Fall 12-2019

Recruitment of Polycomb-Group Proteins at giant in Drosophila Embryos

Elnaz Ghotbi Ravandi
Southern Methodist University, eghotbiravan@smu.edu

RECRUITMENT OF POLYCOMB-GROUP PROTEINS

AT *GIANT* IN *DROSOPHILA* EMBRYOS

Approved by:

_____

Prof. Richard Jones
Professor of Biology

_____

Prof. Bill Orr
Professor of Biology

_____

Prof. Robert Harrod
Associate Professor of Biology

_____

Prof. Tae Hoon Kim
Professor of Biology

RECRUITMENT OF POLYCOMB-GROUP PROTEINS AT *GIANT* IN *DROSOPHILA* EMBRYOS

A Dissertation Presented to the Graduate Faculty of the

Dedman College

Southern Methodist University

in

Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

with a

Major in Molecular and Cell Biology

by

Elnaz Ghotbi Ravandi

B.S., Plant Biology, Bahonar University of Kerman
M.S., Plant Cellular Biology, Bahonar University of Kerman

December 21, 2019

ACKNOWLEDGMENTS

I would like to express my appreciation and gratefulness to my advisor, Professor Richard Jones, for giving me the amazing opportunity to pursue my Ph.D. in his lab. Dr. Jones always inspired me with his amazing analytical thinking abilities, honesty, humbleness and sincere affection for science. He will be my role model for the rest of my scientific life. My deepest gratitude goes to my committee members, Professors Bill Orr, Robert Harrod and Tae Hoon Kim for providing me with scientific advices and suggestions throughout the last five and half years.

I am very thankful to my lab colleague, Piao Ye, and fly lab manager, Judith Benes, for all of their support, help and guidance with my project. I have been very blessed for being surrounded by amazing friends: Lacin Yapindi, Tetiana Bowley, Lena Odnokoz, Amila Nanayakkara and Aditi Malu, who made my graduate years very memorable.

I would like to extend my special thanks to my parents, Eshrat and Kayvan, and my siblings for their constant support and motivation. Finally, the main gratitude goes to my husband, Ishmael Dehghan, whose patience, love and support continuously motivated and encouraged me through my graduate years.

Elnaz Ghotbi Ravandi B.S., Plant Biology, Bahonar University of Kerman
M.S., Plant Cellular Biology, Bahonar University of Kerman

Recruitment of Polycomb-Group proteins at *giant* in *Drosophila* embryos

Advisor: Richard S. Jones, Ph.D.
Doctor of Philosophy conferred December 21, 2019
Dissertation completed December 2, 2019

Polycomb Group (PcG) proteins are evolutionarily conserved epigenetic transcriptional regulators that maintain the transcriptional repression of silenced genes. PcG proteins are mainly recognized as negative regulators of Hox genes but they target hundreds of other developmental decision makers and signaling factors. Mammalian PcG proteins are involved in maintaining the pluripotent state of stem cells and their misexpression may lead to a number of human cancers.

Recruitment of *Drosophila* PcG proteins requires the presence of specific DNA sequences called Polycomb Response Elements (PREs). PREs are complex elements which vary in sequence composition and contain binding sites for different DNA binding proteins. We have previously mapped two PREs for PcG target gene, *giant* (*gt*): PRE1 overlaps with *gt* promoter region, and PRE2, located approximately 6 kb upstream of *gt* transcription start site (TSS). However, whether these two PREs are functionally redundant for the recruitment of PcG proteins and PcG-mediated repression or contribute to independent aspects of *gt* regulation needed to be determined. Here, we showed that mutation of predicted extended consensus Pleiohomeotic (Pho) binding sites, led to the loss of Pho binding as well as significantly reduced recruitment of Polycomb repressive complex 1 (PRC1) and Polycomb repressive complex 2 (PRC2) to PRE1 in embryos from cellular blastoderm and mid-embryogenesis stages. We further showed that recruitment of PcG proteins and maintenance of the transcriptional repression of *gt* by PRE2 is independent from PRE1 activity. Pho-like (Phol) is partially redundant with Pho during larval development and binds to

the same DNA sequences in vitro. Surprisingly, we showed that Phol binding is less dependent on the presence of consensus Pho/Phol binding sites and appears to play a minimal role in recruiting other PcG proteins to *gt*. Moreover, our results suggested that PRE binding proteins, Spps and Dsp1, are differently dependent on the presence of Pho for PRE1 binding. Spps binding to PRE1 is dependent on Pho. On the other hand, Dsp1 binds PRE1 independently of Pho.

PcG mediated silencing is divided into two phases: initiation and maintenance. During the initiation phase, PcG proteins initially recognize and bind to their target genes. Once PcG proteins are recruited to their target genes, they can maintain transcriptional repression through an unlimited number of cell cycles. Most studies on PcG proteins have been focused on the maintenance phase of PcG silencing, and the molecular mechanisms by which PcG proteins are initially recruited to their target genes remained unknown. Two models have been proposed for the initial recruitment of PcG proteins to their target genes: instructive and responsive models. Instructive model suggests that transcription factors regulate recruitment of PcG proteins to the chromatin of their target genes. On the other hand, the responsive model suggests that recruitment of PcG proteins is dependent on the transcriptional state of the target genes and repressed chromatin is more compatible for the assembly and stable binding of PcG proteins. In order to experimentally test these two models, we examined the recruitment of PcG silencing complexes at a transcriptionally inert *gt* transgene in a background in which endogenous *gt* is transcriptionally active. We demonstrated that PcG proteins do not respond to *gt* transcriptional state. Furthermore, we provided evidence for the inhibitory effect of the *gt* transcriptional activator, Caudal (cad), on the recruitment of PcG proteins and proposed that this inhibitory effect is antagonized by the presence of the *gt* repressor, Hunchback (Hb).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

This dissertation is dedicated to my husband and best friend, Ishmael.

CHAPTER 1:

BACKGROUND

1.1 Polycomb group proteins

PcG proteins are essential regulators of development and differentiation that maintain the transcriptional repression of silenced genes by altering chromatin structure (Simon and Kingston, 2013).

Polycomb (*Pc*) was found in *Drosophila melanogaster*, as the first PcG gene. While performing genetic studies, Lewis (1947) isolated a dominant *Pc* mutation with the phenotype of additional sex combs, a row of modified bristles, on the second and third pairs of legs of adult males. Additional dominant and recessive mutations with the same phenotype were identified in the following years. Lewis (1978) reported that mutation of *Pc* results in the transformation of thoracic and first seven abdominal segments into the eighth abdominal segment and proposed that Pc is a global repressor of all the Bithorax complex (BX-C) genes. Genetic screenings identified a number of genes in which mutations caused the extra-sex-combs phenotype, resembling the phenotype of weak Pc mutant. These genes were collectively referred to as PcG proteins (Jürgens, 1985).

Genome-wide studies in *Drosophila*, like chromatin immunoprecipitation (ChIP)-chip and ChIP-seq, demonstrated the accumulation of PcG proteins at hundreds of genomic loci showing that they have a role in silencing many target genes encoding transcription factors, receptors,

signaling proteins, morphogens and regulators. PcG proteins are present and conserved in plants, *Drosophila* and mammals (Oktaba et al., 2008).

1.2 PcG proteins complexes

Drosophila PcG proteins assemble into three distinct complexes: Polycomb repressive complex 1 (PRC1), which can be subdivided into canonical (cPRC1) and non-canonical (ncPRC1) complexes, Polycomb repressive complex 2 (PRC2) and Pho-repressive complex (PhoRC) (**Figure 1.1**).

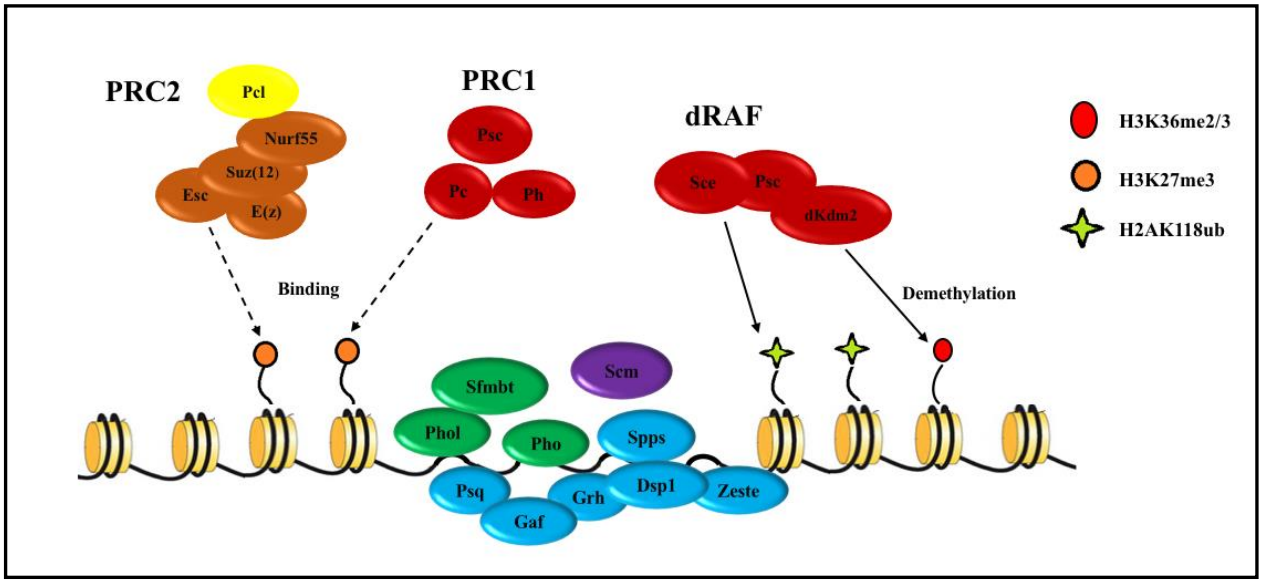

Figure 1.1- **PcG core complexes and their associating proteins in *Drosophila*.**

1.2.1 Polycomb repressive complex 1

Drosophila PRC1 is composed of the core components; Polycomb (Pc or its mammalian homolog Cbx2, 4, 6, 7 and 8), Polyhomeotic (Ph or its mammalian subunit Phc1-3), Posterior sex combs (Psc, or its six mammalian homologs, Polycomb group RING fingers [PCGFs] including Bmi1/PCGF4) and Sex combs extra (Sce, also known as dRing1 and its mammalian homologous subunits, Ring1A and Ring1B) (Di Croce and Helin, 2013; Schwartz and Pirrotta, 2013). Mutations in genes that encode components of PRC1 result in mis-expression of PcG target genes and embryonic lethality. Embryos lacking both maternal and zygotic Pc show the homeotic phenotype, which is transformation of all thoracic and abdominal segments to an eighth abdominal segment identity (Haynie, 1983; Lawrence et al., 1983).

Pc contains a chromodomain which can recognize and bind H3K27me3 histone modification. Studies showed that chromodomain is required and sufficient to anchor Pc to PcG-regulated genes (Messmer et al., 1992; Platero et al., 1995; Cao et al., 2002; Fischle et al., 2003; Min et al., 2003; Wang et al., 2004).

Ph contains a protein interaction sterile alpha motif (SAM) domain, which is also present in other PcG proteins, Sfmbt and Scm (Isono et al., 2013). Ph subunits multimerize through their SAM domains and undergo clustering that enhances binding and spreading of PcG proteins (Isono et al., 2013). Mutations in SAM domain of Ph result in the disruption of polymerization and complete loss of Ph activity and PcG-mediated repression of target genes (Gambetta and Müller, 2014). These findings proposed that SAM domain of Ph is essential for the silencing activity of PRC1.

The Psc subunit is involved in the PRC1-mediated chromatin compaction and inhibition of nucleosome remodeling complexes (King et al., 2005). Psc can be replaced by its functional

homolog, Su(z)2, in PRC1 complex (Lo et al., 2009). Psc and Su(z)2 share similar amino-acid compositions in their C-terminal and contain a 200-amino acid homology region (Brunk et al., 1991; van Lohuizen et al., 1991; Emmons et al., 2009). The C-terminal non-homologous region is required for chromatin compaction and inhibition of chromatin remodeling, while the homology region mediates incorporation of Psc/Su(z)2 into the PRC1 complex (King et al., 2005; Lo et al., 2009).

Sce, also known as dRing, monoubiquitylates lysine 118 of histone H2A (H2AK118ub1) (Wang et al., 2004). Pengelly et al. (2015) showed that the catalytic activity of Sce is not necessary for repression of canonical PcG target genes.

Various ncPRC1s have been identified in both flies and vertebrates. Mammalian PRC1 complexes are very complex and diverse because each *Drosophila* subunit has several mammalian homologs and can form combinational complexes. Mammalian ncPRC1 complexes containing different PCGFs, occupy distinct genomic loci and have specific enzymatic activities compared to canonical complexes. PRC1 diversity indicates that each PCGF, or the accessory proteins associated with it, may drive targeting by a different mechanism (Gao et al., 2012).

Drosophila ncPRC1 complex, dRAF, contains Psc, dRing and the histone H3K36 demethylase, Kdm2 (Gao et al., 2012). Lagarou et al. (2008), showed that dRAF complex in *Drosophila,* but not cPRC1 complex, is responsible for catalyzing the H2AK118ub1 both in vitro and in vivo. The ubiquitylation mark recruits PRC2 complex and inhibits elongation of RNA polymerase II (RNAPII) (Blackledge et al., 2014). However, other studies showed that cPRC1 and ncPRC1 mediated H2AK118ub resulted in the recruitment of PRC2 and deposition of H3K27me3 (Blackledge et al., 2014; Kalb et al., 2014; Cooper et al., 2016). Moussa et al. (2019) reported that ectopic recruitment of both cPRC1 and ncPRC1 subunits in mouse embryonic stem cells (mESCs)

resulted in the formation of a functional PRC1 and transcriptional silencing. However, the same study further showed that only cPRC1-dependent gene repression can be maintained through genome replication.

1.2.2 Polycomb repressive complex 2

PRC2 core complex is composed of Enhancer of zeste (E(z), or its mammalian homologs; Ezh1 or Ezh2), Suppressor of zeste 12 (Su(z)12), or its mammalian homolog with the same name), Extra sex combs (Esc), or its mammalian homolog Eed) and p55 (Nurf55 or Caf1, or its mammalian homologs; RBBP4 and RBBP7) (Di Croce and Helin, 2013; Schwartz and Pirrotta, 2013).

E(z) contains a SET domain which mono di and trimethylates lysine 27 of histone H3 (H3K27) (Cao et al., 2002; Czermin et al., 2002; Kuzmichev et al., 2002; Muller et al., 2002; reviewed in Cao and Zhang, 2004). Deposition of H3K27me3 is required for PcG-mediated repression.

Su(z)12 interacts with E(z) and is required for the activity and stable formation of PRC2 complex through a conserved VEFS-box domain (O'Meara and Simon, 2012).

The Esc subunit facilitates protein-protein interactions through its WD (tryptophan and aspartic acid residues) repeats (O'Meara and Simon, 2012), and allosterically enhances PRC2 repression by preferential binding to H3K27me3 and allowing local spreading of silencing (Margueron et al., 2009; O'Meara and Simon, 2012). E(z) possesses a low enzymatic activity in the absence of Esc and Su(z)12 (reviewed in O'Meara and Simon, 2012).

P55 is the only subunit of PRC2 complex present in a number of other chromatin remodeling complexes (O'Meara and Simon, 2012). P55 can physically interact with Su(z)12 as well as H3 and H4, but little is known about the functional significance of these interactions,

because loss of p55 has subtle consequence on PRC2 activity (Nowak et al., 2011; O'Meara and Simon, 2012; Wen et al., 2012).

1.2.3 Pho-repressive complex

Two Pho-containing PcG complexes exist in *Drosophila*. Pho repressive complex (PhoRC), consists of Pho or Phol bound to Scm-related gene containing four mbt domains (Sfmbt) (Klymenko et al., 2006; Grimm et al., 2009). A second Pho-containing complex is Pho-INO80 that contains the INO80 nucleosome remodeling complex (Klymenko, 2006).

Among all the PcG proteins, Pho and Phol are the only subunits with verified DNA-binding activity. They bind DNA in a sequence-specific manner and have central roles in the recruitment of PcG complexes to their target genes (Grossniklaus and Paro, 2014). Pho and Phol are partially redundant for the recruitment of PcG proteins at *bxd* PRE in wing imaginal discs (Wang et al., 2004).

Phol and Pho share 80% sequence identity and all amino acids involved in making DNA contacts are conserved in these two proteins. They also share a conserved spacer domain which binds to Sfmbt (Alfieri et al., 2013). Brown et al. (2003) showed that in gel shift assays, Phol binds to the oligonucleotide containing a Pho-binding site but not to the one with a mutated Pho binding site, suggesting that Pho and Phol can recognize and bind to the same sequence of DNA.

Kahn et al. (2014) showed that both Pho and Phol can form a complex with Sfmbt, but Pho is favored and outcompetes Phol in PhoRC complex. Consistent with it, the same authors reported that the interaction of Pho and Phol with Sfmbt is mutually exclusive, as they showed that Phol binding was increased upon RNAi knockdown of Pho. Furthermore, genome-wide analysis suggested different genomic distributions for Pho and Phol. Although ChIP-chip peaks for Pho

correspond to PREs and colocalize with Sfmbt, Phol peaks reside outside PcG domains, within 300 bp of TSS of a subset of active genes in Sg4 cells (Kahn et al., 2014).

Both Phol and Pho contain four zinc fingers that are 96% and 80% identical to those of Yin Yang1 (YY1), a mammalian transcription factor, respectively (Brown et al., 2003). Despite YY1 binding to the similar DNA sequnces as Pho and Phol, it does not bind to mammalian PcG target genes. YY1 transgene is able to rescue homeotic phenotypes of Pho mutant flies (Atchison et al., 2003), suggesting an efficient interaction of YY1 with Sfmbt. However, the interactions involved in the recruitment of PcG proteins to PREs in *Drosophila* are not conserved in mammals, as YY1 does not play a role in the mammalian PcG-mediated repression (Kahn et al., 2014).

1.3 Polycomb response elements

1.3.1 Introduction

Zink and Paro (1989) detected PcG proteins at specific bands on *Drosophila* salivary gland polytene chromosomes. Some of these bands represented locations of Hox genes suggesting the presence of PcG-recruiting DNA sequences in these genes. In 1991, Muller and Bienz identified a DNA fragment from the regulatory region of the homeotic gene, *Ultrabithorax* (*Ubx*), which was able to maintain the β-galactosidase expression in a pattern recapitulating of the endogenous *Ubx* throughout embryogenesis (Müller and Bienz, 1991). Soon after, more DNA fragments were discovered that could recruit PcG proteins in transgenic experiments. These specific cis-regulatory sequences were called “Polycomb response elements” (PREs).

Recruitment of *Drosophila* PRC2 and PRC1 to their target genes needs the presence of one or more PREs (Simon et al., 1993). PREs are complex elements which vary in sequence composition and size (Kassis and Kennison, 2010). Some PREs are located at the promoter of genes, depleted of nucleosomes. These PREs are well-positioned to regulate the transcriptional activity of the PcG target genes. Other PREs, such as Hox PREs, are located many kilobases far

from the promoters they control, and form looping interactions with the promoters of their repressed target genes (Oktaba et al., 2008).

Genome-wide studies on the PcG proteins, estimated the presence of two hundred PREs in the *Drosophila* genome (Schwartz et al., 2010). Although many PREs are known, very few PREs have been studied in detail. The *bxd* PRE from the cis-regulatory region of *Ubx* was the first PRE to be identified by several research groups in 1990s (Simon et al., 1993; Chan et al., 1994).

Although studies of different *Drosophila* PREs have identified a number of binding sites to be important for PRE activity (reviewed in Kassis and brown, 2013), the exact sequence composition required for PRE activity remain unknown (Kassis and Brown, 2013). This is attributable to the presence of different combinations of PRE binding proteins and low conservation of their consensus binding site sequences. PREs contain binding sites for many different DNA binding proteins including Pho, GAGA factor (Gaf), Pipsqueak (Psq), Dorsal switch protein 1 (Dsp1), Zeste, Sp1 factor for pairing-sensitive silencing (Spps) and Grainyhead (Grh) (Kassis and Kennison, 2010).

Previous studies showed that not all identified PRE binding proteins are present at all PREs. Pho is the only known PRE binding protein, identified at all characterized PREs (Kassis and Brown, 2013). Moreover, PRE binding proteins can bind to many sites that are not within PREs (Kassis and brown, 2013). Several research groups have attempted to develop alignment algorithms to predict location of PREs based on the consensus sequences of PRE binding proteins (Ringrose et al., 2003; Fielder and Rehmsmeier, 2006). However, prediction programs have failed due to the heterogeneous nature of PREs (Schwartz et al., 2006; Oktaba et al., 2008; Cunningham et al., 2010).

1.3.2 Mapping *Drosophila* PREs

PRE fragments were discovered in transgenic *Drosophila* using three approaches. First, the majority of PREs produce a phenomenon called "pairing-sensitive silencing" (PSS) in which PRE-containing transgene represses the expression of the P element reporter gene, mini-*white*, and the degree of repression is stronger in flies homozygous for PRE–containing P element at same or nearby sites compared to the flies heterozygous for it (Kassis, 1994). However, not all characterized PREs produce PSS (Muller et al.,1999), and minimal PRE fragments may need additional sequences to act as a pairing-sensitive silencer.

Second, when incorporated in a transgenic construct and integrated elsewhere in *Drosophila* genome, PREs are able to prevent the ectopic expression of reporter genes, and this repression is dependent on PcG proteins (Müller and Bienz, 1991; Simon et al., 1993; Chan et al., 1994; Chiang et al., 1995). Simon et al. (1993) showed that PRE–including transgenes containing *Ubx* regulatory regions were able to maintain PcG repression of a *lac*Z reporter gene in a pattern recapitulating that of the endogenous *Ubx* expression. Moreover, loss of the *bxd* PRE activity of *Ubx* and ectopic expression of the transgene was observed in a *Pc* mutant background (Fristch et al., 1999). Additional characterization of the *bxd* PRE indicated that although a short ~560 bp fragment was able to maintain PcG repression, improved maintenance could be obtained with larger constructs. The latter finding proposed that sequences flanking the strictly delimited PREs may contribute to PRE activity, or alternatively, PREs could cooperate with one another in the regulatory regions (Simon et al., 1993; Chan et al., 1994; Cunningham et al., 2010).

Third, PRE-containing transgenes were able to form new PcG protein binding sites in the polytene chromosomes of the larval salivary glands (Zink et al., 1991).

1.3.3 PRE binding proteins implicated in PRE function

Multiple studies suggested that sequence-specific PRE binding proteins play important roles in the recruitment of PcG complexes to PREs. Consensus sequences for the identified PRE binding proteins are listed in **Table 1.1.**

Table 1.1- **Consensus binding sites for known PRE binding proteins** (Pirrotta et al., 2017).

Protein	**Sequence Specifity**
Pho	GCCAT(T/A)TT
Phol	GCCATTAC
Gaf	GAGAG
Psq	GAGAG
Spps	(G/A)(G/A)GG(C/T)G(C/T)
Dsp1	GAAAA
Grh	TGTTTTTT

1.3.3.1 Pho and Phol

Pho is the only PRE binding protein that is present at all well characterized PREs. Pho binding sites are required for the PRE activity of the endogenous *Ubx* gene (Kozma et al., 2008), but are not sufficient to recruit PcG proteins (Americo et al., 2002).

Wang et al. (2004) reported that Pho and Phol play an important but redundant role in the recruitment of PcG proteins at *bxd* PRE, other genetic experiments also demonstrated stronger derepression of *Ubx* in wing discs in *pho; phol* double mutants compared to *pho* mutants (Brown et al., 2003).

Genome–wide studies have shown that binding of Pho is strongly correlated with genomic location of PRC1 subunits, strengthening the consensus view of Pho's central role in the recruitment of PcG proteins (Oktaba et al., 2008). Despite strong correlation of Pho binding with Pc and Ph, Phol peaks do not show a high correlation (only 21%) with genomic positions bound by PRC1 (Schuettengruber et al., 2009). However, Brown et al. (2003) proposed that Pho and Phol recognize and bind to the same sequence of DNA.

1.3.3.2 Gaf and Psq

GAGAG sequences are the binding sites for two proteins, Gaf and Psq (Lehmann et al., 1998). Gaf/Psq-binding sites are important for the function of many PREs (reviewed in Kassis, 2002; Fujioka et al., 2008). In vitro studies suggest that binding of Gaf makes the DNA more accessible for Pho binding (Mahmoudi et al., 2003). However, mutated Gaf sites showed that Gaf had no effect on *bxd* PRE activity (Brown et al., 2003). ChIP studies have shown that Gaf binds to about 50% of sites bound by PcG proteins (Negre et al., 2006; Schuettengruber et al., 2009). Studies showed that Psq mutation increases derepression of *Ubx* in larvae heterozygous for a Pc allele, indicating that Psq might play an important role in PcG repression (Huang et al., 2002).

1.3.3.3 Spps

Spps is a member of Sp1/KLF zinc finger protein family. Studies showed that Sp1/KLF binding sites are required for PSS and PRE activity of a 181 bp PRE from *engrailed* (*en*) gene (Americo et al., 2002; Brown et al., 2005). Spps colocalizes with PRC1 in polytene chromosomes, and its mutation enhances *pho* mutant phenotype, indicating that Pho and Spps either function cooperatively or in different parts of the same pathway to repress PcG target genes (Brown and Kassis, 2010). In a recent study, Brown et al. (2017) showed loss of recruitment of Pho, E(z), and PRC1 subunits and a global reduction of H3K27me3 in PcG domains in mutant larvae lacking

both maternal and zygotic Spps. The latter study suggested Spps is involved in PcG stable binding or recruitment.

1.3.3.4 Zeste

Zeste is involved in both activating and silencing activities of PREs depending on the context (Saurin et al., 2001). Genome-wide binding data showed small overlap between Zeste and Pho and/or Ph (Oktaba et al., 2008; Schuettengruber et al., 2009). *Zeste* mutants do not show PcG phenotypes and are homozygous viable and fertile (Goldberg et al., 1989).

1.3.3.5 Dsp1

Dsp1 plays an important role in the recruitment of PcG complexes to polytene chromosomes. *Dsp1* mutants die prematurely as adults with homeotic phenotypes (Decoville et al., 2001). Dejardin et al. (2005) showed that mutations of Dsp1 binding sites within the PREs of *en* and *Fab7*, abrogated PSS of mini-*white* gene, suggesting that Dsp1 is required for the silencing activity of PREs. Genome-wide studies on *Drosophila* embryos demonstrated that Dsp1 binds to about 50% of the Ph/Pc sites, while its consensus binding sites are absent at these PREs (Schuettengruber et al., 2009). Dsp1 contains two high mobility group (HMG) domains which bind minor groove of DNA, and introduces a prominent bend into the DNA (Agresti and Bianchi, 2003; Stros, 2010), facilitating long-range interactions or binding of PcG proteins.

1.3.3.6 Grh

Grh was first identified at the *iab-7* PRE, and shown to interact with Pho in vitro (Blastyak et al., 2006). There are discrepancies in the literature about the consensus binding sites for Grh. Studies have suggested that binding and presence of Grh may be important for only a subset of PREs. Tuckfield et al. (2002) have found that interaction of mammalian Grh-family member, CP2, with a mammalian Ring protein, DinG, was necessary for transcriptional repression by PcGs. *Drosophila* Grh can act as either a transcriptional repressor or activator in the regulation of many

genes. Hur et al. (2002) showed that Grh binding sites were not able to maintain PcG-mediated repression of a *Ubx-lac*Z reporter construct in *Drosophia* embryos. The role of Grh in PcG-mediated repression is not clear yet.

1.3.4 PREs in vertebrates

In recent years, few mouse or human DNA fragments were shown to act like *Drosophila* PREs in recruiting PRC2 or PRC1 subunits, when incorporated in transgenic constructs and integrated the in genome (Mendenhall et al., 2010; Arnold et al., 2013; Woo et al., 2010). Some of these elements appear to autonomously recruit both PRC complexes, while others recruit only one of the two complexes.

Little is known about DNA binding proteins or sequence composition involved in the recruitment of mammalian PcG proteins. Studies showed that unmethylated CpG islands are the functional equivalent of PREs in vertebrates (Farcas et al., 2012; Klose et al., 2013). CpG islands are regions of high density of cytosine and guanine dinucleotides compared to the rest of the genome. Since methylated cytosine tends to mutate to thymine, methylation of CpG dinucleotides elsewhere in the genome leads to their eventual depletion over evolutionary time. The majority of CpG islands do not undergo DNA methylation, and thus maintain a high density of CpG dinucleotides (Deaton and Bird, 2011). CpG islands, lacking bound-transcriptional activators, have been reported to target PRC2 (Mendenhall et al., 2010; Lynch et al., 2012; Jermann et al., 2014).

1.4 Trithorax group proteins and TrxG response elements

Trithorax group (TrxG) proteins are required to maintain the active state of their target genes. Several of these proteins were initially identified in *Drosophila* as suppressors of homeotic phenotype in a screen for suppressors of *Pc* or *Antennapedia* (*Antp*) mutations (Kennison and Tamkun, 1988).

Members of the TrxG complex are involved in chromatin remodeling, histone modification and transcription initiation and elongation (Kingston and Tamkun, 2014; Steffen and Ringrose, 2014). For examples; Set1, Trithorax (Trx) and Trithorax-related (Trr) are involved in methylation of H3K4. Set1 is responsible for global gene activation, while Trx and Trr target specific genes (Schuettengruber et al., 2011; Shilatifard, 2012). CBP acetylates H3K27, while UTX demethylates H3K27. Ash1, another subunit of the TrxG, contains histone methyl transferase activity specific for H3K36 (Schuettengruber et al., 2011).

Studies have shown that TrxG proteins also act through PREs, which in this case are called TrxG response elements (TREs) (reviewed in Ringrose and Paro, 2004). The first evidence showing that PREs could maintain either the active or silenced chromatin state comes from testing a DNA fragment from the BX-C, *Fab-7*, within a vector with a Gal4-inducible *lac*Z gene followed by the mini-*white* marker (Cavalli and Paro, 1998). Inclusion of PRE-containing *Fab-7* fragment, resulted in the transcriptional repression of both *lac*Z and mini-*white* genes. However, after Gal4-mediated activation of *lac*Z, expression of both genes, *lac*Z and *mini-white*, were maintained in the developing embryos (Cavalli and Paro, 1998). Further studies indicated that the transcriptional state of a DNA element, determines whether they act as PREs or TREs (Schmitt et al., 2005).

An enhancer-trap assay (reviewed by Kassis and Brown, 2013), using an *en-lac*Z transgene which contains two PREs of *en*, showed that each PRE could interact with flanking regulatory DNA as either PRE or TRE based on the genomic context. The latter finding suggests that PREs form looping interactions with enhancers or silencers surrounding the transgene insertion site.

1.5 Mechanisms of recruitment of PcG proteins

1.5.1 PcG-associated proteins

1.5.1.1 Polycomb-like

In addition to the core components of PcG complexes, additional proteins interact with these complexes and contribute to PcG-mediated repression. Polycomb-like (Pcl) is a substoichiometric subunit of PRC2. Nekrasov et al. (2007) reported that in *Drosophila* embryos, a fraction of PRC2 contains Pcl as a stable subunit and inclusion of Pcl in PRC2 is necessary for deposition of high levels of H3K27me3 at PcG target genes. In larvae, Pcl exists in a complex that is distinct from PRC2 or PRC1 and contributes to target site binding by the latter complexes (Savla et al., 2008). The same authors further showed that Pcl binds to *bxd* PRE in the absence of PRC2 and PRC1, while its binding was dependent on the presence of Pho and Phol. Choi et al. (2017) showed that Pcl contains a winged helix DNA binding domain, through which it interacts with DNA and extends PRC2 residence time on chromatin and therefore stimulates deposition of H3K27me3.

Mammalian homologs of Pcl; PHF1, MTF2, and PHF19, all contain one Tudor domain and two PHD fingers (Perino et al., 2018). MTF2 and PHF19 both recruit PRC2 to its target genes in mESCs, but through different ways. PHF19 needs its Tudor domain for PRC2 targeting to chromatin, while MTF2 uses its second PHD finger for PRC2 recruitment (Ballaré et al., 2012; Brien et al., 2012; Casanova et al., 2011; Hunkapiller et al., 2012; Walker et al., 2010). Furthermore, PHF1 was reported to trigger the activity of PRC2, both in HeLa cells and recombinant nucleosomes, but it does not seem to be involved in recruiting PRC2 to chromatin (Cao et al., 2008; Sarma et al., 2008). Perino et al. (2018) showed that MTF2 was essential for DNA-driven recruitment of PRC2 complex in mouse ESCs. MTF2 interacts with DNA backbone,

and the associated helical structure dictate MTF2 binding to unmethylated CpGs (Perino et al., 2018).

1.5.1.2 Jarid2

Jarid2 is a member of Jumonji family of proteins, which does not possess histone demethylase activity unlike other members of this family (Sanulli et al., 2015). Grijzenhout et al. (2016) reported identification of a PRC2 complex containing Jarid2. A more recent study showed that Jarid2 is required for both the initial targeting and stable binding of PRC2 to chromatin (Oksuz et al., 2018). Interestingly, genome-wide localization studies also showed that Jarid2 is needed for the binding of PcG proteins to their target genes in ESCs. Jarid2 has two DNA binding domains, an AT-rich interaction domain (ARID), required for the recruitment of PRC2 to its target genes (Pasini et al., 2010), and a Zn-finger domain, which specifically binds GC-rich regions (Li et al., 2010). Considering that PcG proteins are associated with GC-rich promoters, these data suggest a role for Jarid2 in the recruitment of PcGs to their target genes.

Jarid2 was reported to be di- and tri-methylated by PRC2 complex on lysine 116 in mammalian cells (Sanulli et al., 2015). Although methylated Jarid2 was shown to be dispensable for the recruitment of PRC2, it promotes enzymatic activity of E(z) and is required for the proper deposition of PRC2-mediated H3K27me3 (Sanulli et al., 2015).

1.5.1.3 Sex Comb on Midleg (Scm)

Scm is conserved between *Drosophila* and mammals, and is largely found in an uncharacterized protein complex that is distinct from PRC1 in *Drosophila* embryos (Peterson et al., 2004). However, its co-purification with PRC1 in substochiometric amounts was also reported (Peterson et al., 2004). Scm and Sfmbt proteins contain mbt domains, through which they interact with one another (Kassis and Kennison, 2010). Scm, Sfmbt and Ph contain both Zn finger motifs and SPM-type SAM domains, leading to the physical interaction of Scm with these two protein

(Saurin et al., 2001; Peterson et al., 2004; Grimm et al., 2009). The polymerization capacity of Scm though its SAM domain and its interaction with other PcG proteins may play an important role in the propagation of PcG domain from target sites to the neighboring chromatin environment (Kang et al., 2015)

By using RNAi and mutations to knock down various PcG subunits, Wang et al. (2010) demonstrated that binding of Scm to *bxd* PRE is not dependent on the presence of Pho, PRC1 or PRC2. However, Scm depletion dislodges both PRC1 and PRC2, while it has no effect on the association of Pho with chromatin (Wang et al., 2010). A more recent study showed that Scm binding to *bxd* PRE was greatly reduced upon mutation of Pho binding sites in this PRE fragment, suggesting the dependence of Scm binding on the presence of Pho (Frey et al., 2016).

1.5.2 Histone marks and histone modifying enzymes

Histone modifications, deposited by histone modifying complexes, have been shown to regulate PcG recruitment. Among the histone modifications, TrxG deposited active histone marks can inhibit PRC2 activity and counteract PcG-mediated silencing (Klymenko and Müller, 2004; Papp and Müller, 2006). Yuan et al. (2010) reported that in HeLa cells, H3K27me3 does not co-exist with H3K36me2/me3 on the same histone H3 tail. Moreover, the activity of human, mouse and *Drosophila* PRC2 is inhibited by preinstalled H3K4 and H3K36 methylation in vitro, indicating a conserved mechanism for inhibiting PcG-mediated silencing (Paro et al., 2015).

Schmitges et al. (2011) found that a complex of PRC2 subunits, Nurf55-Su(z)12, binds the N-terminus of unmodified H3 and this binding is prevented by H3K4me3. In vitro assays showed that impaired nucleosome binding does not inhibit the PRC2 catalytic activity by H3K4me3-containing nucleosomes, but it is rather the reduced catalytic turnover that drives H3K4me3-mediated inhibition (Schmitges et al., 2011). These data suggest that PRC2 can act like a control module which inhibits deposition of H3K27me3 on genes with pre-existing active histone marks.

The Tudor domains of all three mammalian homologs of Pcl (PHF1, MTF2, and PHF19) bind H3K36me3 with high affinity. PHF19 recruits NO66, the H3K36 demethylase (Alhaj Abed and Jones, 2012; Brien et al., 2012). KDM2b, another H3K36me2/me3 demethylase, also colocalizes to PHF19-contaning PRC2 in ESCs (Alhaj Abed and Jones, 2012; Ballaré et al., 2012), suggesting that Pcl-family proteins facilitate the recruitment of PRC2 into newly repressed target genes to change the expression state of these genes.

E(z) is the only methyltransferase for H3K27, and catalyzes both H3K27me1/me2, in addition to H3K27me3 (Ferrari et al., 2014). However, the three forms of methylated H3K27 are mutually exclusive. H3K27me1 is enriched within the gene body of active genes and shows a positive correlation with H3K36me3. De et al. (2019) also reported that H3K36me3 domain inhibited the interactions between PREs and the neighboring chromatin as well as spreading of the PcG domain. H3K27me2 accumulates within intergenic and intragenic domains and serves as a protection against non-specific H3K27 acetylation (Ferrari et al., 2014). The recent findings suggest that PRC2 transiently binds to most of the genomic chromatin and samples the chromatin regardless of the DNA sequence, leading to the H3K27me1/me2. Additional factors, including non-coding RNAs (ncRNAs), transcription factors, transcriptional state, or histone modifications can then stabilize binding of PRC2 to chromatin where it can catalyze H3K27me3.

1.5.3 Hierarchical versus cooperative model

Using RNAi and mutations to knock down various PcG subunits, Wang et al. (2004) proposed a hierarchical model for the recruitment of PcG proteins at *Ubx* in larval wing imaginal discs. According to this model, Pho binds PREs and directly interacts with components of PRC2, E(z) and Esc, recruiting them to PREs. The E(z) subunit of PRC2 tri-methylates H3K27 in nucleosomes adjacent to PRE. H3K27me3 then serves as a docking site for PRC1 (Cao et al., 2002; Czermin et al., 2002; Fischle et al., 2003; Kuzmi- chev et al., 2002; Min et al., 2003). Pc

chromodomain binds to H3K27me3 (Muller et al., 2002; Cao and Zhang, 2004), helping to recruit PRC1, which compact arrays of nucleosomes and inhibit nucleosome remodeling in vitro (**Figure 1.2**). However, this early sequential model was based on the recruitment to a PRE of a single gene and not all other studies supported this model. Alternative mechanisms have been suggested for the establishment of PcG-mediated repression at other PREs. Genome-wide studies revealed that PRC1 and PRC2 do not always colocalize and PRC1-bound regions, devoid of H3K27me3, are present at a large subset of genomic regions (Schwartz et al., 2006; Schaaf et al., 2013; Loubie`re et al., 2016). Schaaf et al. (2013) found that PRC1 is present at most active genes bound by cohesion and facilitates phosphorylation of RNAPII to the elongating form, phosphorylated RNAPII at serine 2 (RNAPII S2p).

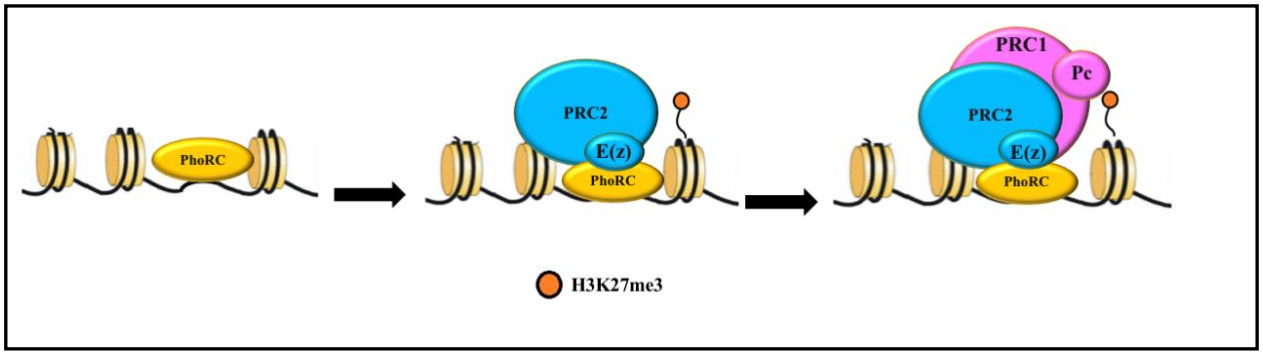

Figure 1.2- **Hierarchical binding of PcG proteins at PREs** (adopted from Wang et al., 2004).

Using *Drosophila* cultured cell lines lacking critical subunits of PRC1 or PRC2, Kahn et al. (2016) showed that at most PREs, PRC1 or dRAF was required for the recruitment of PRC2, while at some PREs, PRC1 and PRC2 are recruited independently of each other. However, their

data suggested that the presence of both PRC1 and PRC2 is required for effective repression at most PREs.

In another study, Kahn et al. (2014) showed that binding of Sfmbt and Pho to the majority of PREs is disrupted in cells lacking Psc/Su(Z)2 subunits of PRC1. Scm directly interacts with both PRC1 and PRC2, and is suggested to be a connecting mediator between these two complexes (Kang et al., 2015). Scm also interacts with Sfmbt and Ph through its SAM domain and mediates the recruitment of PRC1 independently of PRC2, stabilizing the binding of PhoRC through a positive feedback loop (Kang et al., 2015). Therefore, PRC2-independent recruitment of PRC1 might be mediated by Scm, which acts as a molecular bridge between PRC1 and PhoRC.

1.5.4 Involvement of ncRNAs in PcG recruitment

Multiple lines of evidence implicated the role of ncRNAs in targeting mammalian PcG complexes. Plath et al. (2003) and Zhao et al. (2008) demonstrated that a 1.6 kb ncRNA transcribed from the Xist locus, called RepA, recruited PRC2 to the X chromosome resulting in the X chromosome inactivation. Surprisingly, RepA and mammalian PRC2 were shown to be required for the initiation and spreading of X chromosome inactivation, but not for the maintenance of the silencing. In another study, Yap et al. (2010), demonstrated that Cbx7, one of the mammalian homologs of Pc, binds ANRIL ncRNA and recruits PRC1 to ink4a/ARF locus and silences Ink4a. Another evidence for the role of ncRNAs in the regulation of PcG target genes, was provided by HOTAIR ncRNA. HOTAIR is required for the transcriptional repression of HoxD cluster by recruiting mammalian PRC2 in trans to it (Rinn et al., 2007). There are few reports linking ncRNAs to PcG recruitment in *Drosophila*. Young et al. (2012) showed the enrichment of *Drosophila* long intergenic noncoding RNAs (lincRNA) loci with PcG-associated chromatin. This observation might argue for a role of ncRNAs in PcG recruitment in cis but could also be interpreted as preferential association of lincRNA loci with developmentally regulated genes.

1.5.5 Instructive versus responsive models for the recruitment of PcG proteins

PcG-mediated repression of target genes can be divided into two phases: initiation and maintenance phases. During the initiation phase, PcG proteins recognize the repressed state of their target gene and take over the repression from gene-specific transcription factors. Once the maintenance phase is established, PcG proteins maintain repression through an indefinite number of cell cycles. Within the topic of initiation of PcG-mediated repression, the mechanisms by which PcG proteins distinguish between repressed versus active states of target genes are not understood.

Instructive and responsive models have been proposed for the recruitment of PcG complexes to unmethylated CpG islands/*Drosophila* PREs (Klose et al., 2013). The instructive model suggests that biochemical interactions of sequence-specific transcription factors and/or lncRNAs with PcG complexes may target these complexes to their transcriptionally repressed target promoters (**Figure 1.3A**). At active gene promoters, transcription factors may directly recruit TrxG proteins to the promoters of their transcriptionally active target genes (Klose et al., 2013).

Few studies reported the direct recruitment of mammalian PRC1 by transcription factors, providing support to the instructive model (Dietrich et al., 2012; Yu et al., 2012). Genome-wide analysis of Pc binding in Rest mutant mESCs showed that Rest, a repressor of neuronal genes in non-neuronal cells, was required for PRC1 recruitment to a subset of PcG target neuronal genes (Dietrich et al., 2012). In another report, direct physical and functional interaction between Runx1/CBFβ and PRC1 has been observed (Yu et al., 2012). CBFβ is a Core-binding transcription factor with roles in stem cell self-renewal and tissue differentiation in mammalian cells and Runx1 is a subunit of CBFβ. Yu et al. (2012) provided evidence that Runx1 can directly recruit PRC1 to chromatin independently of PRC2. However, recruitment of PRC1 in the absence of PRC2 and H3K27me3 may indicate PcG-independent roles for PRC1. PRC1 has been shown to regulate the

transition of paused RNAP II to elongating form at active genes independently of other PcG proteins (Schaaf et al., 2013).

The results of at least one study support the instructive model in *Drosophila*. Hb, a repressive transcription factor of Hox genes, can recruit dMi-2, a subunit of a nucleosome remodeling and deacetylation complex (NuRD) (Kehle et al., 1998). Hb-mediated recruitment of NuRD complex to Hox genes, may result in local chromatin changes and nucleosome remodeling that allow binding of PcG proteins to the nucleosome template. Alternatively, the Hb-dMi-2 complex may directly interact with a PcG protein and recruit it to DNA (Kehle et al., 1998). However, a main challenge to the instructive model is to explain how PcG proteins can interact with the wide range of transcription factors or ncRNAs at different genes. On the other hand, the responsive model proposes that both PcG and TrxG sample the chromatin environment at PREs/CpG islands irrespective of their transcriptional state, which presumably would initially involve transient interactions. Potential stable targeting of PcG complexes would then be responsive to the transcriptional state of their target genes (Klose et al., 2013; **Figure 1.3B**). Conversely, in the presence of RNAPII and ongoing transcription at transcribed genes, accumulation of TrxG complexes, which antagonize the function of PcG complexes, would be favored. Therefore, the capacity of PcG proteins to sample CpG islands/PREs would permit them to respond to the transcriptional state of their target genes without a requirement for direct interactions with transcription factors or ncRNAs (Klose et al., 2013).

Riising et al. (2014) demonstrated that blocking transcription with two different inhibitors; 5,6-dichloro-1-beta-D-ribofurano-sylbenzimidazole riboside (DRB), which inhibits serine 2 phosphorylation of RNAPII by inhibiting the kinase activity of cyclin-dependent kinase 9 (CDK9) subunit of P-TEFb, and Triptolide, which induces proteasomal degradation of RNAPII by

inhibiting the ATPase activity of the XPB helicase subunit of TFIIH (Bensaude, 2011), resulted in the recruitment of Suz12, a PRC2 subunit, and deposition of H3K27me3 on the promoters of a subset of PcG-target genes but not on those of PcG-independent genes in mouse ESCs.

Another study consistent with the responsive model showed that constitutive transcription through the *Fab*-7 PRE, induced by an actin promoter, resulted in the activation of *Fab*-7 PRE and prevented establishment of PcG-mediated repression (Schmitt et al., 2005). This observation suggested that transcription through endogenous PREs, producing either sense or antisense RNA, must be active continuously to prevent the access and recruitment of PcG complexes to the chromatin. Several studies investigated if transcriptional activation could preclude PcG proteins from PREs. Kaneko et al. (2014) proposed that PRC2 binds nascent transcripts in vitro and this interaction in turn inhibits its histone methyltransferase activity. When transcription is silenced, and therefore no nascent transcript is produced, PRC2 is relieved from RNA-mediated inhibition, allowing the deposition of H3K27me2/3. Therefore, presence of PRC2 at actively transcribed promoters, lacking H3K27me2/3, could be explained as a consequence of RNA-mediated inhibition of PRC2 activity. An in vitro histone methyltransferase assay showed that RNA binding to PRC2 inhibited H3K27 methylation but had only a small effect on Ezh2 auto-methylation. This data suggested that RNA is not an active-site inhibitor of PRC2 activity, but rather inhibits its activity though other ways (Wang et al., 2017). Beltran et al. (2016) proposed that nascent RNAs and nucleosomes compete to bind to PRC2 and association of PRC2 with chromatin antagonizes its interaction with RNA. They further showed that release of PRC2 from chromatin and RNA, increases its binding to the RNA and nucleosomes, respectively.

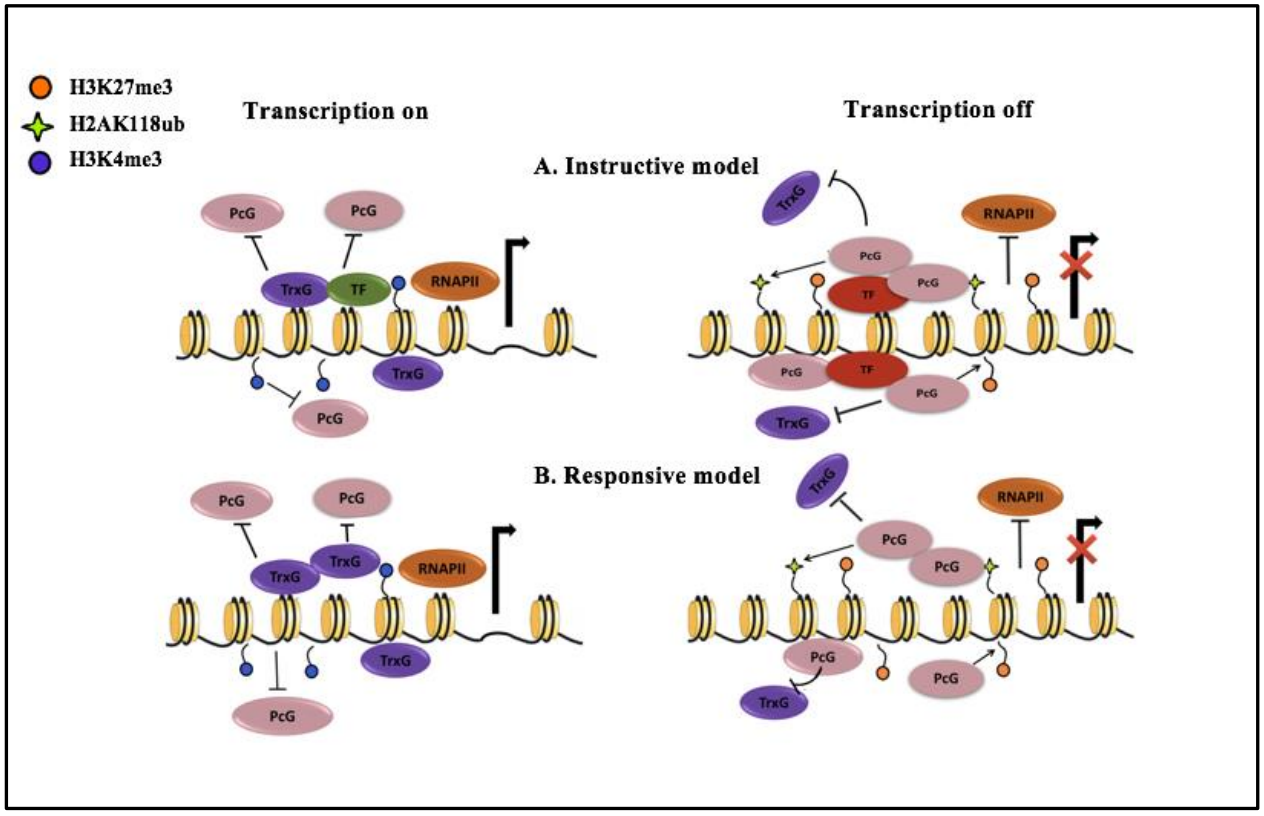

Figure 1.3- **Comparison of instructive and responsive models for recruitment of PcG complexes to PREs at target gene promoters.** (**A**) The instructive model, shows the interaction of sequence-specific transcription factors with TrxG complexes to recruit them to their target genes. At repressed gene promoters, TFs directly recruit PcG proteins. (**B**) The responsive model proposes that stable binding of PcG and TrxG to PREs, is dependent on the transcriptional state of the target gene.

1.6 Maintenance of PcG silencing through cell division

PcG-mediated repression can persist through numerous cell cycles and chromatin reorganizations that accompany cell cycle progression. A potential mechanism to maintain repression, and binding by PcG proteins, is the retention of the H3K27me3 mark (Hansen et al., 2008). Maintenance of H3K27me3 on chromatin during S phase has been shown by immunofluorescence, ChIP, and proximity ligation studies (Hansen et al., 2008; Lanzuolo et al., 2011). Retention of at least some H3K27me3 is in agreement with the retention of E(z) (Petruk et

al., 2012). Since PRC2 can bind the H3K27me3 mark and stimulate further H3K27 methylation (Margueron et al., 2009), even low levels of H3K27me3 remained following each replication can trigger efficient K27me3 of surrounding nucleosomes. However, Petruk et al. (2012) reported that nucleosomes behind the replication fork lack significant amount of H3K27me3 histone mark, suggesting that parental histone methyl transferases, and not the methylated histones, remain associated with targets through DNA replication. It has been shown that quantitatively less PcG proteins remain bound to chromatin during mitosis than in interphase (Follmer et al., 2012), therefore, it is likely that both PRC1 and PRC2 remain associated with newly replicated daughter strands during S phase of interphase. Components of both complexes, Pc from PRC1 or E(z) from PRC2, are physically near the replication fork (Petruk et al., 2012) but there is no evidence about the physical contacts between PcG components and replication fork features (Probst et al., 2009). ChIP studies on homogenous population of cells showed retention of PcG subunits on replicating DNA (Francis et al., 2009; Lanzuolo et al., 2011). The maintenance of PRC1 on the template after fork passage has been shown for mammalian SV40 replication system (Francis et al., 2009) and even after replication of naked DNA by the bacteriophage T7 system (Lengsfeld et al., 2012). Psc, subunit of PRC1, is able to maintain stable Psc-Psc association with the template DNA as the fork passes (Lo et al., 2012).

H3K27me3 is proposed to propagate during cell replication through two mechanisms: first, the parental H3K27me3 nucleosomes are locally re-deposited during replication. Second, these modified parental nucleosomes can be served as templates for PRC2 to copy the H3K27me3 mark into newly incorporated nucleosomes (Margueron et al., 2009; Jiao and Liu, 2015; Justin et al., 2016).

Muller et al. (2017) provided evidence that PREs are needed for both generation and propagation of H3K27me3 in daughter strand chromatin. PRC2 is first recruited to PREs and then binds to parental H3K27me3 nucleosomes in flanking chromatin after replication (Muller et al., 2017). Studies suggesting that parental PRC components anchored at PREs, are responsible for reestablishing the mark, predict that removal of a PRE would result in the loss of silencing and H3K27me3 after the first cell cycle, a prediction that is not supported by the findings of other studies. Coleman and Struhl (2017) showed that upon excision of a PRE, the local availability of PRC2 reduces and this in turn allows some of the newly incorporated nucleosomes escape being modified, resulting in the serial dilution of H3K27me3 histone mark through cell divisions. The same authors further reported a ~10-12% decrease in H3K27me3 levels following each replication cycle, while a 50% reduction was expected if PRC2 first bound the PRE and copied the mark. However, given that the free PRC2, not anchored at PRE(s), rapidly exchange with chromatin-bound PRC2 in vivo, the possible contribution of free PRC2 in binding and propagation of H3K27me3 can explain the latter results.

1.7 PcG proteins and cancer

PcG proteins are involved in the development of tumors through misspecification of cells towards a stem cell state, which will be the main driving force behind tumor proliferation, and transcriptional repression of tumor-suppressor genes (Pardal et al., 2000). The ability of stem cells to evade proliferative restriction allows the accumulation of mutations and epigenetic changes, leading to the malignancy. Studies have shown that PcG proteins are involved in inhibition of differentiation and promotion of stem cell self-renewal and thus play an important role in tumor progression and development.

PRC1 and PRC2 genes are aberrantly expressed in a broad range of human cancers (Sparmann and van Lohuizen, 2006; Rajasekhar and Begemann, 2007). Activity of Bmi1, a

homolog of *Drosophila* Psc, is crucial for the maintenance of normal and cancer stem cells and its overexpression was reported in several human cancers. Elevated levels of Bmi1 protein were found in squamous cell carcinomas (He et al., 2009), neuroblastomas (Nowak et al., 2006), and bladder tumors (Shafaroudi et al., 2008). Furthermore, deregulation of Bmi1 was reported in leukemia (Lessard and Sauvageau, 2003). Bmi1 was also reported to play a pivotal role in lung carcinogenesis and loss of Bmi-1 reduced the self-renewal ability and thus limited progression of lung cancer stem cells (Dovey et al., 2008).

Ezh2, Suz12 and Pcl3, subunits of mammalian PRC2, are also overexpressed in a variety of different human cancers (Rajasekhar and Begemann, 2007; Sparmann and van Lohuizen, 2006). The *Ezh2* gene is amplified in many human prostate cancer cell lines, and indicates aggressiveness in lymphoma, melanoma, bladder tumors, prostate and breast cancers (Varambally et al., 2002; Kleer et al., 2003; Sarama ki et al., 2006). Zingg et al. (2014) reported that a tumor suppressor, Adenosylmethionine decarboxylase 1 (AMD1), is among Ezh2 targets. It was further shown that Ezh2 function is essential for the initiation of melanoma and its inactivation prevents metastatic spread of melanoma (Zingg et al., 2014). Mimori et al. (2005) and Kidani et al. (2009) provided evidence that the emergence of colorectal cancer and oral squamous carcinomas was linked to the deregulation of *Ezh2* gene.

Bmi1 and Ezh2 are both involved but play different roles in the progression of glioblastomas. Although downregulation of Ezh2 protein decreased the self-renewal and the tumor-initiating capacity of cancer stem cells (Suva et al., 2009), knockdown of Bmi1 levels in glioblastomas reduced the proliferation of the glioma stem cells (Godlewski et al., 2008).

1.8 PcG and differentiation

Hox genes, the most prominent PcG targets in *Drosophila*, are organized in two complexes: The Antennapedia complex (ANT-C) and BX-C. ANT-C comprises five Hox genes (*lab*, *pb*, *Dfd*,

Scr, and *Antp*), involved in the development of head and the anterior thorax (Kaufman et al., 1990). BX-C contains three Hox genes (*Ubx*, *abd-A*, and *Abd-B*), which specify parts of the posterior thorax and the abdomen (Duncan, 1987; Lewis et al., 2003). PcG genes maintain repression of *Drosophila* Hox genes in the patterns initially established by the products of the segmentation genes (Pirrotta, 1997a). Mutations in the PcG genes lead to derepression of Hox gene expression and homeotic transformations (Lewis, 1978).

Additional PcG targets are genes encoding key developmental regulators and signaling proteins in ESCs. Knockout of PRC2 and PRC1 genes leads to the loss of H3K27me3 and H2AK119ub1 marks, expression of developmental genes and differentiation defects in ESCs (Chamberlain et al., 2008; Shen et al., 2008). ESCs lacking Ezh2, Eed and Suz12 are not able to differentiate due to the inability of these cells to repress pluripotent genes during differentiation (Pasini et al., 2007; Chamberlain et al., 2008). Jarid2, a substochiometric subunit of PRC2, is required for the differentiation of ESCs and proper embryo development (Pasini et al., 2010). In accordance with the latter study, Landeira et al. (2010) reported that while presence of Jarid2 is not necessary for PRC2-mediated silencing, Jarid2 knockout ESCs are not able to properly differentiate (Landeira et al., 2010).

1.9 Embryogenesis in *Drosophila*

Following fertilization, *Drosophila* nuclei synchronously divide, while cytokinesis does not occur (**Figure 1.4**). The latter results in the formation of syncytium in which nuclei are not separated by cell membranes. This allows morphogen gradients to play a key role in the patterning of early *Drosophila* embryos. Initiation of first nuclear cycle (nc1) happens after fusion of the male and female pronuclei. Through nc 1–3, nuclei divide at the anterior of the embryo (Gilbert, 2000). During nc 4–6, nuclei spread throughout the whole embryo along the anterior–posterior axis. The first 8 nuclear divisions take place rapidly, averaging 8 minutes for each nuclear cycle (Gilbert,

2000). At nc9, pole cells form at the posterior end of the embryo, which later give rise to the germ cells in the adults (Gilbert, 2000). After the ninth nuclear cycle, the nuclei move to the periphery of the embryo, forming the syncytial blastoderm which lasts untill nc13 (Zolakar and Erk, 1976; Foe et al., 1993). Through nuclear divisions 10-13 duration of interphase increases slightly with each cycle (Farrell and O'Farrell, 2014). Nonetheless, the overall rate of nuclear division is still rapid through these stages. The first prolonged interphase happens at nc14, mainly due to the lengthening of S phase and introduction of G2 at this stage embryos (Farrell and O'Farrell, 2014). During nc14, cell membranes partition the 6000 or so nuclei, forming the cellular blastoderm (**Figure 1.4**; Zolakar and Erk, 1976). This coincides with the midblastula transition (MBT), in which widespread zygotic gene expression starts (Farrell and O'Farrell, 2014).

Gastrulation starts just after completion of cellularization and includes coordinated movements of cells that give rise to endoderm, mesoderm and ectoderm layers. These layers serve as the tissue primordia for the larval organs. Following this stage, further development forms an embryo with distinct morphological segments.

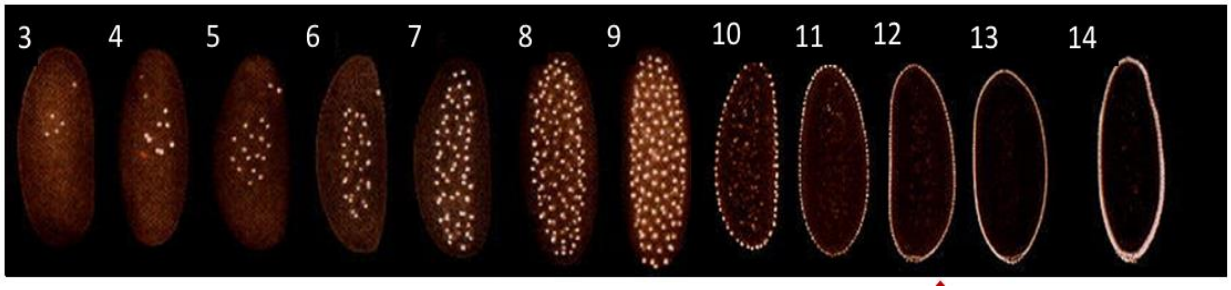

Figure 1-4- **Nuclear divisions during early *Drosophila* embryogenesis.** During divisions 3-9, nuclei divide and spread throughout embryo. Nuclei migrate to the periphery of the embryo during divisions 10–13 (Syncytial blastoderm). Cellularization happens after nc13 (Kotadia et al., 2010).

1.10 The origins of anterior-posterior polarity in *Drosophila*

1.10.1 Maternal effect genes

The anterior-posterior axis of *Drosophila* is established by localized maternally expressed mRNAs. patterned before the formation and function of nuclei. The *bicoid* (*bcd*) and *nanos* (*nos*) mRNAs are synthesized in nurse cells, loaded into the oocyte and localized in the anterior and posterior portion of the unfertilized egg, respectively. *hunchback* (*hb*) and *cad* mRNAs are distributed uniformly throughout the oocyte. Upon fertilization, translation of localized *bcd* and *nos* mRNA produces opposing gradients of their protein products; *bcd* highest at the anterior and *nos* highest at the posterior end of the embryo. Bcd protein inhibits the translation of the *cad* mRNA, resulting in the posterior localization of Cad protein (Dubnau and Struhl, 1996; Rivera-Pomar et al., 1996). Nos protein binds to maternal *hb* mRNA, and prevents its translation in the posterior portion of the embryo (Tautz, 1988; Wharton and Struhl, 1991).

1.10.2 Zygotically expressed genes

Maternally-expressed proteins can activate a hierarchy of zygotic genes, starting with gap genes. Gap genes are a group of genes which were initially defined by mutant phenotypes in which multiple adjacent body segments are missing, resembling a gap in the normal body (Petschek et al., 1987). The examples of gap genes are *hb*, *Kruppel* (*Kr*), *knirps* (*kni*), *gt*, *tll* and *huckebein* (*hkb*) (Campos-Ortega and Hartenstein, 1985). These genes are expressed in single or double domains along the anterior-posterior axis, and decay around mid-embryogenesis.

Gap genes encode transcription factors which control the expression of other gap genes and pair-rule genes. Pair-rule genes divide the embryo into striped pattern of seven parasegments and encode transcription factors regulating the expression of other pair-rule genes, segment polarity genes and hometic genes (Gilbert, 2000). Segment polarity genes define the anterior and posterior polarities within each parasegment, resulting in 14 stripes (Sanders et al., 2019).

1.11 *giant*, our PcG-target gene

gt is a zygotic gap gene that affects the development of the head and abdominal regions in *Drosophila.* Identification of *gt* as a PcG target gene was based on the isolation of dominant *E(z)* allele that suppress the *nos* maternal effect (Pelegri and Lehmann, 1994). Later, Negre et al. (2006) also showed the derepression of *gt* in PcG mutant background.

gt is initially transcribed in two broad anterior and posterior stripes at nc12. During cellular blastoderm stage, the anterior domain splits into two stripes, and finally, a fourth stripe develops at the anterior terminus (**Figure 1.5A**).

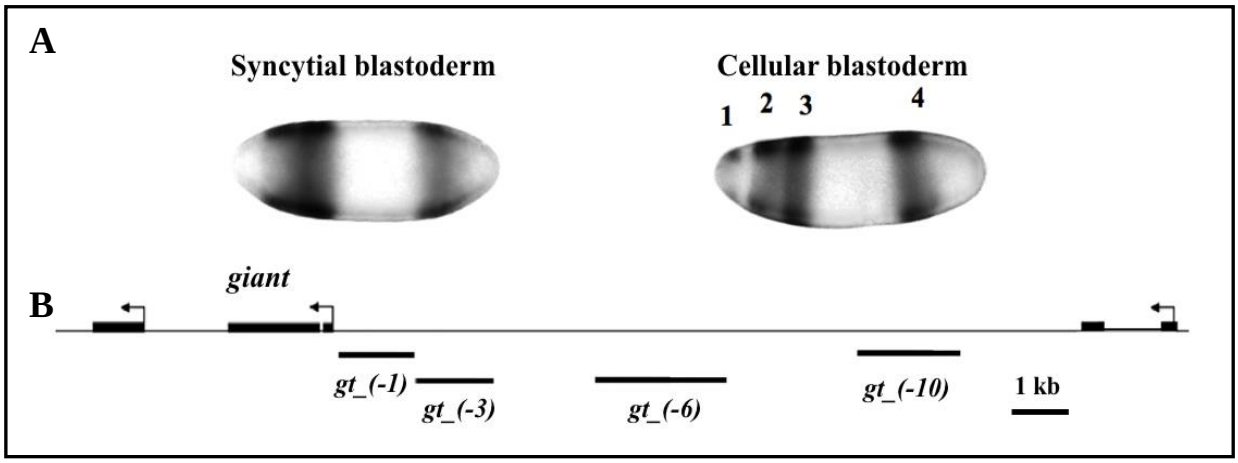

Figure 1.5- **Pattern of *gt* expression and schematic map of *gt* upstream regulatory region**. (**A**) In situ hybridization of transgenic syncytial and cellular blastoderm embryos with *lac*Z probe. Lateral views of embryos are shown, anterior to the left, dorsal up. (**B)** Upstream regulatory region showing locations of *gt* enhancers; *gt*_(*-1*), *gt*_(*-3*), *gt*_(*-6*) and *gt*_(*-10*).

Expression of gt is regulated by four characterized enhancers located within the *gt* upstream regulatory region. *gt_(-1),* (-163 to -1402) which is bifunctional and produces both the anterior and posterior expression domains. *gt_(-3)* (-1410 to -2596) and *gt_(-10)* (-8904 to -10649)

produce the posterior and the anterior tip expression, respectively. *gt_(-6)* (-4438 to -6620) produces the most anterior expression stripe (**Figure 1.5B**; Berman et al., 2002; Schroeder et al., 2004). The role played by the apparent redundancy of *gt_(-1)* with *gt_(-10)* and *gt_(-3)* in regulation of *gt* is not clear (Berman et al., 2002).

Hb and Bcd play important roles in regulation of *gt* expression. The expression of the anterior stripe is activated by Bcd, while Cad is the activator of the posterior stripe. Gt regulates its own expression, by binding and auto-activating *gt_(-1)* enhancer (Hoermann et al., 2016). Kr and Hb restrict the anterior and posterior boundaries of *gt* posterior stripe, respectively (Mohler et al., 1989; Eldon and Pirrotta, 1991). The terminal system, which acts through Tsl, activates the zygotic gap genes *tll* and *hkb* in the terminal regions (Weigel et al., 1990). Eldon and Pirrotta (1991) showed that Tsl has contrasting effects on the expression of *gt*, showing a repressive effect posteriorly and an activating effect anteriorly. Low concentrations of Bcd activates *Kr* (Hülskamp and Tautz, 1991), while both Bcd and Cad are activators of *kni* (Schulz and Tautz, 1995). High concentrations of Hb represses expression of *kni*, *Kr* and *gt*, whereas lower concentrations allow *Kr* activity but still prevent *kni* and *gt* expression (Struhl et al., 1992). Kni negatively regulates *Kr* expression (Jaeger et al., 2004). Gt has been demonstrated to repress *kni* and *Kr* (Kraut and Levine, 1991; Eldon and Pirotta, 1991; Small et al., 1992).

Among all the PcG-target genes in *Drosophila*, the *gt* gene is selected as our model gene because regulation of *gt* by its specific activating and repressive transcription factors is well characterized. This permits construction of genetic systems in which *gt* is uniformly repressed by PcG proteins after the decay of its repressor at cellular blastoderm-stage embryos, or is ubiquitously expressed by knocking down its inhibitor and ubiquitous expression of its activator.

hb is expressed both maternally and zygotically. Maternal Hb protein forms an anterior to posterior slope in the fertilized embryo and degrades at nc14 (Tautz, 1988; Wharton and Struhl, 1991). Maternal Hb is substituted by zygotic Hb at blastoderm stage. Proper posterior localization of *nos* mRNA in the oocyte requires the activity of Oskar (Osk) (Ephrussi et al., 1991). Bcd and Tll are activators of zygotic *hb* at anterior and posterior ends, respectively. *bcd osk tsl* females produce embryos in which maternal *hb* is ubiquitously expressed as the result of mutant allele of *osk*. Zygotic *hb* is not activated as the result of mutant alleles of *bcd* and *tsl*. In these embryos, maternal Hb is degraded at cellular blastoderm stage and not replaced by zygotic Hb, yet *gt* is not expressed. Maternal *cad* mRNA is ubiquitously translated in embryos derived from *bcd osk tsl* females due to the absence of Bcd at syncytial blastoderm. However, it fails to activate expression of *gt* due to the presence of maternal Hb and its repressive effects. Terminal system, which is absent from *bcd osk tsl* embryos, is required for the activation of zygotic expression of *cad* (Mlodzik and Gehring, 1987). Therefore, in the embryos from *bcd osk tsl* females both maternal Hb and Cad are distributed ubiquitously, while neither of them is zygotically expressed. In the absence of zygotic Hb, maintenance of *gt* repression requires PcG activity.

Embryos produced by *bcd osk tsl* females, offer an in vivo model of a homogenous population of cells in which repression of a target gene is transitioning from gene-specific transcription factors to PcG proteins. Using the embryos derived from *bcd osk tsl* females, the molecular mechanisms and sequence of events led to the establishment of PcG-mediated repression was studied (Alhaj abed et al., 2018).

1.12 Transcription: Initiation and elongation

Eukaryotic transcription starts upon binding of an activator to the enhancer DNA sequences which promotes the recruitment of general transcription factors and RNAPII to the target gene

promoters (Thomas and Chiang, 2006). The latter results in the formation of a functional pre-initiation complex (PIC) and initiation of transcription.

The core promoter elements (CPEs) are located either upstream or downstream of TSS, or within the coding region (Juven-Gershon et al., 2008). Major CPEs include the TATA box, initiator (Inr) and downstream promoter element (DPE) (Juven-Gershon et al., 2008). The TATAA sequence is recognized by TATA binding protein (TBP) and several TBP-associated factors (TAFs), forming the TFIID complex. TFIIA interacts with TBP subunit of TFIID and stabilizes the interaction between TBP and TATA element.

The sequences immediately flanking the TATA box can contain the TFIIB recognition elements (BRE). TFIIB is involved in the formation of a stable complex with TBP-bound DNA and positioning RNAPII (Deng and Roberts, 2007). RNAPII binds to the core promoter along with TFIIF. TFIIH and TFIIE are the next general transcription factors recruited to the PIC. TFIIH is composed of ten subunits that regulate the formation of ATP-dependent open form of PIC, required for productive transcription initiation (Laine and Egly, 2006). The majority of eukaryotic genes have promoters that lack a canonical TATA sequence. In genes lacking canonical TATA sequences, Inr and DPE can be recognized by TFIID and efficiently form a functional PIC (Baumann et al., 2010).

Deposition of specific modifications on the C-terminal domain (CTD) of RNAPII and N-terminal histone tails of the nucleosomes is an important step for the initiation of transcription. Acetylation of histone H3 and H4 at lysines generates an open chromatin state competent for transcription (Ansari et al., 2009). In contrast, tri-methylation of H3K9 by methyltransferases leads to the formation of a repressive chromatin state (Li et al., 2007; Hager et al., 2009).

RNAPII loaded onto an active promoter is mostly in a hypo-phosphorylated state. Serines at positions 2 and 5 of RNAPII-CTD are phosphorylated at specific stages of transcription (Gomes et al., 2006). The kinase activity of the cyclin-dependent kinase 7 (CDK7) subunit of TFIIH is responsible for the phosphorylation of Serine 5 residue of the RNAPII (RNAPII S5p) (Ohkuma and Roeder, 1994), while CDK9 subunit of positive transcription elongation factor-b (P-TEFb) phosphorylates CTD at Serine 2 (RNAPII S2p).

Progression of RNAPII towards the 3′ of the gene is accompanied by a reduction in the phosphorylation status of RNAPII S5 and an increase in the phosphorylation of RNAPII S2. AS CDK9 phosphorylates RNAPII S2, phosphatases SSU72 and Rtr1, or its human homolog RPAP2, dephosphorylate RNAPII S5 along with transcription elongation (Reyes-Reyes and Hampse, 2007).

1.13 Rationale

gt has two PREs, a proximal PRE1 and a distal PRE2 (Abed et al. 2013). We have previously described a temporal hierarchy of PcG protein de novo recruitment at the *gt* locus in syncytial blastoderm through cellular blastoderm/gastrulating embryos (Abed et al. 2018). PhoRC binding to *gt* PRE1 precedes its stable binding to PRE2 (Abed et al. 2018). Both Pho and Sfmbt are weakly present at both PREs in syncytial blastoderm stage embryos, but do not stably bind to PRE2 until embryos reach cellular blastoderm/gastrulation, approximately 30 minutes later than their stable binding to PRE1. This observation raised the possibility that PhoRC binding to PRE2 may be at least partially dependent on the presence of PhoRC at PRE1. In order to determine whether PhoRC must first bind to PRE1 to assist PhoRC binding to PRE2, and to examine recruitment of other PcG proteins to both PREs in the absence of Pho at PRE1, we compared the recruitment of PcG proteins in reporter transgenes that contain either the entire wild type *gt* regulatory region or the same *gt* fragment with mutant Pho binding sites in PRE1. This study shed

light on the operation of relatively uncharacterized *gt* PREs and presented an in-depth analysis of the functions of individual PRE binding proteins at these PREs.

A major gap in our understanding of PcG proteins, is the molecular mechanisms which lead to the initiation of PcG mediated-silencing. Within the topic of initiation, the mechanisms by which PcG proteins discern the transcriptionally repressed and active states of their target genes are not understood. Instructive and responsive models have been proposed for the initial recruitment of PcG complexes to PREs/CpG islands (Klose et al., 2013). Despite the few studies which supported either models in mammalian cells (Kehle et al., 1998; Dietrich et al., 2012; Yu et al., 2012; Riising et al., 2014; Beltran et al., 2016), the mechanisms triggering the initial recruitment of PcG proteins have been left unraveled. By investigating the recruitment of PcG proteins at the transcriptionally inert *gt* transgene in a background in which endogenous *gt* is transcriptionally active, we experimentally tested these two models for PcG targeting in vivo and answered a very important question on the mechanisms by which PcG proteins discern the transcriptionally repressed and active states of their target gene. The results of our study will be relevant to stem cell biology and oncogenic effects resulting from misexpression of PcG proteins.

CHAPTER 2:

MATERIALS AND METHODS

2.1 Generation of transgenic lines

2.1.1 Construction of SD10 plasmids

The SD10 P element reporter vector was converted into a Gateway Destination Vector by ligating C.1 cassette (Invitrogen), containing the Chloramphenicol resistance gene (CmR) and the *ccdB* gene flanked by attR1 and attR2 sites, into the SD10 *Sph*I site. SD10-PRE1.1 was constructed by PCR amplifying a 555 bp fragment extending from 543 upstream to 12 bp downstream of the *gt* TSS from the CH322-05H16 BAC clone (BACPAC resources: http://www.pacmanfly.org/) using primers with attB sites and ligating the PCR product into the pENTR1A vector using BP Clonase II enzyme mix (Invitrogen). The *gt* PRE1.1 fragment was then recombined into the SD10 destination vector using LR Clonase II enzyme mix (Invitrogen). Pho consensus binding sites within PRE1.1 fragment in pENTR were sequentially mutated using Phusion II Site-Directed Mutagenesis kit (Thermo Scientific). The mutations were confirmed by sequencing, and the mutant and wild type (wt) fragments were then digested with *Nsp*I and ligated to the *Sph*I site of SD10 P element reporter vector, located between *en* enhancer and promoter. P element constructs were injected into embryos by BestGene (Chino Hills, CA). Five homozygous lines were obtained for SD10-PRE1.1-wt. A single line was obtained for SD10-PRE1.1-mut. Eight additional SD10-PRE1.1-mut lines were generated by remobilizing the original transgene using the genomic Δ2-3 P element transposase source (Robertson et al., 1988).

2.1.2 Construction of Pelican plasmids

The Pelican P element reporter vector (Barolo et al., 2000) was converted into a Gateway Destination vector by ligating Gateway Cassette C.1 into the Pelican *Stu*I site. The attB sequences, required for phage φC31 integrase-mediated recombination, were amplified from pUASP-attB vector (Drosophila Genome Resource Center, DGRC) and then inserted into Pelican *Avr*II site, between the gypsy insulator and the left P sequence. A 10,781 bp *gt* genomic fragment that extends from 10,421 bp upstream to 360 bp downstream of *gt* TSS was first assembled from fragments from overlapping BAC clones. A *Nru*I-*Hind*III fragment from the CH322-101F02 BAC clone was first ligated to a *Hind*III-*Not*I fragment from the CH322-05H16 BAC clone within the pBS-KS vector. A *BsrG*I fragment that included the *gt* fragment was blunt-ended with Klenow and then inserted into the *Dra*I and *EcoR*V sites of the pENTR1A vector. In order to be able to distinguish ChIP signals from the *gt* fragment and endogenous *gt* gene, the PCR priming sites for regions 4, 6, and 9 were mutated by sequential rounds of site directed mutagenesis using the Phusion II Site-Directed Mutagenesis Kit (Thermo Scientific). The *gt* fragment that includes only these primer site mutations will be referred as wild type (wt) in subsequent descriptions.

2.1.2.1 Pelican-*gt*-wt/mut

The same Pho sites that were mutated in the smaller PRE1.1 fragment, were mutated in the larger ~10.8 kb *gt* fragment. The entire inserts of both the wt and Pho-Phol mutant pENTR1A plasmids were sequenced to confirm that no unintended mutation had been introduced prior to recombining them into the Pelican-attB vector. Both Pelican-*gt*-wt and Pelican-*gt*-mut constructs were integrated into the attP40 landing site on the second chromosome through the φC31 integrase and transgenic flies were generated by BestGene (Chino Hills, CA).

2.1.2.2 Pelican-*gt*-promoter mut

To construct transgenic flies containing a transcriptionally inert *gt* gene, TATA box, Initiator region (Inr) and downstream promoter region (DPE) of 10,781 bp *gt* fragment in pENTR1A-*gt* were consecutively mutated using Phusion II Site-Directed Mutagenesis Kit (Thermo Scientific) and primers listed in **Table 2.1**. Both wt and the transcriptionally inert *gt* fragment, were then transferred into the modified version of Pelican destination vector, as previously described, through LR Clonase II. The resulting expression vectors were sent to BestGene, Inc. for embryo injections. The constructs were integrated into the attP40 landing site on the second chromosome through the φC31 integrase and Pelican-*gt*-promoter mut (Pelican-*gt*-pm) were generated.

Table 2.1- **Primers used to mutate DPE, TATA box and Inr sites and their corresponding annealing temperatures.**

Amplicon	**Primer**	**Primer Sequence (5'-3')**	**Annealing temp (C∘)**
DPE Mutagenic	Forward	CGCAGGTCAGTCTCCTCTCGTCGGA	72
	Reverse	TTCGACATCGTCAGCGTGACTCATGTGCAACA TCAGAGCTAGATCACCAG	
TATA Mutagenic	Forward	TCGGAACCCCCTCGACCCGCGCCGAAAATCCG CAGGCTGCG	65
	Reverse	CGAGAGGAAATCAGTTTGCG	
Inr Mutagenic	Forward	AACCCCCTGCATTTATACCG	66
	Reverse	CCGACGAGAGGAAACTGACCTGCGTTCGACAT CGTCAGCG	

2.2 ChIP experiments

2.2.1 Embryo collection for ChIP assays

About 1,800-2,000, 4,000-5,000, and 6,000-8,000 females were added to small, medium and large size cages, respectively. Embryos were collected from cages using apple or grape juice agar plates supplemented with yeast paste. To minimize the number of older embryos, a 2-hour pre-lay collection was done before any collection. Agar plates were placed, usually upside down, in the cage for exactly 30 or 60 minutes at 25°C and then removed and aged at 25°C for the time After Egg Lay (AEL) to target the appropriate developmental stages (**Table 2.2**). The embryos were washed off the plates, followed by dechorionation with 50% bleach for 2 minutes. Embryos were then washed thoroughly with deionized water and PBST (0.5 % Triton X-100, 1X PBS).

Table 2.2- **Developmental stages and the corresponding mass amount in ChIP assays.**

Embryo Stage	**Time AEL (min)**	**Mass (mg)**
nc10-13	80-140	30
nc13	110-140	20
nc14a	140-170	10
nc14b	170-200	10
nc14b-nc15	170-230	10
Stage 10	260-320	10

The embryos were then fixed as previously described (Ahaj Abed et al., 2018; Blythe and Wieschaus, 2015). Embryos were crosslinked using 108 µl of formaldehyde (Fisher #BP531) in 2

ml PBST and 6 ml heptane. After 15 minutes, 0.125 M glycine (pH=7) in PBST was added to quench the f ixation for 3 minutes. Embryos were then washed with ice-cold PBST at least three times, prior to adding 1X protease inhibitor (Sigma #P8340). The embryos were kept on ice at 4°C until they were hand-sorted.

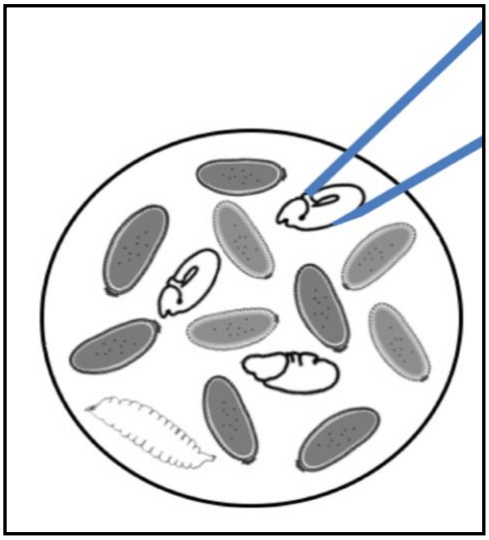

Figure 2.1- **Elimination of older *Drosophila* embryos by hand sorting based on the morphology** (adopted from Chen et al., 2013).

Since *Drosophila* females often retain fertilized eggs, crosslinked embryos were manually sorted using a Zeiss Primovert microscope to eliminate older embryos, based on their morphology. Older embryos would significantly affect ChIP signals, due to greater amount of chromatin. The percentage of contaminating older embryos is dependent on the genotype of the embryos, ranging from 10-50% of the collected embryos. Carefully staged embryos were then weighed, flash frozen in liquid nitrogen and stored at 80°C. In order to approximately equalize the chromatin amount per

ChIP using embryos at different developmental stages, we used different masses of embryos for each embryonic stage per tube (**Table 2.2**). Each embryo aliquot has enough chromatin for nine ChIP samples and two aliquots of input genomic DNA.

2.2.2 ChIP Protocol

For each ChIP sample, 50 µl of Protein A magnetic beads (Pierce #88846) was added, left at magnetic stand for 2-3 minutes, and then washed with 500 µl PBST-3% BSA. The beads were then blocked with 1 ml of PBST (0.5 % Triton X-100, 1X PBS)-3% BSA for 1 hour.

Dilutions of antibodies were prepared in 250 µl volume of PBST-3% BSA and then were added to the magnetic beads while rotating at room temperature for 20 minutes, then at 4°C for 2.5 hours. During this time, flash frozen embryos, while still not completely thawed, were homogenized in 100 µl of RIPA buffer supplemented with 1mM DTT, 1X protease inhibitor (Sigma-Aldrich #P8340). Once embryos were homogenized, 600 µl of RIPA was added to get the final volume to 700 µl. Homogenized embryos were then centrifuged in a refrigerated microfuge at full speed at 4°C for 5 minutes. Supernatant was then discarded and the pellet was resuspended first in 100 µl of RIPA buffer and then sufficient amount of RIPA was added to obtain the final concentration of 10 mg of embryos/ml for nc14b, nc14b-nc15 and stage 10, 20 mg/ml for nc13 and 30 mg/ml for nc10-13. The chromatin was then sonicated on ice using a microtip probe and Misonix sonicator 3000: 30% power, 15 sec pulses, 45 sec pauses, total of 4.5 minutes. After each 3 pulses series, the sonicated chromatin was gently mixed, and spun down quickly to collect the liquid. Ice bath was regularly renewed to make sure that the tube did not get warm during sonication. This sonication should give a size of $\leq$ 1 kb for 90% of the resulting DNA fragments with a 30-35% peak at 1 kb.

The sonicate was then centrifuged at high speed, and supernatant was transferred to a new tube and then pre-adsorbed with 80 µl Protein A agarose/Salmon Sperm DNA (Millipore #16-157) at 4°C for 1 hour. After 2.5 hours of incubation of Ab-beads and 1 hour of sonicate pre-clearing, tubes containing Ab-beads were spun down, placed on the magnetic stand and supernatant was removed. The mixture of chromatin and Protein A agarose beads was also centrifuged and 100 µl aliquots of supernatant were added to each of magnetic bead containing tubes. 500 µl RIPA buffer was added to each sample and the samples were then incubated on a rotating wheel at 4°C overnight. A 20 µl aliquot of sonicated chromatin was used as input DNA, and 0.5 µl of 10 mg/ml RNAse A was added to it and incubated for 30 minutes at 37°C. 77.5 µl of elution buffer (50 mM Tris-HCl pH 7.5, 10 mM EDTA, 1% SDS, 300 mM NaCl) and 2.5 µl of 10 mg/ml Proteinase K (final concentration of 0.25 mg/ml) was added to each input tube, to bring the volume to 100 µl. Input tubes were then incubated at 65°C overnight. After overnight incubation, ChIP samples were centrifuged and then washed on a magnetic stand with a series of the following buffers for 2 minutes each: 3x low salt wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0, 150 mM NaCl), 1x high salt wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0, 500 mM NaCl), 2x with LiCl wash buffer (0.25 M LiCl, 1% NP-40, 1% SDC, 1 mM EDTA, 10 mM Tris-HCl pH 8.0), and 1x with TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA).

After the last wash with TE, beads were resuspended with 100 µl elution buffer and incubated at 65°C for 15 minutes. Tubes were vortexed every 5 minutes, followed by brief spinning down. After 15 minutes, beads were collected on the magnetic stand, and supernatant was transferred to new tubes, supplemented with 95 µl TE and 5 µl of 10 mg/ml Proteinase K, and incubated at 65°C for 4 hours to reverse crosslinking. 100 µl of TE was added to the input tubes,

and input DNA was left at room temperature during the 4-hour incubation. 10 µl of Sodium acetate (3M) was added to each of the ChIP samples and input DNA, prior to extraction with equal amounts of phenol/chloroform and chloroform. 40 µg of glycogen (2 µl of 20 mg/ml glycogen) was added to each tube and ethanol precipitation, using 2X volume ethanol, was performed at -20°C overnight. All samples were centrifuged for at least 15 minutes at full speed and the pellets were washed once with 80% ethanol, SpeedVac-dried for 10 minutes and resuspended in 33 µl of water and 50 µl (1.5 X) volume of Agencourt Ampure beads (Beckman Coulter #A63880). Thoroughly-mixed samples were incubated for 15 minutes at room temperature, followed by leaving on magnetic stand for another 15 minutes. Supernatant was removed from the tubes, and the beads were washed with 200 µl of 80% ethanol twice, 30 seconds each time. Beads were then left on the magnetic stand to dry for 15 minutes. Dried beads were then resuspended in appropriate amount of PCR water, 15 µl for each primer set, and incubated for another 5 minutes before transferring the supernatant to the new tubes. An additional equal volume of water was added to the input DNA to make the 10% input sample as the qPCR standard.

2.2.3 Antibodies used in ChIP assays and the corresponding volumes

Pc, 10 µl (Wang et al., 2004)

Pho, 5 µl (Brown et al., 2003)

E(z), 10 µl (Carrington and Jones, 1996)

H3, 0.5 µl (Abcam #ab1791)

Mock, 0.5 µl (Anti-IgG, Cell Signaling #2729)

Pcl, 10 µl (O'Connell et al., 2001)

Dsp1, 5 µl (R. Jones lab, unpublished)

Spps, 2.5 µl (Brown et al., 2010)

Sfmbt, 5 µl (Alhaj abed, 2018)

Pho-like, 15 µl (Wang et al., 2004)

Hb, 5 µl (Ahaj Abed et al., 2018)

Cad, 5 µl (Ahaj Abed et al., 2018)

H3K27me3, 0.2 µl (Millipore #07-449)

H3K27ac, 0.2 µl (Abcam #ab4729)

H3K4me3, 0.2 µl (Abcam, ab8580)

H3K36me3, 0.2 µl (Abcam #9050)

H2AK119ub (Cell Signaling #D27C4)

P-Rpb1 CTD (S2) (Cell Signaling #E1Z3G)

P-Rpb1 CTD (S5) (Cell Signaling #D9N51)

Scm, 10 µl (Peterson et al., 2004)

Psc, 10 µl (Provided by Nicole Francis)

dMi-2 0.5, 1, 2 µl (Provided by Alexander Brehm)

Kr 5, 10, 15 µl (R. Jones lab, unpublished)

F(s)1h, 4 µl (Provided by Der-Hwa-Hwang)

Lsd1, 5 µl (R. Jones lab, unpublished)

Kni, 5, 10, 15 µl (R. Jones lab, unpublished)

CBP, 10, 15 µl (Provided by Mattias Mannervik)

RBP4, 1, 2 µl (abcam #ab51462)

2.2.4 PCR cycling

For each reaction tube, the reaction master mixture was prepared as described in **Table 2.3**. 5 µl of DNA was added to each PCR reaction and a series of 10%, 1% and 0.2% dilutions of

input genomic DNA were used as standards for each primer set. All amplifications were performed in triplicates using Rotor Gene RG3000 thermocycler for a total number of 40 cycles. Sequences of PCR primers and their annealing temperature are listed in **Table 2.4**.

Table 2.3- **qPCR reaction master mixture.**

	For 20 µl
H_2O	6.5
Quanta Biosciences Perfecta SYBR green supermix	12.5
Forward primer, 10 µM	0.5
Reverse primer, 10 µM	0.5

Rotor Gene 5 software was used to determine Ct values and sample concentrations were calculated using the ΔΔCT method (Rao et al., 2013), and then normalized to the negative regions, 12 or *Pka-C1*. The *Pka-C1* gene is not bound by PcG proteins (modENCODE) and is both maternally and zygotically expressed (Lott et al., 2011; Fisher et al., 2012). The p-values for all the statistical analyses were calculated using unpaired two-tailed Student's t-test. All graphs were plotted using GraphPad Prism 7.

Table 2.4- **Primers used in ChIP-qPCR assays and their corresponding annealing temperatures.**

Amplicon	**Primer**	**Primer Sequence (5'-3')**	**Annealing temp (C∘)**
3-endogenous	Forward	GGAGTCTTCCTGGGTGTCTCTACGC	55
	Reverse	CCACTTGCCGCACAGCCAAT	

3-transgene	Forward	GGAGTCTTCCTGGGTGTCTCTACGC	55
	Reverse	CTTGTTGGTCAAAGTAAACGACAT GGT	
4-endogenous *gt*	Forward	GGAGTCTTCCTGGGTGTCTCTACGC	55
	Reverse	CCACTTGCCGCACAGCCAAT	
4-Pelican-*gt*-wt	Forward	GGAGTCTTCCTGGGTGTCTCTACGC	55
	Reverse	CTTGTTGGTCAAAGTAAACGACAT GGT	
4-Pelican-*gt*-mut	Forward	TCTCGTGCATTAGCATGGTGTT	58
	Reverse	CCTACCTGCTTCGTACCGCTTC	
6-endogenous *gt*	Forward	TTTCTCGTGCATTAGCATTCACAA	58
	Reverse	GCCTACCTGCTTCGTACTCAGAG	
6-Pelican-*gt*-wt/mut	Forward	TTTCTCGTGCATTAGCATTCACAA	58
	Reverse	CCTACCTGCTTCGTACCGCTTC	
9-endogenous *gt*	Forward	CGTATAGCCCAGCCCAATC	56
	Reverse	GCTCATTATGGCGAAGGAACA	
9-Pelican-*gt*-wt/mut	Forward	CACCAGGTTCCCGAGTTCTAC	56
	Reverse	ATTGTGGCGAAGGCTGTATG	
Pka- C1	Forward	CCGGGCCATGCAATAAAGTA	58
	Reverse	CGCTTCCTCCAACTCCCTATATTC	
SD10-∆gt1	Forward	CCGGGCCATGCAATTTTCAT	58
	Reverse	CGCTTCCTCCAACTGGGATTTAAG	
SD10-PRE1.1-wt/mut	Forward	ATATGCCACGCCATCTTAGCAC	55
	Reverse	CCTCAGTTCTCAGTCCGCTTCTAAT	
4-endogenous *gt*	Forward	CCA AAT GCC ACA CAC AAC ACA	55
	Reverse	GCC AGT TTC ACA TGC ACA TCA A	
4-Pelican-*gt*-wt	Forward	CAATCAGCAGATTCTCCGGCT	58
	Reverse	AGCCGCACTCGCGCTTCTAC	
4-Pelican-*gt*-mut	Forward (*en* enhancer)	ATTGGCATTGTTATTGCCCAG	55
	Reverse (*en* promoter)	GGGCTTGTTAGGCAGCAATATG	
6-endogenous *gt*	Forward (within PRE1.1)	GAGCGGGACAGAGTCAGAAG	56

	Reverse (*en* promoter)	GGGGCTTGTTAGGCAGCAAT	
7	Forward	CTGACCAGCCAAGCGAAAAG	55
	Reverse	GGCCGGTGCAAACTTAAGATAG	
10	Forward	ATATGCCACGCCATCTTAGCAC	55
	Reverse	CCTCAGTTCTCAGTCCGCTTCTAAT	
12	Forward	CCA AAT GCC ACA CAC AAC ACA	55

2.3 Immunostaining of embryos

Embryos were collected in 4-hour time period at 25° C, fixed and processed as previously described (AlHaj Abed et al., 2013) with the following modifications. Fix solution contained 4% formaldehyde (750 µl of 16% formaldehyde or 324 µl of 37% formaldehyde) in 1X PEM in the total volume of 3 ml. Embryos were crosslinked for 20 minutes, followed by at least 3 washes with heptane to remove residuals of fixative.

Rabbit anti-Hb antibody was diluted 1:250, guinea pig anti-Gt was diluted 1:500 (Kosman et al., 1998), rat anti-Kni antibody (Stephen Small) was diluted 1/500, rat anti-Kr antibody (Yu and Small, 2008) was diluted 1/150, rabbit anti-Cad antibody was diluted 1/400. Rabbit anti-□-galactosidase antibodies (Cappel) were diluted 1:1500. Biotin-SP–conjugated goat-anti-rabbit secondary antibodies (Jackson Immunoresearch) were diluted 1:10,000. Biotin-SP–conjugated goat-anti-rat secondary antibodies (Jackson Immunoresearch) were diluted 1:500. Biotin-SP–conjugated goat-anti-guinea pig secondary antibodies (Jackson Immunoresearch) were diluted 1:5000. Streptavidin-horseradish peroxidase (Jackson Immunoresearch) was diluted 1:5000. Images were obtained using a Ziess Axiovert 200M microscope. Staining were developed by incubating embryos in 1 mg/ml diaminobenzidine (Sigma-Aldrich) in 0.1 M Tris-HCl (pH 7), 0.02% $NiCl_2$, 0.001% H_2O_2. Time required for the signals to be detected is mostly dependent on the antibody and the dilution used. If signals are not detected after 10-15 minutes of staining, trace

amount of H_2O_2 (0.0005%) can be gradually added to the embryos. Adding too much H_2O_2 results in a very high background. Reactions were stopped by washing with PBST (0.01% Trition-X-100). Embryos were dehydrated in increasing ethanol series (30%, 60%, 90% and 100%) and mounted in Permount (Fisher Scientific). Images were obtained using a Ziess Axiovert 200 M microscope.

2.4. In situ hybridization of embryos

Embryos were processed as described for immunostaining. In situ hybridization of embryos were performed using previously described protocol with following modifications (Lehmann R, Tautz D. 1994). Single-stranded digoxigenin-labeled RNA probes for *gt* and *lac*Z were synthesized using an in vitro RNA transcription kit (DIG RNA Labeling kit from Roche) and listed primers (**Table 2.5**), from CH322-05H16 BAC clones and Pelican P element reporter vector (Barolo et al., 2000), respectively. To make *gt* probes, first a 1.8 kb fragment was amplified using *gt*-out primers and then the PCR product was re-amplified using *gt*-T7 primers. Probes were hydrolyzed to make 300-700 bp RNA fragments using 0.2M sodium carbonate buffer, prior to hybridizing to the fixed embryos. Proteinase K digestion step is dispensable, and should be followed only if the probe gives a very good staining ratio of signal to background. Otherwise, this step can result in a high staining background. Probe denaturation was performed at 80°C for 5 minutes, followed by hybridization at 45°C overnight. Anti-Digoxigenin-antibody conjugated to alkaline phosphatase (Roche) was used with the final working dilution of 1:1000 and detection was done by incubating embryos in NBT/BCIP solution. Stained embryos were washed in PBST and dehydrated in ethanol series (30%, 60%, 90%, and 100%), and mounted in Permount (Fisher Scientific #SP15). Images were obtained using a Zeiss Axiovert 200 M microscope.

Table 2.5- **Primers used to make gt and lacZ probes and their corresponding annealing temperatures.**

Amplicon	Primer	Primer Sequence (5'-3')	Annealing temp (C∘)
***gt*-out**	Forward	CCA GGA GGC GAC CAA CGA GA	72
	Reverse	CCA GAG ACC ATA CAC CGA ACA CCA T	
***gt*-T7**	Forward	AAG TAATACGACT CACTATAG GGCTATGAAAAA	70
	Reverse	CAAGCCCCTGATGCACCACCAC	
***lac*Z-T7**	Forward	GGAAAACCCTGGCGTTACCCAACTTAATC	72
	Reverse	AAGTAATACGCCTCCGCCGCCTTCATACTGC	

2.5 Extraction of genomic DNA from adult flies

Each fly was gently homogenized in 100 µl of lysis buffer containing 0.1 M Tris-HCl pH 9, 0.1 M EDTA, 1% SDS. Homogenate was then incubated for 30 minutes at 70°C. 1 M Potassium acetate (14 µl of 8 M Potassium acetate) was added to each tube, thoroughly mixed and incubated on ice for 30 minutes. Following centrifugation for 15 minutes, supernatant was transferred to a new tube and DNA was precipitated by adding 0.5 volume isopropanol. The pellet was washed with 70% ethanol twice, air dried and resuspended in 50 µl of PCR water. Extracted DNA was directly used in PCR reactions.

2.6 Preparation of *Drosophila* embryo extracts

Embryos were dechorionated in 50% bleach for 2 minutes and washed extensively with 0.03% Trtiton-X-100. Embryos were then homogenized in 2X volume of lysis buffer containing 40 mM HEPES pH 7.5, 350 mM NaCl, 10% glycerol, 0.1% tween-20, and 1X protease inhibitor cocktail (Sigma #P8340), using appropriate pestles. The homogenized extract was centrifuged at 13,000 rpm for 30 min at 4°C. The supernatant was flash frozen and stored at -80°C.

2.7 Western Blot analysis

Extracts of *Drosophila* embryos were run on a SDS-PAGE and the proteins were then transferred to nitrocellulose membrane. Blocking of membrane was performed in 5% milk diluted in TBS-0.1%Tween for 2 hours. Primary rabbit antibodies were diluted 1:1000 in 5% milk TBS-0.1%Tween and incubated with the membrane overnight at 4°C. Membranes were washed three times with TBS-0.1%Tween followed by incubation with horseradish peroxidase conjugated to goat anti-rabbit secondary antibodies (1:25,000, Jackson Immunoresearch). Membranes were washed three additional times with 5% milk TBS-0.1%Tween. Protein-antibody complexes were detected using Amersham ECL Prime Western Blotting Detection Reagent (GE Healthcare). Molecular weights of proteins were roughly calculated by comparison to standards of known molecular weight (Bio-Rad).

2.8 RT-qPCR

Embryos were collected for one hour, and then aged at 25°C for 110 minutes to reach nc13-nc14a. Embryos were then dechorinated in 50% bleach for 2 minutes and then manually sorted. Total RNA was isolated using manufacturer's protocol provided by Aurum total RNA mini kit (Bio-Rad). To quantitate RNA concentration, spectrophotometer (Eppendorf No. 613102570) was used and reading was taken at wave lengths of 260 nm. For all samples, 3 µg of RNA was reverse transcribed using Maxima First Strand cDNA synthesis kit (Thermo Scientific). Reaction mixture was prepared according to the manufacturer's protocol and then incubated for 10 minutes at 25°C followed by 30 minutes at 55°C. Reaction was terminated by heating at 85°C for 5 minutes. cDNA was analyzed by qPCR using gene-specific primers (**Table 2.6**). PCR reactions were performed in triplicates. To calculate the triplicate mean values, ΔΔCT quantification method with *rp49* transcript as normalization reference, was used.

Table 2.6- **Primers used in RT-qPCR and their corresponding annealing temperatures.**

Amplicon	Primer	Primer Sequence (5'-3')	Annealing temp (C°)
rp49	Forward	ATGACCATCCGCCCAGCATACA	55
	Reverse	GTCGATACCCTTGGGCTTGCG	
lacZ	Forward	GGAGTCTTCCTGGGTGTCTCTACGC	55
	Reverse	CTTGTTGGTCAAAGTAAACGACATGGT	

2.9 Affinity purification of 6X Histidine-tagged proteins

2.9.1 Protein extraction

Bacteria containing vectors expressing 6X Histidine-tagged E(z), Pc, Dsp1, Kni and Kr were cultured in 2XYT media supplemented with 100 μg/ml Ampicillin at 25°C overnight. Saturated overnight cultures were diluted 1:20 with 100 μg/ml Ampicillin. IPTG was added to a final concentration of 1 mM to induce protein expression after an additional 3-4 hours of growth. The cultures were then incubated at 37°C for another 3 hours or at 25°C overnight. Bacterial lysates containing polyhistidine-tagged fusion proteins were isolated after centrifugation for 10 minutes at 6000 rpm at 4°C. Cells were lysed in 8 M urea, 0.1 M NaH_2PO_4, 0.01 M Tris-HCl pH 8. Ni^{+2}-NTA agarose beads (Invitrogen, R901-01) was used as an affinity chromatography matrix for purifying recombinant proteins. 1:4 ratio of Ni^{+2}-NTA: lysate was used and the mixture was incubated for 1 hour at room temperature. The mixture was then transferred to the column and Histidine-tagged proteins were eluted in elution buffer containing 8 M urea, 0.1 M NaH_2PO_4, 0.01 M Tris-HCl pH 4.5. The eluent was then dialyzed against the coupling buffer (0.1 M $NaHCO_3$, 0.5 M NaCl pH 8.3) overnight.

2.9.2 Medium preparation

Each gram of CN-Br-activated sepharose powder (GE Healthcare) gives a volume of 3.5 ml which has the capacity for binding to up to 5-10 mg of protein. Appropriate amount of CN-Br beads was suspended in 1 mM HCl and rotated to get separated. The beads were then washed though the sintered glass filter with the vacuum suction with 1 mM HCl and coupling buffer (0.1 M NaHCO3, 0.5 M NaCl, pH 8.3). CN-Br beads were then suspended with recombinant proteins for 1 hour at room temperature. The ligand- CNBr coupled beads were centrifuged at 2200 rpm for 5 minutes and then washed with coupling buffer, followed by blocking with 0.1 M Tris-HCl, pH 8 at room temperature for 2 hours. The mixture was washed 3X with the wash buffers (0.5 M NaCl, 0.1 M NaOAc) with alternating pHs, 2 and 8. The mixture of ligand-CNBr coupled beads was left in the pH 8 wash and then stored at 4°C. Aliquots of the supernatant were run on a SDS-PAGE gel, followed by coomassie blue staining before and after coupling to determine coupling efficiency. BCA assay was performed using BCA protein assay kit (Thermo Scientific) to determine the total concentration of proteins.

2.9.3 Affinity purification of antibodies

Histidine-tagged recombinant proteins-CNBr beads, whether used as pre-adsorption or test beads, were washed with PBTN (1x PBS, 0.3 M NaCl, 0.1%Triton-X-100). The anti-serum was diluted with equal amount of PBTN, added to the pre-adsorption beads and incubated for 1 hour at room temperature. The mixture was poured into a column and the flow through was collected. Approximately, 4-5 mg of proteins attached to CNBr beads were used, considering that the total volume of CNBr beads does not exceed 2 ml. The pre-adsorbed anti-serum was incubated with Histidine-tagged protein-CNBr coupled beads at 4°C overnight. The mixture was then poured into a column. The flow through was collected and beads were washed with 100 ml PBTN. Elution

step was performed with 2X 5 ml fractions of elution buffer containing 0.2 M glycine, 0.8 M NaCl pH 2.5. The flow rate of elution buffer was kept very slowly and the elution was done one ml at a time. After collecting the eluate, pH was equilibrated by adding 50 µl of 2 M Tris-HCl pH 8 to each 1 ml of purified antibodies. Eluate was dialyzed in buffer containing 1X PBS, 0.02% NaN_3 at 4°C for at least 4 hours then 50 or 100 µl aliquots were taken, flash frozen and stored in -80. The affinity purified antibody was tested in ChIP and/or western blot assays. E(z), Pc, Dsp1, Kni and Kr antibodies were affinity-purified using this protocol.

2.10 Affinity purification of GST-tagged proteins

Bacteria containing pGEX plasmids expressing GST-tagged proteins were cultured in 2XYT media supplemented with 100 µg/ml Ampicillin as described for purification of 6X Histidine-tagged proteins. Cultures were induced by 1 mM IPTG. Bacterial lysates were isolated after centrifugation for 10 minutes at 6000 rpm at 4°C and then resuspended in 30 ml ice-cold 1X PBS to wash. After centrifugation, the pellet was resuspended in 10 ml ice-cold STE buffer (10 mM Tris-Hcl pH 8, amM EDTA,150 mM NaCl). 100 µl of freshly made lyozyme solution added, and incubated on ice for 15 minutes. Sonication was performed for the total time of 1 minute after adding 100 µl of 1 M DTT and 1.4 ml of 10% Sarkosyl to the lysate. Debris was pelleted and supernatant was saved. 4 ml of 10% Triton X-100 was added and STE buffer was added to 20 ml, leaving the final concentrations of Sarkosyl and Triton X-100 to 0.7% and 2%, respectively. The solution was then incubated at room temperature for 30 minutes. To make 50% slurry of Glutathione Sepharose beads, 2 ml of Glutathione Sepharose beads was thoroughly mixed with 48 ml of PBS. Supernatant was discarded after the centrifugation and the pellet was resuspended in 2 ml of PBS. 1 ml of 50% slurry of Glutathione Sepharose beads was added to the lysate and the mixture was incubated for 1 hour at room temperature with agitation. The beads were washed with

50 ml PBS for 3 times, resuspended in 5 ml of PBS and loaded into a column. The beads were then washed with 100 ml of PBS and eluted with 2 X 5 ml fractions of elution buffer. Eluent was then dialyzed in dialysis buffer containing 1X PBS, 0.02% NaN_3 at 4°C for at least 4 hours then 100 µl aliquots were taken, flash frozen and stored in -80. The affinity purified antibody was tested in ChIP and western blot assays. E(z) anti-sera were affinity-purified using this protocol.

2.11 Transgenic RNAi Project (TRiP) lines

The TRiP Project at Harvard University has generated over 12,000 Gal4-inducible short hairpin RNA expressing transgenic *Drosophila* lines that may be used for the targeted depletion of thousands of fly proteins. TRiP reagents and fly stocks are transferred to the Bloomington *Drosophila* Stock Center (BDSC) for distribution to the fly community. Maternally expressed Gal4 drivers in conjunction with TRiP UAS-shRNA transgenes were used to knock down the levels of proteins of interest. In order to test the efficiency of UAS-shRNA-mediated knock down of specific proteins, I looked at the mortality rate and cuticle patterns of the embryos. Transgenic lines tested for the efficient knockdown is listed in **Table 2.7.**

Table 2.7- **TRiP lines tested for the efficient knockdown using maternally expressed Gal4 driver.**

Name	**BDSC Stock ID**	**Vector**	**Insertion site**
Ash1	36803	VALIUM22	attp40
Hb	54478	VALIUM22	attP2
Cad	57546	VALIUM20	attP40
Set1	40931	VALIUM20	attp40
Fs(1)h	41693	VALIUM20	attp40
Fs(1)h	41943	VALIUM20	attp40
Fs(1)h	44009	VALIUM20	attp40
Cdk9	41932	VALIUM20	attp40
Set2	42511	VALIUM20	attp40
Haywire	53345	VALIUM20	attp40
Set2	55221	VALIUM20	attp40

Cdk7	57245	VALIUM20	attp40
Cdk7	62304	VALIUM20	attp40
Ash2	64942	VALIUM20	attp40
Lsd-1	65020	VALIUM20	attp40
Wds	60399	VALIUM20	attp40
Mat1	57312	VALIUM20	attp40
Xpd	65883	VALIUM20	attp40
CycT	32976	VALIUM20	attp2
Ash2	35388	VALIUM22	attp2
Cdk9	35323	VALIUM22	attp2
Lilli	34592	VALIUM20	attp2
Cdk7	53199	VALIUM22	attp2
dMi-2	35398	VALIUM22	attp2
CBP	37489	VALIUM20	attp2
Hopped CBP -2.1		VALIUM20	2nd chromosome
Hopped CBP-2.2		VALIUM20	2nd chromosome
Hopped CBP-3.1		VALIUM20	2nd chromosome
Hopped CBP-17		VALIUM20	2nd chromosome
Hopped CBP-36.3		VALIUM20	2nd chromosome
Hopped CBP-36.2		VALIUM20	2nd chromosome
Hopped CBP-10		VALIUM20	2nd chromosome

2.12 Establishment of *bcd osk tsl Drosophila* genetic system

All the preliminary work to generate and verify the *bcd osk tsl Drosophila* genetic system has been done by the former graduate student, Jumana Alhaj Abed. *bcd*$_6$ *osk*$_6$ *tsl*^{PZRev32} on the right arm of the third chromosome (3R), was gifted from Leslie Stevens. Another stock containing *bcd*$_7$ *osk*$_6$ *tsl*$_4$, was obtained from crossing the *tsl*$_4$ allele (Bloomington Drosophila Stock Center #3289) to the *bcd*$_7$*osk*$_6$ (Bloomington Drosophila Stock Center #3252) chromosome (R. Jones). Both stocks were not able to homozygose because of the presence of unidentified recessive lethal mutations and thus were crossed together to generate trans-heterozygous females containing *bcd*$_7$ *osk*$_6$ *tsl*$_4$/*bcd*$_6$ *osk*$_6$ *tsl*^{PZREV32} females. Embryos derived from *bcd osk tsl* females, offer an in vivo

model of a homogenous population of cells in which *gt* is uniformly repressed by PcG proteins during early embryogenesis.

2.13 Establishment of *Drosophila* genetic system which expresses *gt* uniformly

In order to produce embryos in which *gt* is ubiquitously active, the levels of maternal Hb, a repressor of *gt*, was knocked down in a *bcd osk tsl* maternal background. Zygotic Hb is not expressed in the resulting embryos due to *bcd* and *tsl*. Cad, the activator of *gt*, is uniformly expressed.

2.14 Knock down of maternal levels of Cad in *bcd osk tsl* maternal background

Maternally expressed Gal4 drivers in conjunction with UAS-shRNA, was used to knock down the levels of maternal Cad, an activator of *gt*, in a *bcd osk tsl* maternal background. Zygotic Hb is not expressed due to *bcd* and *tsl*. Hb, the repressor of *gt*, is uniformly expressed in the resulting embryos.

2.15 Knock down of maternal levels of Cad and Hb proteins in *bcd osk tsl* maternal background

In order to generate a genetic system in which all known activators and repressors of *gt* are absent, maternal levels of both Hb and Cad proteins were knocked down in a *bcd osk tsl* maternal background. Activators of *gt*, Cad, Bcd and Gt, are absent in the resulting embryos. Furthermore, *kni* and *Kr*, repressors of *gt*, are not expected to be expressed in this genetic system due to the absence of their activators, Cad and Bcd, respectively. Hb, another known repressor of *gt*, is present at very low levels due to the maternally expressed shRNAs for this protein.

2.16 Cuticle Preparation

Embryos were collected in 24-hour time period and aged for another 24 hours at 25°C. Embryos were then washed off the plates and dechorinated for 2 minutes in 50% bleach. To de-vitellinize, embryos were vigorously shaken in a mixture of 1:1 heptane: ice-cold methanol for 30

seconds. De-vitellinized embryos sank to the bottom of the tube and were washed three times with methanol to remove traces of heptane. Embryos were then washed with 0.1% Triton three times to remove methanol. Embryos were transferred onto a microscope slide and mountant (Polysciences) was added. Slides were covered by cover glasses and incubated at 70°C for overnight before inspection under microscope.

CHAPTER 3:

THE REQUIREMENT OF PHO FOR THE RECRUITMENT OF PCG PROTEINS AND MAINTENANCE OF PCG-MEDIATED REPRESSION AT GT PRES

3.1 Introduction

Drosophila PREs have been shown to be required for the recruitment of PcG proteins and formation of PcG domains. Studies of different *Drosophila* PREs have identified a number of consensus protein binding sites, including those for Pho, Phol, Dsp1 and Spps, to be important for PRE activity (reviewed in Kassis and brown, 2013). However, the exact sequence motifs required for PRE activity remain elusive due to the presence of different combinations of PRE binding proteins and low conservation of their consensus binding site sequences (Kassis and Brown, 2013).

Pho has been shown to bind to at all characterized PREs (Kassis and Brown, 2013), and thus is a good indicator of the presence of a PRE. Genome-wide studies have shown that binding of Pho is strongly correlated with PRC1 subunits (Oktaba et al. 2008) and the hierarchical model for the recruitment of PcG proteins, places Pho at the base of the recruitment (Wang et al., 2004). Phol is partially functionally redundant with Pho during larval development and all amino acids involved in making DNA contacts are conserved in these two proteins (Alfieri et al., 2013). Pho binding sites are demonstrated to be required for *bxd* PRE activity of the endogenous *Ubx* gene (Kozma et al. 2008). Mutations of Pho binding sites within *bxd* PRE, demonstrated that recruitment of PRC1 and PRC2 is dependent on the binding of Pho to Pho binding sites (Frey et al., 2016). However, a number of studies have published evidence that Pho is not sufficient to recruit PcG

proteins (Americo et al., 2002; Brown et al., 1998; Fritsch et al., 1999; Shimell et al., 2000; Poux et al., 2001a).

Several *Drosophila* PcG target genes, such as *engrailed-invected* (*en-inv*), *dachshund* (*dac*), *vestigial* (*vg*), and *gt*, have a minimum of two PREs (Abed et al. 2013; Ahmad and Spens 2018; Cunningham et al. 2010; DeVido et al. 2008; Ogiyama et al. 2018). At some genes, such as *dac*, the two PREs are at least partially non-redundant and cooperate to form PcG domains (Ogiyama et al. 2018). However, deletion of four strong PREs of *en/inv* region did not result in the alteration of the PcG domain or the derepression of *en* (De et al. 2016).

Two PREs have been mapped for the *gt* locus; PRE1, a 1.9 kb fragment, which overlaps with the *gt_(-1)* enhancer and encompasses the *gt* promoter. PRE2 localized 6 kb upstream from the *gt* TSS and overlaps with *gt_(-6)* enhancer. We have previously described the temporal sequence of de novo recruitment of PcG proteins at *gt* two PREs, in syncytial blastoderm through cellular blastoderm/gastrulating embryos (Abed et al. 2018). PhoRC binding to PRE1 precedes its stable binding to PRE2 by 30 minutes (Abed et al. 2018). This raises key questions about possible dependency of PhoRC binding to PRE2 on the presence of PhoRC at PRE1 and the requirement of Pho binding to PRE1 for the recruitment of PcG proteins to both PREs.

In order to determine whether PhoRC binding to PRE2 is dependent on its presence at PRE1 and the consequence of absence of Pho on the binding of PcG proteins to both PREs, we have generated flies carrying reporter transgenes which contain either the entire wild type *gt* regulatory region or the same *gt* fragment with mutations in Pho/Phol binding sites in PRE1. We found that PRE2 is redundant with PRE1 for the recruitment of PcG proteins and maintenance of transcriptional repression. Two other PRE binding proteins, Dsp1 and Spps, are differently dependent on Pho for PRE1 binding. Surprisingly, we find that Phol binding to PRE1 is unaffected

by the mutation of Pho consensus binding sites that eliminate binding by Pho. Our findings shed light on the functions of individual PRE binding proteins and the modulating effect of PcG proteins at transcriptionally active genes.

3.2 Results

3.2.1 Mutations of Pho binding sites eliminated PRE1.1 activity

In order to more narrowly define the boundaries of *gt* PRE1, a smaller fragment extending from 543 upstream to 12 bp downstream from *gt* TSS (PRE1.1), was tested for the PRE activity within the context of the SD10 P element reporter vector (**Figure 3.1A**; DeVido et al., 2008). SD10 includes *en* enhancer and promoter that produces *lacZ* expression in 14 stripes, recapitulating the expression pattern of endogenous *en* at this developmental stage (**Figure 3.1C**). Inclusion of a fragment with PRE activity results in the maintenance of *lacZ* expression from SD10 transgenes in 14 stripes, while lack of a PRE, results in the ectopic expression of *lacZ* between the stripes (Abed et al., 2013; Cunningham et al., 2010; DeVido et al., 2008).

The PRE activity of PRE1.1 was assessed by immunostaining of stage 14 embryos from five SD10-PRE1.1 transgenic lines using anti-β-galactosidase (anti-β-gal) antibody. All tested five lines exhibited the PRE activity in the maintenance assays (a representative embryo is shown in **Figure 3.1D**).

Two predicted Pho/Phol consensus binding sites, CGCCATTT, are present within PRE1.1. These binding sites resemble the extended Pho/Phol binding sites (CGCCAT(T/A)TT). In order to investigate the requirement of Pho for the activity of PRE1.1, we mutated the core sequence of Pho binding sites (GCCAT; **Figure 3.1B**) and cloned this fragment into the SD10 vector and transgenic lines were generated. Nine out of nine lines demonstrated ectopic β-gal expression in

the maintenance assays (a representative embryo is shown in **Figure 1D**), indicating the disruption of PRE1.1 activity due to the mutation of the Pho/Phol consensus binding sites.

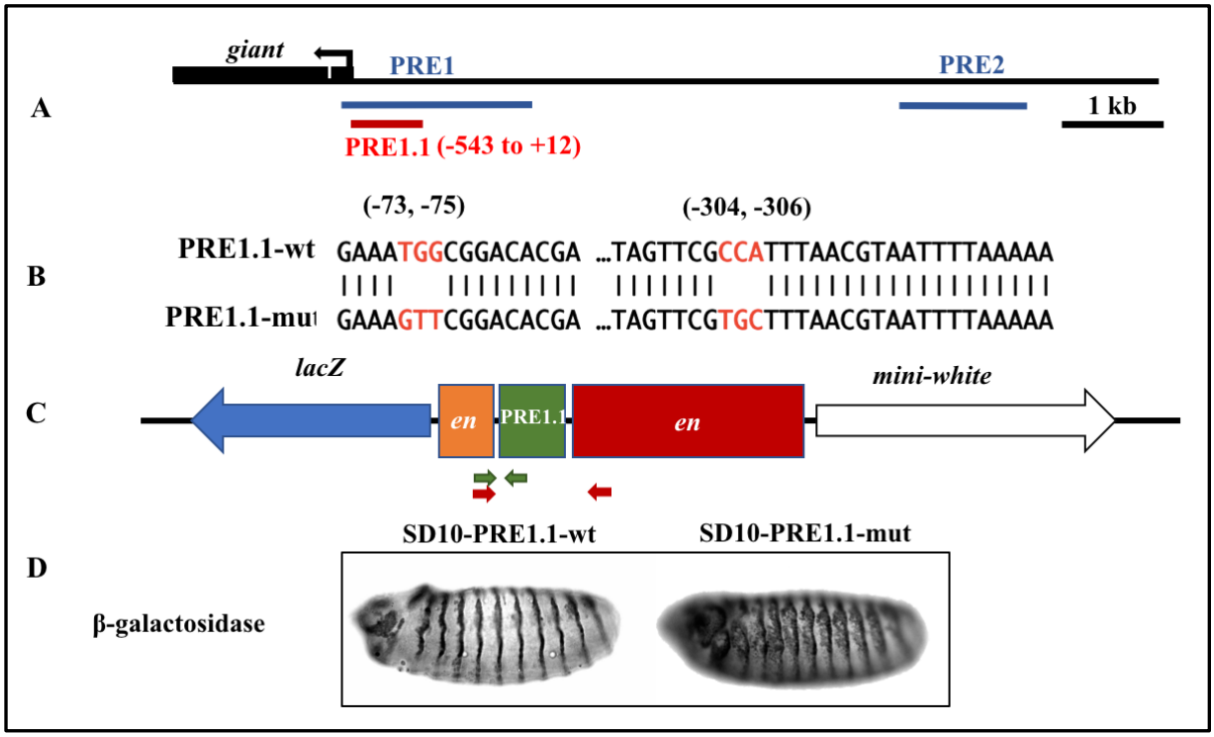

Figure 3.1- **Mutations of Pho-Phol consensus binding sites disrupt PRE1.1 activity**. (**A**) Schematic above shows the *gt* genomic region containing two PREs, shown as blue bars. PRE1.1 fragment (+12 to -543) is shown as red bar. (**B**) Sequence alignment of the wild type PRE1.1 region containing two extended Pho-Phol consensus binding sites with PRE1.1 containing mutant Pho-Phol consensus sequences; coordinates of the mutated sites are indicated above. (**C**) Schematic representation of the SD10 P element reporter vector. Orange and red rectangles labeled "*en*" represent the *engrailed* promoter and enhancer, respectively. Arrows indicate the direction of transcription of the reporter gene (*lac*Z) and marker (mini-*white* gene). Green and red arrows below indicate the approximate locations of PCR priming sites for detection of ChIP signals from SD10-PRE1.1 and SD10-Δgt1 transgenes, respectively. (**D**) Representative stage 14 embryos from lines carrying SD10-PRE1.1-wt or SD10-PRE1.1-mut transgenes immunostained with anti-β-galactosidase. Embryos are shown as lateral views, dorsal up, anterior to the left.

3.2.2 Binding of Phol and Sfmbt was not abrogated by the mutation of Pho binding sites in SD10 vector

In order to confirm that mutations of Pho/Phol binding sites can abrogate Pho binding in SD10 vector, we performed ChIP assays on nc14b-nc15 SD10 transgene carrying embryos. Pho binds to the endogenous PRE1.1 and PRE1.1-wt transgene, while its binding to SD10-PRE1.1-mut transgene was reduced to background levels (**Figure 3.2**). To determine if Phol binds to PRE1.1 and if its binding is also affected by the mutations of Pho/Phol binding sites, we looked at Phol binding in SD10-PRE1.1-wt and SD10-PRE1.1-mut embryos. Phol binds to the endogenous and SD10-PRE1.1-wt transgene. However, Phol binding to SD10-PRE1.1-mut transgene was decreased to half, but not eliminated by the mutations of Pho/Phol binding sites (**Figure 3.2**). Given that Phol and Phol were shown to bind to the exact DNA sequences in vitro (Brown et al., 2003), the latter observation was surprising, and raised three possibilities: 1) Mutation of Pho/Phol bindings sites reduces but does not abrogate binding of Phol. 2) Phol binds to other sites within PRE1.1. 3) Phol binds to sites within SD10 vector. To test which of these possibilities was correct, we looked at Phol binding to SD10 vector. Our lab previously used the modified version of SD10 vector, in which *gt* inserts were flanked by FRT sites (Alhaj Abed et al., 2013). In the latter study, a series of *gt* fragments were cloned into the SD10 vector to be tested for PRE activity in the context of SD10 vector. FRT-flanked fragments were later deleted by FLP recombinase, leaving only the SD10 construct (SD10-Δgt1), resulting in the ectopic β-gal expression and loss of maintenance of *en*-like *lac*Z expression (Alhaj Abed et al., 2013). Unexpectedly, Phol binding to the transgene of SD10-Δgt1 embryos was at the same level as that to the SD10-PRE1.1-mut transgene. These results suggested that mutation of Pho binding sites abrogated binding of Phol to mutated sites and the detected Phol signal resulted from Phol binding to the SD10. Phol was shown to bind outside PcG domains, within 300 bp of TSS of a subset of active genes (Kahn et al., 2014).

Given that *en* promoter is located just downstream of PRE1.1, we suggest that Phol may bind to the *en* promoter.

Both Pho and Phol make a dimer with Sfmbt and interact with it in a mutually exclusive manner (Kahn et al. 2014). In order to test if the detected Phol at the transgene of SD10-PRE1.1-mut and SD10-Δgt1 embryos is present as a subunit of PhoRC, we looked at Sfmbt signals at SD10 containing embryos. Sfmbt binds to the transgene of SD10-PRE1.1-wt embryos at the levels comparable to endogenous PRE1.1. However, Sfmbt binding to SD10-PRE1.1-mut was surprisingly at the same level of that at SD10-Δgt1 transgene and showed about 60% reduction compared to the endogenous PRE1.1 (**Figure 3.2**).

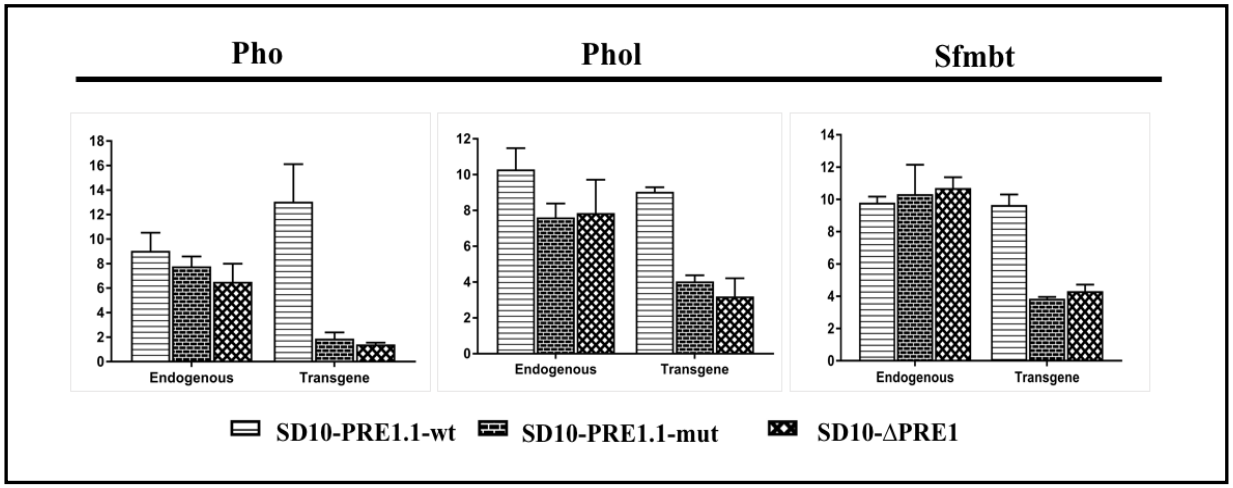

Figure 3.2- **Effects of Pho-Phol consensus binding sites mutations on PhoRC recruitment.** ChIP-qPCR was performed with anti-Pho, anti-Phol and anti-Sfmbt antibodies, as indicated above, using nc14b-nc15 SD10-PRE1.1-wt, SD10-PRE1.1-mut and SD10-Δgt1 embryos. ChIP signals are presented as fold enrichment relative to *Pka-C1*, shown at 1. Error bars show standard deviations for three biological replicates.

3.2.3 PRE1 is redundant with PRE2 for the maintenance of PcG-mediated repression of *gt*

In order to test *gt* PRE activities within the entire *gt* upstream regulatory region, a 10.8 kb *gt* fragment (-10421 to +360), including all four enhancers and two PREs, was used (**Figure 3.3A**). To distinguish signals of the transgenes from the endogenous *gt* in ChIP-qPCR experiments, we introduced base pair substitution in the PCR priming sites for regions 4, 6 and 9. This *gt* fragment was then inserted into a modified Pelican reporter vector (**Figure 3.3B,** Barolo et al., 2000). The transgene was integrated into the attP40 docking site and transgenic flies were generated. In order to examine the recruitment of PcG proteins to PRE1 and PRE2, in the absence of Pho at PRE1, the same sequence changes that were made in PRE1.1 fragment in the context of SD10 vector, were introduced in this larger fragment.

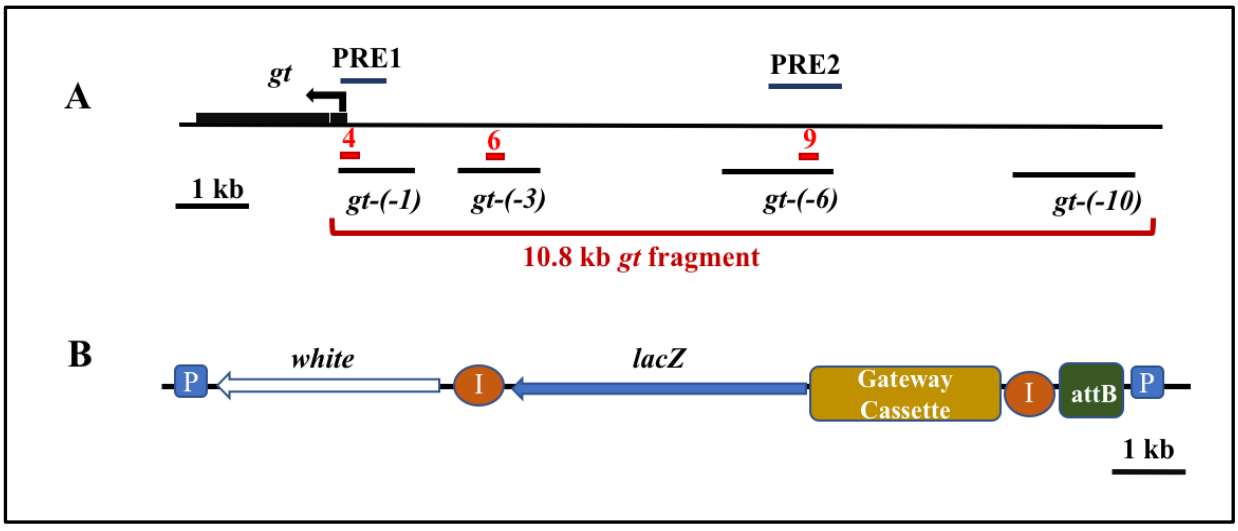

Figure 3.3- **Pelican-*gt* reporter construct.** (A) Schematic map of the *gt* genomic region including 4 enhancers and two PREs. Regions amplified by qPCR (4, 6 and 9) in ChIP assays are represented as small red bars. Enhancer-containing regions are indicated with black bars. (B) Diagram of modified Pelican P element reporter vector. "I" represents *gypsy* insulators. "P" represents P element transposition sequences. Locations of added Gateway cassette and attB sequence indicated. Arrows indicate the direction of transcription of the reporter gene (*lac*Z) and marker (*white*).

PcG-mediated repression of *gt* is redundant with transcription factors, such as Hb and Kr, during early embryonic stages. Negre et al. (2006) showed the ectopic expression of *gt* in PcG mutant background in stage 10 embryos, which indicates that *gt* repression is PcG-mediated in this developmental stage (Negre et al., 2006). In situ hybridization of Pelican-*gt*-wt embryos showed *lac*Z expression in a pattern recapitulating endogenous *gt* expression (**Figure 3.4**). This observation demonstrated that the ~10.8 kb fragment contains all the regulatory elements necessary for regulating *gt* expression. Maintenance of *gt*-like pattern of *lac*Z expression in Pelican-*gt*-mut transgene through stage 10 showed that PcG-mediated repression of *gt* was not affected by the mutations of Pho binding sites and the activity of PRE2 alone was sufficient for the maintenance of PcG-mediated repression (**Figure 3.4**). However, the possible presence of weak small PREs cannot be excluded.

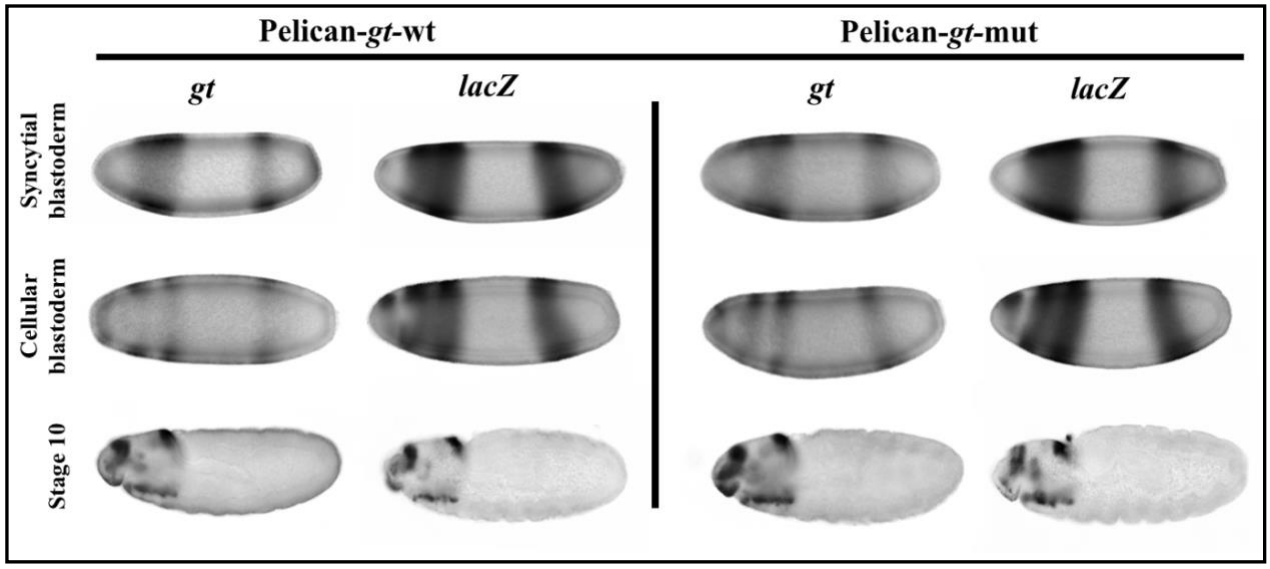

Figure 3.4- **Maintenance of *gt*-like *lacZ* expression from Pelican-*gt*-wt and Pelican-*gt*-mut transgenes.** Pelican-*gt*-wt (left) and Pelican-*gt*-mut (right) embryos hybridized with *gt* or *lac*Z RNA probes, as indicated. Embryonic stages are indicated on the left.

RT-qPCR analysis showed small but significant increase of *lacZ* transcription due to the mutations of Pho binding sites from Pelican-*gt*-mut transgene compared to that from Pelican-*gt*-wt transgene at syncytial and cellular blastoderm embryonic stages (**Figure 3.5**). At stage 10, there is still an increase in *lac*Z expression from Pelican-*gt*-mut transgene but this increase is not significant.

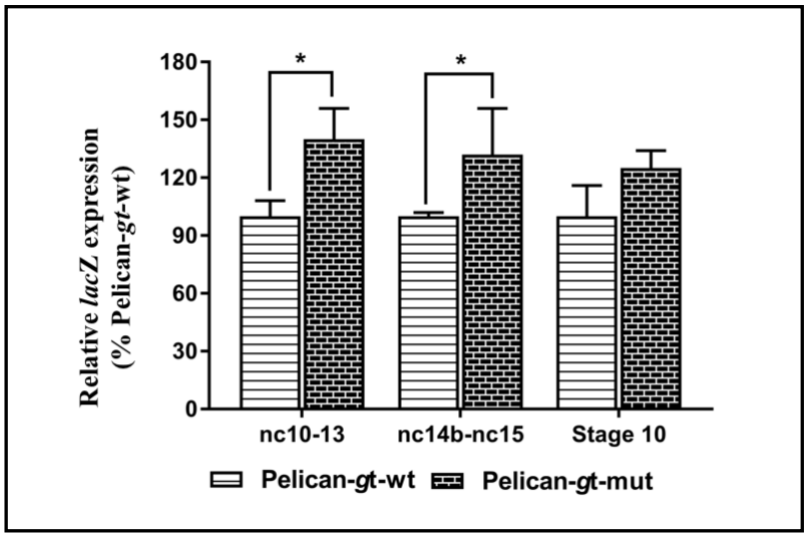

Figure 3.5- **Mutation of PRE1 Pho-Phol sites resulted in a greater amplitude of *lac*Z expression.** Genotype and developmental stages of embryos are indicated below. RT-qPCR signals were normalized to *rp49* transcript. Y-Axis shows the relative expression of *lac*Z to Pelican-*gt*-wt embryos. Error bars show the standard deviation for three biological replicates. *$p \leq 0.05$, unpaired two-tailed Student's t test.

3.2.4 Distribution of PcG proteins in Pelican transgene-containing embryos

Binding of PcG proteins at Pelican-*gt*-mut transgene was compared to that at Pelican-*gt*-wt transgene and the endogenous *gt* at three time points: syncytial blastoderm (nc10-13), cellular blastoderm-early germband elongation (nc14b-nc15), and stage 10.

Both Pho and Phol are strongly bound to PRE1 (region 4) at *gt* and Pelican-*gt*-wt transgene in nc10-13 embryos, while their binding to PRE2 (region 9) is very weakly positive at this stage. Pho and Phol PRE1 signals increase from nc10-13 to nc14b-nc15. However, Pho binding did not change from nc14b-nc15 to stage 10 at PRE1 of Pelican-*gt*-wt and endogenous *gt*, while signal of Phol dropped to one third in stage 10 embryos (**Figure 3.6**). Pho signals at PRE2 of endogenous *gt* and Pelican-*gt*-wt transgene increased across the time points, while Phol signals at PRE2 stayed weakly positive at all time points. Sfmbt binding to PRE1 of endogenous *gt* and Pelican-*gt*-wt transgene increased from nc10-13 to stage 10, however, its binding to PRE2 stayed weak but well above the background at all stages (**Figure 3.6**).

Binding of Pho is abrogated at Pelican-*gt*-mut transgene due to the mutation of Pho binding sites at all embryogenic stages. Interestingly, signal of Pho at PRE2 of Pelican-*gt*-mut transgene is not affected by the lack of Pho binding at PRE1 (**Figure 3.6**).

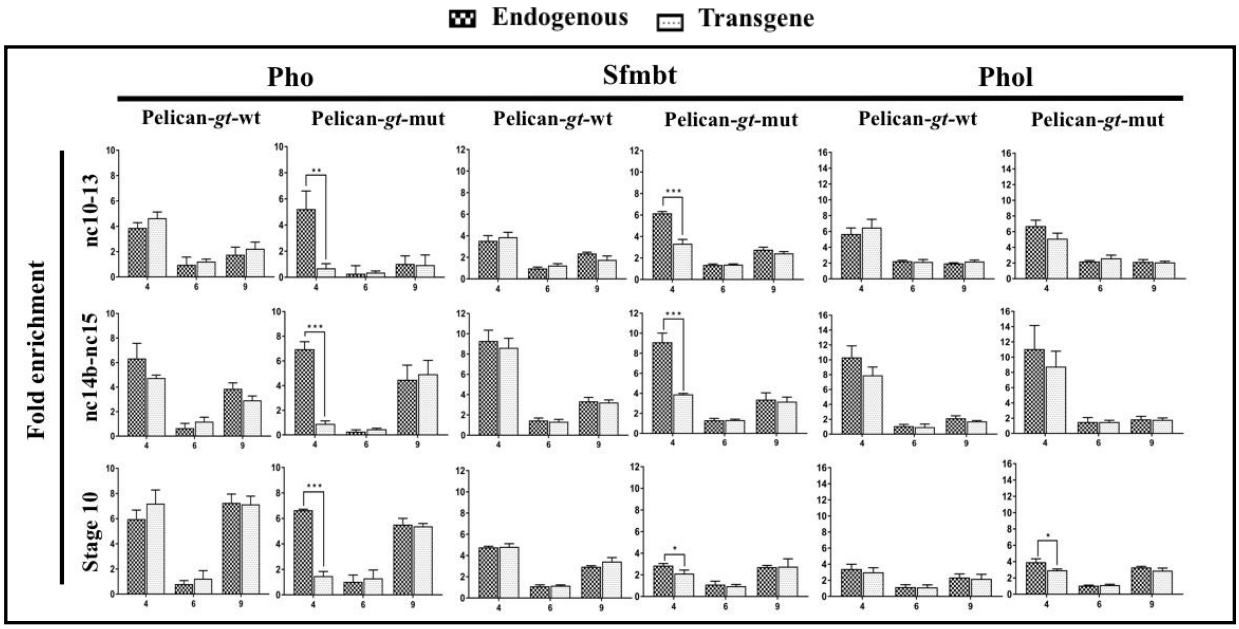

Figure 3.6- **Effects of mutant PRE1 Pho-Phol consensus binding sites on PhoRC recruitment to PRE1 and PRE2.** Time course ChIP-qPCR assays using Pelican-*gt*-wt and Pelican-*gt*-mut embryos at the developmental stages indicated on the left using anti-Pho, anti-Sfmbt and anti-Phol antibodies. ChIP signals are presented as fold enrichment relative to *Pka-C*, shown at 1. Error bars show the standard deviations for three biological replicates. *$p \leq 0.05$, **$p \leq 0.01$ and ***$p \leq 0.001$, unpaired two-tailed Student's t test.

Surprisingly, Phol binding level to PRE1 of Pelican-*gt*-mut transgene was comparable to that at Pelican-*gt*-wt transgene and endogenous *gt* in nc10-13 and nc14b-nc15 embryos. At stage 10, Phol binding showed a small but statistically significant reduction at PRE1 of Pelican-*gt*-mut transgene compared to the endogenous *gt* (**Figure 3.6**).

Sfmbt signals at PRE1 of Pelican-*gt*-mut transgene showed a significant reduction (~50% reduction) at all time points, but did not drop to the background levels. Sfmbt binding to PRE2 was not affected by the absence of Pho at PRE1 of Pelican-*gt*-mut transgene.

Our lab previously showed that E(z) and Pc, respective components of PRC2 and PRC1, are weakly associated with *gt* in nc10-12 and nc13 embryos, prior to their stable binding in nc14b

embryos (Abed et al., 2018). We have interpreted this weak interaction of E(z) and Pc with chromatin as "sampling", in which PRC1 and PRC2 sample the chromatin environment at PREs independently of Pho, prior to the strong and prolonged interaction of PRC2 with chromatin and thus deposition of H3K27me3, and stable binding of PRC1.

Similar to our previous report, E(z) was weakly positive and negative at *gt* PRE1 and PRE2, respectively, at endogenous *gt* and Pelican-*gt*-wt transgene in nc10-13 embryos. E(z) binding at PRE1 and PRE2 of Pelican-*gt*-wt transgene and endogenous *gt* increases from nc10-13 to nc14b-nc15 (**Figure 3.7**). Interestingly, mutation of Pho binding sites, resulted in the significant reduction of E(z) binding at PRE1 of Pelican-*gt*-mut transgene compared to the endogenous PRE1, as early as nc10-13 (**Figure 3.7**). This indicates that even the weak association of E(z) with chromatin or "sampling" was dependent on the presence of Pho. E(z) signal remained significantly reduced, dropped to half, at PRE1 of Pelican-*gt*-mut transgene through later time points. On the other side, E(z) binding at PRE2 was not affected by the Pho mutations in PRE1 and its signal at PRE2 of Pelican-*gt*-mut transgene embryos was comparable to that of the endogenous *gt* gene (**Figure 3.7**).

In accordance with the weak association of PRC2 with chromatin, H3K27me3 signals were close to background levels at endogenous *gt* and Pelican-*gt*-wt transgene in nc10-13 embryos. H3K27me3 signals increased across the *gt* region from nc10-13 to nc14b-nc15 consistent with the stable binding of PRC2 at this stage. However, H3K27me3 did not drastically increase through stage 10 (**Figure 3.7**). Deposition of H3K27me3 was a little stronger at PRE1 compared to PRE2 of the endogenous *gt* and Pelican-*gt*-wt transgene. Given that H3K27me3 signal at PRE1 of Pelican-*gt*-mut transgene was already close to background level at nc10-13, absence of Pho did not have an effect on the deposition of H3K27me3 at this stage (**Figure 3.7**). However, H3K27me3

signal sank to near the background levels at nc14b-nc15 and stage 10 embryos, due to the lack of Pho binding. Deposition of H3K27me3 at regions 6 and 9 (PRE2) of Pelican-*gt*-mut transgene was not affected by mutations of Pho sites (**Figure 3.7**).

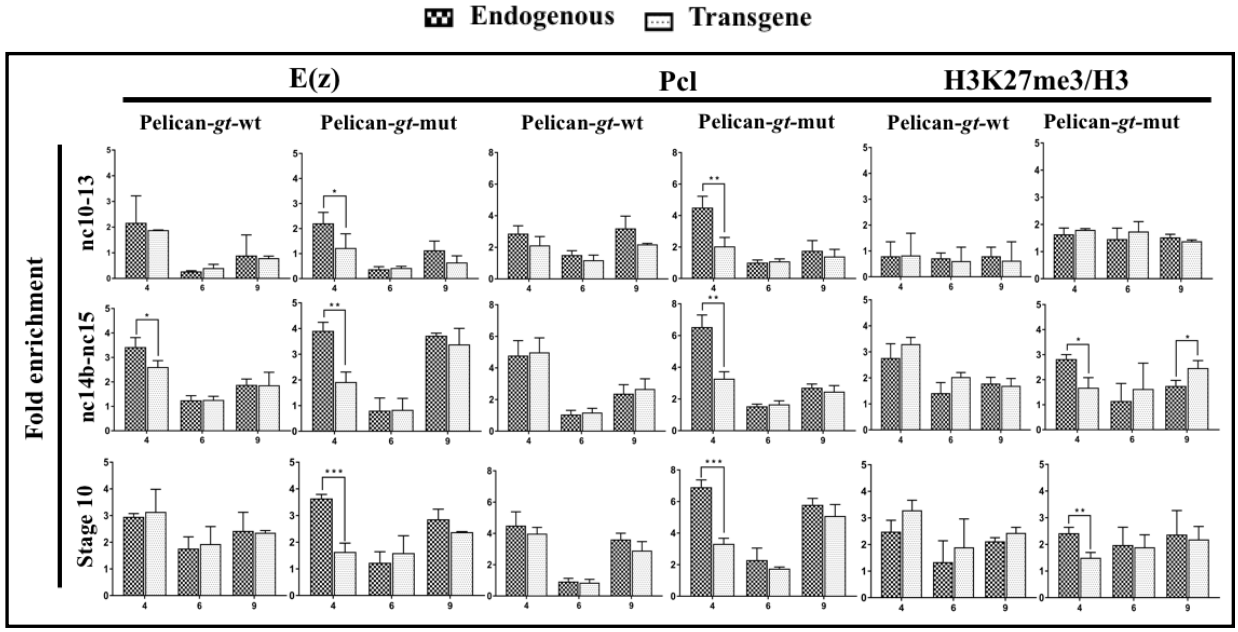

Figure 3.7- **Effects of mutant PRE1 Pho-Phol consensus binding sites on recruitment of PRC1, PRC2 and deposition of H3K27me3 at *gt*.** Time course ChIP-qPCR assays were performed as in figure 3.6, except using anti-E(z), anti-Pcl and anti-H3K27me3. ChIP signals from three biological replicates are presented as in figure 3.6.

Pcl, a substochiometric subunit of PRC2, is required for high levels of trimethylation of H3K27 at PcG target genes (Nekrasov et al., 2007). Pcl is detected at PRE1 of the endogenous gt and Pelican-gt-wt transgene as early as nc10-13. Pcl signal at endogenous and Pelican-*gt*-wt PRE1 increases further at nc14b-nc15 and stays at the same level at stage 10. Pcl binding at PRE2 is relatively lower than endogenous PRE1 at all developmental stages. Pcl recruitment at PRE2

slightly increases from nc10-13 to nc14b-nc15 and noticeably rises at stage 10 (**Figure 3.7**). Pcl binding to Pelican-*gt*-mut PRE1 was significantly reduced, but did not sink to background levels, as the result of absence of Pho. Binding of Pcl to PRE2 of Pelican-*gt*-mut transgene was not affected by the absence of Pho at PRE1 (**Figure 3.7**)

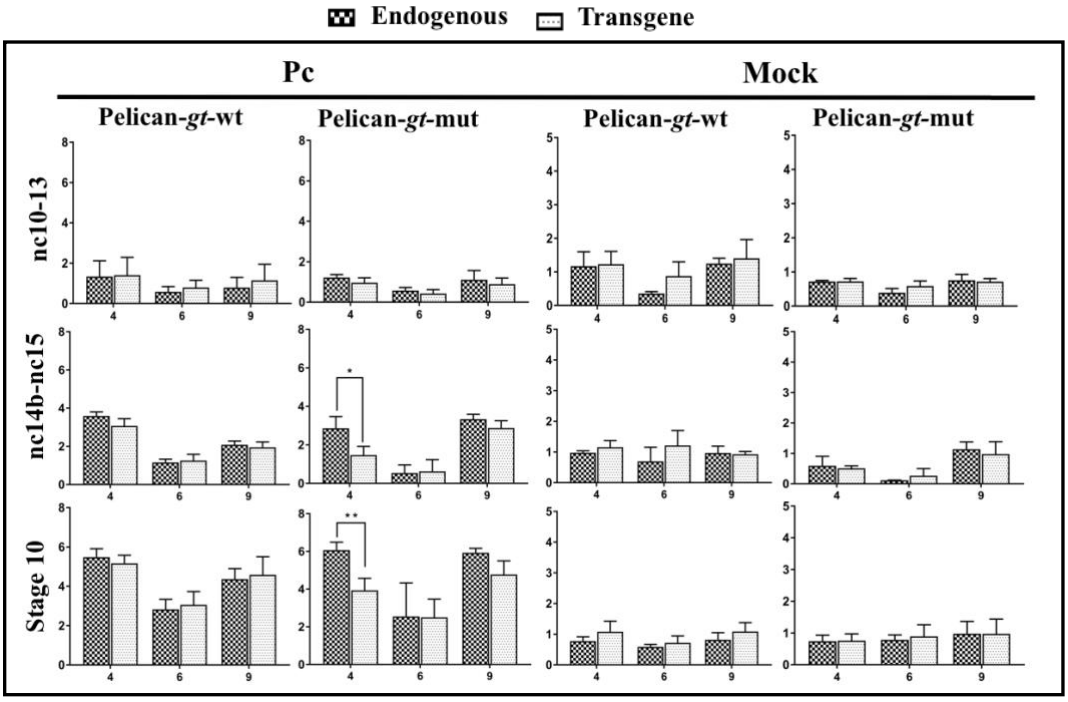

Figure 3.8- **Effects of mutant PRE1 Pho-Phol consensus binding sites on recruitment of PRC1 at *gt*.** Time course ChIP-qPCR assays were performed as in figure 3.6, except anti-Pc and IgG (mock) antibodies. ChIP signals from three biological replicates are presented as in figure 3.6.

Mock (rabbit IgG) was used as our negative control which gave a negative to mildy positive signals at all tested region in three embryonic stages (**Figure 3.8**). Concurrent with unstable binding of PRC2 and negative H3K27me3 signal, Pc signal was close to background level in nc10-

13 embryos. Binding of Pc at PRE1 of endogenous *gt* and Pelican-*gt*-wt transgene increased from nc10-13 to stage 10 (**Figure 3.8**).

Pc binding to PRE2 was weakly positive at nc14b-nc15 and increased drastically at stage 10 embryos. Pc signal dropped to background levels at PRE1 of Pelican-*gt*-mut transgene at nc14b-nc15. At stage 10, Pc signal was significantly reduced, but still well above the background, at PRE1 of Pelican-*gt*-mut. Binding of Pc at region 6 and PRE2 (region 9) of Pelican-*gt*-mut was comparable with endogenous *gt* and Pelican-*gt*-wt transgene (**Figure 3.8**). Furthermore, we examined the requirement for the presence of Pho on the binding of PRE binding proteins, Dsp1 and Spps, to PREs. Spps was weakly detected at PRE1 of the endogenous *gt* and Pelican-*gt*-wt transgene at nc10-13. Its signal increased through nc14b-nc15, but remained at the same level at stage 10. Spps signal stayed at background levels at PRE2 and region 6 across all time points, indicating the differential binding of Spps at two PREs (**Figure 3.9**). We surprisingly observed the significant reduction in Spps binding at Pelican-*gt*-mut PRE1 in all three embryonic stages. Reduction of Spps signals to background levels due to the absence of Pho, indicated that Spps binding to PRE1 was dependent on Pho (**Figure 3.9**).

Dsp1 signal increased at PRE1 of the endogenous *gt* and Pelican-*gt*-wt transgene from nc10-13 to nc14b-nc15 and decreased somehow at stage 10. Dsp1 was very weakly detected at PRE2 at nc10-13 and nc14b-nc15 embryos and did not increase to slightly over the background level till stage 10 (**Figure 3.9**). Despite Spps, Dsp1 binding to PRE1 was not affected by the mutations of Pho binding sites. Dsp1 bound to PRE2 of Pelican-*gt*-mut transgene at the same level of Pelican-*gt*-wt transgene and endogenous *gt* (**Figure 3.9**).

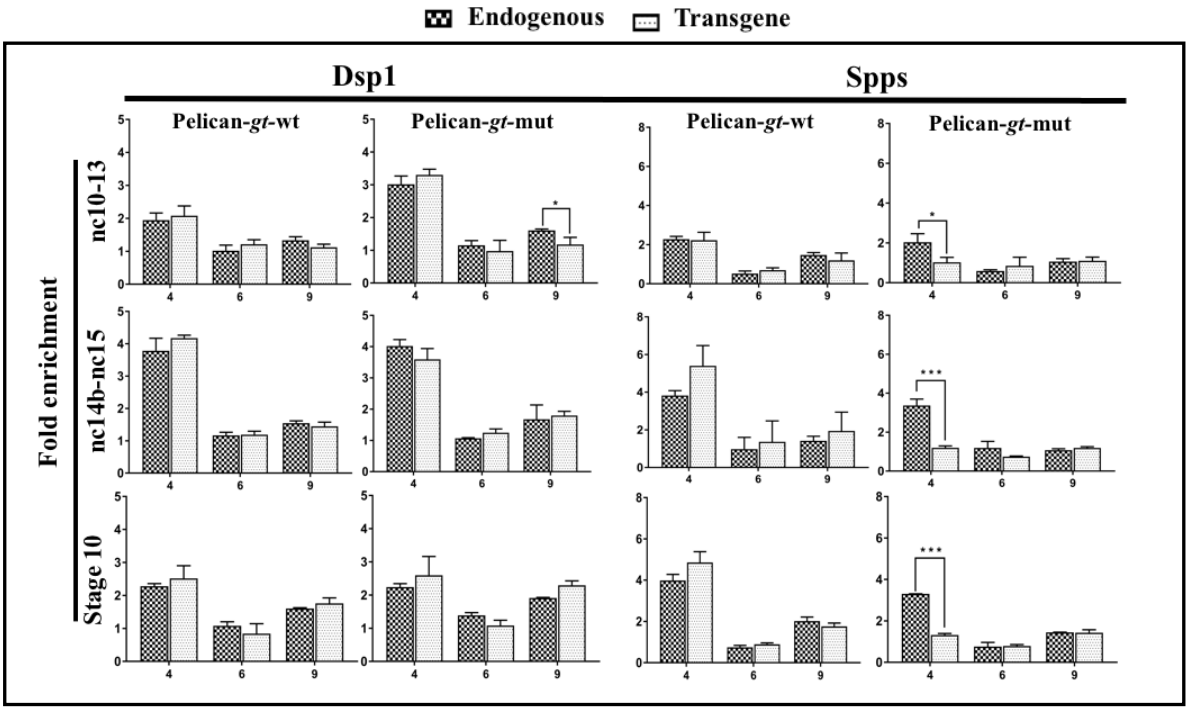

Figure 3.9- **Recruitment of PRE binding proteins, Dsp1 and Spps, to PRE1 was differentially affected by the absence of Pho.** Time course ChIP-qPCR assays were performed as in figure 3.6, except using anti-Dsp1 and anti-Spps antibodies. ChIP signals from three biological replicates are presented as in figure 3.6.

3.3 Discussion

3.3.1 Phol binding plays a minimal role in PRE activity

Several Drosophila genes have more than one PRE that may contribute to their PcG dependent transcriptional repression (Abed et al., 2013; Ahmad and Spens, 2018; Cunningham et al., 2010; DeVido et al., 2008; Ogiyama et al., 2018). However, the functional implication of having two or more PREs at a genomic locus is not clear and may not be consistent at all loci. Some may form the bases of chromatin loops that contribute to the establishment or maintenance of repressive domains (De et al., 2016; Eagen et al., 2017; Ogiyama et al., 2018). In other cases,

individual PREs may differently contribute to the regulation of associated genes and/or the establishment of H3K27me3 domains (Ahmad and Spens, 2018; Ogiyama et al., 2018).

Two predicted Pho/Phol binding sites were present within the delimited PRE1 (PRE1.1). Base pair substitution of the core sequences of binding sites resulted in the disruption of PRE activity of PRE1.1 fragment within the context of SD10 vector. Pho and Phol can recognize and bind to the same sequence of DNA (Brown et al., 2003; Grossniklaus and Paro, 2014). However, the genome wide analysis showed different genomic distributions of Pho and Phol. Although, ChIP-chip peaks for Pho correspond to PREs and colocalize with Sfmbt, a subset of Phol peaks are located outside PcG domains, within 300 bp of TSS of a number of active genes (Kahn et al., 2014). SD10 vector includes the *en* promoter region which serves as a candidate for Phol binding. However, the *en* promoter does not include a matching sequence with the core (GCCAT) or extended Pho/Phol binding sites (CGCCAT(T/A)TT), while the extended binding site has been proposed to permit binding of Phol near TSSs (Kahn et al. 2014). Furthermore, ChIP assays on SD10 transgene-containing embryos suggested that Phol is able to make a dimer with Sfmbt in the absence of Pho, which is counter to previous observations that Phol binds near TSSs independently of Sfmbt and other PcG proteins. Kahhn et al. (2014) reported that both Pho and Phol can form a dimer with Sfmbt, but Pho is favored and outcompetes Phol. The interaction of Pho and Phol with Sfmbt is mutually exclusive, as Phol binding to chromatin increases upon RNAi knockdown of Pho (Kahn et al., 2014). The presence of Phol-containing PhoRC is of interest because SD10 vector has been previously used to map PREs for *gt* and *en* (Cunningham et al., 2009; Alhaj Abed, 2013). Moreover, the activity of PRE1.1 was lost in SD10-PRE1.1-mut embryos, which indicates that Phol-containing PhoRC does not contribute to PcG-mediated repression.

ChIP experiments performed on Pelican-*gt*-wt and Pelican-*gt*-mut with anti-Phol showed that Phol binding to PRE1 of Pelican-*gt*-mut did not significantly reduce compared to that of Pelican-*gt*-wt. ChIP results from SD10-PRE1.1-mut and SD10-Δgt1 excludes the possibility of presence of unmutated Pho/Phol binding sites in PRE1.1. Level of Phol binding to PCR amplicons surrounding PRE1, regions 3 (+297, +560) and 5 (-362, -580), was lower than that of PRE1 (data not shown). The nearest upstream core Pho consensus site is located approximately 1 kb from *gt* TSS. Therefore, we propose that a region downstream of PRE1.1 (+12) and upstream of region 3 (+297) might contain the additional Phol binding sites. We noted that a core Pho/Phol consensus site (ATGG) is located at +137, which is well within the ±300 bp of TSS suggested as targets for Phol binding (Kahn et al., 2014). Whether Phol may contribute to the transcriptional activity of *gt* remains elusive.

3.3.2 PcG proteins dampen the transcription of transcriptionally active genes

PcG-mediated repression of *gt* is redundant with transcription factors, Hb and Kr, during early embryonic stages. Negre et al. (2006) showed the ectopic expression of *gt* in PcG mutant background in stage 10 embryos, which indicated that maintanence of *gt* repression is PcG-dependent in this stage. Transgene expression in Pelican-*gt*-mut and Pelican-*gt*-wt embryos was assessed by in situ hybridization using a *lac*Z probe. The absence of ectopic expression of *lac*Z in these embryos through stage 10 was indicative of PcG-mediated repression by PRE2 alone, although we cannot rule out possible contributions from additional unidentified weak PREs.

Surprisingly, RT-qPCR analysis of the *lac*Z expression from Pelican-*gt*-mut and Pelican-*gt*-wt embryos showed small but significant increase of *lacZ* transcription due to the mutations of Pho/Phol binding sites at nc10-13 and nc14b-pnc14b embryos. At stage 10, the increase in the *lac*Z expression from Pelican-*gt*-mut is less. Therefore, we propose that disruption of PRE1

activity results in a higher transcription level from the transgene in those cells in which *gt*, and the transgene, are active. Our data suggest that binding of PcG proteins to PRE1 of the active state of *gt* has a moderating effect on transcription even under conditions of transcriptional activation. Our observation is inconsistent with the related study which showed that deletion of four major PREs in *en-inv* domain did not result in changes of level or pattern of *en* expression (De et al., 2016).

Enderle et al. (2011) showed a significant positive correlation between the higher level of PcG binding and lower level of mRNA production and proposed a model in which PcG proteins reduce the transcriptional level once present at the active genes. Our findings of elevated *lac*Z expression upon mutation of Pho binding sites directly support their model.

TrxG proteins, which maintain the active state of their target genes, also act through PRE/TREs. One possibility may be that mutation of Pho binding sites results in TRE activity of PRE1.1.

3.3.3 Low signals of PcG proteins at PRE1 of the Pelican-*gt*-mut transgene

Pho is located at the base of the hierarchical model for the recruitment of PcG proteins. Pho directly recruits E(z) (Wang et al., 2004), which in turn methylates H3K27 (Jenuwein et al., 1998; Rea et al., 2000), providing a scaffold for the binding of Pc through its chromodomain (Cao et al., 2002; Fischle et al., 2003; Min et al., 2003; Wang et al., 2004). Mutation of Pho binding sites in PRE1 results in the reduced E(z), H3K27me3 and Pc signals at PRE1 of the mutant transgene. The reduction in Pc and E(z) binding and deposition of H3K27me3 at PRE1 of Pelican-gt-mut was statistically significant but the signals did not drop to the background level at nc14b-pnc14b and stage 10 embryos. Presence of E(z) and Pc in the absence of Pho can be explained through different possible mechanisms. 1) PREs, or their mammalian equivalent CGIs, act as nucleation sites for the recruitment of PcG proteins and formation of PcG domains. PRC2 can

spread from nucleation sites (PREs) and its stable binding and residence time on the chromatin decreases with distance from PREs (Oksuz et al., 2018). E(z) may spread from PRE2 or other unidentified weak PREs within *gt* regulatory region to region 6 and PRE1. Binding of E(z) and Pc as well as deposition of H3K27me3 at the mutant transgene increases across the time points along with the increase of the signals at PRE2. However, it is worth noting that signals of E(z), H3K27me3 and Pc at PRE1 is higher than region 6, which is closer to PRE2 (~3kb downstream of PRE2). 2) Low binding of PcG proteins at PRE1 of Pelican-*gt*-mut transgene, may also be attributed to the presence of weak unidentified PREs. 3) Another possible explanation for the weak PcG signals at PRE1 of Pelican-*gt*-mut is the recruitment by Phol-containing PhoRC complex. Sfmbt interacts with Scm through its SAM domain (Frey et al., 2016), and Scm plays the connecting role between RhoRC and PRC1 (Sfmbt-SAM: Scm-SAM: SAM-Ph) (Kim et al., 2005; Frey et al., 2016). 4) The alternative recruitment mechanisms and other PRE binding proteins which can bind PREs independently of Pho, such as Dsp1, may contribute to the low signals of PcG proteins. 5) Detection of PcG proteins, Pc, E(z), Pcl and Sfmbt, at mutant PRE1 may be due to the chromatin looping between PREs. The results from a number of studies suggest that PREs form chromatin loops with promoters, enhancers or silencers in the flanking regulatory DNA (Lanzuolo et al., 2007; Bantignies et al., 2011). Looping between *gt* PRE1 and PRE2 can explain the low signals of PcG proteins at PRE1 of Pelican-*gt*-mut, however, lack of Pho signal at PRE1 of the mutant transgene cannot be explained through chromatin looping between the two PREs.

3.3.4 Recruitment of PRE binding proteins, Dsp1 and Spps, is differentially dependent on the presence of Pho

The role played by PRE binding proteins, such as Dsp1 and Spps, in the recruitment of PcG proteins is less studied. In this study, we showed that Spps and Dsp1 only bind to PRE1 through nc14b-nc15 and their signals become weakly positive at PRE2 at stage 10. Our results

also suggested that Spps binding to PRE1 is dependent on Pho. Reduced binding of Pho has been reported in *Spps* mutants (Brown et al., 2018). Therefore, we suggest a cooperative binding of Spps and Pho to a subset of PREs. On the other hand, Dsp1 binding to the mutant PRE1 was not affected by the absence of Pho and Spps, indicating its independence of both Pho and Spps for PRE1 binding.

3.3.5 The temporal difference in PcG binding to *gt* PREs

Our time course ChIP assays showed that PhoRC and PRC2 bind to PRE1 as early as nc10-13, while their binding to PRE2 is at background level at nc10-13, and increases at later embryonic stages (nc14b-pnc14b or stage 10). The temporal difference of PcG binding to two PREs may be due to changes in chromatin accessibility through earlier to later time points. ATAC-seq analysis showed that PRE2 region of *gt* becomes accessible at late nc13 (Blythe and Wieschaus, 2016), which is in accordance with the weaker binding of PcG proteins in nc10-13 embryos in ChIP assays (Alhaj abed et al., 2018). Binding of pioneer transcription factors, Zelda (Zld) and Gaf, to enhancers and promoters at earlier developmental stages facilitate chromatin accessibility (Moshe and Kaplan, 2017). Zld is present at PRE1 and not PRE2 of nc14a-nc14b wt embryos (Alhaj abed, unpublished data), while Gaf has not been detected at *gt* locus in embryonic stage (Negre et al., 2006). Binding of Zld to PRE1 may well explain the difference in chromatin accessibility of PRE1 and PRE2 in earlier embryonic stage.

The general increase in the signals of PcG proteins from nc10-13 to nc14b-nc15 is in accordance with the lengthening of interphase. First nuclear divisions are very rapid, however, duration of interphase progressively increases slightly with each cycle in *Drosophila* embryos (Farrell and O'Farrell, 2014). The dramatic lengthening of interphase happens from nc13 to nc14 through which interphase increases from 20 minutes to 70 minutes, mostly because of the

lengthening of S phase and introduction of G2 at nc14 (Edgar and Farrell, 1989; Foe, 1989; Edgar and Farrell, 1990). Both changes result from downregulation of Cyclin-dependent kinase 1 (Cdk1), a cell cycle regulator, at this stage (Edgar and Farrell, 1990; Edgar et al., 1994; Farrell et al., 2012). Ezh2, mammalian homologue of E(z), was shown to be phosphorylated by CDK1 at threonine residues 345 and 487 (Kaneko et al., 2010). Phosphorylation of Ezh2 impedes the recruitment of PRC2 to chromatin and thus affects the global level of H3K27me3 at target loci (Chen et al., 2010). Based on the literature and our time-course ChIP results, we suggest that CDK1 is downregulated, and the first prolonged interphase occurs at nc14. Therefore, E(z) is not phosphorylated by CDK1 and is able to bind to chromatin, resulting in the deposition of H3K27me3 and recruitment of PRC1 at this stage.

CHAPTER 4:

WHAT DICTATES THE INITIAL RECRUITMENT OF POLYCOMB-GROUP PROTEINS TO A *DROSOPHILA* TARGET GENE?

4.1 Introduction

PcG proteins are essential regulators of development and differentiation that do not initiate transcriptional silencing but maintain the transcriptional repression of silenced genes by altering chromatin structure (Simon and Kingston, 2013). After initial recognition and binding to their repressed target genes, PcG proteins are able to maintain the transcriptional repression through an unlimited number of cell cycles. However, the mechanisms that these proteins initially distinguish between the active and repressed states of their target genes remain a major gap in our understanding of the PcG-mediated repression.

Klose et al. (2013) have proposed two models for the recruitment of PcG complexes to the PREs/CpG islands: instructive and responsive models. The instructive model suggests that biochemical interactions of sequence-specific transcription factors with PcG complexes may recruit these complexes to their transcriptionally repressed target promoters. Repressive transcription factors may directly target PcG proteins to PREs/CpG islands, or recruit co-repressors which make chromatin changes and facilitate binding of PcG proteins to their target genes.

Susceptibility of CpG islands, equivalent of *Drosophila* PREs, to acquire H3K27 methylation is dependent on the absence of binding sites for transcriptional activators (van Heeringen et al., 2014). Mutation in the binding site of the specificity protein 1 transcription factor (SP1) results in the recruitment of PcG proteins to CpG islands, indicating that binding of an

activating transcription factor leads to the exclusion of PcG proteins (Caputo et al. 2013). These observations suggest an alternative version of the instructive model proposing that activators or the co-activators can antagonize binding of PcG proteins.

The results of at least one study support the instructive model in *Drosophila* (Kehle et al., 1998). Hb, a repressive transcription factor of Hox genes, recruits dMi-2 as a co-repressor (Kehle et al., 1998). dMi-2 is the ATPase subunit of the NuRD histone deacetylase complex with nucleosome remodeling activity. Derepression of *Ubx* was observed in *hb* and PcG mutant embryos and became more extensive in the absence of *dMi-2* (Kehle et al., 1998). The latter study suggested a linking role for dMi-2 between Hb and the recruitment of PcG proteins. Hb-dMi-2 complex may directly recruit PcG proteins to their target genes. On the other hand, the nucleosome remodeling and deacetylation activity of NurD complex may result in chromatin changes leading to the recruitment of PcG proteins.

The responsive model proposes that PcG complexes sample the chromatin environment at permissive chromatin sites, e.g. PREs or CGIs, and make weak association with their targets irrespective of their transcriptional state. Stable targeting of PcG complexes is then based on the transcriptional state of their target genes and the molecular features of the chromatin in repressed genes are compatible with, or stabilize, recruitment of PcG complexes. Riising et al. (2014) demonstrated that blocking transcription with chemical inhibitors, DRB and Triptolide, resulted in the recruitment of PRC2 and deposition of H3K27me3 at three different mammalian PcG target genes, while PRC2 recruitment was lost upon DRB removal.

By construction of embryos with a transcriptionally inert *gt* transgene, in a background in which endogenous *gt* is transcriptionally active, we experimentally tested the instructive and responsive models for the recruitment of PcG proteins and provided promising evidence to support

the instructive model. We find that recruitment of PG proteins at *gt* is not dependent on the transcriptional state and, therefore, follows the instructive model. We further showed that the default state of the chromatin is compatible with the strong binding of PcG proteins and the binding of activating transcription factors makes the chromatin less favorable for the recruitment of PcG proteins.

4.2 Results

4.2.1 Establishment of *Drosophila* genetic system in which *gt* is uniformly expressed

Based on the knowledge that *gt* is a PcG target gene (Pelegri and Lehmann, 1994) and the well-characterized maternal regulation of it by specific activating and repressive transcription factors, we developed the *bcd osk tsl* genetic system.

In embryos from *bcd osk tsl* females, *gt* is uniformly repressed due to the ubiquitous expression of maternal Hb, a repressor of *gt*, and the mutant allele of *bcd*, the activator of the anterior stripe of *gt* (Abed et al., 2018). In these embryos, zygotic *hb* is not activated as the result of *bcd* and *tsl*. Consequently, maternal Hb is degraded at cellular blastoderm stage and not replaced by zygotic Hb. In the absence of zygotic Hb, maintenance of *gt* repression requires PcG activity (**Figure 4.1**). Maternal Cad, an activator of the posterior stripe of *gt*, is uniformly distributed in embryos from *bcd osk tsl* females, due to the absence of Bcd. Expression of zygotic *cad* is dependent on the terminal system, which is inactive in this genetic system due to *tsl*.

To produce embryos in which *gt* is ubiquitously expressed in all cells of the embryos, the levels of maternal Hb in *bcd osk tsl* background was knocked down by introducing maternally expressed Gal4 driver in conjunction with TRiP *UAS-shRNA-hb* **(Figure 4.2)**. These female flies will be referred to as HbKD *bcd osk tsl*.

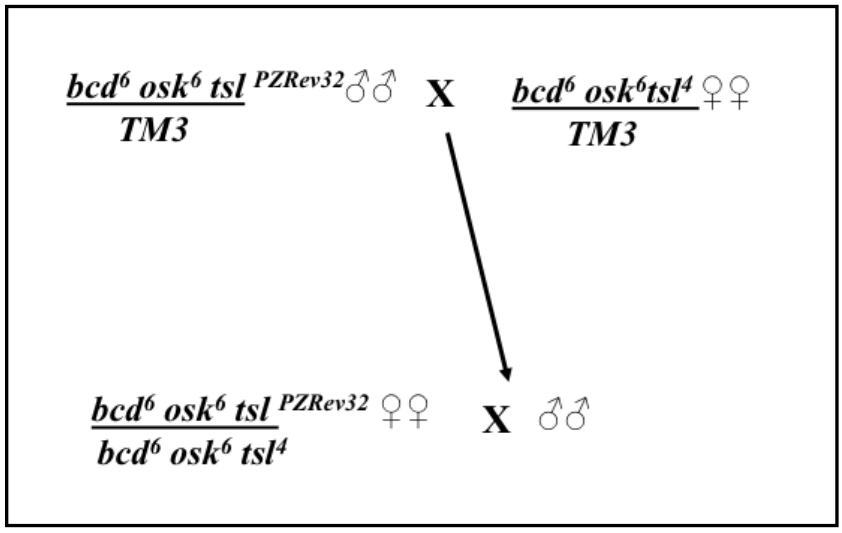

Figure 4.1- **Genetic crosses to produce embryos in which *gt* is ubiquitously repressed.**

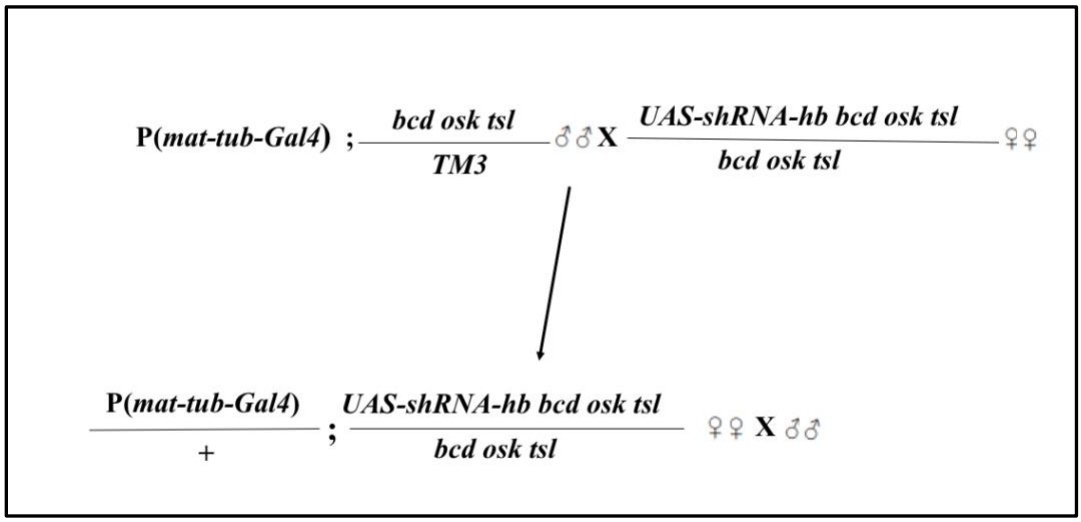

Figure 4.2- **Genetic crosses to produce embryos in which *gt* is ubiquitously expressed.**

The first few population cages of HbKD *bcd osk tsl* flies were set up at 25 °C and failed to show efficient knock down level of maternal Hb in immunostaining and ChIP assays (data not shown). Transferring the cage to 27 °C for at least one day before the collection, helped reach an efficient Gal4-driven knock down of maternally expressed Hb. The embryos from HbKD *bcd osk tsl* females, express minuscule amount of maternal Hb, due to the activity of *UAS-shRNA-hb*. Zygotic Hb is absent in this genetic system due to *bcd* and *tsl*. Similar to *bcd osk tsl* embryos, maternal *cad* is uniformly translated in HbKD *bcd osk tsl* embryos, and zygotic *cad* is absent due to *tsl* (**Figure 4.3**).

Using RT-qPCR analysis, the level of *gt* and *Kr* expression in sorted nc13-14a HbKD *bcd osk tsl* and *bcd osk tsl* embryos was investigated (**Figure 4.4**). Expression of *gt* in *bcd osk tsl* embryos increased from ~8% to 295% of wt embryos upon the knock down of maternal Hb.

Kr, a repressor of *gt_(-3)* enhancer, is expressed at ~50% and ~70% of its wt expression level in *bcd osk tsl* and HbKD *bcd osk tsl* embryos, respectively (**Figure 4.4**). The mildly increased expression level of *Kr* in HbKD *bcd osk tsl* embryos, may be due to the reduced level of Hb. Hb represses and activates *Kr* at high and low concentrations, respectively.

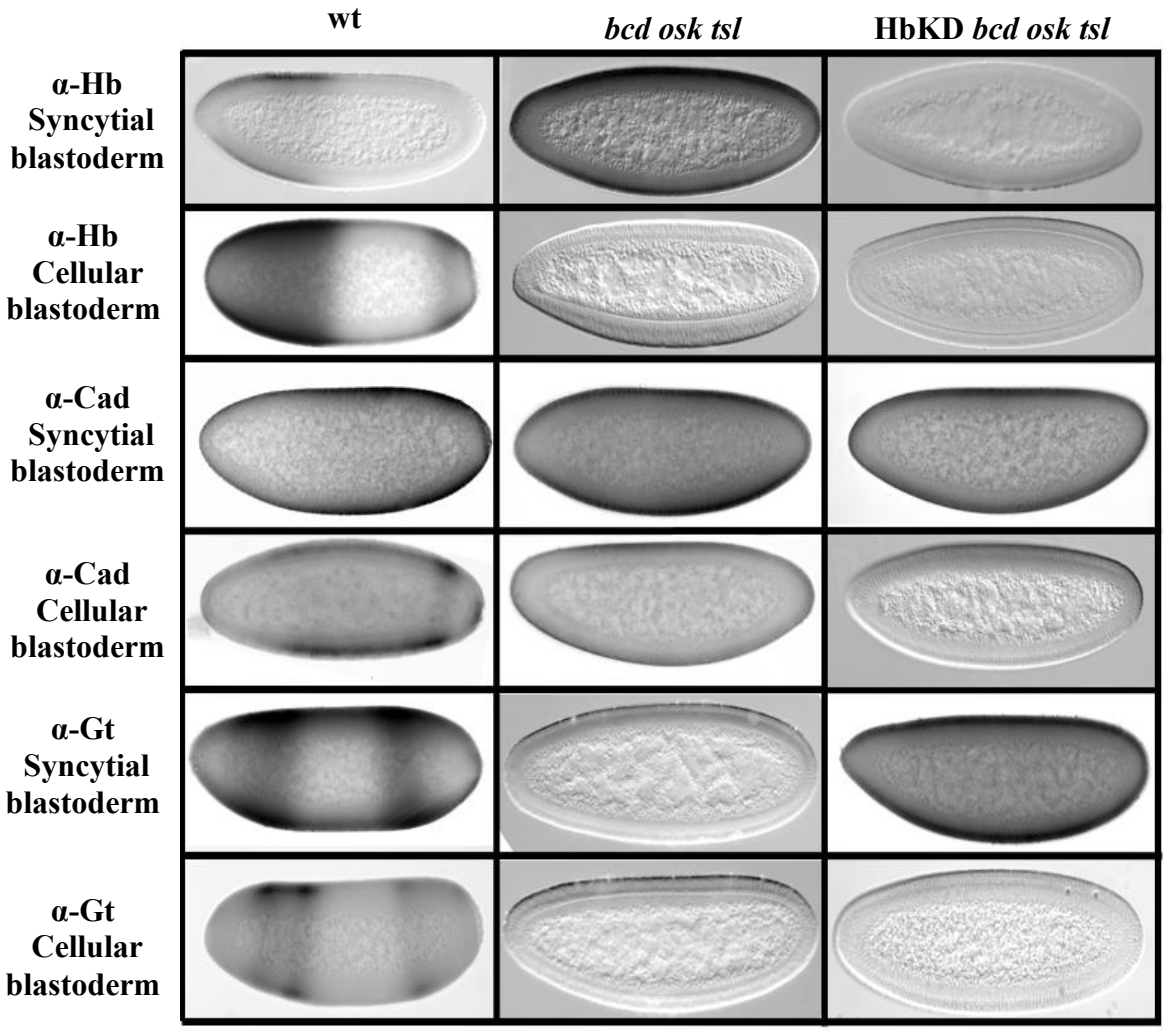

Figure 4.3- **Characterization of *bcd osk tsl* and HbKD *bcd osk tsl* embryos.** Embryos from *bcd osk tsl* and HbKD *bcd osk tsl* females immunostained with anti-Hb, anti-Cad and anti-Gt antibodies. Embryonic stages and the antibodies used are stated on the left.

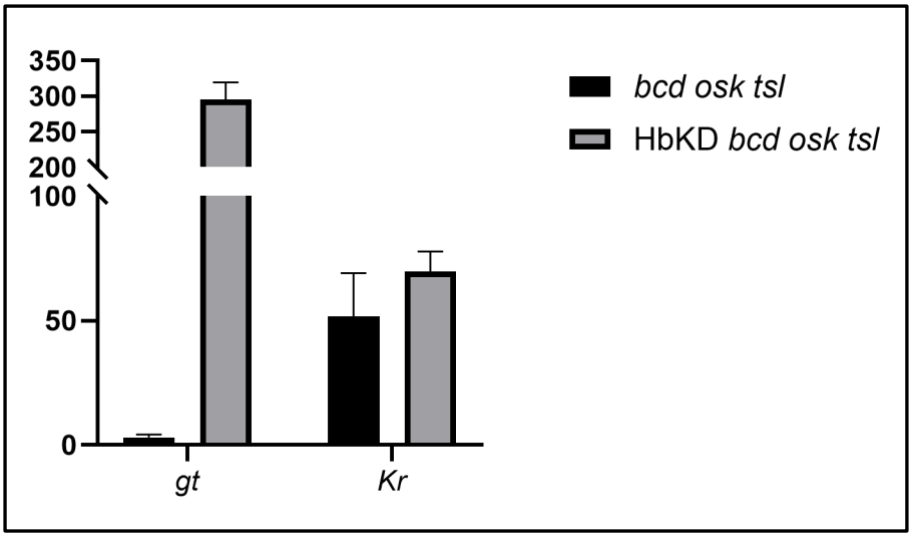

Figure 4.4- **Expression of *gt* and *Kr* in nc13-14a embryos derived from HbKD; *bcd osk tsl* females.** Genotype of embryos is indicated above. RT-qPCR signals were normalized to *rp49* transcript. Y-Axis shows the relative expression of *lac*Z to wt embryos. Error bars show the standard deviation for three biological replicates.

4.2.2 Binding of transcriptional repressor, Hb, and activator, Cad, at *gt*

Early expression of *gt* is regulated by maternally expressed transcription factors which bind to early acting enhancers, *gt_(-3)* and *gt_(-10)*. Hb and Cad are known repressor and activator of *gt_(-3)* enhancer, respectively.

Previous time-course ChIP assays on *bcd osk tsl* embryos showed the strong binding of both Hb and Cad at PCR region 6, which amplifies the *gt_(-3)* enhancer, in sorted nc10-12 embryos (Abed et al., 2018). Signals of Hb and Cad increased more through nc13 embryos, and dropped down to near background levels at nc14a-nc14b (Abed et al., 2018).

To confirm the efficient knock down of maternal Hb in *bcd osk tsl* embryos, binding of Hb and Cad to *gt_(-3)* enhancer in sorted nc13 HbKD *bcd osk tsl* embryos was compared to that in *bcd osk tsl* embryos. As expected, Cad binds to region 6 of HbKD *bcd osk tsl* embryos at the same

level shown for *bcd osk tsl* embryos. Binding of Hb is greatly reduced at region 6 of HbKD *bcd osk tsl* embryos, showing ~79% decrease compared to that in *bcd osk tsl* counterparts (**Figure 4.5**B). The levels of maternally expressed Hb and Cad decrease to near-backgound levels at region 6 in nc14b HbKD *bcd osk tsl* and *bcd osk tsl* embryos, consistent with the timing of degradation of maternally deposited Cad and Hb (**Figure 4.5**C). However, Cad signal at the *gt* promoter region of nc14b HbKD *bcd osk tsl* and *bcd osk tsl* embryos, remained at the same level as that of nc13 embryos, raising the possibility for the unspecific binding of Cad to the promoters, which are the targets for binding of many known proteins. To investigate the latter possibility, we looked at Cad signal at the promoter of the *miR-9a* gene, at which Cad has been shown to be absent (modENCODE). The strong Cad signal at *miR-9a* promoter of nc14b HbKD *bcd osk tsl* embryos, suggested the unspecific binding of affinity-purified Cad antibody to the promoter region of genes.

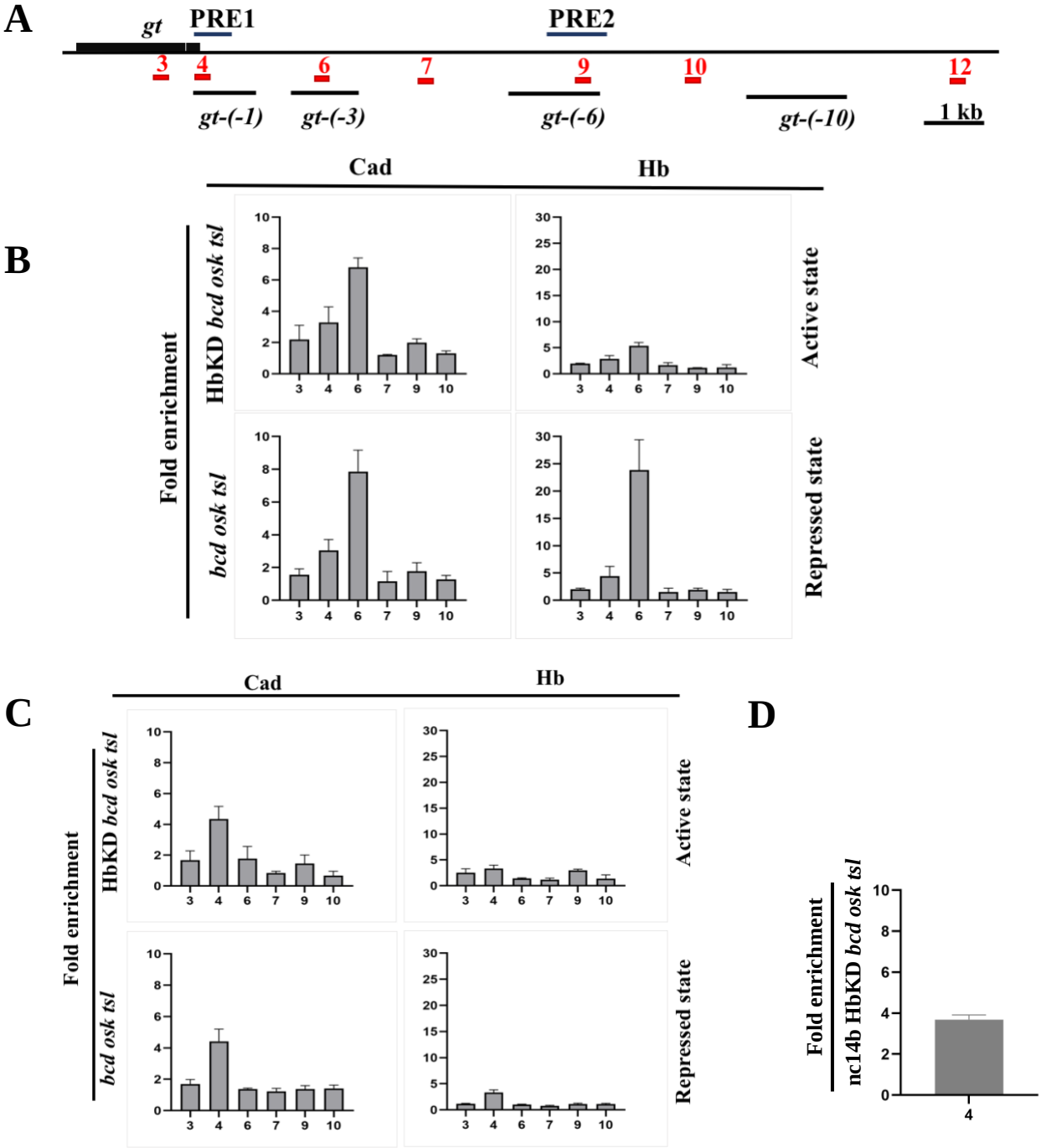

Figure 4.5- **Binding of activator Cad and repressor Hb at *gt* upstream regulatory region of HbKD *bcd osk tsl* and *bcd osk tsl* embryos.** (**A**) Schematic map of *gt* genomic region containing two PREs, shown as blue bars, and four enhancers, shown as black bars. PCR regions are shown as red bards (3-10) along the *gt* map. ChIP-qPCR performed with anti-Hb and anti-Cad antibodies in sorted nc13 (**B**) and nc14b embryos (**C**). (**D**) Cad signal at the promoter of *miR-9a* in nc14b embryos. Genotype of embryos is indicated on the left. ChIP signals are presented as fold enrichment to *Pka-C1* on 2L, shown at 1. Error bars show the standard deviation for three biological replicates.

4.2.3 Differential recruitment of PcG proteins to repressed versus active *gt*

In order to determine how PcG proteins differentially bind an active versus repressed *gt* gene, recruitment of PcG proteins and deposition of histone marks at *gt* gene in embryos from *bcd osk tsl* (reprssed state) and HbKD *bcd osk tsl* (active state) females were compared. Previous studies of *bcd osk tsl* embryos have demonstrated increased levels of H3K27me3 and stable binding of PRC1 and PRC2 at nc14b embryos following the weak association of PRC1 and PRC2 complexes in earlier stages, nc10-12 to nc14a (Alhaj Abed et al., 2018). Therefore, nc14b stage was targeted to look at the possible differences in binding of PcG proteins at active and reprssed states of *gt*.

Both subnuits of PhoRC, Pho and Sfmbt, are strongly detected at region 4 of both active and repressed states of *gt*. Pho binding at region 9 shows a clear reduction in the active state compared to the repressed one, decreasing from seven-fold enrichment in *bcd osk tsl* embryos to approximately four-fold enrichment in HbKD *bcd osk tsl*. Sfmbt signals also showed a similar trend, decreasing from six-fold enrichment at PRE2 of the repressed state to half at that of the active state. Binding of Pho and Sfmbt at other tested regions (3-12) of active state was roughly comparable to that of the repressed *gt* (**Figure 4.6**).

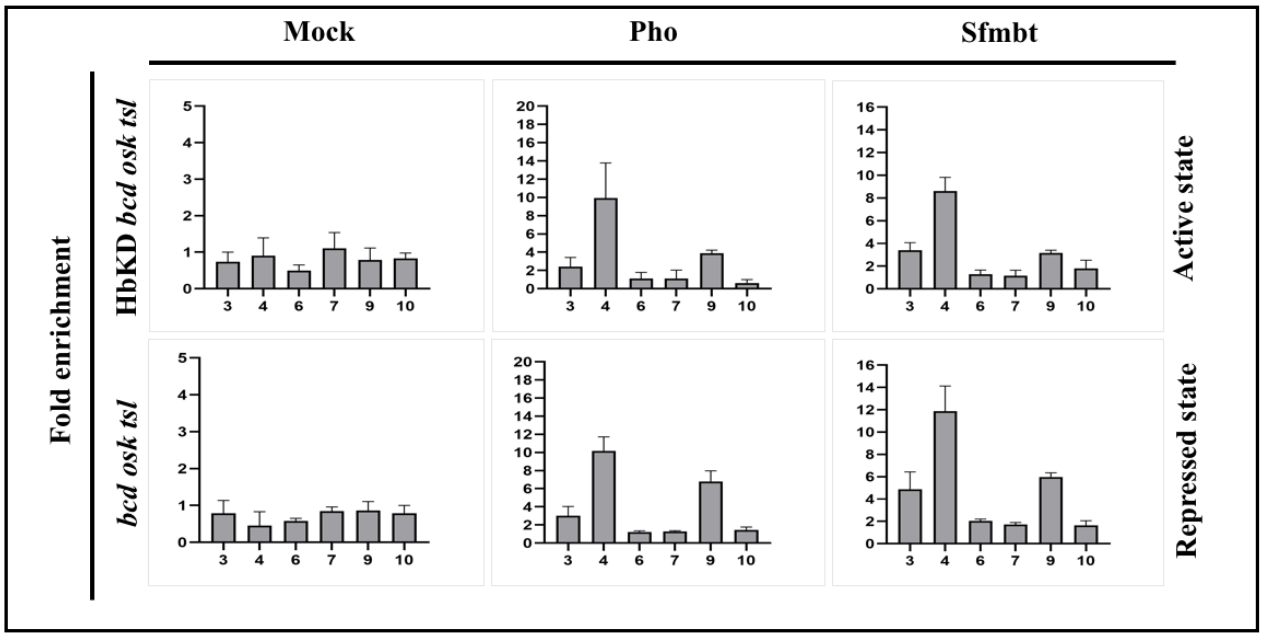

Figure 4.6- **ChIP-qPCR performed in sorted nc14b embryos with IgG (mock), anti-Pho and anti-Sfmbt antibodies, as indicated above.** Genotype of embryos is indicated on the left. ChIP signals from three biological replicates are presented as in figure 4.5.

To detect PRC2 and PRC1, anti-E(z) and anti-Pc antibodies were used, respectively (**Figure 4.7**). E(z) signal is detected above background at all tested *gt* regions of *bcd osk tsl* embryos, with strong peaks at regions 3 and 4. PRC2-dependent H3K27me3 signal is well above the background at most tested regions of the repressed state, with peaks of approximately three-fold enrichment at regions 3, 4 and 9. Accordingly, strong Pc signals were detected at PRE1 and PRE2 of the repressed state.

E(z) was detected at lower levels at most *gt* regions in the active state compared to the repressed state. In concert with the decreased presence of E(z), deposition of H3K27me3 is mostly close to or below the background level at almost all tested regions of the active state with a small peak at PRE1, slightly above two-fold enrichment. Concurrent with the significant reduction of H3K27me3 in the active state, Pc signal is slightly above the background in most *gt* regions of the

active state (**Figure 4.7**). Pcl, a substochiometric subunit of PRC2, was strongly detected at regions 4 and 9 of the repressed state. Although Pcl binding to the PRE1 of the active state is higher than that of the repressed state, its binding to PRE2 was greatly reduced to two-fold enrichment (**Figure 4.7**). High level of Pcl binding at PRE1 of the active gene, while E(z) signal was slightly above the background at the same region, indicates the recruitment of Pcl independently of the canonical PRC2 at this region.

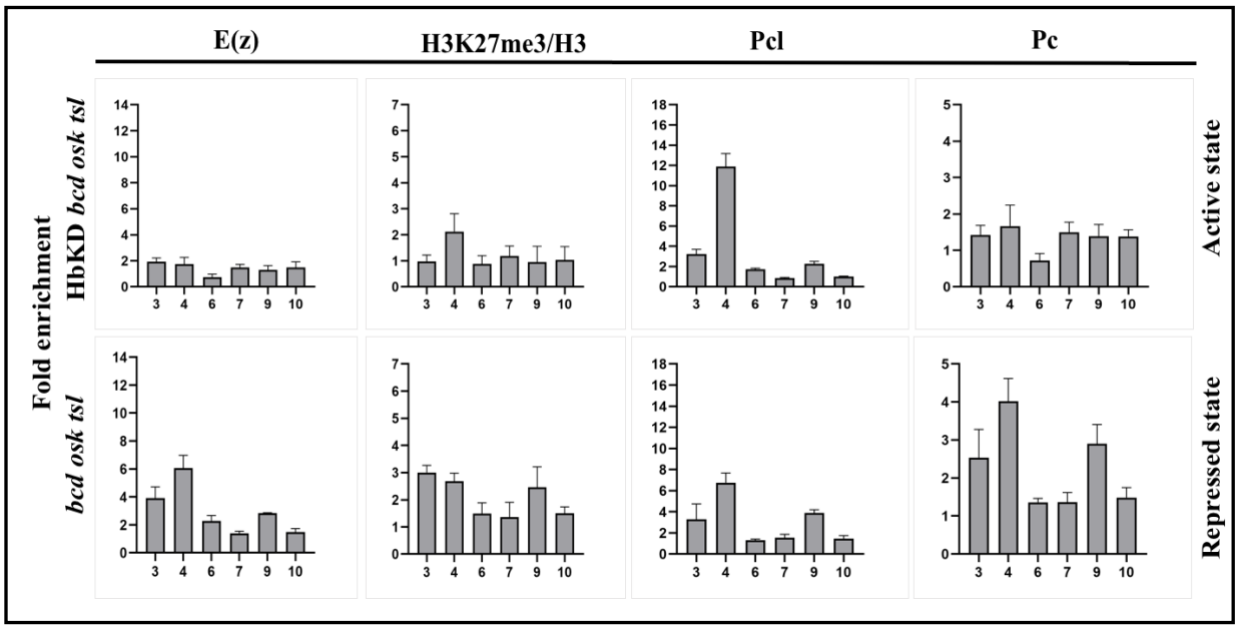

Figure 4.7- **ChIP-qPCR in sorted nc14b embryos with anti-E(z), anti-H3K27me3, anti-Pc and anti-Pcl antibodies, as indicated above.** Genotype of embryos is indicated on the left. ChIP signals from three biological replicates are presented as in figure 4.5.

H3K4me3, a histone mark associated with the promoter of active genes, was strongly enriched at the promoter region of the active state. However, its signal sinks to slightly above the background level at the promoter of the repressed *gt* (**Figure 4.8**). H3K27ac is another histone

mark deposited at promoters and enhancers of actively transcribed genes (Wang et al., 2008; Karlić et al., 2010). H3K27ac was shown to be mutually exclusive with H3K27me3 (Tie et al., 2009). Signals for H3K27ac are negative at almost all tested regions of repressed *gt* with a roughly two-fold peak at region 4. On the other hand, deposition of H3K27ac was weakly positive at most regions (6, 7, 9 and 10) of active *gt* with a relatively strong peak at region 4 (**Figure 4.8**). RNAPII S5P is associated with transcription initiation. ChIP assays showed the presence of RNAPII S5p at the promoter region of both active and repressed states, however, RNAPII S5p signal was slightly lower at the promoter of the repressed *gt* (**Figure 4.8**). Detection of RNAPII S5P along with the absence of H3K4me3 at the promoter region of the *bcd osk tsl* embryos, may indicate the onset of transcription initiation that failed to progress to the downstream active histone mark.

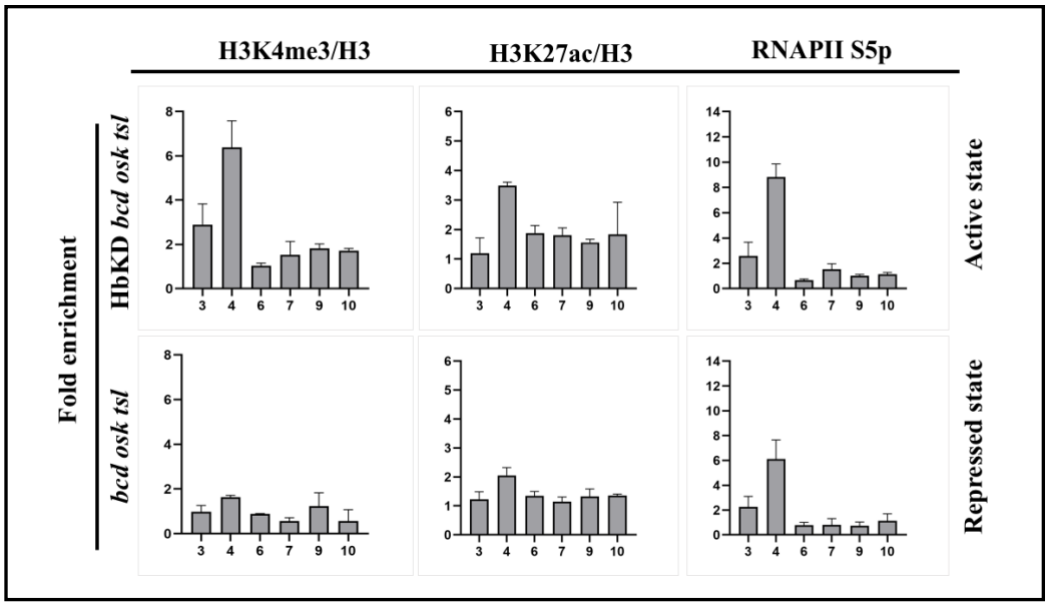

Figure 4.8- **ChIP-qPCR in sorted nc14b embryos with anti-H3K4me3, anti-H3K27ac, anti-RNAPII S5p antibodies, as indicated above.** Genotype of embryos is indicated on the left. ChIP signals are presented as fold enrichment to region 12, on X chromosome, shown at 1. Error bars show the standard deviation for three biological replicates.

We investigated binding of Dsp1, Spps and Phol, PRE binding proteins, in the active versus repressed states. Surprisingly, Spps robustly binds to PRE1 of the active *gt*, while it is weakly detected at the same region of the repressed state (**Figure 4.9**). Spps does not bind to PRE2 at both states of *gt*. Dsp1 binds more strongly to PRE1 compared to PRE2 and its binding level to both PRE1 and PRE2 is comparable at both *gt* states (**Figure 4.9**).

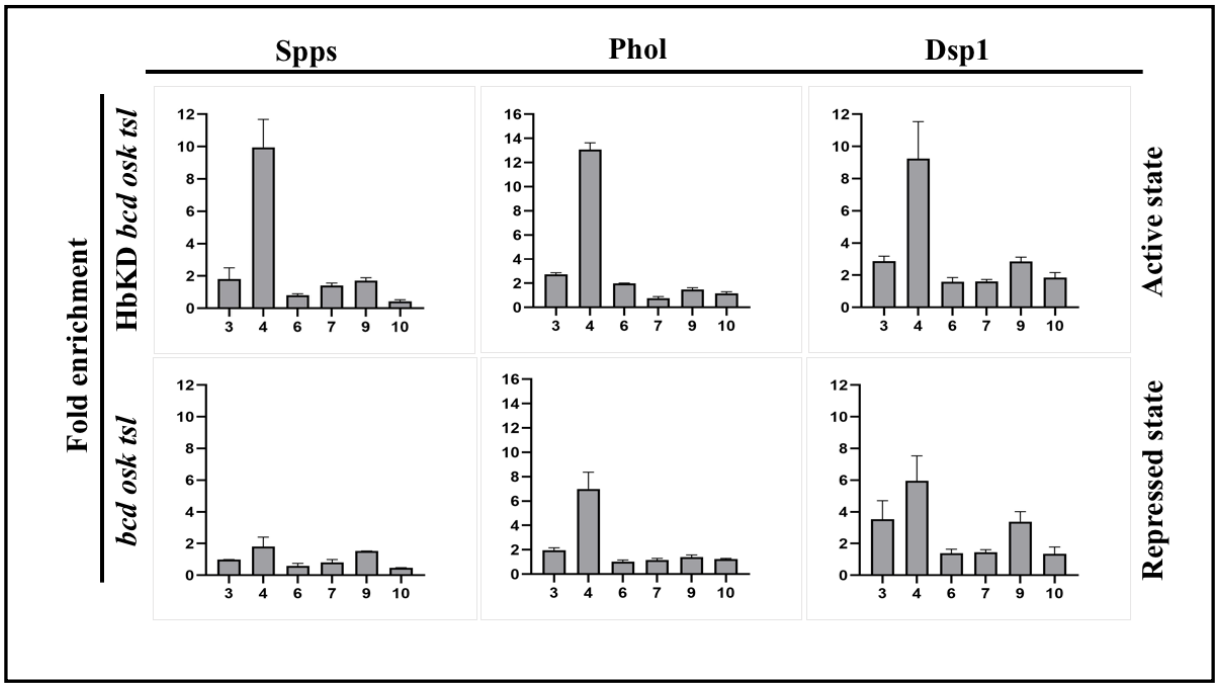

Figure 4.9- **ChIP-qPCR in sorted nc14b embryos with anti-Spps, anti-Phol and anti-Dsp1 antibodies, as indicated above.** Genotype of embryos is indicated on the left. ChIP signals from three biological replicates are presented as in figure 4.5.

Phol, a subunit of PhoRC, binds to PRE1 of both active and repressed *gt.* Nevertheles, it is not present at other tested regions of *gt*, including region 9 (PRE2) at both states (**Figure 4.9**).

Comparing the binding of PcG proteins to the active and repressed states of *gt*, we showed that Pho, Sfmbt and Pcl bind at PRE1 of the active state at levels comparable or higher than those of the repressed state. However, their binding to PRE2 was strongly reduced in the active state. Beltran et al. (2016), proposed that nascent RNAs and nucleosomes compete to bind to PRC2. They further showed that association of PRC2 with chromatin antagonizes its interaction with RNA and release of PRC2 from chromatin and RNA, increases its binding to the RNA and nucleosomes, respectively. The latter observation suggests that RNA transcripts serve as an obstruction to PRC2 binding to the actively transcribed genes, therefore, PRC2 is released to bind to PREs/CpG islands in the absence of transcription. Using RT-qPCR analysis, we investigated the possibility of transcription from the upstream regulatory region of *gt* in three different genotypes; wt *Oregon-R*, *bcd osk tsl* and HbKD *bcd osk tsl* (**Figure 4.10**). The results showed the absence of transcription in all tested upstream regulatory regions of *gt* (4, 6, 7, 9, 10), compared to region 3, which is located within the transcribed body of gene. Therefore, the possibility of reduced binding of PRC1 and PRC2 at PRE2 due to the presence of transcripts was excluded.

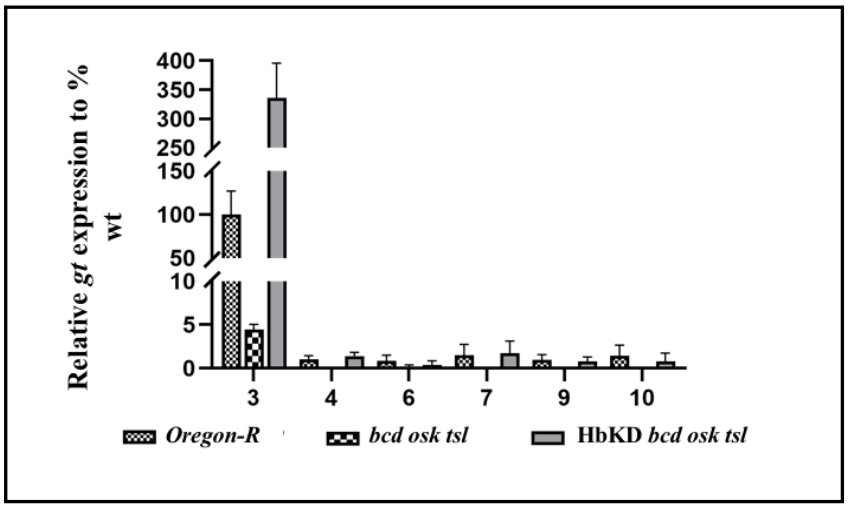

Figure 4.10- **Upstream regulatory region of *gt* is not transcribed in *bcd oks tsl, Oregon-R* and HbKD *bcd osk tsl embryos*.** Genotype of embryos are indicated below. RT-qPCR signals on nc13-14a embryos were normalized to *rp49* transcript. Y-Axis shows the relative expression of *lac*Z to *Oregon-R* embryos. Error bars show the standard deviation for three biological replicates.

4.2.4 Recruitment of PcG proteins to the transcriptionally inert *gt* transgene

A transcriptionally inert *gt* transgene was generated by introducing base pair substitutions in the TATA box, Inr and DPE of the *gt* fragment in the Gateway entry clone (**Figure 4.11A**). The *gt* fragment was then inserted into a modified Pelican reporter vector (**Figure 3.3**). The Pelican-*gt*-promoter mutant (pm) transgene was integrated into the attP40 docking site. Sequencing of the genomic DNA from adult flies confirmed the presence of the Pelican-*gt*-pm transgene in the latter flies.

Transgene expression in Pelican-*gt*-pm was assessed by in situ hybridization of the resulting embryos with a *lac*Z probe. Mutations in the promoter region of the *gt* transgene, resulted in the absence of *lac*Z expression from Pelican-*gt*-pm embryos (**Figure 4.11B**). The effect of the promoter mutations was further validated by RT-qPCR (**Figure 4.11C**), showing that the level of

*lac*Z expression in these embryos was the same as that in wt flies, which do not have the *lac*Z gene.

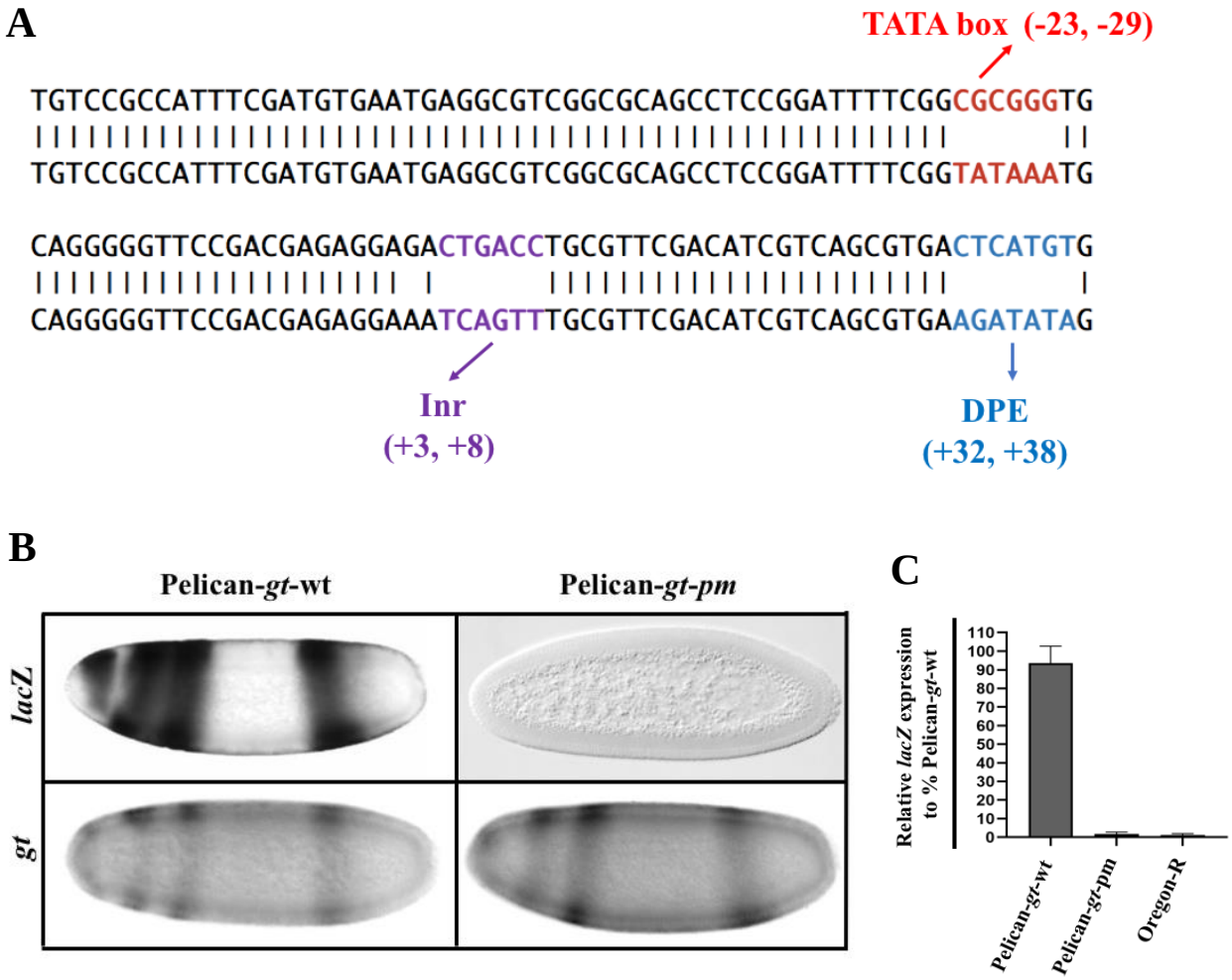

Figure 4.11- **Absence of *lac*Z expression in embryos from Pelican-*gt*-pm flies.** (**A**) Sequence alignment of the wild type promoter region containing TATA box, Inr and DPE sites with mutated promoter sequences; coordinates of the mutated sites are indicated above. (**B**) In situ hybridization of the cellular blastoderm Pelican-*gt*-wt and Pelican-*gt*-pm embryos by *lac*Z and *gt* probes, as indicated on the left. Genotype of embryos is indicated above. (**C**) *lac*Z expression in nc13-14a wt and Pelican-*gt*-pm embryos. Y-Axis shows the relative expression of *lac*Z to Pelican-*gt*-wt embryos. RT-qPCR signals were normalized to *rp49* transcript. Error bars show the standard deviation for three biological replicates.

4.2.4 Binding of repressor, Hb, and activator, Cad, to the transcriptionally *gt* inert transgene

In order to investigate the mechanisms that PcG proteins initially distinguish between the active and repressed states of their target genes, we constructed embryos with a transcriptionally inert *gt* transgene in a background in which endogenous *gt* is transcriptionally active. In the resulting embryos, the same transcription factors bind to the *gt* enhancers of the transcritionally inert *gt* transgene and trascriptionally active endogenous *gt* gene. Therefore, the only differentiating factor between the endogenous and transgene *gt* would be the transcriptional state.

To determine if binding of repressor, Hb, and activator, Cad, to *gt_(-3)* enhancer of the transcriptionally inert transgene was affected by the mutations of the promoter region, we looked at the binding pattern of these transcription factors in ChIP assays (**Figure 4.12**). In sorted nc13 embryos, binding of Cad and Hb to regions 3 and 4 of the transcriptionally inert transgene was reduced compared to the endogenous *gt* and control transgene. Nonetheless, their binding to *gt_(-3)* enhancer was not affected by the mutations introduced in the promoter region (**Figure 4.12**).

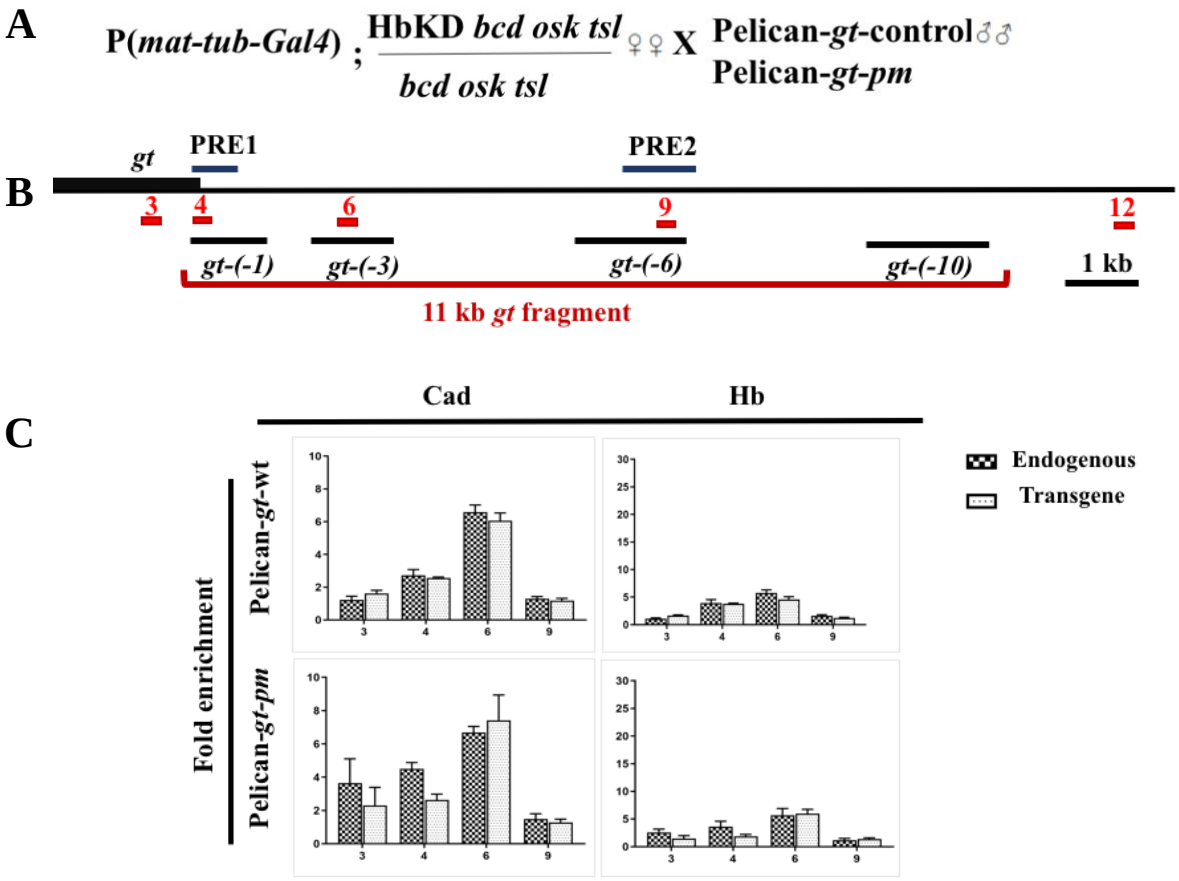

Figure 4.12- (**A**) Genetic crosses performed to obtain the embryos with the transcriptionally inert transgene in the active *gt* background. (**B**) Schematic map of *gt* genomic region. PCR regions are shown as red bards (3-10) along the *gt* map. PRE1 and PRE2 are represented as blue bars above. (**C**) ChIP-qPCR in sorted nc13 embryos with IgG (mock), anti-Cad and anti-Hb antibodies, as indicated above. Genotype of embryos is indicated on the left. ChIP signals from three biological replicates are presented as in figure 4.5.

4.2.5 Binding of PcG proteins and deposition of histone marks to the inert *gt* transgene in the transcriptionally active background

In order to investigate if the inert *gt* transgene would be distinguished as active or repressed by PcG proteins in the active endogenous *gt* background, we compared the recruitment of PcG proteins to the transcriptionally inert and control transgenes.

Pho and Sfmbt bind to the control *gt* transgene at the approximately same level as the endogenous *gt* (**Figure 4.13**). However, their binding to the inert transgene showed a different pattern. Surprisingly, binding of both proteins were greatly reduced at region 3 and 4 of the inert transgene, while their binding to region 6 and 9, was similar to that of endogenous *gt* and control transgene (Pelican-*gt*-wt), which resembled binding of these proteins to the active *gt*.

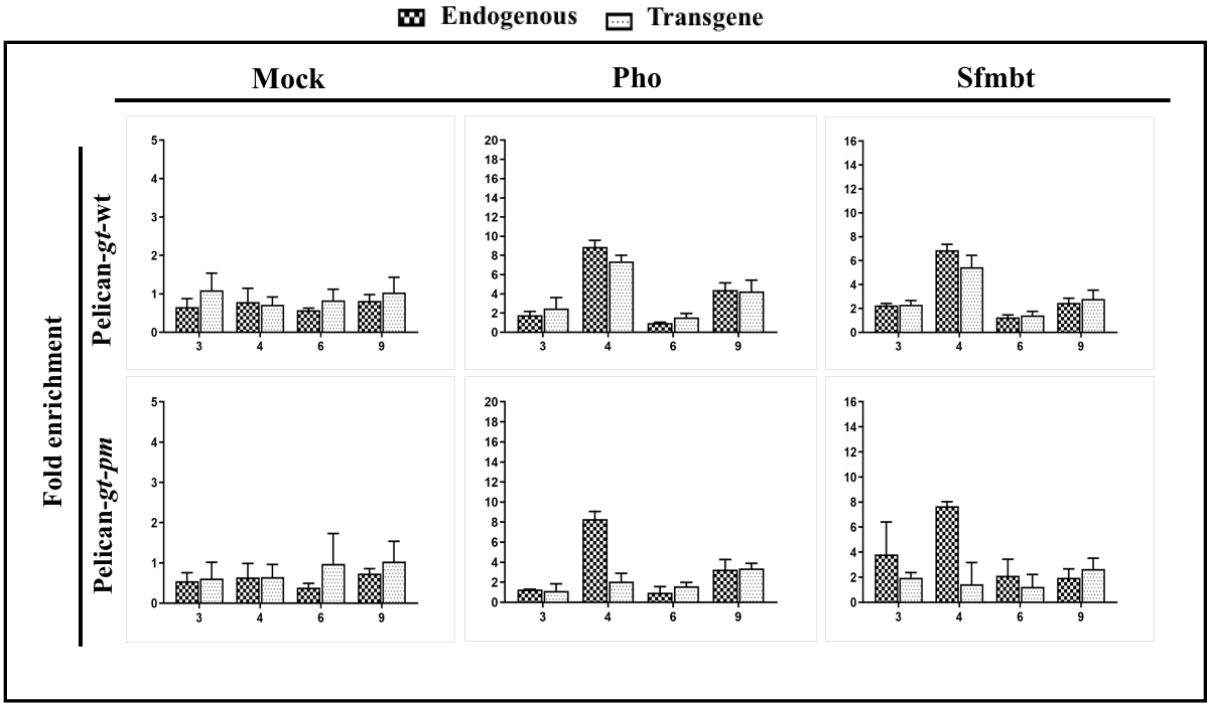

Figure 4.13- **ChIP-qPCR in sorted nc14b embryos with IgG (mock), anti-Pho, anti-Sfmbt antibodies, as indicated above.** Genotype of embryos is indicated on the left. ChIP signals from three biological replicates are presented as in figure 4.5.

The same binding pattern was observed for Pc, E(z) and Pcl. The latter proteins showed the similar binding level to the active endogenous *gt* as well as control and inert transgenes at regions 6 and 9. However, their binding to regions 3 and 4 of the inert transgene was greatly reduced compared with the active endogenous *gt* and control transgene (**Figure 4.14**).

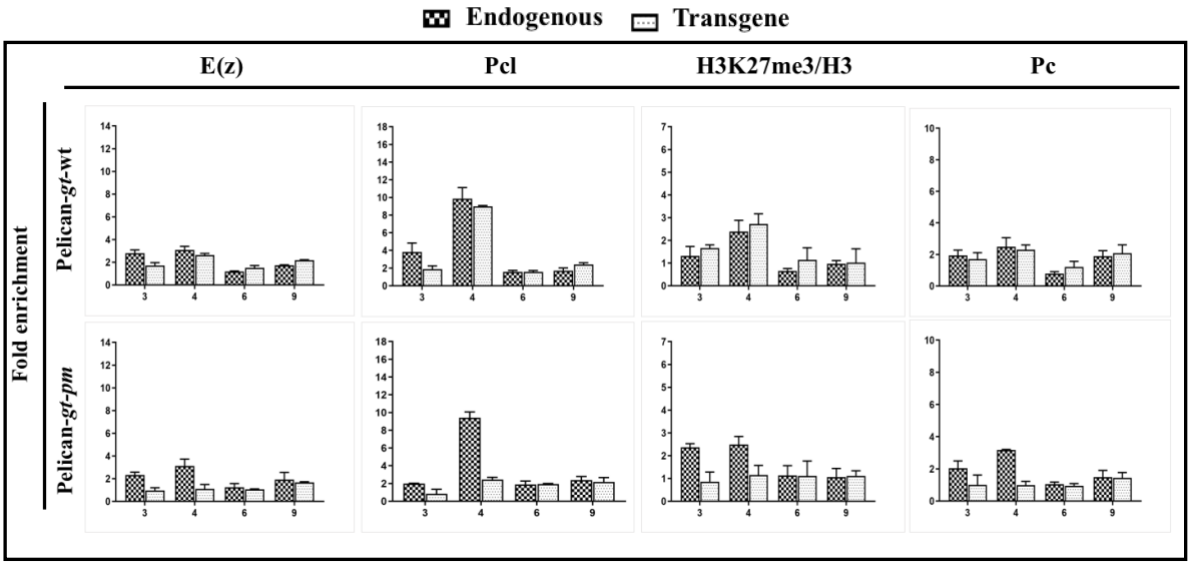

Figure 4.14- **ChIP-qPCR in sorted nc14b embryos with anti-E(z), anti-H3K27me3, anti-Pc and anti-Pcl antibodies, as indicated above.** Genotype of embryos is indicated on the left. ChIP signals from three biological replicates are presented as in figure 4.5.

Deposition of H3K27me3 at regions 3 and 4 was also affected negatively by the mutations in the promoter region. H3K27me3 was enriched at the same level as actively transcribed *gt* at regions 6 and 9 (**Figure 4.14**). Spps binding at regions 6 and 9 of the inert transgene resembled that of the active *gt*, while its binding to regions 3 and 4 of the inert transgene was greatly reduced (**Figure 4.15**).

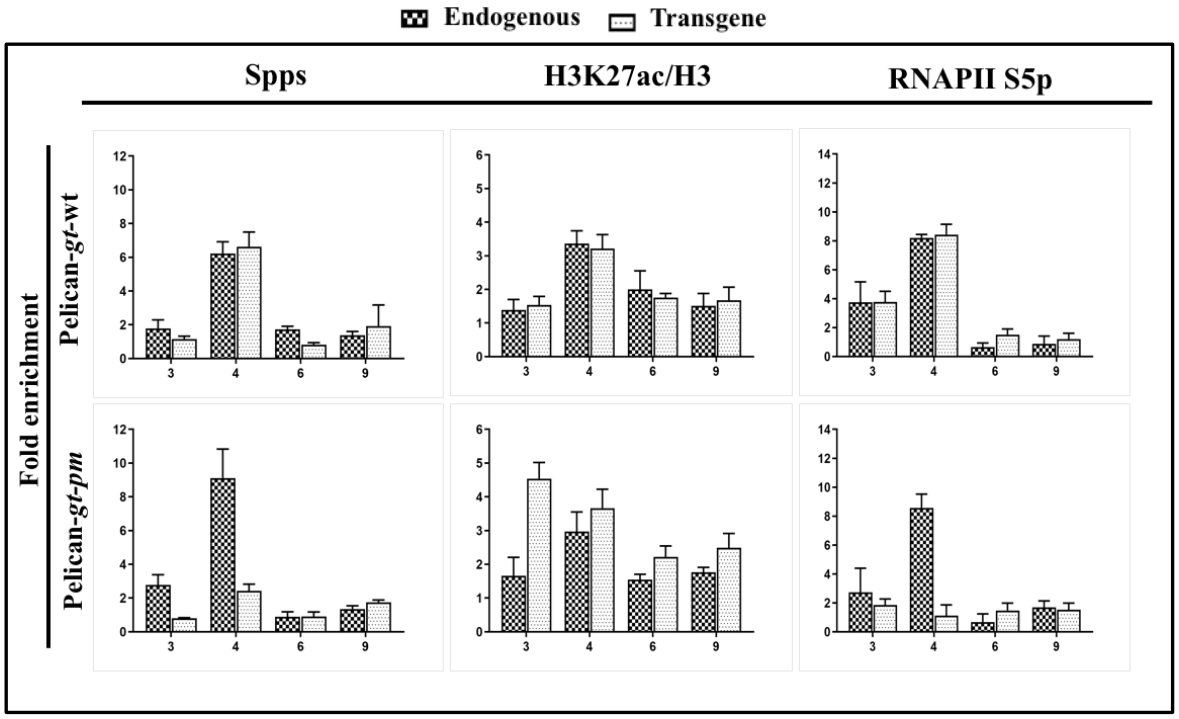

Figure 4.15- **ChIP-qPCR in sorted nc14b embryos with anti-Spps, anti-H3K27ac and anti-RNAPII S5p antibodies, as indicated above.** Genotype of embryos is indicated on the left. ChIP signals are presented as in figure 4.5 for Spps and figure 4.8 for H3K27ac and RNAPII S5p. Error bars show the standard deviation for three biological replicates.

Deposition of H3K27ac at regions 4, 6 and 9 of the inert transgene showed the similar pattern to the endogenous active *gt* and the control transgene. Interestingly, H3K27ac was enriched at higher levels at region 3 of the inert transgene compared to the endogenous *gt* and the control transgene (**Figure 4.15**).

Comparison of RNAPII S5p signals at the endogenous *gt*, as well as inert and control transgenes, indicated its enriched signal at the promoter region (region 4) of the endogenous *gt* and control transgene, which sank to the background level in the inert transgene (**Figure 4.15**).

Overal, PcG binding to transcriptional inert transgene was reduced at regions 3 and 4 (PRE1), and no difference was observed in PcG signals to PRE2 of the inert transgene and

endogenous *gt*. These data support the conclusion that PcG recognition of *gt* as repressed or active is independent from its transcriptional state. As the transcriptionally inert gene transgene was expected to be more compatible for the recruitment of PcG proteins but shown to recruit lower level of PcG proteins to regions 3 and 4 and same level at PRE2 compared to the actively transcribed endogenous *gt* and Pelican-*gt*-wt transgene.

4.2.6 Nucleosome density at *gt* of transcriptionally inert *gt* transgene

As shown above, we surprisingly observed a reduction in the binding of all tested PcG proteins and even transcription factors, Hb and Cad, to regions 3 and 4 of the transcriptionally inert transgene. We further noticed an approximately two-fold increase of H3 signals, in these two regions of the transcriptionally inert transgene compared to the endogenous *gt* gene and control transgene (**Figure 4.16**).

Signals of H3 may be considered as an indication of nucleosome density. Increase of H3 signals at the inert transgene indicates that mutations of promoter region resulted in the more compact chromatin region and consequently reduced binding of the tested proteins.

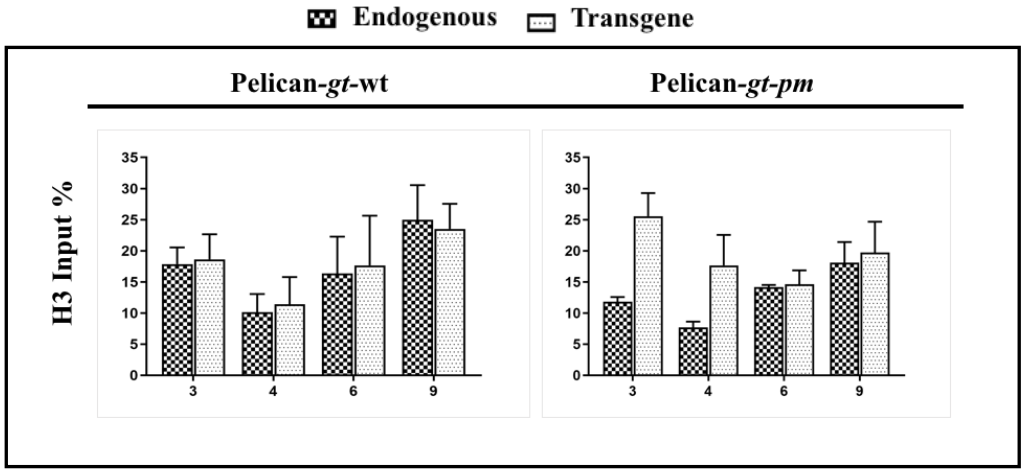

Figure 4.16- **H3 signals are shown as percentage of input in sorted nc14b embryos.** Genotype of embryos is indicated above. Error bars show the standard deviation for three biological replicates.

4.2.7 PcG binding to *gt* in Cad KD; *bcd osk tsl* and Cad; HbKD *bcd osk tsl* embryos

Comparing the recruitment of PcG proteins at the transcriptionally inert *gt* transgene with the ubiquitously expressed control transgene and endogenous *gt*, we showed that the differential binding of PcG proteins to the *gt* regulatory region in the active (HbKD *bcd osl tsl*) and repressed (*bcd osl tsl*) states was due to the presence of different transcription factors in the two *gt* states.

We have previously shown that in *bcd osk tsl* embryos, where Cad and Hb are uniformly distributed, the level of PRC1 and PRC2 binding at *gt* PREs is higher than that in HbKD *bcd osk tsl* embryos. This observation raised two possibilities; 1) Hb, and/or a corepressor it recruits, trigger recruitment of PcG proteins to *gt* in *bcd osk tsl* embryos. 2) In the absence of Hb, Cad inhibits binding of PcG proteins in HbKD *bcd osk tsl* embryos. To determine which of these two possibilities was correct, we looked at PcG binding to *gt* upon knocking down the levels of solely maternal Cad and both maternal Hb and Cad in *bcd osk tsl* background. Maternally expressed Gal4

driver was used to knock down the maternal levels of Cad and both Hb and Cad proteins in *bcd osk tsl* background (**Figure 4.17; 4.18**).

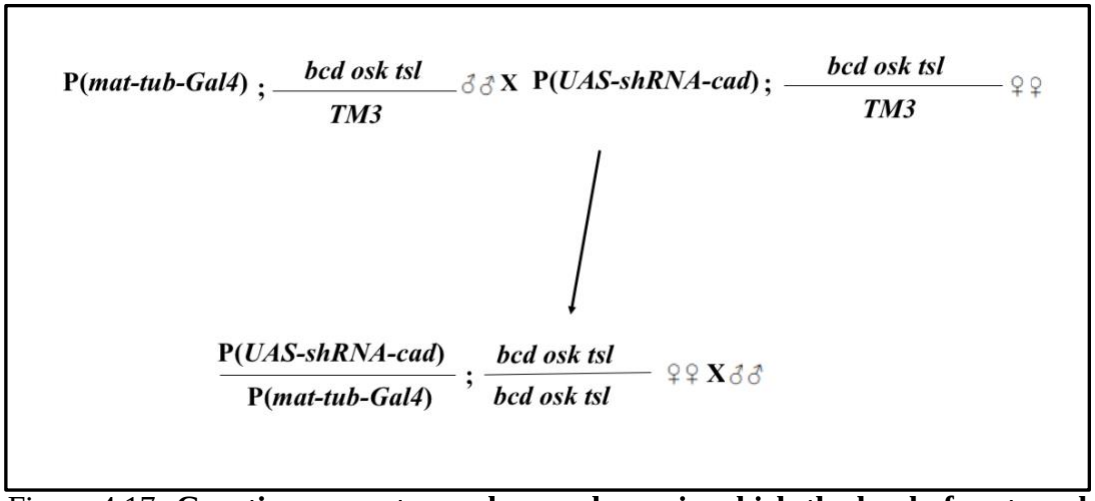

Figure 4.17- **Genetic crosses to produce embryos in which the level of maternal Cad is knocked down in *bcd osk tsl* embryos.**

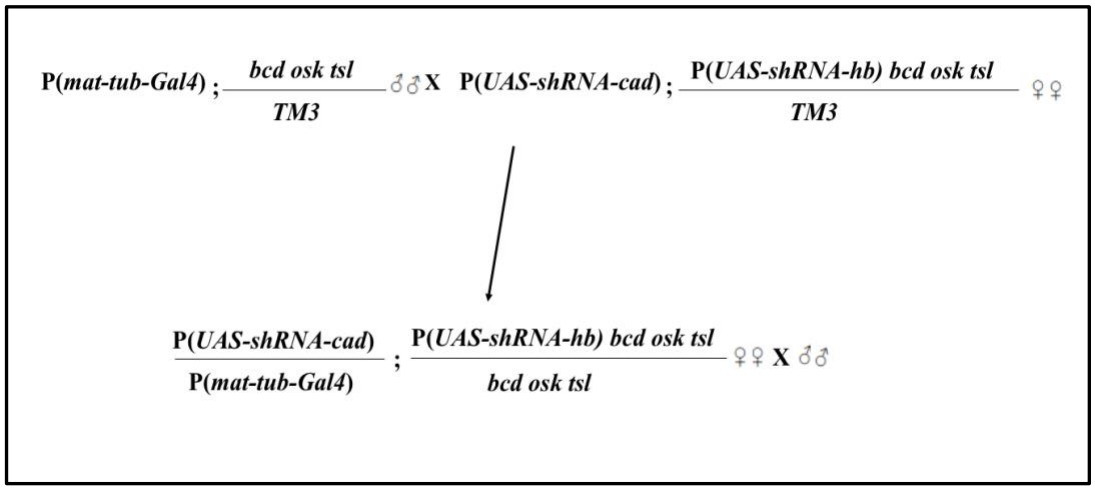

Figure 4.18- **Genetic crosses to produce embryos in which the levels of maternal Cad and Hb are knocked down in *bcd osk tsl* embryos.**

ChIP assays on embryos from CadKD; *bcd osk tsl* females showed a 82% reduction in Cad binding to *gt_(-3)* enhancer (**Figure 4.19**). RT-qPCR assays on nc10-13 CadKD; *bcd osk tsl* embryos showed that *gt* was expressed in miniscule levels, while *Kr* was expressed three times more than its expression level in wt embryos (**Figure 4.20**).

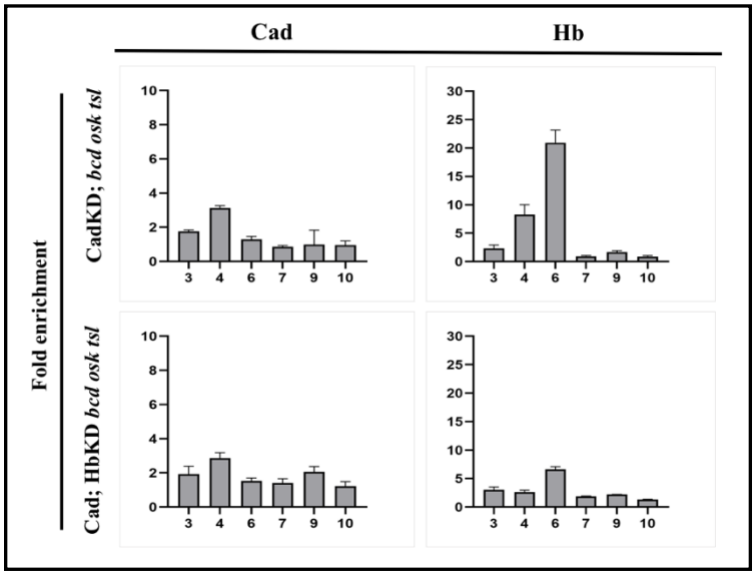

Figure 4.19- **ChIP-qPCR in sorted nc13 embryos with anti-Cad and anti-Hb antibodies, as indicated above.** Genotype of embryos is indicated on the left. ChIP signals from three biological replicates are presented as in figure 4.5.

In embryos from Cad; HbKD *bcd osk tsl* females, Cad and Hb binding to *gt_(-3)* enhancer showed 78% and 73% reduction, respectively (**Figure 4.19**). The latter embryos express *Kr* and *gt* 25.6% and 35.5% relative to expression levels in wt embryos, respectively (**Figure 4.20**). Bcd and low levels of Hb are activators for *Kr* expression, therefore, expression of *Kr* is reduced in the

resulting embryos from Cad; HbKD *bcd osk tsl* females (**Figure 4.20**). All the known activators and repressors of *gt* are either absent (Bcd, Tll) or expressed in very low levels (Hb, Cad, Kr, Gt) in the latter genetic system, hence, *gt* was expected to be uniformly repressed in the resulting embryos. Immunostaining of CadKD; *bcd osk tsl* embryos confirmed the results of RT-qPCR assays for *Kr* and *gt* expression (**Figure 4.21**). Interestingly, Cad; HbKD *bcd osk tsl* embryos did not show any staining using anti-Kr and anti-Gt antibodies, indicating the low expression level of these proteins per nuclei (**Figure 4.22**).

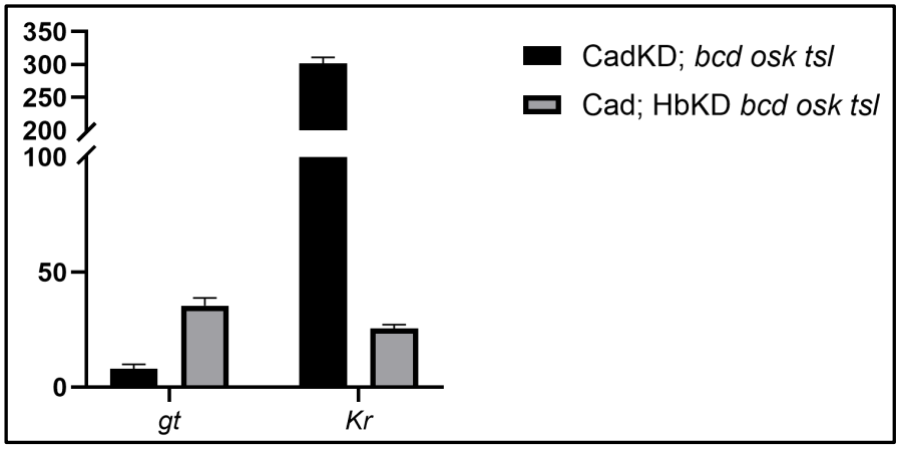

Figure 4.20- **Expression of *gt* and *Kr* in nc13-14a embryos derived from CadKD; *bcd osk tsl* and Cad; HbKD *bcd osk tsl* females.** Genotype of embryos is indicated above. RT-qPCR signals were normalized to *rp49* transcript. Y-Axis shows the relative expression of *lac*Z to wt embryos. Error bars show the standard deviation for three

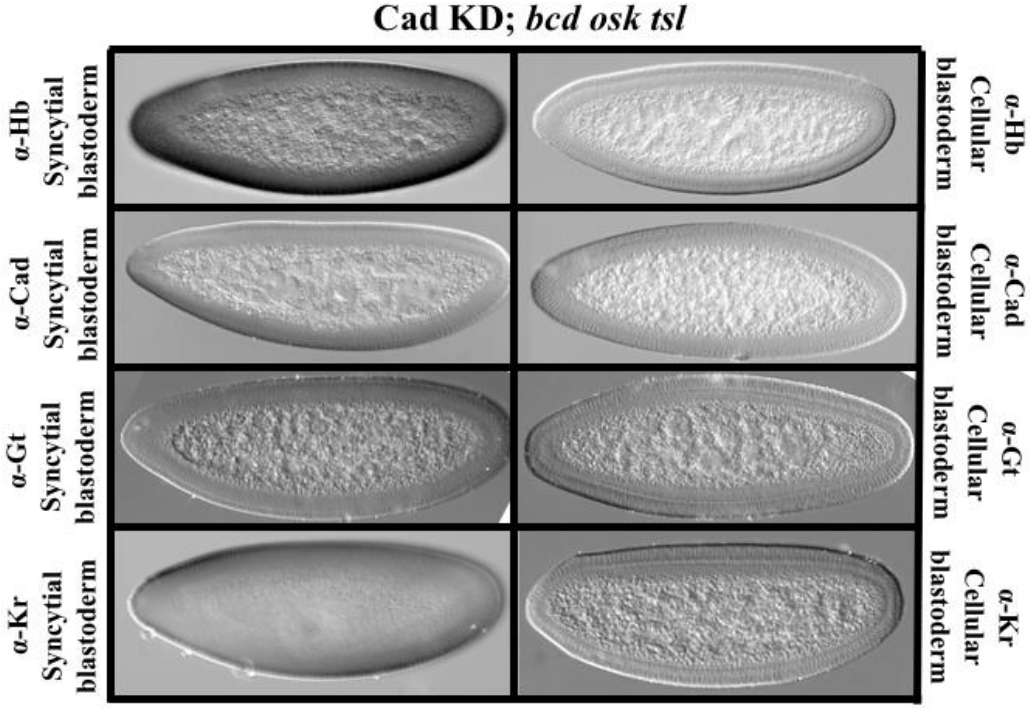

Figure 4.21- **Embryos from CadKD; *bcd osk tsl* females immunostained with anti-Hb, anti-Cad, anti-Gt and anti-Kr antibodies.** Embryonic stages and the antibodies used are stated on the left.

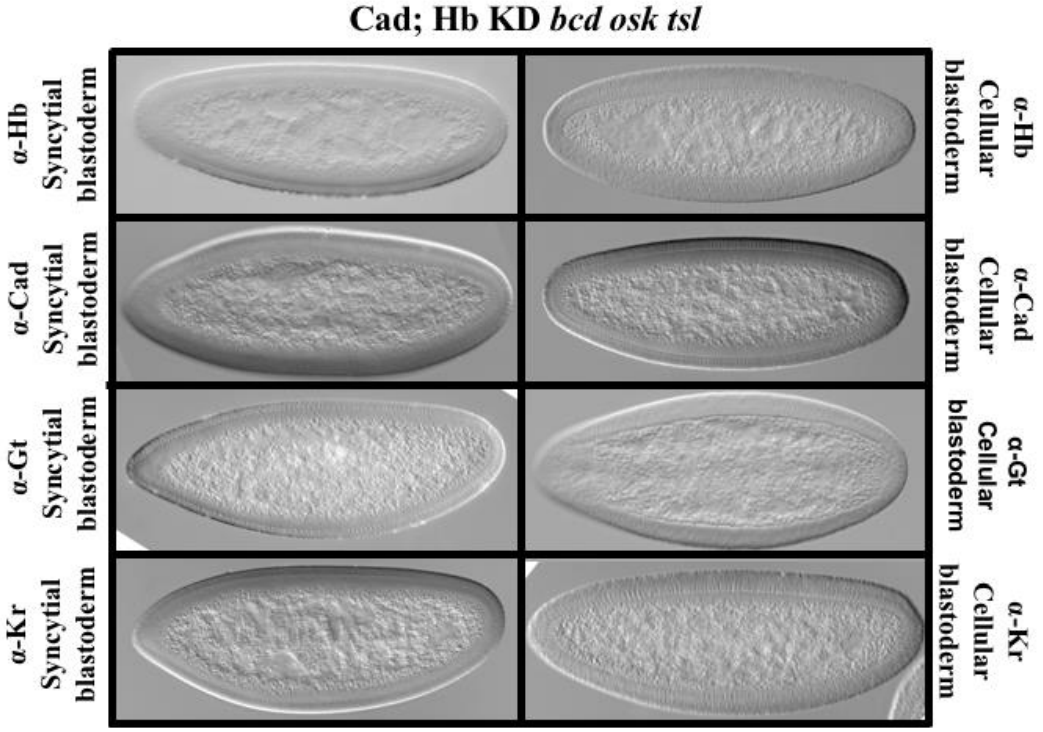

Figure 4.22- **Embryos from Cad; HbKD *bcd osk tsl* females immunostained with anti-Hb, anti-Cad, anti-Gt and anti-Kr antibodies.** Embryonic stages and the antibodies used are stated on the left.

Pho binding to the *gt* regulatory region of CadKD; *bcd osk tsl* embryos, mirrored its binding to that of the *bcd osk tsl* females, with strong peaks at PRE1 and PRE2 and background signals at other *gt* region (3, 6, 7 and 10) (**Figure 4.23**). Signals of H3K27ac were weakly above the background and comparable at all tested *gt* regions of CadKD; *bcd osk tsl* and *bcd osk tsl* embryos (**Figure 4.23**). Although, *gt* is not expressed in embryos from *bcd osk tsl* and is expressed at very low levels in CadKD; *bcd osk tsl* embryos, RNAPII S5p, associated with transcription initiation, was detected at the promoter region of both genetic systems (**Figure 4.23**). Additionally, the signal

of RNAPII S5p at the promoter region (six-fold enrichment), is similar in embryos from both *bcd osk tsl* and CadKD; *bcd osk tsl* females.

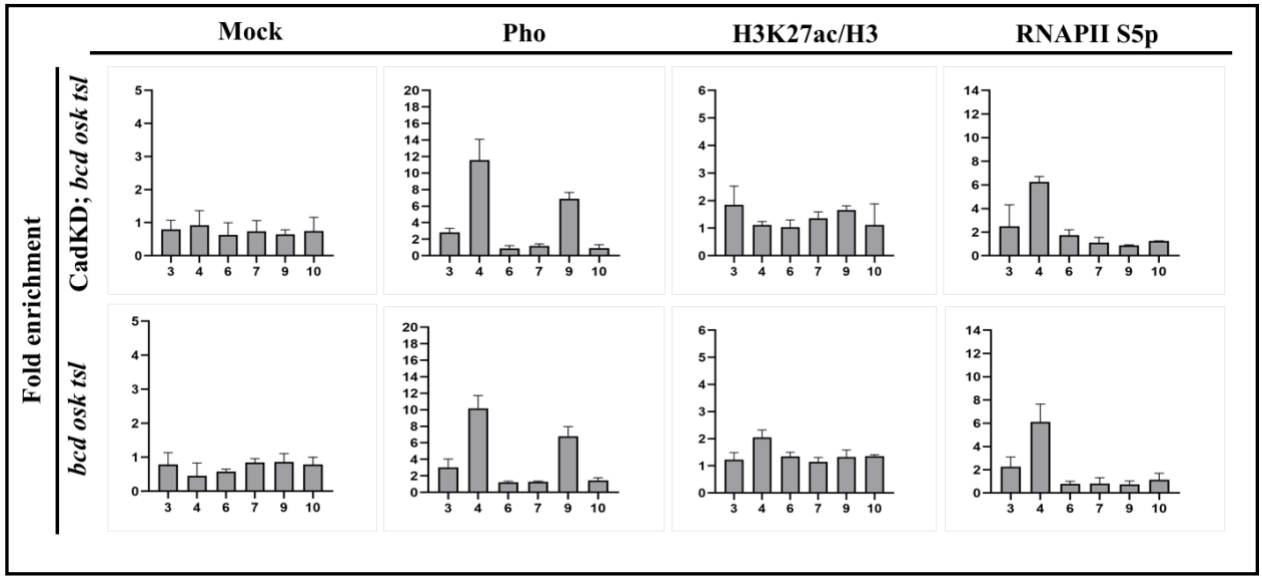

Figure 4.23- **ChIP-qPCR in sorted nc14b embryos with IgG and anti-Pho, anti-H3K27ac and anti-RNAPII S5p antibodies, as indicated above.** Genotype of embryos is indicated on the left. ChIP signals are presented as figure 4.5 for mock and Pho or figure 4.8 for H3K27ac and RNAPII S5p. Error bars show the standard deviation for three biological replicates.

E(z) binding to regions 3, 4, 7, 9 and 10 of CadKD; *bcd osk tsl* embryos was slightly higher than that of the corresponding regions in *bcd osk tsl* embryos (**Figure 4.24)**. In concert with E(z) signal, deposition of H3K27me3 showed slight increase at regions 3, 4 and 6 and robust elevation at regions 7, 9 and 10 of CadKD; *bcd osk tsl* embryos compared with embryos from *bcd osk tsl* females (**Figure 4.24)**. Signals of Pcl at regions 3 and 4 of CadKD; *bcd osk tsl* embryos, showed an increase compared to those at the same regions of *bcd osk tsl* embryos. However, Pcl binding to regions 6, 7, 9 and 10 remained at the same level in both genetic systems (**Figure 4.24)**. Pc

binding to the all tested *gt* regions (3-10) of CadKD; *bcd osk tsl* embryos increased by at least two times compared to *bcd osk tsl* embryos (**Figure 4.24)**.

We further investigated the binding of PcG proteins in the embryos in which all the known *gt* activators and repressors are either absent or expressed in low levels. Signals of RNAPII S5p, Pho and Pc increased dramatically at region 4 upon double knock down of Cad and Hb in *bcd osk tsl* embryos, while the latter signals at other tested regions of *gt* were comparable to those of CadKD; *bcd osk tsl* (**Figure 4.25; Figure 4.26)**. H3K27ac signals did not show considerable changes at all tested regions in Cad; HbKD *bcd osk tsl* compared to CadKD; *bcd osk tsl* embryos (**Figure 4.25**).

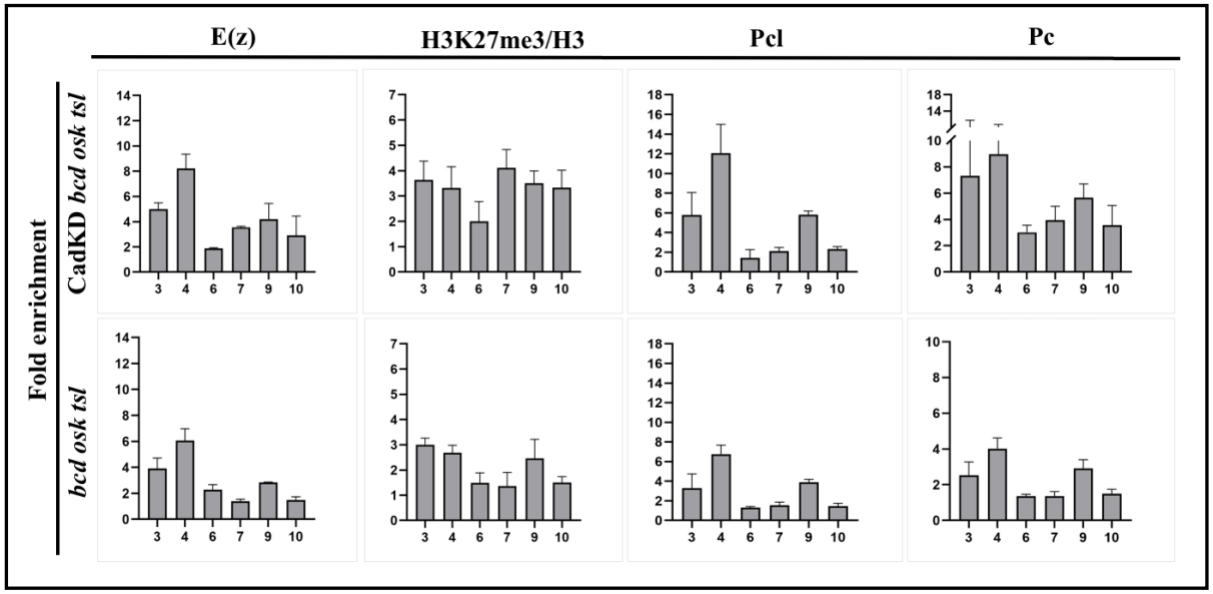

Figure 4.24- **ChIP-qPCR in sorted nc14b embryos with anti-E(z) and anti-H3K27me3, anti-Pcl and anti-Pc antibodies, as indicated above.** Genotype of embryos is indicated on the left. ChIP signals from three biological replicates are presented as in figure 4

Binding of Pcl and E(z) and deposition of H3K27me3 showed a mild elevation at region 4 in Cad; HbKD *bcd osk tsl*, however, their binding to other tested regions was comparable with CadKD; *bcd osk tsl* (**Figure 4.26**). Our ChIP data in Cad; HbKD *bcd osk tsl* excluded the possibility of a positive role for Hb in recruiting PcG proteins, providing solid evidence for the inhibitory role of Cad on the recruitment of PcG proteins to *gt*. ChIP assays on CadKD; *bcd osk tsl* embryos also supported our hypothesis, as binding of PcG proteins to *gt* genomic region showed a mild increase compared to that in *bcd osk tsl* embryos.

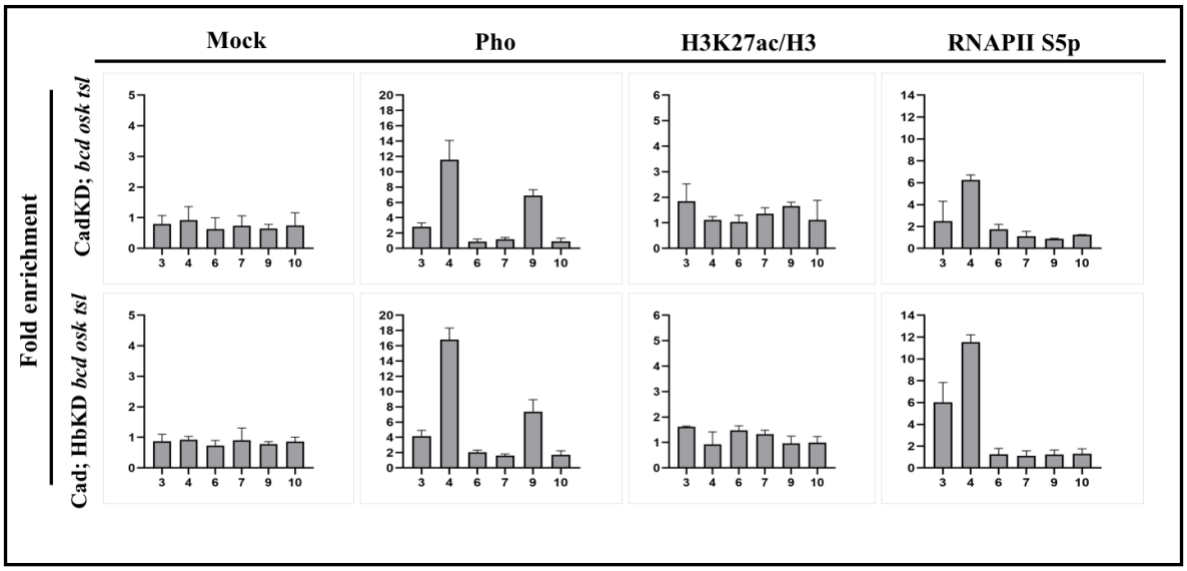

Figure 4.25- **ChIP-qPCR in sorted nc14b embryos with IgG, anti-Pho, anti-H3K27ac and anti-RNAPII S5p antibodies, as indicated above.** Genotype of embryos is indicated on the left. ChIP signals from three biological replicates are presented as in figure 4.5.

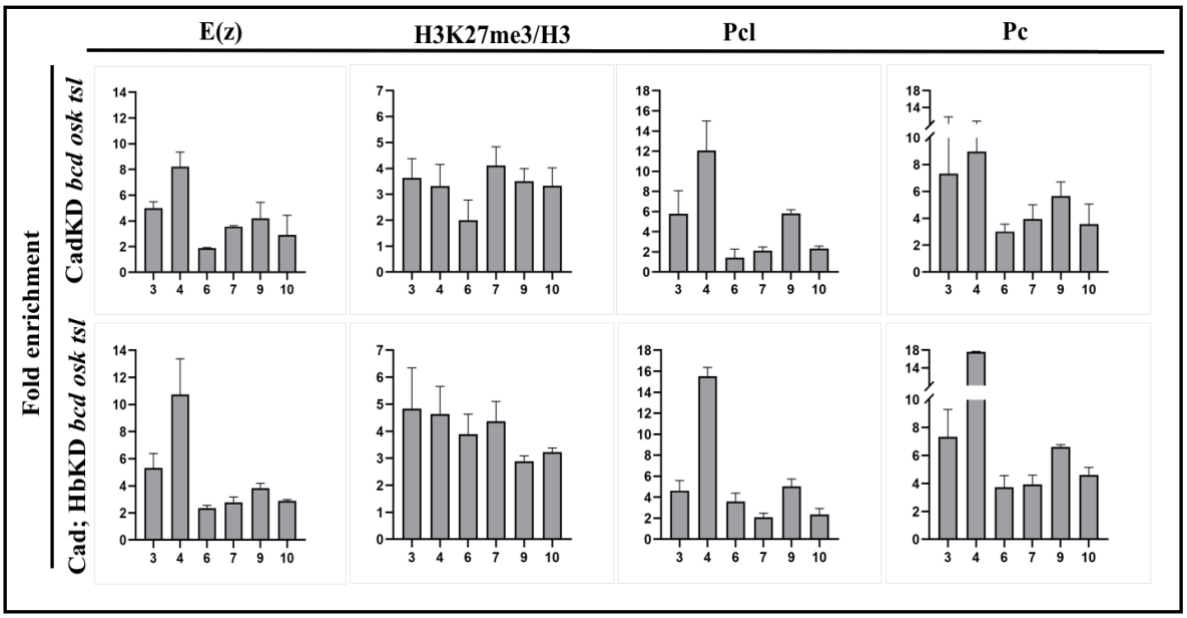

Figure 4.26- **ChIP-qPCR in sorted nc14b embryos with anti-E(z), anti-H3K27me3, anti-Pcl and anti-Pc antibodies, as indicated above**. Genotype of embryos is indicated on the left. ChIP signals from three biological replicates are presented as in figure 4.5.

4.3 Miscellaneous data

4.3.1 Testing TRiP knock down lines

Maternally expressed Gal4 driver was used to test the efficiency of TRiP UAS-shRNAs for proteins involved in the initiation or elongation of transcription. In order to test the efficiency of UAS-shRNA-mediated knock down of specific proteins, mortality rate and cuticle patterns of the embryos were examined. The results of examining the knockdown efficiency of tested transgenic lines are listed in **Table 4.1**.

Table 4.1- **Efficiency of various TRiP RNAi fly stocks was tested based on cuticle patterns and mortality rate.**

Name	BDSC Stock ID	Vector	Insertion site	Temperature	Phenotype
Hb	54478	VALIUM22	attP2	25°C	Head defect, lacking of T2, T3 and A1 segments
Cad	57546	VALIUM20	attP40	25°C	Severely deformed head, missing abdominal segments
Ash1	36803	VALIUM22	attp40	25°C	Normal, grew to adults
Set1	40931	VALIUM20	attp40	25°C	Normal, grew to adults
Fs(1)h	41693	VALIUM20	attp40	25°C	Eggs died very early or unfertilized eggs
Fs(1)h	41943	VALIUM20	attp40	25°C	Normal, grew to adults
Fs(1)h	44009	VALIUM20	attp40	25°C	Normal, grew to adults
Cdk9	41932	VALIUM20	attp40	25°C	Normal, grew to adults

Set2	42511	VALIUM20	attp40	25°C	Normal, grew to adults
Haywire	53345	VALIUM20	attp40	25°C	70-80% died as embryos with defects such as less number of denticle belts
Set2	55221	VALIUM20	attp40	25°C	Normal, grew to adults
Cdk7	57245	VALIUM20	attp40	25°C	Normal, grew to adults
Cdk7	62304	VALIUM20	attp40	25°C	Normal, grew to adults
Ash2	64942	VALIUM20	attp40	25°C	Normal, grew to adults
Lsd-1	65020	VALIUM20	attp40	25°C	Normal, grew to adults
Wds	60399	VALIUM20	attp40	25°C	Very few eggs, didn't hatch
Mat1	57312	VALIUM20	attp40	25°C	1/4th of the eggs died very early, the rest grew normally
Xpd	65883	VALIUM20	attp40	25°C	Normal, grew to adults
CycT	32976	VALIUM20	attp2	25°C	Normal, grew to adults
Ash2	35388	VALIUM22	attp2	25°C	Normal, grew to adults
Cdk9	35323	VALIUM22	attp2	25°C	Most embryos did not hatch, showed embryonic defects
Lilli	34592	VALIUM20	attp2	25°C	Normal, grew to adults
Cdk7	53199	VALIUM22	attp2	25°C	Normal, grew to adults

dMi-2	35398	VALIUM22	attp2	25°C	Most embryos died, shorter embryos with missing denticle belts
CBP	37489	VALIUM20	attp2	25 and27°C 29°C	30-40% of embryos hatched and developed into adults. 60-70% died as embryos. Very few eggs were laid.
Hopped CBP -2.1	N/A	VALIUM20	2nd chromosome	25-27-29°C	Few eggs were laid and 50-60% of embryos hatched and died as first and second instar larvae.
Hopped CBP-2.2	N/A	VALIUM20	2nd chromosome	25-27-29°C	Normal, grew to adults.
Hopped CBP-3.1	N/A	VALIUM20	2nd chromosome	25-27-29°C	Few eggs were laid and 50-60% of embryos hatched and died as first and second instar larvae.
Hopped CBP-17	N/A	VALIUM20	2nd chromosome	25-27-29°C	Normal, grew to adults.
Hopped CBP-36.3	N/A	VALIUM20	2nd chromosome	25-27-29°C	Normal, grew to adults.
Hopped CBP-36.2	N/A	VALIUM20	2nd chromosome	25-27°C 29°C	Laid few eggs, grew to adults. No eggs
Hopped CBP-10	N/A	VALIUM20	2nd chromosome	25-27-29°C	Normal, grew to adults.

4.3.2 Various tested antibodies in ChIP assays

Scm, a substochiometric shared subunit of PRC1 and PRC2, was tested for the possible detection at *gt* PRE1 and PRE2. However, Scm was not detected across *gt* genomic region in nc14b

embryos from *bcd osk tsl* females. Scm signal was negative across the region with two significant peaks at regions 7 and 10, in nc14b embryos from HbKD *bcd osk tsl* females (**Figure 4.27**). Since no regulatory element is located at or close to regions 7 and 10, and these regions have consistently showed low signals for all tested proteins in both active and repressed states of *gt*, the observed peaks are suggested to be unspecific at these regions.

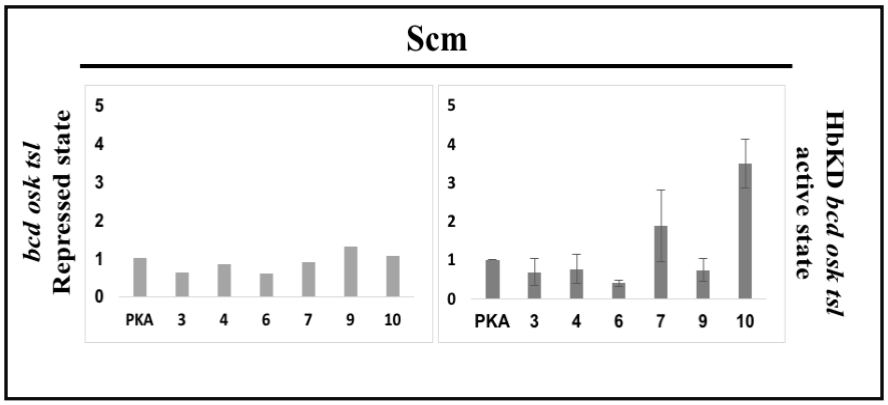

Figure 4.27- **ChIP-qPCR in sorted nc14b embryos with anti-Scm antibody, as indicated above.** Genotype of embryos is indicated on the left and right. ChIP signals are presented as in figure 4.5. Error bars show the standard deviation for two biological replicates in the active state. ChIP on the repressed state was performed in one replicate.

dRing, a shared subunit of both canonical and non-canonical PRC1 complexes, was detected at PRE1 and PRE2 of both repressed and active states of *gt* in nc14b embryos. However, its binding at the active state was lower compared to the repressed state at both PREs (**Figure 4.28**). Ring monoubiquitylates lysine 118 of histone H2A (H2AK118ub) in *Drosophila* or the corresponding lysine 119 in mammals (Wang et al., 2004). Using H2AK119ub antibody,

previously used successfully in *Drosophila* (Pengelly et al., 2015), the considerable deposition of this histone mark across the region of active *gt* was observed (**Figure 4.28**). However, the signals of H2AK118ub is not strongly correlated with dRing, as two peaks for this histone mark were localized at regions 3 and 6, which are not supported by the low binding of dRing at these regions.

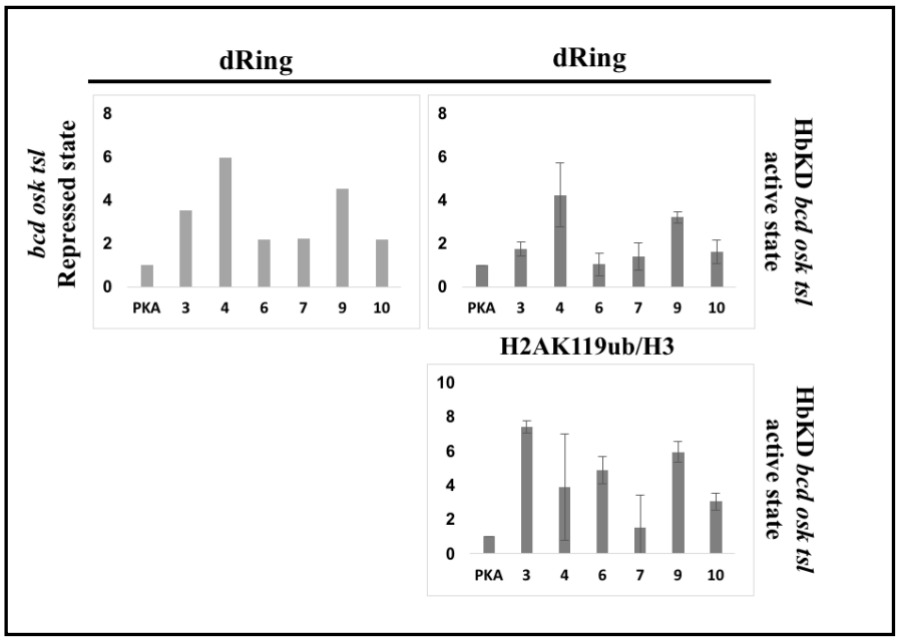

Figure 4.28- **ChIP-qPCR in sorted nc14b embryos with anti-dRing and anti-H2AK119ub antibodies, as indicated above.** Genotype of embryos is indicated on the left and right. ChIP signals are presented as in figure 4.5. Error bars show the standard deviation for three biological replicates in the active state. ChIP on the repressed state was performed in one replicate.

H3K36me2/me3 is the active histone mark associated with transcription elongation and is enriched at the gene body of actively transcribed genes. Using H3K36me3 antibody, a sharp peak

was observed at region 4, which encompasses promoter region, at active state of *gt*. H3K36me3 signal was weakly positive at region 3, which is ~300 to 500 bp downstream of TSS. Given that higher signal of H3K36me3 was expected at region 3 compared to region 4, this antibody was not verified for further use in ChIP assays (**Figure 4.29**). H3K36me3 signal was negative across the *gt* genomic region in the repressed state (**Figure 4.29**).

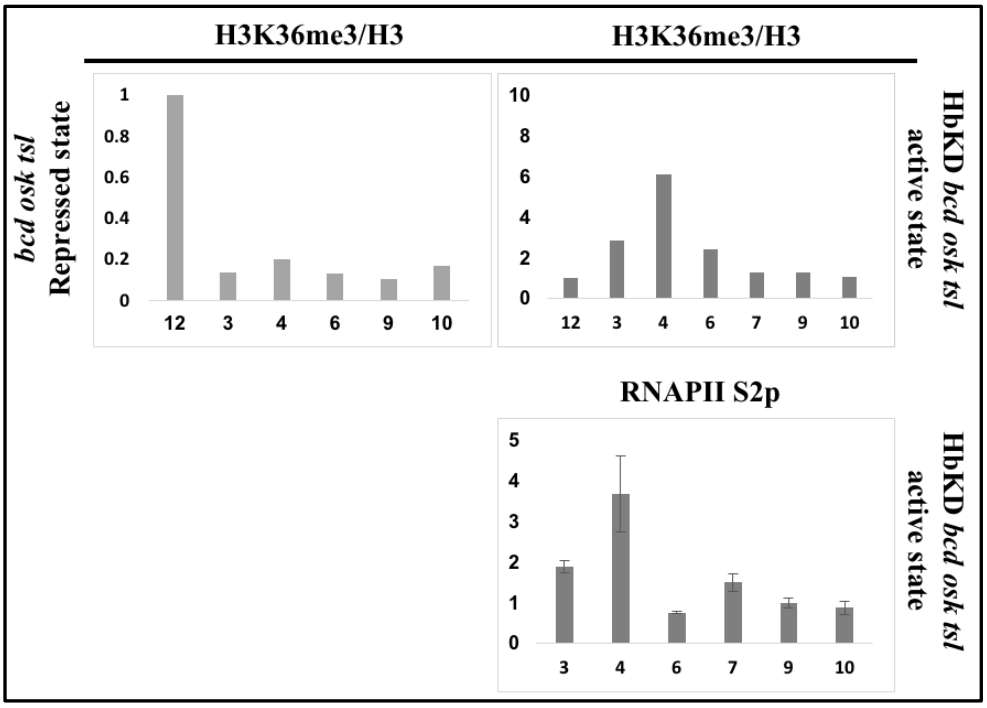

Figure 4.29- **ChIP-qPCR in sorted nc14b embryos with anti-H3K36me3 and RNAPII S2p antibodies, as indicated above.** Genotype of embryos is indicated on the left and right. ChIP signals are presented as in figure 4.8. ChIP experiments were performed in one replicate for H3K36me3 in both active and repressed states. ChIP on RNAPII S2p was performed in two replicates.

RNAPII S2p is associated with transcription elongation. Therefore, highest signal of RNAPII S2p was expected to be detected within the gene body at region 3. However, a peak was detected at region 4, but not 3, in ChIP assays (**Figure 4.29**), resembling what has been observed using RNAPII S5p (**Figure 4.29**). This antibody was also excluded from further use in ChIP assays.

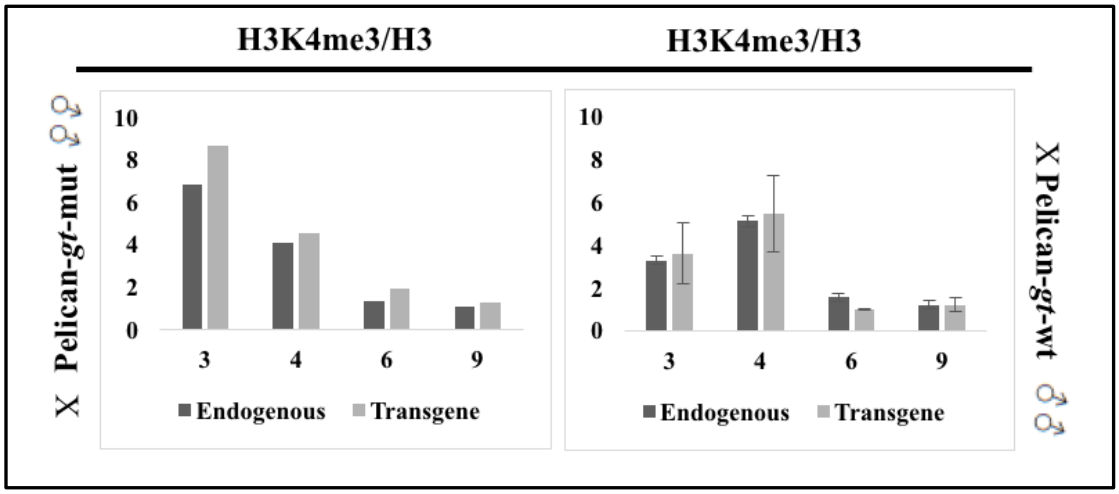

Figure 4.30- **ChIP-qPCR in sorted nc14b embryos from HbKD *bcd osk tsl* females crossed to either Pelican-*gt*-wt or Pelican-*gt*-pm with anti-H3K4me3.** Genotype of embryos is indicated on the left and right. ChIP signals are presented as in figure 4.8. ChIP experiments were performed in one and two replicates for Pelican-*gt*-pm and Pelican-*gt*-wt, respectively.

Signal of H3K4me3 showed peaks at regions 3 and 4 of transcriptionally active endogenous *gt* as well as Pelican-*gt*-wt and Pelican-*gt*-pm transgenes (**Figure 4.30**). Enriched signal of H3K4me3 at regions 3 and 4 of Pelican-*gt*-pm transgene, indicated that the observed histone mark is not associated with the active transcription at *gt* and may be due to the binding of TrxG proteins.

Hb and Cad are the repressor and activator of *gt_(-3)* enhancer. dMi-2 and CBP are the possible candidates as the co-repressor and co-activator of Hb and Cad, respectively. CBP acetylates H3K27 and dMi-2 is a subunit of the NuRD histone deacetylase complex with nucleosome remodeling activity. The constitutive ribosomal gene *rp49* was shown to be enriched for both dMi-2 and CBP within 1 kb upstream of TSS (modENCODE). Nevertheless, antibodies against dMi-2 and CBP failed to show any specific signal cross the *gt* region and 600 bp upstream of TSS of *rp49* (**Figure 4.31**).

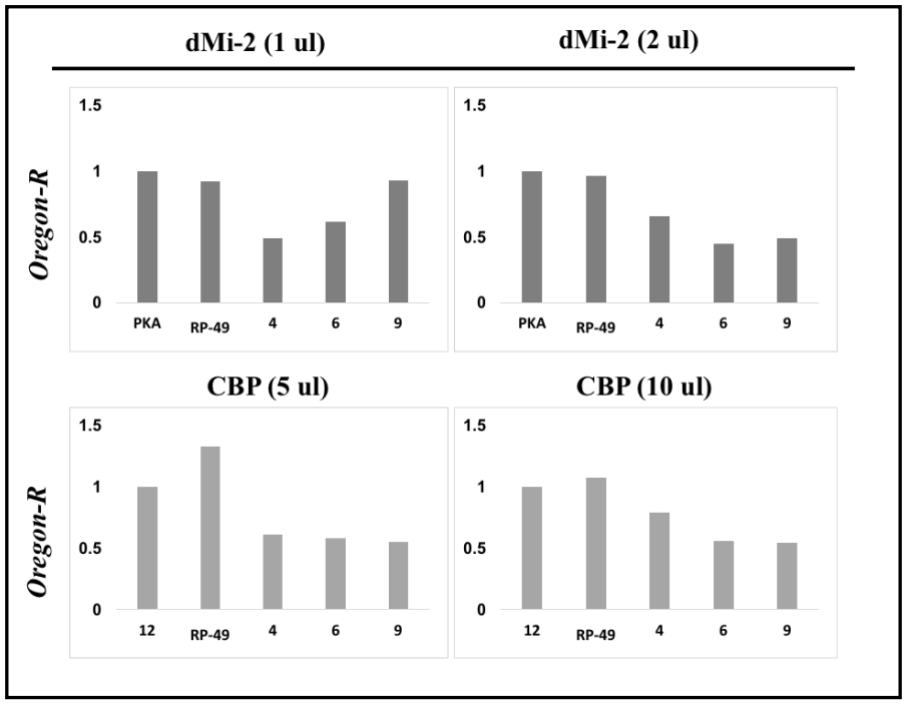

Figure 4.31- **ChIP-qPCR in sorted nc10-13 *Oregon-R* embryos.** Antibodies and amounts used are indicated above. Genotype of embryos is indicated on the left. ChIP signals are presented as fold enrichment as in figure 4.5 for dMi-2 and figure 4.8 for CBP. ChIP experiments were performed in one replicate.

Kni recruits CtBP as a co-repressor at some genes (Nibu et al. 1998). The histone lysine demethylase Lsd1 is a subunit of CtBP complex which can confer transcriptional activation, through demethylation of H3K9me3, or transcriptional repression through demethylation of H3K4me3 (Ray et al., 2014). Kni crude antiserum and Lsd1 purified antibody and antisera (462 and 463) were used to determine the presence of these proteins at *gt*. The signals were negative across the *gt* genomic region (**Figure 4.32**). Since the signals at a positive region were not investigated, we are not able to determine if the negative signals of Lsd1 and Kni were due to the antibody unspecificity or the absence of these protein at *gt*.

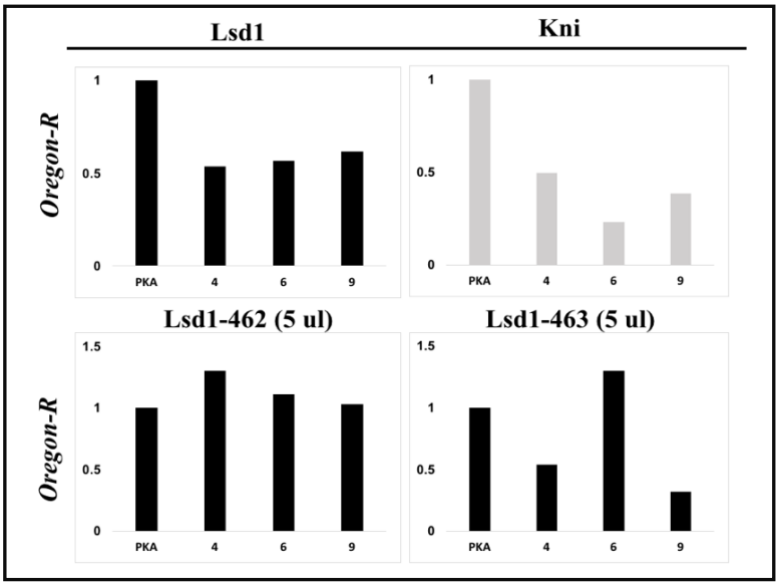

Figure 4.32- **ChIP-qPCR in sorted nc14b-nc15 *Oregon-R* embryos.** Antibodies and amounts used are indicated above. ChIP signals are presented as in figure 4.5. ChIP experiments were performed in one replicate.

RNAPII is consisted of 12 subunits. Rpb4 is the fourth largest subunit and an antibody against it, is commonly used to detect the presence of all forms of RNAPII. Using Western blot-verified Rbp4 antibody, a negative signal was detected at *gt* promoter region (region 4), which showed the incapability of this antibody to recognize Rbp4 in ChIP assays (**Figure 4.33**).

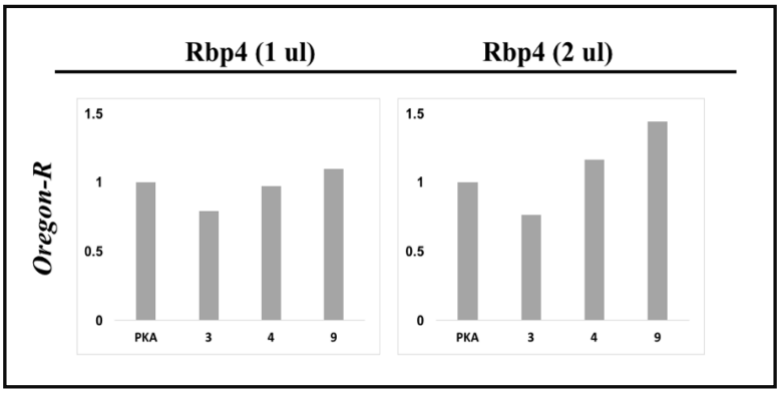

Figure 4.33- **ChIP-qPCR in sorted nc14b-nc15 *Oregon-R* embryos.** Amounts used for anti-Rbp4 are indicated above. . ChIP signals are presented as in figure 4.5. ChIP experiments were performed in one replicate.

4.4 Discussion

4.4.1 Characterization of HbKD *bcd osk tsl* genetic system

As mentioned previously (section **4.2.1**), our first attempts for an efficient knock down of Hb failed with young females (2-4 days old) at 25 °C. Using slightly older females (5-9 days old) and higher temperature (27 °C) increased the knock down level from 50% to 79%. Ni et al. (2011) reported the dependence of the efficiency of shRNA-mediated knock downs on the maternal age and temperature. The age-dependent efficiency of knock down levels, can be explained through the faster rate of egg production in young females which results in the overall lower loading of maternally-deposited shRNAs into eggs (Ni et al., 2011). Gal4 is a yeast transcription factor, which

serves as a transcriptional activator in UAS-Gal4 gene expression system. Gal4 has the highest activity at 30 °C, which is the optimal temperature for yeast. Therefore, at higher temperatures, 27 or 29 °C, the Gal4-driven knock down is expected to be more effective. However, we observed higher rate of mortality and unfertilized eggs upon shifting the flies into higher temperatures. The mentioned observation may be due to the more efficient depletion of maternal Hb or the general Gal4-induced ectopic expression of other genes at higher temperatures.

4.4.2 Binding of Hb and Cad to *gt_(-3)* enhancer in *bcd osk tsl* and HbKD *bcd osk tsl* embryos

Maternal Hb and Cad proteins are uniformly expressed in syncytial blastoderm *bcd osk tsl* embryos and bind to *gt_(-3)* enhancer. The results of both immunostaining and ChIP assays, showed the miniscule amount of maternal Hb due to the efficient knock down of Hb. However, the level of Cad binding to *gt_(-3)* enhancer of nc13 HbKD *bcd osk tsl* was comparable to that of *bcd osk tsl* embryos at the same stage. There are a number of overlapping and non-overlapping binding sites for Hb and Cad within *gt_(-3)* enhancer (Schroeder et al., 2004), but the independent binding of these proteins to *gt_(-3)* enhancer in different nuclei, cannot be supported by the uniform repression of *gt* in *bcd osk tsl* embryos. Furthermore, we did not notice an increase in the signals of Cad binding to *gt_(-3)* enhancer in HbKD *bcd osk tsl,* where binding of Hb was greatly reduced by ~80%. The latter observation excludes the possibility of competitive binding of these transcription factors to their overlapping binding sites.

4.4.3 Differential recruitment of PcG proteins to the PREs of active versus repressed states of *gt*

ChIP assays on HbKD *bcd osk tsl* (active state) and *bcd osk tsl* (repressed state) embryos, indicated comparable binding levels of PhoRC to PRE1 of active and repressed states, while binding of these proteins to PRE2 of the active state was greatly reduced. Moreover, signals of

E(z), H3K27me3 and Pc were reduced at both PRE1 and PRE2 of the active state compared to the repressed *gt*.

Pcl signals, a substochiometric subunit of PRC2, were even higher, at PRE1 of the active state compared to the repressed state. H3K27me3 signal showed a small peak at PRE1 despite the low signal of E(z) at this region. Canonical PRC2 occupies the chromatin for a short time. Pcl binds DNA and increases the dwelling time of PRC2 on chromatin and enhances histone methyl transferase activity of mammalian PHF1-PRC2 (Choi et al., 2017). The small peak of H3K27me3 at PRE1, and its negative signal across other tested *gt* regions may be explained through the strong signal of Pcl at PRE1 and its weak presence at other regions.

4.4.4 Reduced binding of PcG proteins to PRE1 of transcriptionally inert *gt* transgene

Comparing the recruitment of PcG proteins at the transcriptionally inert transgene and active endogenous *gt*, we observed a consistent reduction of signals for all tested PcG proteins and H3K27me3 at regions 3 and 4. Furthermore, H3K27ac signal increased at region 3 of the inert transgene compared to the active endogenous *gt*. Region 3 represents ~300-500 bp downstream of *gt* TSS. Given that *lacZ* is not transcribed in the inert transgene, the absence of histone turnover may result in the buildup of H3K27ac histone mark at this region.

The surprising reduction of PcG proteins at regions 3 and 4, which encompasses promoter and PRE1, can be explained by the observed increase in H3 signals at the latter regions of the transcriptionally inert *gt* transgene compared to the active endogenous *gt* and control transgene in ChIP assays. Although, H3 signals are not direct indicators of the chromatin accessibility, the consistent reduction in binding of all tested PcG proteins and even activator Cad and repressor Hb at regions 3 and 4, may suggest the presence of a compact chromatin environment. Recruitment of Gaf and TFIID is required for the generation of chromatin accessibility at the promoter of *hsp26*

(Boris et al., 2002). Base pair substitution of *gt* TATA box, inhibits binding of TFIID, and consequently transcription pre-initiation complex does not form. Lack of TFIID binding possibly results in a compact chromatin structure, precluding binding of transcription factors, Cad and Hb, and PhoRC proteins. Reduced presence of PhoRC may result in the decreased recruitment of PRC1 and PRC2. We also noted that base pair changes introduced to *gt* promoter region, do not overlap with the consensus binding sites of Zld and Gaf.

Comparing the recruitment of PcG proteins at the transcriptionally inert *gt* transgene in a genetic background in which endogenous *gt* is actively and uniformly expressed, we determined that binding of PcG proteins is not dictated by the transcriptional state of the target gene, but rather is affected by the presence of repressive or activating transcription factors. This observation supported the "instructive model", which suggests that transcription factors, and not the transcriptional state, play the central role in the regulation of PcG binding. Multiples studies have provided the evidence that the recruitment of PcG proteins to their targets is not responsive to the transcriptional state, contradicting the "responsive model". Papp and Muller, (2006) compared the recruitment of PcG proteins to the *Ubx* gene in the haltere and third-leg imaginal discs, where it is expressed in all cells, and in the wing imaginal disc, where it is uniformly stably repressed. Their quantitative analysis of X-ChIP, showed that PhoRC binding at both *bx* PRE (located ~32 kb downstream of TSS) and *bxd* PRE (30 kb upstream of TSS) was similar in both transcriptional states. PRC1 and PRC2 subunits were shown to bind at comparable levels to the *bxd* PRE at both active and repressed states, while their binding was lower at the active *bx* PRE compared to its repressed state. The aforementioned study revealed the constitutive binding of PcG proteins at *Ubx* PREs irrespective of the transcriptional state.

Langlais et al. (2012), used UAS-driven-flag-tagged PcGs in conjugation with *engrailed* (*en*)-Gal4 and *cubitus interruptus* (*ci*)-Gal4 drives in specific cell populations that express *en* and those that do not (express *ci*, instead). Using this system, PcG binding to PRE2, at TSS, and PRE1, ~1 kb upstream of TSS, in the on and off states of the *en* gene in 3rd instar wing imaginal discs was compared. Pho was bound at comparable levels in both the on and off transcriptional states of *en* PRE1. Furthermore, the binding level of Pho, PRC1 and PRC2 at *en* PRE2 in the off state was higher than those in the on state, while the binding levels were comparable at *en* PRE1 in both states. However, it is worth noting that the number of cells expressing *ci* (off state of *en*) was greater than the *en*-expressing cells, and thus the detected higher level of PcG proteins in the off state cannot be thoroughly verified through this study.

Multiple studies support the instructive model in mammalian systems. Transcription factors such as Rest and Runx1 have been proposed to recruit PcG proteins through direct biochemical interactions. Rest interacts with PRC1 complex and triggers its recruitment to a subset of PcG-regulated neuronal genes in mESCs (Dietrich et al., 2012).

Core binding transcription factor Runx1 directly recruits PRC1 to chromatin independently of PRC2 in mammalian cells (Yu et al., 2012). However, PRC2-independent recruitment of PRC1 may be distinct from PRC1 role in PcG silencing. PRC1 has been shown to regulate the transition of paused RNAP II to elongating form at active genes independently of other PcG proteins (Schaaf et al., 2013).

4.4.5 Transcriptional activator Cad negatively affects binding of PcG proteins to *gt*

ChIP assays on Cad KD; *bcd osk tsl* and Cad; HbKD *bcd osk tsl* embryos, provided evidence for the negative role of transcriptional activator Cad on the recruitment of PcG proteins to *gt*.

Maternal Hb and Cad are ubiquitously expressed in embryos from *bcd osk tsl* females. However, Cad fails to activate *gt* in the presence of Hb. A number of studies, have suggested an antagonistic role for Hb and Cad at other genes (Clyde et al., 2003; Small et al.,1992; Vincent et al., 2018). Hb is converted from a repressor into an activator upon binding of Bcd or Cad (Janssens et al., 2006; He et al., 2010; Kim et al., 2013; Samee et al., 2014). The repressive effect of Hb on stripe 2+7 enhancer of *even-skipped*, is counteracted by the Cad binding sites in the nearby sequences within the enhancer (Vincent et al., 2018). We suggest that the maternal Hb can counter-activate Cad in embryos from *bcd osk tsl* females, resulting in the uniform repression of *gt* in these embryos. Despite the uniform expression of both Hb and Cad in *bcd osk tsl* embryos, a general reduction in binding of PcG proteins to *gt* regulatory region was observed upon the knock down of Hb in *bcd osk tsl* embryos. This observation proposes that Hb represses the inhibitory effect of activator Cad on the recruitment of PcG proteins, therefore, in the absence of Hb (HbKD; *bcd osk tsl*), Cad was then able to confer its inhibitory effects on the recruitment of PcG.

H3K27ac is mutually exclusive with H3K27me3 and was shown to antagonize PRC2-mediated H3K27me3 and consequently PcG silencing (Tie et al., 2009). Mutation of replication-independent H3.3K27 in *Drosophila*, indicated that H3.3K27 is dispensable for transcription, while it is required for PcG-mediated repression (Leatham-Jensen et al., 2019). Furthermore, H3K27ac does not play a central role for gene activation but mainly functions to antagonize H3K27me (Pengelly et al., 2013).

H3K27ac is deposited at all the tested regions of the *gt* regulatory region of the active state with a peak at region 4 (*gt* promoter and PRE1). The positive signal of H3K27ac at the promoter of active *gt,* may be due to the recruitment of CBP, a H3K27 acetyltransferase. CBP has been shown to function as a Bcd and Cad co-activator at *gt* and *fushi tarazu* loci, respectively (Shapira

et al., 2014). Furthermore, CBP genomic sites have been found at active promoters and enhancers as well as PREs of both active and silenced PcG-target genes (Tie et al., 2009; Philip et al., 2015). CBP binds to all *gt* enhancers in wt embryos (modENCODE). CBP can possibly be recruited as a TrxG subunit to *gt_(-1)* and *gt_(-6)* enhancers, which overlap with TRE/PREs. Alternatively, It may also be recruited as a co-activator for Bcd to *gt_(-6)* and *gt_(-10)* enhancers and for Cad to *gt_(-1)* and *gt_(-3)* enhancers. However, the presence of TrxG complex at *gt* has not been studied.

Hb was shown to interact with dMi-2, a member of NuRD nucleosome remodeling and deacetylation complex, in two-hybrid assays (Kehle et al., 1998). *Drosophila* NuRD complex comprises helicase-containing nucleosome remodeling ATPases Mi-2, histone deacetylases Rpd3, histone chaperones p55 and GATA-type zinc finger proteins such as MBD-like, MTA-like and Simjang (Zhang et al., 2016). Reduced binding of Hb to *gt_(-1)* and *gt_(-3)* enhancers upon the knock down of Hb in *bcd osk tsl* embryos, may result in the absence of histone deacetylase activity of NuRD and mildly positive signal for H3K27ac at these regions which could inhibit PcG-mediated methylation. The antagonizing effect of Hb on the Cad-mediated negative regulation of the recruitment of PcG proteins at *gt*, may be well explained by the deacetylation activity of NuRD complex on the H3K27ace histone mark deposited by CBP. However, the presence of NuRD has never been investigated at *gt* and our attempts to identify NuRD complex by using dMi-2 antibody failed due to the inefficiency of the latter antibody in our ChIP protocol.

The only study supporting the instructive model in *Drosophila* showed the derepression of *Ubx* in *hb* and PRC1 mutant embryos, and the ectopic expression was more extensive in the absence of dMi-2 (Kehle et al., 1998). The result of this study was interpreted as an evidence on the positive role of a repressive transcription factor or its co-repressor in the recruitment of PcG proteins. However, based on the results of our study, we suggest that the deacetylation activity of

dMi-2 containing NuRD complex, contributes to the facilitated binding of PcG proteins in the presence of an activator.

Despite the efficient knock down of Cad, and the absence of Bcd, *gt* was still expressed in Cad; HbKD *bcd osk tsl* embryos at low levels (33% relative to wt embryos). However, the increased recruitment of PcG proteins in Cad; HbKD *bcd osk tsl* embryos, where *gt* is expressed at low levels, compared to *bcd osk tsl* embryos, where *gt* is uniformly repressed, provides another evidence on the irresponsiveness of PcG proteins to the transcriptional state of *gt*.

Zld not only facilitates activation by transcriptional activators, it can also cause low levels of transcription in earlier stages of embryogenesis. Zld was shown to activate low expression level of *miR-309* in the absence of its known activator, *bcd,* in blastoderm stage (Fu et al., 2014). Low *gt* expression level in Cad; HbKD *bcd osk tsl* embryos may be triggered by Zld which binds to *gt* promoter (Alhaj Abed J, unpublished data).

A dramatic increase in the binding of Pho, Pcl, Pc and RNAPII S5p and a mile elevation in the signals of E(z) and H3K27me3 were observed at region 4 of Cad; HbKD *bcd osk tsl* compared to CadKD; *bcd osk tsl* embryos. We have determined that Hb binds strongly to region 4 of nc13 CadKD; *bcd osk tsl* and *bcd osk tsl* embryos although at much lower levels compared to *gt*_(-3) enhancer. NuRD complex increases the local nucleosome density at regulatory sequences, which in turn interferes with the association of transcription factors, and co-activators with the chromatin. NuRD-mediated nucleosome rearrangements results in the decrease of RNAPII S5p occupancy at TSS, irrespective of the transcriptional state of the associated gene (Bornelov et al., 2018). The latter study explains the increased signal of RNAPII S5p at promoter region of Cad; HbKD *bcd osk tsl*, despite the low level of *gt* expression in these embryos. We suggest that the decreased Hb, and consequently NuRD binding, to region 4 in Cad; HbKD *bcd osk tsl* embryos,

may contribute to the generation of more accessible or open chromatin and the increased binding of PcG proteins to this region.

In conclusion, the ChIP assays on Cad; HbKD *bcd osk tsl* embryos suggested that the default state of native chromatin is to recruit PcG proteins and the presence of an activator may directly or indirectly, through the recruitment of a co-activator, result in the deposition of active histone marks which antagonize stable binding or catalytic activity of PcG complexes.

CHAPTER 5:

CONCLUSIONS AND FUTURE DIRECTIONS

5.1 PRE binding proteins contribute differentially to PRE activity at *gt*

Recruitment of *Drosophila* PRC2 and PRC1 to their target genes needs the presence of one or more PREs (Simon et al., 1993). PREs are complex DNA elements which vary greatly in sequence composition and size (Kassis and Kennison, 2010). Individual PREs may be bound by different combinations of various PRE binding proteins. Furthermore, their surrounding chromatin environment or topology can also add to the uniqueness of a PRE.

We studied the binding of PcG proteins at *gt* PRE1 in the absence of Pho binding and also the dependency of PRE2 on PRE1 for the recruitment of PcG proteins and the maintenance of PcG-mediated repression. We found that PRE1 is redundant with PRE2 for the recruitment of PcG proteins and maintenance of PcG-mediated repression and Pho plays the major role in PRE1 activity. We further showed that Phol binding is less dependent on the presence of consensus Pho-Phol binding sites and appears to play a minimal role in recruiting other PcG proteins and maintenance of transcriptional repression of *gt*. PRE binding proteins Spps and Dsp1 show differential dependence on the presence of Pho for PRE1 binding. Dsp1 binds PRE1 independently from Pho binding, whereas Spps binding is dependent on Pho. Phol, Spps and Dsp1 mainly bind strongly to PRE1 and very weakly to PRE2. Moreover, stable binding of most tested PcG proteins to PRE1 precedes that to PRE2. Our results demonstrated the heterogeneity and complexity of PREs, albeit for the same gene, in terms of function and the proteins which contribute to their activity and how much we have yet to learn.

5.2 Initial recruitment of PcG proteins to *gt* is dictated by the absence of a transcriptional activator

PcG proteins do not initiate transcriptional silencing but maintain the transcriptional repression of silenced genes by altering chromatin structure (Simon and Kingston, 2013). After initial recognition and binding of PcG proteins to their repressed target genes, they are able to maintain the transcriptional repression through an unlimited number of cell cycles. How these proteins initially distinguish between the active and repressed states of their target genes remains a major gap in understanding of PcG proteins.

One important problem in studying the initial recruitment of PcG proteins to target genes has been the inability to obtain temporally synced populations of cells in which a target gene is uniformly repressed by PcG proteins. PcG-target genes are usually expressed heterogeneously in *Drosophila* embryos. Therefore, their presence in both active and PcG-mediated silenced states would result in the confusing data from ChIP assays of chromatin proteins. We have circumvented this technical problem by using *Drosophila* genetics to produce embryos in which the PcG target gene, *gt*, is either ubiquitously transcriptionally repressed or active.

Within the topic of initiation, the mechanisms by which PcG proteins distinguish between repressed versus active states of target genes are not understood. Construction of embryos with a transcriptionally inert *gt* transgene in a background in which endogenous *gt* is transcriptionally active, uniquely positioned us to answer a very important but unanswered question concerning the mechanisms by which PcG proteins discern the transcriptionally repressed and active states of their target genes. The results of our studies showed that PcG proteins do not respond to the transcriptional state of *gt* and their recruitment is solely dictated by the transcription factors. Moreover, we demonstrated that the activating transcription factor Cad confers inhibitory effects

on the recruitment of PcG proteins and this inhibition is impeded by the repressive transcription factor Hb.

5.3 Future directions

- Investigating if Cad recruits CBP at *gt* and the recruitment of PcG proteins at *gt* in embryos in which CBP level is knocked down in *bcd osk tsl* background. Comparing PcG binding to *gt* in CBP KD; *bcd oks tsl* and CadKD; *bcd osk tsl* embryos, will provide us with more information about the mechanism that Cad reduces binding of PcG proteins to *gt*.
- Determining whether NuRD is present at *gt*. NuRD is a co-repressor of Hb, but it is not known whether Hb recruits it to *gt*. The NuRD complex includes dMi-2 and histone deacetylases. The antagonistic effect of Hb on Cad inhibition of PRC2 recruitment may involve deacetylation of H3K27ac. In case of the presence of NuRD at *gt*, it will be interesting to investigate the recruitment of PcG proteins in embryos in which the level of dMi-2 is knocked down in *bcd osk tsl* background. This study may shed light on the mechanism by which Hb represses *gt* and prevents the inhibitory effects of Cad on the recruitment of PcG proteins.
- Hb and Cad are the repressor and activator of *kni*, a PcG target gene. *kni* is uniformly repressed or expressed at levels comparable to *gt* in *bcd osk tsl*, HbKD; *bcd osk tsl*, CadKD; *bcd osk tsl* and Cad; HbKD *bcd osk tsl* embryos (data not shown). Surveying the recruitment of PcG proteins at *kni* in the aforementioned genetic systems, will provide us with more insights into how general the results of our study is and if they can be applied to other PcG target genes.
- Determining the effects of the absence of PRE binding proteins Spps and/or Dsp1 on the recruitment of PcG proteins to *gt*.

BIBLIOGRAPHY

Abed, J. A., Ghotbi, E., Ye, P., Frolov, A., Benes, J., Jones, R. S. (2018) De novo recruitment of Polycomb-group proteins in *Drosophila* embryos. Development 145(23), dev165027.

Agresti, A., Bianchi, M. E. (2003) HMGB proteins and gene expression. Current Opinion in Genetics and Development 13(2),170-178.

AlHaj Abed, J., and Jones, R. S. (2012) H3K36me3 key to Polycomb-mediated gene silencing in lineage specification. Nature Structural and Molecular Biology 19(12), 1214-1215.

AlHaj Abed, J., Cheng, C. L., Crowell, C. R., Madigan, L. L., Onwuegbuchu, E., Desai, S., Benes, J., and Jones, R. S. (2013) Mapping Polycomb Response Elements at the *Drosophila melanogaster giant* locus. G3: Genes, Genomes, Genetics 3(12), 2297-2304.

Americo, J., Whiteley, M., Brown, J. L., Fujioka, M., Jaynes, J. B., Kassis, J. A. (2002) A complex array of DNA-binding proteins required for pairing-sensitive silencing by a Polycomb group response element from the *Drosophila engrailed* gene. Genetics 160(4), 1561-1571.

Ansari, K. I., Mishra, B. P., Mandal, S. S. (2009) MLL histone methylases in gene expression, hormone signaling and cell cycle. Frontiers in Bioscience 14, 3483-3495.

Arnold, P., Scholer, A., Pachkov, M., Balwierz, P. J., Jorgensen, H., Stadler, M. B., Nimwegen, E. V., Schubeler, D. (2013) Modeling of epigenome dynamics identifies transcription factors that mediate Polycomb targeting. Genome Research 23(1), 60-73.

Atallah, J., Vurens, G., Mavong, S., Mutti, A., Hoang, D., and Kopp, A. (2014) Sex-specific repression of *dachshund* is required for *Drosophila* sex comb. Developmental Biology 386(2), 440-447.

Atchison, L., Ghias, A., Wilkinson, F., Bonini, N., Atchison, M. L. (2003) Transcription factor YY1 functions as a PcG protein in vivo. The EMBO Journal 22(6), 1347-1358.

Ballaré, C., Lange, M., Lapinaite, A., Martin, G. M., Morey, L., Pascual, G., Liefke, R., Simon, B., Shi, Y., Gozani, O., Carlomagno, T., Benitah, S. A., and Croce L. D. (2012) Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity. Nature Structural and Molecular Biology 19(12), 1257-1265.

Bantignies, F., Roure, V., Comet, I., Leblanc, B., Schuettengruber, B., Bonnet, J., Tixier, V., Mas, A., and Cavalli, G. (2011) Polycomb-dependent regulatory contacts between distant Hox loci in *Drosophila*. Cell 144(2), 214-226.

Baumann, M., Pontiller, J., Ernst, W. (2010) Structure and basal transcription complex of RNA polymerase II core promoters in the mammalian genome: an overview. Molecular Biotechnology 45(3), 241-247.

Beltran, M., Yates, C. M., Skalska, L., Dawson, M., Reis, F. P., Viiri, K., Fisher, C. L., Sibley, C. R., Foster, B. M., Bartke, T., Ule, J., and Jenner, R. G. (2016) The interaction of PRC2 with RNA or chromatin is mutually antagonistic. Genome research 26(7), 896-907.

Bensaude, O. (2011) Inhibiting eukaryotic transcription: Which compound to choose? How to evaluate its activity? Which compound to choose? How to evaluate its activity? Transcription 2(3), 103-108.

Berleth, T., Burri, M., Thoma, G., Bopp, D., Richstein, S., Frigerio, G., Nell, M., Niisslein-Volhard, C. (1988) The role of localization of *bicoid* RNA in organizing the anterior pattern of the *Drosophila* embryo. The EMBO Journal 7(6),1749-1756.

Berman, B. P., Nibu, y., Pfeiffer, B. D., Tomancak, P., Celniker, S. E., Levine, M., Rubin, G. M., and Eisen, M. B. (2002) Exploiting transcription factor binding site clustering to identify cis regulatory modules involved in pattern formation in the *Drosophila* genome. Proceeding of the National Academy of Sciences. USA 99(2), 757-762.

Black, L. K., Petruk, S., Fenstermaker, T. K., Hodgson, J. W., Caplan, J. L., Brock, H. W., and Mazo, A. (2016) Chromatin proteins and RNA are associated with DNA during all phases of mitosis. Cell Discovery 2, 16038.

Blackledge, N. P., Farcas, A. M., Kondo, T., King, H. W., McGouran, J. F., Hanssen, L. L. P., Ito, S., Cooper, S., Kondo, K., Koseki, Y., Ishikura, T., Long, H. K., Sheahan, T. W., Brockdorff, N., Kessler, B. M., Koseki, H., Klose, R. J. (2014) Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and Polycomb domain formation. Cell 157(6), 1445-1459.

Blastyak, A., Mishra, R. K., Karch, F., Gyurkovics, H. (2006) Efficient and specific targeting of Polycomb group proteins requires cooperative interaction between Grainyhead and Pleiohomeotic. Molecular and Cellular Biology 26(4),1434-1444.

Blythe, S. A., and Wieschaus, E. F. (2016) Establishment and maintenance of heritablechromatin structure during early *Drosophila* embryogenesis. Elife 5, e20148.

Bornelöv, S., Reynolds, N., Xenophontos, M., Gharbi, S., Johnstone, E., Floyd, R., Ralser, M., Signolet, J., Loos, R., Dietmann, S., Bertone, P., Hendrich, B. (2018) The Nucleosome Remodeling and Deacetylation Complex Modulates Chromatin Structure at Sites of Active Transcription to Fine-Tune Gene Expression. Molecular Cell 71(1), 56-72, e4.

Bornemann, D., Miller, E., Simon, J. (1998) Expression and properties of wild-type and mutant forms of the *Drosophila* sex comb on midleg (SCM) repressor protein. Genetics 150(2), 675-686.

Breen, T. R., Duncan, I. M. (1986) Maternal expression of genes that regulate the *bithorax* complex of *Drosophila melanogaster*. Developmental Biology 118(2), 442-456.

Brien, G. L., Gambero, G., O'Connell, D. G., Jerman, E., Turner, S. A., Egan, C. M., Dunne, E, J., Jurgens, M. C., Wynne, K., Piao, L., Lohan, A. J., Ferguson, N., Shi, X., Sinha, K. M., Loftus, B. J., Cagney, G., and Bracken, a. p. (2012) Polycomb PHF19 binds H3K36me3 and recruits PRC2 and demethylase NO66 to embryonic stem cell genes during differentiation. Nature Structural and Molecular Biology 19(12), 1273-1281.

Brockdorff, N. (2013) Noncoding RNAs and Polycomb recruitment. RNA 19(4), 429-442.

Brown, J. L., Fritsch, C., Müller, J., Kassis, J. A. (2003) The *Drosophila pho-like* gene encodes a YY1-related DNAbinding protein that is redundant with *pleiohomeotic* in homeotic gene silencing. Development 130(2), 285-294.

Brown, J. L., Grau, D. J., DeVido, S. K., Kassis, J. A. (2005) An Sp1/KLF binding site is important for the activity of a Polycomb group response element from the *Drosophila* engrailed gene. Nucleic Acids Research 33(16), 5181-5189.

Brown, J. L., Kassis, J. A. (2010) Spps, a *Drosophila* Sp1/KLF family member binds to PREs and is required for PRE activity late in development. Development 137(15), 2597-2602.

Brown, J. L., Mucci, D., Whiteley, M., Dirksen, M. L., Kassis, J. A. (1998) The *Drosophila* Polycomb group gene *pleiohomeotic* encodes a DNA binding protein with homology to the transcription factor YY1. Molecular Cell 1(7), 1057-1064.

Brown, J. L., Sun, M. A., Kassis, J. A. (2018) Global changes of H3K27me3 domains and Polycomb group protein distribution in the absence of recruiters Spps or Pho. Proceeding of the National Academy of Sciences 115(8), 1839-1848.

Brunk, B. P., Martin, E. C., and Adler, P. N. (1991) *Drosophila* genes *Posterior sex combs* and *Suppressor two of zeste* encode proteins with homology to the murine *bmi-1* oncogene. Nature 353(6342), 351-353.

Calvo, O., Manley, J. L. (2003) Strange bedfellows: Polyadenylation factors at the promoter. Genes and Development 17(11), 1321-1327.

Campos-Ortega, J. A., Hartenstein, V. (1985) The embryonic development of *Drosophila melanogaster*. Berlin: Springer-Verlag Berlin and Heidelberg Gmbh and Co. K., Print.

Cang, Y., Auble, D. T., Prelich, G. (1999) A new regulatory domain on the TATA-binding protein. The EMBO Journal 18(23), 6662-6671.

Cao, R., and Zhang, Y. (2004) The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. Current Opinion in Genetics and Development 14(2), 155-164.

Cao, R., Wang, H., He, J., Erdjument-Bromage, H., Tempst, P., Zhang, Y. (2008) Role of hPHF1 in H3K27 methylation and Hox gene silencing. Molecular and cellular Biology 28(5), 1862-1872.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S., and Zhang, Y. (2002) Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science 298(5595), 1039-1043.

Caputo, V. S., Costa, J. R., Makarona, K., Georgiou, E., Layton, D. M., Roberts, I., Karadimitris, A. (2013) Mechanism of Polycomb recruitment to CpG islands revealed by inherited disease-associated mutation. Human Molecular Genetics. 22(16), 3187–3194

Casanova, M., Preissner, T., Cerase, A., Poot, R., Yamada, D., Li, X., Appanah, R., Bezstarosti, K., Demmers, J., Koseki, H., Brockdorff, N. (2011) Polycomb like 2 facilitates the recruitment of PRC2 Polycomb group complexes to the inactive X chromosome and to target loci in embryonic stem cells. Development 138(8), 1471-1482.

Cavalli, G., Paro, R. (1998) The *Drosophila* Fab-7 chromosomal element conveys epigenetic inheritance during mitosis and meiosis. Cell 93(4), 505-518.

Cavalli, G., Paro, R. (1999) Epigenetic inheritance of active chromatin after removal of the main transactivator. Science 286(5441), 955-958.

Chamberlain, S. J., Yee, D., Magnuson, T. (2008) Polycomb repressive complex 2 is dispensable for maintenance of embryonic stem cell pluripotency. Stem Cells 26(6),1496-1505.

Chan, C. S., Rastelli, L., and Pirrotta, V. (1994) A Polycomb response element in the *Ubx* gene that determines an epigenetically inherited state of repression. The EMBO Journal 13(11), 2553-2564.

Chen, S., Bohrer, L. R., Rai, A. N., Pan, Y., Gan, L., Zhou, X., Bagchi, A., Simon, J. A., and Huang, H. (2010) Cyclin-dependent kinases regulate epigenetic gene silencing through phosphorylation of EZH2. Nature Cell Biology 12(11), 1108-1114.

Chiang, A., O'Connor, M. B., Paro, R., Simon, J., and Bender, W. (1995) Discrete Polycomb-binding sites in each parasegmental domain of the bithorax complex. Development 121(6),1681-1689.

Clyde, D. E., Corado, M. S. G., Wu, X., Pare, A., Papatsenko, D., and Small, S. (2003) A self organizing system of repressor gradients establishes segmental complexity in *Drosophila*. Nature 426, 849-853.

Cooper, S., Grijzenhout, A., Underwood, E., Ancelin, K., Zhang, T., Nesterova, T. B., Anil-Kirmizitas, B., Bassett, A., Kooistra, S. M., Agger, K., Helin, K., Heard, E., and Neil Brockdorff, N. (2016) Jarid2 binds mono-ubiquitylated H2A lysine 119 to mediate crosstalk between Polycomb complexes PRC1 and PRC2. Nature Communications 7,13661.

Cunningham, M. D., Brown, J. L., Kassis, J. A. (2010) Characterization of the Polycomb group response elements of the *Drosophila melanogaster invected* Locus. Molecular and Cellular Biology 30(3), 820-828.

Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A., and Pirrotta, V. (2002) *Drosophila* enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. Cell 111(2), 185-196.

De, S., Cheng, Y., Sun, M., Gehred, N. D., Kassis, J. A. (2019) Structure and function of an ectopic Polycomb chromatin domain. Science Advances 5(1), eaau9739.

Deaton, A. M., Bird, A. (2011) CpG islands and the regulation of transcription. Genes and Development 25(10), 1010-1022.

Decoville, M., Giacomello, E., Leng, M., Locker, D. (2001) DSP1, an HMG-like protein, is involved in the regulation of homeotic genes. Genetics 157(1), 237-244.

Dejardin, J., Rappailles, A., Cuvier, O., Grimaud, C., Decoville, M., Locker, D., Cavalli, G. (2005) Recruitment of *Drosophila* Polycomb group proteins to chromatin by DSP1. Nature 434(7032), 533-538.

Deng, W., Roberts, S. G. (2007) TFIIB and the regulation of transcription by RNA polymerase II. Chromosoma 116(5), 417-429.

DeVido, S. K., Kwon, D., Brown, J. L., Kassis, J. A. (2008) The role of Polycomb-group response elements in regulation of engrailed transcription in *Drosophila*. Development 135(4), 669-676.

Di Croce, L., Helin, K. (2013) Transcriptional regulation by Polycomb group proteins. Nature Structural and Molecular Biology 20(10), 1147-1155.

Dietrich, N., Lerdrup, M., Landt, E., Agrawal-Singh, S., Bak, M., Tommerup, N., Rappsilber, J., Södersten, E., Hansen, K. (2012) REST–Mediated Recruitment of Polycomb Repressor Complexes in Mammalian Cells. PLOS Genetics 8(3), e1002494.

Dovey, J. S., Zacharek, S. J., Kim, C. F., Lees, J. A. (2008) Bmi1 is critical for lung tumorigenesis and bronchioalveolar stem cell expansion. Proceeding of the National Academy of Sciences USA 105(33),11857-11862.

Driever, W., Thoma, G., and Nusslein-Volhard, C. (1989) Determination of spatial domains of

zygotic gene expression in the *Drosophila* embryo by the affinity of binding sites for the bicoid morphogen. Nature 340(6232), 363-367.

Dubnau, J., and Struhl, G. (1996) RNA recognition and translational regulation by a homeodomain protein. Nature 379(6567), 694-699.

Dubois, M. F., Nguyen, V. T., Bellier, S., Bensaude, O. (1994) Inhibitors of transcription such as 5, 6- dichloro-1-beta -D- ribofuranosylbenzimidazole and isoquinoline sulfonamide derivatives (H-8 and H-7) promote dephosphorylation of the carboxyl-terminal domain of RNA polymerase II largest subunit. Journal of Biological Chemistry 269(18), 13331-13336.

Edgar, B. A., O'Farrell, P. H. (1989) Genetic control of cell division patterns in the *Drosophila* embryo. Cell 57(1), 177-187.

Edgar, B. A., O'Farrell, P. H. (1990) The three postblastoderm cell cycles of *Drosophila* embryogenesis are regulated in G2 by string. Cell 62(3), 469-480.

Edgar, B. A., Sprenger, F., Duronio, R. J., Leopold, P., O'Farrell, P. H. (1994) Distinct molecular mechanism regulate cell cycle timing at successive stages of *Drosophila* embryogenesis. Genes andDevelopment 8(4), 440-452.

Eldon, E. D., and Pirrotta, Z. (1991) Interactions of the *Drosophila* gap gene *giant* with maternal and zygotic pattern-forming genes. Development 111(2), 367-378.

Emmons, R. B., Genetti, H., Filandrinos, S., Lokere, J., and Wu, C. T. (2009) Molecular genetic analysis of suppressor 2 of zeste identifies key functional domains. Genetics 182(4), 999-1013.

Enderle, D., Beisel, C., Stadler, M. B., Gerstung, M., Athri, P., Paro, R. (2011) Polycomb preferentially targets stalled promoters of coding and noncoding transcripts. Genome Research 21(2), 216-226.

Farcas, A. M., Blackledge, N. P., Sudbery, I., Long, H. K., McGouran, J. F., Rose, N. R., Lee, S., Sims, D., Cerase, A., Sheahan, T. W., Koseki, H., Brockdorff, N., Ponting, C. P., Kessler, B. M., Klose, R. J. (2012) KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands. Elife 1, e00205.

Farrell, J. A., O'Farrell, P. H. (2014) From egg to gastrula: how the cell cycle is remodeled during the *Drosophila* mid-blastula transition. Annual Review of Genetics 48, 269-294.

Farrell, J. A., Shermoen, A. W., Yuan, K., O'Farrell, P. H. (2012) Embryonic onset of late replication requires Cdc25 down-regulation. Genes and Development 26(7), 714-725.

Ferrari, K. J., Scelfo, A., Jammula, S., Cuomo, A., Barozzi, I., Stützer, A., Fischle W., Bonaldi, T., Pasini, D. (2014) Polycomb-dependent H3K27me1 and H3K27me2 regulate active transcription and enhancer fidelity. Molecular Cell 53(1), 49-62.

Fiedler, T., Rehmsmeier, M. (2006) jPREdictor: a versatile tool for the prediction of cis-regulatory elements. Nucleic Acids Research 34(supple-2), W546-W550.

Fischle, W., Wang, Y., and Allis, C. D. (2003) Histone and chromatin cross-talk. Current Opinion in Cell Biology 15(2), 172-183.

Fischle, W., Wang, Y., Jacobs, S. A., Kim, Y., Allis, C. D., and Khorasanizadeh, S. (2003) Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. Genes and Development 17(15), 1870-1881.

Foe, V. E. (1989) Mitotic domains reveal early commitment of cells in *Drosophila* embryos. Development 107(1), 1-22.

Foe, V. E., Odell, G. M., Edgar, B. A. (1993) Mitosis and morphogenesis in the *Drosophila* embryo: point and counterpoint. The Development of *Drosophila* melanogaster. Cold Spring Harbor NY: Cold Spring Harbor Laboratory (1), 149-300.

Foo, S. M., Sun, Y., Lim, B., Ziukaite, R., O'brien, K., Nien, C. Y., Kirov, N., Shvartsman, S. Y., and Rushlow, C. A. (2014) Zelda potentiates morphogen activity by increasing chromatin accessibility. Current Biology 24(12), 1341-1346.

Fraenkel, E., Cantor, A. B. (2012) Direct recruitment of Polycomb Repressive Complex 1 to chromatin by core binding transcription factors. Molecular Cell 45(3), 330-343.

Frey, F., Sheahan, T., Finkl, K., Stoehr, G., Mann, M., Benda, C., Muller, J. (2016) Molecular basis of PRC1 targeting to Polycomb response elements by PhoRC. Genes andDevelopment 30(9), 1116-1127.

Fritsch, C., Brown, J. L., Kassis, J. A., and Muller, J. (1999) The DNA-binding polycomb group protein Pleiohomeotic mediates silencing of a *Drosophila* homeotic gene. Development 126(17), 3905-3913.

Fu, S., Nien, C. Y., Liang, H. L., and Rushlow, C. (2014) Co-activation of microRNAs by Zelda is essential for early *Drosophila* development. Development 141(10), 2108-2118.

Fujioka, M., Yusibova, G. L., Zhou, J., Jaynes, J. B. (2008) The DNA-binding Polycomb-group protein Pleiohomeotic maintains both active and repressed transcriptional states through a single site. Development 135(24), 4131- 4139.

Gambetta, M. C., and Müller, J. (2014) O-GlcNAcylation prevents aggregation of the Polycomb group repressor polyhomeotic. Developmental Cell 31(5), 629-639.

Gardini, A., Baillat, D., Cesaroni, M., Hu, D., Marinis, J. M., Wagner, E. J., Lazar, M. A., Shilatifard, A., Shiekhattar, R. (2014) Integrator regulates transcriptional initiation and pause release following activation. Molecular Cell 56(1), 128-139.

Garrick, D., Higgs D. R. (2012) An interspecies analysis reveals a key role for unmethylated CpG dinucleotides in vertebrate Polycomb complex recruitment. The EMBO Journal 31(2), 317-329.

Geisler, S. J., and Paro, R. (2015) Trithorax and Polycomb group-dependent regulation: a tale of opposing activities, Development 142(17), 2876-2887.

Gilbert, S. F. (2000) Developmental Biology (6th ed.). Sunderland (MA): Sinauer Associates. pp. The origins of anterior-posterior polarity. Retrieved 23 October 2015.

Godlewski, J., Nowicki, M. O., Bronisz, A., Williams, S., Otsuki, A., Nuovo, G., RayChaudhury, A., Newton, H. B., Chiocca, E. A., and Lawler, S. (2008) Targeting of the Bmi-1 oncogene/stem cell renewal factor by microRNA-128 inhibits glioma proliferation and self-renewal. Cancer Research 68(22), 9125-9130.

Goldberg, M. L., Colvin, R. A., Mellin, A. F. (1989) The *Drosophila zeste* locus is non-essential. Genetics 123(1),145-155.

Gomes, N. P., Bjerke, G., Llorente, B., Szostek, S. A., Emerson, B. M., Espinosa, J. M. (2006) Gene-specific requirement for P-TEFb activity and RNA polymerase II phosphorylation within the p53 transcriptional program. Genes and Development 20(5), 601-612.

Grijzenhout, A., Godwin, J., Koseki, H., Gdula, M. R., Szumska, D., McGouran, J. F., Bhattacharya, S., Kessler, B. M., Brockdorff, N., and Cooper, S. (2016) Functional analysis of AEBP2, a PRC2 Polycomb protein, reveals a Trithorax phenotype in embryonic development and in ESCs. Development 143(15), 2716-2723.

Grimm, C., de Ayala Alonso, A. G., Rybin, V., Steuerwald, U., Ly-Hartig, N., Fischle, W., Müller, J., Müller, C. W. (2007) Structural and functional analyses of methyl-lysine binding by the malignant brain tumour repeat protein sex comb on midleg. EMBO Reports 8(11), 1031-1037.

Grimm, C., Matos, R., Ly-Hartig, N., Steuerwald, U., Lindner, D., Rybin, V., Muller, J., Muller, C. W. (2009) Molecular recognition of histone lysine methylation by the Polycomb group repressor dSfmbt. The EMBO Journal 28(3), 1965-1977.

Grossniklaus, U., Paro, R. (2014) Transcriptional silencing by polycomb-group proteins. Cold Spring Harbor Perspectivies in Biology 6(11), a019331.

Hager, G. L., McNally, J. G., Misteli, T. (2009) Transcription dynamics. Molecular Cell 35(6), 741-753.

Han, J., Yuan, P., Yang, H., Zhang, J., Soh, B. S., Li, P., Lim, S. L., Cao, S., Tay, J., Orlov, Y. L., Lufkin, T., Ng, H. H., Tam, W. L., Lim, B. (2010) Tbx3 improves the germ-line competency of induced pluripotent stem cells. Nature 463(7284), 1096-1100.

Harrison, M. M., Li, X. Y., Kaplan, T., Botchan, M. R., Eisen, M. B. (2011) Zelda binding in the early *Drosophila melanogaster* embryo marks regions subsequently activated at the maternal-to-zygotic transition. PLOS Genetics 7(10), e1002266.

Haynie, J. L. (1983) The maternal and zygotic roles of the gene Polycomb in embryonic determination in *Drosophila melanogaster*. Developmental Biology 100(2), 399-411.

He, X. T., Cao, X. F., Ji, L., Zhu, B., Lv, J., Wang, D. D., Lu, P. H., and Cui, H. G. (2009) Association between Bmi1 and clinicopathological status of esophageal squamous cell carcinoma. World Journal of Gastroenterol 15(19), 2389-2394.

He, X., Samee, M. A. H., Blatti, C., Sinha, S. (2010) Thermodynamics-based models of transcriptional regulation by enhancers: the roles of synergistic activation, cooperative binding and short-range repression. PLOS Computational Biology 6(9), e1000935

Henry, K. W., Wyce, A., Lo, W. S., Duggan, L. J., Emre, N. C., Kao, C. F., Pillus, L., Shilatifard, A., Osley, M. A., Berger, S. L. (2003) Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. Genes and Development 17(21), 2648-2663.

Hoermann, A., a,b, Cicin-Sain, D., a,b, Jaeger, J. (2016) A quantitative validated model reveals two phases of transcriptional regulation for the gap gene *giant* in *Drosophila*. Developmental Biology 411(2), 325-338.

Huang, D. H., Chang, Y. L. (2004) Isolation and characterization of CHRASCH, a Polycomb-containing silencing complex. Methods in Enzymology 377, 267-282.

Huang, D. H., Chang, Y. L., Yang, C. C., Pan, I. C., King, B. (2002) pipsqueak encodes a factor essential for sequence-specific targeting of a polycomb group protein complex. Molecular and Cellular Biology 22(17), 6261-6271.

Hülskamp, M., Pfeifle, C., Tautz, D. (1990) A morphogenetic gradient of hunchback protein organizes the expression of the gap genes *Krüppel* and *knirps* in the early *Drosophila* embryo. Nature 346(6284), 577-580.

Hulskamp, M., Tautz, D. (1991) Gap genes and gradients–the logic behind the gaps. BioEssays 13(6), 261-268.

Hunkapiller, J., Shen, Y., Diaz, A., Cagney, G., McCleary, D., Ramalho-Santos, M., Krogan, N., Ren, B., Song, J. S., Reiter, J. F. (2012) Polycomb-like 3 promotes polycomb repressive complex 2 binding to CpG islands and embryonic. PLOS Genetics 8(3), e1002576.

Hur, M. W., Laney, J. D., Jeon, S. H., Ali, J., Biggin, M. D. (2002) Zeste maintains repression of *Ubx* transgenes: Support for a new model of Polycomb repression. Development 129(6),1339-1343.

Isono, K., Endo, T. A., Ku, M., Yamada, D., Suzuki, R., Sharif, J., Ishikura, T., Toyoda, T., Bernstein, B. E., Koseki, H. (2013) SAM domain polymerization links subnuclear clustering of PRC1 to gene silencing. Developmental Cell 26(6), 565-577.

Itzen, F., Greifenberg, A. K., Bosken, C. A., Geyer, M. (2014) Brd4 activates P-TEFb for RNA polymerase II CTD phosphorylation. Nucleic Acids Research 42(12), 7577-7590.

Jaeger, J., Blagov, M., Kosman, D., Kozlov, K. N., Myasnikova, E., Surkova, S., Vanario-Alonso, C. E., Samsonova, M., Sharp, D. H., and Reinitz, J. (2004) Dynamical analysis of regulatory interactions in the gap gene system of *Drosophila melanogaster*. Genetics 167(4), 1721-1737.

Jenuwein, T., Laible, G., Dorn, R., and Reuter, G. (1998) SET domain proteins modulate chromatin domains in eu- and heterochromatin. Cellular and Molecular Life Sciences CMLS 54(1), 80-93.

Jermann, P., Hoerner. L., Burger, L., Schubeler, D. (2014) Short sequences can efficiently recruit histone H3 lysine 27 trimethylation in the absence of enhancer activity and DNA methylation. Proceeding of the National Academy of Sciences. USA 111(33), E3415-E3421.

Jiao, L., Liu, X. (2015) Structural basis of histone H3K27 trimethylation by an active Polycomb repressive complex 2. Science 350(6285), aac4383.

Jürgens, G. (1985) A group of genes controlling the spatial expression of the bithorax complex in *Drosophila*. Nature 316(6024), 153-155.

Justin, N., Zhang, Y., Tarricone, C., Martin, S. R., Chen, S., Underwood, E., Marco, V. D., Haire, L. F., Walker, P. A., Reinberg, D., Wilson, J. R., Gamblin, S. J. (2016) Structural basis of oncogenic histone H3K27M inhibition of human Polycomb repressive complex 2. Nature Communications 7, 11316.

Juven-Gershon, T., Hsu, J. Y., Theisen, J. W., Kadonaga, J. T. (2008) The RNA polymerase II core promoter - the gateway to transcription. Current Opinion in Cell Biology 20(3), 253-259.

Kahn, T. G., Dorafshan, E., Schultheis, D., Zare, A., Stenberg, P., Reim, I., Pirrotta, V., and Schwartz, Y. B. (2016) Interdependence of PRC1 and PRC2 for recruitment to Polycomb response elements. Nucleic Acids Research 44(21), 10132-10149.

Kahn, T. G., Stenberg, P., Pirrotta, V., and Schwartz, Y. B. (2014) Combinatorial interactions are required for the efficient recruitment of Pho repressive complex (PhoRC) to Polycomb response elements. PLOS Genetics 10(7), e1004495.

Kalb, R., Latwiel, S., Baymaz, H. I., Jansen, P. W. T. C., Müller, C. W., Vermeulen, M., and Müller, j. (2014) Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression. Nature Structural and Molecular Biology 21(6), 569-571.

Kaneko, S., Li, G., Son, J., Xu, C. F., Margueron, R., Neubert, T. A., and Reinberg, D. (2010) Phosphorylation of the PRC2 component Ezh2 is cell cycle-regulated and up-regulates its binding to ncRNA. Genes and Development 24(23), 2615-2620.

Kaneko, S., Son, J., Bonasio, R., Shen, S. S., Reinberg, D. (2014) Nascent RNA interaction keeps PRC2 activity poised and in check. Genes and Development 28(18), 1983-1988.

Kaneko, S., Son, J., Shen, S. S., Reinberg, D., Bonasio, R. (2013) PRC2 binds active promoters and contacts nascent RNAs in embryonic stem cells. Nature Structural and Molecular Biology 20(11), 1258-1264.

Kang, H., McElroy, K. A., Jung, Y. L., Alekseyenko, A. A., Zee, B. M., Park, P. J., and Kuroda, M. I. (2015) Sex comb on midleg (Scm) is a functional link between PcG-repressive complexes in *Drosophila*. Genes andDevelopment 29(11), 1136-1150.

Karlić, R., Chung, H. R., Lasserre, J., Vlahovicek, K. and Vingron, M. (2010) Histone modification levels are predictive for gene expression. Proceeding of the National Academy of Sciences USA 107(7), 2926-2931.

Kassis, J. A. (1994) Unusual properties of regulatory DNA from the *Drosophila engrailed* gene: three "pairing-sensitive" sites within a 1.6-kb region. Genetics 136(3), 1025-1038.

Kassis, J. A. (2002) Pairing-sensitive silencing, Polycomb group response elements, and transposon homing in *Drosophila*. Advances in Genetics 46, 421- 438.

Kassis, J. A., Kennison, J. A. (2010) Recruitment of Polycomb Complexes: a Role for SCM. Molecular and Cellular Biology 30(11), 2581-2583.

Kehle, J., Beuchle, D., Treuheit, S., Christen, B., Kennison, J. A., Bienz, M., Müller, J. (1998) dMi-2, a Hunchback-interacting protein that functions in Polycomb repression. Science 282(5395), 1897-1900.

Kennison, J. A., Tamkun, J. W. (1988) Dosage-dependent modifiers of Polycomb and Antennapedia mutations in *Drosophila*. Proceedings of the National Academy of Sciences U. S. A. 85(21), 8136-8140.

Kidani, K., a,b, Osaki, M., a,c, Tamura, T., a, Yamaga, K., a, Shomori, K., a, Ryoke, K., b, Ito, H., a. (2009) High expression of EZH2 is associated with tumor proliferation and prognosis in human oral squamous cell carcinomas. Oral Oncology 45(1), 39-46.

Kim, A. R., Martinez, C., Ionides, J., Ramos, A. F., Ludwig, M. Z., Ogawa, N., Sharp, D. H., Reinitz, J. (2013) Rearrangements of 2.5 kilobases of noncoding DNA from the *Drosophila even-skipped* locus define predictive rules of genomic cis-regulatory logic. PLOS Genetics 9(2), e1003243.

Kim, C. A., Gingery, M., Pilpa, R. M., Bowie, J. U. (2002) The SAM domain of polyhomeotic forms a helical polymer. Nature Structural and Molecular Biology 9(6), 453-457.

Kim, C. A., Sawaya, M. R., Cascio, D., Kim, W., Bowie, J. U. (2005) Structural organization of a Sex-comb-on-midleg/polyhomeotic copolymer. Journal of Biological Chemistry 280(30), 27769-27775.

King, I. F. G., Emmons, R. B., Francis, N. J., Wild, B., Müller, J., Kingston, R. E., Wu, C. T. (2005) Analysis of a Polycomb group protein defines regions that link repressive activity on nucleosomal templates to in vivo function. Molecular and Cellular Biology 25(15), 6578-6591.

Kingston, R. E., Tamkun, J. W. (2014) Transcriptional regulation by trithorax- group proteins. Cold Spring Harbor Perspective in Biology 6(10), a019349.

Klose, R. J., Cooper, S., Farcas, A. M., Blackledge, N. P., Brockdorff, N. (2013) Chromatin sampling an emerging perspective on targeting polycomb repressor proteins. PLOS Genetics 9(8), e1003717.

Klymenko, T., and Müller, j. (2004) The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. EMBO Reports 5(4), 373-377.

Klymenko, T., Papp, B., Fischle, W., Köcher, T., Schelder, M., Fritsch, C., Wild, B., Wilm, M., Müller, J. (2006) A Polycomb group protein complex with sequence-specific DNA-binding and selective methyl-lysine-binding activities. Genes and Development 20(9), 1110-1122.

Komarnitsky, P., Cho, E. J., Buratowski, S. (2000) Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes and Development 14(19), 2452-2460.

Kosman, D., and Small, S. (1997) Concentration-dependent patterning by an ectopic expression domain of the *Drosophila* gap gene *knirps*. Development 124, 1343-1354.

Kotadia, S., Crest, J., Tram, U., Riggs, B., Sullivan, W. (2010) Blastoderm Formation and Cellularisation in *Drosophila melanogaster*. eLS.

Kozma, G., Bender, W., Sipos, L. (2008) Replacement of a *Drosophila* Polycomb response element core, and in situ analysis of its DNA motifs. Molecular Genetics and Genomics 279(6), 595-603.

Kraut, R., and Levine, M. (1991) Spatial regulation of the gap gene *giant* during *Drosophila* development. Development 111(2), 601-609.

Krogan, N. J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D. P., Beattie, B. K., Emili, A., Boone, C., Shilatifard, A., Buratowski, S., Greenblatt, J. (2003) Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. Molecular and Cellular Biology 23(12), 4207-4218.

Kuras, L., Struhl, K. (1999) Binding of TBP to promoters in vivo is stimulated by activators and requires Pol II holoenzyme. Nature 399(6736), 609-613.

Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2002) Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. Genes and Development 16(22), 2893-2905.

Laine, J. P., Egly, J. M. (2006) When transcription and repair meet: a complex system. Trends in Genetics 22(8), 430-436.

Landeira, D., Sauer, S., Poot, R., Dvorkina, M., Mazzarella, L., Jørgensen, H. F., Pereira, C. F., Leleu, M., Piccolo, F. M., Spivakov, M., Brookes, E., Pombo, A., Fisher, C., Skarnes, W. C., Snoek, T., Bezstarosti, K., Demmers, J., Klose, R. J., Casanova, M., Tavares, L., Brockdorff, N., Merkenschlager, M., and Fisher, A. G. (2010) Jarid2 is a PRC2 component in embryonic stem cells required for multi-lineage differentiation and recruitment of PRC1 and RNA Polymerase II to developmental regulators. Nature Cell Biology 12(6), 618-624.

Langlais, K. K., Brown, J. L., Kassis, J. A. (2012) Polycomb Group proteins bind an *engrailed* PRE in both the "ON" and "OFF" transcriptional states of *engrailed*. PLOS one 7(11), e48765.

Janssens, H., Hou, S., Jaeger, J., Kim, A. R., Myasnikova, E., Sharp, D., Reinitz, j. (2006) Quantitative and predictive model of transcriptional control of the *Drosophila melanogaster even skipped* gene. Nature Genetics 38(10), 1159.

Lanzuolo, C., Lo Sardo, F., Diamantini, A., Orlando, V. (2011) PcG complexes set the stage for epigenetic inheritance of gene silencing in early S phase before replication. PLOS Genetics 7(11), e1002370.

Lanzuolo, C., Roure, V., Dekker, J., Bantignies, F., and Orlando, V. (2007) Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. Nature Cell Biology 9(10), 1167-1174.

Laprell, F., Finkl, K., Müller, J. (2017) Propagation of Polycomb-repressed chromatin requires sequence-specific recruitment to DNA. Sience 356(6333), 85-88.

Laue, E. D. (2016) The nucleosome remodeling and deacetylase complex NuRD is built from preformed catalytically active sub-modules. Journal of Molecular Biology 428(14), 2931-2942.

Lee, T. H., and Young, R. A. (2000) Transcription of eukaryotic protein-coding genes. Annual Review of Genetics 34(1), 77-137.

Lehmann, M., Siegmund, T., Lintermann, K. G., Korge, G. (1998) The Pipsqueak protein of *Drosophila melanogaster* binds to GAGA sequences through a novel DNA-binding domain. The Journal of Biological Chemistry 273(43), 28504-28509.

Leatham-Jensen, M., Uyehara, C. M., Strahl, B. D., Matera, G., Duronio, R. J., McKay, D. J. (2019) Lysine 27 of replication-independent histone H3.3 is required for Polycomb target gene silencing but not for gene activation. PLOS Genet 15(1):e1007932.

Lessard, J., and Sauvageau, G. (2003) Bmi-1 determines the proliferative capacity of normal and leukaemic stem cells. Nature 423(6937), 255-260.

Lewis, E. B. (1968) Genetic control of developmental pathways in *Drosophila melanogaster*. in Proceedings of the Twelfth International Congress of Genetics 2, 96-97. edited by C. Oshima. Science Council of Japan, Tokyo.

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

Li, B., Carey, M., Workman, J. L. (2007) The role of chromatin during transcription. Cell 128,707-719.

Li, G., Margueron, R., Ku, M., Chambon, P., Bernstein, B. E., and Reinberg, D. (2010) Jarid2 and PRC2, partners in regulating gene expression. Genes and Development 24(4), 368-380.

Lis, J. T., Mason, P., Peng, J., Price, D. H., Werner, J. (2000) P-TEFb kinase recruitment and function at heat shock loci. Genes and Development 14(7), 792-803.

Lo, S. M., Ahuja, N. K., and Francis, N. J. (2009) Polycomb group protein suppressor 2 of zeste is a functional homolog of Posterior sex combs. Molecular and Cellular Biology 29(2), 515-525.

Lo, S. M., Follmer, N. E., Lengsfeld, B. M., Madamba, E. V., Seong, S., Grau, D. J., Francis, N. J. (2012) A bridging model for persistence of a Polycomb group protein complex through DNA replication in vitro. Molecular Cell 46(6), 784-796.

Loubiere, V., Delest, A., Thomas, A., Bonev, B., Schuettengruber, B., Sati, S., Martinez, A. M., and Cavalli, G. (2016) Coordinate redeployment of PRC1 proteins suppresses tumor formation during *Drosophila* development. Nature Genetics 48(11), 1436-1442.

Lynch, M. D., Smith, A. J. H., De Gobbi, M., Flenley, M., Hughes, J. R., Vernimmen, D., Ayyub, H., Sharpe, J. A., Solane-Stanley, J. A., Sutherland, L., Meek, S., Burdon, T., Gibbons, R. J., Messmer, S., Franke, A., and Paro, R. (1992) Analysis of the functional role of the Polycomb chromo domain in *Drosophila melanogaster*. Genes and Development 6(7),1241-1254.

Margueron, R., Justin, N., Ohno, K., Sharpe, M. L., Son, J., Drury Lll, W. J., Voigt, P., Martin, S. R., Taylor, W. R., Marco, V. D., Pirrotta, V., Reinberg, D., Gamblin, S. J. (2009) Role of the polycomb protein EED in the propagation of repressive histone marks. Nature 461(7265), 762-767.

Mazumdar, A., and Mazumdar, M. (2002) How one becomes many: blastoderm cellularization in *Drosophila melanogaster*. BioEssays 24(11),1012-1022.

Mendenhall, E. M., Koche, R. P., Truong, T., Zhou, V. W., Issac, B., Chi, A. S., Ku, M., Bernstein, B. E. (2010) GC-rich sequence elements recruit PRC2 in mammalian ES cells. PLOS Genetics 6(12), e1001244.

Mihaly, J., Hogga, I., Gausz, J., Gyurkovics, H., Karch, F. (1997) In situ dissection of the Fab-7 region of the bithorax complex into a chromatin domain boundary and a Polycomb-response element. Development 124(9), 1809-1820.

Mimori, K., Ogawa, K., Okamoto, M., Sudo, T., Inoue, H., Mori, M. (2005) Clinical significance of enhancer of zeste homolog 2 expression in colorectal cancer cases. European Journal of Surgical Oncology 31(4), 376-380.

Min, J., Zhang, Y., and Xu, R. M. (2003) Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. Genes and Development 17(15), 1823-1828.

Mlodzik, M., Fjose, A., and Gehring, W. J. (1985) Isolation of *caudal*, a *Drosophila* homeo box-containing gene with maternal expression, whose transcripts form a concentration gradient at the pre-blastoderm stage. The Embo journal 4(11), 2961-2969.

Mlodzik, M., Gehring, W. J. (1987) Expression of the caudal gene in the germ line of *Drosophila*: formation of an RNA and protein gradient during early embryogenesis. Cell 48(3), 465-478.

Mohd-Sarip, A., Cleard, F., Mishra, R. K., Karch, F., and Verrijzer, C. P. (2005) Synergistic recognition of an epigenetic DNA element by Pleiohomeotic and a Polycomb core complex. Genes and Development 19(15), 1755-1760.

Mohd-Sarip, A., Venturini, F., Chalkley, G. E., and Verrijzer, C. P. (2002) Pleiohomeotic can link Polycomb to DNA and mediate transcriptional repression. Molecular and Cellular Biology 22(21), 7473-7483.

Mohler, J., Eldon, E. D., and Pirrotta, V. (1989) A novel spatial transcription pattern associated with the segmentation gene, *giant*, of *Drosophila*. The EMBO Journal 8(5), 1539-1548.

Moshe, A., and Kaplan, T. (2017) Genome-wide search for Zelda-like chromatin signatures identifies GAF as a pioneer factor in early fly development. Epigenetics Chromatin 10(1), 33.

Moussa, H. F., Bsteh, D., Yelagandula, R., Pribitzer, C., Stecher, K., Bartalska, K., Michetti, L., Wang, J., Zepeda-Martinez, J. A., Elling, U., Stuckey, J. I., James, L. I., Frye, S. V., and Bell, o. (2019) Canonical PRC1 controls sequence-independent propagation of Polycomb-mediated gene silencing. Nature Communications 10(1), 1931.

Müller, J., and Bienz, M. (1991) Long range repression conferring boundaries of Ultrabithorax expression in the Drosophila embryo. The EMBO Journal 10(11), 3147-3155.

Müller, J., Hart, C. M., Francis, N. J., Vargas, M. L., Sengupta, A., Wild, B., Miller, E. L., O'Connor, M. B., Kingston, R. E., Simon, J. A. (2002) Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. Cell 111(2), 197-208.

Muller, M., Hagstrom, K., Gyurkovics, H., Pirrotta, V., Schedl, P. (1999) The Mcp element from the *Drosophila melanogaster* bithorax complex mediates long-distance regulatory interactions. Genetics 153(3), 1333-1356.

Nagy, P. L., Griesenbeck, J., Kornberg, R. D., Cleary, M. L. (2002) A trithorax-group complex purified from Saccharomyces cerevisiae is required for methylation of histone H3. Proceeding of the National Academy of Sciences. USA 99(1), 90-94.

Negre, N., Hennetin, J., Sun, L. V., Lavrov, S., Bellis, M., White, K. P., Cavalli, G. (2006) Chromosomal distribution of PcG proteins during Drosophila development. PLOS Biology 4(6), e170.

Nekrasov, M., Klymenko, K., Fraterman, S., Papp, B., Oktaba, K., Köcher, T., Cohen, A., Stunnenberg, H. G., Wilm, M., and Müllera, J. (2007) Pcl-PRC2 is needed to generate high levels of H3-K27 trimethylation at Polycomb target genes. The EMBO Journal 26(18), 4078-4088.

Ng, H. H., Robert, R. A., Struhl, K. (2003) Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. Mollecular Cell 11(3), 709-719.

Nibu, Y., Zhang, H., Levine, M. (1998) Interaction of short-range repressors with *Drosophila* CtBP in the embryo. Science 280(5360),101-104.

Nowak, K., Kerl, K., Fehr, D., Kramps, C., Gessner, C., Killmer, K., Samans, B., Berwanger, B., Christiansen, H., Lutz, W. (2006) BMI1 is a target gene of E2F-1 and is strongly expressed in primary neuroblastomas. Nucleic Acids Research 34(6), 1745-1754.

O'Connell, S., Wang, L., Robert, S., Jones, C. A., Saint, R., Jones, R. S. (2001) Polycomblike PHD fingers mediate conserved interaction with Enhancer of zeste protein. Journal of Biological Chemistry 276(46), 43065-43073.

Ogiyama, Y., Schuettengruber, B., Papadopoulos, G. L., Chang, J. M., Cavalli, G. (2018) Polycomb-dependent chromatin looping contributes to gene silencing during *Drosophila* development. Molecular Cell 71(1), 73-88.

Ohkuma, Y., Roeder, R. G. (1994) Regulation of TFIIH ATPase and kinase activities by TFIIE during active initiation complex formation. Nature 368(6467), 160-163.

Oksuz, O., Narendra, V., Lee, C. H., Descostes, N., LeRoy, G., Raviram, R., Blumenberg, L., Karch, K., Rocha, P. P., Garcia, B. A., Skok, J. A., Reinberg, D. (2018) Capturing the onset of PRC2-mediated repressive domain formation. Molecular Cell 70(6), 1149-1162. e5.

Oktaba, K., Gutierrez, L., Gagneur, J., Giradot, C., Sengupta, A. K., Furlong, E. E. M., and Môller, J. (2008) Dynamic regulation by Polycomb group protein complexes control pattern formation and the cell cycle in *Drosophila*. Developmental Cell 15(6), 877-889.

O'Meara, M. M., Simon, J. A. (2012) Inner workings and regulatory inputs that control Polycomb repressive complex 2. Chromosoma 121(3), 221-234.

Papp, B., Muller, J. (2006) Histone trimethylation and the maintenance of transcriptional ON and OFF states by TrxG and PcG proteins. Genes and Development 20(15), 2041-2054.

Pasini, D., Bracken, A. P., Hansen, J. B., Capillo, M., Helin, K. (2007) The Polycomb group protein Suz12 is required for embryonic stem cell differentiation. Mollecular and Cellular Biology 27(10), 3769-3779.

Pasini, D., Cloos, P. A. C., Walfridsson, J., Olsson, L., Bukowski, J. P., Johansen, J. V., Bak, M., Tommerup, N., Rappsilber, J., and Helin, K. (2010) JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells. Nature 464(7286), 306-310.

Pelegri, F., and Lehmann, R. (1994) A role of Polycomb group genes in the regulation of gap gene expression in *Drosophila*. Genetics 136(4), 1341-1353.

Peng, J. C., Valouev, A., Swigut, T., Zhang, J., Zhao, Y., Sidow, A., Wysocka, J. (2009) Jarid2/Jumonji coordinates control of PRC2 enzymatic activity and target gene occupancy in pluripotent cells. Cell 139(7), 1290-1302.

Pengelly, A. R., Copur, Ö., Jäckle, H., Herzig, A., and Müller, J. (2013) A histone mutant reproduces the phenotype caused by loss of histone-modifying factor Polycomb. Science 339(6120), 698-699.

Pengelly, A. R., Kalb, R., Finkl, K., and Müller, J. (2015) Transcriptional repression by PRC1 in the absence of H2A monubiquitylation. Genes and Development 29(14), 1487-1492.

Perino, M., Mierlo, G. V., Karemaker, I. D., Genesen, S. V., Vermeulen, M., Marks, H., Van Heeringen, S. J., and Veenstra, G. J. C. (2018) MTF2 recruits Polycomb Repressive Complex 2 by helical-shape-selective DNA binding. Nature Genetics 50(7), 1002-1010.

Peterson, A. J., Mallin, D. R., Francis, N. J., Ketel, C. S., Stamm, J., Weller, R. K., Kingston, R. E., Simon, J. A. (2004) Requirement for Sex comb on midleg protein interactions in *Drosophila* Polycomb group repression. Genetics 167(3), 1225-1239.

Petruk, S., Sedkov, Y., Johnston, D. M., Hodgson, J. W., Black, K. L., Kovermann, S. K., Beck, S., Canaani, E., Brock, H. W., Mazo, A. (2012) TrxG and PcG proteins but not methylated histones remain associated with DNA through replication. Cell 150(5), 922-933.

Petschek, J. P., Perrimon, N., Mahowald, A. P. (1987) Region-specific defects in l(1) *giant* embryos of *Drosophila melanogaster*. Developmental Biology 119 (1), 175-189.

Pherson, M., Misulovin, Z., Gause, M., Mihindukulasuriya, K., Swain, A., Dorsett, D. (2017) Polycomb repressive complex 1 modifies transcription of active genes. Science Advances 3(8), e1700944.

Philip, P., Boija, A., Vaid, R., Churcher, A. M., Meyers, D. J., Cole, P. A., Mannervik, M., and Stenberg, P. (2015) CBP binding outside of promoters and enhancers in *Drosophila melanogaster*. Epigenetics and Chromatin 8(1), 48.

Phrussi, A., Dickinson, L. K., and Lehmann, R. (1991) Oskar organizes the germ plasm and directs localization of the posterior determinant *nanos*. Cell 66(1), 37-50.

Platero, J. S., Hartnett, T., and Eissenberg, J. C. (1995) Functional analysis of the chrome domain of HP1. The EMBO Journal 14(16), 3977-3986.

Rajasekhar, V. K., and Begemann, M. (2007) Concise review: roles of Polycomb group proteins in development and disease: a stem cell perspective. Stem Cells 25(10), 2498-2510.

Rank, G., Prestel, M., Paro, R. (2002) Transcription through intergenic chromosomal memory elements of the *Drosophila* bithorax complex correlates with an epigenetic switch. Molecular and Cellular Biology 22(22), 8026-8034.

Ray, S. K., Li, H. J., Metzger, E., Schüle, R., Leiter, A. B. (2014) CtBP and associated LSD1 are required for transcriptional activation by neurod1 in gastrointestinal endocrine cells. Molecular and Cellular Biology 34(12), 2308-2317.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z. W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C. P., Allis, C. D., and Jenuwein, T. (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature 406(6796), 593-599.

Reyes-Reyes, M., Hampsey, M. (2007) Role for the Ssu72 C-terminal domain phosphatase in RNA polymerase II transcription elongation. Molecular and Cellular Biology 27(3), 926-936.

Reynolds, N., Salmon-Divon, M., Dvinge, H., Hynes-Allen, A., Balasooriya, G., Leaford, D., Behrens, A., Bertone, P., and Hendrich, B. (2012) NuRD-mediated deacetylation of H3K27 facilitates recruitment of Polycomb Repressive Complex 2 to direct gene repression. The EMBO Journal 31(3), 593-605.

Ringrose, L., Paro, R. (2004) Epigenetic regulation of cellular memory by the Polycomb and trithorax group proteins. Annual Review of Genetics 38, 413-443.

Ringrose, L., and Paro, R. (2007) Polycomb/Trithorax response elements and epigenetic memory of cell identity. Development 134(2), 223-232.

Ringrose, L., Rehmsmeier, M., Dura, J. M., Paro, R. (2003) Genome-wide prediction of Polycomb/Trithorax response elements in *Drosophila melanogaster*. Developmental Cell 5(5), 759-771.

Rivera-Pomar, R., Niessing, D., Schmidt-Ott, U., Gehring, W. J., and Jackle, H. (1996) RNA binding and translational suppression by bicoid. Nature 379(6567), 746-749.

Robinson, A. K., Leal, B. Z., Chadwell, L. V., Wang, R., Ilangovan, U., Kaur, Y., Junco, S. E., Schirf, V., Osmulski, P. A., Gaczynska, M., Hinck, A. P., Demeler, B., McEwen, D. G., and Kim, C. A. (2012) The growth-suppressive function of the Polycomb group protein Polyhomeotic is mediated by polymerization of its sterile Alpha motif (SAM) domain. Journal of Biological Chemistry 287(12), 8702-8713.

Samee, M. A. H., Sinha, S. (2014) Quantitative modeling of a gene's expression from its intergenic sequence. PLOS Computational Biology 10(3), e1003467.

Sanders, M. F., Bowman, J. L. (2019) Genetic analysis: an integrated approach (Third ed.). New York: Pearson Education, Inc. ISBN 0134605179.

Sanulli, S., Justin, N., Teissandier, A., Ancelin, K., Portoso, M., Caron, M., Michaud, A., Lombard, B., da Rocha, S. T., Offer, J., Loew, D., Servant, N., Wassef, M., Burlina, F., Gamblin, S. J., Heard, E., and Margueron, R. (2015) Jarid2 methylation via the PRC2 complex regulates H3K27me3 deposition during cell differentiation. Molecular Cell 57(5), 769-783.

Saramaki, O. R., Tammela, T. L., Martikainen, P. M., Vessella, R. L., and Visakorpi, T. (2006) The gene for Polycomb group protein enhancer of zeste homolog 2 (EZH2) is amplified in late-stage prostate cancer. Genes Chromosome and Cancer 45(7), 639-645.

Sarma, K., Margueron, R., Ivanov, A., Pirrotta, V., Reinberg, D. (2008) Ezh2 requires PHF1 to efficiently catalyze H3 lysine 27 trimethylation in vivo. Molecular and cellular Biology 28(8), 2718-2731.

Saurin, A. J., Shao, Z., Erdjument-Bromage, H., Tempst, P., Kingston, R. E. (2001) A *Drosophila* Polycomb group complex includes Zeste and dTAFII proteins. Nature 412(6847), 655-660.

Sauvageau, M., and Sauvageau, G. (2010) Polycomb group proteins: multi-faceted regulators of somatic stem cells and cancer. Cell Stem Cell 7(3) 299-313.

Savla, U., Benes, J., Zhang, J., and Jones, R. S. (2008) Recruitment of *Drosophila* Polycomb-group proteins by Polycomblike, a component of a novel protein complex in larvae. Development 135(5), 813-817.

Schaaf, C. A., Misulovin, Z., Gause, M., Koenig, A., Gohara, D. W., Watson, A., and Dorsett, D. (2013) Cohesin and polycomb proteins functionally interact to control transcription at silenced and active genes. PLOS Genetics 9(6), e1003560.

Schmitges, F. W., Prusty, A. B., Faty, M., Stu¨tzer, A., Lingaraju, G. M., Aiwazian, J., Sack, R., Hess, D., Li, L., Zhou, S., Bunker, R. D., Wirth, U., Bouwmeester, T., Bauer, A., Ly-Hartig, N.,
Zhao, K., Chan, H., Gu, J., Gut, H., Fischle, W., Mu¨ ller, J., and Thoma¨, N. H. (2011) Histone methylation by PRC2 is inhibited by active chromatin marks. Molecular Cell 42 (3), 330-341.

Schmitt, S., Prestel, M., and Paro, R. (2005) Intergenic transcription through a Polycomb group response element counteracts silencing. Genes and Development 19(6), 697-708.

Schroeder, M. D., Pearce, M., Fak, J., Fan, H., Unnerstall, U., Emberly, E., Rajewsky, N., Siggia, E. D., Gaul, U. (2004) Transcriptional control in the segmentation gene network of Drosophila. PLOS Biology 2(9), E271.

Schuettengruber, B., Bourbon, H. M., Di Croce, L., and Cavalli, G. (2017) Genome regulation by Polycomb and Trithorax: 70 Years and Counting. Cell 171(1), 34-57.

Schuettengruber, B., Cavalli, G. (2009) Recruitment of Polycomb group complexes and their role in the dynamic regulation of cell fate choice. Development 136(21), 3531-3542.

Schuettengruber, B., Cavalli, G. (2010) The DUBle life of Polycomb complexes. Developmental Cell 18(6), 878-880.

Schuettengruber, B., Ganapathi, M., Leblanc, B., Portoso, M., Jaschek, R., Tolhuis, B., Lohuizen, M. V., Tanay, A., Cavalli, C. (2009) Functional anatomy of Polycomb and Trithorax chromatin landscapes in Drosophila embryos. PLOS Biology 7(1), e13.

Schulz, C., Tautz, D. (1994) Autonomous concentration-dependent activation and repression of Kruppel by *hunchback* in the *Drosophila* embryo. Development 120(10), 3043-3049.

Schulz, C., and Tautz, D. (1995) Zygotic caudal regular and its role in abdominal segment formation of the *Drosophila* embryo. Development 121(4), 1023-1028.

Schulz, C., Tautz, D. (1995) Zygotic caudal regulation by hunchback and its role in abdominal segment formation of the *Drosophila* embryo. Development 121(4), 1023-1028.

Schulz, K. N., Bondra, E. R., Moshe, A., Villalta, J. E., Lieb, J. D., Kaplan, T., Mckay, D. J., and Harrison, M. M. (2015) Zelda is differentially required for chromatin accessibility, transcription factor binding, and gene expression in the early *Drosophila* embryo. Genome Research 25(11), 1715-1726.

Schwartz, Y. B., Kahn, T. G., Nix, D. A., Li, X. Y., Bourgon, R., Biggin, M., and Pirrotta, V. (2006) Genome-wide analysis of Polycomb targets in *Drosophila melanogaster*. Nature Genetics 38(6), 700-705.

Schwartz, Y. B., Kahn, T. G., Stenberg, P., Ohno, K., Bourgon, R., Pirrotta, V. (2010) Alternative epigenetic chromatin states of Polycomb target genes. PLOS Genetics 6(1), e1000805.

Shafaroudi, A. M., Mowla, S. J., Ziaee, S. A., Bahrami, A. R., Atlasi, Y., Malakootian, M. (2008) Overexpression of BMI1 a Polycomb group repressor protein in bladder tumors: a preliminary report. Urology Journal 5(2), 99-105.

Shi, Y., Sawada, J. I., Sui, G., Affar, E. B., Whetstine, J. R., Lan, F., Ogawa, H., Luke, M. P. S., Nakatani, Y., and Shi, y. (2003) Coordinated histone modifications mediated by a CtBP co-repressor complex. Nature 422(6933), 735-738.

Shilatifard, A. (2012) The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. Annual Review of Biochemistry 81, 65-95.

Shir-Shapira, H., Sharabany, J., Filderman, M., Ideses, D., Ovadia-Shochat, A., Mannervik, M., and Juven-Gershon, T. (2015) Structure-function analysis of the *Drosophila melanogaster* Caudal transcription factor provides insights into core promoter-preferential activation. The Journal of Biological Chemistry 290(28), 17293-17305.

Simon, J. A., and Kingston, R. E. (2013) Occupying chromatin: Polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put. Molecular Cell 49(5) 808-824.

Simon, J., Chiang, A., Bender, W., Shimell, M. J., and O'Conner, M. (1993) Elements of the *Drosophila* Bithorax complex that mediate repression by Polycomb group products. Developmental Biology 158(1), 131-144.

Sims III, R. J., Belotserkovskaya, R., and Reinberg, D. (2004) Elongation by RNA polymerase II: The short and long of it. Genes and Development 18(20) 2437-2468.

Sipos, L., Kozma, G., Molnar, E., Bender, W. (2007) In situ dissection of a Polycomb response element in *Drosophila melanogaster*. Proceeding of the National Academy of Sciences. U.S.A. 104(30), 12416-12421.

Small, S., Blair, A., and Levine, M. (1992) Regulation of *even-skipped* stripe 2 in the *Drosophila* embryo. The EMBO journal 11(11), 4047-4057.

Sparmann, A., and van Lohuizen, M. (2006) Polycomb silencers control cell fate, development and cancer. Nature Reviews Cancer 6(11), 846-856.

Steffen, P. A., Fonseca, J. P., Ganger, C., Dworschak, E., Kockmann, T., Beisel, C., and Ringrose, L. (2013) Quantitative in vivo analysis of chromatin binding of Polycomb and Trithorax group proteins reveals retention of ASH1 on mitotic chromatin. Nucleic Acids Research 41(10), 5235-5250.

Steffen, P. A., Ringrose, L. (2014) What are memories made of? How Polycomb and Trithorax proteins mediate epigenetic memory. Nature Reviews Molecular Cell Biology 15(5), 340-356.

Stros, M. (2010) HMGB proteins: Interactions with DNA and chromatin. Biochimica et Biophysica Acta (BBA)- Gene Regulatory mechanisms 1799 (1-2), 101-113.

Struhl, G., Johnston, P., Lawrence, P. A. (1992) Control of *Drosophila* body pattern by the *hunchback* morphogen gradient. Cell 69(2), 237-249.

Sun, Y., Nien, C. Y., Chen, K., Liu, H. Y., Johnston, J., Zeitlinger, J., and Rushlow, C. (2015) Zelda overcomes the high intrinsic nucleosome barrier at enhancers during *Drosophila* zygotic genome activation. Genome Research 25(11), 1703-1714.

Suva, M. L., Riggi, N., Janiszewska, M., Radovanovic, I., Provero, P., Stehle, J. C., Baumer, K., Le Bitoux, M. A., Marino, D., Cironi, L., Marquez, V. E., Clément, V., and Stamenkovic, I. (2009) stem cell maintenance. Cancer Research 69(24), 9211-9218.

Takahashi, K., Yamanaka, S. (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126(4), 663-676.

Tautz, D. (1988) Regulation of the *Drosophila* segmentation gene *hunchback* by two maternal morphogenetic centres. Nature 332(6161), 281-284.

Thomas, M. C., Chiang, C. M. (2006) The general transcription machinery and general cofactors. Critical Reviews in Biochemistry and Molecular 41(3), 105-178.

Tie, F., Banerjee, R., Stratton, C. A., Prasad-Sinha, J., Stepanik, V., Zlobin, A., Diaz, M. O., Scacheri, P. C., and Harte, P. J. (2009) CBP-mediated acetylation of histone H3 lysine 27 antagonizes *Drosophila* Polycomb silencing. Development 136(18),1331-3141.

Tram, U., Riggs, B., Sullivan, W. (2002) Cleavage and gastrulation in *Drosophila* embryos. The cytoskeleton guides early embryogenesis in *Drosophila*. In: Encyclopedia of Life Sciences. New York: Macmillan Publishers Ltd, Nature Publishing Group. pp. 1-7.

Tuckfield, A., Clouston, D. R., Wilanowski, T. M., Zhao, L. L., Cunningham, J. M., Jane, S. M. (2002) Binding of the RING Polycomb proteins to specific target genes in a complex with the *grainyhead*-like family of developmental transcription factors. Molecular and Cellular Biology 22(6),1936-1946.

van Heeringen, S. J., Akkers, R. C., van Kruijsbergen, I., Arif, M. A., Hanssen, L. L., Sharifi, N., andVeenstra, G. J. (2014) Principles of nucleation of H3K27 methylation during embryonic development. Genome research 24(3), 401–410.

van Lohuizen, M., Frasch, M., Wientjens, E., and Berns, A. (1991) Sequence similarity between the mammalian *bmi-1* proto-oncogene and the *Drosophila* regulatory genes Psc and Su(z)2. Nature 353(6342), 353-355.

Vincent, B. J., Staller, M. V., Lopez-Rivera, F., Bragdon, M. D. J., Pym, E. C. G., Biette, K. M., Wunderlich, Z., Harden, T. T., Estrada, J., and DePace, A. H. (2018) Hunchback is counter repressed to regulate *even-skipped* stripe 2 expression in *Drosophila* embryos. PLOS Genetics 14(9), e1007644.

Wade, P. A., Jones, P. L., Vermaak, D., Wolffe, A. P. (1998) A multiple subunit Mi-2 histone deacetylase from *Xenopus laevis* cofractionates with an associated Snf2 superfamily ATPase. Current Biology 8(14), 843-846.

Walker, E. Chang, W. Y., Hunkapiller, J., Cagney., G., Garcha, K., Torchia, J., Krogan, N. J., Reiter, J. F., Stanford, W. L. (2010) Polycomb-like 2 associates with PRC2 and regulates transcriptional networks during mouse embryonic stem cell self-renewal and differentiation. Cell Stem Cell 6(2), 153-166.

Walker, E., Manias, J. L., Chang, W. Y., Stanford, W. L. (2011) PCL2 modulates gene regulatory networks controlling self-renewal and commitment in embryonic stem cells. Cell Cycle 10(1), 45-51.

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

Wang, L., Brown, J. L., Cao, R., Zhang, Y., Kassis, J. A., and Jones, R. S. (2004) Hierarchical recruitment of Polycomb group silencing complexes. Molecular Cell 14(5), 637-646.

Wang, L., Jahren, N., Miller, E. L., Ketel, C. S., Mallin, D. R., Simon, J. A. (2010) Comparative analysis of chromatin binding by Sex Comb on Midleg (SCM) and other Polycomb group repressors at a *Drosophila* Hox gene. Molecular and Cellular Biology 30(11), 2584-2593.

Wang, X., Goodrich, K. J., Gooding, A. R., Naeem, H., Archer, S., Paucek, R. D., Youmans, D. T., Cech, T. R., Davidovich, C. (2017) Targeting of polycomb repressive complex 2 to RNA by short repeats of consecutive guanines. Molecular Cell 65(6), 1056-1067. e5.

Wang, X., Paucek, R. D., Gooding, A. R., Brown, Z. Z., Eva, J. G., Muir, T. W., and Cech, T. R. (2017) Molecular analysis of PRC2 recruitment to DNA in chromatin and its inhibition by RNA. Nature Structural and Molecular Biology 24(12), 1028-1038.

Wang, Z., Zang, C., Rosenfeld, J. A., Schones, D. E., Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Peng, W., Zhang, M. Q., Zhao, K. (2008) Combinatorial patterns of histone acetylations and methylations in the human genome. Nature Genetics 40(7), 897-903.

Weigel, D., Seifert, E., Reuter, D., and Jackle, H. (1990) Regulatory elements controlling expression of the *Drosophila* homeotic gene fork head. The EMBO Journal 9(4), 1199-1207.

Wharton, R. P., and Struhl, G. (1991) RNA regulatory elements mediate control of *Drosophila* body pattern by the posterior morphogen *nanos*. Cell 67(5), 955-967.

Woo, C. J., Kharchenko, P. V., Daheron, L., Park, P. J., Kingston, R. E. (2010) A region of the human HOXD cluster that confers Polycomb-group responsiveness. Cell 140(1), 99-110.

Yu, M., Mazor, T., Huang, H., Huang, H. T., Kathrein, K. L., Woo, A. J., Chouinard, C. R., Labadorf, A., Akie. T. E., Moran, T. B., Xie, H., Zacharek, S., Taniuchi, I., Roeder, R. G., Kim, C. F., Zon, L. I., Fraenkel, E., Cantor, A. B. (2012) Direct recruitment of Polycomb repressive complex 1 to chromatin by core binding transcription factors. Molecular Cell 45(3), 330-343.

Yuan, W., Wu, T., Fu, H., Dai, C., Wu, H., Liu, N., Li, X., Xu, M., Zhang, Z., Niu, T., Han, Z., Chai, J., Zhou, X. J., Gao, S., Zhu, B. (2012) Dense chromatin activates Polycomb repressive complex 2 to regulate H3 lysine 27 methylation. Science 337(6097), 971-975.

Yuan, W., Xu, M., Huang, C., Liu, N., Chen, S., and Zhu, B. (2011) H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. Journal of Biological Chemistry 286(10), 7983-7989.

Zhang, W., Aubert, A., de Segurade, J. M. G., Karuppasamy, M., Basu, S., Murthy, A. S., Diamante, A., Drury, T. A., Balmer, J., Cramard, J., Watson, A. A., Lando, D., Lee, S. F., Palayret, M., Kloet, S. L., Smits, A. H., Deery, M. J., Vermeulen, M., Hendrich, B., Klenerman, D., Schaffitzel, C., Berger, I., Laue, E. D. (2016) The Nucleosome remodeling and deacetylase complex NuRD is built from preformed catalytically active sub-modules. Journal of Molecular Biology 428(14), 2931-2942.

Zhang, Y., LeRoy, G., Seelig, H. P., Lane, W. S., Reinberg, D. (1998) The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities. Cell 95(2), 279-289.

Zink, B., Engstrom, Y., Gehring, W. J., and Paro, R. (1991) Direct interaction of the Polycomb protein with Antennapedia regulatory sequences in polytene chromosomes of *Drosophila melanogaster*. The EMBO Journal 10(1), 153-162.

Zink, B., and Paro, R. (1989) In vivo binding pattern of a transregulator of homoeotic genes in *Drosophila melanogaster*. Nature 337(6206), 468-471.

Zolakar, M. (1976) Division and migration of nuclei during early embryogenesis of *Drosophila melanogaster*. J. Microsc. Biol. Cell 25, 97-106.