencing. In flies, the underlying DNA sequence plays a necessary role in the maintenance of H3K27me. The specific DNA sequences that promote Polycomb recruitment and inheritance have been identified and named "Polycomb Response Elements" (PREs). The PREs are 1-3kb nucleosome depleted DNA sequences that contain motifs for sequence-specific DNA-binding proteins to recruit PcGs[69]. Deletion of PREs results in loss of inheritance of H3K27me3 leading to gene de-repression[70,71]. This suggests that similar to *S. pombe,* specific DNA sequences can play a role in epigenetic inheritance of gene silencing.

Inheritance of Polycomb-Mediated Gene Silencing in Mammalian Cells:

Role of CGIs in Recruitment of Polycomb complexes

The quest to find "PRE"-like sequences and specific DNA binding factors that would be conserved from *D. melanogaster* to mammalian cells has not yielded promising results. Most of the specific DNA binding factors are not conserved, and the ones which are conserved have functionally divergent properties[6]. Sequences containing DNA Binding protein motifs associated with PcG silencing have also not been found. However, it was observed that H3K27me3

localization correlates with CGIs that are short 1-2kb CpG-rich DNA[72]. CGIs occur at 70-80% of genes in the mammalian genome and most are methylated in constitutive heterochromatin[26]. However, the CGIs that are associated with Polycomb seem to be devoid of DNA methylation. Mendenhall et al. showed that artificial CG-rich element devoid of activating Transcription Factor (TF)-binding sites can ectopically recruit PcG proteins, suggesting that CGIs have PRE-like properties in mammalian cells[73]. Recently, accessory proteins of PRC2 and PRC1, such as the PCLs/JARID2 and KDM2B (which interacts with vPRC1) have been shown to be enriched at CG-rich DNA in vivo and bind CG-rich DNA in vitro, which supports the hypothesis that CGIs contribute to the recruitment of PRCs to target sites[19,21,74–76]. However, there are constitutively active genes that also have CGIs devoid of methylation but are not targeted by Polycomb. This observation supports the speculation that active transcription and its associated chromatin modifications counteract Polycomb silencing.

Besides CGIs, sequence-specific DNA-binding factors such as MAX, MGA, E2F6, DP1/2 seem to bind and target PCGF6-vPRC1 to certain sites in the genome[77–79]. Other examples include PCGF3-vPRC1 that can recognize some of its target sites in embryonic stem cells (ESCs) by interacting with the transcription factors upstream stimulatory factor 1 (USF1) and USF2[72,78]. REST, RUNX1 and SNAIL1 are DNA Binding proteins that associate with PRC1/PRC2 for its targeting in certain cell types[80]. Therefore, in a few cases, Polycomb recruitment appears to depend on interactions with site-specific DNA binding proteins.

Are CGIs necessary for Inheritance of Polycomb Domains?

Experimental evidence has suggested that CGIs are important for recruitment of Polycomb complexes as described above; however, an open question is whether CGIs are necessary for the inheritance of Polycomb domains once they have been established.

1.3.2.2. Transcriptional Activity Impacts Epigenetic Inheritance

Antagonistic Transcriptional Activity on Formation of Polycomb Domains

There is evidence that transcriptional activity restricts the activity of Polycomb complexes. This is directed by transcription factors (master regulators) that recruit the Trithorax complex (includes MLL1-2, SETD1A-B) to create transcription-permissive, H3K4me3-containing domains through feedback mechanisms that are linked to active transcription[6]. Mechanistically, active chromatin marks which include H3K4me3, H3K36me3, H3K27ac amongst others have been shown to counteract PRC2 activity[81].

Initial studies using RNA polymerase II (Pol II) inhibitors found that upon transcription inhibition, Polycomb was recruited to CGI-containing genes that were previously highly transcribed[27]. Furthermore, if transcription termination signal is inserted adjacent to an active CGI-associated promoter, it stabilizes PRC2 binding and H3K27me levels[82]. It has been shown that CG-rich elements devoid of activating transcription factor-binding sites can ectopically recruit PcG proteins[73]. Ectopic H3K27me3 enrichment was also observed at CGIs that are devoid of transcriptional activators or proximal enhancer elements[26,83]. These experiments suggest that transcriptional activity counteracts Polycomb domain formation at CG-rich domains.

Antagonistic Transcriptional Activity Influences Polycomb Inheritance:

In a recent study by Holoch and colleagues, transient disruption and restoration of EZH1/2 in a differentiated cell line was shown to lead to the loss of H3K7me3 Polycomb domains, but this loss was reversible at genes that were generally lowly transcribed and not at those that were highly transcribed[84]. This suggests that high transcriptional activity can prevent the re-establishment of Polycomb-mediated silencing and contributes to epigenetic memory at Polycomb target genes. Furthermore, transient activation of silenced Polycomb targets led to stable loss of silencing and epigenetic switching, mimicking activation of bivalent Polycomb target genes during ESC differentiation. These observations are consistent with the irreversible activation of subsets of Polycomb target genes in certain lineages during development but do not address possible locus-specific contributions to the mechanism of epigenetic inheritance via read and write-based mechanisms.

Bistable model of Chromatin: Incorporating the Contributions of Locus-Dependent cis and trans Factors to Epigenetic Inheritance

Mathematical modeling studies have recently emerged that incorporate the role of cis pathways that include Polycomb feedback loops and CGIs that bind DNA binding proteins and trans pathways that include antagonistic transcription factor networks to determine the epigenetic memory of a target gene to regulate its expression[85–89].

The model proposes that the antagonistic relationship between the Polycomb and Trithorax complex should essentially create two bistable states, active and repressive (Figure 4). When the target gene is in a predominantly repressive state with Polycomb domains that are maintained by PcGs feedback loops and contact with CGIs, it creates an epigenetic memory that buffers against low-level or spurious transcription activation signals. Similarly, when the gene is

in an active state with the Trithorax complex creating a transcriptionally active domain through activity of Pol II and transcription factor networks, it counteracts Polycomb activity. This mechanism can create bistable chromatin states that can switch depending on the strength of the Polycomb and Trithorax complexes. This mechanism ensures that during propagation of a Polycomb domain, stochastic and transient activation signals do not lead to a switch in the epigenetic state. Only if the activation signal crosses a certain threshold, would the gene switch to an active state.

There is experimental evidence that supports this bistable chromatin states and switching model from plants, but evidence in mammalian cells is limited[84,88]. Experimental evidence for bi-stability has been observed for decades in the variegation of reporter gene expression, whereby a reporter gene placed near a heterochromatic region shows spontaneous silencing in a subset of cells that can be stably inherited during mitosis[90,91]. In *Arabidopsis,* it has been shown that during the vernalization process (the perception and memory of winter in plants), the *FLC* reporter in the cold is PRC2-silenced and slowly gains H3K27me3[86]. This silenced epigenetic state can then be stably inherited in a subset of cells after switching to warm conditions. This study demonstrates that the mechanism of inheritance of Polycomb-mediated gene silencing in plants is dependent on the length of the cold that determines the accumulation and spreading of H3K27me3 in cis. Therefore, epigenetic memory has been demonstrated in plants that is dependent on a histone modification. Furthermore, in another study in *Arabidopsis,* the authors added different fluorescent reporters to the two copies of the FLC genes that were silenced during the cold. When switched to warm conditions, one copy could retain silencing while the other copy was active in the same cell[91]. This provided evidence that epigenetic memory is

stored in cis. However, both studies do not address other factors, such as DNA sequences or opposing cis transcriptional activity that could be required for inheritance of the silenced state.

In mammalian cells, as previously described, Holoch et al, showed stable epigenetic switching of a Polycomb silenced gene when transiently activated[84]. They proposed that that the mutual antagonism between PRC2 activity and cis-acting transcriptional inputs (in the absence of trans-acting gene regulatory network) determines the expression state of the gene which can then be stably inherited. However, the experimental design of this study did not allow the role of CGIs or other locus-specific factors to the inheritance of a silenced state to be studied. Therefore, experimental evidence that takes into account the role of PRC1/PRC2 activity, CGIs, and antagonistic transcriptional activity in formation of bi-stable Polycomb domains is still lacking.

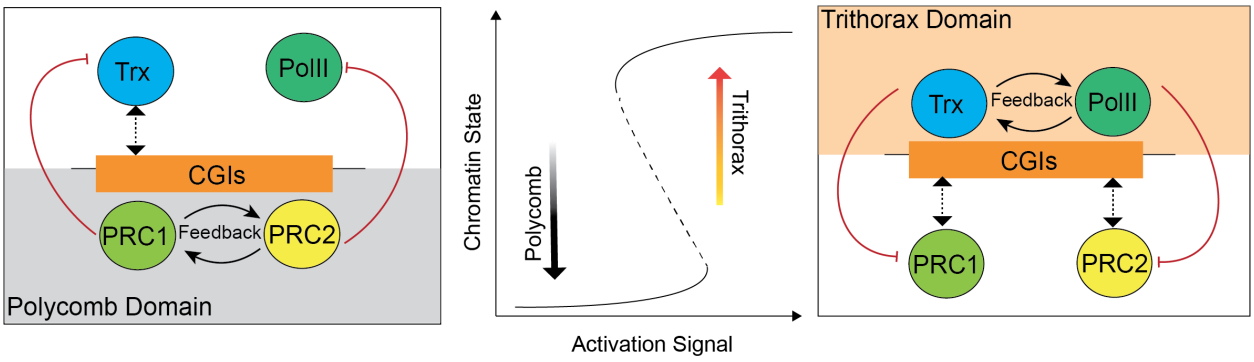


Figure 4. Chromatin Bi-stability model.

This model proposes that when transcriptional activity is low, the feedback loops of PRC1 and PRC2 cooperatively work together through contact with CGIs (CpG rich DNA) to form Polycomb domains in silencing the gene. This prevents the activity of RNA Polymerase II (PolII) and Trithorax group of proteins (Trx)(Left). On the other hand, high persistent transcriptional signals can direct the feedback of PolII and Trx to inhibit the activity of PRC1/PRC2, leading to formation of an active state (Right). This mutual antagonism of PcGs and Transcriptional activity can form bi-stable expression states, silenced or active. This mechanism could prevent inappropriate switching of epigenetic state if exposed to stochastic transcriptional signal.

Image adapted from Blackledge et al, 2021, Nat Rev Mol Cell Biol[47].

1.4 Current Challenges and Overview of this Dissertation

Over the last few decades, a substantial amount of knowledge has been gained about the Polycomb system in mammalian cells. However, many of the studies were performed in stem cells and early development context to study the recruitment of PcGs and the formation and establishment of Polycomb modifications. CGIs and antagonistic transcriptional activity have been proposed to mediate the establishment of Polycomb domains, but their role in the maintenance of Polycomb silencing is poorly understood. In general, a limited number of studies have been conducted to understand the role of the Polycomb complexes in maintenance of gene silencing in cells that have differentiated and committed to a specific lineage. This dissertation is focused on the mechanism of maintenance or inheritance of Polycomb-mediated gene silencing in a differentiated cell line. I have used an experimental strategy to address the following questions: are cis elements, such as CGIs and antagonistic transcriptional networks, necessary for maintenance of an "epigenetic memory"? Or is the self-reinforcing positive feedback read-write mechanism sufficient to restore H3K27me3 and H2AK119ub1 after each replication cycle?

Studies by others have demonstrated that DNA binding proteins that bind CG-rich DNA in vitro could increase residency time of PRC2 on chromatin[92], but experimental evidence is lacking as to the contribution of these DNA binding proteins, in particular their ability to bind CG-rich DNA to maintain Polycomb domains and silencing. Additionally, modeling studies have led to the hypothesis that the antagonistic relationship between Polycomb and active transcriptional machinery is necessary to maintain normal cell identity in mammals. However, experimental evidence showing that Polycomb feedback loops, CGIs and proteins that bind to them, and antagonistic transcriptional activity play necessary roles in the maintenance of Polycomb domains is lacking.

Therefore, to specifically study the inheritance of Polycomb-mediated gene silencing, I utilized an inducible system to ectopically establish Polycomb domains of H3K27me3 and H2AK119ub1 at reporter genes that were inserted adjacent to endogenous active Polycomb target and housekeeping target genes. This allowed me to uncouple the establishment/formation of the Polycomb domain from its maintenance in differentiated cells. In Chapter 2, I demonstrate that silencing of reporters and Polycomb domains at Polycomb target genes is heritable and displays a slower decay rate compared to silencing of reporters at housekeeping targets which is rapidly lost. These findings suggest that there is a locus-dependent contribution to the inheritance of Polycomb-mediated gene silencing. In Chapter 3, I investigate the contributions of this locus-independent mechanism, i.e., the roles of the positive feedback read-write mechanism, and locus-dependent factors such as DNA binding proteins, CGIs, and proximal transcriptional activity to the inheritance of Polycomb domains. I show that the PRC2 read-write mechanism contributes to epigenetic inheritance, but surprisingly is not essential for inheritance over shorter time periods. Furthermore, PRC1, the CGI-DNA binding ability of the PRC2 accessory subunit MTF2, and PHF1 are necessary for maintenance of Polycomb domains. Finally, I show that proximal transcriptional activity plays a role preventing the epigenetic inheritance of Polycomb silencing. In Chapter 4, I present future directions to further investigate the inheritance of Polycomb-mediated gene silencing. The thesis concludes with the methods section in Chapter 5.

Chapter 2: Developing Tools to Study Possible Locus-Dependent Effects on Epigenetic Inheritance

2.1 Introduction

To specifically study the possible role of locus-dependent factors in inheritance of Polycomb domains, I set up a system to uncouple the establishment and maintenance phases of silencing at two types of genes: genes that are targeted for silencing by Polycomb in some cell types and housekeeping genes which are never targeted by Polycomb. This strategy allowed me to ask whether specific sequences or features of Polycomb target genes contribute to the epigenetic maintenance of silencing.

Several previous studies have attempted to investigate the establishment and maintenance of an ectopically induced domain of heterochromatin, including H3K9me and H3K27me domains. Ragunathan et al. inserted a *10XtetO-reporter* sequence at a euchromatic site to establish an ectopic domain of H3K9me through tethering of TetR-Clr4 in fission yeast (Clr4 is the H3K9me HMT in the fission yeast *S. pombe*). Upon release of the TetR-Clr4, they observed inheritance of the H3K9me domain, but only in cells in which the putative histone demethylase *Epe1* was deleted[65]. This observation suggested that DNA sequence-independent inheritance of H3K9me at an ectopic locus was conditional and gave rise to the possibility that specific DNA sequences at endogenous heterochromatic loci contribute to epigenetic inheritance. Wang et al. then inserted a *10XtetO-reporter* cassette at an endogenous heterochromatic locus in fission yeast to establish an inducible domain of H3K9me independently of endogenous establishment factors. Upon release of the TetR-Clr4, they observed inheritance of the H3K9me domain even in *epe1*$^{+}$ cells, demonstrating that specific DNA sequences contribute to epigenetic inheritance. In fact, deletion of specific DNA sequences that bind DNA binding proteins resulted in loss of

inheritance[67,68]. In mammalian cells, Hansen et al. showed that an inducible GAL4-EED when bound to a reporter locus could deposit H3K27me3, which was maintained when GAL4-EED expression was turned off. However, the tetracycline inducible GAL4-EED used in this study displayed leaky expression that could have contributed to maintenance[93]. Bintu et al. and Blackledge et al. successfully used TetR-EED and TetR-PRC1 subunits to establish H3K27me3/H2AK119ub1 domains[75,94]. In the former study, rTetR-EED was tethered to a *5XtetO-CITRINE* reporter locus on a human artificial chromosome (HAC, devoid of any active genes) in Chinese Hamster Ovarian Cells (CHO). The authors observed maintenance of H3K27me3 upon release of rTetR-EED. In the latter study, a *tetO* sequence was inserted in a gene desert region in mESCs to establish a domain of H3K27me3 and H2AK119ub1 using various TetR-PRC1 fusion proteins, but inheritance of the silenced state following the release of TetR-PRC1 was not tested. Moussa et al, inserted a *7XtetO-GFP* reporter in a gene desert in mESCs to establish an ectopic domain of H3K27me3 and H2AK119ub1 through tethering of rTetR-CBX7. They observed inheritance of the domain upon release of the initiator which they concluded to be DNA sequence-independent but relies on the positive feedback loops of PRC1 and PRC2[56].

Therefore, these previous studies have successfully established an ectopic domain of H3K27me3 and H2AK119ub1 to study the mechanism of inheritance. However, investigating possible locus-dependent contributions to epigenetic inheritance requires an experimental strategy that can compare the requirements for epigenetic inheritance at evolved Polycomb target genes versus non-target genes. Toward this goal, I designed experiments to establish ectopic Polycomb domains at two distinct types of loci: genes that are targeted for silencing by Polycomb in some cell types (Polycomb target genes), but are expressed in the cell line used for

my experiments, and housekeeping genes that are not Polycomb targets and are expressed in all cell line types. Thus, I integrated a *5XtetO-pEF1-H2B-CITRINE* reporter at multiple Polycomb target and housekeeping genes. For Polycomb target loci, I chose *WT1, EN2, HOXD11,* and *HOXB4*, which are expressed in HEK293FT cells. For the housekeeping target loci, I used *TFRC* and *B2M*. In addition, I attempted to create cell lines with the reporter inserted at two distinct gene desert regions. Using lentiviral transduction, I expressed a reversible rTetR-CBX7 in the above reporter cell lines, which upon doxycycline addition would bind to the *tetO* sites and can be released from *tetO* by removal of doxycycline from the growth medium. I observed maintenance of silencing of the reporter inserted at Polycomb target loci, *WT1, EN2* and *HOXD11,* and corresponding maintenance of H3K27me3/H2AK119ub1 domains, upon release of the rTetR-CBX7. However, at the housekeeping target loci, *TFRC* and *B2M*, silencing of the reporter and H3K27me3/H2AK119ub1 domains were rapidly lost upon release of the rTetR-CBX7. These findings at Polycomb versus non-target genes suggest that there is a locus-dependent contribution to the inheritance of Polycomb-mediated gene silencing. This system also provides tools to test the possible role of locus-dependent factors in the epigenetic inheritance of the Polycomb-mediated gene silencing (Chapter 3).

2.2 Results

2.2.1 Site-specific Integration of the *5XtetO-H2B-CITRINE* and expression of rTetR-CBX7

In order to test for locus-dependent epigenetic inheritance of Polycomb-mediated gene silencing, *5XtetO* with a human *EF1* promoter and *H2B-CITRINE* was independently integrated a few kb upstream of four Polycomb target genes, *WT1, EN2, HOXD11* and *HOXB4* and two housekeeping target genes, *TFRC* and *B2M* in a human embryonic kidney cell line, HEK293FT. These Polycomb target genes are actively expressed in HEK293FT cells, but both WT1 and EN2 are silenced in liver and skin lineage cells, while HOXD11 is silenced in brain tissue and HOXB4 in retinal tissue[95,96]. These genes therefore have the ability to undergo Polycomb-mediated gene silencing in specific cell lineages that requires them to be transcriptionally shut off[6,43,97,98]. The reporter lines were genotyped and southern blot was carried out to ensure proper and single integration at the desired site.

The H3K27me3/H2AK119ub1 domain was established by tethering the reverse Tet repressor (rTetR) protein fused to several Polycomb group proteins, which recruit the PRC2 or PRC1 complex, to a *tetO* array. rTetR only binds the *tetO* array in the presence of doxycycline and is released upon removal of doxycycline from the culture medium (Figure 5, TET-ON). This allows me to control the association of the initiator with DNA and assess silencing and H3K2me/H2AK119ub1 with and without DNA sequence-dependent initiation. This system can therefore separate the establishment and inheritance phases of silencing, allowing me to specifically study the requirements for inheritance. Several mCherry-rTetR fusion proteins were tested for their silencing capabilities, including PRC1 components: CBX7, RING1B or RYBP, and PRC2 component: EED. Among the rTetR fusion candidates, rTetR-CBX7 induced robust silencing after eight days in doxycycline-containing (+Dox) medium, while rTetR-RING1B,

rTetR-RYBP and rTetR-EED were less efficient at establishing silencing. Therefore, I chose rTetR-CBX7 for my studies.

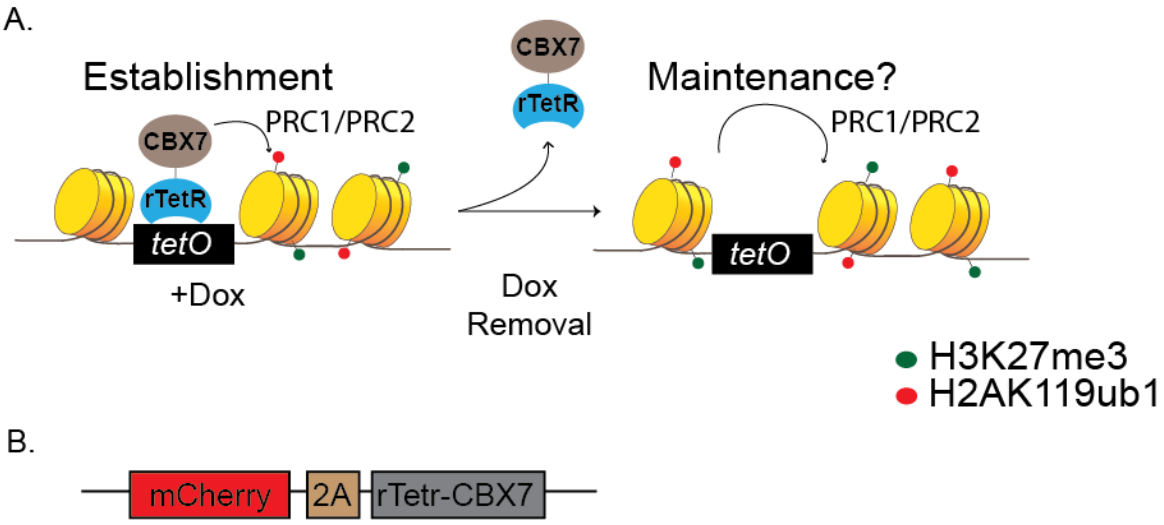


Figure 5. Diagram of experimental scheme for rTetR-CBX7 mediated H3K27me3 and H2AK119ub1 at the *5xtetO-H2B-CITRINE* reporter.

A. Doxycycline (+Dox) promotes the binding of rTetR-CBX7 to *tetO* sites leading to establishment of silencing. Dox Removal releases rTetR-CBX7 so that initiator-independent maintenance could be tested.

B. The *mCherry-2A-rTetR-CBX7* was integrated into *5xtetO-H2B-CITRINE* reporter cell lines using lentiviral transduction. The *2A* sequence enables ribosomal skipping between mCherry and rTetR-CBX7 during translation, which results in the mRNA being separated into two distinct proteins.

2.2.2. Observation of Locus-Dependent Epigenetic Inheritance

Inheritance of Polycomb Domains and Silencing of Reporter at Polycomb Target Genes

To investigate inheritance of Polycomb-mediated gene silencing at Polycomb target *WT1, 5xtetO-H2B-CITRINE* reporter was inserted 3.6kb upstream of the endogenous *WT1* gene. In HEK293FT cells, *WT1* is devoid of repressive H3K27me3 and H2AK119ub1, and is expressed , while in hESCs, it is associated with H3K27me3 and is not expressed (Figure 6A). The cells were cultured in +Dox medium to tether rTetR-CBX7 and establish silencing for 8 days (~8 cell

divisions), while cells in doxycycline-free (-Dox) medium served as controls. Fluorescent activated cell sorting (FACS) analysis showed that the *CITRINE* reporter was silenced in >95% of cells grown in +Dox medium but was fully expressed in cells grown in -Dox medium (Figure 6B). Following establishment, doxycycline was removed to release the rTetR-CBX7 from the *5XtetO* site (Dox Removal). At the *WT1* locus, I observed stable maintenance of the *CITRINE* reporter silencing at 2, 4 and 8 days after release of the rTetR-CBX7 initiator, with 74% of the cells maintaining silencing at 8 days (Figure 6B). The cells were followed to 40 days after release, at which time 27% of the cells still maintained silencing (Figure 7A). The inheritance of silencing was additionally verified through immunofluorescence imaging (Figure 7B).

Tethering rTetR alone did not lead to silencing at the reporter, indicating that silencing was not caused by rTetR-mediated steric inhibition (Figure 7C). To determine if the observed inheritance was due to leaky binding of rTetR-CBX7 when doxycycline was removed from the cell culture medium, I deleted rTetR-CBX7 from the cell line after establishment of silencing. The results showed that cells maintained silencing when rTetR-CBX7 was excised after 8 days of establishment, demonstrating that leaky binding of rTetR-CBX7 was not responsible for epigenetic inheritance at the *WT1* locus (Figure 7D).

Tethering of rTetR-CBX7 to ectopic loci has been reported to deposit H3K27me3 and H2AK119ub1 through the recruitment of PRC1-RING1A/B and PRC2 complex[56]. To assess whether H3K27me3 and H2AK119ub1 was established when rTetR-CBX7 bound to the *tetO* sites at the *WT1* site, I performed ChIP-qPCR and ChIP-seq experiments. The results showed that a domain of H3K27me3 and H2AK119ub1 was established by rTetR-CBX7, which was maintained 8 days after the release of rTetR-CBX7 (Figure 6C). The H3K27me3 domain extended to ~7kb on both sides of the *tetO* sites (Figure 6C, D). This was confirmed by ChIP-

Seq of H3K27me3 at the *WT1* locus, which showed that the H3K27me3 signal spread to the promoter and first exon of the *WT1* gene (Figure 6D). These results demonstrate that at the *WT1* site, rTetR-CBX7 mediates the establishment of silencing and H3K27me3 and H2AK119ub1 modifications, which can then be epigenetically maintained upon release of the rTetR-CBX7 initiator.

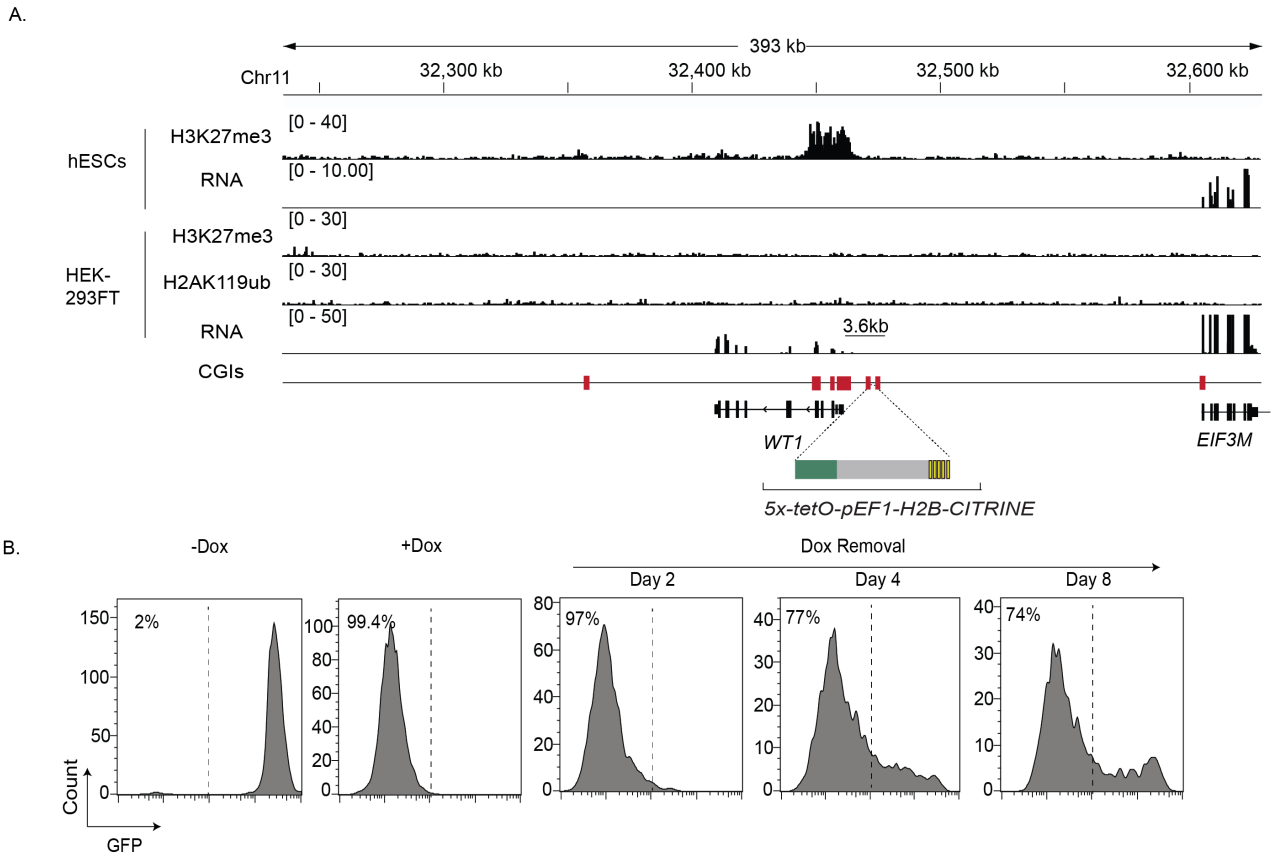


Figure 6 (Continued)

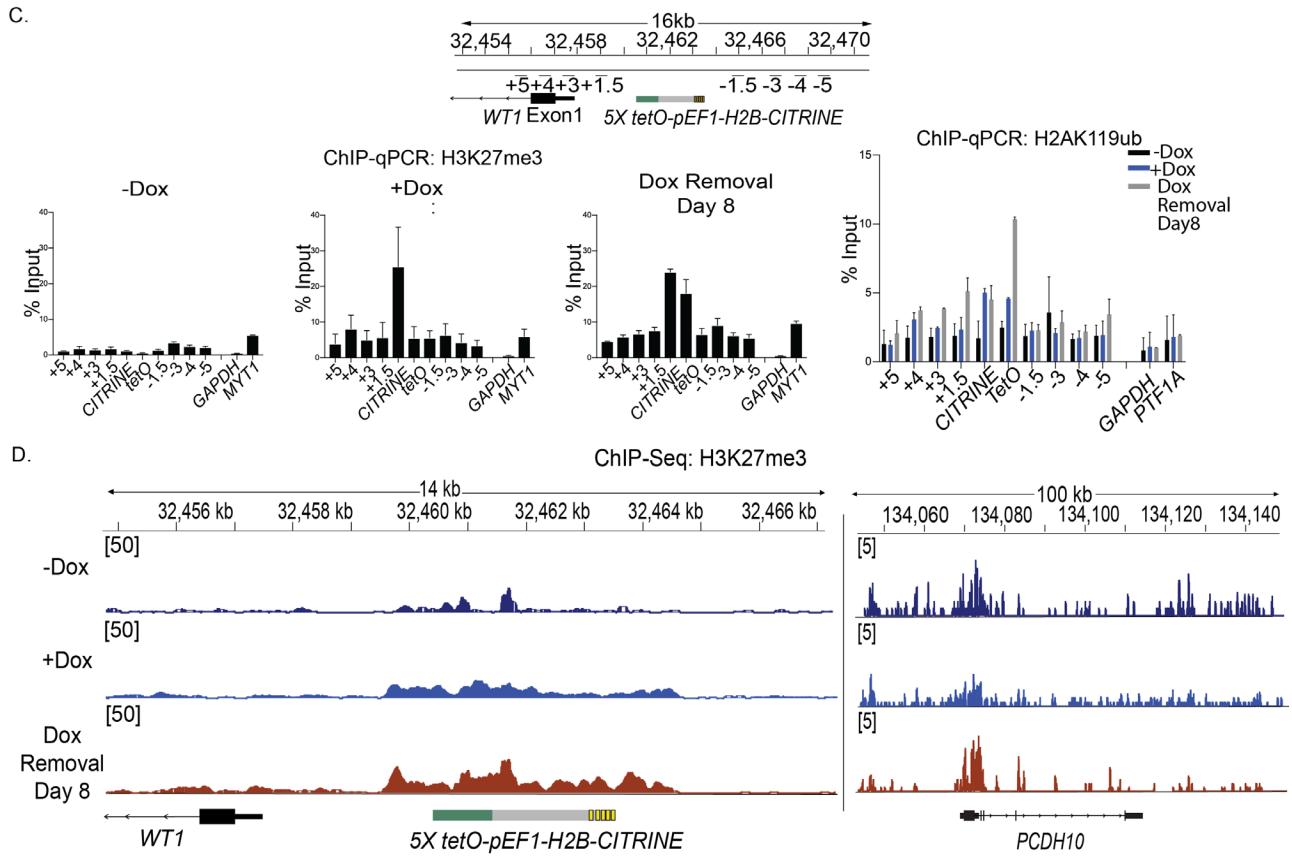


Figure 6. Establishment and maintenance of silencing of the *5xtetO-H2B-CITRINE* reporter and Polycomb domains at the *WT1* Locus.

(A) H2AK119ub1 and/or H3K27me3, RNA expression and CGIs at the integration site of the *5xteto-H2B-CITRINE* reporter construct proximal to the *WT1* gene on chromosome 11 in HEK293FT. Data from hESCs serve as a control cell line that shows a Polycomb silenced *WT1*. CGI locations are based on data acquired from the UCSC genome browser.

(B) Representative Flow Cytometry histograms showing the percentage of silenced cells resulting from tethering of the rTetR-CBX7 at Day8 in doxycycline-containing medium (+Dox) and then after transfer to doxycycline free medium at Day 2, 4, 8 (Dox Removal). Cells grown in doxycycline-free medium serve as controls (-Dox). Percentages (%) show fraction of silenced cells.

(C) ChIP-qPCR of H3K27me3 (left) and H2AK119ub1 (right) at the *WT1* locus. Cells that express rTetR-CBX7 but were propagated on -Dox medium to prevent binding the reporter act as controls for cells that were propagated in +Dox medium to establish silencing (+Dox) and cells that were transferred to -Dox medium to release rTetR-CBX7 (Dox Removal Day 8). Primer positions are indicated. *GAPDH* is a negative control for both HPMs and *MYT1* is a positive control for H3K27me3 and PTF1A is a positive control for H2AK119ub1. Error bars are deviation from the mean of at least two replicates.

Figure 6 (Continued)

(D) ChIP-Seq of H3K27me3 at the *WT1* locus (left) and a positive control locus *PCDH10* (right). Cells that express rTetR-CBX7 but have not had it tethered to the reporter (cultured in -Dox) act as controls for cells that have tethered rTetR-CBX7 in doxycycline medium (+Dox) to establish silencing and for cells that have released rTetR-CBX7 (Dox Removal Day8).

CGIs: CpG-rich islands; hESCs= Human Embryonic Stem Cells; Dox=Doxycycline

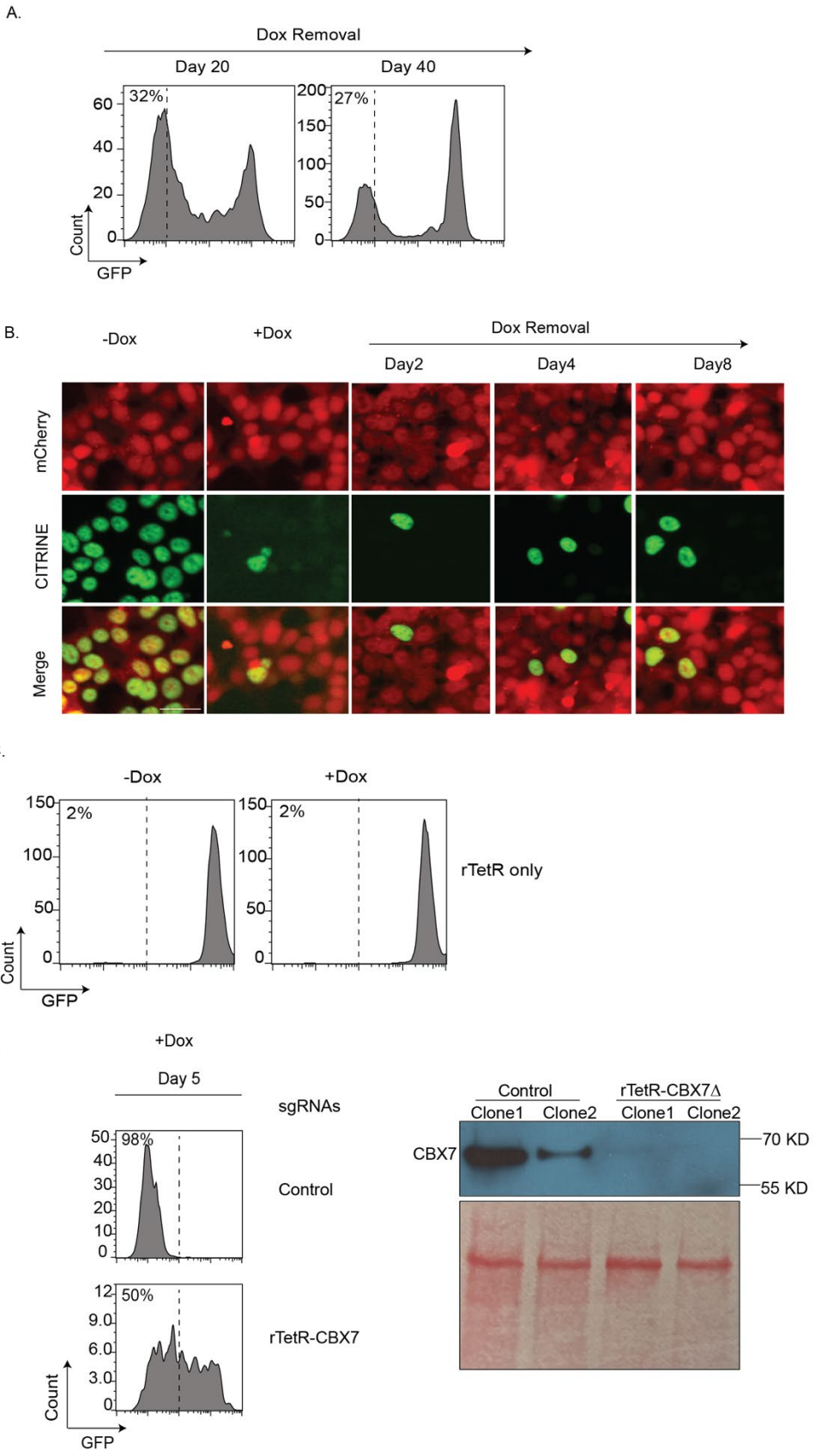

A.
Dox Removal
Day 20
Day 40
32%
27%
Count
GFP
B.
-Dox
+Dox
Dox Removal
Day2
Day4
Day8
mCherry
CITRINE
Merge
C.
-Dox
+Dox
2%
2%
rTetR only
Count
GFP
D.
+Dox
Day 5
sgRNAs
98%
Control
50%
rTetR-CBX7
Count
GFP
Control
rTetR-CBX7Δ
Clone1
Clone2
Clone1
Clone2
CBX7
70 KD
55 KD

Figure 7 (Continued)

Figure 7. Additional control experiments conducted at the *WT1* locus.

(A) Representative Flow Cytometry histograms to show the percentage of silenced cells at the *WT1* locus that were expanded in doxycycline containing media and then cultured in doxycycline free media at Day 20 and 40. Percentages (%) show fraction of silenced cells.

(B) Representative Immunofluorescence images of cells grown in doxycycline-free medium as control (-Dox) and doxycycline-containing medium (+Dox), to allow tethering of rTetR-CBX7, showing silencing of the *5X tetO-H2B CITRINE* reporter inserted at the *WT1* locus at DAY8. Cells were switched into doxycycline-free medium (Dox Removal) and imaged at Day 2, 4 and 8. mCherry indicates the expression of rTetR-CBX7 the reporter cell line.

(C) Representative Flow Cytometry histogram to show tethering of rTetR only in doxycycline-containing medium (+Dox).

(D) Representative Flow Cytometry histograms to show the percentage of silenced cells at the *WT1* locus 5 days after the excision of rTetR-CBX7 from cells in which silencing had been established by 8 days of growth in +Dox medium (left). Western Blot indicates absence of rTetR-CBX7 protein after CRISPR-Cas9 deletion (right).

I then sought to explore whether this phenomenon can be recapitulated at other Polycomb target genes. I inserted the *5xtetO-H2B-CITRINE* reporter at *EN2*, and two *HOX* loci, *HOXD11* and *HOXB4*. At the transcriptionally active endogenous *EN2* locus, which is devoid of H3K27me3 but has low levels of H2AK119ub1, the reporter was inserted 3.1kb upstream of the promoter region (Figure 8A). Similar to the *WT1* reporter locus, the *CITRINE* reporter at *EN2* was silenced in >95% of the cells (Figure 8B), and silencing was maintained in 56% and 27% of the cells at 4 and 8 days after the release of rTetR-CBX7, respectively (Figure 8B). However, compared to the *WT1* reporter, the rate of decay of silencing at this locus was faster. This faster rate of decay may be attributed to distinct locus-dependent factors that regulate *WT1* and *EN2* expression in HEK293FT cells (See Chapter 3).

At the endogenous *HOXD11* gene, the reporter was inserted 4.5kb upstream of the gene's promoter region. In HEK293FT cells, this locus lacks H3K27me3 but is associated with

H2AK119ub1, and is expressed (Figure 8C). The reporter near *HOXD11* was silenced in >95% of the cells during establishment, and 68% of the cells maintained silencing 8 days after the release of rTetR-CBX7 (Figure 8D). The rate of decay of silencing of the reporter at this locus was similar to the *WT1* site.

Interestingly, for the reporter inserted 3.1kb upstream of the endogenous *HOXB4* gene, I observed little or no rTetR-CBX7-mediated silencing in +Dox medium (Figure 8E-F). Therefore, inheritance could not be studied at this locus. It has previously been suggested that transcriptional activity opposes the formation of Polycomb domains[27]. Therefore, a possible explanation for the weak establishment of silencing at this locus may be the presence of the highly expressed proximal *HOXB* genes to the reporter gene (Figure 8E). In order to promote establishment at the *HOXB4* locus, I separately inserted three different silencing factors fused with rTetR-CBX7 into the genome of HEK293FT cells. Even when rTetR-CBX7 was expressed in combination with rTetR-EED or rTetR-HDAC1/4, the *CITRINE* reporter was not completely silenced. Instead, the expected tethering of two proteins at the *tetO* sites led to a reduction of the partial silencing observed with rTetR-CBX7 alone.

Together, the above results show that rTetR-CBX7 can induce reporter gene silencing at some, but not all, Polycomb target loci and that when silencing is established, it can be epigenetically inherited for several cell divisions.

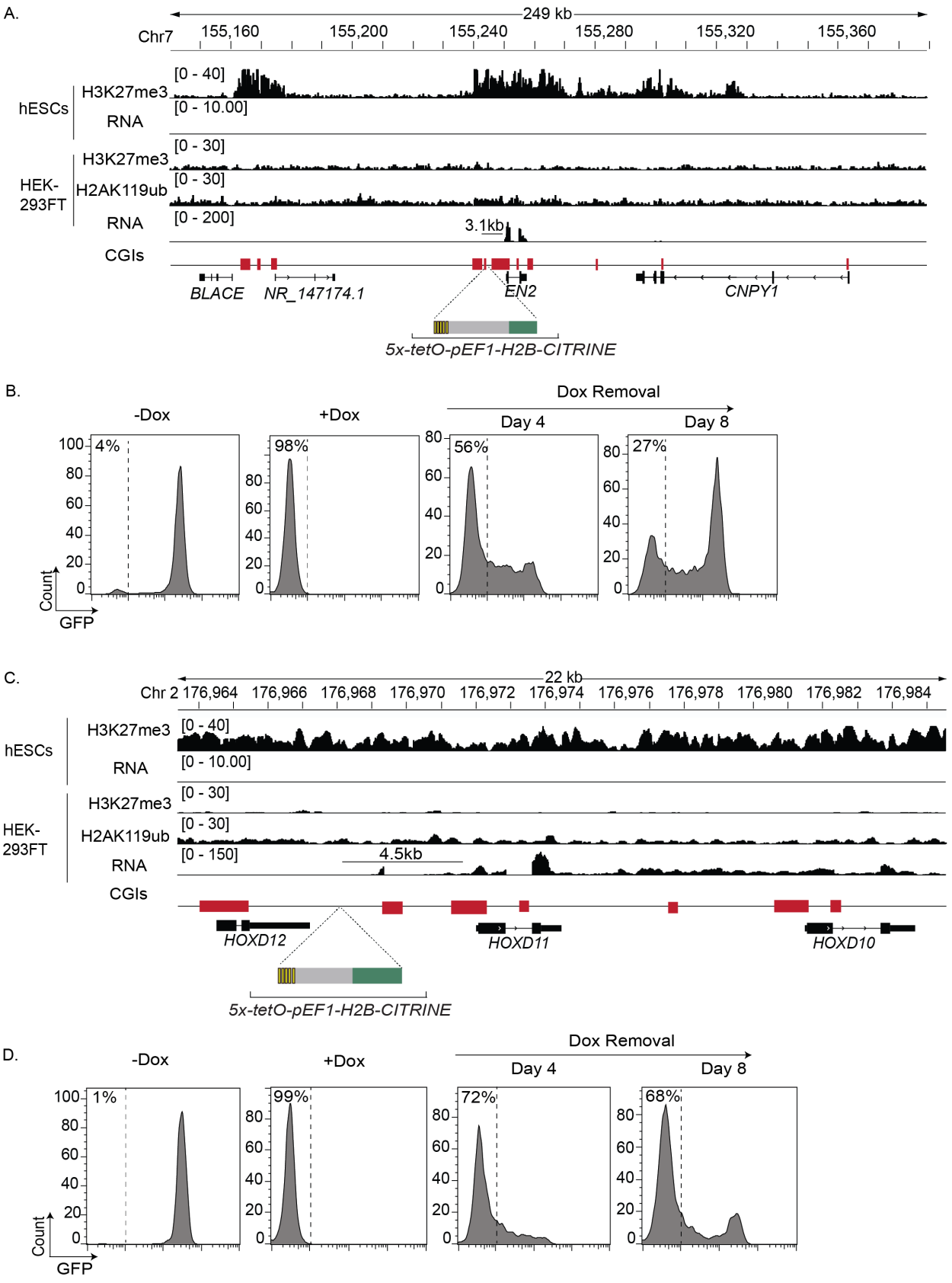

A.
249 kb
Chr7
155,160
155,200
155,240
155,280
155,320
155,360
hESCs
H3K27me3
[0 - 40]
RNA
[0 - 10.00]
HEK-293FT
H3K27me3
[0 - 30]
H2AK119ub
[0 - 30]
RNA
[0 - 200]
3.1kb
CGIs
BLACE
NR_147174.1
EN2
CNPY1
5x-tetO-pEF1-H2B-CITRINE
B.
-Dox
+Dox
Dox Removal
Day 4
Day 8
4%
98%
56%
27%
Count
GFP
C.
22 kb
Chr 2
176,964
176,966
176,968
176,970
176,972
176,974
176,976
176,978
176,980
176,982
176,984
hESCs
H3K27me3
[0 - 40]
RNA
[0 - 10.00]
HEK-293FT
H3K27me3
[0 - 30]
H2AK119ub
[0 - 30]
RNA
[0 - 150]
4.5kb
CGIs
HOXD12
HOXD11
HOXD10
5x-tetO-pEF1-H2B-CITRINE
D.
-Dox
+Dox
Dox Removal
Day 4
Day 8
1%
99%
72%
68%
Count
GFP

Figure 8 (Continued)

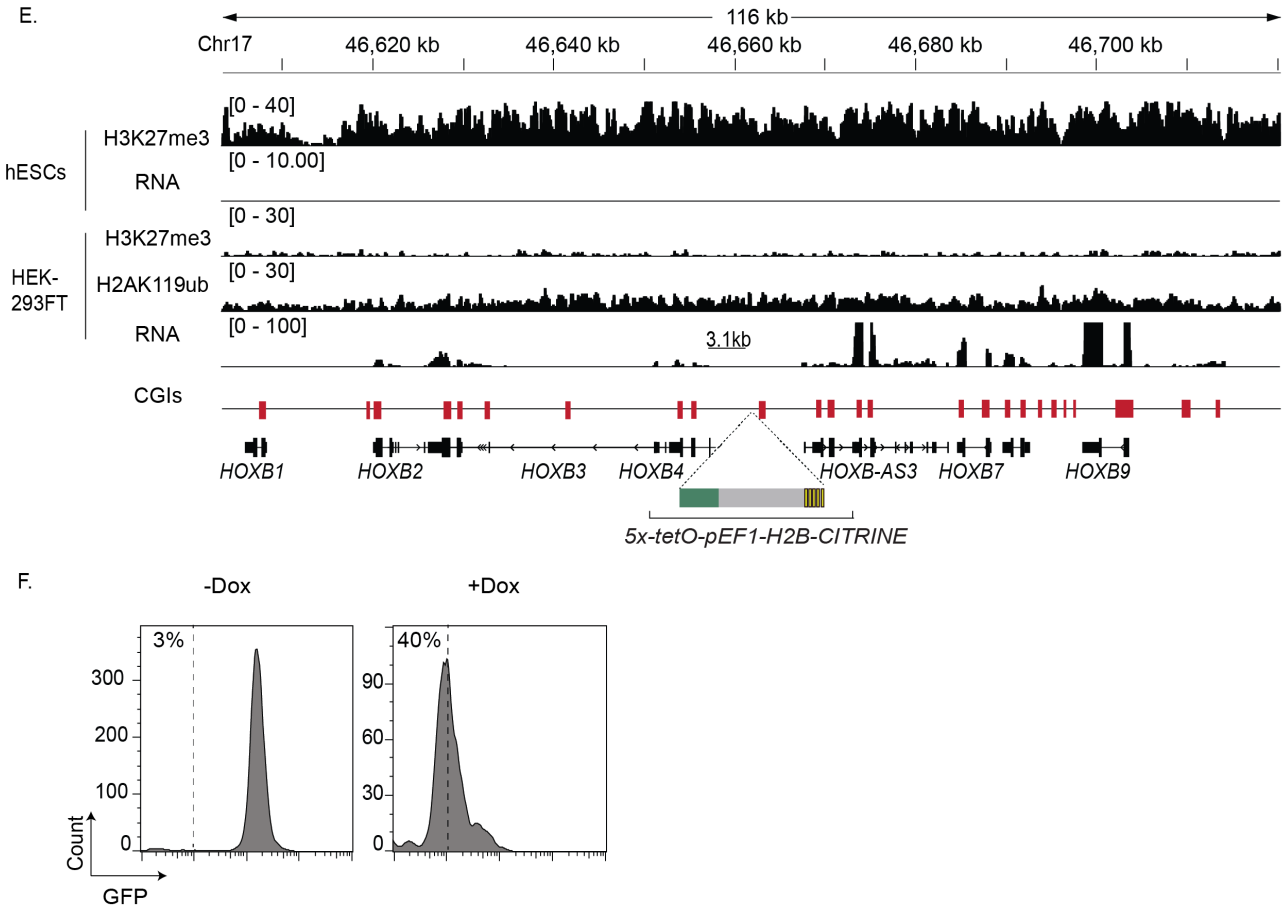


Figure 8. Establishment and maintenance of silencing of *5xtetO-H2B-CITRINE* reporter inserted at *EN2, HOXD11* and *HOXB4* loci.

(A) H2AK119ub1 and/or H3K27me3, RNA expression and CGIs at the integration site of the *5xtetO-H2B-CITRINE* reporter construct proximal to the *EN2* gene on chromosome 7 in HEK293FT. Data from hESCs serve as a control cell line that shows a Polycomb silenced *EN2.* CGIs locations are based on data acquired from UCSC genome browser.

(B) Representative Flow Cytometry histograms showing the percentage of cells with silenced reporter inserted in the *EN2* locus resulting from tethering of the rTetR-CBX7 at Day 8 in doxycycline-containing medium (+Dox), and after transfer to doxycycline-free medium at Day 4 and 8 (Dox Removal). Cells grown in doxycycline-free medium serve as controls (-Dox). Percentages (%) show fraction of silenced cells.

(C) H2AK119ub1 and/or H3K27me3, RNA expression and CGIs at the integration site of the *5xtetO-H2B-CITRINE* reporter construct proximal to the *HOXD11* gene on chromosome 2 in HEK293FT. Data from hESCs serve as a control cell line that shows a Polycomb silenced *HOXD11.* CGIs locations are based on data acquired from UCSC genome browser

Figure 8 (Continued)

(D) Representative Flow Cytometry histograms showing the percentage of cells with silenced reporter inserted in the *HOXD11* locus resulting from tethering of the rTetR-CBX7 at Day 8 in doxycycline-containing medium (+Dox) and after transfer to doxycycline-free medium at Day 4 and 8 (Dox Removal). Cells grown in doxycycline-free medium serve as controls (-Dox). Percentages (%) show fraction of silenced cells.

(E) H2AK119ub1 and/or H3K27me3, RNA expression and CGIs at the integration site of the *5xtetO-H2B-CITRINE* reporter construct proximal to the *HOXB4* gene on chromosome 17 in HEK293FT. Data from hESCs serve as a control cell line that shows a Polycomb silenced *HOXB4.* CGIs location are based on data acquired from UCSC genome browser.

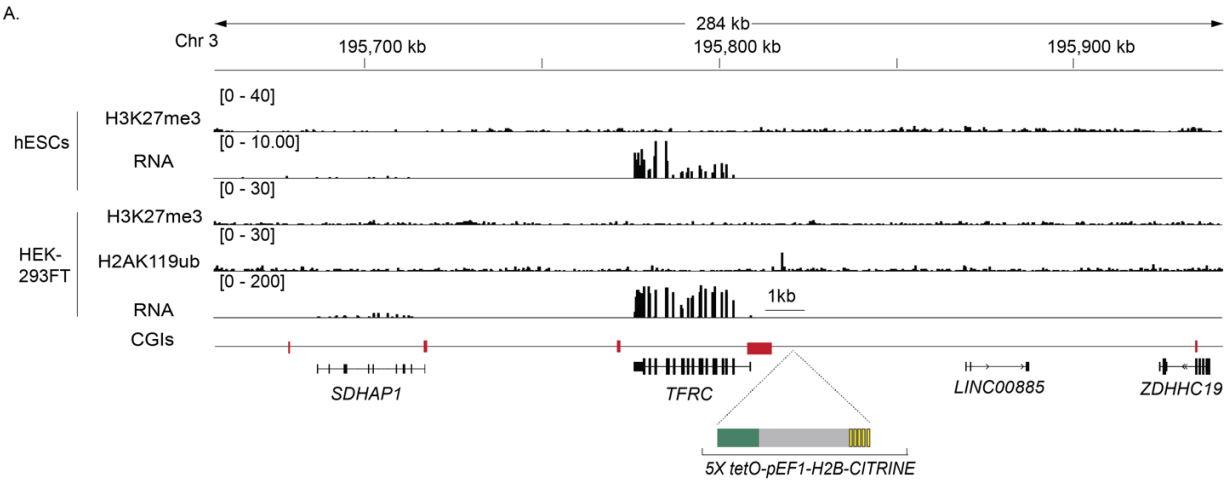
(F) Representative Flow Cytometry histograms showing the percentage of cells with silenced reporter inserted in the *HOXB4* locus resulting from tethering of the rTetR-CBX7 at Day 8 in doxycycline-containing medium (+Dox). Cells grown in doxycycline-free medium serve as controls (-Dox). Percentages (%) show fraction of silenced cells.

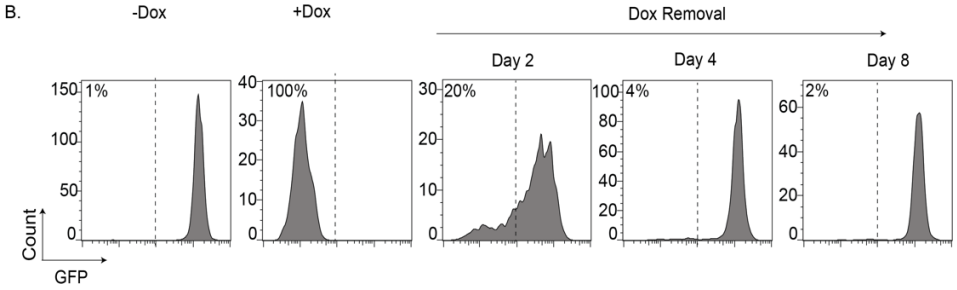
CGIs: CpG-rich islands; hESCs= Human Embryonic Stem Cells; Dox=Doxycycline

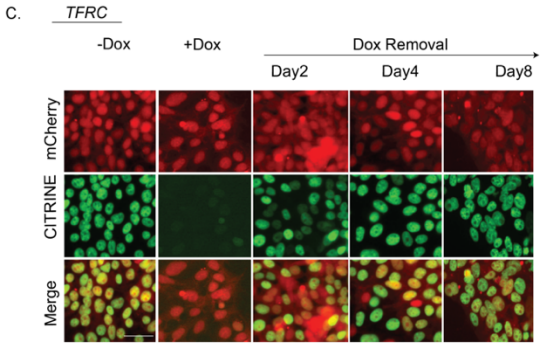
Polycomb-mediated Silencing of the Reporter at Housekeeping Genes cannot be Epigenetically Maintained

I next inserted the *5xtetO-H2B-CITRINE* reporter upstream of two housekeeping genes, *TFRC* and *B2M* in HEK293FT cells. Housekeeping genes are expressed in all cell types and are devoid of Polycomb modifications[99]. The *5xtetO-H2B-CITRINE* reporter was inserted 1kb upstream of the endogenous *TFRC* gene. In HEK293FT cells, H3K27me3 and H2AK119ub1 are absent at this locus and the *TFRC* gene is expressed (Figure 9A). The establishment and maintenance assays were carried out similarly to the Polycomb target genes. The results indicated that silencing was robustly established at the *TFRC* locus as the *CITRINE* reporter was silenced in >95% of cells. However, in contrast to the Polycomb target genes, reporter gene silencing was rapidly lost at this locus. While about 20% of the cells maintained silencing 2 days

after the release of rTetR-CBX7, only 4% and 2% maintained silencing by 4 and 8 days after release, respectively (Figure 9B, C). ChIP-qPCR of H3K27me3 and H2AK119ub1 showed that both HPMs were deposited when rTeTR-CBX7 was bound to the *tetO* sequence, but were lost in the maintenance phase, 8 days after the release of rTetR-CBX7 (Figure 9D). The size of the H3K27me3 domain at the *TFRC* locus was smaller relative to the *WT1* reporter site. The spreading at this locus was limited to the reporter cassette. However, the intensity of the H3K27me3 signal was similar at both *TFRC* and *WT1* reporter loci (Figure 6C, 9D).


A.
Chr 3
284 kb
195,700 kb
195,800 kb
195,900 kb
hESCs
H3K27me3
[0 - 40]
RNA
[0 - 10.00]
HEK-293FT
H3K27me3
[0 - 30]
H2AK119ub
[0 - 30]
RNA
[0 - 200]
1kb
CGIs
SDHAP1
TFRC
LINC00885
ZDHHC19
5X tetO-pEF1-H2B-CITRINE


B.
-Dox
+Dox
Dox Removal
Day 2
Day 4
Day 8
1%
100%
20%
4%
2%
Count
GFP


C.
TFRC
-Dox
+Dox
Dox Removal
Day2
Day4
Day8
mCherry
CITRINE
Merge

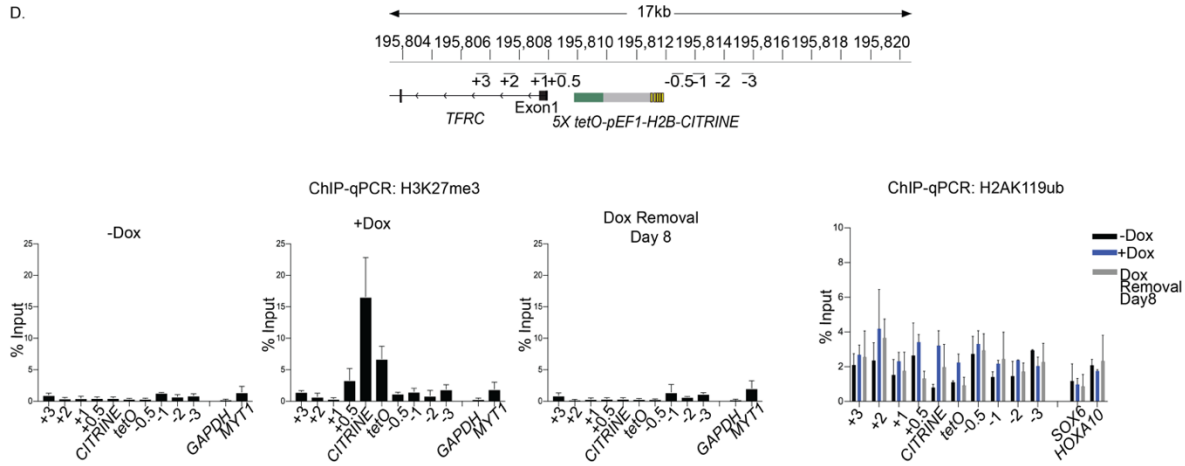

D.
17kb
195,804 195,806 195,808 195,810 195,812 195,814 195,816 195,818 195,820
+3 +2 +1 +0.5
-0.5 -1 -2 -3
TFRC
Exon1
5X tetO-pEF1-H2B-CITRINE
ChIP-qPCR: H3K27me3
-Dox
+Dox
Dox Removal Day 8
ChIP-qPCR: H2AK119ub
% Input
-Dox
+Dox
Dox Removal Day8
+3 +2 +1 +0.5 CITRINE tetO -0.5 -1 -2 -3 GAPDH MYT1
+3 +2 +1 +0.5 CITRINE tetO -0.5 -1 -2 -3 SOX6 HOXA10

Figure 9 (Continued)

Figure 9. Establishment and loss of silencing of the *5xtetO-H2B-CITRINE* reporter and Polycomb domains at the *TFRC* Locus.

(A) H2AK119ub1 and/or H3K27me3, RNA expression and CGIs at the integration site of the *5xtetO-H2B-CITRINE* reporter construct proximal to the *TFRC* gene on chromosome 3 in HEK293FT. Data from hESCs serve as a control cell line. CGIs locations are based on data acquired from UCSC genome browser

(B) Representative Flow Cytometry histograms showing the percentage of silenced cells resulting from tethering of the rTetR-CBX7 at Day8 in doxycycline-containing medium (+Dox) and after transfer to doxycycline free media at Day 2, 4, 8 (Dox Removal). Cells grown in doxycycline-free medium serve as controls (-Dox). Percentages (%) show fraction of silenced cells.

(C) Representative Immunofluorescence images of cells grown in doxycycline-free medium as control (-Dox) and doxycycline-containing medium (+Dox), to allow tethering of rTetR-CBX7, showing silencing of the *5X tetO-H2B CITRINE* reporter inserted at the *TFRC* locus at DAY8. Cells were switched into doxycycline-free medium (Dox Removal) and imaged at Day 2, 4 and 8. mCherry indicates rTetR-CBX7 expression.

(D) ChIP-qPCR of H3K27me3 (left) and H2AK119ub1 (right) at the *TFRC* locus. Cells that are expressing rTetR-CBX7 but have not had it tethered to the reporter serve as controls (-Dox) for cells that were grown in doxycycline-containing medium to tether rTetR-CBX7 and establish silencing (+Dox) and cells that have released rTetR-CBX7 (Dox Removal Day 8). Primer positions are indicated. *GAPDH* is a negative control and *MYT1* is a positive control for H3K27me3. *SOX6* is a negative control and *HOXA10* is a positive control for H2AK119ub1. Error bars are deviation from the mean of at least two replicates.

CGIs: CpG-rich islands; hESCs= Human Embryonic Stem Cells; Dox=Doxycycline

To investigate if this phenomenon can be recapitulated at another housekeeping target genes, the *5xtetO-H2B-CITRINE* reporter was inserted at the housekeeping gene, *B2M*, 4.7kb upstream of the endogenous gene promoter region (Figure 10A). The site is devoid of repressive H3K27me3 and H2AK119ub1 marks, and the *B2M* gene is highly expressed in HEK293FT cells (Figure 10A). The kinetics of establishment and maintenance of the silencing for the reporter at *B2M* were similar to the observations at the *TFRC* locus. With the rTetR-CBX7 tethered in +Dox medium, >95% of the cells silenced the reporter, but by 4 and 8 days after rTetR-CBX7 release,

the reporter was de-repressed in ~85% and ~93% of the cells, respectively (Figure 10B). ChIP-qPCR for H3K27me3 and H2AK119ub1 showed the deposition of these marks during the establishment phase and their absence 8 days after release of rTetR-CBX7 (Figure 10C). Similar to the *TFRC* gene, the H3K27me3 domain was limited to the reporter cassette.

These results indicate that Polycomb-mediated silencing and H3K27me3/H2AK119ub1 domains established near housekeeping target genes, *TFRC* and *B2M,* could not be epigenetically inherited.

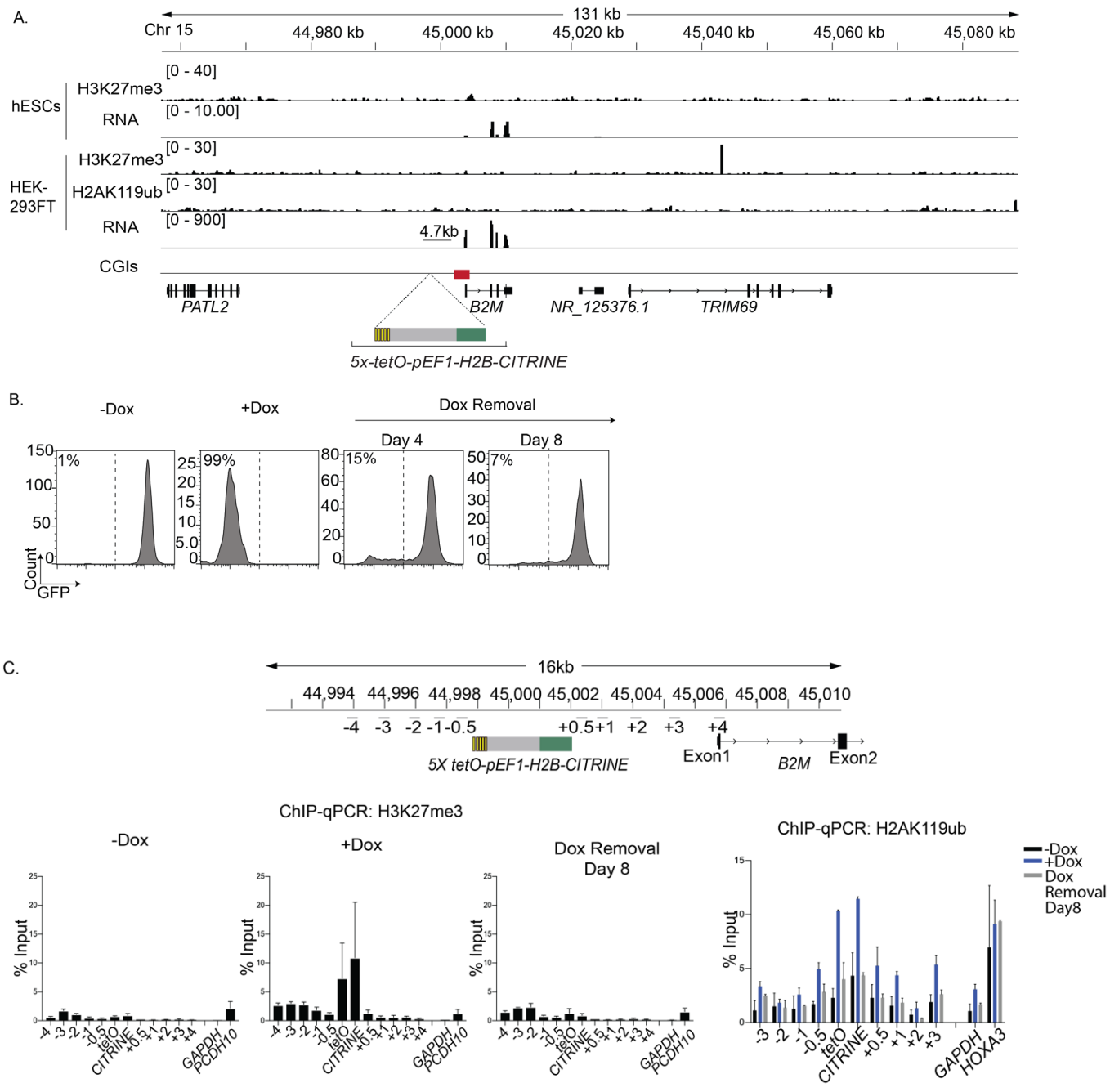


Figure 10. Establishment and loss of silencing of the *5xtetO-H2B-CITRINE* reporter and Polycomb domains at the *B2M* Locus.

(A) H2AK119ub1 and/or H3K27me3, RNA expression and CGIs at the integration site of the *5xtetO-H2B-CITRINE* reporter construct proximal to the *B2M* gene on chromosome 15 in HEK293FT. Data from hESCs serve as a control cell line. CGI locations are based on data acquired from UCSC genome browser.

Figure 10 (Continued)

(B) Representative Flow Cytometry histograms showing the percentage of silenced cells resulting from tethering of the rTetR-CBX7 at Day8 in doxycycline-containing medium (+Dox) and after transfer to doxycycline-free medium at Day 4 and 8 (Dox Removal). Cells grown in doxycycline-free medium serve as controls (-Dox). Percentages (%) show fraction of silenced cells.

(C) ChIP-qPCR of H3K27me3 (left) and H2AK119ub1 (right) at the *B2M* locus. Cells that express rTetR-CBX7 but have not had it tethered to the reporter (-Dox) serve as controls for cells with tethered rTetR-CBX7 in doxycycline-containing medium (+Dox) to establish silencing and to cells that have released rTetR-CBX7 (Dox Removal Day 8). Primer positions are indicated. *GAPDH* is a negative control and *PCDH10* is a positive control for H3K27me3. GAPDH is a negative control and HOXA3 is a positive control for H2AK119ub1. Error bars are deviation from the mean of at least two replicates.

CGIs: CpG-rich islands; hESCs= Human Embryonic Stem Cells; Dox=Doxycycline

A trivial explanation for the difference in epigenetic inheritance at Polycomb versus housekeeping reporter cell lines may involve differences in the expression of the rTetR-CBX7 protein. Western blot was carried out to assess the expression of rTetR-CBX7 in all the reporter cell lines studied above. The results showed that rTetR-CBX7 was overexpressed at similar levels in each cell line (Figure 11A). Therefore, different expression levels of rTetR-CBX7 cannot explain the difference in inheritance observed.

The establishment and maintenance assay at these reporter cell lines were conducted with an untagged rTetR-CBX7. Therefore, experiments to determine that rTetR-CBX7 binding to the *tetO* sequence at *WT1* and *TFRC* reporters are similar could not be conducted. To resolve this, I remade the *WT1* and *TFRC* reporter cell lines with a 3XFLAG-tagged rTetR-CBX7. I ensured that the rTetR-CBX7-FLAG proteins were expressed to similar levels in the two reporter cell lines and exhibited kinetics of establishment and maintenance near to the untagged version (Figure 11B). ChIP-pPCR for FLAG showed that rTetR-CBX7-FLAG bound the *tetO* sequence under establishment conditions at comparable levels at both the *WT1* and *TFRC* reporters (Figure

11C). At the *TFRC* reporter, no rTetR-CBx7 binding was detected 4 days after its release by growth in -Dox medium. Interestingly, at the *WT1* locus, there was binding of the rTetR-CBX7 under maintenance conditions, albeit to lower levels than in the establishment phase. A likely explanation is that rTetR-CBX7 is incorporated into the endogenous PRC1 complex and can be recruited to the locus during maintenance phase via Polycomb histone modifications. As expected, in cells grown in -Dox medium, rTetR-CBX7 did not bind to the reporter locus, as ChIP-qPCR signals were similar to the cell line where rTetR-CBX7 was absent (Figure 11C). This result again confirmed that there was no leaky binding of rTetR-CBX7 to the *tetO* sites in -Dox growth medium.

Another explanation for the differences in epigenetic inheritance at Polycomb versus housekeeping reporter cell lines may be the distance of reporter from the TSS of the endogenous gene and their directionality. At most loci, the reporters were inserted 3-4kb upstream of the TSS of the endogenous gene, except *TFRC* (which was inserted ~1kb from TSS). However, *TFRC* and *B2M* exhibited similar loss of maintenance, suggesting that differences in reporter distance does not play a role in inheritance of the silenced state. Furthermore, the directionality of the reporter does not affect inheritance, since *WT1* and *TFRC* have the same directionality while *B2M, EN2* and *HOXD11* do not, and the reporter insertions at the latter two genes exhibit heritable silencing.

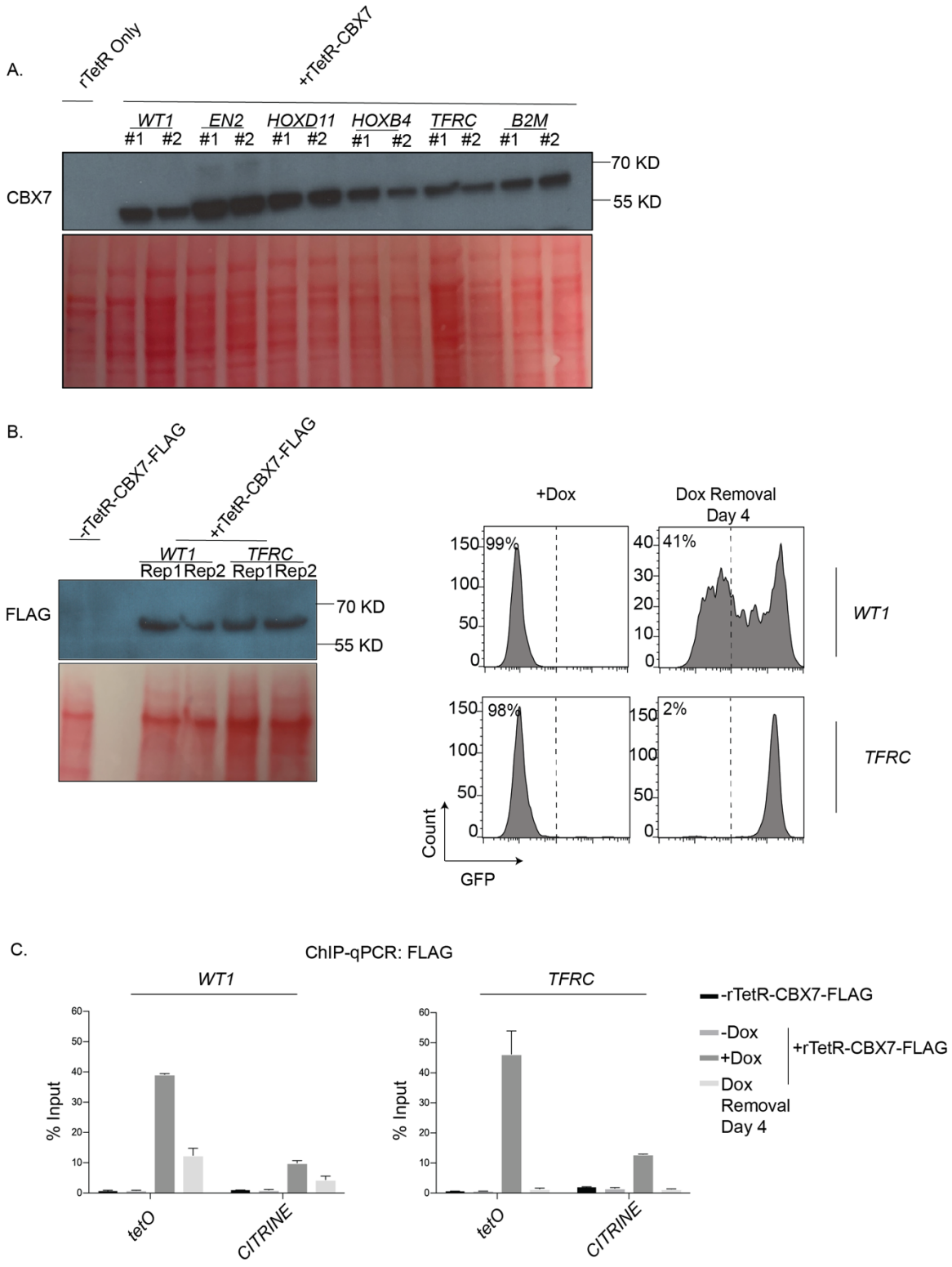


Figure 11. rTetR-CBX7 is expressed and bound to *tetO* sequences at comparable levels at the *WT1* and *TFRC* loci

Figure 11 (Continued)

(A)Western blot detecting rTetR-CBX7 with anti-CBX7 antibody in cells in which the reporter cassette was inserted at *WT1, EN2, HOXD11, HOXB4, TFRC* or *B2M.* Cells with only rTetR serve as a control.

(B) Western blot detecting rTetR-CBX7-FLAG with anti-FLAG antibody in cells in which the reporter cassette was inserted at *WT1* or *TFRC* (left). Representative Flow Cytometry histograms showing the percentage of silenced cells resulting from tethering of the rTetR-CBX7-FLAG to the reporter at *WT1* and *TFRC* loci on Day8 in doxycycline-containing medium (+Dox) and after transfer to doxycycline-free medium at Day 4 (Dox Removal). Cells grown in doxycycline-free medium serve as controls (-Dox). Percentages (%) show fraction of silenced cells.

(C) ChIP-qPCR to analyze the binding of rTetR-CBX7-FLAG to the reporter inserted at the *WT1* and *TFRC* loci. Cells that do not express rTetR-CBX7-FLAG and cells that express rTetR-CBX7 but have not had it tethered to the reporter (-Dox) act as controls for cells that have tethered rTetR-CBX7 (+Dox) to establish silencing and cells that have released rTetR-CBX7 (Dox Removal Day 4). Primers used are for the *tetO* and *CITRINE* sequences of the reporter construct. Error bars are deviation from the mean of at least two replicates.

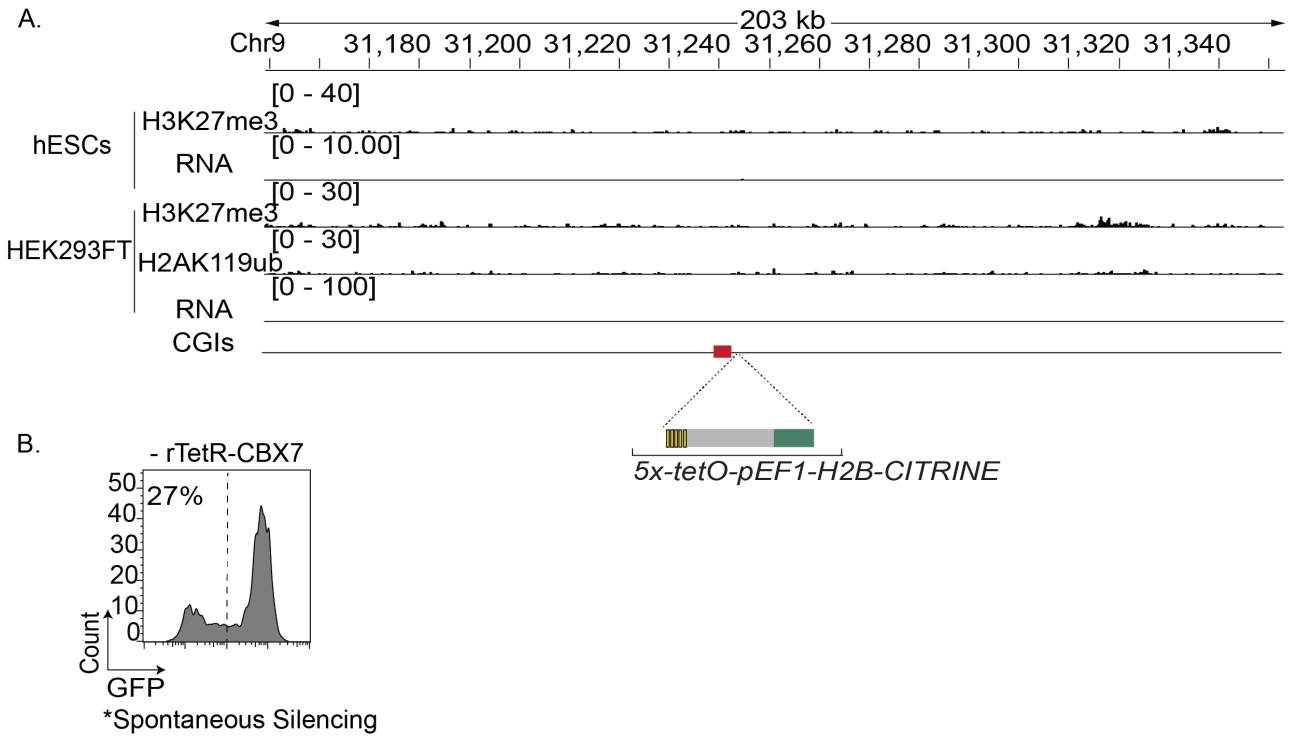
Investigating Epigenetic Inheritance at Gene Desert Regions

To further investigate locus-dependent effects on inheritance, I inserted the *5xtetO-H2B-CITRINE* reporter at two distinct gene desert regions, one with and one lacking CGIs. I hoped that these insertions would allow me to separate the proposed role of CGIs in Polycomb inheritance from the effect of cell type-specific transcription factor networks (TF network) that may control transcriptional activity at Polycomb target genes. Most CGIs in the genome contain 5mC DNA methylation, except the ones at Polycomb target genes, which depending on cell type are either bound by Polycomb or are transcriptionally active. Deliberately, I searched the human genome to find a gene desert that harbors a CGI but lacks DNA methylation. Thus, I inserted the reporter at a gene desert on Chromosome 9 that harbors a 1kb CGI (Figure 12A). I verified that this CGI is not methylated in HEK293 cells (UCSC genome browser, Reduced Representation

Bisulfite Sequencing). Unexpectedly, the reporter at this gene desert underwent spontaneous silencing under normal cell culture conditions in the absence of rTetR-CBX7 (Figure 12B).

Similarly, the reporter that I inserted at the second gene desert on Chromosome 7, which lacked CGIs and DNA methylation, also underwent spontaneous silencing (Figure 12C-D). A similar phenomenon was observed by another member of our lab who inserted a GFP reporter at a gene desert in mESCs (unpublished data from Moazed Lab). Therefore, due to the spontaneous silencing detected at the two gene deserts, they could not be used in my studies.

The spontaneous silencing at these two gene deserts may be mediated by the HUSH silencing complex that has been reported to recognize and transcriptionally repress a broad range of intronless transgenes in human cells[100]. Although in the case of HUSH, short intronless genes like the *CITRINE* reporter used in my experiments were refractory to silencing. In this regard, Moussa et al. reported that a *tetO-GFP* reporter cassette inserted at a gene desert in mESCs, was silenced in cells expressing rTetR-CBX7 and this silencing was maintained in ~77% of the cells after the release of rTetR-CBX7 (10-12 cellular divisions)[56]. It is possible that some gene desert locations allow reporter gene expression and can be used to study epigenetic inheritance.


A.
203 kb
Chr9
31,180 31,200 31,220 31,240 31,260 31,280 31,300 31,320 31,340
[0 - 40]
hESCs
H3K27me3
[0 - 10.00]
RNA
[0 - 30]
H3K27me3
HEK293FT
[0 - 30]
H2AK119ub
[0 - 100]
RNA
CGIs
5x-tetO-pEF1-H2B-CITRINE
B.
- rTetR-CBX7
27%
50
40
30
20
10
0
Count
GFP
*Spontaneous Silencing

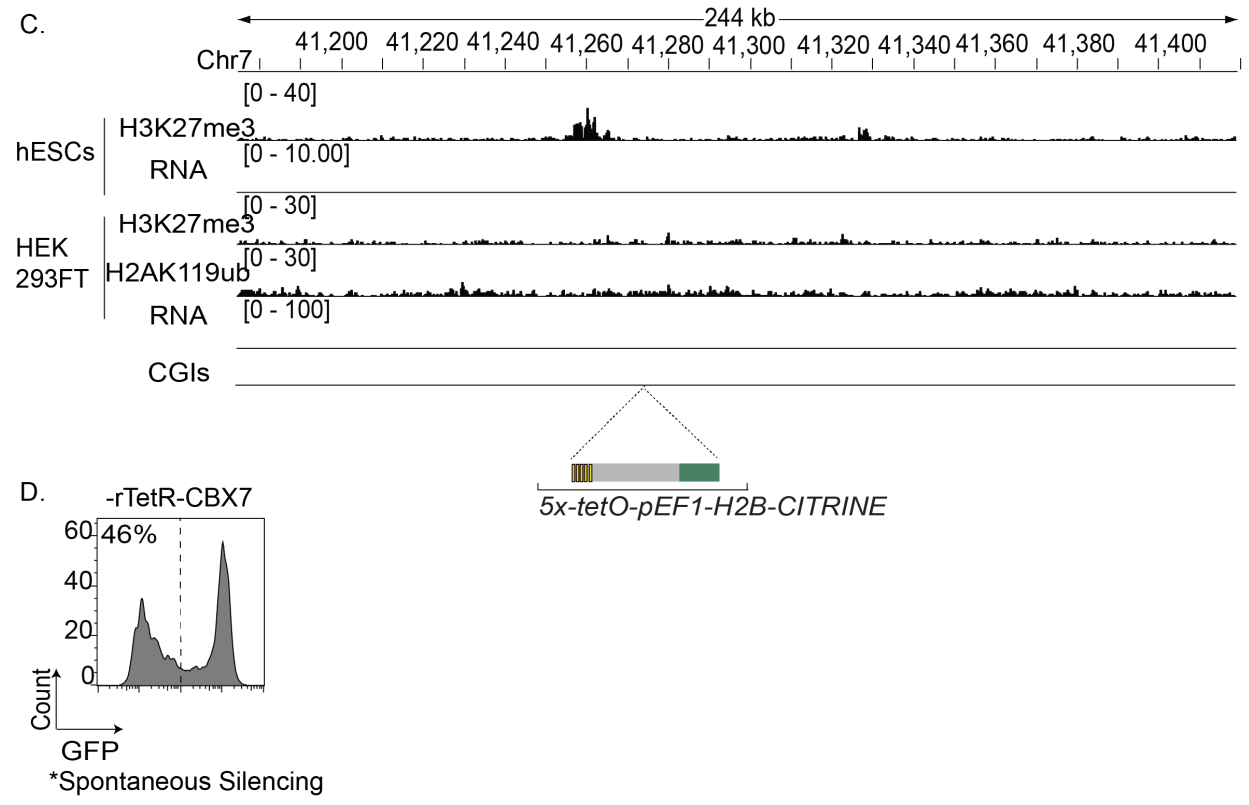

C.
244 kb
Chr7
41,200 41,220 41,240 41,260 41,280 41,300 41,320 41,340 41,360 41,380 41,400
[0 - 40]
hESCs
H3K27me3
[0 - 10.00]
RNA
[0 - 30]
H3K27me3
HEK 293FT
[0 - 30]
H2AK119ub
[0 - 100]
RNA
CGIs
5x-tetO-pEF1-H2B-CITRINE
D.
-rTetR-CBX7
46%
60
40
20
0
Count
GFP
*Spontaneous Silencing

Figure 12 (Continued)

Figure 12. *5xtetO-H2B-CITRINE* reporter integration at two distinct gene desert regions in HEK293FT cells.

(A) H2AK119ub1 and/or H3K27me3, RNA expression and CGIs at the integration site of the *5xteto-H2B-CITRINE* reporter construct in a gene desert region on chromosome 9 in HEK293FT and in hESCs. CGI location is based on data acquired from UCSC genome browser.

(B) Representative Flow Cytometry histogram showing the percentage of silenced cells resulting from spontaneous silencing under normal tissue culture conditions at the reporter integrated at indicated gene desert region on Chr9 in HEK293FT cells (shown in A). rTetR-CBX7 is absent in this cell line.

(C) H2AK119ub1 and/or H3K27me3, RNA expression and CGIs at the integration site of the *5xteto-H2B-CITRINE* reporter construct in a gene desert region on chromosome 7 in HEK293FT and in hESCs. No CGIs were annotated for this region based on date from the UCSC genome browser.

(D) Representative Flow Cytometry histogram showing the percentage of silenced cells resulting from spontaneous silencing under normal tissue culture conditions at the Chr 7 gene desert reporter in HEK293FT cells. rTetR-CBX7 is absent in this cell line.

CGIs: CpG-rich islands; hESCs= Human Embryonic Stem Cells

2.3 Discussion

By examining the heritability of Polycomb-mediated gene silencing for a reporter gene inserted at developmental targets of PcG versus housekeeping genes, I uncovered a role for locus-specific effects on epigenetic inheritance of Polycomb silencing.

In contrast to previous studies in mammalian cells[56,93,94] my results indicate that the positive feedback activities associated with the PRC1 and PRC2 complexes cannot maintain epigenetic states independently of DNA sequence at genic loci but rely on locus-specific inputs. I investigate the contributions of PRC positive feedback to the epigenetic inheritance of Polycomb silencing and the nature of locus-specific elements that impact inheritance in chapter 3.

Recently, Moussa and colleagues attempted to study the epigenetic inheritance of Polycomb silencing by integrating a reporter gene in an isolated gene desert in mESCs[56]. They observed robust maintenance of silencing of the reporter and concluded that the maintenance of silencing was sequence-independent and only relied on the contributions of the positive feedback loops of PRC1 and PRC2. Although their study demonstrates that Polycomb positive feedback can maintain silencing, a caveat of their experimental design is that it does not allow the role of locus-dependent factors that may contribute to epigenetic inheritance at endogenous loci to be tested. Their findings are consistent with recent results demonstrating that PcG silencing may be the default state at genes whose transcription is artificially repressed[27,73,84]. By contrast, my experimental strategy enabled the investigation of locus-dependent factors that may participate in epigenetic inheritance. The integration of reporter genes at two distinct types of loci, Polycomb target genes, where possible locus-dependent inheritance factors would be expected to have evolved, versus non-targe/housekeeping genes, provides an opportunity to dissect the locus-dependent factors. Using this system, I have demonstrated that Polycomb-mediated gene silencing at the *5xtetO-H2B-CITRINE* reporter can be stably inherited for multiple cellular divisions when the reporter is inserted near Polycomb target genes, *WT1, EN2* and *HOXD11*. By contrast, when the same reporter is inserted near housekeeping target genes, *B2M* and *TFRC*, silencing is rapidly lost (Figure 13). These results indicate that locus-dependent factors contribute to the epigenetic inheritance of Polycomb-mediated gene silencing.

Previous studies provide clues as to the nature of locus-dependent factors that may contribute to epigenetic inheritance of heterochromatin. With regards to H3K9me3 inheritance, it has been demonstrated that DNA sequence elements that bind site-specific DNA binding proteins are necessary for its propagation[67,68]. Furthermore, the necessity of DNA sequence

elements has also been reported in *Drosophila*, where PRE excision leads to the loss of Polycomb-mediated gene silencing. Therefore, it is reasonable to hypothesize that DNA sequences, such as CGIs that are thought to be "PRE-like" in mammalian cells, contribute to epigenetic inheritance observed at the Polycomb target genes that I have studied. It has previously been shown that PRC2 accessory factors, MTF2/PHF1/PHF19 have the ability to bind CG-rich DNA in vitro and that their genome-wide localization correlates with CGIs in vivo. Therefore, the reporter cell lines described in this chapter allowed me to test the contribution of the CGI binding ability of these proteins to the inheritance of Polycomb domains (chapter 3).

An additional locus-dependent factor that has been implicated in the formation of Polycomb domains is transcriptional activity that opposes Polycomb (See Chapter 1). Mathematical modeling and simulations have suggested that antagonistic transcriptional activity contributes to Polycomb mediated silencing. However, little to no experimental data exists that demonstrate the role of antagonistic transcriptional activity in inheritance of Polycomb domains. In plants, studies of the *FLC* reporter during vernalization suggests that trans-factors that regulate gene expression may play a role in switching of the epigenetic state[86,91]. In mammalian cells, transient activation of a Polycomb silenced gene can lead to stable epigenetic switching to an active state[84], suggesting a contribution of *cis* acting transcriptional activity in counteracting the maintenance of Polycomb-mediated gene silencing. The reporters inserted at housekeeping loci that show rapid decay in silencing (this chapter) will allow me to test for the contribution of antagonistic transcriptional inputs to inheritance of Polycomb-mediated gene silencing (chapter 3).

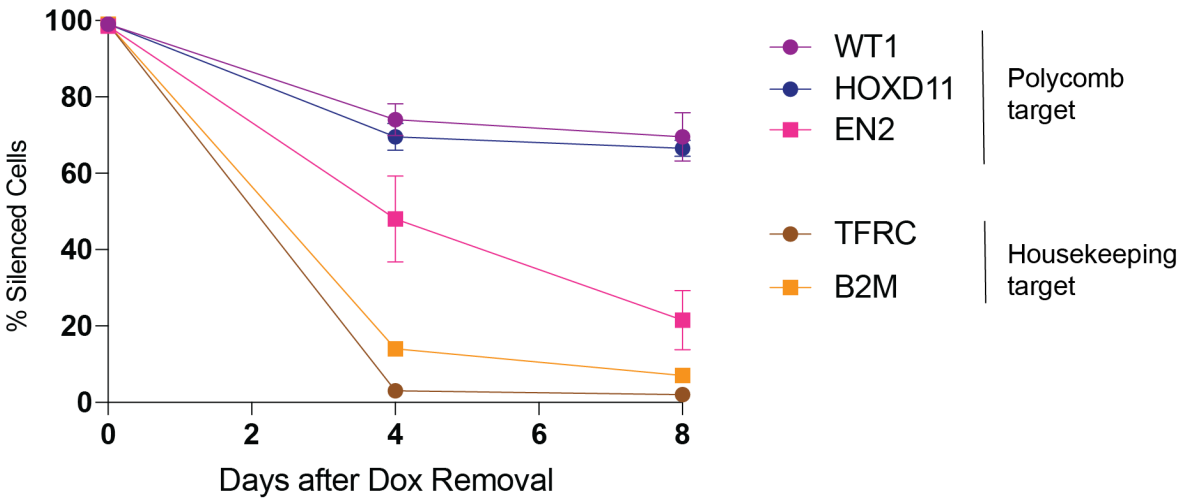


Figure 13. The rate of decay of silencing after release of rTetR-CBX7 at Polycomb and housekeeping target genes.

Plot showing the decay rate of silencing for the *5x-tetO-H2B-CITRINE* reporter inserted at Polycomb target genes, *WT1, HOXD11* and *EN2*, and housekeeping genes, *TFRC* and *B2M,* after release of rTetR-CBX7 (Dox Removal) for the indicated days.

Chapter 3: Requirements for Locus-Dependent Epigenetic Inheritance of Polycomb-Mediated Gene Silencing

3.1 Introduction

In Chapter 2, I showed that epigenetic inheritance of Polycomb-mediated gene silencing was strongly influenced by the chromosomal location at which silencing was induced. To summarize, I observed that at developmental targets of Polycomb, *WT1*, *EN2* and *HOXD11* genes, induced silencing of a reporter gene and Polycomb-associated histone modifications, H3K27me3 and H2AK119ub1, were maintained after the release of rTetR-CBX7. Silencing at these loci still decays, but at a relatively slow rate. By contrast, at the housekeeping genes, *TFRC* and *B2M,* silencing of the reporter and the Polycomb modification are rapidly lost (Figure 13). In the studies described in this chapter, I investigate the factors that determine the differences in inheritance observed at these two distinct types of loci. These factors could be subdivided into locus-independent (or DNA sequence-independent) and -dependent (or DNA sequence-dependent) factors.

Locus-independent factors include the PRC1 and PRC2 complexes and their ability to recognize and catalyze H3K27me3 and H2K119ub1(read-write). As mentioned in Chapter 1, the positive feedback read-write capability of HMTs has been proposed to help mediate epigenetic inheritance. In fission yeast, the read-write capability of Clr4 (H3K9me HMT) is necessary for conditional epigenetic inheritance of H3K9me when the putative H3K9me demethylase is deleted[65,66]. Clr4 read and write capabilities both exist in Clr4 itself, which has a chromodomain that can recognize H3K9me and a SET domain that catalyzes H3K9me. PRC2's reader function is located in EED, which contains WD40 repeats that form an aromatic cage and binds H3K27me3. This binding in turn allosterically activate EZH2, the writer protein to catalyze

H3K27me3[12]. Oksuz et al, investigated the positive feedback mechanism of EED-PRC2 in mESCs in the formation of a Polycomb domain but not inheritance of the domain[15]. They expressed a WT tagged or aromatic cage mutant tagged from the endogenous EED. Their aromatic cage mutant may have been hypomorphic since they only changed one of the three residues that form the aromatic cage[12]. However, they were able to observe that H3K27me3 levels are significantly reduced (H3K27me2 is still present) and the aromatic cage is not necessary for the recruitment of PRC2 but for the efficient catalysis of di to tri methylation at nucleation sites. These results suggest that the primary role of the EED aromatic cage is to bind di and tri methylated nucleosomes to help recruit EED-PRC2, while binding to the tri-methylated state allows for allosteric activation of EZH2 and H3K27 tri-methylation. Therefore, these findings indicate that EED's read-write function is necessary for nucleation; however, the evidence that this positive read-write mechanisms of EED-PRC2 is necessary for inheritance of an established H3K27me3 Polycomb domain is lacking.

Evidence for PRC1 and PRC2 positive feedback loops being necessary for epigenetic inheritance of Polycomb-mediated gene silencing in mammalian cells was provided by Moussa et al. through examination of an ectopic Polycomb domain that could be maintained upon release of the initiator[56]. Their findings demonstrated that the feedback loop between H3K27me3 and cPRC1 (with its H2AK119ub1 activity) mediate the inheritance, since inheritance was lost with an EZH1/2 inhibitor that reduces H3K27me3 levels, a chromodomain mutant of CBX7 that is unable to recognize H3K27me3 and overexpression of de-ubiquitinase enzymes that reduces H2AK119ub1 levels. cPRC1s have lower H2AK119 ubiquitination activity (major function is thought to be chromatin compaction and/or LLPs [37,101,102]), yet this study demonstrated that in certain contexts cPRC1 can catalyze H2AK119ub1 and link H3K27me3 to H2AK119ub through

its chromodomain. However, this study did not demonstrate the necessity of EED-PRC2 feedback loop for inheritance.

Besides the linked feedback loops between cPRC1 and PRC2 as described above, other feedback loops could include PRC2.2-JARID2 that recognizes H2AK119ub1 deposited by the vPRC1. Furthermore, vPRC1 reinforces the histone modification it catalyzes, since its RYBP subunit is a reader protein that binds H2AK119ub1 and allosterically activates RING1A/B, the writer subunit that catalyzes H2AK119ub1. This model is supported by evidence demonstrating that PRC2 occupancy is severely reduced upon complete catalytic inactivation of PRC1-RING1A/B in ESCs[49,50]. Despite these important observations, limited evidence exists that feedback loops are necessary for the inheritance of Polycomb-mediated gene silencing in differentiated cells.

Locus-dependent factors include elements that vary from one genomic region to another. This includes DNA binding proteins and their contact with CGIs. As previously described, DNA sequences have been found to be necessary for the inheritance of H3K9me domains in fission yeast and H3K27me3 in flies[67,68,70,71]. The Polycomb responsive DNA elements (PREs), described in flies, do not appear to be conserved in mammals, but it is generally thought that CGIs act like PREs in mammals[103]. Several studies have shown that Polycomb domains correlate with unmethylated CGIs and several PRC accessory subunits have the ability to bind CG-rich DNA[19,21]. There is anecdotal evidence that Polycomb target genes have a higher density of CGIs compared to non-Polycomb targets[26]. However, a difference in motifs at these CGIs between Polycomb and non-Polycomb target loci has not been described. KDM2B, which associates with vPRC1, and JARID2 and AEBP2, which associate with PRC2.2, have been shown to bind CG-rich sequences in vitro and are located at CGIs in vivo[25,33,76,104,105]. In addition, PCL proteins,

MTF2, PHF1 and PHF19 are accessory proteins of PRC2.1 that can bind DNA through their extended homologous (EH) domain that fold into a winged-helix structure. Deletion of PCL proteins in ESCs results in reduced H3K27me3 and PRC2 occupancy[33,105]. Several studies demonstrated that PCLs selectively binds CG sequences in vitro and regions with high density of unmethylated CGIs in vivo[19,21]. Furthermore, computational modeling has shown that the DNA shape at CGIs may influence PCL binding. These studies observe a correlation between Polycomb target genes with CGIs, where the DNA is unwound for optimal contact with PCL's winged helix structure, while non-Polycomb target sites with less CGIs have more tightly wound DNA[21]. It remains to be seen whether this can be experimentally confirmed. PCL proteins also have Tudor domains that can bind to H3K36me3, a modification which is known to counteract Polycomb[20,106]. PHF1 and PHF19 are more efficient at H3K36me3 binding than MTF2 to facilitate PRC2 recruitment and H3K36me demethylase to active regions[107,108]. Besides their function in proper recruitment of PRC2 to Polycomb targets through contact with CGIs and enhancing EZH2 HMT activity, other functions have been proposed[19,21,109]. One is that PCLs increase the residency time of PRC2 on chromatin to enhance its HMT activity, which has been experimentally demonstrated through in vitro biochemical assays with PHF1[92]. However, the role of these PCL proteins in inheritance of Polycomb domains has not been studied.

Locus-dependent factors also include those that specifically impact transcriptional activity at the target site. It is well known that genes have different expression levels, even genes located adjacent to one another. Therefore, the factors that regulate transcriptional activity may have a potential impact on epigenetic inheritance of gene silencing at the locus. As mentioned in Chapter 1, it has been demonstrated that inhibition of transcription leads to de novo Polycomb domain formation[27]. In addition, CGIs that lacks TF binding sites can recruit Polycomb[73].

However, limited evidence exists on the role of antagonistic transcriptional activity in inheritance of Polycomb domains. In *Arabidopsis,* the *FLC* reporter can stably inherit a silent or active state when switched from cold (establishment) to warm (maintenance) temperatures indicating that trans-factors that regulate gene expression could play a role in switching of the epigenetic state[86,91]. In mammalian cells, transient transcriptional activation of a Polycomb silenced gene can lead to an epigenetic switch[84], suggesting the contribution of *cis* acting transcriptional activity in opposing the maintenance of Polycomb-mediated gene silencing. However, this study did not examine locus-specific contributions directly. Furthermore, the study focused on genes in which the endogenous TF network that regulates them was absent.

In this chapter, I describe the contributions of locus-independent and locus-dependent factors to the inheritance of Polycomb-mediated gene silencing. To test the role of locus-independent PRC2 read-write and the role of RING1A/B, I used the cell line with the reporter cassette inserted at the *WT1* locus which exhibited robust inheritance of silencing of the reporter and Polycomb domains. I found that the ability of the EED subunit of PRC2 to recognize H3K27me3 and the RING1A/B subunits of PRC1 were necessary for epigenetic inheritance at the *WT1* locus. More importantly, my findings show that the ability of the PRC2 accessory factors, MTF2 to bind to CG-rich DNA, was necessary for epigenetic inheritance at the *WT1* locus. Finally, I demonstrate that at the housekeeping target loci, *TFRC* and *B2M*, the rapid loss of reporter gene silencing upon the release of the rTetR-CBX7 initiator was partially reversed by DNA deletions that reduced transcriptional activity at these loci. Together with the results described in Chapter 2, my findings suggest that epigenetic inheritance of Polycomb-mediated gene silencing is mediated by both locus-independent factors, which include read-write mechanisms and positive feedback loops, and locus-dependent factors, which includes DNA

binding ability of DNA binding proteins to CGIs and antagonistic transcriptional activity mediated by the TF network of target sites.

3.2 Results

3.2.1 DNA-Sequence Independent Read-Write Positive Feedback Mechanism is required for Heritable Silencing at Polycomb Target Genes

The Role of PRC2 Read-Write in Maintenance of Silencing

To investigate whether the read-write ability of PRC2 was required for the inheritance of Polycomb-mediated gene silencing, I deleted EED, the component of the PRC2 complex that recognizes H3K27me3, in cells carrying the *CITRINE* reporter at the *WT1* locus. As previously described, I observed epigenetic inheritance of the silenced state and Polycomb-associated histone modifications at the *WT1* locus. In *EED*$^{-/-}$ KO cells, establishment of the reporter silencing (in doxycycline-containing medium) was unaffected, but maintenance of silencing was lost (in doxycycline-free medium) (Figure 14A). I then attempted to rescue the maintenance defect of *EED*$^{-/-}$ cells with either *WT HA-EED* or an aromatic cage mutant (F97A, Y148A, Y365A), *HA-EED-3A*), which does not bind to H3K27me3, and thus cannot allosterically activate EZH2[12,15] (Figure 14B). It has previously been shown that H3K27me3 is lost in *EED*$^{-/-}$ cells[15]. Using western blotting, I verified that H3K27me3 was absent from my *EED*$^{-/-}$ cells and was rescued in cells transfected with the *WT HA-EED* but not the mutant *HA-EED-3A* (Figure 14C).

Expression of WT HA-EED fully rescued the maintenance defect of the *CITRINE* reporter in the *EED*$^{-/-}$ cells, while expression of the HA-EED-3A mutant version showed a partial rescue (Figure 14D). A likely explanation for the partial rescue observed in *HA-EED-3A* mutant

cells is that the overexpressed HA-EED-3A subunit is still able to assemble into a stable PRC2 complex, which is not the case in *EED*$^{-/-}$ cells. The PRC2 complex containing the mutant EED appears to promote maintenance of silencing, likely through the propagation of PRC1-mediated H2AK119ub1. In support of this hypothesis, ChIP-qPCR experiments showed that H2AK119ub1 was present in cells transfected with the *HA-EED-3A* construct during the maintenance phase (Figure 15A).

ChIP-qPCR and ChIP-Seq for H3K27me3 showed that H3K27me3 was restored at the *WT1* locus during establishment and maintenance in the *WT HA-EED* but not in *HA-EED-3A* cells (Figure 14E-F). These results suggest that in the absence of H3K27me3, H2AK119ub1 was not sufficient for stable maintenance of the silenced state. However, a substantial fraction of cells (~23%) maintained silencing 8 days after the release of rTetR-CBX7 in *HA-EED-3A* , demonstrating that epigenetic maintenance occurs in the absence of detectable H3K27me3.

In an independent experiment, SUZ12, another core component of the PRC2 complex was deleted from cells with the *CITRINE* reporter at the *WT1* locus. As expected, H3K27me3 was abolished from the *SUZ12*$^{-/-}$ cells (Figure 16A). The establishment of *CITRINE* reporter silencing (in doxycycline-containing medium) was unaffected, but maintenance of silencing was lost (in doxycycline-free medium), as was the case in *EED*$^{-/-}$ cells (Figure 16B).

These results indicate that H3K27me3 and the read-write capability of PRC2 contribute to the stable maintenance of Polycomb-mediated silencing. Surprisingly, they also reveal that substantial epigenetic maintenance can occur in the absence of H3K27me3 or a wild-type EED reader domain.

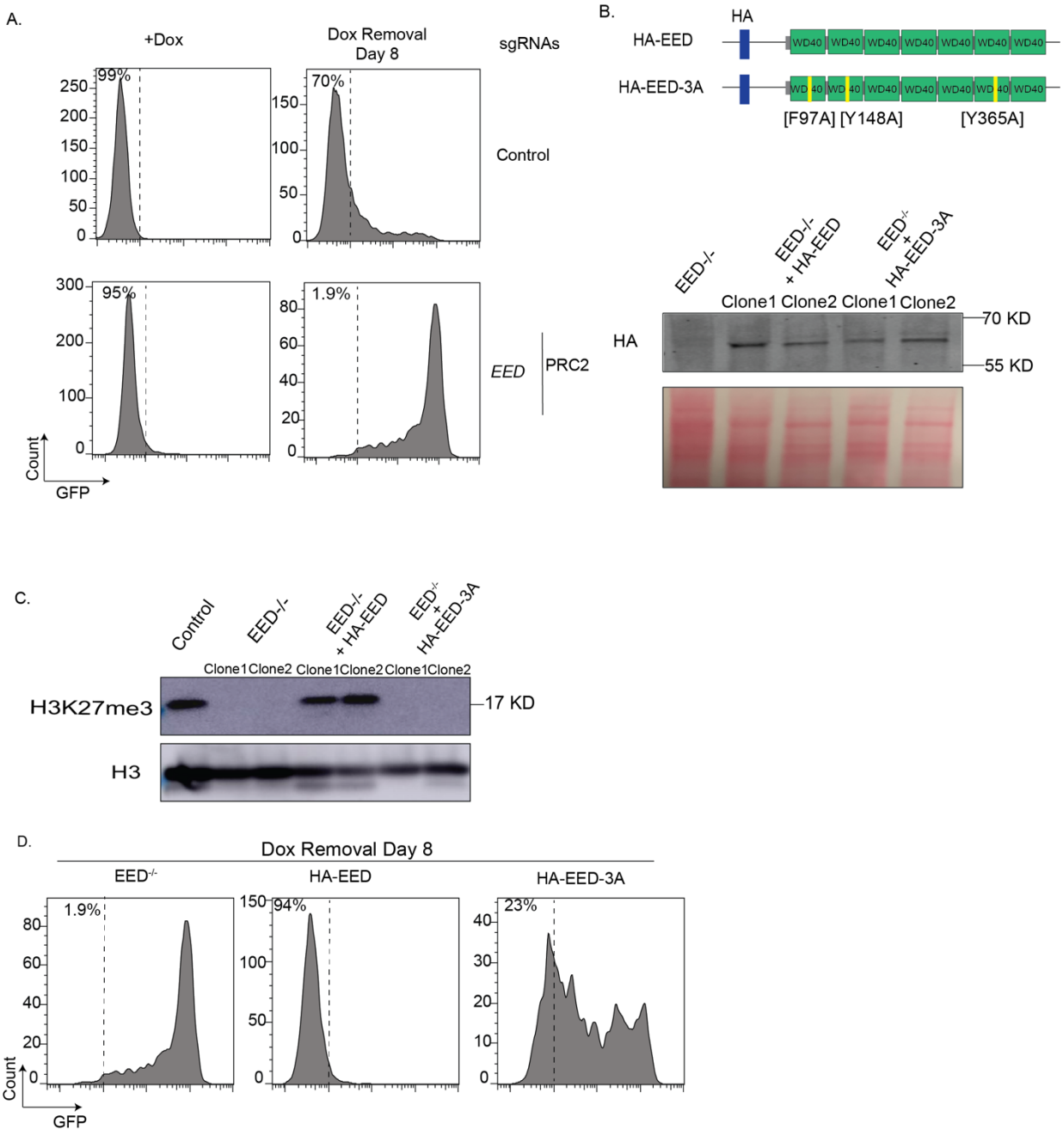

A.
+Dox
Dox Removal Day 8
sgRNAs
99%
70%
Control
95%
1.9%
EED
PRC2
Count
GFP
B.
HA
HA-EED
HA-EED-3A
WD40
[F97A] [Y148A]
[Y365A]
EED-/-
EED-/- + HA-EED
EED-/- + HA-EED-3A
Clone1 Clone2 Clone1 Clone2
HA
70 KD
55 KD
C.
Control
EED-/-
EED-/- + HA-EED
EED-/- + HA-EED-3A
Clone1 Clone2 Clone1 Clone2 Clone1 Clone2
H3K27me3
17 KD
H3
D.
Dox Removal Day 8
EED-/-
HA-EED
HA-EED-3A
1.9%
94%
23%
Count
GFP

Figure 14 (Continued)

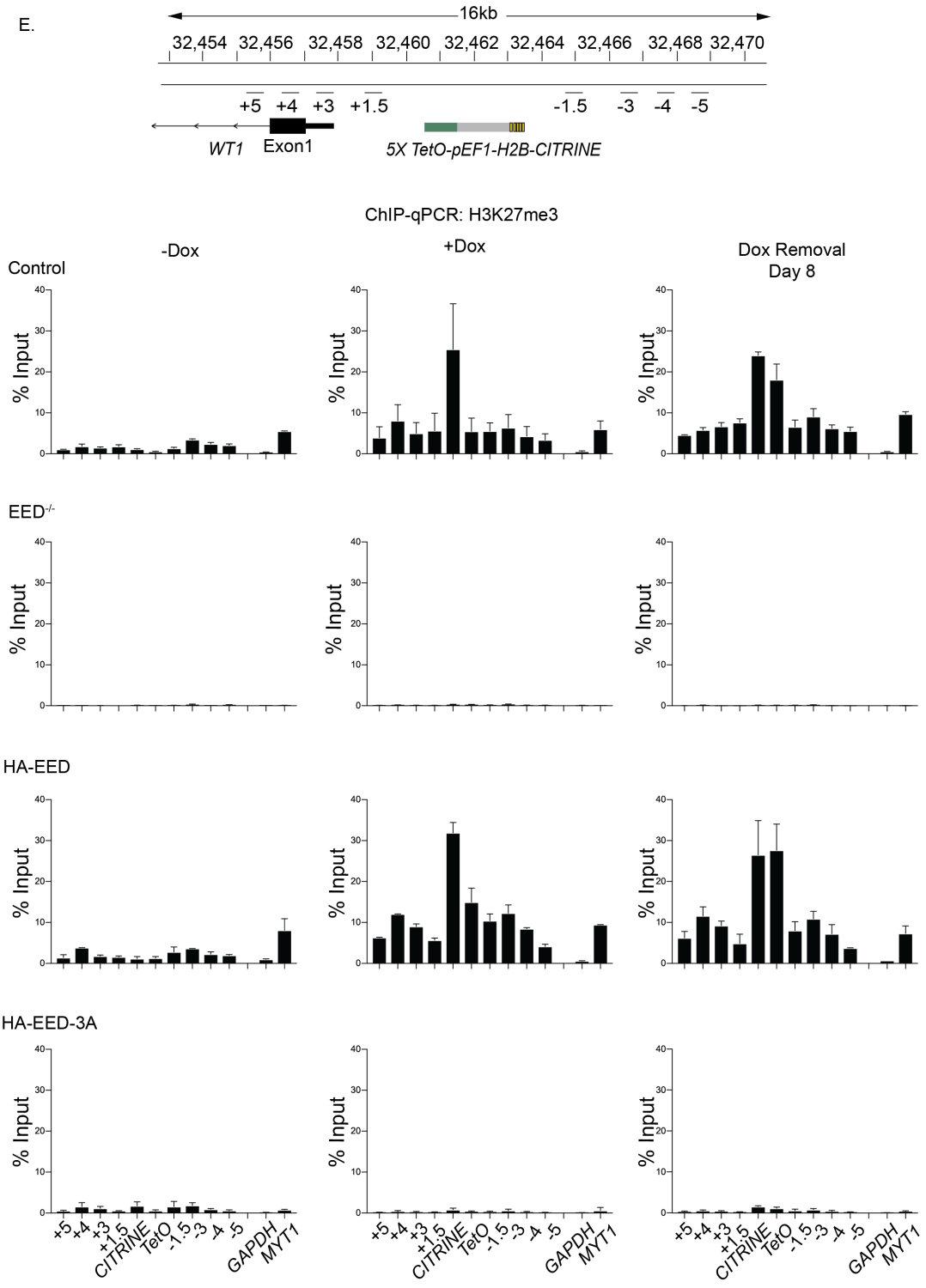


Figure 14 (Continued)

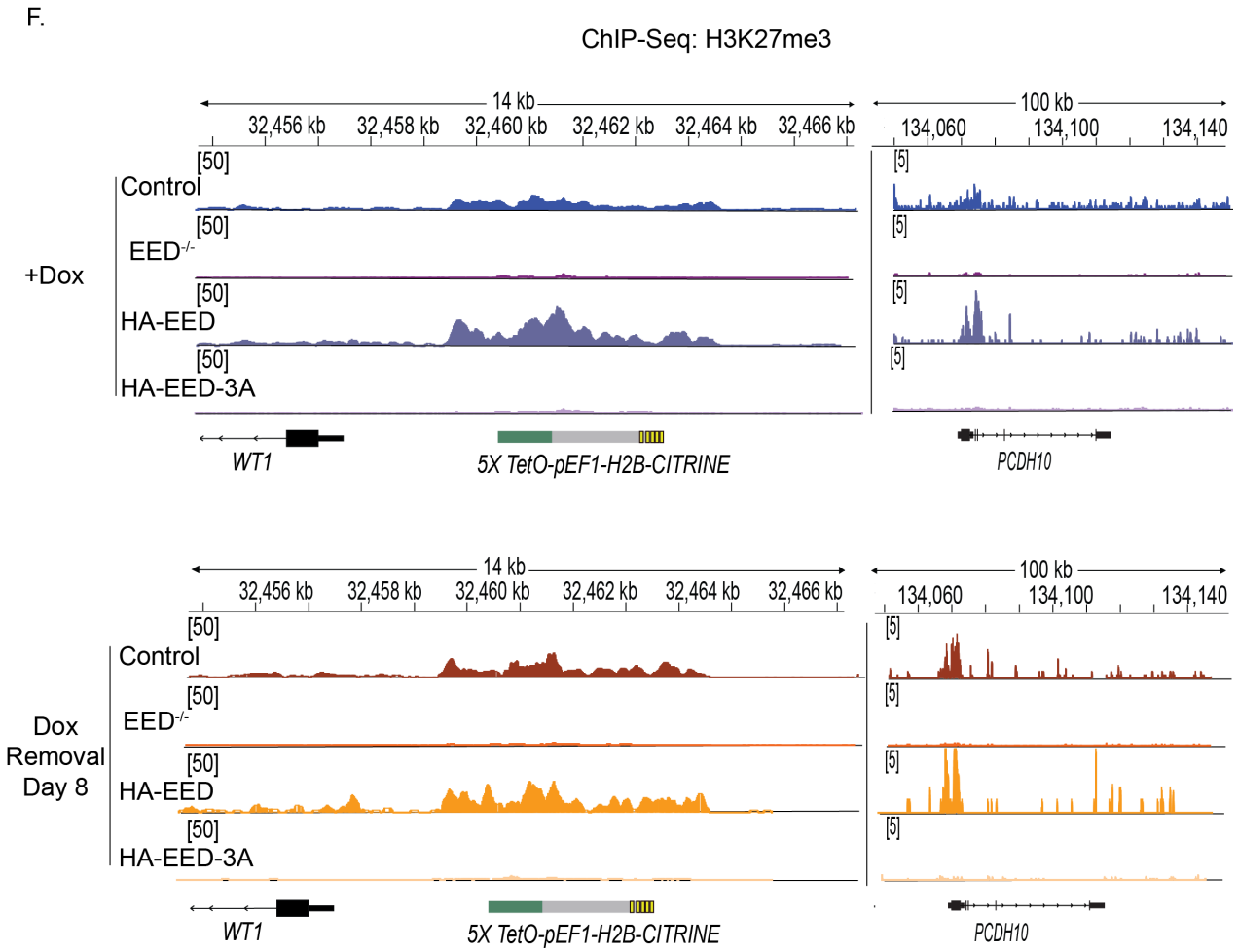


Figure 14. Read-Write mechanism of EED-PRC2 is required for inheritance of Polycomb-Mediated Gene Silencing.

(A) Representative Flow Cytometry histograms of control and *EED*$^{-/-}$ cells to show the percentage of silenced cells resulting from tethering of the rTetR-CBX7 to the reporter inserted at the *WT1* locus at Day8 in doxycycline-containing medium (+Dox) and after transfer to doxycycline free media at Day 8 (Dox Removal). Percentages (%) show fraction of silenced cells.

(B) Western blot with anti-HA antibody to detect the expression of HA-tagged WT (HA-EED) or aromatic cage mutant EED (HA-EED-3A) in *EED*$^{-/-}$ Cells. The aromatic cage mutant has three residue changes: F97A, Y148A and Y365A.

(C) Western Blot showing the absence of H3K27me3 in whole cell *EED*$^{-/-}$ extracts and its restoration by the expression of EED-HA but not EED-HA-3A cells. H3 is used as a loading control.

Figure 14 (Continued)

(D) Representative Flow Cytometry histograms of *EED*$^{-/-}$, *EED-HA* and *EED-HA-3A* cells to show the percentage of silenced cells resulting from the removal of doxycycline from the growth medium to release rTetR-CBX7 at Day 8 (Dox Removal). Percentages (%) show fraction of silenced cells.

(E) ChIP-qPCR of H3K27me3 in WT, *EED*$^{-/-}$, *EED-HA* and *EED-HA-3A* cells. Cells that are expressing rTetR-CBX7 but have not had it tethered to the reporter (-Dox) serve as controls for cells that have tethered rTetR-CBX7 in doxycycline medium (+Dox) to establish silencing and cells that have released rTetR-CBX7 (Dox Removal Day 8). Primer positions are indicated. *GAPDH* is a negative control and *MYT1* is a positive control for H3K27me3. Error bars are deviation from the mean of at least two replicates.

(F) ChIP-Seq of H3K27me3 at the *WT1* locus (left) and a positive control locus *PCDH10* (right) in WT, *EED*$^{-/-}$, *EED-HA* and *EED-HA-3A* cells . Cells with rTetR-CBX7 tethered are in doxycycline medium to establish silencing (+Dox). Switch to doxycycline free media releases tethering of rTetR-CBX7 (Dox Removal Day8).

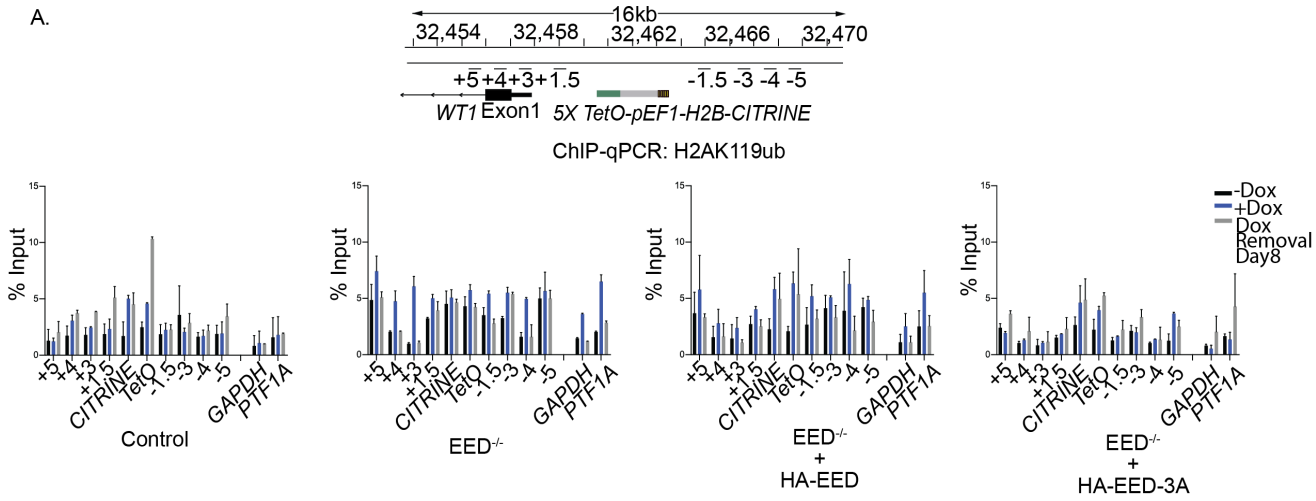


Figure 15. ChIP-qPCR of H2AK119ub1 in *EED*$^{-/-}$ and rescue cells. (A) ChIP-qPCR of H2AK119ub1 in *WT, EED*$^{-/-}$, *EED-HA* and *EED-HA-3A* cells. Cells that are expressing rTetR-CBX7 but have not had it tethered to the reporter serve as controls (-Dox) for cells that have tethered rTetR-CBX7 in doxycycline medium (+Dox) to establish silencing and cells that have released rTetR-CBX7 (Dox Removal Day 8). Primer positions are indicated. *GAPDH* is a negative control and *PTF1A* is a positive control for H2AK119ub1. Error bars are deviation from the mean of at least two replicates.

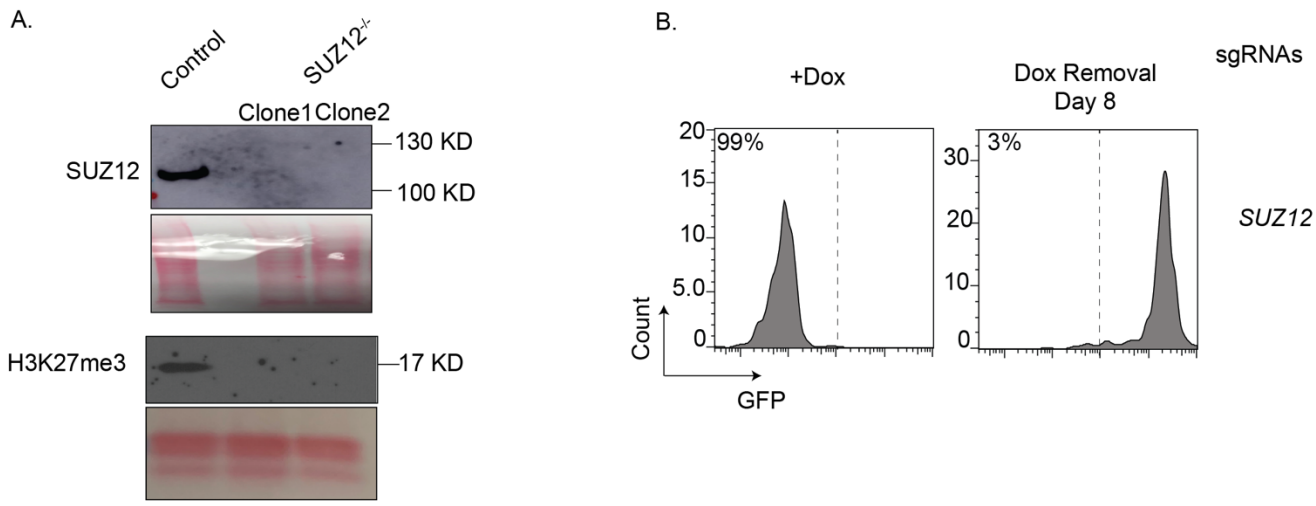


Figure 16. SUZ12 is necessary for the inheritance of Polycomb-Mediated Gene Silencing.

(A) Western blot with anti-SUZ12 and anti-H3K27me3 antibody showing the absence of SUZ12 protein and H3K27me3 in *SUZ12*$^{-/-}$ cells, respectively.

(B) Representative Flow Cytometry histograms of *SUZ12*$^{-/-}$ cells to show the percentage of silenced cells resulting from tethering of the rTetR-CBX7 to the reporter inserted at the *WT1* locus at Day8 in doxycycline-containing medium (+Dox) and after transfer to doxycycline free media at Day 8 (Dox Removal). Percentages (%) show fraction of silenced cells.

The Role of PRC1 in Epigenetic Inheritance of Polycomb-Mediated Gene Silencing

To investigate the role of the PRC1 complex in epigenetic inheritance, I deleted both the *RING1A* and *RING1B* subunits of PRC1 from the cells carrying the *CITRINE* reporter at the *WT1* locus. As expected, deletion of *RING1A* and *RING1B* (*RING1A/B*$^{-/-}$) abolished H2AK119ub1 (Figure 17A). The establishment of *CITRINE* reporter silencing was unaffected in *RING1A/B*$^{-/-}$ cells, but the maintenance of silencing was greatly diminished (Figure 17B). ChIP-qPCR and ChIP-Seq of H3K27me3 showed that during establishment, H3K27me3 was deposited but was greatly decreased by 8 days after release of the rTetR-CBX7 concomitant with de-repression of the *CITRINE* reporter (Figure 17C-D). This suggests that in the absence of H2KA119ub1, the residual H3K27me3 was not sufficient to maintain the silenced state. These observations are also

consistent with the previously reported requirement for H2AK119ub1 in maintenance of H3K27me3 in mESCs and with the possibility of self-reinforcing interactions between the PRC2 and PRC1 complex to maintain the inheritance of the silenced state[49,50]. Additional experiments are needed to test for the presence of specific feedback interactions between PRC2 and PRC1 complex to maintain the inheritance of silencing at this reporter locus.

Altogether, the above experiments show that both H3K27me3, PRC2-EED read-write, and PRC1-RING1A/B contribute to the stable maintenance of Polycomb-mediated gene silencing at the Polycomb target gene *WT1*, and reveal a role for PRC2 in epigenetic inheritance that is independent of its read-write capability and H3K27me3.

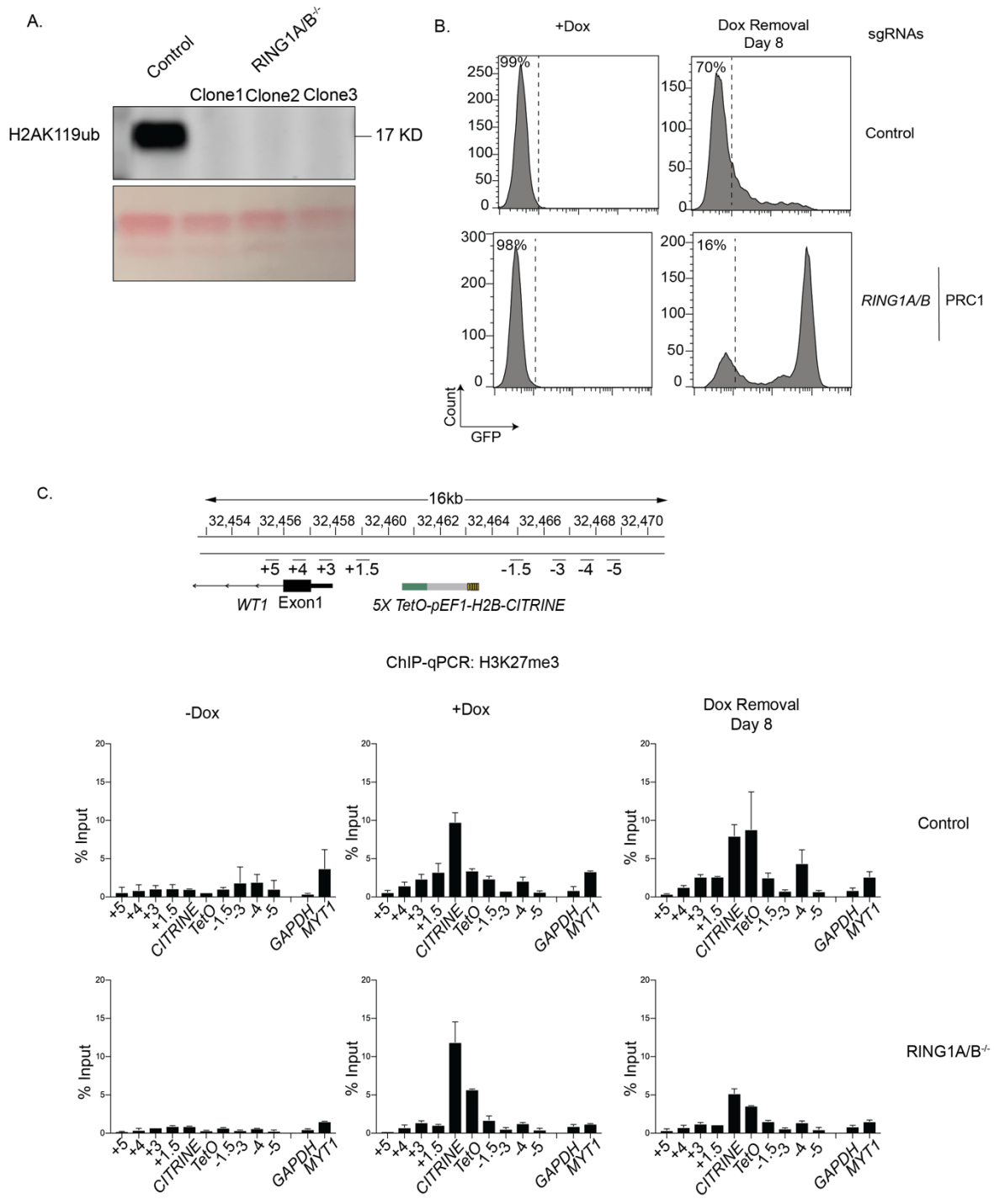

A.
Control
RING1A/B-/-
Clone1 Clone2 Clone3
H2AK119ub
17 KD
B.
+Dox
Dox Removal
Day 8
sgRNAs
99%
70%
Control
98%
16%
RING1A/B
PRC1
Count
GFP
C.
16kb
32,454 32,456 32,458 32,460 32,462 32,464 32,466 32,468 32,470
+5 +4 +3 +1.5 -1.5 -3 -4 -5
WT1 Exon1
5X TetO-pEF1-H2B-CITRINE
ChIP-qPCR: H3K27me3
-Dox
+Dox
Dox Removal
Day 8
% Input
Control
RING1A/B-/-
+5 +4 +3 +1.5 CITRINE TetO -1.5 -3 -4 -5 GAPDH MYT1

Figure 17 (Continued)

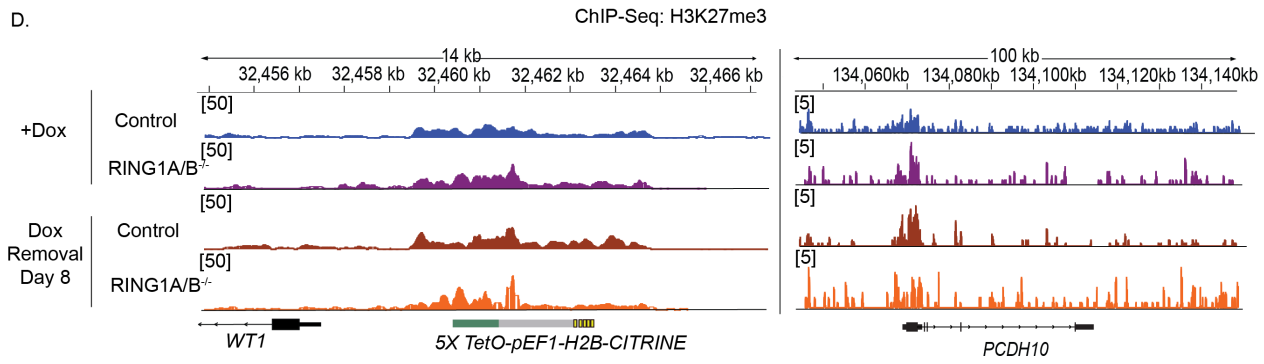


Figure 17. RING1A/B is necessary for inheritance of Polycomb-Mediated Gene Silencing .

(A) Western blot showing the loss of H2AK119ub1 in *RING1A/B*$^{-/-}$ *WT1* reporter cell line.

(B) Representative Flow Cytometry histograms of control and *RING1A/B*$^{-/-}$ cells to show the percentage of silenced cells resulting from tethering of the rTetR-CBX7 to the reporter inserted at the *WT1* locus at Day8 in doxycycline-containing medium (+Dox) and after transfer to doxycycline free media at Day 8 (Dox Removal). Percentages (%) show fraction of silenced cells.

(C) ChIP-qPCR of H3K27me3 in WT and *RING1A/B*$^{-/-}$ cells. Cells that express rTetR-CBX7 but have not had it tethered to the reporter (-Dox) serve as controls for cells that have tethered rTetR-CBX7 in doxycycline medium (+Dox) to establish silencing and cells that have released rTetR-CBX7 (Dox Removal Day 8). Primer positions are indicated. *GAPDH* is a negative control and *MYT1* is a positive control for H3K27me3. Error bars are deviation from the mean of at least two replicates.

(D) ChIP-Seq of H3K27me3 at the *WT1* locus (left) and a positive control locus *PCDH10* (right) in WT and *RING1A/B*$^{-/-}$ cells. Cells with rTetR-CBX7 tethered are in doxycycline medium to establish silencing (+Dox). Switch to doxycycline free media releases rTetR-CBX7 (Dox Removal Day8).

3.2.2 Investigating the Role of Locus-Dependent Factors in Epigenetic Inheritance of Polycomb-Mediated Gene Silencing

Contribution of DNA-Sequence to Epigenetic Inheritance at Polycomb Target Genes

As previously described, PRC2 has multiple accessory subunits, which include DNA binding proteins that have been shown to bind to CG-rich DNA in vitro and are localized to CGIs along with PRC2 in vivo. MTF2 is one such DNA binding protein that interacts with PRC2 through the SUZ12 subunit and has the ability to bind to CG-rich DNA through its Extended Homology Domain. To test the possible role of MTF2 in inheritance of Polycomb silencing, I deleted *MTF2* from cell with the reporter inserted at the *WT1* locus (Figure 18A). *MTF2*$^{-/-}$ cells showed partial loss of reporter silencing by 8 days after release of the rTetR-CBX7 and complete loss by 16 days after release (Figure 18A). These results indicate that the PRC2.1 subunit MTF2 is required to maintain silencing at the *WT1* locus.

To determine whether the DNA binding activity of MTF2 was necessary for the maintenance of *CITRINE* silencing, I introduced 3xFLAG-tagged wild-type MTF2 (*FLAG-MTF2*) or mutant Extended Homology Domain MTF2 (*FLAG-MTF2-EH*), which does not bind DNA, into *MTF2*$^{-/-}$ cells (Figure 18B). The *MTF2*$^{-/-}$ cells overexpressing FLAG-MTF2 restored maintenance of *CITRINE* reporter silencing, while the FLAG-MTF2-EH mutant did not (Figure 18C). ChIP-qPCR and ChIP-Seq for the FLAG tag in *FLAG-MTF2* rescue cells demonstrated that MTF2 was recruited during establishment and maintenance. In *FLAG-MTF2-EH* rescue cells, where the DNA Binding activity of MTF2 is abolished, MTF2 was not recruited during establishment and maintenance (Figure 18D-E). ChIP-qPCR and ChIP-Seq for H3K27me3 in *MTF2*$^{-/-}$ showed reduced levels during maintenance at the reporter which was restored in the

FLAG-MTF2 cells during maintenance. However, in *FLAG-MTF2-EH* cells, H3K27me3 levels were not restored, and were even more reduced likely due to a dominant negative effect of the mutant (Figure 18F-G). H2AK119ub1 was present at the *MTF2*$^{-/-}$, *FLAG-MTF2* and *FLAG-MTF2-EH* cells 8 days after release of rTetR-CBX7, demonstrating maintenance of H2AK119 ubiquitination in the reduced levels of H3K27me3 (Figure 19A). These results demonstrate that the DNA binding activity of MTF2 is required for maintenance of rTetR-CBX7-induced silencing at the *WT1* locus.

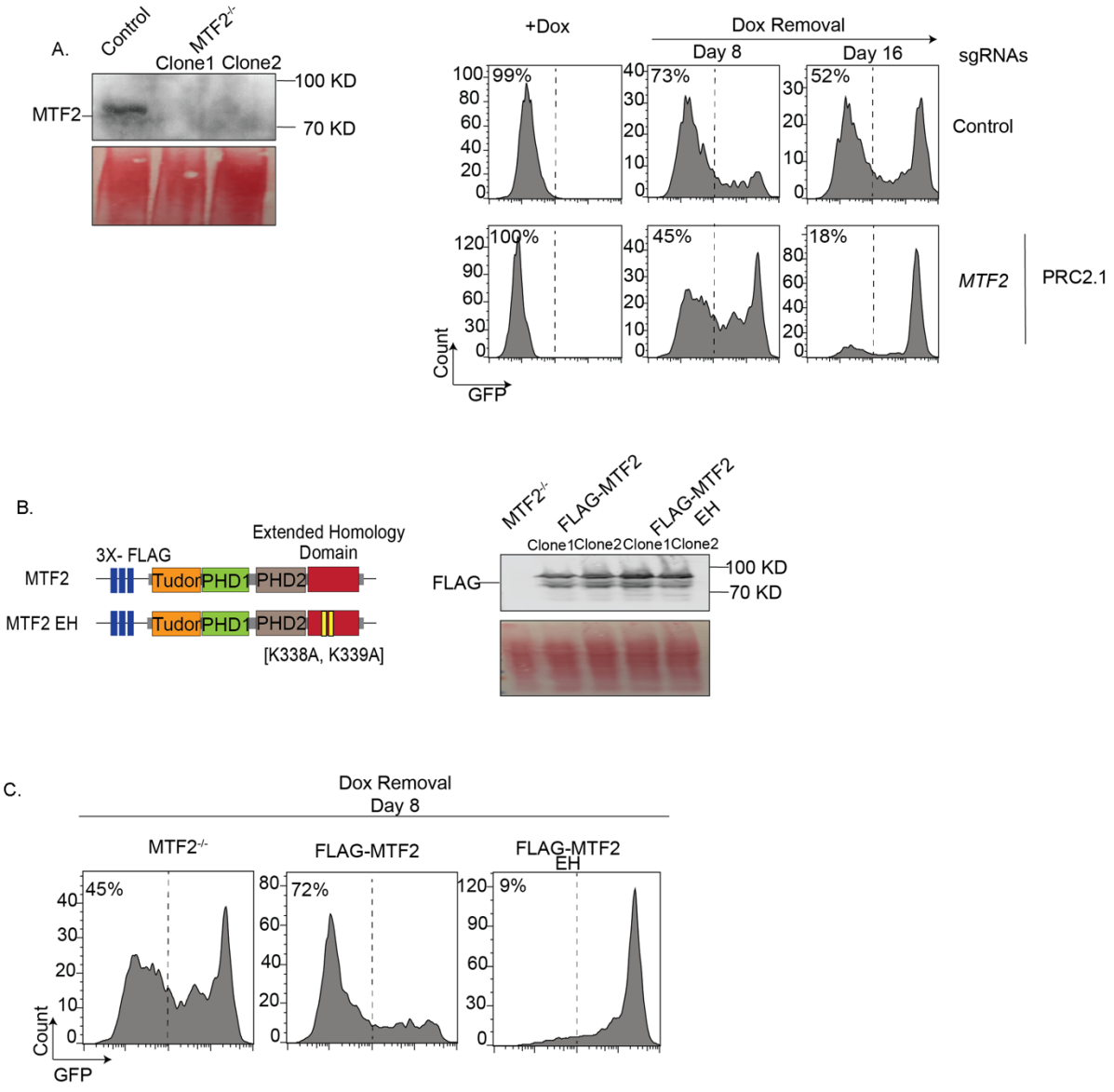

A.
Control
MTF2-/-
Clone1 Clone2
MTF2
100 KD
70 KD
+Dox
Dox Removal
Day 8
Day 16
sgRNAs
99%
73%
52%
Control
100%
45%
18%
MTF2
PRC2.1
Count
GFP
B.
3X- FLAG
Extended Homology Domain
MTF2
Tudor
PHD1
PHD2
MTF2 EH
[K338A, K339A]
MTF2-/-
FLAG-MTF2
FLAG-MTF2 EH
Clone1 Clone2 Clone1 Clone2
FLAG
100 KD
70 KD
C.
Dox Removal
Day 8
MTF2-/-
FLAG-MTF2
FLAG-MTF2 EH
45%
72%
9%
Count
GFP

Figure 18 (Continued)

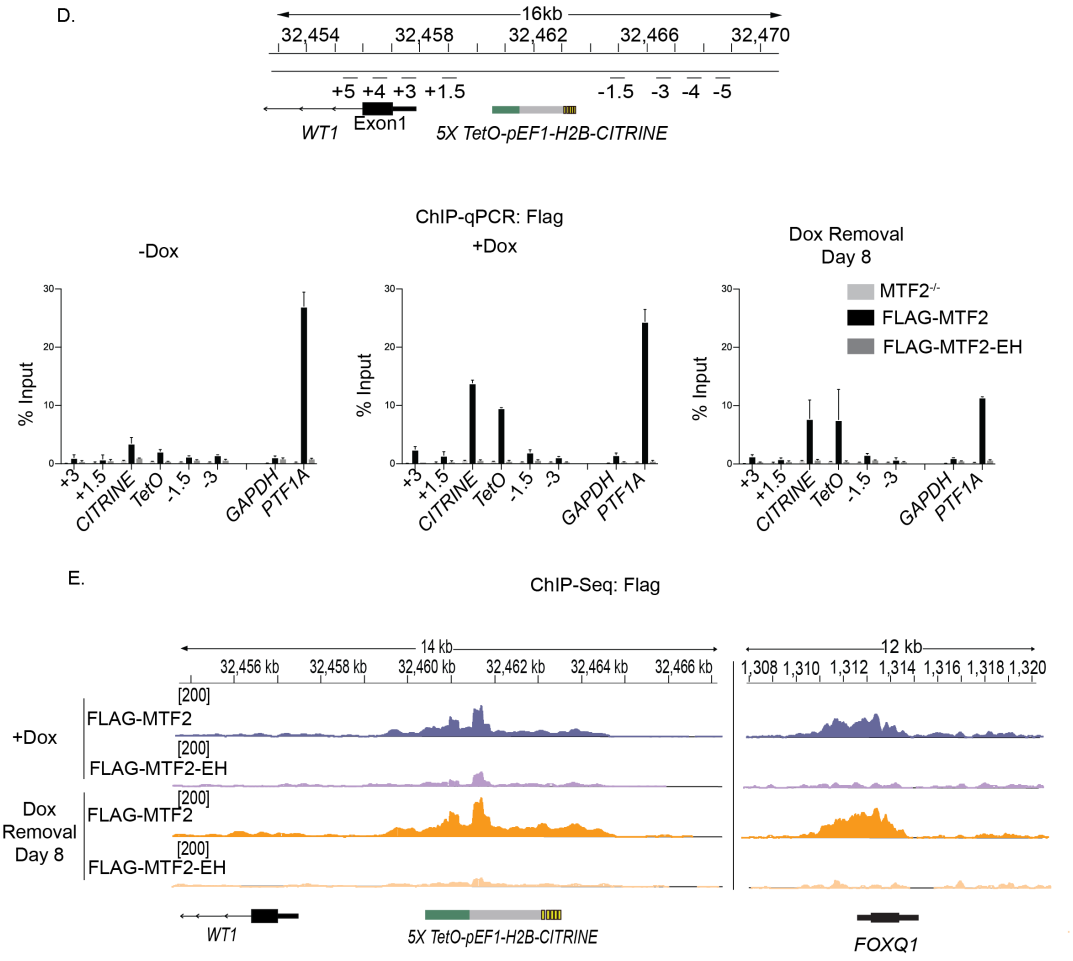


Figure 18 (Continued)

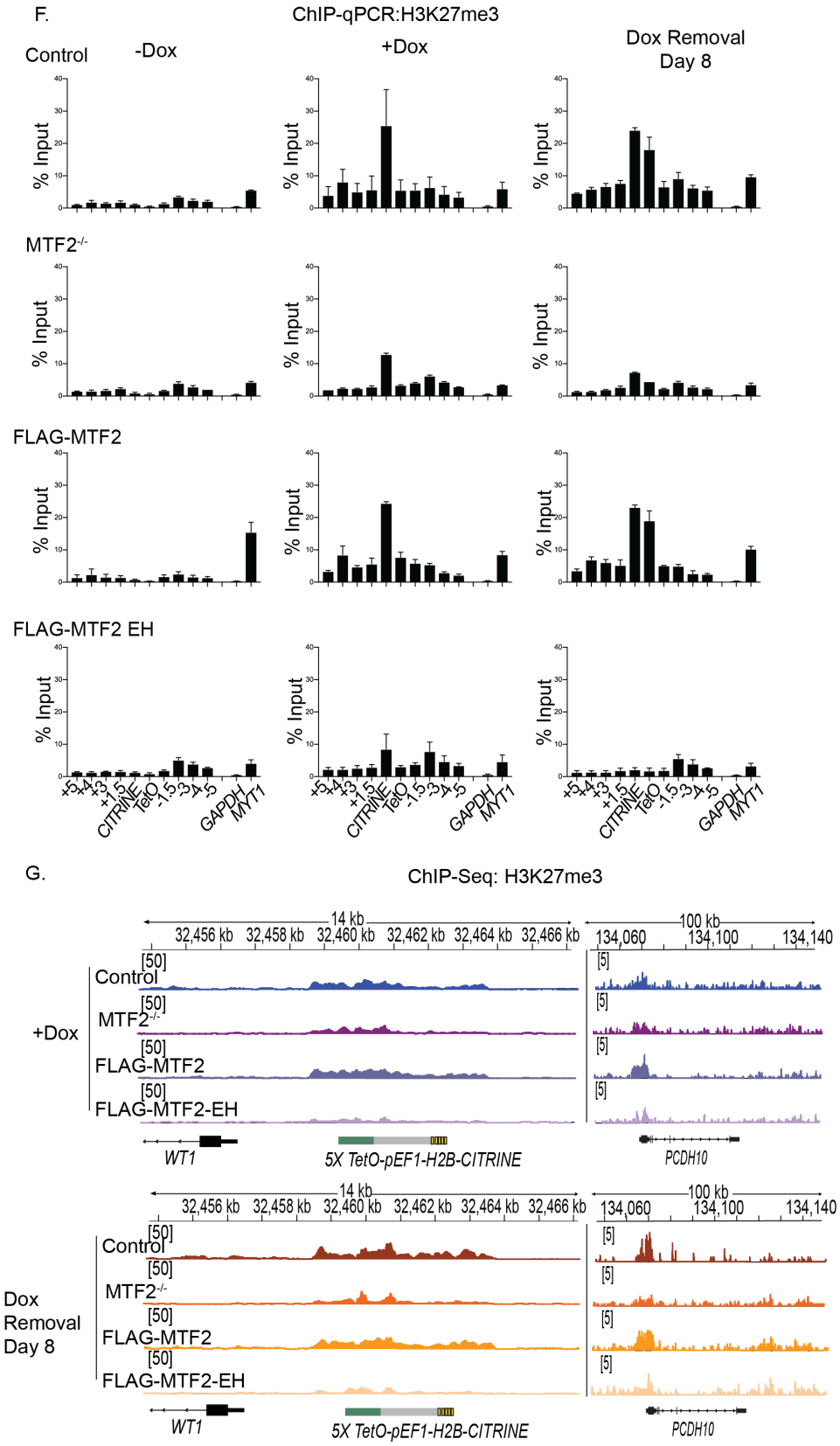


Figure 18 (Continued)

Figure 18. The DNA binding domain of MTF2 is required for the maintenance of *5x-tetO-H2B-CITRINE* silencing at the *WT1* locus.

(A) Western Blot with anti-MTF2 antibody showing the absence of the MTF2 protein in *MTF2*$^{-/-}$ cells (top). Representative Flow Cytometry histograms of control and *MTF2*$^{-/-}$ cells showing the percentage of silenced cells resulting from tethering of the rTetR-CBX7 to the reporter inserted at the *WT1* locus at Day8 in doxycycline-containing medium (+Dox) and after transfer to doxycycline free media at Day 8 and 16 (Dox Removal). Percentages (%) show fraction of silenced cells (bottom) .

(B) Western blot with anti-FLAG antibody to detect the overexpression of FLAG-tagged WT (FLAG-MTF2) or DNA binding mutant (FLAG-MTF2-EH) MTF2 in *MTF2*$^{-/-}$ Cells. The DNA-binding mutant has two residue changes: K338A, K339A.

(C) Representative Flow Cytometry histograms of *MTF2*$^{-/-}$, *FLAG-MTF2* and *FLAG-MTF2-EH* cells showing the percentage of silenced cells resulting from the removal of doxycycline from the growth medium to release rTetR-CBX7 at Day 8 (Dox Removal). Percentages (%) show fraction of silenced cells.

(D) ChIP-qPCR of FLAG in *MTF2*$^{-/-}$, *FLAG-MTF2* and *FLAG-MTF2-EH* cells. Cells that express rTetR-CBX7 but have not had it tethered to the reporter (-Dox) serve as controls for cells that have tethered rTetR-CBX7 in doxycycline medium to establish silencing (+Dox) and cells that have released rTetR-CBX7 (Dox Removal Day 8). Primer positions are indicated. *GAPDH* is a negative control and *PTF1A* is a positive control for MTF2. Error bars are deviation from the mean of at least two replicates.

(E) ChIP-Seq of FLAG at the *WT1* locus (left) and a positive control locus *FOXQ1* (right) in *FLAG-MTF2* and *FLAG-MTF2-EH* cells. Cells with rTetR-CBX7 tethered are in doxycycline medium to establish silencing (+Dox). Switch to doxycycline free media releases tethering of rTetR-CBX7 (Dox Removal Day8).

(F) ChIP-qPCR of H3K27me3 in WT, *MTF2*$^{-/-}$, *FLAG-MTF2* and *FLAG-MTF2-EH* cells. Cells that express rTetR-CBX7 but have not had it tethered to the reporter (-Dox) serve as controls for cells that have tethered rTetR-CBX7 in doxycycline medium (+Dox) to establish silencing and cells that have released rTetR-CBX7 (Dox Removal Day 8). Primer positions are indicated. *GAPDH* is a negative control and *MYT1* is a positive control for H3K27me3. Error bars are deviation from the mean of at least two replicates.

(G) ChIP-Seq of H3K27me3 at the *WT1* locus (left) and a positive control locus *PCDH10* (right) in WT, *MTF2*$^{-/-}$, *FLAG-MTF2* and *FLAG-MTF2-EH* cells. Cells with rTetR-CBX7 tethered are in doxycycline medium to establish silencing (+Dox). Switch to doxycycline free media releases tethering of rTetR-CBX7 (Dox Removal Day8).

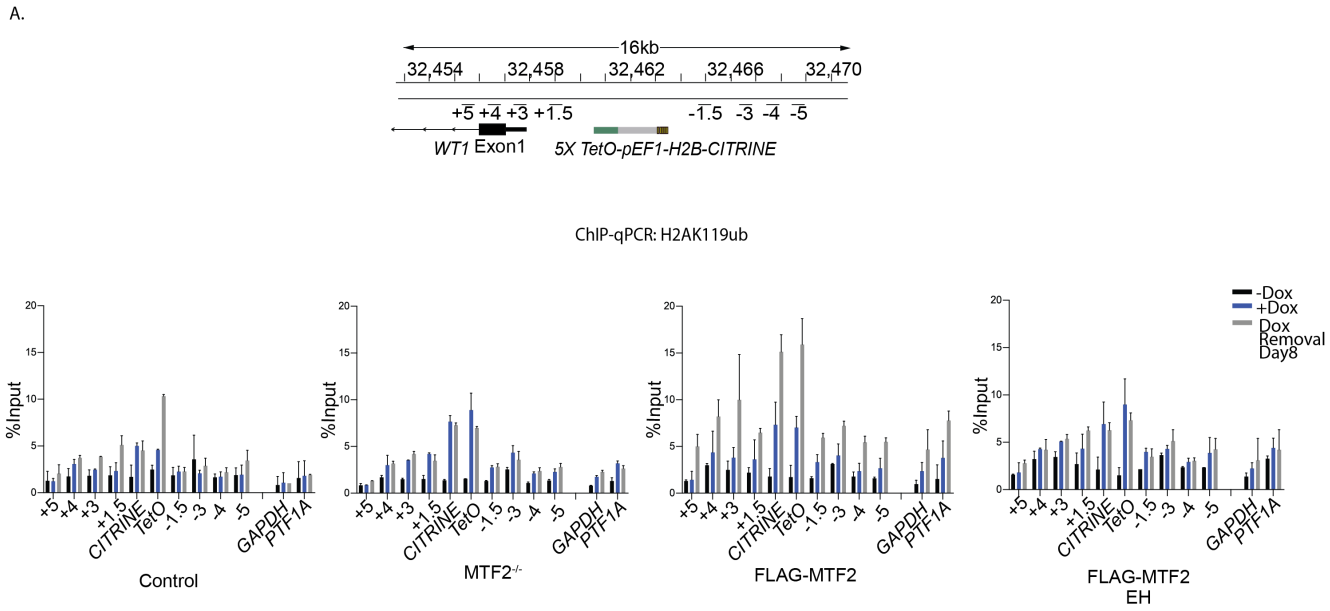


Figure 19. ChIP-qPCR of H2AK119ub1 in *MTF2*$^{-/-}$ and rescue cells . (A) ChIP-qPCR of H2AK119ub1 in WT, *MTF2*$^{-/-}$, *FLAG-MTF2* and *FLAG-MTF2-EH* cells. Cells that express rTetR-CBX7 but have not had it tethered to the reporter (-Dox) serve as controls for cells that have tethered rTetR-CBX7 in doxycycline medium to establish silencing (+Dox) and cells that have released rTetR-CBX7 (Dox Removal Day 8). Primer positions are indicated. *GAPDH* is a negative control and *PTF1A* is a positive control for H2AK119ub1. Error bars are deviation from the mean of at least two replicates.

Comparing the decay rate of *CITRINE* reporter silencing after the release of rTetR-CBX7 in *MTF2*$^{-/-}$ cells with *PRC2*$^{-/-}$ or *PRC1*$^{-/-}$ cells (*EED*$^{-/-}$, *SUZ12*$^{-/-}$ and *RING1A/B*$^{-/-}$) showed a slower rate of decay in *MTF2*$^{-/-}$ cells. This was also reflected in the slower decay of H3K27me3 and H2AK119ub1 modifications. These results imply that there might be a redundant DNA binding protein that can partially compensate for MTF2 loss in HEK293FT cells. To investigate this possibility, I deleted three other DNA-binding proteins that are accessory to PRC2. I made deletions of PRC2.2 subunit JARID2 and PRC2.1 subunits PHF1 and PHF19 in the cell line containing the *CITRINE* reporter at the *WT1* locus. I found that deletion of each *JARID2* or *PHF19* had no effect on establishment and maintenance (Figure 20A-B), but deletion of *PHF1*

led to a slow loss in maintenance by Day 16 after rTetR-CBX7 release, similar to what I observed in *MTF2*$^{-/-}$ cells (Figure 20C).

There are several likely explanations for lack of a maintenance phenotype in *JARID2*$^{-/-}$ cells. It has been hypothesized that JARID2 is necessary for differentiating ESCs into specific lineages but is dispensable in committed cells[31]. Furthermore, JARID2's major function is to recognize H2AK119ub1 to link the feedback loop of H2AK119ub1 and PRC2.2. It has been hypothesized that Polycomb proteins SCML1 and SCML2 can form bridges between PRC1 and PRC2 in mammals, as demonstrated in flies[110]. This may explain why JARID2 is dispensable in differentiated cells[110]. Moreover, the feedback loop of JARID2-PRC2.2-H2AK119ub1-PRC1 may be redundant and its loss may be compensated by the cPRC1-vPRC1-PRC2.1 feedback loops. It has also been reported that PRC2.1 and PRC2.2 compete, and that the loss of PRC2.2-JARID2 increases PRC2 occupancy in hESCs[35]. This model of competition may explain the increase in the percentage of silenced cells observed during maintenance in *JARID2*$^{-/-}$ relative to wild-type cells (compare Figure 20A to Figure 6B).

My findings suggest that MTF2 and PHF1 can compensate for one another to a limited extent, but epigenetic maintenance is eventually lost in the absence of either factor. However, the other CG-rich DNA-binding protein, PHF19, is not required for epigenetic maintenance in HEK293FT cells. It is still unclear if the three PCLs can functionally compensate for one another or not, but they have overlapping localization patterns at majority of target sites in mESCs[19]. However, we do know that they are expressed at different levels in different lineages, for instance MTF2 is the dominant PCL in mESCs while in neural progenitor cells, MTF2 is downregulated and PHF1 and PHF19 are upregulated[111–113]. It is possible that MTF2 and PHF1 are the dominant PCLs in HEK293FTs. Furthermore, these results point to a possible function of

MTF2 and/or PHF1 in maintenance of the silenced state, possibly by increasing the residency time of PRC2 on chromatin through contact with CG-rich DNA[92].

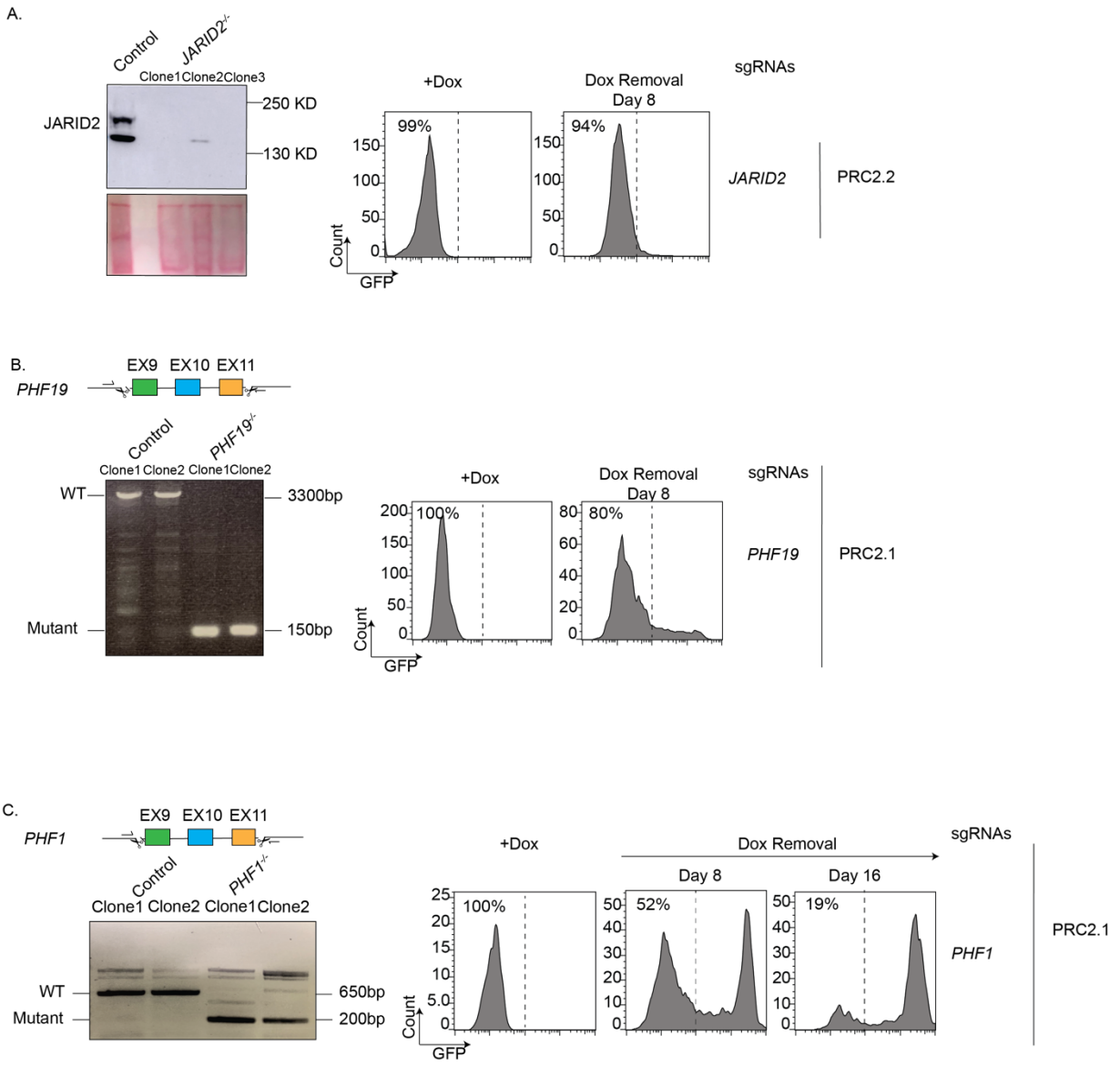


Figure 20. PHF1 is required for the maintenance of *5x-tetO-H2B-CITRINE* silencing at the *WT1* locus.

(A) Western Blot with anti-JARID2 antibody showing the absence of JARID2 protein in *JARID2*$^{-/-}$ cells (left). Representative Flow Cytometry histograms of control and *JARID2*$^{-/-}$ cells showing the percentage of silenced cells resulting from tethering of the rTetR-CBX7 to the reporter inserted at the *WT1* locus at Day8 in doxycycline-containing medium (+Dox) and after transfer to doxycycline free media at Day 8 (Dox Removal). Percentages (%) show fraction of silenced cells (right) .

(B) PCR genotyping to detect the excision of Exon 9-11 in *PHF19*$^{-/-}$ cells. Expected WT band is ~3.3kb while Mutant is ~0.15kb (left). Representative Flow Cytometry histograms of control and *PHF19*$^{-/-}$ cells showing the percentage of silenced cells resulting from tethering of the rTetR-CBX7 to the reporter inserted at the *WT1* locus at Day8 in doxycycline-containing medium

Figure 20 (Continued)

(+Dox) and after transfer to doxycycline free media at Day 8 (Dox Removal). Percentages (%) show fraction of silenced cells (right).

(C) PCR genotyping showing the excision of Exon 9-11 in *PHF1*$^{-/-}$ cells. Expected WT band is ~0.65kb and Mutant is ~0.2kb. (left). Representative Flow Cytometry histograms of control and *PHF1*$^{-/-}$ cells to show the percentage of silenced cells resulting from tethering of the rTetR-CBX7 to the reporter inserted at the *WT1* locus at Day8 in doxycycline-containing medium (+Dox) and after transfer to doxycycline free media at Day 8 and 16 (Dox Removal). Percentages (%) show fraction of silenced cells (right).

It has previously been reported that Polycomb targets have a high CGI density relative to non-Polycomb targets[26,103]. Examining the CGI density at Polycomb targets, *WT1*, *EN2* and *HOXD11* and housekeeping targets, *TFRC* and *B2M*, showed that these Polycomb targets indeed have multiple CGIs while the housekeeping loci have one CGI overlapping the promoter region (Figure 6, 8, 9 and 10). The reporter cassette with the EF1 promoter contains a CG-rich promoter region as well. My findings above suggest that CG-Rich DNA binding activity of MTF2/PHF1 are necessary for epigenetic inheritance. This raises the question as to whether a lack of MTF2/PHF1 recruitment to the housekeeping loci, which have only one promoter-associated CGI, may explain the rapid loss of inheritance at these genes. ChIP-qPCR for MTF2 showed that as expected, it was bound to the *WT1* locus during both the establishment and maintenance of the silenced state (Figure 21A) and was also present at the *TFRC* reporter locus during the establishment of silencing (Figure 21B). The spreading of MTF2 correlated with the spreading of H3K27me3 (i.e limited to the reporter cassette) at the *TFRC* locus. However, this result does not exclude the possibility that the ability of MTF2 to interact with additional CGIs outside the promoter region is critical for epigenetic maintenance. Alternatively, the loss of *CITRINE*

reporter maintenance at the housekeeping genes could be due to proximity to continuously active transcription, which may package their CGI promoters in nucleosomes that repel PRC2/MTF2.

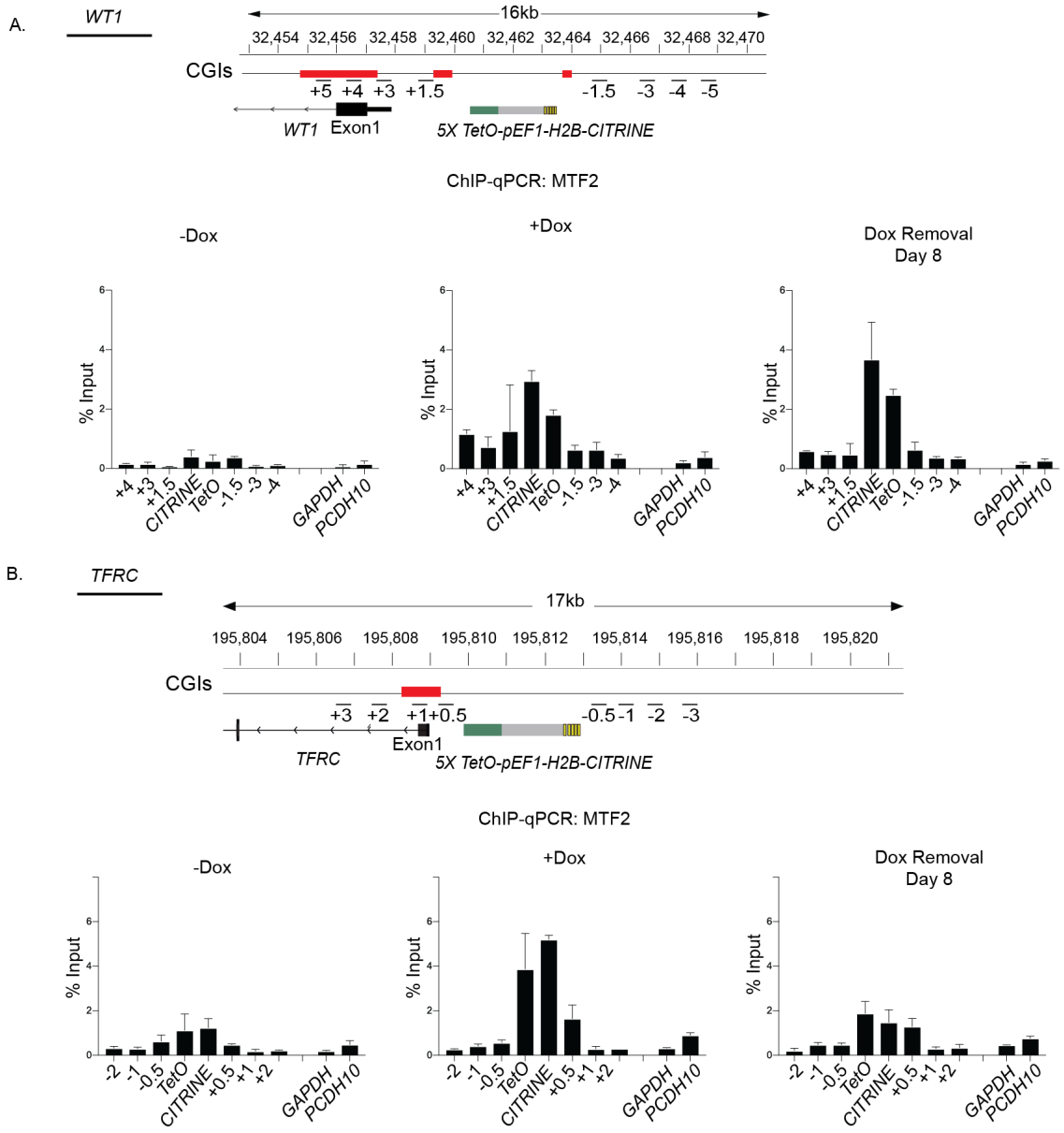


Figure 21. MTF2 is recruited during establishment to *5xtetO-H2B-CITRINE* at the *WT1* and *TFRC* loci.

(A) ChIP-qPCR of MTF2 in the *WT1* reporter cell line. Cells that express rTetR-CBX7 but have not had it tethered to the reporter (-Dox) serve as controls for cells that have tethered rTetR-CBX7 in doxycycline medium to establish silencing (+Dox) and cells that have released rTetR-CBX7 (Dox Removal Day 8). Primer positions are indicated. *GAPDH* is a negative control and

Figure 21 (Continued)

PCDH10 is a positive control for MTF2. Error bars are deviation from the mean of at least two replicates.

(B) ChIP-qPCR of MTF2 in *TFRC* reporter cells. Cells that express rTetR-CBX7 but have not had it tethered to the reporter (-Dox) serve as controls for cells that have tethered rTetR-CBX7 in doxycycline medium to establish silencing (+Dox) and cells that have released rTetR-CBX7 (Dox Removal Day 8). Primer positions are indicated. *GAPDH* is a negative control and *PCDH10* is a positive control for MTF2. Error bars are deviation from the mean of at least two replicates.

Role of Proximal Transcriptional Activity in Opposing Epigenetic Inheritance

Cis-acting proximal transcriptional activity may affect the epigenetic inheritance of Polycomb-mediated gene silencing. Upon examining the decay kinetics of the *CITRINE* reporter silencing at Polycomb and housekeeping target loci, I observed a correlation between the level of transcription at the target gene and the decay rate in loss of silencing of reporter. The Polycomb target genes have significantly lower steady state RNA levels relative to housekeeping targets, as verified by RNA-Seq, which may explain the slower decay of silencing at the proximal reporter during the maintenance phase. By 8 days after the release of rTetR-CBX7, 74%, 68% and 27% of cells maintained reporter gene silencing at the *WT1*, *HOXD11,* and *EN2* genes*,* respectively (Figure 13). On the other hand, housekeeping genes with higher steady state RNA levels showed a faster decay rate for silencing of the proximal reporter gene, with the reporter at the *TFRC* and *B2M* loci silenced in 2% and 7% of cells, respectively, by 8 days after release of rTetR-CBX7 (Figure 13).

To determine if proximal transcriptional activity has a role in the inheritance of the silenced state, I deleted specific sequences in cells with *CITRINE* reporters at the housekeeping target loci to attenuate transcriptional activity of the nearby housekeeping gene. At the *TFRC* locus, I deleted a 5.6 kb region from the end of the reporter to the end of Exon 2 of the *TFRC*

gene (Figure 22A). This deletion excised the majority of the promoter region and some regulatory elements. This is likely to eliminate some binding sites for transcriptional factors that regulates *TFRC* expression[114,115]. However, I noted that this specific sequence deletion does not completely excise all the transcription factor binding sites, specifically of c-Myc binding motifs that activates TFRC (as viewed through UCSC genome browser). Deletion of the 5.6kb sequence resulted in two heterozygote clones and one homozygote clone (Figure 22A). A similar approach was undertaken at the *B2M* housekeeping gene where a 4.4kb region was deleted from the end of the reporter to the end of Exon 1 (Figure 23A) to restrict transcription factor binding[116,117]. This resulted in one heterozygote clone and two homozygote clones. I note that in the heterozygote clones, the isolated deletions occurred only on chromosome with the reporter allele (the reporter is inserted at one out of possibly three alleles of *TFRC* and one of possibly two alleles of *B2M*). The heterozygote clones have intact sequences in their unmodified alleles that lack the reporter. The homozygote clones with the above sequence deletions at both housekeeping genes showed reduced steady state mRNA levels (Figure 22B, 23B).

At the *TFRC* gene, upon release of the rTetR-CBX7 from the sequence deleted clones, the heterozygote promoter region deletions showed a modest delay in the loss of maintenance of silencing 4 days after release of rTetR-CBX7, with 24% of the cells silenced in the heterozygous deletions relative to 4% at in control cells. The homozygote deletion clone showed a longer delay in the loss of silencing with 53% and 35% of the cells silenced at Day 4 and Day 8 after rTetR-CBX7 release, respectively (Figure 22C-D). Interestingly at the *B2M* locus, the heterozygote deletion clone did not show any delay in the loss of maintenance, but the homozygote deletion clones did, with 29% and 17% of the cells silenced compared to 13% and 7% in controls at Day

4 and Day 8 after rTetR-CBX7 release, respectively (Figure 23C-D). These results show that a reduction in the transcriptional activity at the proximal housekeeping genes could partially reverse the loss of the silenced state at the *CITRINE* reporter, suggesting that *cis* proximal transcriptional activity counteracts the epigenetic inheritance of Polycomb-mediated gene silencing. Proximal transcriptional activity is therefore likely to oppose the epigenetic inheritance of Polycomb-mediated gene silencing.

Comparing the decay rate of the silenced state at *WT1* and the *TFRC* homozygote clone at Day 8 (74% and 35% respectively) may provide additional insight into the contribution of TF-activity and CGI densities. Even though the *TFRC* homozygote cis deletion clones have reduced steady state mRNA levels, compared to wild-type intact *TFRC*, these RNA levels may still be higher than the endogenous expression levels of *WT1*, providing a possible explanation for the still faster loss of maintenance at these deletion clones relative to *WT1*. Additional qRT-PCR experiments will be carried out to test this hypothesis. An alternative explanation is that the *TFRC* homozygote deletion clone and endogenous *WT1* expression levels are comparable but the lack of CGIs at the *TFRC* locus is responsible for the faster rate of decay of silencing. These two possibilities are not mutually exclusive.

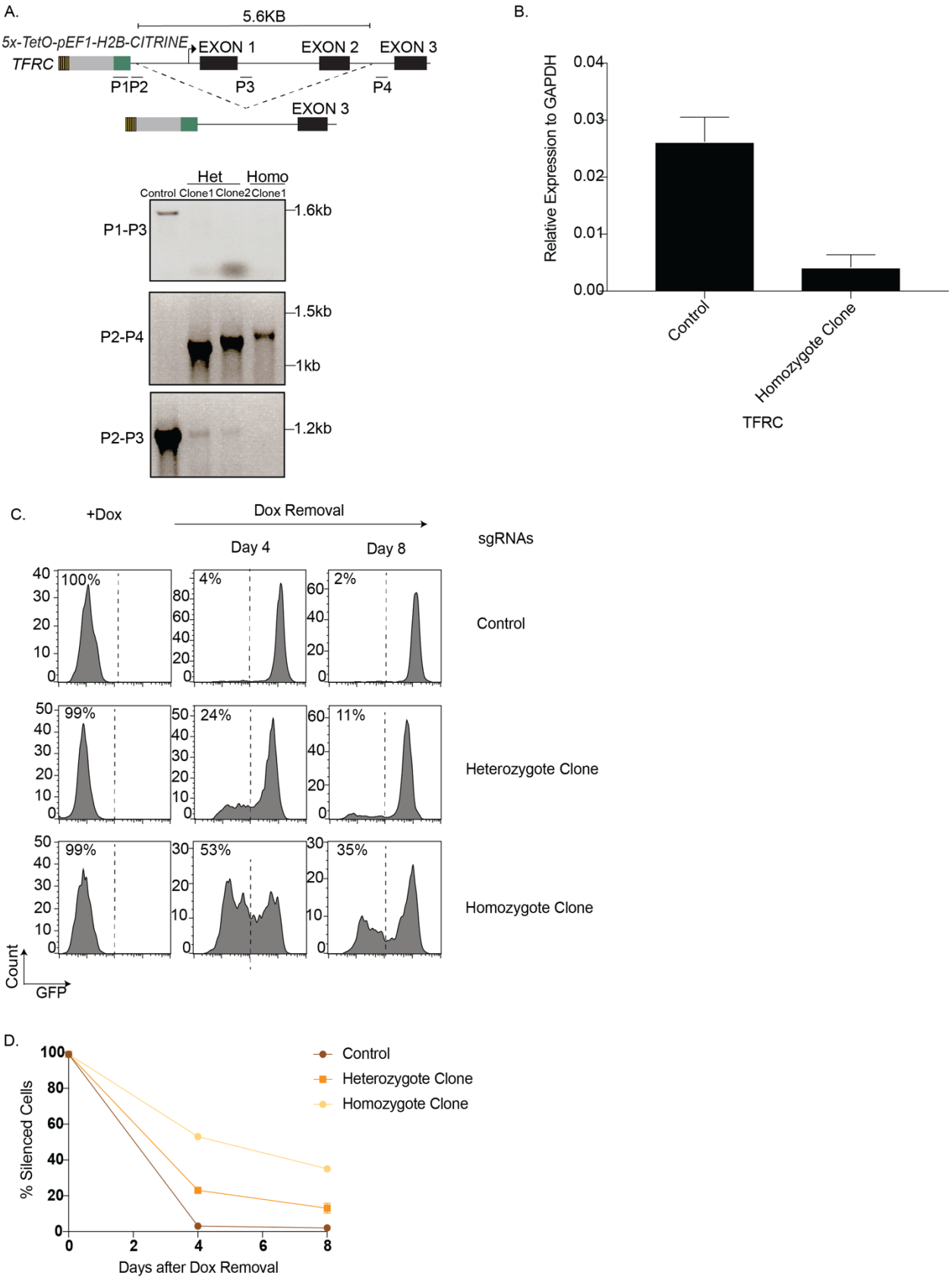

A.
5.6KB
5x-TetO-pEF1-H2B-CITRINE
EXON 1
EXON 2
EXON 3
TFRC
P1
P2
P3
P4
EXON 3
Het
Homo
Control
Clone1
Clone2
Clone1
P1-P3
1.6kb
P2-P4
1.5kb
1kb
P2-P3
1.2kb
B.
Relative Expression to GAPDH
0.04
0.03
0.02
0.01
0.00
Control
Homozygote Clone
TFRC
C.
+Dox
Dox Removal
Day 4
Day 8
sgRNAs
100%
4%
2%
Control
99%
24%
11%
Heterozygote Clone
99%
53%
35%
Homozygote Clone
Count
GFP
D.
% Silenced Cells
Days after Dox Removal
Control
Heterozygote Clone
Homozygote Clone

Figure 22 (Continued)

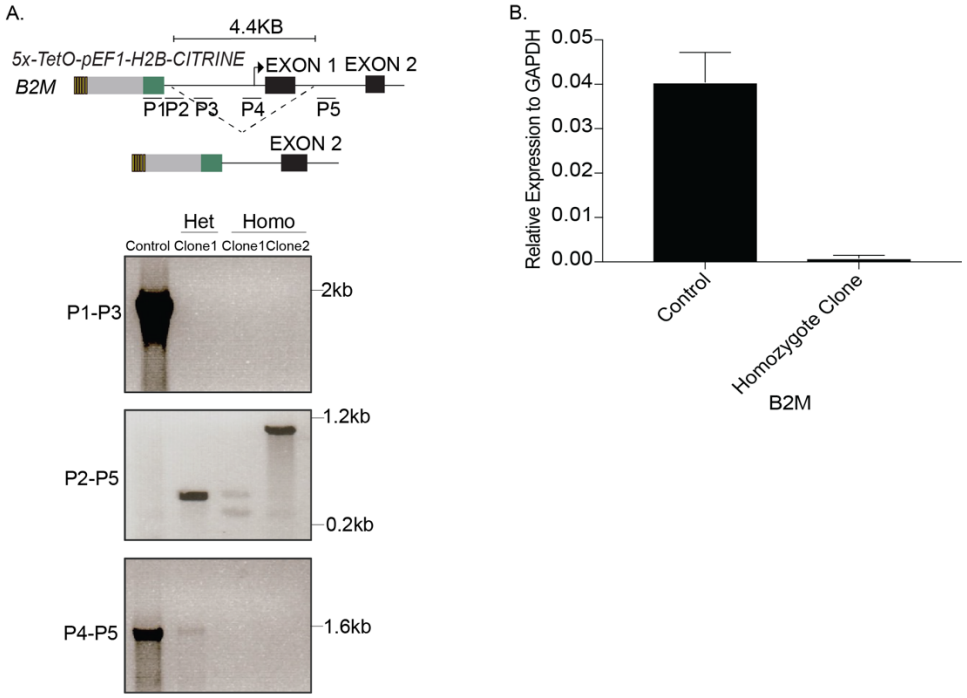
Figure 22. Attenuation of transcriptional activity partially reverses the loss of silencing of the *5xtetO-H2B-CITRINE* at the *TFRC* locus.

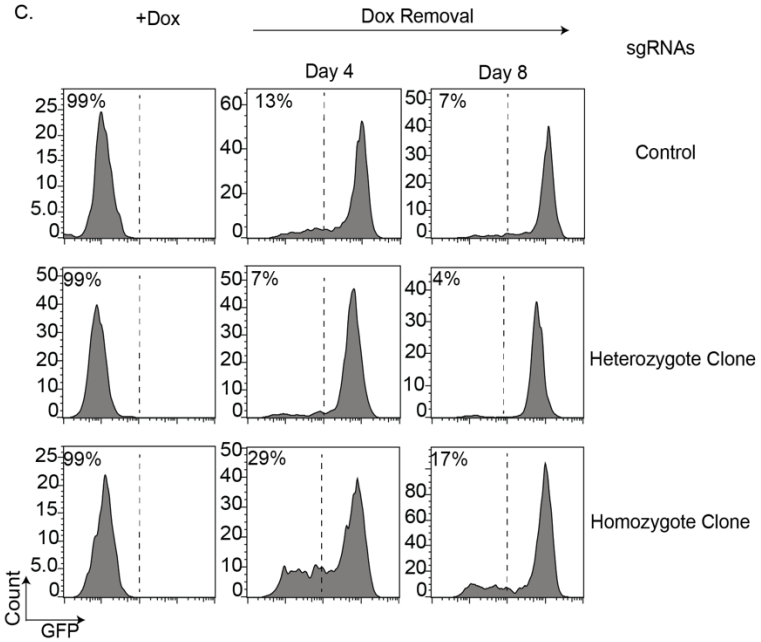
(A) PCR genotyping to detect the excision of the 5.6kb sequence as shown. Primer pairs are as indicated. Expected band for P1-P3 pair is ~1.6kb for WT and absence for Mutant. For P2-P4, expected WT band is too large to be detected (~6.5kb), but the presence of a small band ~1-1.5kb (depending on the sgRNAs used for deletion) indicates excision. For P2-P3, expected WT band is ~1.2kb, and its presence indicates a heterozygote clone, while absence of band indicates homozygous clone.

(B) qRT-PCR to detect steady state TFRC mRNA levels in control and homozygote deletion clones. The data from heterozygote deletion clones is pending. Data is normalized to GAPDH and error bars are deviation from the mean of at least two replicates.

(C) Representative Flow Cytometry histograms of control, heterozygote and homozygote deletion cells showing the percentage of silenced cells resulting from tethering of the rTetR-CBX7 to the reporter inserted at the *TFRC* locus at Day8 in doxycycline-containing medium (+Dox) and after transfer to doxycycline free media at Day 4 and 8 (Dox Removal). Percentages (%) show fraction of silenced cells.

(D) Plot showing quantification of the Flow Cytometry histograms of control, heterozygote and homozygote deletion cells to show the percentage of silenced cells after release of rTetR-CBX7 from the reporter in doxycycline free media at Day 4 and 8 (Dox Removal). Percentages (%) show fraction of silenced cells. Error bars are deviation from the mean of at least two replicates.


A.
4.4KB
5x-TetO-pEF1-H2B-CITRINE
B2M
EXON 1 EXON 2
P1P2 P3 P4 P5
EXON 2
Het Homo
Control Clone1 Clone1Clone2
P1-P3
2kb
P2-P5
1.2kb
0.2kb
P4-P5
1.6kb
B.
Relative Expression to GAPDH
0.05
0.04
0.03
0.02
0.01
0.00
Control
Homozygote Clone
B2M


C.
+Dox
Dox Removal
sgRNAs
Day 4
Day 8
99%
13%
7%
Control
99%
7%
4%
Heterozygote Clone
99%
29%
17%
Homozygote Clone
Count
GFP

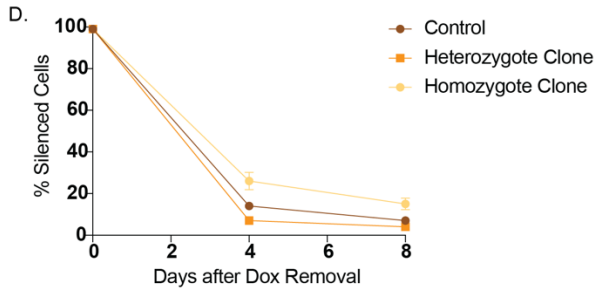

D.
% Silenced Cells
Days after Dox Removal
Control
Heterozygote Clone
Homozygote Clone

Figure 23 (Continued)

Figure 23. Attenuation of transcriptional activity partially reverses the loss of silencing of the *5xtetO-H2B-CITRINE* at the *B2M* locus.

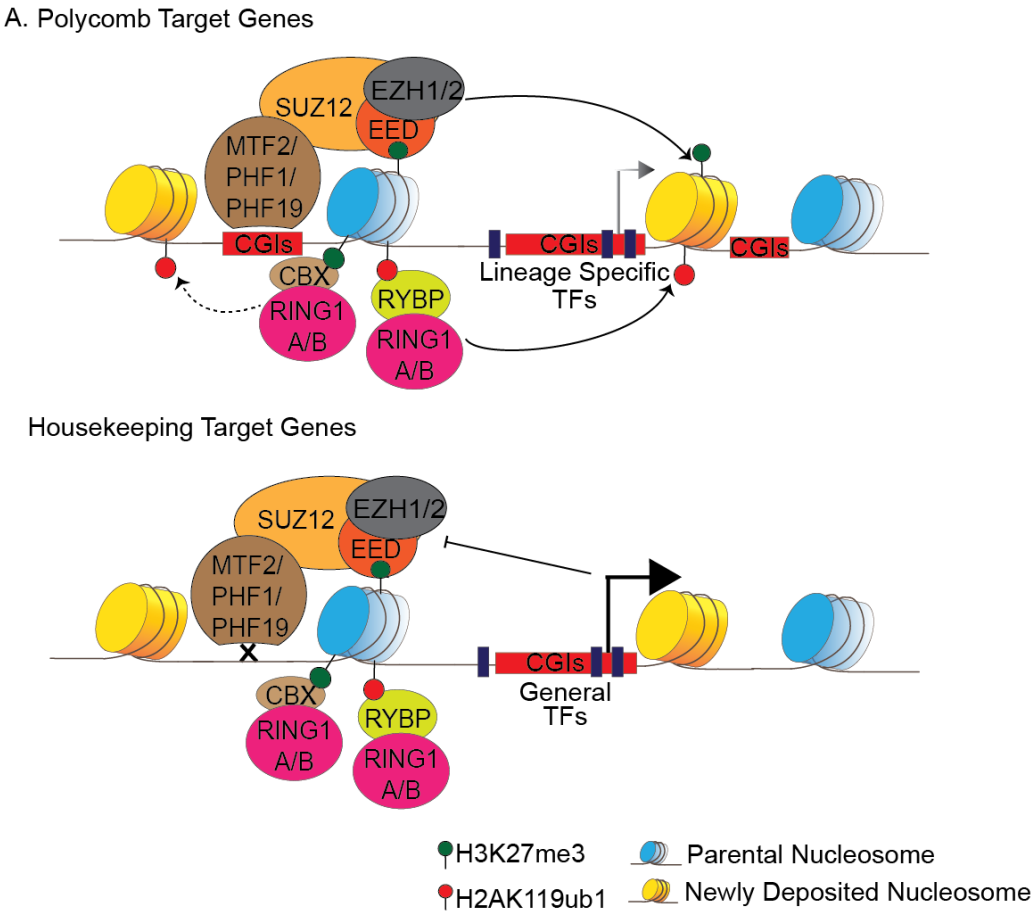
(A) PCR genotyping detecting the excision of the 4.4kb sequence as shown. Primer pairs are as indicated. Expected band for P1-P3 pair is ~1.8kb for WT and absence for Mutant. For P2-P5, WT band is too large to be detected (~4.6kb), but the presence of a small band ~0.3kb/1.1kb (depending on the sgRNAs used for deletion) indicates excision. For P4-P5, expected WT band is ~1.6kb and its presence indicates a heterozygote clone, while absence of band indicates homozygous clone.

(B) qRT-PCR showing steady state B2M mRNA levels in control and homozygote deletion clones. The data from heterozygote deletion clones is pending. Data is normalized to GAPDH and error bars are deviation from the mean of at least two replicates.

(C) Representative Flow Cytometry histograms of control, heterozygote and homozygote deletion cells showing the percentage of silenced cells resulting from tethering of the rTetR-CBX7 to the reporter inserted at the *B2M* locus at Day8 in doxycycline-containing medium (+Dox) and after transfer to doxycycline free media at Day 4 and 8 (Dox Removal). Percentages (%) show fraction of silenced cells.

(D) Plot showing quantification of the Flow Cytometry histograms of control, heterozygote and homozygote deletion cells showing the percentage of silenced cells resulting from release of rTetR-CBX7 from the reporter in doxycycline free media at Day 4 and 8 (Dox Removal). Percentages (%) show fraction of silenced cells. Error bars are deviation from the mean of at least two replicates.

3.3 Model for Locus-Dependent Epigenetic Inheritance of Polycomb-Mediated Gene Silencing



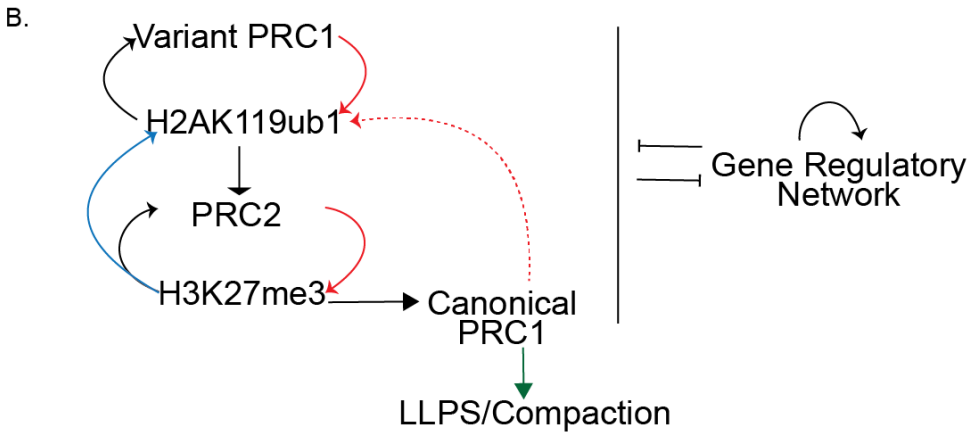


Figure 24 (Continued)

Figure 24. Model for epigenetic inheritance of Polycomb-mediated gene silencing.

A. At Polycomb target loci, where transcriptional activity is relatively low, inheritance of the silenced state requires the PRC2-EED's Read-Write positive feedback loop which recognizes H3K27me3 on the parental histone and allosterically activates EZH1/2 to catalyze H3K27 methylation on newly deposited histones. cPRC1 can recognize H3K27me3 through its CBX subunit and can potentially catalyze H2AK119ub1 through RING1A/B subunit (dotted arrow). vPRC1 recognizes H2AK119ub1 and allosterically activates RING1A/B for catalysis. The Polycomb machinery inhibits transcription as directed by the lineage specific transcription factors (TFs). Additionally, the maintenance of the silenced state requires the recognition of CGIs (CpG-rich DNA) by the DNA binding accessory proteins of PRC2- MTF2/PHF1/PHF19.

In contrast at the housekeeping target loci, where transcriptional activity is very strong relative to Polycomb target genes, an induced Polycomb domain and silencing cannot be maintained. The lack of multiple CGIs and/or chromatin features associated with transcription may obstruct the DNA-binding subunits of PRC2 and prevent the inheritance of Polycomb domains. Notably, H2AK119ub1 and RING1A/B, likely via vPRC1, can mediate short term epigenetic memory in the absence of H3K27me3.

B. Overall summary of reinforcement PRC1 and PRC2 feedback loops and gene regulatory networks (GRNs) that mutually antagonize each other. Red arrows denote catalysis (write) and black arrows denote recognition (read).

3.4 Discussion

In this chapter, I have demonstrated that both locus-independent (read-write positive feedback) and locus-specific cis-acting mechanisms contribute to the epigenetic inheritance of Polycomb-mediated gene silencing.

My findings reveal that the read-write positive feedback capability of EED-PRC2 plays a role in epigenetic inheritance; however, the partial rescue displayed by the aromatic cage EED mutant suggests that EED-PRC2 can also act independently of its read-write function. The EED-PRC2 WD40 mutant is still able to form a stable PRC2[15] and is potentially recruited to the reporter locus, but cannot catalyze H3K27me3 (as shown by western blot and ChIP-Seq data). Therefore, a likely explanation is that PRC2 functions in H2AK119ub1 epigenetic maintenance, which mediates partial silencing, independently of its ability to recognize and catalyze H3K27me3. Further studies are needed to understand the mechanism of this PRC2 Read-Write independent function in epigenetic inheritance. One hypothesis is that PRC2 physically bridges to PRC1 independently of H3K27me3, which increases the residency time of PRC1 to catalyze H2AK119ub1. To further investigate this, RYBP/YAF2 can be abolished in EED aromatic cage mutant cells. This will ensure that without H3K27me3, cPRC1 catalytic activity may be limited; therefore, the contribution of H2AK119ub1 catalyzed by vPRC1 can be appropriately studied in the inheritance of the silenced state. Furthermore, I observe that in the absence of H3K27me3, during rescue with the mutant EED, the silenced state is less stable. This suggests that to maintain stable silencing, both H3K27me3 and H2AK119ub1 are required. This is also supported by observations in *RING1A/B*$^{-/-}$ cells, where the rapid decay in silencing was correlated with the loss in H2AK119ub1 and significantly reduced levels of H3K27me3. Additionally, the partial rescue results indicate that H2AK119ub1 has “epigenetic memory” and

can be propagated in the absence of H3K27me3. The inheritance of H2AK119ub1 during DNA replication has been controversial in the field. It has been shown that parental H3-H4 tetramers can be inherited during DNA replication. By contrast, nucleosomal H2A–H2B undergoes rapid exchange with free H2A–H2B, which suggests that H2AK119ub1 cannot be stably inherited[116–120]. It remains possible that special mechanisms have evolved that facilitate the maintenance of H2-H2B dimers with H2AK119ub1 during DNA replication[121]. Further studies are needed to address how H2AK119ub1 can be retained during cell division. An alternative hypothesis as to the mechanism of PRC2's read-write independent function is that PRC2 can mediate chromatin compaction which inhibits transcriptional activity to maintain the silenced state; this is supported by in vitro compaction of chromatin templates by EZH1[63].

There are open questions in the field on the role of both cPRC1 and vPRC1 in maintaining gene repression in differentiated cells. Studies focused on the function of PRC1 in gene repression has mostly been conducted in mESCs, where it has been shown that cPRC1 is dispensable but H2AK119ub1 is necessary for ESC gene repression.[49,50] However, loss of cPRC1 can cause differentiation defects, suggesting it has a role in maintenance of gene silencing. My experiments demonstrate that when RING1A/B, which is a subunit of both the cPRC1 and vPRC1, is knocked out, there is a rapid loss in the maintenance of silencing at the *CITRINE* reporter. My findings cannot distinguish if cPRC1 or vPRC1 or both is necessary for epigenetic inheritance. To distinguish their roles, the requirement for vPRC1-specific proteins, RYBP/YAF2, need to be analyzed. In addition, to investigate the importance of vPRC1 Read-Write, rescue experiments with wild-type RYBP and YAF2, and their mutant versions that are defective in H2AK119ub1 binding need to be performed.

Studies in flies have demonstrated that PREs are necessary for the inheritance of Polycomb domains[70,71]. Unmethylated CGIs are thought to act as PRE-like sequences in mammalian cells[26,73,103]. However, the requirement for the ability of PRC2 accessory factors, MTF2/PHF1/PHF19 to bind CG-rich sequences, in epigenetic inheritance in mammalian cells had not been addressed. My findings reveal that the CGI-DNA binding ability of MTF2/PHF1 is necessary for the inheritance of Polycomb-mediated gene silencing. Additional experiments need to be conducted on the function of MTF2/PHF1 binding to CGIs in epigenetic inheritance. One hypothesis is that MTF2/PHF1 increase the residency time of PRC2, recruited to H3K27me3 nucleosomes to increase the chances that it can methylate newly deposited nucleosomes. In vitro biochemical evidence with PHF1 provide support for this notion[92].

There are other outstanding questions as to if and how CGIs differ from a Polycomb target locus to a housekeeping/non-target site. It has been reported that Polycomb target genes have a high CGI density relative to non-Polycomb targets[26,103]. Additionally, it has been proposed through computational simulations that the CGIs at Polycomb target genes have DNA that is unwound for optimal contact with PCL's winged helix structure, while non-Polycomb target sites with less CGIs have more tightly wound DNA[21]. Unlike PREs in flies, CGIs in mammalian cells do not seem to have motifs for binding of site-specific DNA-binding proteins. Examining the CGI density at Polycomb targets, *WT1*, *EN2* and *HOXD11* and housekeeping targets, *TFRC* and *B2M*, shows that the Polycomb targets have a higher density of CGIs compared to the housekeeping genes that have one CGI overlapping the promoter region (Figure 6, 8, 9 and 10). While MTF2 is recruited to the housekeeping target gene *TFRC* during establishment, it is likely that the lack of multiple CGIs or the packaging of that one CGI at the

promoter could block its interaction with MTF2, providing an explanation for the rapid loss in maintenance of silencing.

An alternative explanation for the rapid loss of silencing observed at housekeeping loci is the effect of proximal transcriptional activity (though this is not mutually exclusive with other explanations as in the above). At the housekeeping loci, a reduction in proximal transcriptional activity partially reverses the rapid decay in silencing. To speculate, it is possible that transcription factors binding to housekeeping genes strongly recruit downstream proteins that direct the feedback loop of the Trithorax proteins[122]. These include SWI/SNF complex and COMPASS proteins. SWI/SNF has been shown to regulate chromatin structure of genes and COMPASS proteins (including MLL proteins) catalyze active HPMs such as H3K4me3, H3K27ac, H3K36me etc[6,123–125]. Furthermore, histone demethylases/ deubiquitylases to remove H3K27me and H2AK119ub could also be recruited[126–128] to counteract Polycomb inheritance. Alternatively, nascent RNA at target loci may interact with PRC2, termed the "RNA-bridging model"[129] in ways that affect inheritance. The bridging model proposes that nascent RNA from lowly transcribed genes promotes PRC2 occupancy and function, while nascent RNA from highly transcribed genes repels PRC2. However, experimental evidence to support this model is limited. To speculate, it is possible that at housekeeping targets, the high levels of proximal nascent RNA repels PRC2 binding, while at the lowly transcribed Polycomb targets, the low levels of proximal nascent RNA further promotes PRC2 occupancy or at least does not interfere with PRC2 binding (See Chapter 4).

Interestingly, I observed that homozygote promoter region deleted clones at both housekeeping genes promoted increased inheritance compared to heterozygote clones. This suggests that at the heterozygote clones, the WT allele with its intact sequences influences the

inheritance of silencing at the reporter allele, with the promoter region deletion. One possible hypothesis is that the protein being expressed from the WT allele acts in trans to affect the reporter allele. An alternative hypothesis is that a transvection-like phenomenon contributes to the observed differences. Transvection is the process whereby homologous chromosomes pair up so that regulatory sequence on one homolog can affect transcription on the other. This phenomenon was first described in flies, but evidence in mammalian cells is limited[130] (See Chapter 4).

Additionally, my results show that induced silencing at Polycomb target genes in HEK293FT cells, even though it can be inherited for many cell divisions, is unstable and slowly decays. For instance, at the *WT1* locus, the cells lose maintenance of the silenced state with 74% of the cells silenced at Day 8 and only 27% of the cells silenced at Day 40. This is in contrast to the stable silencing of Polycomb target genes in many cell types and raises the possibility that the cell type-specific transcription factor network that maintains *WT1* expression in HEK293FT cells can slowly chip away at the induced Polycomb domain during the inheritance phase[96,131–135].

Altogether, my results suggest that antagonistic transcriptional activity in *cis* affects the epigenetic inheritance of Polycomb silencing. This provides support to the proposed model of chromatin bi-stability, whereby the mutual antagonism of Polycomb activity and opposing transcriptional activity defines the epigenetic state of the gene. Additional experiments to test this hypothesis could include deleting components of the TF network in HEK293FT cells that activates the *WT1* gene, which would be expected to increase stable inheritance of silencing at the locus.

To further study the contribution of CGIs and cis antagonistic transcriptional activity, experiments to insert additional CGIs at the sequence-deleted homozygote clones at

housekeeping loci could be conducted. Insertion of additional CGIs could potentially enhance the maintenance of silencing of the reporter, providing experimental evidence that the density of CGIs plays a role in epigenetic inheritance.

A potential caveat of my experiments is that the H3K27me3 Polycomb domains that are established at the Polycomb and housekeeping target loci are not similar in size. At the *WT1* locus, the domain size is approximately 14kb but at the housekeeping loci it is limited to 3kb. The antagonistic high transcriptional activity at the housekeeping loci could be a reason for the limited spreading. However, the H3K27me3 signal is comparable at both loci. The smaller size of this domain at the housekeeping genes could be a potential reason for the rapid loss in maintenance of silencing because a minimum number of H3K27me3/H2AK119ub1 nucleosomes may be required for silencing. Additional experiments need to be carried out to test Polycomb epigenetic inheritance at housekeeping genes where the domain size is similar to Polycomb target genes.

Chapter 4: Conclusions and Future Directions

In conclusion, in this dissertation I have demonstrated that the inheritance of Polycomb-mediated gene silencing requires contributions from locus-independent and locus-dependent factors. Using an inducible reporter system that was inserted at Polycomb target and housekeeping loci, I observed that the Polycomb domain and the silent state are maintained through multiple cell divisions at Polycomb target genes but rapidly lost at housekeeping genes, suggesting that maintenance of the silent state was locus-dependent. Maintenance of the silenced state at Polycomb target genes required locus-independent factors, which include the "Read-Write" mechanism of PRC2-EED, and PRC1-mediated H2AK119ub1 modification. Additionally, the positive feedback loop between PRC1 and PRC2 contribute to the inheritance of the domain. In terms of locus-dependent factors, maintenance at the Polycomb target was disrupted by point mutations in the DNA-binding domain of MTF2, suggesting the contribution of DNA sequence (CGIs) in epigenetic inheritance. Furthermore, I found that proximal transcriptional activity plays a role in counteracting Polycomb, as the loss of the silent state at housekeeping genes was partially reversed by DNA deletions that attenuated their transcription. These findings suggest that the heritability of Polycomb-mediated gene silencing requires DNA sequence-independent histone modification positive feedback, but is also dependent on the ability of PRC2 accessory factors to bind to CG rich DNA and is opposed by proximal transcription activity.

My findings highlight outstanding questions and concepts that need to be addressed in future studies.

Investigating the Contributions of CGIs versus Transcriptional Activity on Inheritance

To further investigate the role of CGIs and proximal transcriptional activity on epigenetic inheritance, one could take a synthetic biology approach. For instance, my reporter tool can be inserted at a gene desert region which is devoid of any CGIs and transcriptional activity as a blank slate to test the contribution of each factor. For example, different densities of CGIs and proximal transcriptional inputs can be inserted adjacent to the reporter to understand the contributions of each locus-dependent factor on the inheritance of Polycomb-mediated gene silencing. This strategy could provide us with insights into the balance that is required from each locus-dependent factor to either silence or activate the gene. Recent insights into the mechanisms that silence transgenes in mammalian cells based on their length or intron content should facilitate the generation of appropriate reporter genes that would be resistant to spontaneous silencing[100].

Investigating the Role of the Nuclear Envelope in Epigenetic Inheritance

It has previously been shown that interactions between the genome and nuclear envelope regulates gene expression, especially during differentiation[136,137]. The nuclear lamina, which is a filamentous protein network underlying the nuclear envelope, can interact with genomic domains to sequester silenced genes, including Polycomb silenced genes such as the inactive X chromosome[138]. However, further studies have to be done to understand the mechanism in mammalian cells. Furthermore, the nuclear envelope includes the nuclear pore complex (NPC), which comprises of multiple copies of approximately 30 different proteins termed nucleoporins (Nups). It has been shown that Nup93 and its interactors mediates the silencing of the *HOXA* gene cluster[139]. Another study demonstrated that Nup153 binds around the transcriptional start

site (TSS) of developmental genes in mESCs and mediates the recruitment of PRC1 to a subset of its target loci[140]. Therefore, mechanistic experiments can be carried out to understand the role of NPCs and the nuclear lamina in the epigenetic inheritance of Polycomb-mediated gene silencing using reporter genes.

Investigating the Role of Nascent RNA and the Rixosome Complex in Epigenetic Inheritance

It has previously been demonstrated that nascent RNA can prevent PRC2 binding and its HMTase activity, while others have shown contradicting evidence that nascent RNA promotes PRC2 occupancy[82,129]. This gave rise to the "RNA-bridging model" which proposes that at lowly transcribed genes, nascent RNA promotes PRC2 occupancy while at highly transcribed genes, it repels PRC2 binding. Previous studies have demonstrated that the general transcription machinery and Pol II is present at promoters of Polycomb repressed genes[6,141,142]. Recently, Zhou et al. showed that the rixosome complex (involved in RNA processing and ribosome biogenesis) is recruited to repressed genes through PRC1, and proposed that at target sites where Polycomb-mediated repression is weak and Pol II enters early elongation, the complex interacts with nascent RNA to process it for degradation and thus terminate transcription resulting in Polycomb-mediated gene silencing[53]. Therefore, one can hypothesize that the Rixosome complex contributes to epigenetic inheritance of Polycomb-mediated gene silencing. To speculate, it is possible that at the housekeeping genes used in my experiments, the high levels of proximal nascent RNA cannot be degraded efficiently by the rixosome, allowing the RNA to antagonize Polycomb, leading to the rapid loss of silencing at those loci. Further studies are needed to understand the role of the rixosome and nascent RNA in epigenetic inheritance of Polycomb-mediated gene silencing.

Role of Topologically Associated Domains (TADs) in Epigenetic Inheritance

Throughout my study I have not explored the role of TADs and their possible effect on epigenetic inheritance. TADs are units of three-dimensional (3D) nuclear organization that have been identified though chromosome conformation techniques (HiC). TAD boundaries may contribute to the regulation of gene expression by preventing inappropriate interactions of cis-regulatory sequences with target genes. Previously, it has been shown that the repressed *HOX* genes in ESCs are contained in one TAD, but during differentiation, active *HOX* genes and repressed *HOX* genes are segregated into different TADs for their proper gene regulation, suggesting that TADs play a role in the appropriate expression of *HOX* genes in lineage specific cells[143,144]. Additionally, mutations in TAD boundaries result in relocalization of PcG proteins outside the boundaries leading to an increase in expression of the genes within the TAD domain[145]. Therefore, investigating whether TADs impact the inheritance observed at the Polycomb versus housekeeping targets may provide additional insights into the mechanism of epigenetic inheritance.

Transvection

My results indicate that relative to heterozygote deletion, the homozygote sequence deletion at the two housekeeping loci lead to a more robust epigenetic inheritance of induced silencing. This suggests that the WT allele in heterozygote clones influences the inheritance of the silenced state at the reporter allele. One hypothesis as to the mechanism of this phenomenon is the process of transvection. Transvection is defined as the ability of one allele to interact with its paired allele on the homologous chromosome; one mechanism is through physical pairing of

homologous chromosomes. This phenomenon has been demonstrated in *Drosophila,* certain plants and fungi[146,147]. Experimental evidence in mammalian cells is limited, and has mostly been shown to be present during Cre-LoxP recombination, X-chromosome inactivation, during imprinting, whereby imprinted control regions (ICRs) in one allele physically interact to influence the other, VD(J) recombination, and during ES cell differentiation[148–155]. Furthermore, transvection has been seen in cancer and diseased cell lines[156,157]. Since HEK293FT is derived from kidney cancer cells and demonstrates aneuploidy, it is possible that transvection plays a role in the differences I observed in heritability of homozygote versus heterozygote housekeeping gene deletions. It would be interesting to experimentally observe if transvection is occurring at the heterozygote clones and to more fully explore this phenomenon and its possible impact on gene expression in mammalian cells.

Chapter 5: Methods

Cell Culture:

HEK293FT (ThermoFisher R70007) cells were maintained in DMEM medium (Invitrogen) plus 10% FBS (Invitrogen), 1 mM Glutamine and 100 µg/ml penicillin-streptomycin following standard culture conditions. To induce the binding rTetR-CBX7 to the *5XtetO* Site, 1 µg/ml Doxycycline (Sigma, D9891) was added to the culture medium.

Plasmid Construction:

Donor Plasmids for insertion of *5xtetO-H2B-CITRINE* reporter into the genome were constructed by subcloning *5XTetO-H2B-CITRINE-PolyA* from PhiC31-Neo-ins-5xTetO-pEF-H2B-Citrine-ins (Addgene# 78099)[94] with right and left homology arms (500bps each) in CloneSmart HCKan Blunt (Gift from Jichuan Zhang [Genome Editing and Neurodegeneration Core in the Department of Cell Biology at Harvard Medical School], Lucigen# 40704-2).

Plasmid with mCherry-2A-rTetR-CBX7 was created by subcloning mCherry-2A-rTetR (Addgene # 78101)[94] into lentiviral expression vector backbone pLVU-tTR-KRAB (Addgene# 11645)[158]. CBX7 was amplified from pCMV-SPORT6-CBX7 (DFCI Plasmid# HsCD00339744) and subsequently cloned.

Rescue Plasmids were constructed by cloning HA/3X FLAG Tag, EED or MTF2 WT and Mutant cDNAs using Gibson into pdCas9-DNMT3A-2A-PuroR (Addgene# 71667)[159]. The point mutations were built using IDT gBlocks.

CRISPR genome editing:

sgRNAs for reporter cell line construction, gene knockouts and sequence deletions were designed using the CRISPR design tool in https://benchling.com and/or

https://chopchop.cbu.uib.no/ (Table 1, 2 and 4). sgRNAs were either in vitro transcribed using GeneArt™ Precision gRNA Synthesis Kit (ThermoFisher A29377) and electroporated with Neon Transfection System (ThermoFisher MPK1025), along with donor plasmid and Cas9 protein (Gift from Jichuan Zhang, Genome Editing and Neurodegeneration Core in the Department of Cell Biology at Harvard Medical School) or cloned into pSpCas9(BB)-2A-Puro (PX459) V2.0 (Addgene plasmid # 62988) and transfected into HEK293FT using Lipofectamine 2000 (ThermoFisher). sgRNAs for rTetR-CBX7 deletion were cloned into LentiCRISPRv2 (Addgene 52961). CITRINE positive cells were sorted into single cell colonies in 96 well plates, genotyped by PCR and confirmed by Sanger Sequencing (Quintara Bio) or MiSeq (Illumina). Southern Blot using a CITRINE probe was carried out to verify single integration for the reporter cell lines.

Integration of Rescue Constructs:

Rescue Plasmids were transfected into relevant cell lines using Lipofectamine 2000 (Thermofisher). Cells with insertions were selected using Puromycin (ThermoFisher) at 0.6 ug/ml for 2 weeks. Selected cells were subsequently maintained on 0.2 ug/ml puromycin.

Western Blot:

Whole cell extract was obtained by lysis in RIPA buffer (final: 150 mM NaCl, 1% triton, 0.5% sodium deoxy-cholate, 0.1% SDS, 50 mM Tris pH 8.0). The protein concentration was determined by the Bradford assay (Biorad). 10-20 µg/lane total protein was run on 4–15% Mini-PROTEAN® TGX™ Precast Protein Gels (BioRad) with SDS Running Buffer and transferred on Polyvinylidene difluoride (PVDF) membrane. The membrane was blocked (5% non-fat dry

milk in 1× PBS, 0.1% Tween 20) for x hours and then incubated in 5% non-fat dry milk in 1× PBS, 0.1% Tween 20 with the primary antibodies as listed in Table 9 for 2 hours at room temperature or O/N at 4°C. Finally, the membrane was incubated with corresponding secondary Licor IRDye antibody (5% non-fat dry milk in 1× PBS, 0.1% Tween 20) and imaged by Odyssey Clx (Licor) or HRP-conjugated secondary antibodies and imaged on Amersham Imager (GE).

RT-qPCR:

Total RNA was extracted using the RNeasy Mini kit (74104, Qiagen) and reverse transcribed into cDNA using random hexamers (Invitrogen) and reverse transcription kit (18090010, ThermoFisher). cDNA was analyzed using PCR on a QuantStudio 7 Flex Real Time PCR System (Applied Biosystem). PCR parameters were 95°C for 2 min and 40 cycles of 95°C for 15s, 60°C for 30s, and 72°C for 30 s, followed by 72°C for 1 min. All the qPCR data presented were at least two biological replicates and plotted with Prism GraphPad Software. Primer sequences are presented in Table 8.

LentiViral Production and Infection:

Plasmids were purified using a MaxiPrep DNA isolation Kit (Qiagen). For virus packaging, we used psPAX2 (Addgene# 12260) and pMD2.G (Addgene# 12259) which were transfected into HEK293FT cells using Lipofectamine 2000 (Invitrogen). Medium containing the viral particles was collected 72 hr after transfection and viral particles were concentrated using the PEG-it Virus precipitation solution (SBI LV810A-1). Cells were transduced with the virus for 48 h in the presence of 4 μg/ml polybrene (Sigma H9268).

Immunofluorescence:

Cells were plated on chamber slides (ThermoFisher 154526PK). Cells were first washed with PBS, fixed with 4% paraformaldehyde in PBS for 5mins and permeabilized with PBS/0.25% Triton X-100 at room temperature for 5 mins. Cells were mounted with VECTASHIELD HardSet Antifade Mounting Medium with DAPI (Vector Labs) and imaged with a widefield microscope (Nikon Ti2) equipped with a 40× objective lens. Images were post-processed with Image J[160] and photoshop (Adobe) software.

Fluorescence-Activated Cell Sorting and Analysis:

Cells were made into single cell suspension using 0.5% Trypsin (Invitrogen) and suspended in HEK293FT culture medium. Samples for analysis were collected with LSR Fortessa (BD Biosciences) or FACs Calibur (BD Biosciences). Samples were sorted with M AriaII (BD Biosciences). Data was analyzed with FlowJo™ Version 10.5.3 (Ashland, OR: Becton, Dickinson and Company; 2021).

ChIP-qPCR and ChIP-Seq:

ChIP was performed as previously described with minor modifications[161]. Cells for ChIP were cultured in 15 cm plates (~10 million cells). Cell pellets were first washed with cold PBS, crosslinked at room temperature with 1% formaldehyde (ThermoFisher Scientific) for 8 min. Crosslinking reactions were quenched by addition of 125 mM glycine for 10 min. Cell were then resuspended in Swelling Buffer (25mM Hepes pH 7.8, 1.5mM $MgCl_2$, 10mM KCL, 0.1% NP-40, 1mM DTT) followed by Dounce homogenization. Nuclei were pelleted by centrifugation and then resuspended in sonication buffer (0.1% SDS, 1mM EDTA and 10mM Tri-HCL pH 8.0) . The nuclei were sonicated to shear chromatin into ~200-500 bp fragments using a Covaris E220. Sonicated samples were diluted with ChIP dilution buffer (0.1% SDS, 1mM EDTA and

10mM Tri-HCL pH 8.0, 1% Triton X-100, 150mM NaCl). Diluted samples were centrifuged at 13,000 rpm for 10 min. The supernatant was used for immunoprecipitation using antibodies and 25 µl protein A/G beads for 12-16 h at 4°C. For H3K27me3 ChIP-Seq, Drosophila S2 chromatin (Active Motif# 53083) and Histone H2Av antibody (Active Motif# 61686) were added as spike-in controls. ChIP-Seq samples for Flag antibody do not have spike-in controls. The beads were washed twice with high salt wash buffer A (50 mM Hepes pH 7.9, 500 mM NaCl, 1 mM EDTA, 1% Triton, 0.1% Sodium deoxycholate, and 0.1% SDS), twice with wash buffer B (20mM Tris-HCL pH 8.0, 250 mM LiCl, 1 mM EDTA, 0.5% Sodium deoxycholate, 0.5% NP-40) and twice with 1X TE (x mM Tris pH x and x mM EDTA). The bound chromatin fragments were eluted with elution buffer (50 Mm Tris pH 8.0, 1 mM EDTA, 50mM $NaHCO_3$,1% SDS) twice for 10 min each at 65°C. Eluted DNA-proteins complexes were incubated overnight at 65°C to reverse crosslinks. RNAase A followed by Proteinase K was then added to digest RNA and protein. DNA was further purified using Phenol Chloroform/PCR Purification Kit (QIAGEN) and analyzed by PCR on a QuantStudio 7 Flex Real Time PCR System (Applied Biosystem). PCR parameters were 95°Cfor 2 min and 40 cycles of 95°C for 15s, 60°C for 30s, and 72°C for 15 s, followed by 72°C for 1 min. All the ChIP-qPCR data presented include at least two biological replicates. Primer sequences are in Table 7. Results were plotted with Prism GraphPad Software. Error bars represent deviation from the mean (two biological replicates)

For ChIP-seq, sequencing library was constructed using TruSeq DNA sample Prep Kits (Illumina) and adapter dimers were removed by 2% Agarose and Tris-acetate-EDTA gel electrophoresis. Size selected and purified DNA libraries were sequenced on an Illumina Hiseq 2500 machine (Bauer core facility at Harvard University to obtain 75 bp single-end reads). ChIP-

seq reads were quality controlled with fastqc (v0.11.5) and mapped to the human genome reference (Custom 5xtetO-H2B-CITRINE Reporter Inserted at Chr11-hg19) and Drosophila (dm3) using bowtie2 (v2.2.9) with default parameters. Scale factor was calculated as previously described to normalize H3K27me3 signal[162]. Bam files were generated with samtools 1.3.1, which was followed by making bigwig files with deeptools (v/3.0.2). Reads were normalized with scale factor for H3K27me3 or RPGC with deeptool (v/3.0.2) bamCoverage function.

Table 1: List of *5xtetO-H2B-CITRINE* reporter cell lines created and the sgRNAs used for this study

Reporter Cell Lines	sgRNAs for Reporter Integration
Cell Line 1: hUBC-mCherry-rTetR-CBX7; 5XTetO-EF1-H2B-CITRINE-PolyA at WT1 (chr11)	AATTATGCACCTTCGAGGCC
Cell Line 2: hUBC-mCherry-rTetR-CBX7; 5XTetO-EF1-H2B-CITRINE-PolyA at EN2 (Chr7)	AAGCTTCCTGAAACACCTTG
Cell Line 3: : hUBC-mCherry-rTetR-CBX7; 5XTetO-EF1-H2B-CITRINE-PolyA at HOXD11 (Chr2)	TAACGAACAGTCAACACTCG
Cell Line 4: hUBC-mCherry-rTetR-CBX7; 5XTetO-EF1-H2B-CITRINE-PolyA at HOXB4 (Chr17)	AAGGTCCCTGGTTGACGCTA
Cell Line 5: hUBC-mCherry-rTetR-CBX7; 5XTetO-EF1-H2B-CITRINE-PolyA at TFRC (Chr3)	AACTGACCTTCAGGCCCGTA
Cell Line 6: hUBC-mCherry-rTetR-CBX7; 5XTetO-EF1-H2B-CITRINE-PolyA at B2M (Chr15)	CGTGAAGCCAGCATAGTACT
Cell Line 7: hUBC-mCherry-rTetR-CBX7; 5XTetO-EF1-H2B-CITRINE-PolyA at Gene Desert (Chr9)	GCGCATTGTGATCCAACCAT
Cell Line 8: hUBC-mCherry-rTetR-CBX7; 5XTetO-EF1-H2B-CITRINE-PolyA at Gene Desert (Chr7)	CCTATCTGCCATCTTAATCC

Table 2: List of knockout Cell Lines created and the sgRNAs used

Cell Line 1	sgRNAs for Knockouts
Cell Line 1.1. *hUBC-mCherry-rTetR-CBX7Δ; 5XTetO-EF1-H2B-CITRINE-PolyA at WT1* (chr11)	rTetR: GATGTGAGAGGAGAGCACAG; GCCATGACTCGCCTTCCAGG
Cell Line 1.2. *hUBC-mCherry-rTetR-CBX7; 5XTetO-EF1-H2B-CITRINE-PolyA at WT1* (chr11); *ΔEED*	EED: TTGCCACCAGAGTGTCCGTC
Cell Line 1.3. *hUBC-mCherry-rTetR-CBX7; 5XTetO-EF1-H2B-CITRINE-PolyA at WT1* (chr11); *ΔSUZ12*	SUZ12: CGAAGAGTGAACTGCAACGT
Cell Line 1.4. *hUBC-mCherry-rTetR-CBX7; 5XTetO-EF1-H2B-CITRINE-PolyA at WT1* (chr11); *ΔRING1A and RING1B*	RING1A: GTTCTGAATGCAGTGACCGA RING1B: AATTCACTGTGTAGACTTCG
Cell Line 1.5. *hUBC-mCherry-rTetR-CBX7; 5XTetO-EF1-H2B-CITRINE-PolyA at WT1* (chr11); *ΔMTF2*	MTF2: AGAAGAAGAAGCATTTGTTT
Cell Line 1.6. *hUBC-mCherry-rTetR-CBX7; 5XTetO-EF1-H2B-CITRINE-PolyA at WT1* (chr11); *ΔJARID2*	JARID2: ACAGGTGCTATCCCTCGGGG
Cell Line 1.7. *hUBC-mCherry-rTetR-CBX7; 5XTetO-EF1-H2B-CITRINE-PolyA at WT1* (chr11); *ΔPHF1*	PHF1: TCTTCTTACAGCAAACACTG; GTTTGTTTGGTCTCCATGCT
Cell Line 1.8. *hUBC-mCherry-rTetR-CBX7; 5XTetO-EF1-H2B-CITRINE-PolyA at WT1* (chr11); *ΔPHF19*	PHF19: CCTGGCCCTCTATAATCTGG; GCGTCCCACCCAACCCGCCA

Table 3: List of Genotyping Primers for *ΔPHF1* and *ΔPHF19*

Cell Line 1.4. *hUBC-mCherry-rTetR-CBX7; 5XTetO-EF1-H2B-CITRINE-PolyA at WT1* (chr11); *ΔPHF1*	Genotyping Primer
Primer F	CTGGCTCTTAAAATGCCTCTGT
Primer R	TGGACCAGTGACCTGGTGA
Cell Line 1.4. *hUBC-mCherry-rTetR-CBX7; 5XTetO-EF1-H2B-CITRINE-PolyA at WT1* (chr11); *ΔPHF19*	
Primer F	AGGCTGCCACCTCACCTGGTCC
Primer R	CTGACCCAGGCTTGCTCTTTC

Table 4: List of sgRNAs used for sequence deletion at *TFRC* and *B2M* reporter loci

Cell Line #	sgRNAs for Sequence Deletions
Cell Line 5.1: *hUBC-mCherry-rTetR-CBX7; 5XTetO-EF1-H2B-CITRINE-PolyA at TFRC* (Chr3); *Δ5.6KB Sequence*	sgRNA Pair 1: AGGCCAGTGCGCCCATCGCG; AAAGAGTTCAGCCTTCAGTA sgRNA Pair 2: AAAGAGTTCAGCCTTCAGTA; GCAGCCTCAGAAATACAAAA sgRNA Pair 3: GTAAAAAGCAAAGGCTTgcg; AGGCCAGTGCGCCCATCGCG
Cell Line 6.1: *hUBC-mCherry-rTetR-CBX7; 5XTetO-EF1-H2B-CITRINE-PolyA at B2M* (Chr15); *Δ4.4KB Sequence*	sgRNA Pair 1: ACTCACGCTGGATAGCCTCC; tgtggggttagaataccaag sgRNA Pair 2: ACTCACGCTGGATAGCCTCC; ccatgggctgaagtgctctg

Table 5: List of genotyping primers used for sequence deletions at *TFRC* reporter loci

Cell Line 5.1: *hUBC-mCherry-rTetR-CBX7; 5XTetO-EF1-H2B-CITRINE-PolyA at TFRC* (Chr3); *Δ5.6KB Sequence*	Genotyping Primers Sequence
Primer 1	ACGAGCTGTACAAGTAGGTCG
Primer 2	TGGGTTGCTTTCTTGCATTGT
Primer 3	CGTCATTCGTCGCTCTGTGA
Primer 4	AGGCAAGTCTCAAACTCCTCAAG

Table 6: List of genotyping primers used for sequence deletions at *B2M* reporter loci

Cell Line 6.1: *hUBC-mCherry-rTetR-CBX7; 5XTetO-EF1-H2B-CITRINE-PolyA at B2M* (Chr15); *Δ4.4KB Sequence*	Genotyping Primers Sequence
Primer 1	ACGAGCTGTACAAGTAGGTCG
Primer 2	GGCTCTGCAGTAAGCTTGTG
Primer 3	GCAAGATGGCTGAATAGACGC
Primer 4	GTTTAGGCATGGCCTCCACAA
Primer 5	GATCCAGCCCTGGACTAGC

Table 7: ChIP-qPCR primers

Primer Name	Primer Sequence
TetO F	ACGTATGTCGAGGTAGGCGT
TetO R	CTAGGCACCGGTTCAATTGC
CITRINE F	CGACTTCTTCAAGTCCGCCA
CITRINE R	CTTGTAGTTGCCGTCGTCCT
WT1 ChIP-qCPR +1.5F	TGCATAAACGTTGTCGCCATT
WT1 ChIP-qCPR +1.5R	AAGTGCGCCCTTCGAGTAAG
WT1 ChIP-qCPR +3F	CTAAGTGCTGCTGACTCCAAT
WT1 ChIP-qCPR +3R	TTTGTGGGTTCCAGAGGTCG
WT1 ChIP-qCPR +4F	CCAGGCCAGGATGTTTCCTAA
WT1 ChIP-qCPR +4R	GTGTCCTAGAGCGGAGAGTC
WT1 ChIP-qCPR +5F	GGGACCGGGATGTTTTTGGA
WT1 ChIP-qCPR +5R	TAAGGTAGGAGCGGCCTGAA
WT1 ChIP-qCPR -1.5F	GGAATTCCAGATGGTGCGCT
WT1 ChIP-qCPR -1.5R	GGCCAGAGCAGATACGTAGG
WT1 ChIP-qCPR -3F	TATAAACAGCTGCCCTGCCG
WT1 ChIP-qCPR -3R	GTCCAGATGCAGGAAGGGTT
WT1 ChIP-qCPR -4F	TGTGGTAACTCCAGGAAGAGGA
WT1 ChIP-qCPR -4R	AGCGTATGTCAAGGACATTGGT
WT1 ChIP-qCPR -5F	CAGCGTTTGGATTCGGGTTC
WT1 ChIP-qCPR -5R	CGCCCGACCCCGTAATTTT
TFRC ChIP-qPCR +0.5F	TGGGTTGCTTTCTTGCATTGT

Table 7 (Continued)	
TFRC ChIP-qPCR +0.5R	CAATCACACCCTCTCCCTCC
TFRC ChIP-qPCR +1F	CGTACGTGCCTCAGGAAGTG
TFRC ChIP-qPCR +1R	GTTCTAGAAGCCCGCACTCA
TFRC ChIP-qPCR +2F	GGACAAAGCTGTCCCCGATT
TFRC ChIP-qPCR +2R	AGAATCCACACACAAGGCGA
TFRC ChIP-qPCR +3F	CATGGTTCAAAACGTGGGGC
TFRC ChIP-qPCR +3R	TGGTGTTCTCAATGGTGACTGAA
TFRC ChIP-qPCR -0.5F	TAAGGTAGGCCCTCTGTGGAT
TFRC ChIP-qPCR -0.5R	CAGGAGCATTGCTGCACCTTTA
TFRC ChIP-qPCR -1F	AAACCTATGTCCTCCATGAGGCT
TFRC ChIP-qPCR -1R	TGCCTTTCCTTGACTGAAGTATC
TFRC ChIP-qPCR -2F	CCAGGCTCAGGAAAGTTGAGA
TFRC ChIP-qPCR -2R	AAGATACCATGAGCTGTGGGG
TFRC ChIP-qPCR -3F	CCTGCACTGTTGTTTCCAGC
TFRC ChIP-qPCR -3R	GGAGCCTGTGGTGTGTGTTA
B2M ChIP-qPCR +0.5F	CTATGGTAACCACCGCCTGG
B2M ChIP-qPCR +0.5R	AGGGACATAAGCTTGGCTGG
B2M ChIP-qPCR +1F	GAGGCCACTTGGTATTCTAACC
B2M ChIP-qPCR +1R	GTGTGACCCAGCACATTACA
B2M ChIP-qPCR +2F	CACCTCCCCTAGCTATGTCCTTT

Table 7 (Continued)	
B2M ChIP-qPCR +2R	CATGAGGAAACTTTAGGGTTGATGG
B2M ChIP-qPCR +3F	GGACTCCACCACCACGAAAT
B2M ChIP-qPCR +3R	AAGACAAAGGGCTCGGCAAT
B2M ChIP-qPCR +4F	GAGATGTCTCGCTCCGTGG
B2M ChIP-qPCR +4R	AGACTCACGCTGGATAGCCT
B2M ChIP-qPCR -0.5F	CAGTCTGGGCTGTTTGTATCT
B2M ChIP-qPCR -0.5R	GCTTGGTGTGCCCTCTAAT
B2M ChIP-qPCR -1F	TCAAGCTCACTAATTCTTTCTTCCA
B2M ChIP-qPCR -1R	ATACTGACATACATAAGGGTGCAT
B2M ChIP-qPCR -2F	TCCACTGGAAAGTCTGCTGC
B2M ChIP-qPCR -2R	ACAAACTCCCAAGGTCAAGAA
B2M ChIP-qPCR -3F	GTGCCACTATGTCTGGCTAAT
B2M ChIP-qPCR -3R	GGTGAAAGTGCTGTCTCTACAA
B2M ChIP-qPCR -4F	AGTCCCAGCTACTCAGGAGG
B2M ChIP-qPCR -4R	TCTGCACTATAGATCAAATGGCTC
GAPDH ChIP-qPCR F	AACAGCCTCAAGATCATCAGC
GAPDH ChIP-qPCR R	GGATGATGTTCTGGAGAGCC
PCDH10 ChIP-qPCR F	GGATGGCAACCGATTCGCTGA
PCDH10 ChIP-qPCR R	ACCTCCTCCG TCCACCGCGGT
MYT1 ChIP-qPCR F	ACAAAGGCAGATACCCAACG

Table 7 (Continued)	
MYT1 ChIP-qPCR R	GCAGTTTCAAAAAGCCATCC
SOX6 ChIP-qPCR F	GGGATAGCACGTGAGGATGG
SOX6 ChIP-qPCR R	GGACACCAGACAAGCCTACC
PTF1A ChIP-qPCR F	ATGGACGCGGTGTTGCTGGA
PTF1A ChIP-qPCR R	CGTGAAGACTGGTCGGTGAA
HOXA3 ChIP-qPCR F	GTGCCAATGTGCGCCCTCAC
HOXA3 ChIP-qPCR R	GAGCTGTCGTAGTAGGTCGC
HOXA10 ChIP-qPCR F	CTCTTTCGCGCAGAACATCA
HOXA10 ChIP-qPCR R	TGGCCGAGACTTTGGGGCAT

Table 8: qRT-PCR primers

Primer Name	Primer Sequence
GAPDH qRT-PCR F	CAATGACCCCTTCATTGACC
GAPDH qRT-PCR R	TTGATTTTGGAGGGATCTCG
TFRC qRT-PCR F	ACCGGCACCATCAAGCT
TFRC qRT-PCR R	TGATCACGCCAGACTTTGC
B2M qRT-PCR F	ACTGAATTCACCCCCACTGA
B2M qRT-PCR R	CCTCCATGATGCTGCTTACA

Table 9: List of Antibodies used

Antibodies	Source	Cat#	Application
Anti-RING1B	Cell Signaling Technology	5694S	WB, 1:200
Anti-BETA ACTIN	Abcam	mAbcam8224	WB, 1:1000
Anti-FLAG	Sigma	F3165	WB, 1:5000 ChIP, 4ug
Anti-SUZ12	Cell Signaling Technology	3737	IB, 1:1000
Anti-H3K27me3	Cell Signaling Technology	9733	WB, 1:1000 ChIP, 3ug
Anti-H2AK119ub1	Cell Signaling Technology	8240T	WB, 1:1000 ChIP, 2ug
Anti-MTF2	ProteinTech	16208-1-AP	WB, 1:100 ChIP, 4ug
Anti-JARID2	Novus Biologics	NB100-2214SS	WB, 1:500
Anti-HA	ThermoFisher	26183-HRP	WB, 1:2000
Anti-CBX7	Abcam	ab21873	WB, 1:1000
Anti-Histone H2Av antibody	Active Motif	61686	ChIP-Seq Spike in: 2ug

Table 10: List of Data sources from public databases

ChIP-Seq File	**GEO Accession #**
hESCs H3K27me3 ChIP-Seq	GSM1185386
hESCs RNA-seq	GSM672836
HEK293FT H3K27me3 ChIP-Seq	GSM4239945
HEK293FT H2AK119ub1 ChIP-Seq	GSM4502558
HEK293FT RNA-seq	GSM5343713
HEK293 Bisulfite Seq	GSM683769
CGIs	UCSC Genome browser[163]
Human Protein Atlas	proteinatlas.org[95]
Transcription Factor ChIP	UCSC Genome browser[164–166]

Table 11: List of Next-Generation Sequencing tools used

deeptool (v3.0.2)	https://github.com/deeptools/	Reference[167]
Samtool (v1.3.1)	https://github.com/samtools/samtools	Reference[168]

Bibliography

1. Moazed, D. Mechanisms for the inheritance of chromatin states. *Cell* (2011) doi:10.1016/j.cell.2011.07.013.

2. Bonasio, R., Tu, S. & Reinberg, D. Molecular signals of epigenetic states. *Science* (2010) doi:10.1126/science.1191078.

3. Lewis, P. H. New mutants report. *Dros. Inf. Serv.* **21**, 69 (1947).

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

5. Kassis, J. A., Kennison, J. A. & Tamkun, J. W. Polycomb and trithorax group genes in drosophila. *Genetics* (2017) doi:10.1534/genetics.115.185116.

6. Schuettengruber, B., Bourbon, H. M., Di Croce, L. & Cavalli, G. Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell* (2017) doi:10.1016/j.cell.2017.08.002.

7. Margueron, R. & Reinberg, D. The Polycomb complex PRC2 and its mark in life. *Nature* (2011) doi:10.1038/nature09784.

8. Yu, J. R., Lee, C. H., Oksuz, O., Stafford, J. M. & Reinberg, D. PRC2 is high maintenance. *Genes and Development* (2019) doi:10.1101/gad.325050.119.

9. Laugesen, A., Højfeldt, J. W. & Helin, K. Role of the polycomb repressive complex 2 (PRC2) in transcriptional regulation and cancer. *Cold Spring Harb. Perspect. Med.* (2016) doi:10.1101/cshperspect.a026575.

10. Cao, R. *et al.* Role of histone H3 lysine 27 methylation in polycomb-group silencing. *Science (80-.).* (2002) doi:10.1126/science.1076997.

11. Ferrari, K. J. *et al.* Polycomb-Dependent H3K27me1 and H3K27me2 Regulate Active Transcription and Enhancer Fidelity. *Mol. Cell* (2014) doi:10.1016/j.molcel.2013.10.030.

12. Margueron, R. *et al.* Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* (2009) doi:10.1038/nature08398.

13. Jiao, L. & Liu, X. Structural basis of histone H3K27 trimethylation by an active polycomb repressive complex 2. *Science (80-.).* (2015) doi:10.1126/science.aac4383.

14. Lee, C. H. *et al.* Distinct Stimulatory Mechanisms Regulate the Catalytic Activity of Polycomb Repressive Complex 2. *Mol. Cell* (2018) doi:10.1016/j.molcel.2018.03.019.

15. Oksuz, O. *et al.* Capturing the Onset of PRC2-Mediated Repressive Domain Formation. *Mol. Cell* (2018) doi:10.1016/j.molcel.2018.05.023.

16. Lee, C. H. *et al.* Allosteric Activation Dictates PRC2 Activity Independent of Its

Recruitment to Chromatin. *Mol. Cell* (2018) doi:10.1016/j.molcel.2018.03.020.

17. Cao, R. & Zhang, Y. SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. *Mol. Cell* (2004) doi:10.1016/j.molcel.2004.06.020.

18. Chen, S., Jiao, L., Shubbar, M., Yang, X. & Liu, X. Unique Structural Platforms of Suz12 Dictate Distinct Classes of PRC2 for Chromatin Binding. *Mol. Cell* (2018) doi:10.1016/j.molcel.2018.01.039.

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

20. Qin, S. *et al.* Tudor domains of the PRC2 components PHF1 and PHF19 selectively bind to histone H3K36me3. *Biochem. Biophys. Res. Commun.* (2013) doi:10.1016/j.bbrc.2012.11.116.

21. Perino, M. *et al.* MTF2 recruits Polycomb Repressive Complex 2 by helical-shape-selective DNA binding. *Nat. Genet.* (2018) doi:10.1038/s41588-018-0134-8.

22. Zhang, Z. *et al.* PRC2 complexes with JARID2, MTF2, and esPRC2p48 in ES cells to modulate ES cell pluripotency and somatic cell reprograming. *Stem Cells* (2011) doi:10.1002/stem.578.

23. Beringer, M. *et al.* EPOP Functionally Links Elongin and Polycomb in Pluripotent Stem Cells. *Mol. Cell* (2016) doi:10.1016/j.molcel.2016.10.018.

24. Conway, E. *et al.* A Family of Vertebrate-Specific Polycombs Encoded by the LCOR/LCORL Genes Balance PRC2 Subtype Activities. *Mol. Cell* (2018) doi:10.1016/j.molcel.2018.03.005.

25. Kasinath, V. *et al.* Structures of human PRC2 with its cofactors AEBP2 and JARID2. *Science (80-.).* (2018) doi:10.1126/science.aar5700.

26. Lynch, M. D. *et al.* An interspecies analysis reveals a key role for unmethylated CpG dinucleotides in vertebrate Polycomb complex recruitment. *EMBO J.* (2012) doi:10.1038/emboj.2011.399.

27. Riising, E. M. *et al.* Gene silencing triggers polycomb repressive complex 2 recruitment to CpG Islands genome wide. *Mol. Cell* (2014) doi:10.1016/j.molcel.2014.06.005.

28. Sanulli, S. *et al.* Jarid2 Methylation via the PRC2 Complex Regulates H3K27me3 Deposition during Cell Differentiation. *Mol. Cell* (2015) doi:10.1016/j.molcel.2014.12.020.

29. Kasinath, V. *et al.* JARID2 and AEBP2 regulate PRC2 in the presence of H2AK119ub1 and other histone modifications. *Science (80-.).* (2021) doi:10.1126/science.abc3393.

30. Cooper, S. *et al.* Jarid2 binds mono-ubiquitylated H2A lysine 119 to mediate crosstalk between Polycomb complexes PRC1 and PRC2. *Nat. Commun.* (2016) doi:10.1038/ncomms13661.

31. Son, J., Shen, S. S., Margueron, R. & Reinberg, D. Nucleosome-binding activities within JARID2 and EZH1 regulate the function of PRC2 on chromatin. *Genes Dev.* (2013) doi:10.1101/gad.225888.113.

32. Perino, M., Mierlo, G. van, Wardle, S. M. T., Marks, H. & Veenstra, G. J. C. Two distinct functional axes of positive feedback-enforced PRC2 recruitment in mouse embryonic stem cells. *bioRxiv* (2019) doi:10.1101/669960.

33. Healy, E. *et al.* PRC2.1 and PRC2.2 Synergize to Coordinate H3K27 Trimethylation. *Mol. Cell* (2019) doi:10.1016/j.molcel.2019.08.012.

34. Loh, C. H., van Genesen, S., Perino, M., Bark, M. R. & Veenstra, G. J. C. Loss of PRC2 subunits primes lineage choice during exit of pluripotency. *Nat. Commun.* (2021) doi:10.1038/s41467-021-27314-4.

35. Youmans, D. T., Gooding, A. R., Dowell, R. D. & Cech, T. R. Competition between PRC2.1 and 2.2 subcomplexes regulates PRC2 chromatin occupancy in human stem cells. *Mol. Cell* (2021) doi:10.1016/j.molcel.2020.11.044.

36. van Mierlo, G., Veenstra, G. J. C., Vermeulen, M. & Marks, H. The Complexity of PRC2 Subcomplexes. *Trends in Cell Biology* (2019) doi:10.1016/j.tcb.2019.05.004.

37. Gao, Z. *et al.* PCGF Homologs, CBX Proteins, and RYBP Define Functionally Distinct PRC1 Family Complexes. *Mol. Cell* (2012) doi:10.1016/j.molcel.2012.01.002.

38. de Napoles, M. *et al.* Polycomb group proteins ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev. Cell* (2004) doi:10.1016/j.devcel.2004.10.005.

39. Wang, H. *et al.* Role of histone H2A ubiquitination in Polycomb silencing. *Nature* (2004) doi:10.1038/nature02985.

40. Rose, N. R. *et al.* RYBP stimulates PRC1 to shape chromatin-based communication between polycomb repressive complexes. *Elife* (2016) doi:10.7554/eLife.18591.

41. Zepeda-Martinez, J. A. *et al.* Parallel PRC2/cPRC1 and vPRC1 pathways silence lineage-specific genes and maintain self-renewal in mouse embryonic stem cells. *Sci. Adv.* (2020) doi:10.1126/sciadv.aax5692.

42. Kundu, S. *et al.* Polycomb Repressive Complex 1 Generates Discrete Compacted Domains that Change during Differentiation. *Mol. Cell* (2017) doi:10.1016/j.molcel.2017.01.009.

43. Boyle, S. *et al.* A Central Role for Canonical PRC1 in Shaping the 3D Nuclear Landscape.

bioRxiv (2019) doi:10.1101/2019.12.15.876771.

44. Schoenfelder, S. *et al.* Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem cell genome. *Nat. Genet.* (2015) doi:10.1038/ng.3393.

45. Plys, A. J. *et al.* Phase separation of polycomb-repressive complex 1 is governed by a charged disordered region of CBX2. *Genes Dev.* (2019) doi:10.1101/gad.326488.119.

46. Tatavosian, R. *et al.* Nuclear condensates of the Polycomb protein chromobox 2 (CBX2) assemble through phase separation. *J. Biol. Chem.* (2019) doi:10.1074/jbc.RA118.006620.

47. Blackledge, N. P. & Klose, R. J. The molecular principles of gene regulation by Polycomb repressive complexes. *Nature Reviews Molecular Cell Biology* (2021) doi:10.1038/s41580-021-00398-y.

48. Bernstein, B. E. *et al.* A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell* (2006) doi:10.1016/j.cell.2006.02.041.

49. Tamburri, S. *et al.* Histone H2AK119 Mono-Ubiquitination Is Essential for Polycomb-Mediated Transcriptional Repression. *Mol. Cell* (2020) doi:10.1016/j.molcel.2019.11.021.

50. Blackledge, N. P. *et al.* PRC1 Catalytic Activity Is Central to Polycomb System Function. *Mol. Cell* (2020) doi:10.1016/j.molcel.2019.12.001.

51. Zhang, Z. *et al.* Role of remodeling and spacing factor 1 in histone H2A ubiquitination-mediated gene silencing. *Proc. Natl. Acad. Sci. U. S. A.* (2017) doi:10.1073/pnas.1711158114.

52. Lehmann, L. *et al.* Polycomb repressive complex 1 (PRC1) disassembles RNA polymerase II preinitiation complexes. *J. Biol. Chem.* (2012) doi:10.1074/jbc.M112.397430.

53. Zhou, H. *et al.* Rixosomal RNA degradation contributes to silencing of Polycomb target genes. *Nature* 167–174 (2022).

54. Tavares, L. *et al.* RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. *Cell* (2012) doi:10.1016/j.cell.2011.12.029.

55. Frey, F. *et al.* Molecular basis of PRC1 targeting to polycomb response elements by PhoRC. *Genes Dev.* (2016) doi:10.1101/gad.279141.116.

56. Moussa, H. F. *et al.* Canonical PRC1 controls sequence-independent propagation of Polycomb-mediated gene silencing. *Nat. Commun.* (2019) doi:10.1038/s41467-019-09628-6.

57. Pengelly, A. R., Kalb, R., Finkl, K. & Müller, J. Transcriptional repression by PRC1 in the absence of H2A monoubiquitylation. *Genes Dev.* (2015) doi:10.1101/gad.265439.115.

58. Wang, L. *et al.* Hierarchical recruitment of polycomb group silencing complexes. *Mol. Cell* (2004) doi:10.1016/j.molcel.2004.05.009.

59. Shan, Y. *et al.* PRC2 specifies ectoderm lineages and maintains pluripotency in primed but not naïve ESCs. *Nat. Commun.* (2017) doi:10.1038/s41467-017-00668-4.

60. Miller, S. A., Damle, M., Kim, J. & Kingston, R. E. Full methylation of H3K27 by PRC2 is dispensable for initial embryoid body formation but required to maintain differentiated cell identity. *Dev.* (2021) doi:10.1242/dev.196329.

61. Pengelly, A. R., Copur, Ö., Jäckle, H., Herzig, A. & Müller, J. A histone mutant reproduces the phenotype caused by loss of histone-modifying factor polycomb. *Science (80-.).* (2013) doi:10.1126/science.1231382.

62. Jadhav, U. *et al.* Replicational Dilution of H3K27me3 in Mammalian Cells and the Role of Poised Promoters. *Mol. Cell* (2020) doi:10.1016/j.molcel.2020.01.017.

63. Margueron, R. *et al.* Ezh1 and Ezh2 Maintain Repressive Chromatin through Different Mechanisms. *Mol. Cell* (2008) doi:10.1016/j.molcel.2008.11.004.

64. Ardehali, M. B. *et al.* Polycomb Repressive Complex 2 Methylates Elongin A to Regulate Transcription. *Mol. Cell* (2017) doi:10.1016/j.molcel.2017.10.025.

65. Ragunathan, K., Jih, G. & Moazed, D. Epigenetic inheritance uncoupled from sequence-specific recruitment. *Science (80-.).* (2015) doi:10.1126/science.1258699.

66. Audergon, P. N. C. B. *et al.* Restricted epigenetic inheritance of H3K9 methylation. *Science (80-.).* (2015) doi:10.1126/science.1260638.

67. Wang, X. & Moazed, D. DNA sequence-dependent epigenetic inheritance of gene silencing and histone H3K9 methylation. *Science (80-.).* (2017) doi:10.1126/science.aaj2114.

68. Wang, X. *et al.* A composite DNA element that functions as a maintainer required for epigenetic inheritance of heterochromatin. *Mol. Cell* (2021) doi:10.1016/j.molcel.2021.07.017.

69. Schwartz, Y. B. *et al.* Genome-wide analysis of Polycomb targets in Drosophila melanogaster. *Nat. Genet.* (2006) doi:10.1038/ng1817.

70. Laprell, F., Finkl, K. & Müller, J. Propagation of Polycomb-repressed chromatin requires sequence-specific recruitment to DNA. *Science (80-.).* (2017) doi:10.1126/science.aai8266.

71. Coleman, R. T. & Struhl, G. Causal role for inheritance of H3K27me3 in maintaining the off state of a Drosophila HOX gene. *Science (80-.).* (2017) doi:10.1126/science.aai8236.

72. Ku, M. *et al.* Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes

of bivalent domains. *PLoS Genet.* (2008) doi:10.1371/journal.pgen.1000242.

73. Mendenhall, E. M., Koche, R. P., Truong, T., Zhou, V. W. & Issac, B. GC-Rich Sequence Elements Recruit PRC2 in Mammalian ES Cells. **6**, 1–10 (2010).

74. Shin Voo, K., Carlone, D. L., Jacobsen, B. M., Flodin, A. & Skalnik, D. G. Cloning of a Mammalian Transcriptional Activator That Binds Unmethylated CpG Motifs and Shares a CXXC Domain with DNA Methyltransferase, Human Trithorax, and Methyl-CpG Binding Domain Protein 1. *Mol. Cell. Biol.* (2000) doi:10.1128/mcb.20.6.2108-2121.2000.

75. Blackledge, N. P. *et al.* Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. *Cell* (2014) doi:10.1016/j.cell.2014.05.004.

76. Farcas, A. M. *et al.* KDM2B links the polycomb repressive complex 1 (PRC1) to recognition of CpG islands. *Elife* (2012) doi:10.7554/eLife.00205.

77. Stielow, B., Finkernagel, F., Stiewe, T., Nist, A. & Suske, G. MGA, L3MBTL2 and E2F6 determine genomic binding of the non-canonical Polycomb repressive complex PRC1.6. *PLoS Genet.* (2018) doi:10.1371/journal.pgen.1007193.

78. Scelfo, A. *et al.* Functional Landscape of PCGF Proteins Reveals Both RING1A/B-Dependent-and RING1A/B-Independent-Specific Activities. *Mol. Cell* (2019) doi:10.1016/j.molcel.2019.04.002.

79. Liu, M. *et al.* The polycomb group protein PCGF6 mediates germline gene silencing by recruiting histone-modifying proteins to target gene promoters. *J. Biol. Chem.* (2020) doi:10.1074/jbc.ra119.012121.

80. Yu, M. *et al.* Direct Recruitment of Polycomb Repressive Complex 1 to Chromatin by Core Binding Transcription Factors. *Mol. Cell* (2012) doi:10.1016/j.molcel.2011.11.032.

81. Schmitges, F. W. *et al.* Histone Methylation by PRC2 Is Inhibited by Active Chromatin Marks. *Mol. Cell* (2011) doi:10.1016/j.molcel.2011.03.025.

82. Kaneko, S., Son, J., Bonasio, R., Shen, S. S. & Reinberg, D. Nascent RNA interaction keeps PRC2 activity poised and in check. *Genes Dev.* (2014) doi:10.1101/gad.247940.114.

83. Jermann, P., Hoerner, L., Burger, L. & Schübeler, D. Short sequences can efficiently recruit histone H3 lysine 27 trimethylation in the absence of enhancer activity and DNA methylation. *Proc. Natl. Acad. Sci. U. S. A.* (2014) doi:10.1073/pnas.1400672111.

84. Holoch, D. *et al.* A cis-acting mechanism mediates transcriptional memory at Polycomb target genes in mammals. *Nat. Genet.* (2021) doi:10.1038/s41588-021-00964-2.

85. Kar, G. *et al.* Flipping between Polycomb repressed and active transcriptional states introduces noise in gene expression. *Nat. Commun.* (2017) doi:10.1038/s41467-017-

00052-2.

86. Angel, A., Song, J., Dean, C. & Howard, M. A Polycomb-based switch underlying quantitative epigenetic memory. *Nature* (2011) doi:10.1038/nature10241.

87. Klose, R. J., Cooper, S., Farcas, A. M., Blackledge, N. P. & Brockdorff, N. Chromatin Sampling-An Emerging Perspective on Targeting Polycomb Repressor Proteins. *PLoS Genet.* (2013) doi:10.1371/journal.pgen.1003717.

88. Berry, S., Dean, C. & Howard, M. Slow Chromatin Dynamics Allow Polycomb Target Genes to Filter Fluctuations in Transcription Factor Activity. *Cell Syst.* (2017) doi:10.1016/j.cels.2017.02.013.

89. Sneppen, K. & Ringrose, L. Theoretical analysis of Polycomb-Trithorax systems predicts that poised chromatin is bistable and not bivalent. *Nat. Commun.* (2019) doi:10.1038/s41467-019-10130-2.

90. Allshire, R. C., Javerzat, J. P., Redhead, N. J. & Cranston, G. Position effect variegation at fission yeast centromeres. *Cell* (1994) doi:10.1016/0092-8674(94)90180-5.

91. Berry, S., Hartley, M., Olsson, T. S. G., Dean, C. & Howard, M. Local chromatin environment of a Polycomb target gene instructs its own epigenetic inheritance. *Elife* (2015) doi:10.7554/eLife.07205.

92. Choi, J. *et al.* DNA binding by PHF1 prolongs PRC2 residence time on chromatin and thereby promotes H3K27 methylation. *Nat. Struct. Mol. Biol.* (2017) doi:10.1038/nsmb.3488.

93. Hansen, K. H. *et al.* A model for transmission of the H3K27me3 epigenetic mark. *Nat. Cell Biol.* (2008) doi:10.1038/ncb1787.

94. Bintu, L. *et al.* Dynamics of epigenetic regulation at the single-cell level. *Science* (2016) doi:10.1126/science.aab2956.

95. Uhlén, M. *et al.* Tissue-based map of the human proteome. *Science (80-.).* (2015) doi:10.1126/science.1260419.

96. Menke, A. L., Van Der Eb, A. J. & Jochemsen, A. G. The Wilms' tumor 1 gene: Oncogene or tumor suppressor gene? *International Review of Cytology* (1998) doi:10.1016/s0074-7696(08)60418-0.

97. Zhang, L. *et al.* EED, a member of the polycomb group, is required for nephron differentiation and the maintenance of nephron progenitor cells. *Dev.* (2018) doi:10.1242/dev.157149.

98. Vieira, J. M. *et al.* BRG1-SWI/SNF-dependent regulation of the Wt1 transcriptional landscape mediates epicardial activity during heart development and disease. *Nat. Commun.* (2017) doi:10.1038/ncomms16034.

99. Eisenberg, E. & Levanon, E. Y. Human housekeeping genes, revisited. *Trends in Genetics* (2013) doi:10.1016/j.tig.2013.05.010.

100. Seczynska, M., Bloor, S., Cuesta, S. M. & Lehner, P. J. Genome surveillance by HUSH-mediated silencing of intronless mobile elements. *Nature* (2022) doi:10.1038/s41586-021-04228-1.

101. Morey, L., Aloia, L., Cozzuto, L., Benitah, S. A. & Di Croce, L. RYBP and Cbx7 Define Specific Biological Functions of Polycomb Complexes in Mouse Embryonic Stem Cells. *Cell Rep.* (2013) doi:10.1016/j.celrep.2012.11.026.

102. Boyle, S. *et al.* A central role for canonical PRC1 in shaping the 3D nuclear landscape. *Genes Dev.* (2020) doi:10.1101/GAD.336487.120.

103. Bauer, M., Trupke, J. & Ringrose, L. The quest for mammalian Polycomb response elements: are we there yet? *Chromosoma* (2016) doi:10.1007/s00412-015-0539-4.

104. Li, G. *et al.* Jarid2 and PRC2, partners in regulating gene expression. *Genes Dev.* (2010) doi:10.1101/gad.1886410.

105. Petracovici, A. & Bonasio, R. Distinct PRC2 subunits regulate maintenance and establishment of Polycomb repression during differentiation. *Mol. Cell* (2021) doi:10.1016/j.molcel.2021.03.038.

106. Yuan, W. *et al.* H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J. Biol. Chem.* (2011) doi:10.1074/jbc.M110.194027.

107. Gatchalian, J., Kingsley, M. C., Moslet, S. D., Ospina, R. D. R. & Kutateladze, T. G. An aromatic cage is required but not sufficient for binding of Tudor domains of the Polycomblike protein family to H3K36me3. *Epigenetics* (2015) doi:10.1080/15592294.2015.1042646.

108. Brien, G. L. *et al.* Polycomb PHF19 binds H3K36me3 and recruits PRC2 and demethylase NO66 to embryonic stem cell genes during differentiation. *Nat. Struct. Mol. Biol.* (2012) doi:10.1038/nsmb.2449.

109. Sarma, K., Margueron, R., Ivanov, A., Pirrotta, V. & Reinberg, D. Ezh2 Requires PHF1 To Efficiently Catalyze H3 Lysine 27 Trimethylation In Vivo. *Mol. Cell. Biol.* (2008) doi:10.1128/mcb.02017-07.

110. Kang, H. *et al.* Sex comb on midleg (Scm) is a functional link between PcG-repressive complexes in Drosophila. *Genes Dev.* (2015) doi:10.1101/gad.260562.115.

111. Kloet, S. L. *et al.* The dynamic interactome and genomic targets of Polycomb complexes during stem-cell differentiation. *Nat. Struct. Mol. Biol.* (2016) doi:10.1038/nsmb.3248.

112. Smits, A. H., Jansen, P. W. T. C., Poser, I., Hyman, A. A. & Vermeulen, M. Stoichiometry of chromatin-associated protein complexes revealed by label-free

quantitative mass spectrometry-based proteomics. *Nucleic Acids Res.* (2013) doi:10.1093/nar/gks941.

113. Oliviero, G. *et al.* Dynamic protein interactions of the polycomb repressive complex 2 during differentiation of pluripotent cells. *Mol. Cell. Proteomics* (2016) doi:10.1074/mcp.M116.062240.

114. Campisi, A. *et al.* Gene Silencing of Transferrin-1 Receptor as a Potential Therapeutic Target for Human Follicular and Anaplastic Thyroid Cancer. *Mol. Ther. - Oncolytics* (2020) doi:10.1016/j.omto.2020.01.003.

115. O'Donnell, K. A. *et al.* Activation of Transferrin Receptor 1 by c-Myc Enhances Cellular Proliferation and Tumorigenesis. *Mol. Cell. Biol.* (2006) doi:10.1128/mcb.26.6.2373-2386.2006.

116. Madamba, E. V., Berthet, E. B. & Francis, N. J. Inheritance of Histones H3 and H4 during DNA Replication In Vitro. *Cell Rep.* (2017) doi:10.1016/j.celrep.2017.10.033.

117. Serra-Cardona, A. & Zhang, Z. Replication-Coupled Nucleosome Assembly in the Passage of Epigenetic Information and Cell Identity. *Trends in Biochemical Sciences* (2018) doi:10.1016/j.tibs.2017.12.003.

118. Xu, M. *et al.* Partitioning of histone H3-H4 tetramers during DNA replicationdependent chromatin assembly. *Science (80-.).* (2010) doi:10.1126/science.1178994.

119. Jamai, A., Imoberdorf, R. M. & Strubin, M. Continuous Histone H2B and Transcription-Dependent Histone H3 Exchange in Yeast Cells outside of Replication. *Mol. Cell* (2007) doi:10.1016/j.molcel.2007.01.019.

120. Burgess, R. J. & Zhang, Z. Histone chaperones in nucleosome assembly and human disease. *Nature Structural and Molecular Biology* (2013) doi:10.1038/nsmb.2461.

121. Zhao, J. *et al.* RYBP/YAF2-PRC1 complexes and histone H1-dependent chromatin compaction mediate propagation of H2AK119ub1 during cell division. *Nat. Cell Biol.* (2020) doi:10.1038/s41556-020-0484-1.

122. Jozwik, K. M., Chernukhin, I., Serandour, A. A., Nagarajan, S. & Carroll, J. S. FOXA1 Directs H3K4 Monomethylation at Enhancers via Recruitment of the Methyltransferase MLL3. *Cell Rep.* (2016) doi:10.1016/j.celrep.2016.11.028.

123. Kadoch, C. & Crabtree, G. R. Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Science Advances* (2015) doi:10.1126/sciadv.1500447.

124. Kuroda, M. I., Kang, H., De, S. & Kassis, J. A. Dynamic Competition of Polycomb and Trithorax in Transcriptional Programming. *Annual Review of Biochemistry* (2020) doi:10.1146/annurev-biochem-120219-103641.

125. Piunti, A. & Shilatifard, A. Epigenetic balance of gene expression by polycomb and compass families. *Science* (2016) doi:10.1126/science.aad9780.

126. Agger, K. *et al.* UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* (2007) doi:10.1038/nature06145.

127. Zhu, P. *et al.* A Histone H2A Deubiquitinase Complex Coordinating Histone Acetylation and H1 Dissociation in Transcriptional Regulation. *Mol. Cell* (2007) doi:10.1016/j.molcel.2007.07.024.

128. Lavarone, E., Barbieri, C. M. & Pasini, D. Dissecting the role of H3K27 acetylation and methylation in PRC2 mediated control of cellular identity. *Nat. Commun.* (2019) doi:10.1038/s41467-019-09624-w.

129. Long, Y. *et al.* RNA is essential for PRC2 chromatin occupancy and function in human pluripotent stem cells. *Nat. Genet.* (2020) doi:10.1038/s41588-020-0662-x.

130. Apte, M. S. & Meller, V. H. Homologue Pairing in Flies and Mammals: Gene Regulation When Two Are Involved. *Genet. Res. Int.* (2012) doi:10.1155/2012/430587.

131. Dehbi, M. & Pelletier, J. PAX8-mediated activation of the wt1 tumor suppressor gene. *EMBO J.* (1996) doi:10.1002/j.1460-2075.1996.tb00804.x.

132. Discenza, M. T., Vaz, D., Hassell, J. A. & Pelletier, J. Activation of the WT1 tumor suppressor gene promoter by Pea3. *FEBS Lett.* (2004) doi:10.1016/S0014-5793(04)00104-8.

133. McConnell, M. J., Cunliffe, H. E., Chua, L. J., Ward, T. A. & Eccles, M. R. Differential regulation of the human Wilms tumour suppressor gene (WT1) promoter by two isoforms of PAX2. *Oncogene* (1997) doi:10.1038/sj.onc.1201114.

134. Waehle, V., Ungricht, R., Hoppe, P. S. & Betschinger, J. The tumor suppressor WT1 drives progenitor cell progression and epithelialization to prevent Wilms tumorigenesis in human kidney organoids. *Stem Cell Reports* (2021) doi:10.1016/j.stemcr.2021.07.023.

135. Kreidberg, J. A. *et al.* WT-1 is required for early kidney development. *Cell* (1993) doi:10.1016/0092-8674(93)90515-R.

136. Wong, X., Luperchio, T. R. & Reddy, K. L. NET gains and losses: The role of changing nuclear envelope proteomes in genome regulation. *Current Opinion in Cell Biology* (2014) doi:10.1016/j.ceb.2014.04.005.

137. Buchwalter, A., Kaneshiro, J. M. & Hetzer, M. W. Coaching from the sidelines: the nuclear periphery in genome regulation. *Nature Reviews Genetics* (2019) doi:10.1038/s41576-018-0063-5.

138. Rego, A., Sinclair, P. B., Tao, W., Kireev, I. & Belmont, A. S. The facultative heterochromatin of the inactive X chromosome has a distinctive condensed ultrastructure.

J. Cell Sci. (2008) doi:10.1242/jcs.026104.

139. Labade, A. S., Karmodiya, K. & Sengupta, K. HOXA repression is mediated by nucleoporin Nup93 assisted by its interactors Nup188 and Nup205. *Epigenetics and Chromatin* (2016) doi:10.1186/s13072-016-0106-0.

140. Jacinto, F. V., Benner, C. & Hetzer, M. W. The nucleoporin Nup153 regulates embryonic stem cell pluripotency through gene silencing. *Genes Dev.* (2015) doi:10.1101/gad.260919.115.

141. Breiling, A., Turner, B. M., Bianchi, M. E. & Orlando, V. General transcription factors bind promoters repressed by Polycomb group proteins. *Nature* (2001) doi:10.1038/35088090.

142. Dellino, G. I. *et al.* Polycomb silencing blocks transcription initiation. *Mol. Cell* (2004) doi:10.1016/S1097-2765(04)00128-5.

143. Noordermeer, D. *et al.* The dynamic architecture of Hox gene clusters. *Science (80-.).* (2011) doi:10.1126/science.1207194.

144. Noordermeer, D. *et al.* Temporal dynamics and developmental memory of 3D chromatin architecture at Hox gene loci. *Elife* (2014) doi:10.7554/eLife.02557.

145. Narendra, V. *et al.* CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science (80-.).* (2015) doi:10.1126/science.1262088.

146. Lewis, E. B. The Theory and Application of a New Method of Detecting Chromosomal Rearrangements in Drosophila melanogaster. *Am. Nat.* (1954) doi:10.1086/281833.

147. Woodhouse, M. R., Freeling, M. & Lisch, D. Initiation, establishment, and maintenance of heritable MuDR transposon silencing in maize are mediated by distinct factors. *PLoS Biol.* (2006) doi:10.1371/journal.pbio.0040339.

148. Rassoulzadegan, M., Magliano, M. & Cuzin, F. Transvection effects involving DNA methylation during meiosis in the mouse. *EMBO J.* (2002) doi:10.1093/emboj/21.3.440.

149. Sandhu, K. S. *et al.* Nonallelic transvection of multiple imprinted loci is organized by the H19 imprinting control region during germline development. *Genes Dev.* (2009) doi:10.1101/gad.552109.

150. Rodriguez, J. D. *et al.* A model for epigenetic inhibition via transvection in the mouse. *Genetics* (2017) doi:10.1534/genetics.117.201913.

151. Bacher, C. P. *et al.* Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. *Nat. Cell Biol.* (2006) doi:10.1038/ncb1365.

152. Xu, N., Donohoe, M. E., Silva, S. S. & Lee, J. T. Evidence that homologous X-chromosome pairing requires transcription and Ctcf protein. *Nat. Genet.* (2007)

doi:10.1038/ng.2007.5.

153. Xu, N., Tsai, C. L. & Lee, J. T. Transient homologous chromosome pairing marks the onset of X inactivation. *Science (80-.).* (2006) doi:10.1126/science.1122984.

154. Brandt, V. L., Hewitt, S. L. & Skok, J. A. Communication between homologous alleles preserves genomic stability during V(D)J recombination. *Nucleus* (2010) doi:10.4161/nucl.1.1.10595.

155. Hogan, M. S., Parfitt, D. E., Zepeda-Mendoza, C. J., Shen, M. M. & Spector, D. L. Transient pairing of homologous Oct4 alleles accompanies the onset of embryonic stem cell differentiation. *Cell Stem Cell* (2015) doi:10.1016/j.stem.2015.02.001.

156. Atkin, N. B. & Jackson, Z. Evidence for somatic pairing of chromosome 7 and 10 homologs in a follicular lymphoma. *Cancer Genet. Cytogenet.* (1996) doi:10.1016/0165-4608(95)00360-6.

157. Thatcher, K. N., Peddada, S., Yasui, D. H. & LaSalle, J. M. Homologous pairing of 15q11 - 13 imprinted domains in brain is developmentally regulated but deficient in Rett and autism samples. *Hum. Mol. Genet.* (2005) doi:10.1093/hmg/ddi073.

158. Szulc, J., Wiznerowicz, M., Sauvain, M. O., Trono, D. & Aebischer, P. A versatile tool for conditional gene expression and knockdown. *Nat. Methods* (2006) doi:10.1038/nmeth846.

159. Vojta, A. *et al.* Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic Acids Res.* (2016) doi:10.1093/nar/gkw159.

160. Schindelin, J. *et al.* Fiji: An open-source platform for biological-image analysis. *Nature Methods* (2012) doi:10.1038/nmeth.2019.

161. Nowak, D. E., Tian, B. & Brasier, A. R. Two-step cross-linking method for identification of NF-κB gene network by chromatin immunoprecipitation. *Biotechniques* (2005) doi:10.2144/000112014.

162. Egan, B. *et al.* An alternative approach to ChIP-Seq normalization enables detection of genome-wide changes in histone H3 lysine 27 trimethylation upon EZH2 inhibition. *PLoS One* (2016) doi:10.1371/journal.pone.0166438.

163. Gardiner-Garden, M. & Frommer, M. CpG Islands in vertebrate genomes. *J. Mol. Biol.* (1987) doi:10.1016/0022-2836(87)90689-9.

164. Gerstein, M. B. *et al.* Architecture of the human regulatory network derived from ENCODE data. *Nature* (2012) doi:10.1038/nature11245.

165. Wang, J. *et al.* Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. *Genome Res.* (2012) doi:10.1101/gr.139105.112.

166. Wang, J. *et al.* Factorbook.org: A Wiki-based database for transcription factor-binding data generated by the ENCODE consortium. *Nucleic Acids Res.* (2013) doi:10.1093/nar/gks1221.

167. Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. DeepTools: A flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* (2014) doi:10.1093/nar/gku365.

168. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* (2009) doi:10.1093/bioinformatics/btp352.