INSULATOR ELEMENTS MEDIATE LONG-RANGE INTERACTIONS BETWEEN POLYCOMB TARGETS AND BETWEEN ACTIVE ENHANCERS IN DROSOPHILA

by

HUABING LI

A Dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

and

The Graduate School of Biomedical Sciences

University of Medicine and Dentistry of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Biochemistry

Written under the direction of

Professor Vincenzo Pirrotta

And approved by

_____

_____

_____

_____

New Brunswick, New Jersey

MAY, 2011

ABSTRACT OF THE DISSERTATION

Insulator elements mediate long-range interactions between Polycomb targets and between active enhancers in Drosophila

By HUABING LI

Dissertation Director:

Vincenzo Pirrotta, Ph.D.

The genomic binding sites of Polycomb (PcG) complexes have been found to cluster, forming Polycomb "bodies" or foci in mammalian or fly nuclei. These associations are thought to be driven by interaction between PcG complexes and result in enhanced repression. Similar mechanisms were proposed to explain the clustered genes which share the same transcriptional loci, called "transcription factories". However, the elements that mediate the long-range interactions and form those bodies remain elusive.

The insulator element is a new class of DNA elements and indicated to function in the nuclear structure. *Fab-7* and *Mcp*, the boundaries of cis-regulatory domains of the Bithorax Complex in *Drosophila* Genome, each contain two functional parts: the silencer (Polycomb Response Element, PRE) and the insulator. Using in-vivo fluorescence imaging and 3C methods, we show that the interactions between remote

copies of *Mcp* or *Fab-7* elements are dependent on the insulator activities present in these elements and not on their PREs.

However, the interaction ratio is around 6~7%, much lower than the published 20% or 90% interactions. By constructing new transgenic flies, that with enhancer and *Mcp* flanked by removable lox and frt elements, we have observed high interaction ratio (60~90%) in the Drosophila third instar larva eye discs, low interaction ratio (10~20%) with enhancer deleted, and no interaction (<1%) without *Mcp*, demonstrating that enhancer is needed for the high stable trans-interactions. We also show that this high interaction is dependent on the insulator protein dCTCF, *Mcp* PRE part and the Trithorax protein. In addition, a different enhancer could also promote the co-localization.

In the related works we conclude that the insulators mediate long-range interactions of PcG targets to form the PcG bodies in the nucleus, and that the insulators also enhance the transcription of active genes by bringing enhancers into the transcription factories. We clarified the mechanism of PcG bodies, and for the first time prove that insulators also mediate long-range interactions between enhancers and bring active transcription units into the transcription factories. In summary, insulators help both silencers and enhancers to build the higher-order organization in the nucleus to regulate gene expression.

Acknowledgement

First of all, I would like to thank my advisor Dr. Vincenzo Pirrotta, for his exceptional guidance and kind support. I am inspired by his tireless devotement to sciences. This is the best lab where I could get my best scientific training that will benefit my whole future life.

My sincere thanks go to professors serving on my dissertation committee: Marc Gartenberg, J. Don Chen, and Thomas Kusch.

I also want to thank the members of the lab. Donna Mccabe is always there to help, and provide very useful suggestions on both my research and life. Former lab members Drs. Gregory Shanower and Kastuhito Ohno really help me a lot on my projects. Thanks also go to other former and current members: Sung-Yeon Park, Daniela Linder-Basso, Yuri Schwartz, Tatiana Kahn, Hungoo Lee, Parvatha, Simon, Carrado, Dalal.

I felt indebted to my parents, my parents-in-law and my wife. Without their absolute support and love, it is impossible for me to finish my thesis research. My parents brought up my daughter in China and send my daughter to the States until she was three years old when I saw my daughter the first time. My wife takes care of the small family and works to provide the financial support to the family, so I am free of household works and financial pressure, and can focus on my research. I also want to thank my daughter, who brings so much fun into my life.

Table of contents

List of figures

List of tables

Chapter 1. Introduction

1.1 Nuclear Organization and Chromatin Structure

Eukaryotic genomes are organized on linear chromosomes that are compacted together inside the nucleus. The chromatin higher-order structure has still been poorly characterized. FISH / chromosome painting techniques showed that chromosomes occupy discrete territories in the cell nucleus, and determine the three dimensional position of individual active and inactive genes (Cremer et al. 2001). And each chromosome also has preferred positions with respect to the center or periphery of the nucleus and with respect to each other. Their 3D nuclear organization was termed chromosome territory, which is large discrete nuclear body separated from the other chromosome territories by an interchromatin compartment — a more or less continuous space between adjacent chromosomes (Heard et al. 2007). Active euchromatin and inert heterochromatin on the chromosomes also have their own preferential spatial position within the eukaryotic nucleus, and this position reflects distinct gene expression environments. Transcriptionally competent genes are more often found in the interior part of the nuclei, while the nuclei periphery frequently home the inactive genes on heterochromatin (Lunyak 2008).

Many essential biological activities in the nucleus, such as RNA transcription, splicing, DNA repair, replication, RNA and DNA modification and regulation, et al., are aggregated in subnuclear foci termed nuclear bodies (Fig. 1.1), which include

transcription factories, PML bodies, nuclear speckles, Cajal bodies, and PcG bodies, (Taddei et al. 2004; Spector 2001; Chakalova et al. 2005; Kumaran et al. 2008). The function-based foci may arise through the self-assembly of essential components and act to create a local high concentration of essential factors to enhance the functions and facilitate regulation (Chambeyron et al. 2004).

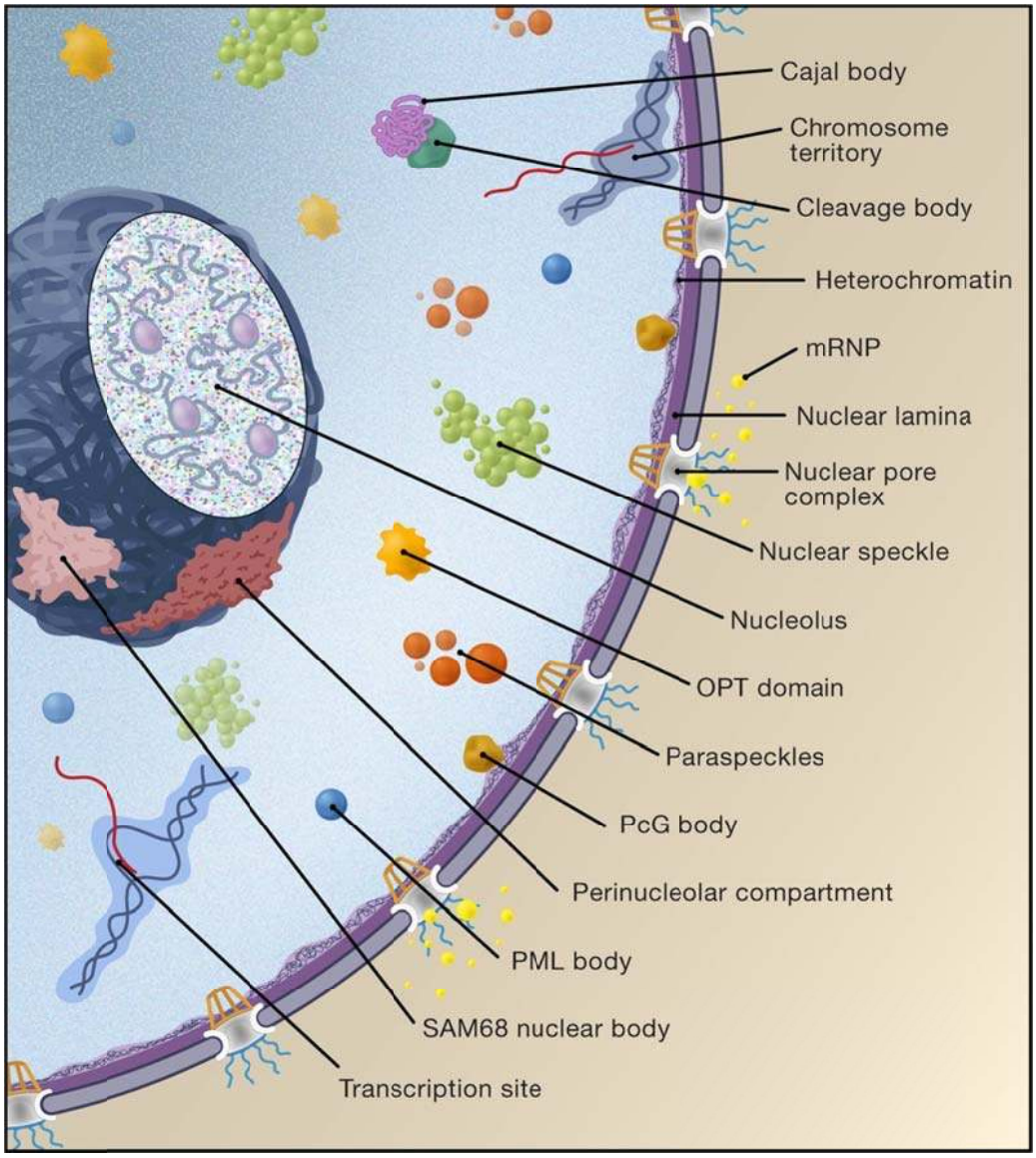


Figure 1.1 Nuclear structure and various kinds of nuclear bodies.

Directly adapted from Kumara et al. (2008).

1.2 Transcription Factories

Early in-vitro biochemistry experiments with whole cell extracts, transcription factors, and recombinant purified polymerases lead us to the traditional views in the text books, that describe the polymerases like a locomotive tracking on the DNA templates, and synthesizing RNAs along the way. Through Br-UTP incorporation and immuno-labeling of nascent RNA or active RNAP Ⅱ, the sites of RNA synthesis were observed under fluorescence microscopy and found that they are not evenly distributed throughout the nucleus, but concentrated in hundreds of discrete sites, which were termed transcription 'factories' (Jackson et al. 1993; Wensink et al. 1993; Iborra et al. 1996). The number of factories in a nucleus range from a few hundred to several thousand, varying with different cell types. Comparing the large number of active transcription units per nucleus, led to an estimate of several to 30 active units per factory. The size of transcription factories was estimated to be 45-100 nm in diameter (Iborra et al. 1996; Osborne et al., 2004; Sutherland & Bickmore, 2009). A typical factory in HeLa cells host a cluster of about 8 RNA polymerase Ⅱ (Faro-Trindade et al. 2006). The transcription factories are not just aggregates of RNAP Ⅱ, but serve as tethering sites for chromatin loops and continue to exist as independent nuclear sub-compartments after transcription termination (Mitchell et al. 2008). Through detergent and salt extraction (Kimura et al. 1999), and 3C analysis of long and short genes activated by TNFα (Papantonis et al. 2010), the active RNA polymerase Ⅱ was found to be immobilized on yet unknown nuclear substructures.

Using live-cell imaging of the GFP-tagged RNA polymerase Ⅱ and fluorescence recovery after photobleaching (FRAP), the elongating fraction of RNAP Ⅱ exchanges slowly with a half-time of ~13 or 20 minutes, which is consistent with the idea that the transcription factories are attached to a substructure (Kimura et al. 2002; Becker et al. 2002).

Current models of transcription factories propose that polymerases concentrate in discrete 'factories', immobilized by attachment to a sub-nuclear structure, and they work together on many different templates to reel in their templates and extrude the newly synthesized RNAs (Fig. 1.2) (Cook, 1999; Cook, 2010; Sutherland et al. 2009; Chakalova et al. 2010). Increasing molecular biology evidences suggest that, rather than the transcription machinery moving along the chromatin fiber, the chromatin template diffused or was dragged into those protein compartments, which have high local concentration of RNA polymerase Ⅱ and factors that could potentially increase the transcription efficiency. Often, higher order organization of the chromosome changes upon activation, and genes extend out of the surface of the chromosome territories (Volpi et al. 2000; Chambeyron et al. 2004; Wiblin et al. 2005; Osborne et al. 2007), which suggests that genes are brought into the transcription factories. *Myc* and *Igh* were found to rapidly relocate into the same pre-assembled transcription factory upon immediate-early gene induction (Osborne et al. 2007). The mouse α-globin locus was found to interact with its neighboring house-keeping genes in a transcription factory (Zhou et al. 2006). Certain distal genes tend to co-localize in the

same factories at high frequency, and the movement into or out of the preassembled transcription factories results in activation or abatement of transcription (Osborne et al. 2004).

The nucleolus is a huge RNA polymerase Ⅰ factory where the all 45S rRNAs are synthesized. And polymerase Ⅱ and Ⅲ are also found to be concentrated in their own dedicated region in Hela cells (Pombo et al. 1999), and possibly transcribe their own class of units. But, does polymerase Ⅱ specialize further, and does each RNAP Ⅱ factory transcribe a certain class of genes? Current evidences indicate that they do and not all factories contain the same kind of factors (Pombo et al. 1998; Carter et al. 2008). 4C experiments using the active β-globin locus and a house-keeping gene *Rad23a* as baits showed that active transcription units are preferentially in contact with their specific sets of other active genes (Simonis et al., 2006). By using 4C, Fraser and colleagues also show that co-regulated genes preferentially cluster at the same specialized factories that optimized for high transcription level (Schoenfelder et al. 2009).

Although the concept of replication factories has been established, there are still a number of controversies regarding the transcription factories (Cook 1999; Sutherland et al. 2009; Chakalova et al. 2010). Except for the RNA polymerase Ⅱ, no other proteins or factors were found to be common constituents of all RNAP Ⅱ factories (Sutherland et al. 2009). And, one of the main gaps in the current transcription model

is that we still do not know how those transcription foci are formed. There are some clues. Murrel et al. (2004) showed that imprinting control region (ICR) between Igf2 and H19 was differentially methylated in maternal and paternal chromatin, partitioning the chromatin into distinct loops, thus switching transcription by bringing the distant Igf2 promoter into contact with the H19 enhancers through CTCF. By a 3C-based Associated Chromosome Trap (ACT) assay, this ICR was found also to loop out of chromosome 7 to interact and promote transcription of the Wsb1/Nfl genes on Chromosome 11, and this long-distance interaction also depended on CTCF (Ling et al. 2006). Those findings suggest CTCF may simply form local loops to bring enhancers in contact with promoters, but also imply that insulator elements and insulator proteins may bring the transcription units into common pre-existing active transcription factories.

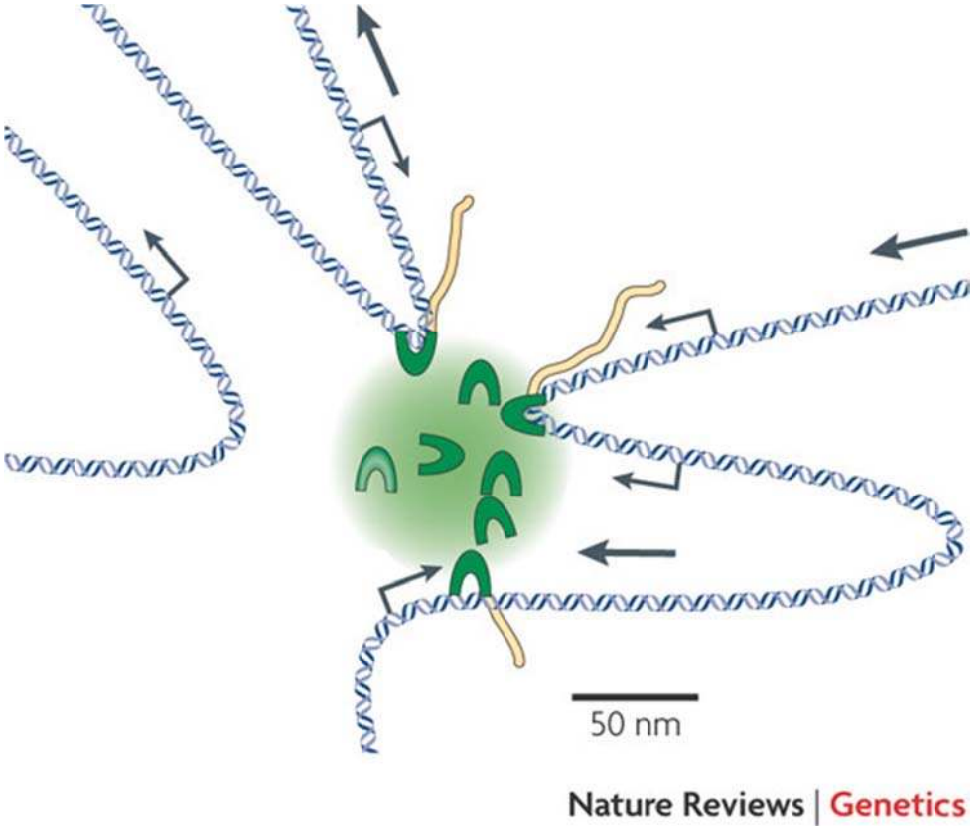


Figure 1.2 The Transcription factories model.

Adapted from Sutherland & Bickmore (2009). Shows that 8 RNA polymerase II (green crescents) constitute the factory, which reel through and transcribe three genes with nascent RNAs (yellow) extruded. Gene from same or different chromosomes could come together to share the same factory.

1.3 Enhancers

Enhancers, or activators, are one important class of DNA regulatory elements that were originally defined as cis-acting elements dramatically increasing gene transcriptional activities that is independent of their relative orientation and distance from its controlled gene start sites (Blackwood & Kadonaga, 1998; Khoury & Gruss, 1983). The first eukaryotic enhancer identified was the SV40 enhancer, which cooperates with the host cell transcriptional machinery upon infection, and is vital for the viral replication and infection (Banerji et al. 1981).

Often, remote enhancers are located tens to hundreds of kilo-bases away from their target genes. The *Drosophila* wing margin enhancer resides 85-kb upstream of *cut* promoter (Jack et al. 1991). The enhancer that activates T cell receptor (TCR) alpha-chain gene reside 69kb downstream of its promoter (Winoto and Baltimore, 1989). And very often, each gene is controlled by multiple autonomous enhancer modules that vary from 50-bp to 1.5-kb in length, and each modules may perform a specific functions and activate the gene at specific tissue or at particular developmental stages. So, each enhancer module contributes, in a cumulative manner, to the overall spatial and temporal regulation of the gene (Blackwood and Kadonaga, 1998).

Most enhancers contain multiple binding sites for several different classes of sequence-specific transcription factors, and use at least two different factors (sometime repressors) to regulate gene expression. Different factors function synergistically to activate gene expression. They may cooperatively occupy the linked sites via protein-protein interaction, or coordinately recruit co-activators such as CBP to the DNA templates (Levine 2010). Once the proper combination of factors and co-activators bind to the enhancers, the next step is how enhancers reach, sometimes far-away, specific promoters of the target genes. The looping model is proposed to explain the mechanism of enhancer action: the specific enhancer-binding proteins and promoter-binding factors interact with each other, thus loop and bring those two far-away DNA elements together. While at the same time, the insulators/boundary elements could be used to block undesired promoter-enhancer communications. It is also possible that insulator could also bring the adjacent enhancer to the cognate promoter, such as in the globin locus or in the Igf2-H19 locus, insulator proteins CTCF may fold the chromatin to bring enhancers to the vicinity of promoters (Murrel et al. 2004; Ling et al. 2006). It is important that the enhancer only activate its cognate promoter sometimes over large distance, and also avoid activating other promoters in its vicinity. To achieve this specificity, the cells may employ several strategies. First, direct protein-protein contact between enhancer-associated proteins and basal transcription machinery components could be formed by DNA looping, then numerous co-activators interact with the DNA-bound factors. Second, transcription factors and cofactors could enzymatically modify the properties and activities of

proteins to regulate the transcription process, such as Pol Ⅱ phosphorylation, histone acetylation. Third, the packaging of long DNA fiber could possibly promote long-range interaction. Fourth, the chromatin remodeling complexes can alter chromatin structure and likely facilitate the enhancer-bound proteins access of the promoter region. Fifth, the super-helical tension in the DNA template could also increase enhancer activity. Sixth, the nuclear structure may play an important role in the enhancer action, which is the core theory of transcription factories (Blackwood and Kadonaga, 1998).

Traditionally, various transgene assays were employed to identify new enhancers, such as the enhancer trap. With advent of ChIP-chip / ChIP-seq, combined with bioinformatics, numerous new enhancers have been identified in various organism and tissues. Zeitlinger et al. determined genome-wide occupancy of Dorsal, Twist, and Snail in the Drosophila embryos, and identified known dorso-ventral enhancers and several hundred additional potential enhancers (Zeitlinger et al. 2007). By using of extreme evolutionary sequence conservation, supported by tests in transgenic mouse enhancer assays, Pennacchio et al. identified enhancers in the human genome by bioinformatics (Pennacchio et al. 2006; Visel et al. 2008). The same group also employed ChIP-seq to map the in vivo binding sites of p300 in mouse embryonic forebrain, midbrain and limb tissue, and heart tissue. Several thousands of tissue specified enhancers were identified, and up to one hundred of those enhancers were shown to be functional in the transgenic mouse assay (Visel et al. 2009; Blow et al.

2010). Heinztman et al. found by ChIP-chip that the active promoters were marked by H3K4me3, while enhancers were marked by mono-methylation of H3K4 in human genome, and predicted several hundreds of promoters and enhancers within 30-MB of human genome (Heintzman et al. 2007).

In the following sections, three typical enhancers of great interest to my research projects will be introduced in detail.

The eye enhancer of the Drosophila *white* gene has been used widely in various transgene assays. It is located 1-kb upstream of the *white* promoter, and the one used in most of transgene assays is about 1kb long and contains separable eye disc specific and testis specific activities (Muravyova et al. 2001; Hagstrom et al. 1996). Both of the enhancers could confer tissue specific expression on *white* gene promoter or heterologous promoter of other genes. The eye-specific part has 5 *zeste* binding sites, while the adjacent testis part does not have any *zeste* sites. Those *zeste* binding sites are important for the *zeste-white* interactions and normally stimulate the enhancer activities (Qian et al. 1992).

The *Ultrabithorax* (*Ubx*) gene provides an excellent model for enhancer action. *Ubx* is one of the best characterized homeotic genes and specifies the identity of the posterior thoracic and anterior abdominal segments. The gene body is 75-kb long, and has a set of complicated regulatory elements that scattered over a region of over

100-kb both upstream and downstream of the promoter, to achieve high temporal and spatial specificity (Pirrotta et al. 1995). There are four principal genetically defined regulatory regions progressively from 50-kb downstream to 40-kb upstream of the transcription start sites: *abx*, *bx*, *bxd* and *pbx*. Several enhancer elements have been characterized by using *lacZ* reporter gene assay from those regions to be responsible for the para-segmental pattern of expression. Four of them, ABX, BRE, BXD and PBX from those genetically defined regions respectively, generate patterns of expression similarly to that of endogenous *Ubx* gene (Qian et al. 1993). The enhancer activities had been mapped down to short acting, autonomous sequences 500-1000 bp in length. Pair-rule gene products bind and activate those enhancers, while the gap gene products, such as *hb* and *tll*, repress those enhancers to limit the *Ubx* domains to its appropriate parasegments (Qian et al. 1991; Poux et al. 1996). The *Ubx* enhancers could also auto-regulate their activities by responding to the *Ubx* level, as well as to genetic background (Crickmore et al. 2009).

Another example of enhancer is the synthetic UAS-Gal4 system developed in Drosophila. Gal4 protein is 881 amino acid long, identified as a regulator in the yeast *Saccharomyces cerevisiae*, and specifically binds to an Upstream Activating Sequence (UAS) element. The Gal4 protein expressed in Drosophila has no deleterious effects, and could induce reporter gene expression under the control of UAS sequences, analogous to enhancer function (Fischer et al. 1988). Brand and Perrimon (1993) developed the powerful UAS-Gal4 system in Drosophila to

temporally and spatially express target genes in vivo. Over the years, the Drosophila community has developed numerous Gal4 driver lines, which allow the selective activation of any UAS-containing transgene in a wide variety of cell and tissue specific patterns.

1.4 Polycomb-group genes

Polycomb group (PcG) genes are highly conserved from flies to mammals, and were first identified as repressors of the *Hox* genes during Drosophila development. Since then, they have been widely recognized as global regulators of gene expression, and involved in a broad range of biological process, including differentiation, cell cycle control, apoptosis, senescence, genetic imprinting, X-inactivation, cancer development, and more recently the stem cell homeostasis and early lineage commitment (Surface et al. 2010; Pirrotta 1998). The PcG family has a diverse set of proteins that assembled into protein complexes, and those chromatin-associated complexes are variable in different tissue and development stages. To date, three distinct *Drosophila* PcG protein complexes have been biochemically purified and characterized: PRC1, PRC2 and, most recently, PhoRC (Fig. 1.3). All three complexes contain multiple subunits that are encoded by PcG genes that are critically required for *Hox* gene silencing. The core of *Drosophila* PRC2 contains E(z), Su(z)12, Esc, Nurf55 (Schwartz et al. 2007), and together this complex functions as a histone methyltransferase (HMTase) that specifically methylates lysine27 in histone H3, which is characteristic of PcG target genes (Czermin et al. 2002; Müller et al. 2002; Cao et al. 2002; Kuzmichev et al. 2002). The core of PRC1 contains Pc, Ph, Psc, Sce/dRing (Shao et al. 1999; Saurin et al. 2001). Many other proteins were co-purified with these core components, and may be used as ways to regulate PcG activities in different tissue and development stages. RING domain-containing subunit

dRing functions as an E3 ubiquitin ligase that mono-ubiquitylates lysine 119 of histone H2A (Wang et al. 2004; Napoles et al. 2004; Cao et al. 2005). The chromodomain of Pc binds specifically to the trimethylated lysine 27 of histone H3 (H3K27me3) (Fischle et al. 2003), deposited by E(z) of PRC2, and it inhibits nucleosome remodeling and transcription (Francis et al. 2004). The third complex, PhoRC, contains Pho/PhoL and dSFMBT (Brown et al. 1998; Brown et al. 2003; Klymenko et al. 2006). Pho is the Drosophila homologue of human Yin-Yang 1(YY1), and the only PcG proteins that binds directly to DNA, dSFMBT contains a MBT-domain and bind specifically to mono- and dimethylated H3K9 and H4K20. Specifically in the PRC2 HMTase, the SET domain protein E(z) is the catalytic subunit, whereas in PRC1, Psc provides most of the activity for inhibition of nucleosome remodeling (Levine et al., 2004; Schwartz & Pirrotta, 2007; 2008). But how those posttranslational histone marks contribute to gene regulation remains to be fully elucidated.

PcG proteins repress their target genes by assembling chromatin complexes at the Polycomb Response Element (PRE), a specific DNA region of several hundred base pairs that can direct silencing of one or more promoters placed in its vicinity. PREs are largely devoid of nucleosomes and PRE DNA serves as an assembly platform for multiple different PcG protein complexes through DNA-protein and protein-protein interactions (Schwartz and Pirrotta, 2007). Of all the three known PcG complexes, only Pho/PhoL interacts directly with PRE. To date, many PREs have been found to

bind several other known DNA-binding proteins, such as Gaf, Psq, Zeste and Dsp1, that might act as recruiters, but none of them are present in all the PREs. Several PREs, such as *Fab-7*, *Mcp*, *Ubx* PREs, have been identified by genetics and functional analysis in Drosophila, and more recently approximately 200 endogenous PREs in the *Drosophila* genome were deduced from ChIP-chip data (Tolhuis et al. 2006; Negre et al. 2006; Schwartz et al. 2006), but no PRE consensus sequences have been found. It is likely that PcG complexes are recruited to PREs by different combinations of DNA-binding proteins in different ways like many complex enhancers. In mammals, only two putative PREs have been recently discovered between HOXD11 and HOXD12 and in the MafB/Kreisler gene (Woo et al. 2010). Often, PREs work at a distance of several tens of kilobases to silence their target gene. Two models were proposed to explain how PcG works. By analogy to yeast SIR complexes, PRE acts as a docking site, PcG were supposed to spread along the chromatin and coat it in a condensed form. However, ChIP-chip experiments show that PcG proteins peak sharply only at known or putative PREs in Drosophila which does not support the model, while the distribution of H3K27me3 extends over the entire transcription units and regulatory regions. To reconcile those observations, a looping model, proposed by Schwartz and Pirrotta (2007), suggests that PcG proteins are recruited to PREs by DNA-binding proteins, and then methylate flanking nucleosomes on either side of the PRE, the methylation domain is extended by looping of the PRE-bound complexes to contact nucleosomes over a broad region, and this looping is mediated by interaction of Pc chromodomain with the H3K27me3,

thus creating and maintaining a broad methylation domain. The complete list of PcG components is as shown in Table 1.1.

Interestingly, Trithorax (TRX) binds constitutively to almost all known or putative PREs (also called TREs) that bind PcG proteins, regardless of the transcriptional activities of the target genes (Chinwalla et al. 1995; Papp et al. 2006). All PcG target genes are positively regulated by TRX, which deposits the active marks H3K4me3 and antagonizes PcG repression (Poux et al. 2002; Klymenko et al. 2004). In Drosophila, Trx function is closely related to Ash1 function, both of which are SET-domain proteins with H3K4 methyltransferase activity (Schwartz et al. 2007). The recruitment and mechanism of TRX and ASH1 are still mysterious.

It have been long noticed that, by immuno-stainning, PcG proteins form distinct foci in both Drosophila and mammalian nuclei, and termed ‘polycomb bodies’ (Saurin et al. 1998; Buchenau et al. 1998; Grimaud et al., 2006). Polytene chromosome staining shows around 120 binding sites for the Pc Protein, and there are approximately 200 (up to 2000 with lower threshold) endogenous PREs in the *Drosophila* genome implicated from ChIP-chip data, but a much smaller number (40-50 PC and PH bodies per nucleus in embryos) of PcG bodies in diploid nuclei can be detected by immunofluorescence (Ficz et al. 2005; Schwartz et al. 2006; Grimaud et al., 2006). This suggests that there may be a significant number of contacts among endogenous PcG target loci. By FISH or in vivo imaging, contacts between PREs have been

observed inside the PcG bodies, which enhance the silencing of the target genes (Bantignies et al., 2003; Grimaud et al., 2006; Vazquez et al., 2006). Although the functional significance is still unclear, the long-distance interaction of target genes seems to be important for the PcG silencing. Analogous to the transcription factories, it is attractive to suppose that PcG bodies are also pre-existing sub-compartments of nuclear structure, and provide silenceing domains, thus silencing multiple targets more efficiently. But fluorescence recovery after photobleaching (FRAP) experiments suggests that PcG bodies (half-life 2-6 minutes) (Ficz et al. 2005) are much more dynamic compared to Pol Ⅱ bodies (half-life 20min)(Kimura et al. 2002), while more stable compared to HP1 (half-time <1min) (Festenstein et al. 2003; Cheutin et al. 2003). Then an understanding of the mechanisms through which genes are targeted to the bodies and how the PcG bodies form becomes imperative to understand PcG silencing.

Table 1.1 Main components of the Polycomb / Trithorax system.

Directly adapted from Schwartz & Pirrotta (2007).

Drosophila protein	Complex	Protein domains	Biochemical activity	Mouse protein homologues
Polycomb group				
PC	PRC1	Chromodomain	Binding to trimethyl H3K27	NPCD, M33 (CBX2), CBX4, CBX6, CBX7, CBX8
PH	PRC1	SAM	?	PHC1, PHC2, PHC3
PSC	PRC1	RING	Cofactor for SCE	BMI1, MEL18
SCE (RING)	PRC1	RING	E3 ubiquitin ligase specific to H2AK119	RING1A, RING1B
SCM	PRC1?	SAM, MBT, Zn-finger	?	SCMH1, SCML2
E(Z)	PRC2	SET	Methylation of H3K9, H3K27	EZH2, EZH1
ESC	PRC2	WD40	Cofactor for E(Z)	EED
ESCL	PRC2	WD40	Cofactor for E(Z)	EED
SU(Z)12	PRC2	Zn-finger	?	SUZ12
PCL	PRC2	PHD, Tudor	?	PHF19, MTF2 (M96)
PHO	PhoRC	Zn-finger	DNA binding	YY1, YY2,
PHOL	?	Zn-finger	DNA binding	YY1, YY2
CG16975 (SFMBT)	PhoRC	MBT, SAM	Binding to mono- and dimethyl H3K9, H4K20	L3MBTL2, MBTD1
SU(Z)2	?	RING,	?	
SXC	?	?	?	?
ASX	?	PHD	?	ASXL1, ASXL2,
MXC	?	LA, RRM	?	Q9CUQ5
E(PC)	?	?	?	EPC1, EPC2
Trithorax group				
TRX	TAC1	PHD, SET,	Methylation of H3K4	WBP7, MLL1
ASH1	?	SET, PHD, BAH	Methylation of H3K4, H3K9, H4K20	ASH1L
ASH2	?	PHD, SPRY	?	ASH2L
BRM	SWI/SNF	SNF2, HELICc, Bromo	ATP-dependent nucleosome sliding	SMARCA4
MOR	SWI/SNF	SWIRM, SAINT	Cofactor for BRM	SMARCC1, SMARCC2
OSA	SWI/SNF?	BRIGHT	?	ARID1B

Table legend of Table 1.1 continued.

ARID1B, AT-rich interactive domain 1B; ASH, absent, small, or homeotic discs; ASX, Additional sex combs; BRM, brahma; CBX, chromobox homologue; EED, embryonic ectoderm development; E(PC), Enhancer of *Polycomb*; ESC, extra sex combs; ESCL, extra sex combs like; E(Z), Enhancer of ***zeste***; MLL1, myeloid/lymphoid or mixed lineage leukaemia; MOR, moira; MTF, metal response element-binding transcription factor; MXC, multi sex combs; NPCD, neuronal pentraxin with chromodomain; PC, Polycomb; PCL, Polycomb-like; PH, polyhomeotic; PHC, polyhomeotic-like; PHF19, PHD-finger protein 19; PHO, pleiohomeotic; PHOL, pleiohomeotic-like; SCE, Sex combs extra; SCM, Sex comb on midleg; SFMBT, Scm-related gene containing four MBT domains; SU(Z), Supressor of ***zeste***; SXC, super sex combs; TRX, Trithorax; WBP7, WW-domain binding protein 7; YY, Yin-Yang transcription factor.

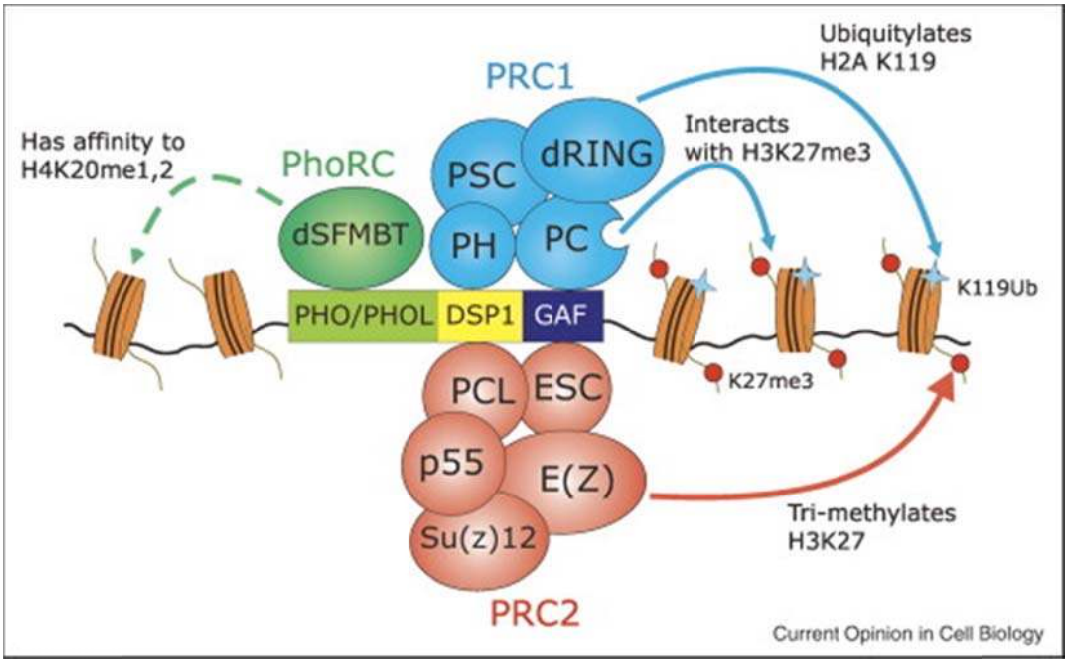


Figure 1.3 The PcG complexes bind on the PRE.

Directly adapted from Schwartz & Pirrotta (2008). PcG proteins form three complexes, PRC1, PRC2 and PhoRC, color coded here in blue, red and green. Those complexes are recruited to the PRE by a combination of DNA binding proteins including Pho, Dsp1 and Gaf. Specifically, dRING monoubiquitylates K119 of H2A, E(z) tri-methylate H3K27, and Pc recognize this H3K27me3 mark. dSFMBT bind to K20me1 and K20me2 of H4.

1.5 Transvection, Pairing Dependent Silencing, Trans-interaction

The somatic homologous chromosomes of Drosophila are intimately paired during interphase, thus significantly influence the gene expression in *trans* by the *cis*-enhancers or silencers. E.B. Lewis introduced the term transvection in 1954 (Lewis, 1954) to describe pairing-dependent complementation in the BX-D complexes in Drosophila, in which the regulatory elements on one homologs could influence the gene expression on the other homologs in *trans*. Since then, over a dozen of loci have been reported to have transvection effects, most in Drosophila, several cases described in other organism as well (Duncan 2002). To detect transvection genetically, one allele needs to have a deficient enhancer or regulatory element, and a second allele has a defect promoter or coding region. Then trasnvection could often be observed as interallelic complementation between two such alleles. One case is the *apterous* locus, where with wing-enhancer deletion on one homolog and *ap* promoter deletion on the other homolog, the *ap* gene still could be activated in *trans* (Gohl et al. 2008).

Another pairing-dependent regulatory interaction is the pairing-dependent silencing mediated by PREs. The mini-*white* gene is the most often used reporter gene for such effects in transgenic flies, it has a weak promoter, and sensitive to regulatory effects or chromosomal context, and causes the mutant white-eyed flies to have colored eyes after expression (Pirrotta et al. 1994). Generally, two copies of the *white* gene will produce darker eye color than a single copy of *white* gene. But when linked with a

silencer, the silencing will often be enhanced in flies homozygous (two copies of *white* gene and lighter eye color) than heterozygous (one copy of *white* gene but have darker eye color) for the construct. And the enhancement of silencing is dependent on the paring of the somatic chromosomes, thus called pairing dependent silencing. In most cases, the regulatory silencers that cause pairing dependent silencing are PREs. However, not all PREs could cause pairing dependent silencing. These DNA elements are hypothesized to be composite sites that contain the sites important for silencing (Polycomb complexes binding sites), and sites important for "pairing" (bring distant DNA elements together) (Kassis 2002). Currently, there are a little more than a dozen DNA elements reported to cause pairing dependent silencing, such as *engrailed*, *bxd* PRE, *Mcp*, *Fab-7*, *Fab-8*. *bxd* PRE is from the *Ubx* gene regulatory region, and could cause efficient pairing dependent silencing and strong silencing of *white* gene even in heterozygotes, depending on the genomic context (Chan et al. 1994). *Mcp* PRE is a weak PRE, but could have efficient pairing dependent silencing effect, even with the 810bp core fragment (Muller et al., 1999). For *Fab-7*, the PRE part also could mediate pairing dependent silencing effect, even with as little as the 260bp core fragment (Mishra et al. 1997).

An interesting phenomenon in the genome is the occurrence of long-range trans-interactions observed in many different organisms, and often playing important regulatory roles in gene expression. As described above in the previous section, the β-globin locus, and Igf2/H19 locus were each found to form intra-chromosomal

trans-interactions (Murrell et al. 2004, Tolhuis et al. 2002). While for the inter-chromosomal trans-interactions, one interesting observation is that olfactory receptor (OR) enhancer element activates only one of various OR genes located in other chromosomes in an olfactory neuron (Lomvardas et al. 2006).

More remarkable is the apparent ability of some PRE-containing DNA fragments to interact in trans with copies of the same construct inserted at remote sites, even on different chromosomes, and again resulting in enhanced repression. At first sight, this appears to be similar to the homologous pairing-dependent silencing effect except that something other than homologous chromosome pairing brings the two remote copies together. This behavior has been observed with constructs containing either *Mcp* or *Fab-7* : *Mcp*-containing constructs could mediate long-distance interaction in 90% of the cells of the eye imaginal disc, causing a lighter eye color when two transgenes recombined on the same chromosomes in heterozygotes, *Fab-7* transgenes could interact with endogenous *Fab-7* in 20% of the cells and cause a more severe wing phenotype (Vazquez et al. 2006, Muller et al. 1999, Bantignies et al. 2003).

1.6 Insulators

To package large eukaryotic genomes into small space of the nuclei while keeping efficient gene transcription at the same time, each gene may be assigned an "address" (the specific position in the genome), and similarly regulated genes are organized into domains which may have similar epigenetic markers to facilitate gene regulation in various development stages and in different types of cells (Lunyak 2008). Regulatory elements such as enhancers and silencers in the genome act bi-directionally over long distances to control spatial and temporal patterns of transcription, and show limited promoter specificity. And heterochromatin could encroach on adjacent domains to perturb gene expression. In addition, highly mobile chromosomes may allow illegitimate interactions between enhancers, silencers, and promoters. To coordinate efficient organization of the genomes, robust mechanisms must exist to prevent inappropriate regulatory interactions between elements so that long-range control is maintained appropriately. To maintain independence of individual domains, specific sequences are needed to establish the well-defined borders to prevent interference of their surroundings. "Insulators" or "boundary elements" are thought to be such kind of DNA elements that could protect genes or domains from activating or repressing signals coming from their adjacent domains (West et al., 2002).

Originally insulators were experimentally defined by two functional properties (Fig. 1.4): 1) the ability to block promoter-enhancer interactions, and 2) the ability to shield

transgenes from position effects caused by surrounding chromatin. Enhancer blocking does not result in inactivation of enhancers, nor does it interfere with promoter function (Gaszner & Felsenfeld, 2006; Kuhn & Geyer, 2003; Valenzuela & Kamakaka, 2006).

Insulator elements have been described in many organisms, including yeast, *Drosophila* and vertebrates. In many cases, an element has been identified as an insulator if it possesses at least one of the two functional properties, and thus is called enhancer blocking insulator or barrier insulator. However, recently, functional dissection of two insulators demonstrated that the enhancer-blocking and protection against position effects are conferred by distinct sequences, uncoupling these activities (West et al., 2002; Valenzuela and Kamakaka, 2006; Gaszner and Felsenfeld, 2006). In yeast, insulators appear to delimit the boundaries of silenced chromatin at telomeres and at the silent mating-type loci HML and HMR. *Drosophila* insulators were the first to be characterized and this organism displays the largest collection of these elements so far, including the *scs* and *scs'* elements flanking one of the *hsp70* gene loci (Udvardy et al. 1985), the Su(Hw) insulator found in the *gypsy* retrovirus (Spana, et al. 1988) , and the *Mcp*, *Fab7* and *Fab8* of the Bithorax complex (Karch et al. 1994; Barges et al. 2000; Hagstrom et al., 1996; Mihaly et al. 1997). In vertebrates, insulators have been found at sites such as the ribosomal RNA genes of Xenopus, the chicken *β-globin* genes and the human T cell receptor-α/δ locus (Robinett et al. 1997; Zhong and Krangel, 1997; Bell et al. 1999). The conservation of

widespread distribution of insulators among different species suggests important roles for these sequences in the nuclear organization and function (West et al., 2002; Mongelard and Corces, 2001).

Currently the mechanism how insulators function is still poorly understood; two models have been proposed to explain the insulator effects based on *gypsy* insulator. There are evidence to support both models, which may not be mutually exclusive. The 'promoter decoy model' suggests that the insulator acts as a barrier which can interact with enhancer-bound transcription factors and then block the activation signal traveling from the enhancer to the promoter, but do no inactivate either. By this way, the insulator traps enhancer, thus the enhancer would not be able to interact with the promoter, and can not activate transcription. Another explanation, termed the 'structural model' assume that insulators can organize the chromatin fiber within the nuclear space, by the way of anchoring the chromatin to fixed subnuclear structure, thus create transcriptionally independent looped domains and then isolate the signal generated with in a domain (Valenzuela & Kamakaka, 2006; Mongelard et al. 2001; Geyer et al. 2002). The direct visualization of *gypsy* insulator bodies and loops provided direct evidence for the structural model (Byrd et al. 2003). With the use of FISH and 3C-related methods, more and more interactions and loops between insulators were found, and the structural model has gained popularity in the current insulator field.

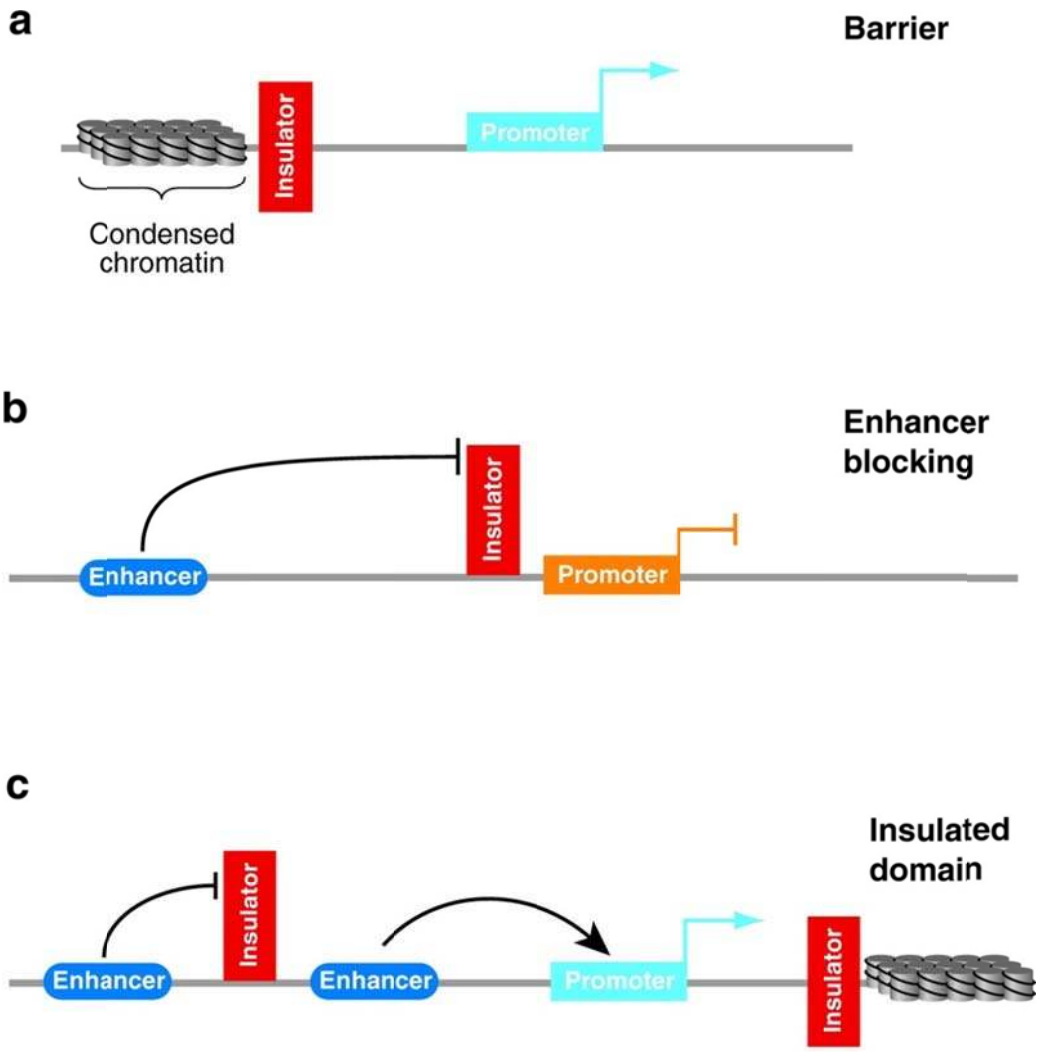


Figure 1.4 Insulators function in a position dependent manner.

Directly adapted from Valenzuela & Kamakaka (2006). (a) A barrier insulator protects a transgene against heterochromatin-mediated silencing. (b) An enhancer-blocking insulator will block transcriptional activation only when it lies between a promoter and an enhancer. (c) Flanking a transgene with two insulators generate a functionally independent domain protected from position effects.

1.6.1 scs and scs'

The first identified insulators were found in the 87A7 region of chromosome 3 of Drosophila, and called specialized chromatin structures, *scs* and *scs'*. This 87A7 locus contains a pair of divergently transcribed *heat shock protein* (*hsp*) *70* genes, and activation of *hsp70* upon heat shock accompanies a puff in the larval salivary gland polytene chromosomes. The nuclease sensitivity analysis on the borders of the puff identified unusual chromatin structures, the *scs* and *scs'* insulators. Both of them have two sets of strong nuclease-hypersensitive sites separated by a nuclease-resistant core, however they do not share DNA sequence similarity (Udvardy et al. 1985). To further characterize *scs* and *scs'*, they were placed on the flanking sides of mini-*white* reporter genes, which caused them to produce uniform eye colors in different transgenic fly lines, while a range of eye color from white to red due to position effects is generally observed without *scs* and *scs'*. The results mean that the insulators generated an independent domain and act as a barrier insulator (Kellum et al. 1991). Enhancer blocking assays were also carried out for both *scs* and *scs'* and they were found to insulate reporter genes from various enhancers (Kellum & Schedl, 1992).

The proteins that bind to *scs* and *scs'* have been identified, and their genomic localization also mapped by ChIP-chip. Zeste-white 5 (Zw5) protein binds *scs* (Gaszner et al. 1999), while boundary-element-associated factors (two isoforms: BEAF32A and BEAF32B) bind *scs'* (Hart et al., 1997). Immuno-localization studies

on polytene DNA showed that both BEAF and Zw5 bind to many sites throughout chromatin (Hart et al. 1997; Zhao et al. 1995). Blanton et al. (2003) showed that BEAF and Zw5 could interact with each in vitro and in vivo, bringing scs and scs' in close proximity detected by 3C, thus supporting the looping model. Genomic localization mapping data by ChIP-chip carried out by the modENCODE Drosophla Chromatin Consortium (http: // www.modencode.org), showed that both BEAF and Zw5 bind thousands of other sites, BEAF binding sites bias to the transcription start sites, and Zw5 binding sites do not fit an insulator pattern, suggesting that that they may have other functions other than insulator (Wallace et al. 2010; Nègre et al. 2010; Schoborg et al. 2010).

1.6.2 The *gypsy* Insulator

The most well-characterized *Drosophila* insulator is the *gypsy* insulator. This element was originally identified as the region of the *gypsy* retrotransposon responsible for causing tissue-specific mutations of several genes. The *gypsy* insulator resides within the 5' untranslated region of the *gypsy* retrotransposon. This 350bp-region contains 12 copies of a degenerate sequence, with a core of TGCATA embedded in AT-rich sequence. These sequences reflect a consensus binding site for the Su(Hw) protein. At least four of these binding sites are required for proper insulator function (Spana et al. 1988). However, most Su(Hw) binding sites in the genome contain a single copy of this sequence and clusters of three or more Su(Hw) binding sites are rare in the

genome (Ramos et al., 2006). The Suppressor of Hairy-wing (Su(Hw)) protein contains 12 zinc fingers and interacts directly with the insulator DNA. This interaction is essential for *gypsy* insulator function. The Su(Hw) protein is expressed throughout *Drosophila* development in most tissues (Spana et al., 1988; Gerasimova et al. 1995). Another protein involved in the complex is a specific isoform of Mod(mdg4) protein, which does not bind DNA directly and is recruited to the *gypsy* insulator complex via interactions between its C-terminal acidic domain and the bHLH-Zip domain of Su(Hw). Mod(mdg4)2.2 contains a BTB/POZ domain at the N-terminus that mediates homodimerizaton or multimerization of this protein (Ghosh et al. 2001), which make the *gypsy* complexes potentially sticky, and could gather together from large distance. The *Mod(mdg4)* gene encodes about 29 different products by complex alternative splicing. These isoforms have different functions, and mutation in this gene results in lethality, but mutations affecting only the isoform *mod(mdg4) 2.2* are viable and result in defective *gypsy* insulator function (Gerasimova et al. 1995).

By immuno-staining with antibodies against Su(Hw) and Mod(mdg4) 2.2, approximately 500 hundred of sites of Mod(mdg4) were found present in the polytene chromosomes from salivary glands, and overlap many Su(Hw) sites. Those sites do not contain copies of the *gypsy* retrotransposon, and are endogenous Su(Hw) binding sites, and might perhaps have similar property as to the one found in *gypsy* (Gerasimova & Corces, 1998). With over 500 endogenous binding sites, one could

expect diffused immuno-staining images within interphase nuclei of diploid cells. Instead, an impressive publication from Gerasimova et al. (2000) showed that only 20 to 25 body-like structure present in specific locations within the nucleus, and were termed insulator bodies. It is possible that several binding sites of those insulator proteins come together in a single location, and form large protein-DNA complexes. Interestingly, those bodies are not randomly located in the nuclei, but preferably to the nuclear lamina. Genetic experiments have shown that proper localization of insulator bodies requires an intact nuclear matrix scaffold, and in particular, the presence of lamin as well as an RNA component. Corces and colleagues suggested that chromatin may be organized into 50~200kb loops that attached to nuclear matrix mediated by special DNA sequences, called matrix attachment regions (MARs) or scaffold attachment regions (SARs) (Byrd et al. 2003). Insulators might be part of MARs that are involved in nuclear organization by bringing chromatin fiber to specific nuclear sub-compartments. *Gypsy* insulators have been shown to create chromatin loop domains by attaching to nuclear matrix (Byrd et al. 2003). However, there might be more complex scenario when considering many other insulators.

CP190, a third component of the Su(Hw) insulator complex, contains both a BTB/POZ domain and three C2H2 zinc fingers. CP190 can bind DNA on its own with low affinity and specificity, but does not interact directly with insulator sequences (Gurudatta & Corces, 2009). CP190 was previously identified as a centrosome-associated protein during mitosis that also associates with chromatin

during interphase. CP190 associates with chromosomes through interaction with both Su(Hw) and Mod(mdg4)2.2, thus is present at the ~500 polytene DNA sites where Su(Hw) and Mod(mdg4) bind, but it is also present at additional genomic locations independent of the other two proteins, perhaps through other DNA binding proteins (Pai et al., 2004). *Drosophila* Topoisomerase I-interacting RS protein (dTopors), which has been previously reported to possess an E3 ubiquitin ligase activity, is also required for *gypsy* insulator activity. dTopors interacts with known *gypsy* insulator proteins and promotes the enhancer blocking function of *gypsy*, also, it associates with the nuclear lamina and facilitates the coalescence of insulator bodies. But dTopors does not ubiquitinate known insulator proteins (Capelson & Corces, 2005).

The activities of the *gypsy* insulator might be regulated by SUMO conjugation. Capelson & Corces (2006) reported that SUMO modification of Mod(mdg4)2.2 and CP190 impairs the enhancer-blocking activities of the *gypsy* insulator without affecting the ability of CP190 and Mod(mdg4)2.2 to bind chromatin.

Another interesting observation from *gypsy* transgenes is that the insulating effects were neutralized when a direct tandem repeat of insulators was used instead of a single copy (Muravyova et al. 2001, Cai et al. 2001). This phenomenon probably suggests that enhancer-blocking may be not the primary function of the insulators. Instead, the primary functions of insulators might be to establish independent gene expression domains, by the way that two insulators physically interact and promote

the looping of the intervening sequence. So the enhancer-blocking properties of insulators may be just a by-product of this primary role, and could be explained by the 'structural model'. However, looping and neutralization of two insulators may by restricted and specific for *gypsy* insulators, since when tested in insulator bypass assays, heterogeneous combination of *gypsy* with *scs* insulators did not cancel the enhancer-blocking activity (Kuhn et al. 2003, Majumder et al. 2003), while *Mcp* insulator could neutralize *gypsy* insulator (Melnikova et al. 2004). This suggest that the interaction between insulators are specific, *scs* and *gypsy* insulators belong to different class of insulators, while *Mcp* and *gypsy* may share the same kind of insulator binding proteins.

1.6.3 *Fab-7* and *Mcp*

The *bithorax* complex (BX-C) is a cluster of homeotic genes in *Drosophila* that controls the morphological diversification of body segments. As one of two *Drosophila* Hox clusters, the *bithorax* complex (BX-C) is responsible for determining the posterior thorax and each abdominal segment of the fly. It does this by regulating the expression of the three BX-C homeotic genes: *Ultrabithorax* (*Ubx*), *abdominal A* (*abd-A*) and *Abdominal B (Abd-B*), which are required respectively for the identity of anterior abdominal and posterior thoracic segments, central abdominal segments, and posterior abdominal segments. Precise parasegmental expression patterns of these homeotic genes are crucial for generating a normal body plan, and mis-regulation of

these genes results in dramatic transformation one body segment into another (reviewed in Lewis 1998). The PS-specific expression patterns of *Ubx*, *abd-A*, *Abd-B* are determined by a complex cis-regulatory region that spans nearly 300kb (Fig. 1.5). For example, *Abd-B* expression in PS10, PS11, PS12 and PS13 is respectively controlled by the *iab-5*, *iab-6*, *iab7, iab8* cis-regulatory domains. Each *iab* domain appears to contain at least one enhancer that initiates *Abd-B* expression in the early embryo, as well as a PRE silencer element that maintains the expression pattern throughout development. It has been proposed that insulators flank each *iab* region and organize the *Abd-B* regulatory DNA into a series of separate chromatin loop domains. *iab-7* is flanked by two insulators, *Fab-7* and *Fab-8*. While *Mcp* reserves the functional autonomy of the *iab-4* and *iab-5* cis-regulatory domains (reviewed in Maeda et al. 2006).

Mcp and *Fab-7* are initially identified by genetic mutations that delete those cis-acting elements resulting in a gain-of-function phenotype that transforms the affected parasegment into a copy of the parasegment immediately posterior. This led to the suggestion that *Mcp* and *Fab-7* may correspond to the boundaries of cis-regulatory domains (Gyurkovics et al. 1990; Karch et al. 1985). *Mcp* was initially mapped to a 2.9-kb *Eco*RⅠ fragment, and it was later shown that a ~800bp core fragment retains all the activities of the full *Mcp* (Li et al. 2010; Muller et al. 1999). Interestingly, this 800bp core *Mcp* functions both as a silencer (PRE) and an insulator. The minimal PRE part that can maintain silencing ability during imaginal disk development is as

small as 138-bp, which contain GAGA and PHO binding sites (Busturia et al. 2001). Immediately adjacent to the PRE part sits the 340-bp insulator, which could block enhancer activities and mediate trans-interaction in transgene assays (Gruzdeva et al. 2005), and even a 210-bp core inside this 340-bp fragment still retains pairing ability for remote *Mcp* transgenes (Kyrchanova et al. 2007). Similarly the 3.6-kb *Hind* Ⅲ *Fab-7* fragment also contains an insulator part and a PRE part (Hagstrom et al. 1996; Zink et al. 1995). Mishra et al. (2001) located a 230-bp fragment (HS3) in the *Fab-7* that could function as PRE in vivo and induce pairing-dependent silencing, this small fragment also contains binding sites for GAGA and PHO. Cavalli and colleagues continued to show that Zeste and TRX could also bind the 219-bp *Fab-7* minimal PRE (Déjardin et al. 2004), and found a novel factor, called DSP1, which could bind to this minimal PRE part and is important for the PcG recruitment (Déjardin et al. 2005). The minimal *Fab-7* insulator was first mapped in the *Pst* Ⅰ-*Apa* Ⅰ 1.2-kb fragment (Schweinsberg et al. 2004), then down to 0.86-kb sub-fragment that still could block enhancers with similar efficiency (Rodin et al. 2007), and the GAGA-factor-binding sites are necessary for the full *Fab-7* insulator activity.

Genetically, *Mcp* and *Fab-7* have been shown to mediate trans-interactions between two copies of a transgene, by the pairing-dependent eye-color assay (Muller et al. 1999; Gruzdeva et al. 2005; Kyrchanova et al. 2007; Rodin et al. 2007). Two interesting reports provide direct evidence that *Mcp* and *Fab-7* have the ability to mediate long-distance communication between transgenes located at distant sites of

the same chromosome arm or even on different chromosomes. Vazquez et al. (2006) showed that the *Mcp* transgene could mediate long-distance interaction in 90% of the eye-disc cells between two, even up to four, copies of *Mcp* transgenes by using LacO Array/GFP-LacI in vivo imaging system. Bantignies et al. reported that by using the technique of 3D-FISH, the *Fab-7* transgene could interact with endogenous *Fab-7*, which is important for the repression of the reporter gene. They went on to show that the *Fab-7* transgene could also co-localize with another *Fab-7* transgene on another chromosome inside PcG bodies, even in the absence of endogenous *Fab-7* (Bantignies et al. 2003). Since the insulator could mediate the trans-interaction from genetic evidence, it is likely that the long-distance interaction between *Fab-7* and *Mcp* observed could be due to the presence of insulator part.

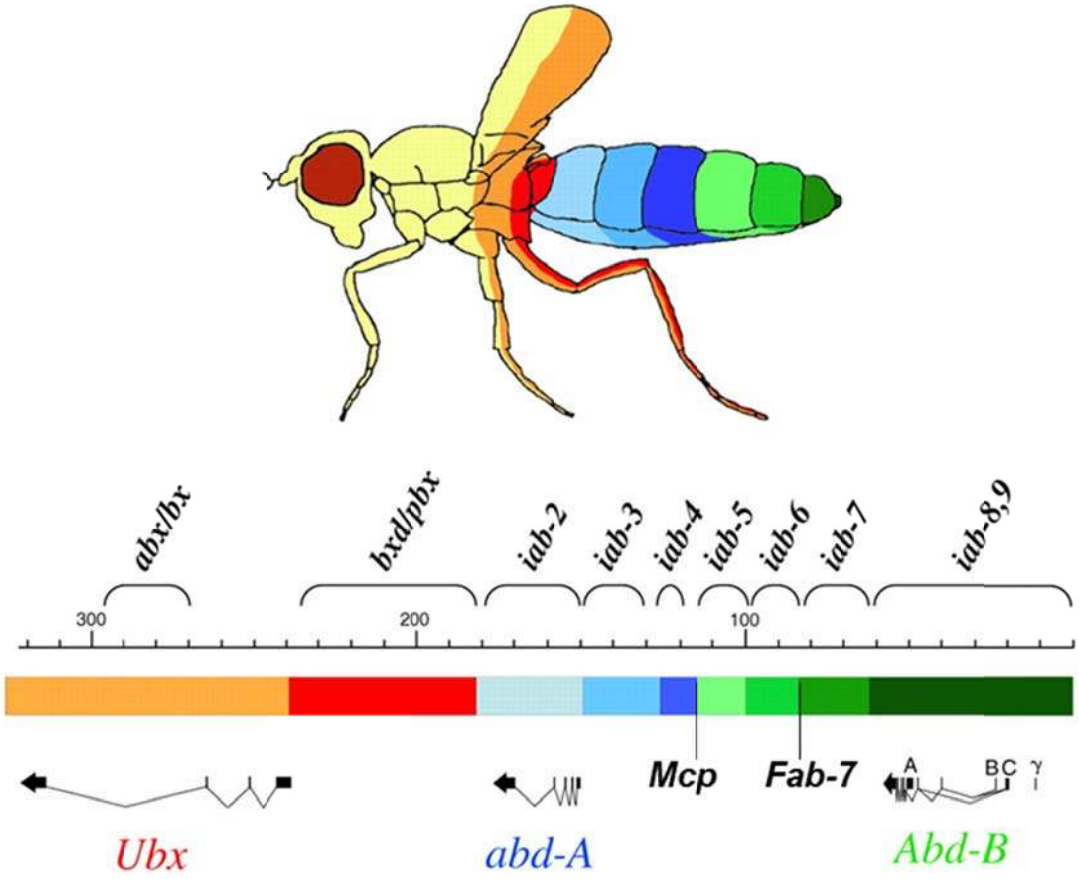


Figure 1.5 Diagram of the BX-C.

Directly adapted from Maeda et al.(2006). The 300kb of genomic DNA represented by multicolored bar cotains three genes draw below the bar: *Ubx, abd-A* and *Abd-B*; and the cis-regulatory regions indicated above the line with different colors, which correspond adult segments. *Mcp* and *Fab-7* are highlighted below the bar.

1.6.4 Genomic Location and Regulation of Insulator Proteins

CCCTC-binding factor (CTCF) is a highly conserved and ubiquitous DNA-binding protein that had been implicated to function both in gene silencing and activation. Felsenfeld and colleagues identified a 1.2-kb DNA element that derived from the 5' end of the chicken β-*globin* locus, and this region sits exactly between an active chromatin and inactive domain. Transgene assays in cultured cells showed that this 1.2-kb fragment (Chung et al. 1993), even down to the 250-bp core (Chung et al. 1997), could both block the enhancer activity and protect the transgene from position effects. Then they identified CTCF, which is the protein that binds to a 42-bp fragment in this 250-bp core, and confers the insulator activity (Bell et al. 1999). Later work showed that CTCF is the only protein so far known in vertebrates that confers enhancer blocking activities for all known vertebrate insulators.

Moon et al. (2005) showed that the orthologous CTCF in Drosophila has similar domain structure and binding site specificity as in vertebrates, and binds *Fab-8* to confer the enhancer blocking activity. In addition to dCTCF, Gerasimova et al. (2007) found that CP190 also binds to *Fab-8*. They also showed that dCTCF co-localizes with a subset of hundreds of CP190 sites on Drosophila polytene chromosomes, and proposed that dCTCF and Su(Hw) may be two different sub-sets of insulators that recruit and share CP190 in the organization of the chromatin fiber in the nucleus.

Cohesins form ring-shaped complexes that mediate sister chromatin cohesion in dividing cells. Recent studies have showed that cohesins also contribute to gene regulation in interphase cells. Wendt et al. (2008) showed that cohesin co-localizes with CTCF in the human genome, and is required for the CTCF to block enhancer activities. Parelho et al. (2008) also found cohesin at most CTCF sites and CTCF is required for cohesin localization to these sites. Cohesins functionally interact with CTCF and mediate enhancer blocking via CTCF-dependent recruitment to insulator sites. But no such kind of association between CTCF and Cohesin has been found in Drosophila, probably because there are other insulator proteins that could be used for Cohesin instead of CTCF.

With advent of the technique of ChIP-chip / ChIP-seq, the human genomic binding sites of CTCF were quickly determined, and showed that there are more than 13,000 CTCF binding sites, and the number varies depending on the cell type (Kim et al. 2007; Barski et al. 2007; Xi et al. 2007). Most of the binding sites are far away from the transcriptional start sites and strongly correlate with genes, indicative of insulator activities. The genomic CTCF binding sites in *Drosophila* have also been determined recently, and the results showed that the insulator elements, *Mcp*, *Fab-7*, *Fab-8*, all bind to dCTCF, in addition to more than 2,000 other sites (Holohan et al., 2007; Bushey et al., 2009; Nègre et al., 2010). More recently, genome-wide ChIP-chip analysis of all insulator proteins in Drosophila, including SU(HW), dCTCF, MOD(MDG4), CP190, BEAF and ZW5, have been carried out by the modENCODE

Drosophila Chromatin Consortium (**http://www.modencode.org**). The analysis from Corces and colleagues showed that SU(HW), dCTCF, and BEAF each bind to a subclass of CP190 binding site, and have their unique distribution patterns (Bushey et al., 2009). ModENCODE results also confirmed that there exist at least 3 classes of insulators in Drosophila.

The insulator proteins identified so far in Drosophila include BEAF32, ZW5, SU(HW), Mod(Mdg4)2.2, CP190 and dCTCF, while only one, CTCF, is known in vertebrates. So, at first, it seems puzzling that, Drosophila has at least 3 different classes of insulators with smaller genome, while human only have one with 20 times bigger genome. One possible reason is that there are more insulator proteins in the vertebrate to be found. It is also possible that the functions of all three classes of insulators in Drosophila converged into one during the course of evolution, since all three classes of insulators in Drosophila share the common factor of CP190, which eventually capture strong DNA-binding abilities, fuse and evolve into CTCF in vertebrates.

Since insulators mediate loop formation and higher-order chromatin organization, they are also important for gene regulation, which is exemplified by several cases discussed above. So there are also existing mechanisms to regulate insulator activities to establish distinct nuclear architecture that is cell type specific. Work from Corces and Colleagues showed that there are 3747 SU(HW), 2266 dCTCF, 2995 BEAF and

5272 CP190 sites in the neural-derived Kc cells,. They also mapped the binding sites of those insulator proteins in Mbn2 cells, a hematopoietic cell lines (Bushey et al., 2009). The results showed that a fraction of the binding sites for each of the insulator proteins is specific for each cell lines, for example, 18% of SU(HW) sites in Kc cells and 5% in Mbn2 cells are cell type specific. It is possible that the cells regulate insulators by controlling the recruitment of DNA binding proteins to their target sites (Gurudatta & Corces, 2009). There are other ways for insulator regulation. dTopors has E3 ubiquitin ligase activity, and this activity is important for *gypsy* insulator function (Capelson & Corces, 2005). SUMO conjugation of Mod(mdg4)2.2 and CP190 could inhibit enhancer blocking activities of the *gypsy* insulator (Capelson & Corces, 2006). Rm62 RNA helicase is a component of the *gypsy* insulator complexes, and mutation of Rm62 enhances the *gypsy* insulator function, while mutation of other components of the RNAi machinery inhibits *gypsy* insulator function (Lei & Corces, 2006). The insulator network is complex in Drosophila, the regulation of insulator function adds more complexity, and the regulation mechanism is elusive and our knowledge of it is far more incomplete.

1.7 RNAi Pathway

RNAi is a highly conserved, post-transcriptional cellular mechanism to silence gene expression in a sequence-specific manner. The RNAi process is initiated by double-stranded RNAs (dsRNA) that are homologous in sequences to the target genes. These dsRNAs are cleaved by the RNase Ⅲ enzyme Dicer into ~21-25 nucleotide fragments. The Argonaute protein-containing complex, called RNA-induced silencing complex (RISC), picks up one RNA strand as a guide to find and degrade the target mRNA, or inhibit the translation of target mRNA (Bellés 2010).

RNAi can also control gene expression by altering the state of chromatin. In *S. pombe*, this involves bi-directional transcription from repetitive sequences, producing dsRNAs which are cleaved into short interfering RNAs (siRNA) of 21-23nt by Dicer-1. siRNAs guide the RNA-induced initiation of transcriptional gene silencing (RITS) complex to homologous sequences in the nucleus. Clr4, the homolog of the histone methyltransferase Su(var)3-9, is recruited along with the RITS complex to chromatin, where it methylates lysine 9 of histone H3 (H3K9). This epigenetic mark promotes the formation of heterochromatin by recruiting the heterochromatin protein Swi6, the homolog of HP1, via its chromodomain (Pal-Bhadra et al. 2004).

Two interesting reports found a close correlation between the RNAi pathway and insulator function. Grimaud et al. (2006) used the 3.6kb full-length *Fab-7* transgenes and found that certain components of the RNAi machinery, *dicer-2*, *piwi*, *argonaute1*, and *aubergine*, are required for efficient *Fab-7* pairing sensitive silencing (PSS) and long distance chromosome interactions with the endogenous *Fab-7* element. Mutation of AGO1, aub, piw and dcr-2 disrupt the long distance colocalization and decrease the repressive effect of *Fab-7* transgenes. Several RNAi proteins localize to distinct nuclear foci and a subset of these foci colocalize with PcG bodies. In addition, 21-23 nt RNAs corresponding to *Fab-7* insulator region have been detected. The authors ascribed the long-distance interaction between *Fab-7* transgenes to PcG proteins, and also showed that two distant PRE-containing loci (BX-C and ANT-C) co-localize frequently in the same PcG bodies only when both loci are transcriptionally repressed (Bantignies et al. 2011). Lei and Corces (2006) used *gypsy* insulators, and found that the insulator binding protein CP190 interacts directly with Rm62 helicase, an RNAi protein involved in double-stranded RNA (dsRNA)-mediated silencing and heterochromatin formation. In addition, Argonaute proteins are required for proper *gypsy* insulator functions. From those studies, a common factor that involved in the RNAi related trans-interaction is that insulators are involved. . Those observations may lead us to hypothesize that the insulator part of *Fab-7* could be responsible for the long-distance interaction, and for the overlap between RNAi bodies and PcG bodies, both of which may occupy the silencing zone sub-compartment of the nuclei.

1.8 Techniques for DNA interactions

Linear chromatin fiber is packed inside the nuclei as a complex three-dimensional structure, and the organization of the chromatin has important roles in the appropriate spatial and temporal regulation of gene expression. To understand how chromatin organize inside nuclei, and how regulatory elements physically interact with genes, the microscopy based Fluorescence in situ hybridization (FISH) and in-vivo live imaging, and the molecular technique of Chromosome Conformation Capture (3C) have been developed, and help us understand more and more of the higher order chromosomal conformation.

1.8.1 FISH

Fluorescence in situ hybridization (FISH) has been greatly developed and extensively used to probe the chromosome structure since its first introduction in 1969 (Gall et al. 1969). The principle of current FISH protocol is simple: the fluorescent probes recognize and hybridize to their endogenous sequences, which then could be visualized under microscope. The FISH technique have been successfully used to address many important questions, such as the positions of genes on chromosomes, the discovery of chromosomes territories, diagnosis of chromosomal abnormalities, and DNA long-distance interaction in interphase. The advantage of this technique is that it provides information on the frequency of interaction in a single cell level. However, the protocol is complex, and difficult to obtain high quality fluorescence

images, in addition its poor resolution (usually > 0.2 μm), which prevent its use in the detection of relatively short-distance DNA interaction.

1.8.2 Live imaging

To investigate the interphase chromosome structure and dynamics while preserving the high-order chromatin structure without fixation, Belmont and colleagues developed the LacO/GFP-LacI based in vivo live imaging (Robinett et al., 1996; Straight et al., 1996). Based on the facts that the bacterial lac repressor (LacI) could highly specifically and efficiently bind to its target sequence lac operator (LacO), even within eukaryotic cells without any detrimental effects, this method has been proven to be successful and widely useful in a variety of systems. This method has two components, the 64 to 256 copies of LacO repeat tagged to the transgene of interest, and the LacI-NLS (nuclear localization signal) fused GFP driven by the promoter/enhancer of interest (Belmont & Straight, 1998). When introduced into the same cells of the two components, the DNA inside live cells could be visualized directly under fluorescence microscope as a dot. A similar method using tet operator-repressor has also been developed, and when those two system combined within the same cells, the DNA interaction from two different genomic loci could be visualized. If combined with other GFP or RFP fused proteins, this method was also used for the study of DNA-protein interaction. Vazquez et al. (2006) adapted this method into *Drosophila*, using a heat shock promoter to drive the expression of GFP-LacI. The problem with the heat shock promoter is that there are only a small

portion of the total tissue cells express GFP. Csink and colleagues then constructed a Ubiquitin promoter driven RFP-LacI, which could express RFP in all the cells (Thakar et al. 2005; Thakar et al. 2006), but the RFP is very unstable and photo-bleached quickly.

The advantage of the LacO/LacI live imaging is that you could study the chromosomal dynamics under its physiological conditions, and acquire time-lapse images. However, the problem with this technique is that the GFP could be easily photo-bleached, rendering the acquisition of high-quality images very difficult. The advent of semiconductor quantum dots (Qdots) promised to improve dramatically the photobleaching problems, in addition to many other benefits (Michalet et al. 2005; Gao et al. 2005). Although Qdots have been applied to lots of research, mainly in cell lines, the problem with it is its still relatively big size, which makes it very difficult to penetrate into the tissue.

1.8.3 Chromosome Conformation Capture

The technique of Chromosome Conformation Capture was first invented by Dekker and colleagues to study the conformation of a complete chromosome in yeast (Dekker et al. 2002), and was quickly adapted to investigate the interactions between complex gene loci and loops. Within a few years, it soon gained popularity, evolved into several 3C based variants and became a standard research tool to study chromosomal interactions.

The principle of 3C technology is based on the formaldehyde fixation of cells which crosslink the protein-DNA complexes in close proximity, followed by proper restriction digestion, then ligated in diluted condition in favor of intra-molecular ligation. After reversion of crosslinks, the ligation products are quantified by PCR, with PCR primers specific for the fragments of interest (Splinter et al. 2004; Hagège et al. 2007). The technique was extensively used to identify the physical interactions between distant DNAs and chromatin loops. This technique could be used to study the genomic region of any size, and between different chromosomes. However, to draw meaningful conclusions, proper controls must be implemented in the experiment design (Dekker 2006). 3C and 3C-based technology only provide information about the DNA interaction frequency of a large population of cells. To explain the function of those 3C interactions, additional information needed, such as RNAi knockdown, transcriptional changes corresponding to the interaction.

To expand the usage of the 3C technique, several 3C-based technologies have been developed for various purposes (Fig 1.6). To investigate a special interaction mediated by a protein of interest among a heterogeneous population of cells, the ChIP-loop assay was developed, which add one additional step before intra-molecular ligation: immune-precipitation using an anti-body against the protein of interest (Horike et al. 2005). To map chromatin interaction in large-scale, Dekker and colleagues developed the 3C-carbon copy (5C) (Dostie et al. 2006). 5C use a

multiplex ligation-mediated amplification step to amplify selected ligation junctions, and then analyzed by microarray or sequencing. This method could analyze a region up to several mega-bases from multiple baits, but not suitable for genome-wide scan because of tens of thousands 5C primers needed. To unbiasly scan the whole genome that interact with the DNA fragments of interest (bait), several groups independently developed 4C (chromosome conformation capture-on-chip, or circular chromosome conformation capture, open-end 3C and 'olfactory receptor' 3C) method (Zhao et al. 2006; Simonis et al., 2006; Würtele and Chartrand 2006; Lomvardas et al., 2006). 4C uses two bait-specific primers to amplify all the fragments that captured by the bait, and puts the PCR products on microarray or high-though sequencing.

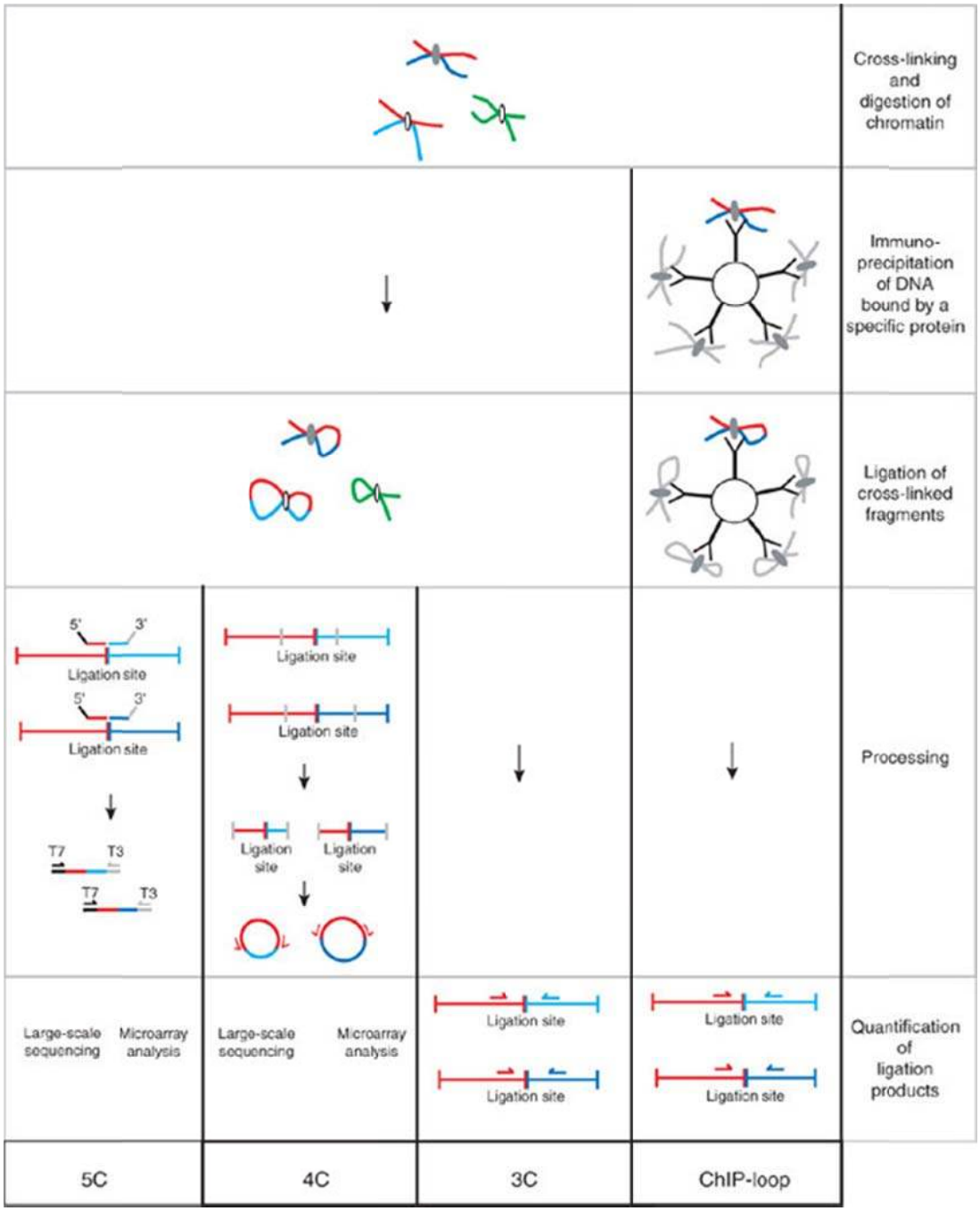


Figure 1.6 Schematic representation of 3C-based technologies.

Directly adapted from Simonis et al. (2007). See text for detailed explanation.

1.9 Perspectives

It has become more and more clear that insulator may be involved in the establishment and maintenance of higher-order chromatin organization. Although great progress have been made, such as identification of several insulator proteins, mapping of genomic binding locations and classification of different insulators, there still may be more insulator proteins that need to be discovered. And little is known about how they are organized, and how they are regulated.

An interesting phenomenon is that PREs are often associated with insulators, at least in the BX-C, so it is possible that the formation of PcG bodies is actually mediated by insulators, not PREs. The model is attractive, since insulators have structural properties and mediate long-distance trans-interaction. The PcG bodies could be a pre-assembled sub-nuclear structure that bring together and attached to nuclear matrix by insulators. They may have their preferred location in the nuclei, the silence zone. The PcG proteins/module could maintain memory of chromatin states, so it is also possible that PcG bodies could memorize their preferred nuclear location after each cellular division.

It seems that insulators not only mediate the formation of silencing bodies, but also involve in gene activation and transcription. It has been shown that CTCF mediate the long-distance enhancer action and loop formation. One major breakthrough of our

understanding of transcription is the theory of transcription factories, which gain several evidence, but still controversial. One gap of the theory is the missing of the factors that recruit genes to the bodies and mediate the formation of the bodies. In regarding to the property of insulators that scatter all over the genome and its involvement in transcription, it is likely that it is insulators that bring genes into Pol Ⅱ bodies, which occupy the active zone of the nuclei.

Then, how and what kind of insulator elements are organized in a single PcG body or Pol Ⅱ factory? That is to say, in a PcG body, do all PREs come together with adjacent insulators, or do all active genes in a Pol Ⅱ factory have adjacent insulators? Then next question is that whether those PREs come into the body because they are close to each other and close to the body, or because of their sequence homology? The active genes come into the same Pol Ⅱ factory because they are in the same chromosome domains, or in distant regions but co-regulated?

So, it would be very interesting to see how the cell put several elements together in a body: does it follow some kind of code (insulator code), or just random? ChIP-chip techniques have helped us to identify most of the potential insulator, PRE elements, enhancers and promoters. And the 4C approach will enable us address these questions directly. FISH is another good tool to investigate the organization of PcG bodies in different tissues and developmental stages. By combining these technologies with bio-informatics, some of these mysteries could possibly be revealed.

Chapter 2. Insulators, not Polycomb complexes, mediate long-range interactions between Polycomb targets in Drosophila

The genetic work decribed in paragraph 2.4.1 and 2.4.2 were done by Martin Müller and Ilham Bahechar, and the genetic work in paragraph 2.4.3 done by Olga Kyrchanova. Katsuhito Ohno injected the plasmid into the embryos to produce the fly lines described in figure 2.2B.

2.1 Abstract

The genomic binding sites of Polycomb (PcG) complexes have been found to cluster, forming Polycomb “bodies” or foci in mammalian or fly nuclei. These associations are thought to be driven by interactions between PcG complexes and result in enhanced repression. Here we show that a Polycomb Response Element (PRE) with strong PcG binding and repressive activity cannot mediate trans-interactions. In the case of the two best studied interacting PcG targets in Drosophila, the Mcp and the Fab-7 regulatory elements, we find that these associations are not dependent on or caused by PcG complexes. Using functional assays and physical co-localization by in vivo fluorescence imaging or 3C methods, we show that the interactions between remote copies of Mcp or Fab-7 elements are dependent on the insulator activities present in these elements and not on their PRE. We conclude that insulator binding proteins and not PcG complexes are responsible for the higher-order organization of PcG targets in the nucleus.

2.2 Introduction

Transgenes containing Drosophila Polycomb Response Elements (PREs) often show a remarkable degree of pairing-enhanced silencing, the increased repression observed when the transgene is present in two allelic copies (Kassis 2002). A PRE can also silence in trans a reporter gene lacking its own PRE but inserted at the same site on the homologous chromosome (Sigrist & Pirrotta, 1997). Both are consistent with a looping model proposed to explain how the *bxd* PRE can produce H3K27 trimethylation and silencing of the *Ubx* promoter many tens of kilobases distant (Kahn et al. 2006). According to this, PcG complexes bound to a PRE can contact and interact with chromatin regions in their physical neighborhood, whether on the same chromatin strand or on a separate strand.

More remarkable is the apparent ability of some PRE-containing DNA fragments to interact in trans with copies of the same construct inserted at remote sites, even on different chromosomes, and again resulting in enhanced repression. At first sight, this appears to be similar to the homologous pairing effect except that something other than homologous chromosome pairing brings the two remote copies together. This behavior has been observed with constructs containing either of two PcG-binding elements from the Bithorax Complex, *Mcp* and *Fab-7* (Grimaud et al., 2006; Muller et al., 1999; Vazquez et al., 2006) and it has been frequently attributed to a general tendency of PcG complexes bound to one genomic site to interact with PcG complexes bound at other sites in the genome.

The idea that PcG-binding sites in the nucleus might tend to cohere is consistent with the observation that, in flies or mammals, staining of diploid nuclei with antibodies against PcG proteins reveals a small number of foci, relative to the many hundreds of binding sites known to be present in the genome. That the PcG complexes might drive this association is supported by the finding that the Drosophila *Antennapedia* gene and the *Abdominal-B* gene, several megabases distant from one another, co-localize when both are repressed but not when one of the two is transcriptionally active (Grimaud et al. 2006) . Furthermore, it has been proposed that this interaction is mediated by RNAi mechanisms (Grimaud et al. 2006), implying their participation in PcG repressive complexes and providing an attractive link to the role of the RNAi machinery in heterochromatin formation (Grewal et al. 2007). In this view, then, PcG complexes are inherently "sticky" and random or RNAi-mediated encounters in the nucleus would cause PcG binding sites to aggregate and, as in pairing-enhanced silencing, result in stronger or more stable repression.

Against the idea that PcG complexes are intrinsically cohesive is the observation that not all PREs have been found to trans-interact with remotely inserted copies. This has never been observed with constructs containing the powerful *bxd* PRE silencer, although these constructs can exhibit pairing-dependent repression when made homozygous (Sigrist et al. 1997; V.P., unpublished). Both *Mcp* and *Fab-7* elements have been shown to contain two distinct and separable activities: a PRE activity and an enhancer-blocking insulator/boundary activity (Gruzdeva et al., 2005; Hagstrom et al., 1996; Sigrist & Pirrotta, 1997; Zhou et al., 1996). This raised the possibility that the ability of these elements to enter

into long-distance interactions might be mediated by their insulator component. Consistent with this, the *bxd* PRE 640 bp fragment, although incapable of remote trans-interactions by itself, acquired this property when associated with the *gypsy* Su(Hw) insulator element (Sigrist & Pirrotta, 1997).

Here, we examine the relationship between PRE activity and the ability to mediate long-distance trans-interactions, comparing the strong silencer *bxd* PRE and the weak silencer *Mcp*. The results show that it is not the PRE that mediates trans-interactions but an insulator activity closely associated with the *Mcp* or *Fab-7* PREs. We show in addition that these insulators but not the PREs mediate the ability of transgenic insertions to become closely juxtaposed with one another and with the corresponding endogenous element in the nucleus.

2.3 Materials and Methods

2.3.1 Transgene constructs

The Flipper constructs were assembled on the pC4YM plasmid backbone. Flipper2$^{Mcp\text{-}bxd}$ was described in Gohl et al.(2008). Unique *Xho*I and *Not*I sites were used to introduce the various *bxd-Mcp* cassettes. The *bxd* PRE is in all cases the 661 bp *Nde*I-*Pst*I fragment used by Sigrist and Pirrotta (1997), flanked by FRT sites. The following Mcp fragments flanked by LoxP sites were tested: Flipper2 contains a 2.9 kb *Eco*RI fragment (Muller et al. 1999); Flipper21 contains a 0.9 kb fragment extending from *Xba*I to the distal *EcoR*I site; Flipper22 contains a 1.2 kb fragment extending from the *Sal*I to the proximal *Eco*RI site; Flipper23 contains the central 0.8 kb *Sal*I-*Xba*I fragment; Flipper24 contains a 210 bp fragment previously described by Kyrchanova et al. (2007); Flipper25 contains a 755 bp *Pst*I–*Pst*I fragment nearly identical to the 0.8 kb fragment in Flipper23, but lacking the 210 bp fragment of Flipper24$^{Mcp\text{-}bxd}$. The orientation of the Mcp fragments was such that the end normally adjacent to *iab4* is closer to the *bxd* PRE. In Flipper23$^{Mcp\text{-}bxd}$, *Mcp* is inserted in the opposite orientation. The *Mcp* M^{Ins210} insulator fragment was obtained by PCR amplification of the DNA fragments between the 5’ aaacttaactcagacttgg 3’ and 5’ cccaatcgttgtaagtgt 3’ primers (fragment from nt 113994-114204). The *Mcp* core carrying a deletion of the 210 bp insulator (*Mcp*^{Ins}) was obtained by ligation of two fragments obtained by PCR amplification between 5’ gacttaaattgatttaaag 3’ and 5’ aatccaagtctgagttaag 3’, between 5’ ctgcagtcaaacgtcaca 3’ and 5’ cttacaacgattggg 3’.

These fragments were cloned between Lox sites and inserted into the FRT-flanked *bxd* PRE cassette and ligated as *Xho*I-*Not*I fragments upstream of the mini-*white* promoter into the pC4YM plasmid.

To assemble the *LacO-Mcp* and *LacO-Fab-7* constructs, the insulator and PRE portions of *Mcp* and *Fab-7* were PCR-amplified from BX-C clone BAC R24L18 (obtained from BACPAC Resources Center, http://bacpac.chori.org/), using PCR primers MI+: gatactgcagctcagagtacataagcg; MI-:tgaggggcccaagcgttgtaagtgtg for the minimal *Mcp* insulator fragment; primers MP+: cttgggatcctcatgtgttagtgcgtgag; MP- : acacaaacgcatctgcagtc for the *Mcp* PRE; primers FI+: caactgcagtgaagacacgaac; FI-: cgacgtgagcgaccgaaactc for the *Fab-7* minimal insulator and FP+: cggggatccgagtttcggtcgctcac; FP- : gaactgcagatgtcggcaattcggattcc for the *Fab-7* PRE. The amplified fragments were ligated into LoxP and FRT cassette plasmids and the resulting plasmids were sequenced to verify the inserted sequence. Fragments containing the mini-white gene, the insulator flanked by LoxP, and PRE flanked by FRT, respectively, were assembled into pBlueScript. The tandem array of 128 copies of LacO was cut from pAFS150 (a gift from J. Vazquez) cut and inserted into pC4Y, and this plasmid was used to accept the LoxP-flanked insulator part, FRT-flanked PRE part and *mini-white* gene.

Plasmid pBSKS-Ubq-mRFP-LacI-NLS containing the LacI repressor fused to mRFP red fluorescent proteins and driven by the ubiquitin promoter was kindly provided by A. Csink (Thakar and Csink 2005). The mRFP sequence in this plasmid was replaced by an EGFP

PCR-amplified fragment cut by *Cla*I-*BamH*I. The *Kpn*I-*Sac*I fragment of Ubq-EGFP-LacI-NLS was inserted into pCaSpeR4.

2.3.2 Fly stocks

Transgenic fly lines were made according to standard procedures (Spradling et al. 1982). Southern blot hybridization was used to verify that the lines contained a single insert and inverse PCR was used to identify the exact insertion sites. The various deletion derivatives were established with the help of Flipase and Cre recombinase-producing stocks (Siegal et al. 1996) as previously described in Gohl et al.(2008) and were verified by PCR analysis. For co-localization studies, two transgene lines on different chromosomes were crossed together through double balancers. In the case of lines with insertions on the same chromosome the two insertions were recombined to obtain a cis-arrangement. PCR was used to verify the presence of both transgenes, all flies were raised at 25℃ or room temperature.

2.3.3 In-vivo imaging and Microscopy

After crosses of transgenic *LacO-Mcp* or *LacO-Fab-7* flies with LacI-EGFP flies, the larvae were raised at 18℃and supplemented with active dry yeast. Third instar larvae were rinsed and dissected in Gibco Schneider's Drosophila media (Invitrogen Co.). The dissected eye and wing imaginal discs were aligned on a bottom dish (MatTek Co.) with a drop of Drosophila medium, and then covered with a coverslip. Z-stack images were taken on a DeltaVision Image Restoration Microscope system (Applied Precision Instrument, LLC Issaquah, WA) with 100 ×/1.35 UplanApo objectives,

deconvoluted and processed with the SoftWoRx software (Applied Precision Instruments). The dots in each nucleus were scored, if only one dot, or two dots touching each other (the center between two dots is less than 0.3 μM), then scored as co-localization; while if two non-overlapping dots are seen separated from each other (>0.3μM), then scored as no no-localization. Chi-square tests were used for pair-wise comparison of any two data sets in each category. All statistical analysis was done using the software JMP (SAS Institute Inc.).

2.3.4 Fluorescence in situ hybridization

To produce the probe for the BX-C locus, the BAC clone BAC R24L18 (obtained from BACPAC Resources Center, http://bacpac.chori.org/) was used as template and labeled by Biotin-Nick Translation mix (Roche). For the probe of transgenes, the LacO repeats containing plasmid pAFS150 (a gift from J. Vazquez) was used as template and labeled by DIG-Nick Translation mix (Roche). For other loci, several PCR fragments (totally up to 20Kb) were used for labeling. The labeled probes were resuspended in FISH Hybridization Buffer (2×SSC, 10% dextransulfat, 50% deionized formamide, 0.5mg/ml Salmon Sperm DNA) at a concentration of 15ng/μl, and 20μl of probe solution was used for each hybridization. The FISH was done according to the Bantignies et al. (2003). The eye or wing discs were dissected from 3rd instar larva in PBS buffer supplemented with 10% Fetal Bovine Serum (Invitrogen). Then the dissected tissue (from around 50 larva) was fixed in PBT (PBS 1×, 0.1% Tween 20), 4% freshly prepared para-formaldehyde solution for 20min at

room temperature. Then the tissue was washed twice with PBT buffer, and incubated with RNase A overnight at 4℃, then with PBS-Tr (PBS 1×, 0.3% Triton X-100) for 2hrs at room temperature. By passing through solutions containing different percentage of PBS-Tr and pre-Hybridization buffer (50% formamide, 4×SSC, 100mM NaH_2PO4, 0.1% Tween 20), the tissue was transfer into pre-Hybridization buffer, and incubate at 80℃ for 15min to denature the genomic DNA, then the denatured probes were added into the buffer, and incubate in a thermomixer at 37℃ overnight. After hybridization, the tissue was extensively washed and gradually transferred to PBS-Tr buffer, and incubate with Blocking solution at 4℃ overnight. Then anti-DIG-Rhodamine and anti-Biotin-FITC antibodies were added into the blocking buffer and incubated for 2 hours at room temperature. After staining, the tissue was washed by PBS-Tr and PBS-T each for three times, then counterstained with DAPI. After wash, and rinse with PBS, the tissue was dissected, and the small discs tissue was put on the slide with Vectashield antifade mounting media (Vector Laboratories), then covered with cover slip, and sealed with clear nail oil. The slides were keep at 4℃ in dark overnight to clear before acquisitions. The slides were imaged on a DeltaVision Image Restoration Microscope system (Applied Precision Instrument, LLC Issaquah, WA) with 100 ×/1.35 UplanApo objectives, deconvoluted and processed with the SoftWoRx software (Applied Precision Instruments), then counted as previously did in in-vivo live imaging.

2.3.5 3C analysis

3C experiments were done as described (Dekker et al. 2002; Hagège et al. 2007) with few modifications. Brain and attached imaginal discs were dissected from 30 third instar larvae in 1×PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4, pH 7.4) containing 10% fetal calf serum. The tissue was then fixed in 2% fresh paraformaldehyde/PBS for 10 min at room temperature. The cells were lysed in lysis buffer (10 mM Tris-HCl, 10 mM NaCl, 0.2% NP-40, pH 8.0, with Roche protease inhibitor cocktail freshly added) on ice for 10 min, followed by 20 strokes of a dounce homogenizer. The nuclei were recovered and washed with 1.2xNEB3 buffer (120 mM NaCl, 60 mM Tris-HCl, 12 mM $MgCl_2$, 1.2 mM Dithiothreitol, pH 7.9), then resuspended in 400µl 1.2×NEB3 buffer with 0.3% SDS. After shaking for 2hrs at 37˚C, Triton X-100 was added to 1.8%, and continue to shake for another 2 hrs at 37˚C. One-third of the nuclear solution (160µl, ~10 larvae) was used for digestion with *Eco*RI or *Hind*III (200 units at 37˚C overnight). Sodium dodecyl sulfate (SDS) was added to 1.5%, and the solution incubated at 65˚C for 25 min to inactivate the enzyme. 80 µl of 10×NEB ligation buffer was added and ddH2O to 950µl. 1% Triton X-100 was used to neutralize the SDS at 37˚C for 1 hr. The DNA was ligated with 8µl ligase (400U/µl, NEB) at 16˚C for 4.5 hrs, then 1 hr at room temperature. The 3C template DNA was then de-crosslinked overnight at 65˚C, and extracted with phenol-chlorophorm. The purified 3C DNA was further digested by a second restriction enzyme which cut outside of the religated region to linearise the circular

DNA, which will minimize potential PCR biases due to differences in template accessibility (Hagège et al. 2007).

3C primers were designed for the regions flanking the religated restriction sites, close to the insertion sites of transgenes. The primers were tested for specificity and efficiency. As a control for the crosslinking and ligation procedure we used Primers K1 (CACGGGAAAAACTACTGAAAG) and K2 (AAGCCGCAGGAGTTTCTAAC). These lie on adjacent *Eco*RI fragments in the *Brk* gene and point in the same direction, close to the *Eco*RI sites. Since the K1/K2 primer pair yield efficient PCR product with 3C templates, thus were chosen as internal positive controls. 3C PCR primers were designed to have the same annealing temperature at 55℃. The 3C template was titrated so that PCR product will be produced while still in linear PCR range. All PCR products were gel purified and sent for sequencing to confirm that they are the chimera molecules coming from two remote parts of the genome. PCR condition: 95℃ 8min, 36 cycles of 95℃ 30s, 55℃ 30s, 72℃ 30s, then 72℃ 10min. The other primers used for the detection of 3C interaction are listed in the tables 2.1:

Table 2.1. Primers used for 3C and ChIP experiments

Name	Sequence
1K	CACGGGAAAAACTACTGAAAG
2K	AAGCCGCAGGAGTTTCTAAC
1F	GCATGGCGGCATAATTTCTG
2R	AGCTCATTAGCCGTTAGTTTC
3F	CTCTCTTGGCCTCGATTAAAC
4R	CGCCGCACTTTTGGTCCAT
5R	TAATCCGCTTTACCCAGTAAG
6F	TGGCGGCAAAGACATTGATG
7F	TTCCCCCAACCATGCACAC
8F	CTGCCGAATCGGTTGAAAGG
9R	ACGTTCTCTAACACTGCAGC
10F	CAGGCATGCAAGCTAGCTTC
11R	ACCACCTCAGATACACCTTC

2.3.6 Cell culture

The *Schneider* cell line S2 were grown at 25℃ in Schneiders Drosophila medium (Invitrogen) supplemented with 10% Fetal Bovine Serum (Invitrogen), 100U/ml of Penicillin G and 100ug/ml of Streptomycin sulfate. The cells were harvested when confluent (~5×10^6 cells /mL).

2.3.7 ChIP analysis

ChIPs on S2 Cell line were done essentially as described in Schwartz et al.(Schwartz et al. 2006). Briefly, 2.5×10^8 cells were crosslinked by adding of 36% fomaldehyde (Sigma) directly into the growing cell culture to a final concentration of 1%, and incubate for 10min at room temperature while shaking gently. The reaction was stopped by adding the glycine to a final concentration of 0.125M. The crosslinked cells were pelleted, and washed once each with 1×PBS, and then ChIP wash buffer A (10mM HEPES pH7.6; 10mM EGTA pH8.0; 0.5mM EGTA pH8.0; 0.25% Triton X-100) for 10min, and then ChIP wash buffer B (10mM HEPES pH7.6; 200 mM NaCl; 1mM EDTA pH 8.0; 0.5mM EGTA pH 8.0; 0.01% Triton X-100) for 10min The washed fixed cells were then pelleted, and frozen in liquid nitrogen, and stored in -80℃. The fixed cells were resuspend into solication buffer (10mM HEPES pH7.6; 1mM EDTA pH8.0; 0.5mM EGTA) to a concentration of 1×10^8 cells/mL, and then subjected to sonication with Biorupter UCD-200TM-EX (Tosho Denki Co., LTD). The power was set “high”, each session with 0.5min on and 0.5min off for 5min long, repeat 5 sessions, and change the iced water in the chamber to prevent overheating

during each sessions. Then N-lauroylsarcosine was added to a final concentration of 0.5%. Proper amount of sonicated chromatin was then precleared by incubation with sepharose beads conjugated to protein A (Sigma), then incubated with proper amount of antibody overnight at 4℃. The Chromatin-antibody complexes were then precipitate by incubation with protein A sepharose beads (Sigma). The beads were washed extensively, 5 times with 1ml RIPA buffer (140mM NaCl; 10mM Tris-HCl pH8.0; 1mM EDTA; 1% Triton X-100; 0.1% SDS; 0.1% sodium deoxycholate, 1mM PMSF), and then 1 time with 1mL LiCl buffer (250mM CiCl; 10mM Tris-HCl pH8.0;, 1mM EDTA; 0.5% NP-40; 0.5% sodium deoxycolate), and then twice with 1mL TE buffer (10mM Tris-HCl pH8.0; 0.1mM EDTA). RNase A was added to a final concentration of 50ug/mL and incubate at 37℃ for 30min. To reverse crosslinking, SDS (final 0.5%) and Proteinase K (final 0.5mg/mL) were added and incubate overnight at 37℃, then transfer to 65℃ for 6hrs. The ChIP DNA was then phenol-chlorophorm purified, and washed and precipitated by ethanol. The immnunoprecipitated DNA was dissolved in 150ul water for following real-time PCR analysis.

ChIP on 3rd instar larva from Fly lines (F9, F9ΔP, M31, M31ΔP) were done essentially the same as above, except the first several steps of preparation of the cells. 300 3rd instar larva (~300mg) were collected, and rinsed in larva wash buffer (0.12M NaCl, 0.04% Triton X-100), then frozen in liquid nitrogen. The frozen larva were then transfer to molt, and grind with pestle, the resulting larva powder was then transfer

into Dounce homogenizer and dounce 10 strokes with 5mL of Crosslinking solution (1.8% fomaldehyde; 10mM HEPES pH7.9; 1mM EGTA; 100mM NaCl). The following neutralization with glycine and washes are the same as described.

All the antibodies used are rabbit polyclonal. The Anti-CP190 antibodies were raised against a peptide containing amino acids 606-742 of the CP190 protein fused to GST. The anti-CTCF was described by Gerasimova et al. (2007) and generously provide by V. Corces. Anti-Pc was described in Horard et al.(B. Horard et al. 2000). No antibody Control ChIPs were done essentially the same way except no antibodies were added in the reaction mixture.

Real-time PCRs were done to quantify the copies of DNA precipitated relative to the input DNA. 5ul of the ChIP DNA were used for each PCR reaction in a 20ul system, with 10ul of 2×Absolute Blue QPCR SYBR Green Mix (Thermo Scientific), 100nM of corresponding primers, 100nM of ROX as reference dye. PCR was performed in 96-well plates with the Mx-3000P machine (Stratagene). 5-point standard curve were used for each primer pairs by amplification of serial dilution of the input DNA isolated from an aliquot of lysate that did not undergo immnunoprecipitation. The specific primers were made to have a annealing temperature of 55℃, and listed in the table 2.2 :

Table 2.2. The PCR primes used for the ChIP experiments

Name	Sequence
Mcp+	ATAAGGGCTTTTCTGGGGAAG
Mcp-	TGTAAGGAGGAAGACTACATC
Fab7+	AGAGAGCGACTGCTTGAATG
Fab7-	GGGTAAGTAACGGTATTTAGG
W+	ATGCCACGACATCTGACC
W-	TGCCCAAGAAAGCTACCC
BP+	GCCATAACGGCAGAACCAAAG
BP-	ATGAGGCCATCTCAGTCGC
Ubx+54	CCGCTGATAATGTGGATAA
Ubx-177	CACCCCGATAAACTTAAAC
CG+	CGTCTAGTGGTTGATTCCAT
CG-	CAGGACCAAAAGTTTAGTGG
FM+	AGCAATTTGTCACCGCAAGG
FM-	GGATTTTGAGTGCGTTCTTCC
M0+	TAGGAACTTCGGAATAGGAAC
M0-	AAGAAGAAGAGGCGAGACAG
M1+	CAATGATCCCACGAGAGATC
M1-	GCGCAAAAGGTTTGGATATTG
M2+	TAGGAACTTCGGAATAGGAAC
M2-	TGATACTTCAAATACCCTTGG
F0+	TAGGAACTTCGGAATAGGAAC
F0-	AAGAAGAAGAGGCGAGACAG
F1+	CAATGATCCCACGAGAGATC
F1-	TCGCTCACTTGGCAACAAAG
F2+	TAGGAACTTCGGAATAGGAAC
F2-	TGATACTTCAAATACCCTTGG

2.4 Results

2.4.1 Comparison of the *bxd* PRE and the *Mcp* PRE (by Martin Müller and Ilham Bahechar)

We begin by comparing the properties of the *bxd* PRE and of the *Mcp* PRE in the same genomic context. For this purpose, we assembled a construct called Flipper$^{Mcp\text{-}bxd}$ in which the 661 bp *bxd* PRE and different fragments from the *Mcp* region are flanked by FRT and LoxP sites, respectively, in a vector containing the two marker genes *yellow* and *mini-white* (Figure 2.1). Hence, for a given Flipper$^{Mcp\text{-}bxd}$ insert, *bxd* and *Mcp* can be individually deleted in situ with the Flp and Cre recombinases, to assess the role of each PRE. The results obtained with the different Mcp fragments show that all the functions are fully contained in the 800bp core fragment. For brevity, we will describe in detail only the experiments done with Flipper23$^{Mcp\text{-}bxd}$. The results for 13 independent lines with inserts on chromosome 3 are summarized in Table 2.3. Based on their *mini-white* phenotypes (see Table 2.3, columns 2 and 3), they could be classified into two groups:

- 2 lines show pairing-dependent silencing of *mini-white*: eye pigmentation is weaker in homozygotes than in heterozygous siblings.
- 11 lines have no eye pigmentation at all in hetero- as well as in homozygous condition. They could only be isolated thanks to their *yellow*[+] phenotype, which is either uniform or variegated.

These results show that, together, the two PREs act as a powerful silencer of the *mini-white* reporter. The *yellow* reporter is less sensitive to PRE silencing, as has been previously observed for other constructs. To assess the contribution of each of the two PREs to *mini-white* silencing, we deleted the *bxd* PRE (yielding Flipper23^{Mcp}) or the *Mcp* PRE (yielding Flipper23^{bxd}). In the Flipper23^{bxd} flies, more than half of the lines (7/12) remain white-eyed while the others become white when homozygous (Table 2.3, columns 4 and 5). None of the Flipper23^{Mcp} derivatives are white-eyed either when hetero- or homozygous. In all cases, loss of the *bxd* PRE greatly decreases the pairing-dependent effects (Table 2.3, columns 8 and 9). These results indicate that, at all insertion sites tested, the *bxd* PRE is a more potent silencer than the *Mcp* core fragment.

To ask whether Flipper23^{bxd} can interact in trans with Flipper23^{Mcp} on the paired homologue, the *Mcp* and *bxd* derivatives were tested in trans to each other and the resulting eye colour was compared to that of Flipper23^{Mcp}/+ and Flipper23^{bxd}/+ flies (see Table 2.3, compare column 11 with columns 5 and 8). If the two constructs act independently, we expect that, for a given insertion site, the eye colours of the two inserts would be approximately additive. The results (Table 2.3, column 11), show that in all tested cases the eye colour of the combination is either white or weaker than that of either heterozygous insert separately, consistent with the conclusion that, at a homologous position, *Mcp* and *bxd* PREs interact efficiently with each other.

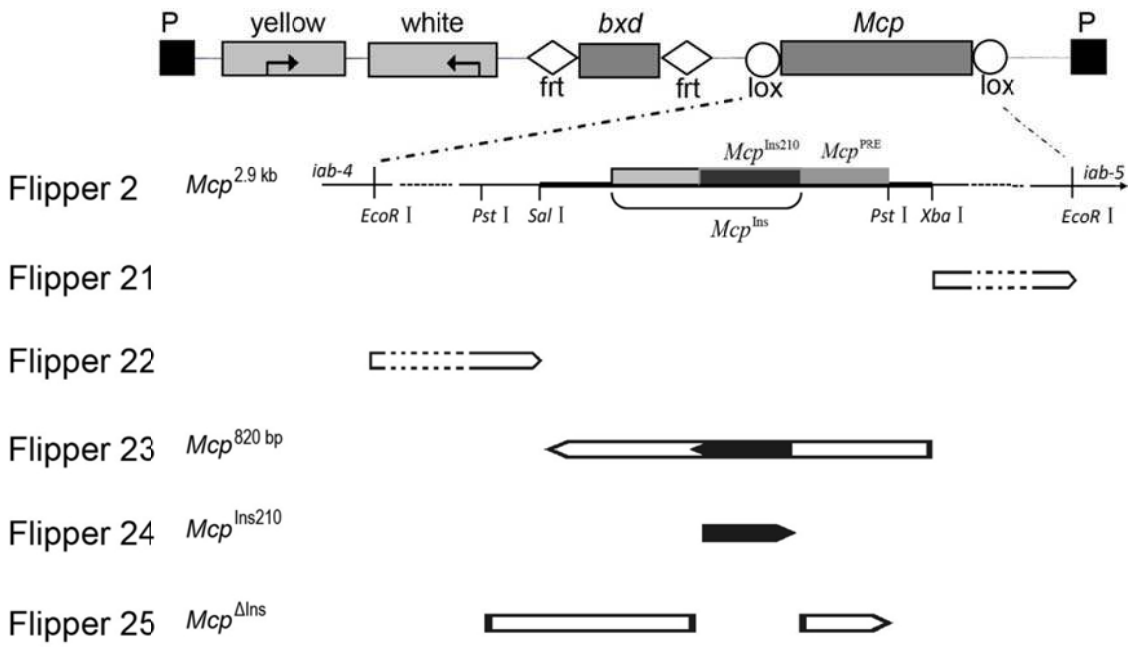


Figure 2.1 Maps of the Flipper2 construct and derivatives.

The transposon construct contains the intronless *yellow* gene with wing and body enhancers and the *mini-white* gene as markers and reporters. The bxd PRE 661 bp fragment is flanked by LoxP sites and the 2.9 kb *Eco*RI *Mcp* fragment is flanked by FRT sites to allow independent excision of either element. The Flipper21 and 22 constructs contain the flanking regions and Flipper 23 contains the core 800bp *Sal*I-*Xba*I *Mcp* fragment that includes both the PRE and the insulator activities. The orientation of the *Mcp* fragments is such that the side nearest the *bxd* PRE fragment is the side normally adjoining the *iab-4* region in the Bithorax Complex. In the Flipper24 and 25 constructs, the *white* eye enhancer (WE) has been added. Mcp^{Ins210} is the minimal *Mcp* insulator fragment and Mcp^{ins} is the 755 bp *Pst*I-*Pst*I *Mcp* core region from which the 210 bp insulator has been deleted. The various *Mcp* sequences used in the constructs have been shown.

2.4.2 Long-distance interactions (by Martin Müller and Ilham Bahechar)

All Flipper23$^{Mcp\text{-}bxd}$ inserts presented in Table 1 can participate in long-distance interactions with a panel of 12 *Mcp*-containing insertions on chromosome 3 previously reported to be good partners for long-distance interaction (Muller et al. 1999). Each Flipper23$^{Mcp\text{-}bxd}$ insert shows interactions with at least 3 of the 12 tester lines (Table 2.3, column 4). However, when we compared the long-distance interactions of the Flipper23^{Mcp} and Flipper23^{bxd} derivatives, the Flipper23^{bxd} derivatives had generally lost the ability to interact in trans with distant *Mcp* testers: only one cross of 40 tested gave an interaction (Table 2.3, column 7). In contrast, only two of 11 Flipper23^{Mcp} lines failed to trans-interact with the *Mcp* testers used (Table 2.3, compare column 10 to column 4). These results confirm that, while the *bxd* PRE is a powerful silencer, it lacks an activity responsible for mediating long-distance trans-interactions. This activity is present in the 800 bp *Mcp* fragment, even though this element contains a much weaker silencing activity. We conclude that the ability of *Mcp* to interact with other *Mcp* constructs inserted at remote genomic sites is not due to the PRE itself but to an associated function present in the *Mcp* fragment but absent in the *bxd* PRE.

Table 2.3. Flipper 23 phenotypes and interactions

Line	Flipper23$^{Mcp-bxd}$			Flipper23^{bxd}			Flipper23^{Mcp}			$\frac{Flipper23^{bxd}}{Flipper23^{Mcp}}$
	P/+ Eye color	P/P Eye color	Interactions with Mcp testers	P/+ Eye color	P/P Eye color	Interactions with Mcp testers	P/+ Eye color	P/P Eye color	Interactions with Mcp testers	
84.7.3	white	white	3/12	pale yellow	white	0/3	orange	dark orange	1/3	white
84.14.1	white	white	4/12	white	white	1/3	orange	dark orange	2/3	white
84.32.3	white	white	4/12	white	white	0/2	orange	dark orange	1/2	white
84.63.1	white	white	4/12	white	white	0/2	orange	dark orange	2/2	white
84.77.1	white	white	6/12	pale yellow	white	0/4	not obtained		nd	nd
84.80.5	pale yellow	white	4/12	not done			not done		nd	nd
84.81.1	white	lethal	6/12	white	lethal	0/2	orange	lethal	0/2	lethal
84.95.2	white	lethal	6/10	white	lethal	0/4	orange	lethal	1/4	lethal
84.95.5	pale yellow	white	4/10	pale yellow	white	0/3	dark orange	red	1/3	white
84.98.2	white	white	6/11	pale yellow	white	0/4	orange	dark orange	0/4	white
84.110.5	white	white	5/12	white	lethal	0/5	orange	dark orange	1/5	lethal
84.131.4	white	white	5/12	pale yellow	white	0/5	orange	dark orange	2/5	faint yellow
84.146.1	white	white	6/12	white	white	0/3	weak yellow	dark orange	2/3	white
Column 1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**

Column 1 lists 13 insertions on chromosome 3 with the eye color of heterozygotes and homozygotes in columns 2 and 3. Column 4 lists the number of *Mcp* tester lines with which a given Flipper23 line interacts phenotypically. Columns 4, 6 and 7 indicate the corresponding phenotypes and interactions for the Flipper23 line with the *Mcp* fragment

Table legend of Table 2.3 continued.

deleted. Columns 8, 9 and 10 list the same features for the Flipper23 lines with the *bxd* PRE deleted. Column 11, the last on the right, lists the eye color of the Flipper23^{bxd} version allelic with the Flipper23^{Mcp} version.

2.4.3 The insulator/boundary component of *Mcp* is the trans-interacting element (by Olga Kyrchanova and Pavel Georgiev)

The functional difference between *Mcp* and *bxd* PRE is also illustrated by the fact that trans-interactions have never been observed among constructs containing the latter even when large fragments of 1.5 to 6 kb were used (V.P. unpublished observations). However, the *bxd* PRE was notably able to enter into trans-interactions when a *gypsy* Su(Hw) insulator was incorporated into the construct (Sigrist & Pirrotta, 1997). We reasoned therefore that the *Mcp* fragment might contain a similar insulator activity responsible for the trans-interactions. Although *Mcp* does not bind SU(HW), the core region has been shown to contain an insulator activity that can be separated from the PRE activity (Gruzdeva et al., 2005).

To test the role of the *Mcp* insulator, we made two new constructs (Figure 2.1). In the Flipper24$^{Mcp\ Ins210\text{-}bxd}$ construct, the *bxd* PRE, flanked by FRT sites, was inserted next to the 210 bp core insulator from the *Mcp* element (Kyrchanova et al., 2007), flanked by LoxP sites. We obtained 10 Flipper24$^{Mcp\ Ins210\text{-}bxd}$ lines on the third chromosome and mapped their insertion sites. In these lines the *bxd* PRE repressed *white* expression, as shown by the fact that its deletion darkened the eye colour of heterozygous flies and pairing-dependent silencing was lost in homozygous flies (Table 2.4, columns 2, 3). Deletion of *Mcp*^{Ins210} (M^{210}) had no effect on the eye colour.

To test trans-interaction we again used the panel of 10 *Mcp* tester lines on the third chromosome[4] and crossed them with ten Flipper24$^{Mcp\ Ins210\text{-}bxd}$ candidate lines in all possible pairwise combinations. Eight out of 10 tested Flipper24$^{Mcp\ Ins210\text{-}bxd}$ lines displayed trans-silencing interactions with two to four of the *Mcp* tester lines (Table 2.4, columns 4, 5). In general, long-distance interactions were more likely to be observed when the insertions were closer to one another but in some cases they were found between inserts located on opposite arms of the third chromosome (66D12 line and 99B tester line). All trans-interactions were lost after excision of the *Mcp*^{Ins210} insulator.

We also constructed a Flipper25$^{Mcp\ \Delta\text{Ins-}bxd}$ transgene containing the 660 bp *bxd* PRE as before and an *Mcp* element that includes the *Mcp* PRE but from which the 210 bp insulator fragment is deleted (Figure 2.1). We obtained nine transgenic lines on the third chromosome, in which the PRE repressed *white* expression (Table 2.4, columns 7, 8). In all possible pairwise combinations with the *Mcp* tester panel, we observed no trans-interactions (Table 2.4, columns 9, 10). Deletion of the insulator-less *Mcp*$^{\Delta Ins}$ had no effect. These results suggest that the *Mcp* insulator is essential for trans-interactions while neither the *Mcp* PRE nor the *bxd* PRE can mediate such interactions.

Table 2.4. Flipper24 and 25 phenotypes and interactions.

Flipper24 *bxd* PRE Mcp^{210}	Eye colour		Interaction with *Mcp* testers			Flipper25 *bxd* PRE $Mcp^{\Delta 210}$	Eye colour		Interaction with *Mcp* testers	
	P/+	P/P	+/-	+			P/+	P/P	+/-	+
100E3	Or	pY	-	3/10		83B4	Y^{var}	L	-	-
100E3-ΔM^{210}	pY^{var}	W	-	-		83B4-$\Delta M^{\Delta 210}$	Or	L	-	-
69A	pY^{var}	W	3/10	1/10		70F4	Y	L	-	-
69A-ΔM^{210}	pY^{var}	W	-	-		70F4-$\Delta M^{\Delta 210}$	Y	L	-	-
95F4	W	W	-	-		*100E3*	Y^{var}	W	-	-
95F4-ΔM^{210}	W	W	-	-		100E3-$\Delta M^{\Delta 210}$	Y^{var}	W	-	-
82A3	pY	W	2/10	1/10		*97E5*	W	W	-	-
82A3-ΔM^{210}	pY	W	-	-		97E5-$\Delta M^{\Delta 210}$	W	W	-	-
66D12	Y^{var}	W	1/10	2/10		72D10	pY^{var}	W	-	-
66D12-ΔM^{210}	pY^{var}	W	-	-		72D10-$\Delta M^{\Delta 210}$	Y	W	-	-
70A2	Y^{var}	W	-	-		*100E3-1*	pY^{var}	W	-	-
70A2-ΔM^{210}	pY^{var}	W	-	-		100E3-1-$\Delta M^{\Delta 210}$	Y	W	-	-
94E8	W	W	-	3/10		89B3	W^{var}	L	-	-
94E8-ΔM^{210}	W	W	-	-		89B3-$\Delta M^{\Delta 210}$	Y	L	-	-
85B7	Y^{var}	W	-	2/10		*92F12*	Y	L	-	-
85B7-ΔM^{210}	Y^{var}	W	-	-		92F12-$\Delta M^{\Delta 210}$	Y	L	-	-
65D5	pY^{var}	L	-	3/10		86A4	dY	W	-	-
65D5-ΔM^{210}	pY^{var}	L	-	-		86A4-Δ $M^{\Delta 210}$	Y	W	-	-
88F3	W	W	-	3/10						
88F3-ΔM^{210}	W	W	-	-						
Column 1	**2**	**3**	**4**	**5**		**6**	**7**	**8**	**9**	**10**

On the left, column 1 lists the Mcp^{Ins210} lines by their cytological insertion site. Each line is followed by its derivative in which the Mcp^{Ins210} fragment has been excised (ΔM^{Ins210}). Columns 2 and 3 indicate the eye colors of the heterozygote (P/+) and homozygote (P/P). The color gradation from white to red is indicated by W, pY, Y, dY, Or, dOr, Br, BrR, R (Gruzdeva et al. 2005). The subsequent columns indicate the number of *Mcp* tester lines with which a given line displays trans-interactions out of 10 tester lines used (Muller et al. 1999). Interactions are scored as "weak" (+/-) when the eye color of the rans-heterozygous combination was the same as of one of the heterozygous parents (if both had colored eyes)

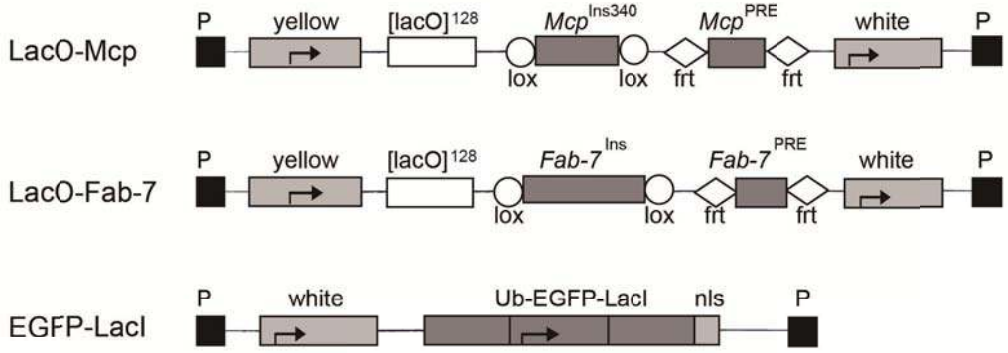
Table legend of Table 2.4 continued.

or was lighter than that of the heterozygous parents (if one had white eyes). Interactions are scored as "strong" (+) when the eye color of the trans-heterozygous combination was lighter than one or both of the heterozygous parents with colored eyes. No interactions were observed in crosses between any combination of tester lines and Mcp^{Ins} lines (on the right), in which the minimal 210 bp insulator is deleted from the core *Mcp* fragment (See appendix 1 for details).

2.4.4 Co-localization of *Mcp* insulator constructs

The ability of two remote constructs to affect one another's expression strongly suggests that they are able to make contact in the nucleus. To demonstrate this physical interaction, we used two methods. In one, co-localization is visualized by in vivo imaging of fluorescence-tagged loci (Vazquez et al., 2006). For this purpose we made constructs containing a 340 bp *Mcp* insulator fragment, flanked by Lox sites, and a 138 bp *Mcp* PRE fragment, flanked by FRTs (Figure 2.2a; appendix 1). We used the 340 bp insulator fragment because it has a stronger insulator activity than the 210 bp fragment used in the preceding section (Gruzdeva et al., 2007). To visualize the insertion site, the transposon contained 128 copies of the Lac operator (*LacO*) and the flies were crossed with a line expressing the LacI repressor fused to EGFP fluorescent protein, driven by the ubiquitin promoter. Three *LacO-Mcp* lines were obtained (Figure 2.2b) and were crossed to test all pairwise combinations.

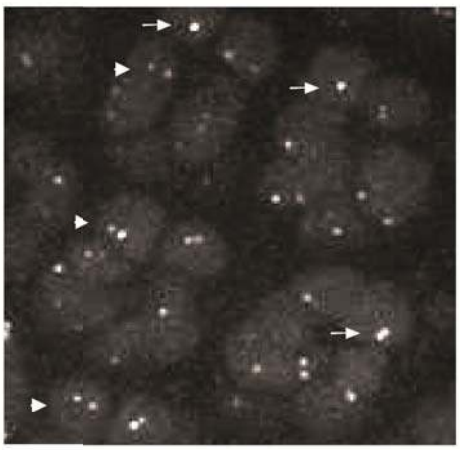
A



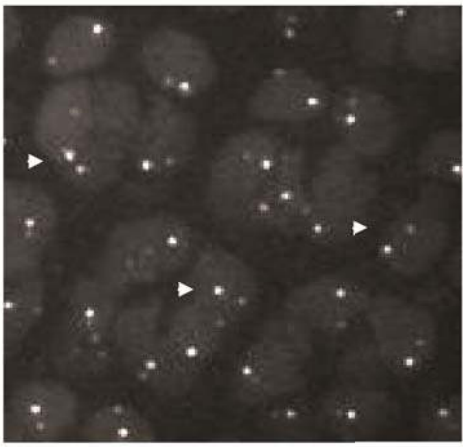
B

Line	Insertion Locus
LacO-Mcp M16	X: 2,195,155
LacO-Mcp M25	X: 21,023,193
LacO-Mcp M31	3R: 27,769,892
LacO-Fab-7 F12	X: 13,022,784
LacO-Fab-7 F4	2L: 12,822,843
LacO-Fab-7 F9	3R: 6,158,884

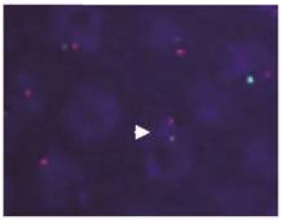
C



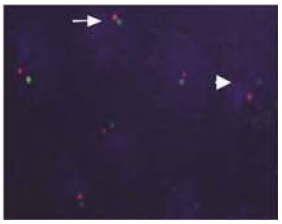
D



E



F



G

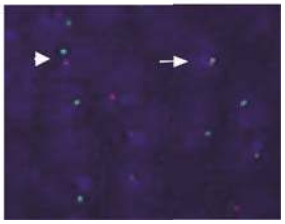


Figure 2.2 Fluorescent tagging of *Mcp* and *Fab-7* components.

Figure legend of Figure 2.3 continued.

A) Structure of the reporter constructs. The *Mcp* insulator fragment is flanked by LoxP sites and the *Mcp* PRE is flanked by FRT sites. A parallel construct contains a similar arrangement of the *Fab-7* insulator and the *Fab-7* PRE. The constructs utilize the yellow and mini-white genes as markers. The tandem array of 128 *lacO* sequences is used to bind the *lacI* repressor fused to EGFP fluorescent protein expressed from a different construct driven by the ubiquitin promoter. B) The insertion sites of the transgenes determined by inverse PCR. C) Image of eye imaginal disc nuclei showing a typical frequency of one-dot (arrows) and two-dot (arrowheads) nuclei from F9ΔP-M31ΔP line. D) Image of eye membrane cell nuclei showing the two-dot nuclei from F9ΔI-M31ΔI line. E-G) Detection of interaction between transgenes and their endogenous partners by FISH. The nuclei were labeled with Dapi (Blue), the BAC clone containing the BX-C sequence was labeled with Biotin and detected with a-biotin-FITC (Green). E) Typical images of the yw67 fly eye discs, with the probes (Red) close to the insertion sites of F9 transgene. F) Typical images of F9 transgenic fly, with the LacO sequences labeled by Digoxigenin (Red). G) Images of M31 transgenic fly eye membrane cells, with the LacO sequences labeled by Digoxigenin (Red). One-dot cells labeled with arrows.

Fluorescence image stacks of eye or wing discs of larvae carrying two transposon insertions were obtained to assemble a three-dimensional representation of the nuclei and each nucleus was scored as “one dot” (co-localization) or “two dot” (no co-localization) when the two signals were non-overlapping (Figure 2.2c and 2.2d). The results show a rather low frequency of co-localization ranging from 6% to 8% of the nuclei (Figure 3a; Appendix Table 2). However, when the insulator element was excised, co-localization dropped to 0.1-0.2% and Chi square tests indicate a $P<0.0001$ (Appendix Tables 2-4). Deletion of the insulator from only one of the two transgenes caused the same drop in co-localization as the deletion from both transgenes (data not shown). In contrast, excision of the PRE fragment from one or both transgenes had no appreciable effect on the frequency of co-localization.

A parallel construct was made using fragments containing the insulator and PRE portions of the *Fab-7* element (Figure 2.2a; appendix 1). Co-localization between different insertions of this construct was observed at a frequency similar to that seen with the *Mcp* construct (Figure 2.3b; Appendix Table 3). In this case also, excision of the insulator caused loss of the co-localization while excision of the PRE fragment had no effect. Bantignies et al. (Bantignies et al., 2003) reported that transgenic *Fab-7* elements interacted with the endogenous *Fab-7* element and two transgenic *Fab-7*s on different chromosomes could also interact with each other without endogenous *Fab-7*. To ask if the co-localization that we detect between *Fab-7* transgenes also do not depend on the endogenous element, we tested two pairs of inserts F4ΔP; F9ΔP and F12ΔP; F9ΔP in a genetic background homozygous for the $Fab\text{-}7^{1}$ deletion of the endogenous element. No significant difference was observed

in the frequency of co-localization. Thus, although, as shown by the 3C experiments below, the transgenes do interact with the endogenous copy, this is not a prerequisite for interaction between transgenic copies.

We also tested whether *Mcp* inserts could interact with *Fab-7* inserts on the same chromosome. The results show that the two inserts do co-localize at a frequency of 5.15%, somewhat lower than that seen between *Mcp* inserts (Figure 2.3c; Appendix Table 4). Deletion of the PRE elements from both constructs gives a slight decrease in the incidence of co-localization to 4.34% but deletion of the insulator elements reduces it to 0.29%. These results show that a degree of interaction can be observed between *Mcp* and *Fab-7* elements.

A few words need to be said about the changes in the expression of the mini-*white* gene as the PRE, insulator, or both are excised from these constructs. As might be expected, excision of the PRE results in higher levels of eye pigmentation: although the PRE element responds to both PcG repressive effects and Trithorax stimulation, the overall effect is an increase in *mini-white* gene expression. More surprising is the fact that excision of the insulator component results in a strong decrease in eye pigmentation in the three *Fab-7* lines tested and in one of the three *Mcp* lines (of the other two, one is completely white to begin with and the third is very weakly pigmented). Deletion of both PRE and insulator elements has little further effect (Appendix Table 5). While these results do not allow a definite conclusion, they raise the possibility that the insulator component may contain an enhancer

or that, by mediating long-range interactions with other insulator elements accompanied by an enhancer, it may stimulate expression.

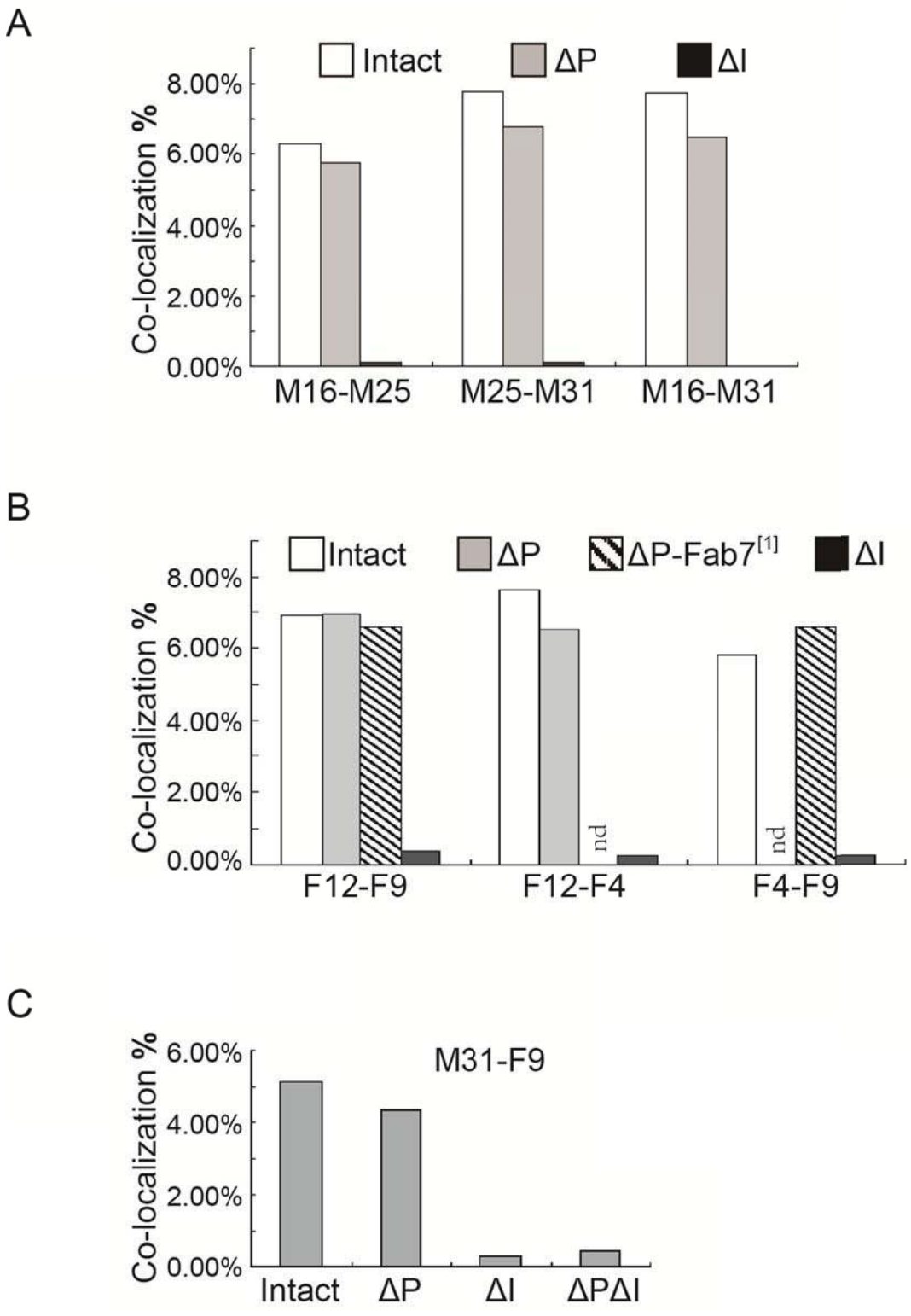


Figure 2.3 Frequency of co-localization of *Mcp* and *Fab-7* transgenes.

a) Interactions between *Mcp* transgenes. For each of the pairwise combinations between three lines bearing the *Mcp* construct, the histogram shows frequencies of

Figure legend of Figure 2.3 continued.

co-localization (one dot nuclei) for the starting lines (white), for the lines after PRE deletion (gray) or for the lines after insulator deletion (black). b) Interactions between *Fab-7* transgenes. The histogram shows the frequencies of co-localization for the pairwise combinations between three lines bearing the *Fab-7* construct. In both sets of experiments, the frequency of co-localization drops more than 30-fold when the insulator is deleted but is not affected by deletion of the PRE (see tables 3-5 for details). Some combinations were tested in a background deleted for the endogenous *Fab-7* element (*Fab-7*^{1}), showing that it is not required for the trans-interactions. c) Interaction between *Mcp* and *Fab-7* trangenes. The histograms show the frequencies of co-localization between the F9 and the M31 transgenes before or after deletion of the PRE or insulator segments. Numerical data are given in Appendix Tables 2-4.

2.4.5 Trans-interactions are detected by FISH and 3C

To ask if the transgenes also interact with their endogenous partners, we employed the FISH techniques to detect the physical interactions. As shown in figure 2.2 E-G, the transgene F9 interact with endogenous Fab-7 at 7.7% of the nuclei (data not shown), while the wild type yw67 fly have no interaction.

However, since the FISH technique do not always yield clear images for counting, and is hard to apply on larva discs, we turn to the 3C method for physical association between transgenes to systematically investigate and confirm the in-vivo imaging results (Dekker et al. 2002). We first tested if any one insertion of the *Mcp* construct was able to interact with the endogenous *Mcp* element that resides in the Bithorax Complex. As shown in Figure 2.4b and 2.5c, the results were unambiguous: for all three insertion sites tested, there was no interaction in the absence of the inserted construct, but insertion of the *Mcp* construct produced interaction with the endogenous *Mcp* dependent on the insulator component and not on the PRE component. Entirely similar results were obtained with the *Fab-7* lines (Figure 2.4c).

We then tested the interaction between remote insertions of the *Mcp* or of the *Fab-7* construct. Clear interactions were observed for two of the pairs tested, producing the predicted PCR product (Figure 2.4c and 2.5b). The identity of the product was confirmed by excising the band from the gel and sequencing the DNA fragment. Several other pairs did not give a detectable 3C PCR band. While this may be partly dependent on the choice of

primers to detect the 3C interaction, it is likely that the interaction with the endogenous element is stronger than interactions between insertions.

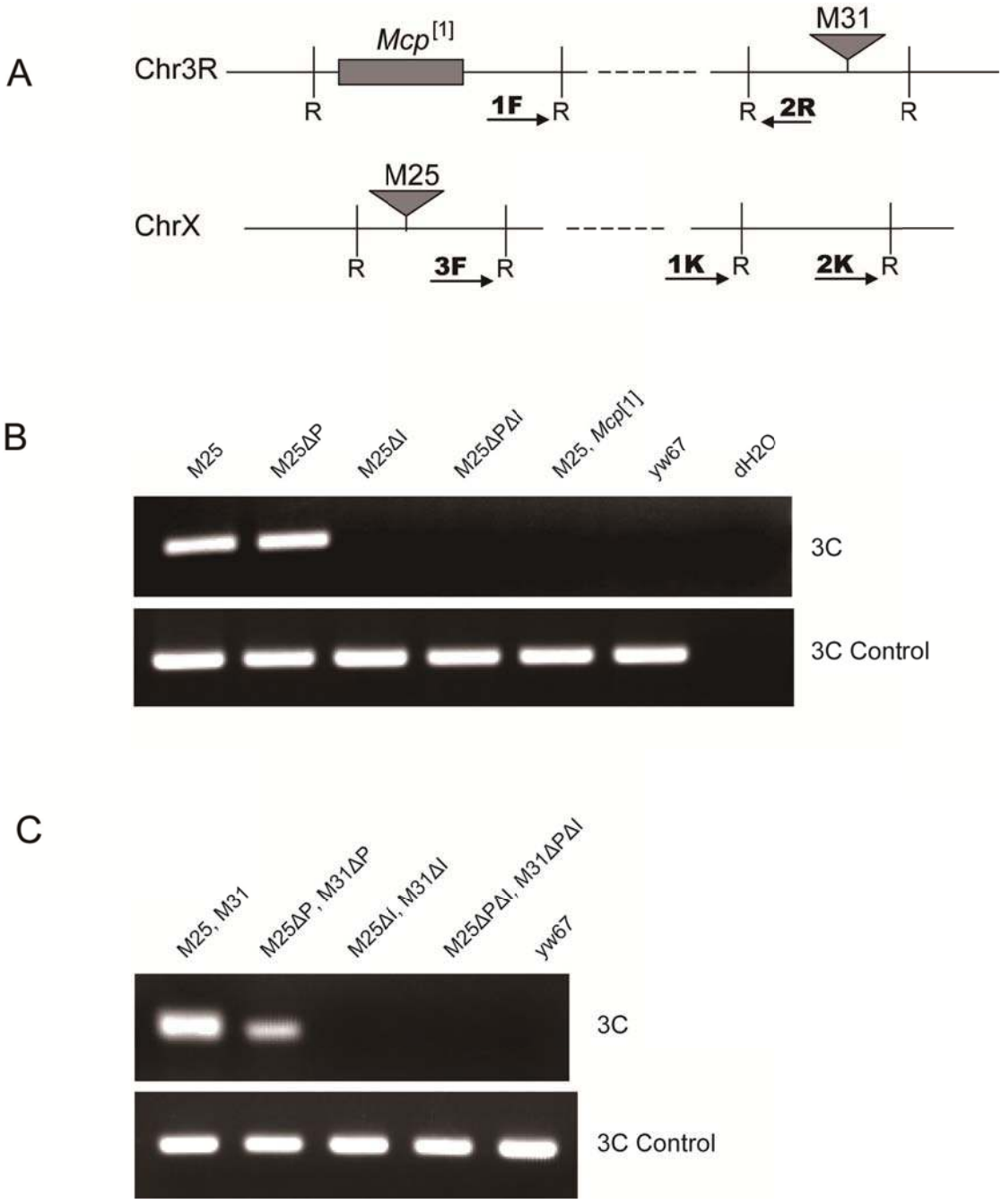


Figure 2.4 3C assay of trans-interactions for LacO-Mcp lines.

a) Maps of 3C primers used. Arrows 1F, 2R, and 3F denote the primer directions; R denotes the *Eco*R1 restriction sites. The map also shows the approximate extent of the *Mcp*[1] deletion of the endogenous *Mcp*, and transgenes insertion sites. b) 3C Assay between the M25 insertion and the endogenous *Mcp*. Upper lanes: 3C primer pair

Figure legend of Figure 2.4 continued.

1F/3F; Lower lanes: 3C control primers K1/K2. The 3C DNA from fly lines used for the assay is labeled above each lane. c) 3C assay between two transgene M25 and M31 insertion sites using primer pair 2R/3F. The fly lines used are labeled above each lane. The lowers lanes are the 3C controls using the primers K1/K2.

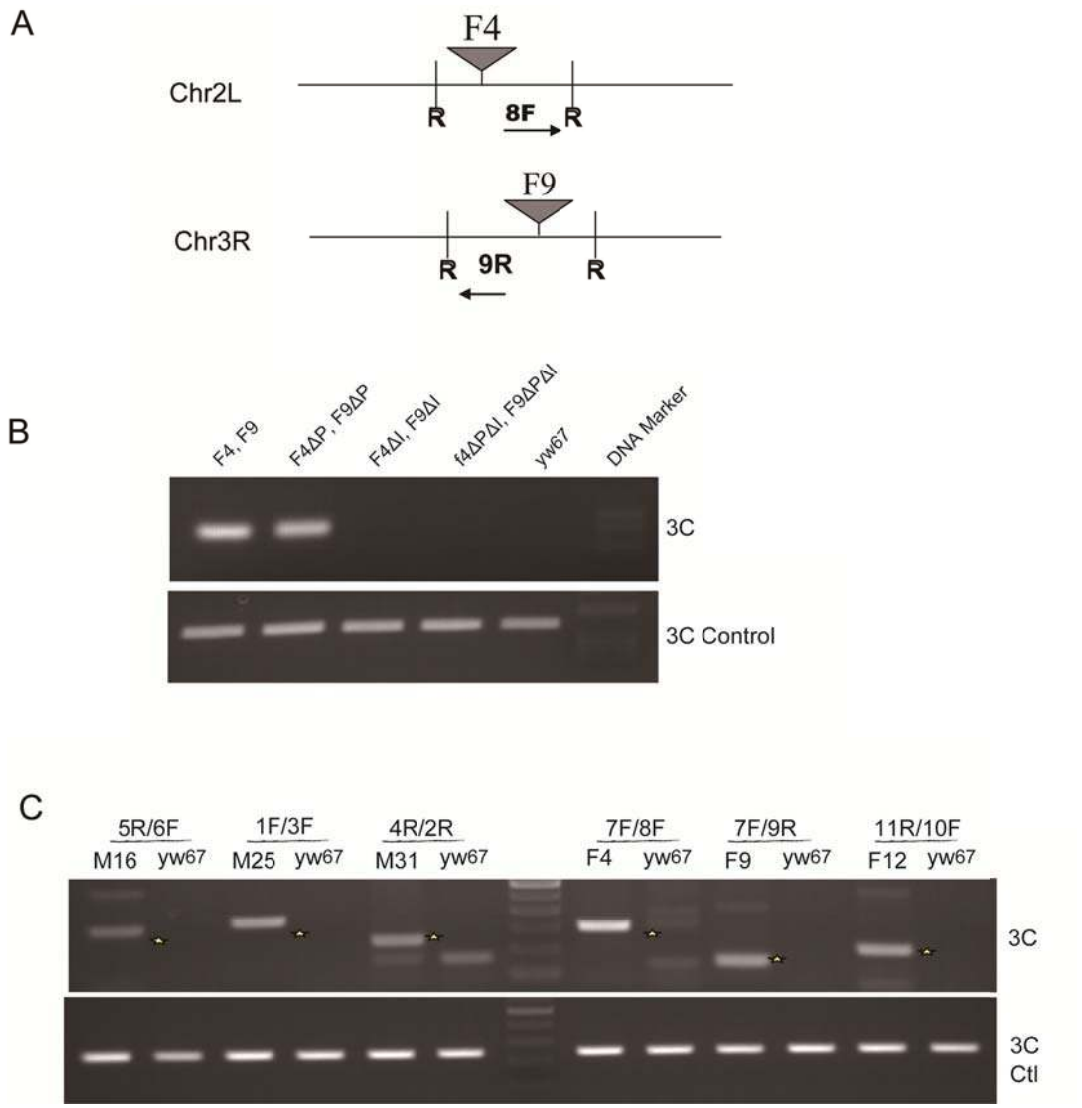


Figure 2.5 3C assay of trans-interactions for LacO-Fab-7 lines.

a) Maps of 3C primers used. Arrows 8F and 9R denote the primer directions; R denotes the *Eco*R1 restriction sites. b) 3C Assay between two transgenes the F4 insertion and the F9 insertion, and their derivatives. Upper lanes: 3C primer pair 8F/9R; Lower lanes: 3C control primers K1/K2. The 3C DNA from fly lines used for the assay is labeled above each lane. c) 3C assay between the transgene insertion sites and their endogenous partners. The primers and the 3C DNA used for each assay are labeled above each lane. The yw67 host lines are served as negative control. The

Figure legend of Figure 2.5 continued.

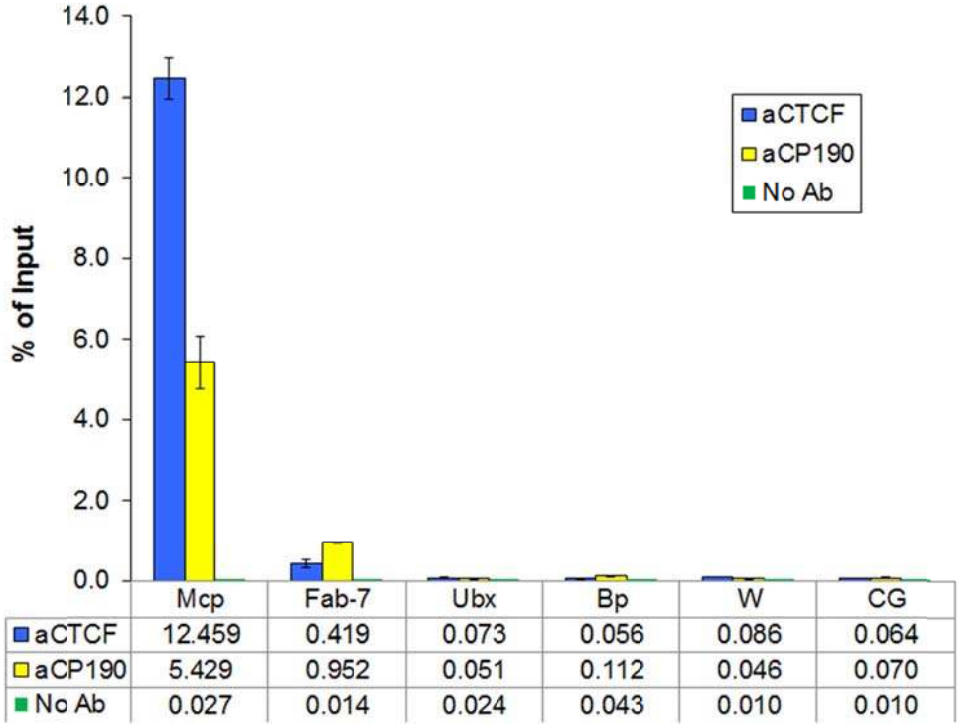
asterisks indicate the band of the expected molecular weight. The lowers lanes are the K1/K2 internal controls for corresponding 3C samples.

2.4.6 Binding of insulator proteins

If the insulators of *Mcp* and *Fab-7* are responsible for the trans-interactions, the ability of these two elements to interact with one another presupposes that their insulators share some common component. The binding of some insulator proteins to the *Mcp* and *Fab-7* regions has been reported (Bushey et al., 2009; Holohan et al., 2007; Mohan et al., 2007; Nègre et al., 2010) and more recently genome-wide ChIP/chip analysis of SU(HW), CTCF, MOD(MDG4), CP190, BEAF and ZW5 has been carried out by the modENCODE Drosophila Chromatin Consortium (2009). These results generally agree in detecting CTCF and CP190 at *Mcp* but only CP190 at *Fab-7*. We repeated this analysis using quantitative real-time PCR and concluded that a low but significant presence of CTCF is in fact detectable also at *Fab-7* (Figure 2.6), as has also been reported by Holohan et al. (2007). Examination of the *Mcp* and *Fab-7* sequences reveals that the *Mcp* insulator region contains two CTCF binding consensus sequences, one excellent and one moderately good. No recognizable CTCF consensus can be found in the *Mcp* PRE portion. No consensus sequence for any of the known insulator binding proteins is present in the *Fab-7* region. The GAGAG consensus for the binding of GAF, BTB/POZ domain protein like CP190, is present in both *Mcp* (once) and *Fab-7* (twice) PREs as well as in the *Fab-7* insulator (6 times) but not in the *Mcp* insulator fragment.

To exclude the possibility that the PcG proteins may mediate the long-distance interaction, we perform the ChIP experiment s with a-Pc antibodies on the 3rd instar larva from fly lines M31 and M31ΔP. As shown in figure 2.7a, there is a low but significant amount of Pc

binding on the PRE portion of the *Mcp* in fly line M31, but no presence of Pc after excision of PRE part in the M31ΔP line. Similar results were obtained with F9 and F9ΔP fly larva regarding the *Fab-7* fragment (fig. 2.7b).

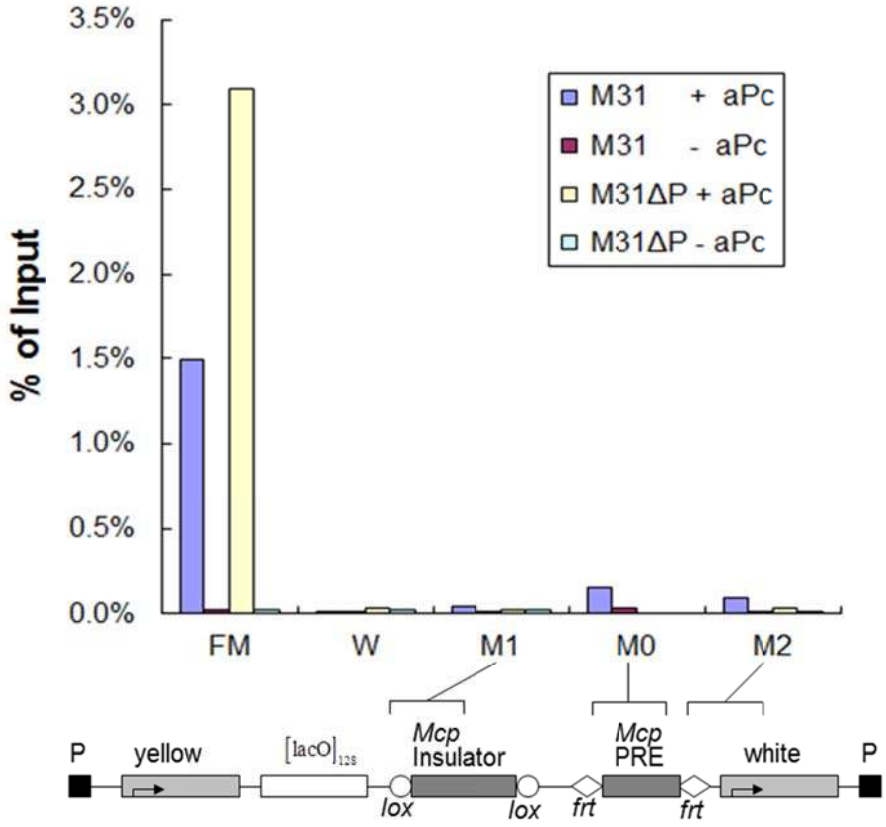


	Mcp	Fab-7	Ubx	Bp	W	CG
aCTCF	12.459	0.419	0.073	0.056	0.086	0.064
aCP190	5.429	0.952	0.051	0.112	0.046	0.070
No Ab	0.027	0.014	0.024	0.043	0.010	0.010

Figure 2.6 ChIP binding assay for CTCF and CP190 at *Mcp* and *Fab-7*.

Quantitative PCR evaluation of ChIP results are expressed as per cent of input DNA. The primers used are listed on Table 2.2. For *Mcp* and *Fab-7*, they represent the insulator regions. The remaining sites are negative controls. *Ubx* and *w* represent the promoters of the corresponding genes. The bxd PRE and gene CG5270 were also tested, the latter chosen as a particularly active gene in the eye disc.

A


3.5%
3.0%
2.5%
2.0%
1.5%
1.0%
0.5%
0.0%
% of Input
M31 + aPc
M31 - aPc
M31ΔP + aPc
M31ΔP - aPc
FM
W
M1
M0
M2
Mcp Insulator
Mcp PRE
P
yellow
[lacO]128
white
P
lox
lox
frt
frt

B

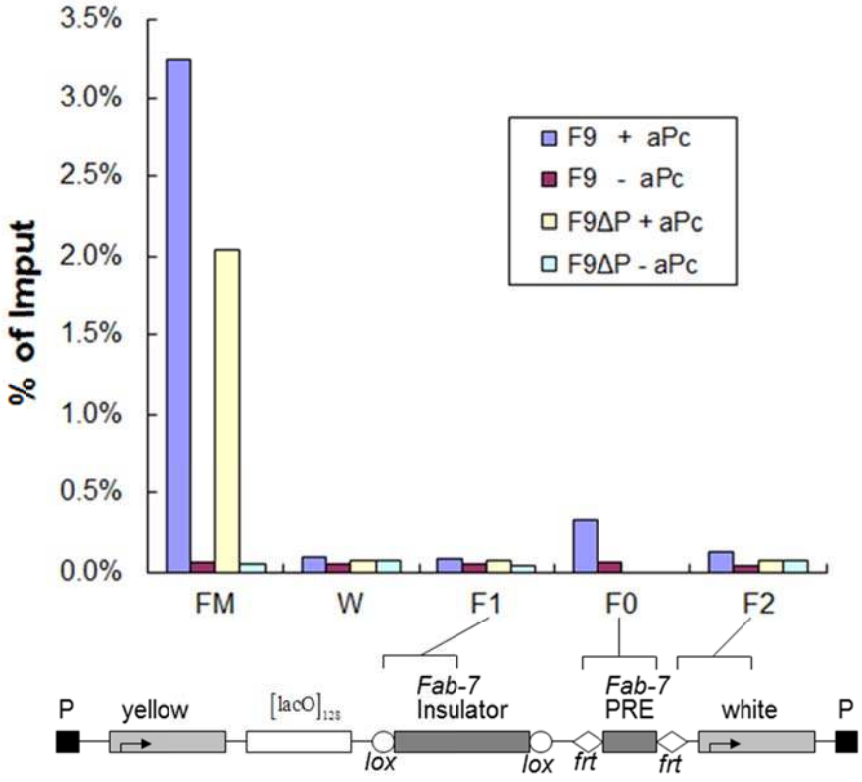

3.5%
3.0%
2.5%
2.0%
1.5%
1.0%
0.5%
0.0%
% of Imput
F9 + aPc
F9 - aPc
F9ΔP + aPc
F9ΔP - aPc
FM
W
F1
F0
F2
Fab-7 Insulator
Fab-7 PRE
P
yellow
[lacO]128
white
P
lox
lox
frt
frt

Figure 2.7 ChIP binding assay for Pc at F9 and M31 fly larva.

The 3rd instar larva are used to prepare the Chromatin used for α-Pc ChIP. The primers FM locate on the *bxd* PRE and serve as positive control, the primers W locate on the third exon of the white gene and serve as the negative control, the location of the primers on the transgenes is indicated. a) the α-PC ChIP assay on the M31 and M31ΔP larva. b) the α-Pc ChIP assay on the F9 and F9ΔP larva.

2.5 Discussion

2.5.1 Polycomb complexes and long-range interactions

As we and others have shown, PcG binding sites (PREs) in the nucleus can interact, often even when they are very distant from one another in the genome, resulting in enhanced repression. We have shown here that PcG complexes bound at three different PREs (*bxd*, *Mcp* and *Fab-7*) are neither necessary nor sufficient to mediate long-distance interaction. The previously described long-range interactions involving the *Mcp* and *Fab-7* elements are not generated by the PcG complexes that bind to their PREs but by the chromatin insulators that flank these PREs. A similar kind of long-range interaction has been proposed for Su(Hw) insulators (Byrd et al., 2004) and, in fact, the *gypsy* Su(Hw) insulator was shown to mediate an analogous interaction between remote constructs containing the *bxd* PRE(Sigrist & Pirrotta, 1997) or between the *yellow* enhancers on one construct and the promoter at a remote site(Kravchenko et al. 2005). Although we cannot exclude the possibility that certain PcG complexes have an intrinsic ability to form clusters, these results suggest that the foci of PcG proteins, the Polycomb "bodies", that have been visualized in the nucleus are brought together primarily by insulator mechanisms and not by PcG interactions alone.

Grimaud et al (2006) found that *Antp* and the *Abd-B* genes, separated by 10 Mb, co-localized in nuclei in which both were repressed but not when one was active and the other repressed. They concluded therefore that the interaction was related to the binding of PcG complexes. Our results indicate that insulators, present in both the Antennapedia

Complex and the Bithorax complex, rather than PcG complexes are responsible for the co-localization. Other factors might also be involved. One of these is the state of activity of the promoter. Active genes have been reported to become associated with “transcription factories” (Cook 1994; Osborne et al. 2004), which are not likely to cohabit with PcG target regions. The different localization of the *Antp* and *Abd-B* loci when one of the two is active and one repressed might therefore reflect the transcriptional state rather than the binding of PcG complexes as such. Preliminary evidence suggests that the transcriptional activity of the associated genes has a powerful influence on nuclear co-localization of *Mcp* elements (Results shown in the following Chapter).

2.5.2 Insulators interactions

The currently preferred model for the action of chromatin insulators is based largely on the behavior of the *gypsy* Su(Hw) insulator (Capelson & Corces, 2004). A DNA-binding protein, Su(Hw), binds to specific DNA sequences found in the insulator; a second layer of proteins capable of extensive protein-protein interactions binds to Su(Hw). These proteins, Mod(mdg4) and CP190, both have POZ/BTB domains that mediate homo- and heterotypic interactions and are thought to be responsible for the association of multiple insulator elements into clusters (Bushey et al., 2009; Ghosh et al., 2004). This clustering and the consequent organization of the chromatin into loops constitute a powerful mechanism that brings together remote chromatin sites.

The binding of some insulator proteins to the *Mcp* and *Fab-7* regions has been reported (Bushey et al., 2009; Holohan et al., 2007; Mohan et al., 2007) and more recently genome-wide ChIP/chip analysis of SU(HW), CTCF, MOD(MDG4), CP190, BEAF and ZW5 has been carried out by the modENCODE Drosophila Chromatin Consortium (2009). Neither *Mcp* nor *Fab-7* bind Su(Hw), however, *Mcp* binds CTCF and CP190, while *Fab-7* binds CP190 and a very small amount of CTCF (Holohan et al. 2007; see also Fig. 2.7).

The sequence of the *Mcp* insulator fragment contains at least one CTCF binding consensus but no obvious match could be found for this consensus in the *Fab-7* insulator sequence. In contrast, the *Fab-7* insulator contains six GAF-binding GAGAG motifs and the *Fab-7* PRE contains two. GAF binding was found to be important for *Fab-7* insulator activity (Schweinsberg et al. 2004) but also for the silencing activity of the *Fab-7* PRE (Mishra et al. 2001). GAF binding was also reported to be required for the *Mcp* silencing activity, although the *Mcp* PRE contains only one GAGAG consensus in the PRE region (Busturia et al. 2001) and none in the insulator region. The GAF protein also contains a POZ/BTB domain that might, in some circumstances, interact with the CP190 POZ/BTB and account for the interaction between *Mcp* and *Fab-7*.

2.5.3 Does homology play a role?

Bantignies et al.(2003) argued that homology plays a role because they saw no interaction between insertions containing their *Fab-7* element and insertions containing the *bxd* PRE and, although they observed interactions between *Fab-7* and *Mcp* insertions, they were

weaker than those between *Fab-7* insertions. We now know, however, that the *bxd* PRE has no associated insulator and our experiments show that, by itself, a PRE is unable to engage in long-range interactions. In our experiments, all interacting *Mcp* constructs share at least the *miniwhite* gene and the 200 bp insulator part of *Mcp.* However, interactions between *Mcp* insulator constructs and the endogenous *Mcp* depend on no homology other than the 200 bp insulator. Homology may well contribute to the stability of the interaction and certainly chromosomal homology is sufficient for the pairing of elements inserted at allelic sites on homologous chromosomes. However, in our constructs, we must conclude that the interaction is largely independent of the extent of homology but depends critically on the presence of the insulator in both interacting partners.

2.5.4 Is the long-distance contact important for gene expression?

In the case of allelic pairing, the proximity of two PREs can have very strong effects on PcG repression, as shown by the fact that the eye color can go from orange in the heterozygous case to entirely white when homozygous for the insertion. The effects of long-range interactions on PRE-dependent repression are generally much subtler. How important physiologically and how wide-spread are such interactions? Several thousand Su(Hw) sites have been mapped in the Drosophila genome and similar numbers of sites have been mapped for CTCF and CP190 (Bushey et al., 2009). Some of these, like those in the Bithorax Complex, may be needed to form higher order folding to bring together PREs and other regulatory elements as has been reported to occur at the mammalian Igf2-H19 locus (Kurukuti et al. 2006) or the globin locus (Tolhuis et al., 2002).

The low level of co-localization detected by the in vivo GFP-tagged imaging requires explanation. The percent co-localization detected in these experiments is one order of magnitude lower than that reported by Vazquez et al.(Vazquez et al., 2006) using the same technique to detect trans-interactions of the 2.9 kb *Mcp* fragment. Our analysis of the fragments flanking the 800 bp *Mcp* core shows that they make no contribution to insulator, PRE, or co-localisation. Another possible explanation is that the construct used by Vazquez et al. contained two additional insulator-like elements, the *scs* and *scs'* elements, placed at the two ends of the construct. These elements have never been specifically tested for their contribution to trans-interactions. Another possible player is the eye enhancer of the *white* gene, which was included in the construct used by Vazquez et al. but not in the constructs used in the present work. Preliminary results suggest that the eye enhancer may make an important contribution to co-localization (H.L. and V.P., following Chapter).

2.5.5 The RNAi connection

If insulators are frequently responsible for the association of remote PREs in the nucleus, this might also account for another puzzling observation. Grimaud et al.(2006) reported that the long-range interactions and degree of silencing of reporter genes produced by constructs containing the *Fab-7* element were affected by mutations in the RNAi machinery. They attributed this to an involvement of the RNAi machinery in regulating PcG silencing activity. An alternative explanation is suggested if trans-interactions are due to insulator elements. Work on the Su(Hw) insulator has revealed that *Argonaute* genes are needed for

efficient insulator activity and *Argonaute* mutations were associated with loss of higher order interactions between insulator elements in the nucleus (Lei & Corces, 2006). An attractive synthesis of these results with our observations would explain the effect of RNAi mutations on PcG silencing as due to the loss of long-range interactions brought about by insulator elements such as those found in the *Mcp* and *Fab-7* elements. Since co-localization results in enhancement of repression, loss of co-localization would account for the modest but significant reduction in PcG function at target genes seen in the presence of RNAi mutations.

Chapter 3. Insulator-dependent long-range contacts between active enhancers

Katsuhito Ohno injected the plasmids into embryos to produce all the fly lines used in this chapter.

3.1 Abstract

Insulator elements have important roles in the organization of higher-order nuclear structure. They could mediate long-distance interaction between Polycomb targets and form the Polycomb bodies in Drosophila, which result in enhanced repression. Here we show that *Mcp* insulator could bring two distant *Mcp*-eye enhancer transgenes into the same transcription factories in up to 80% of eye imaginal disc cells where the eye enhancer is active, but only in around 10% of the wing disc cells where eye enhancer is inactive. The insulator part of *Mcp* is absolutely required for the long-distance interactions, and the PRE part of *Mcp* and the eye enhancer are also required for the high stable interaction in the eye disc cells which also need dCTCF and TRX, but not Pc. *Mcp* could also mediate higher interactions with a different enhancer where this enhancer is active. Those findings suggest that insulators could partner with silencers or activators, and bring genes into distinct sub-nuclear bodies, thus organize chromosome structure and regulate gene expression.

3.2 Introduction

Eukaryotic genomes are organized into chromosomal domains and territories in the nucleus to efficiently regulate gene expression in various developmental stages and in different types of cells. Many biological activities are concentrated in subnuclear foci called nuclear bodies, which include nuclear speckles, PML bodies, PcG bodies, and transcription factories (Chakalova et al., 2005; Kumaran et al. 2008). Those function-based foci may arise through the self-assembly of essential components and act to create a local high concentration of essential factors to enhance the functions (Chambeyron et al. 2004). However, the mechanisms by which genes are targeted to those foci are still poorly understood. Here we present evidence indicating that insulator elements may play a role in the formation of nuclear foci.

Insulator elements are experimentally defined by their ability to block promoter-enhancer communication and shield transgenes from position effects (Gaszner et al. 2006; Kuhn et al. 2003; Valenzuela et al. 2006). Insulators have been implicated to play an important role in chromosome higher-order organization, by attaching chromatin fibers to the nuclear matrix and creating transcriptionally independent looped domains (Byrd et al. 2003). Genomic binding of insulator proteins, including CTCF in mammals, dCTCF, CP190, Su(Hw) and BEAF in Drosophila, have been mapped by ChIP-chip/ ChIP-seq, revealing that these proteins are found at thousands of sites scattered all over the genome (Kim et al. 2007; Bushey et al. 2009; modENCODE *Drosophila* Chromatin Consortium). Therefore, most genes,

with their regulatory elements, such as enhancers and silencers, have at least one nearby insulator, which could potentially bring those genes in proximity to other genomic sites.

PcG bodies have been observed by immuno-staining with antibodies against PcG proteins both in mammalian and Drosophila nuclei. Distant genes in two Hox gene clusters in *Drosophila* were found to interact with each other within the same PcG body, and the co-localization is required for enhanced silencing of both genes and thought to be dependent on the PcG proteins in specific tissues and during specific developmental stages (Bantignies et al. 2011). *Fab-7* and *Mcp* are so-called boundary elements that separate cis-regulatory regions of the Bithorax Complex in Drosophila. Each has been shown to contain two separable functional parts: the core PRE and the insulator. Cavalli and colleagues also showed that transgenes containing the full *Fab-7* boundary region could interact with the endogenous *Fab-7* and co-localize at relatively high frequencies inside PcG bodies and RNAi bodies, and argue that this co-localization requires PcG proteins (Bantignies et al., 2003; Grimaud et al., 2006). Using transgenes containing the *bxd* PRE, *Fab-7* and *Mcp* elements, we showed in a previous report (Li et al. 2011) that PREs have no intrinsic ability to co-localize and that long-distance interactions between PREs are mediated by the insulator elements, not by the PREs, although PcG proteins bound on PREs may stabilize and increase the interactions between Polycomb targets inside the PcG bodies. The frequencies of localization that have been reported in the various published experiments are rather

low, ranging from 6 to 25% of nuclei. A striking exception are the live imaging results of Vazquez et al. (2006) where the frequencies of co-localization of two remote transgenes containing *Mcp* was as high as 90% of the nuclei. A comparison of our results with those of Vazquez et al. (2006) suggested that the critical difference lay in the presence of the *white* eye enhancer in their construct and consequent transcriptional activity in the eye disc. In the present work, we examine this possibility and conclude that transcriptional competence is in fact an important component that drives co-localization, leading to the conclusion that insulators are involved not only in the formation of Polycomb bodies but are likely to play a role in foci of transcriptional activity.

3.3 Materials and Methods

3.3.1 Transgene constructs

Mcp-101B is essentially identical to the construct LacO-Mcp used in Li et al. (2011), except that the LOX and FRT flanked elements were changed into the *white* eye enhancer and 820-bp *Mcp*. Therefore, to assemble Mcp-101 constructs, the *white* eye enhancer was PCR-amplified from the plasmid 3H3pUCP (described in Qian et al., 1992), using primers we+: CTGGGAATTCAGTCAACCCAGACCAACC, and we-: GATATGGATCCGACTGGGACG; the 820bp *Mcp* was PCR amplified from BX-C clone BAC R24L18 (obtained from BACPAC Resources Center, http://bacpac.chori.org/), using PCR primers Mcp+: CTGGGGATCCGGCCGTTTTCCGTTTTATTG, and Mcp820-: CGAAGCATGCTAGAAAAATTCCGCACCAG. To assemble Mcp-100A construct, the insulator part of *Mcp* was PCR-amplified from BX-C clone BAC R24L18 using PCR primers Mcp+: CTGGGGATCCGGCCGTTTTCCGTTTTATTG, and Mcp530-: CACTGCATGCTGAGAAACCCAAGCGTTG. To assemble Mcp-126A, the *Mcp* part is the same 820-bp fragment as in Mcp-101, while the 5×UAS was PCR-amplified from vector pUASTattB (GenBank accession No.: EF362409.1), using PCR primers Gal4+: CTTGGAGCTCCGCGGCACTGGAACTAGGCTAG, and Gal4-: CGTTGAATTCCGGCGCTCGCTAGAG. As a result, the 820-bp *Mcp* fragment corresponds to Chr3R nucleotides 12694651 to 12695464, the insulator part of *Mcp*

fragment Mcp^{Ins} corresponds to Chr3R nucleotides 12694651 to 12695176, and the *white* eye enhancer fragment corresponds to ChrX nucleotides 2692371 to 2691262. All genomic sequences are taken from FlyBase Genome Release 5.

All those constructs were assembled in a similar way. First, the PCR fragments were each inserted into the LOX and FRT cassette vectors, and the resulting plasmids were sequenced to verify the inserted sequences. Then, fragments containing the LOX flanked eye enhancer/ UAS and the FRT flanked *Mcp* were assembled into pBluescript. The tandem array of 128 copies of LacO was cut from pAFS150 (a gift from J. Vazquez) and inserted into pCaSpeR4, and this plasmid was used to accept the Kpn Ⅰ fragments containing LOX flanked enhancer part and FRT-flanked *Mcp* part, thus give rise to two plasmids with two different positions of *Mcp* relative to the *white* reporter gene (Mcp-101B and Mcp-101A).

3.3.2 Fly stocks

Transgenic fly lines were made according to standard procedures (Spradling et al. 1982). Southern blot hybridization was used to verify that the lines contained a single insert and inverse PCR was used to identify the exact insertion sites. The various deletion derivatives were established with the help of Flipase and Cre recombinase-producing stocks (Siegal et al. 1996), as previously described in Gohl et al. (2008) and were verified by PCR analysis. For co- localization studies, two transgene lines on different chromosomes were crossed together through double

balancers. In the case of lines with insertions on the same chromosome the two insertions were recombined to obtain a cis-arrangement. PCR was used to verify the presence of both transgenes. The LacI-EGFP line was described in Li et al. (2011), the mRFP-LacI line was a gift from A. Csink (Thakar et al. 2005; 2006), and the Pc-GFP line was kindly provide by R. Paro (Dietzel et al., 1999).

3.3.3 In-vivo imaging and Microscopy

After crosses of transgenic flies with LacI-EGFP flies, the larvae were raised at 18°C and supplemented with active dry yeast. Third instar larvae were rinsed and dissected in Gibco Schneider's Drosophila media (Invitrogen Co.). The dissected eye and wing imaginal discs were aligned on a bottom dish (MatTek Co.) with a drop of Drosophila medium, and then covered with a coverslip. Z-stack images were taken with a DeltaVision Image Restoration Microscope system (Applied Precision Instrument, LLC Issaquah, WA), using a 100x /1.35 UplanApo objective, deconvoluted and processed with the SoftWoRx software (Applied Precision Instruments). The dots in each nucleus were scored, one dot as co-localization and two non-overlapping dots (center to center distance greater than 0.3 mm) as no co-localization.

3.4 Results

3.4.1 The 820-bp *Mcp* can mediate high frequency co-localization in the eye disc

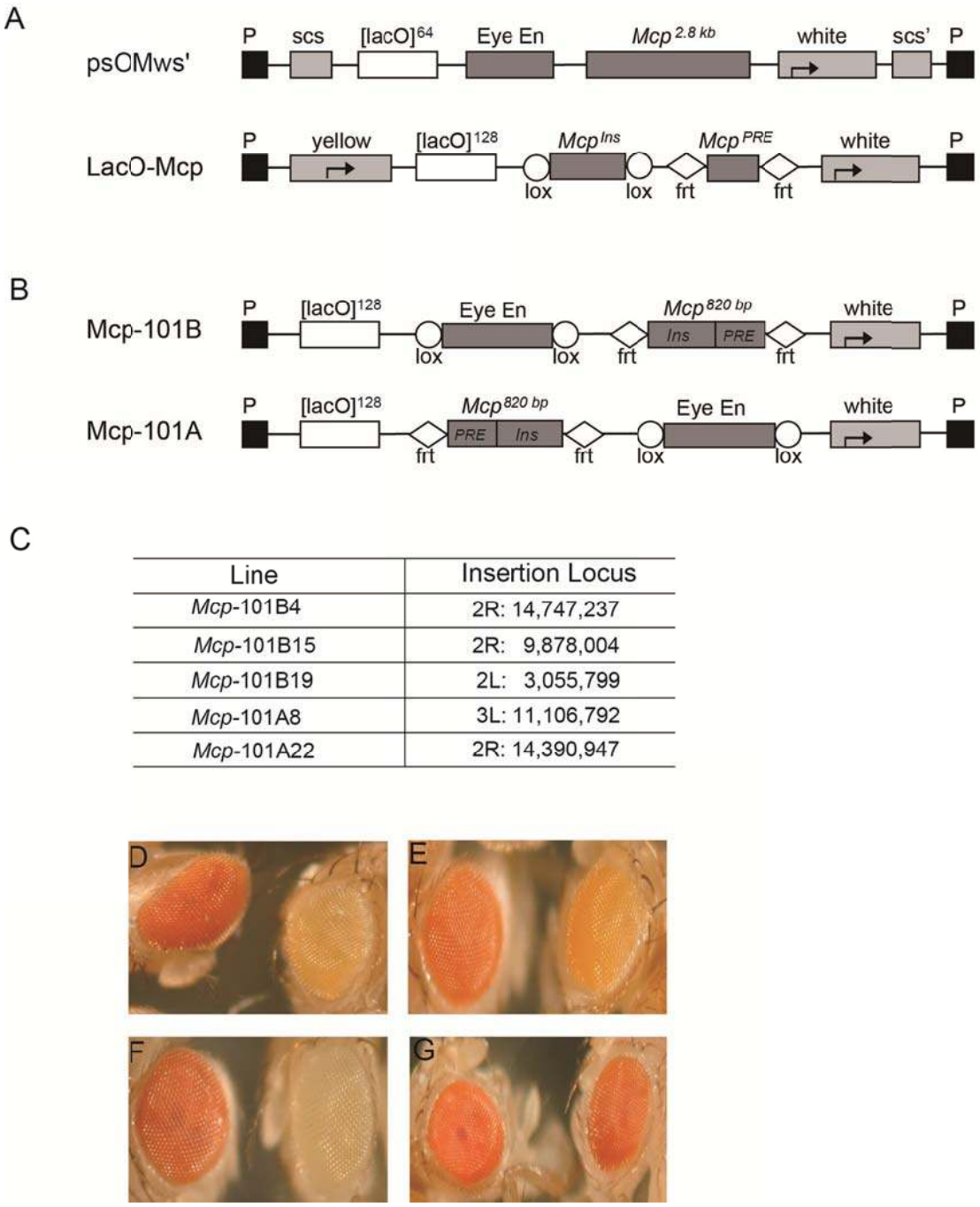
Vazquez et al. (2006) reported that the 2.8-kb *Mcp* could mediate stable interaction of two distant *Mcp* transgenes resulting in up to 90% co-localization in the cells of the eye disc. Our previous results with transgenes containing a minimal *Mcp* element showed only around 7% interaction in both eye and wing imaginal disc cells that dropped to less than 0.5% when the insulator part of *Mcp* was deleted (Li et al. 2011). We repeated Vazquez's experiments using their OM4-OM6 transgenic lines and using our LacI-EGFP expressed from the ubiquitin promoter. The results showed that the *Mcp* transgenes could co-localize at a frequency of around 50% in eye imaginal disc cells, but only around 10% in wing and eye membrane cells. Therefore, the Vazquez *Mcp* construct really co-localizes at a much higher frequency and this high frequency is specific for the eye imaginal disc and not for other tissues.

The obvious discrepancy between those two experiments led us to look into the differences between the two transgene constructs used (Fig. 3.1A): First, Vazquez et al. flanked their construct with the *scs* and *scs'* elements., However, in their control experiments, *scs*/*scs'* had no effects on long-distance interactions between transgenes in the absence of *Mcp* or in tissues other than the eye disc. Second, they used full length 2.8-kb *Mcp*, while we used minimal *Mcp* in which the insulator and PRE

elements were separated by LOX and FRT sites, respectively. Our results showed that the minimal *Mcp* fragment has all functions of the 2.8-kb *Mcp* (Li et al., 2011; Muller et al., 1999). Third, the Vazquez et al. construct included the *white* eye enhancer. Since the high interaction frequency was specifically found in eye disc cells (50%) and not in wing disc cells (10%) in their *Mcp* transgenic lines, we surmised that the eye enhancer might be responsible for stabilizing and enhancing the interaction between *Mcp* transgenes. Surprisingly, this would imply that enhancer/transcriptional activity increases the ability of *Mcp* transgenes to co-localize.

To prove the hypothesis that insulator could mediate high stable interactions between active enhancers/genes, we made two different constructs containing the 820-bp *Mcp* and the *white* eye enhancer, flanked by FRT and LOX respectively, but with different position relative to the *white* reporter gene (Fig. 3.1B). The Mcp-101B construct assumes the same configuration as that used in the construct of Vazquez et al., while in Mcp-101A the *Mcp* insulator is placed distal to the enhancer and *white* gene. The transgenes were tagged with 128 copies of LacO repeats, which could be visualized directly under fluorescence microscope after crossing with EGFP-LacI lines driven by the Ubiquitin promoter (Li et al. 2011). Two independent lines were obtained for Mcp-101A, and three independent lines for Mcp-101B all of which showed typical pairing-dependent silencing effects (Fig. 3.1C-1G), indicating that *Mcp* could insulate the *white* eye enhancer function from activating the *white* gene.

We then combined the two transgenes into one fly line, and visualized the interaction between the two transgenes by live imaging techniques after crossing with EGFP line (Fig. 3.2A). In the re-constituted 3D images, nuclei with one dot were taken to indicate co-localization and two dots as no interaction (Li et al. 2011). The results showed that in all combinations the 820-bp *Mcp* could mediate high frequency interactions (50% to 85%) in the eye imaginal disc cells, but low colocalization (5% to 10%) in wing disc and membrane cells (Fig 3.2b). This result reproduces the observations of Vazquez et al. with the full-length 2.8-kb *Mcp* fragment. In addition, the two relative positions of the Mcp and enhancer with respect to the *white* gene gave similar results as did the interactions between transgenes with different relative positions, which suggests that to get high frequency co-localization it is sufficient to have the enhancer in the transgenes even if it does not stimulate the promoter.



Line	Insertion Locus
Mcp-101B4	2R: 14,747,237
Mcp-101B15	2R: 9,878,004
Mcp-101B19	2L: 3,055,799
Mcp-101A8	3L: 11,106,792
Mcp-101A22	2R: 14,390,947

Figure 3.1 Comparison and maps of *Mcp* Drosophila lines.

(A) Direct comparison of psOMws' and LacO-Mcp constructs. (B) Structure of new *Mcp*-enhancer reporter constructs. The *Mcp* is flanked by FRT sites and white enhancer by LOX sites. The construct utilize the *mini-white* genes as markers. The tandem array of 128 lacO repeats is used to bind the lacI repressor fused toEGFP-NLS expressed from a different construct driven by the ubiquitin promoter. (C) The fly

Figure legend of Figure 3.1 continued.

lines obtained and their insertion sites determined by inverse PCR. (D) – (F) Pairing dependent silencing of *white* in flies lines Mcp-101B4, Mcp-101B15 and Mcp-101B19. The eye on the left shows heterozygous fly eye color, while the right homozygous. (G) Typical eye color of Mcp-101A8 and Mcp-101A22 fly lines, which do not show pairing dependent silencing of eye color.

3.4.2 The enhancer is required for high interaction and accounts for the difference between tissues

The eye imaginal disc of Drosophila 3rd instar larvae contains two layers of cells: the eye columnar cells and the membrane cells that cover it. In the region posterior to the morphogenetic furrow, the columnar cells differentiate and assemble in clusters containing seven photoreceptor cells in which the eye enhancer is activated. The activation of the eye enhancer could possibly account for the large difference between eye imaginal disc and wing disc cells for the co-localization of the transgene.

To test this possibility, we excised the eye enhancer on the transgene using the flanking LOX sites by crossing the fly lines to Cre recombinase fly line, and then recombined two enhancer-deficient lines into one line. As shown in Figure 3.2B, deletion of the eye enhancer (ΔE) dramatically decreases the frequency of co-localization in the eye disc cells (from ~70% down to 20%) in all the five different recombination lines. Interestingly, even after deletion of *white* enhancer, we constantly got higher co-localization frequency in the eye disc cells (~20%) than in wing and membrane cells among all five combinations (vary from 4% to 10%), probably due to the fact that the reporter *white* gene still have residual eye-specific transcriptional activity without the enhancer, as evidenced by the eye color from dark red to light red or yellow. We conclude that the enhancer activity helps to bring two distant transgenes together, and the ratio of interaction in the eye disc cells could be an indicator of the transcription level. While in the case where the insulator blocks the

eye enhancer activities when placed in between the enhancer and *white* genes, we still could get high frequency of co-localization even in low transcription levels, probably because that the *white* eye enhancer help the insulators to stabilize the long distance interaction between the transgenes even without stimulating transcription.

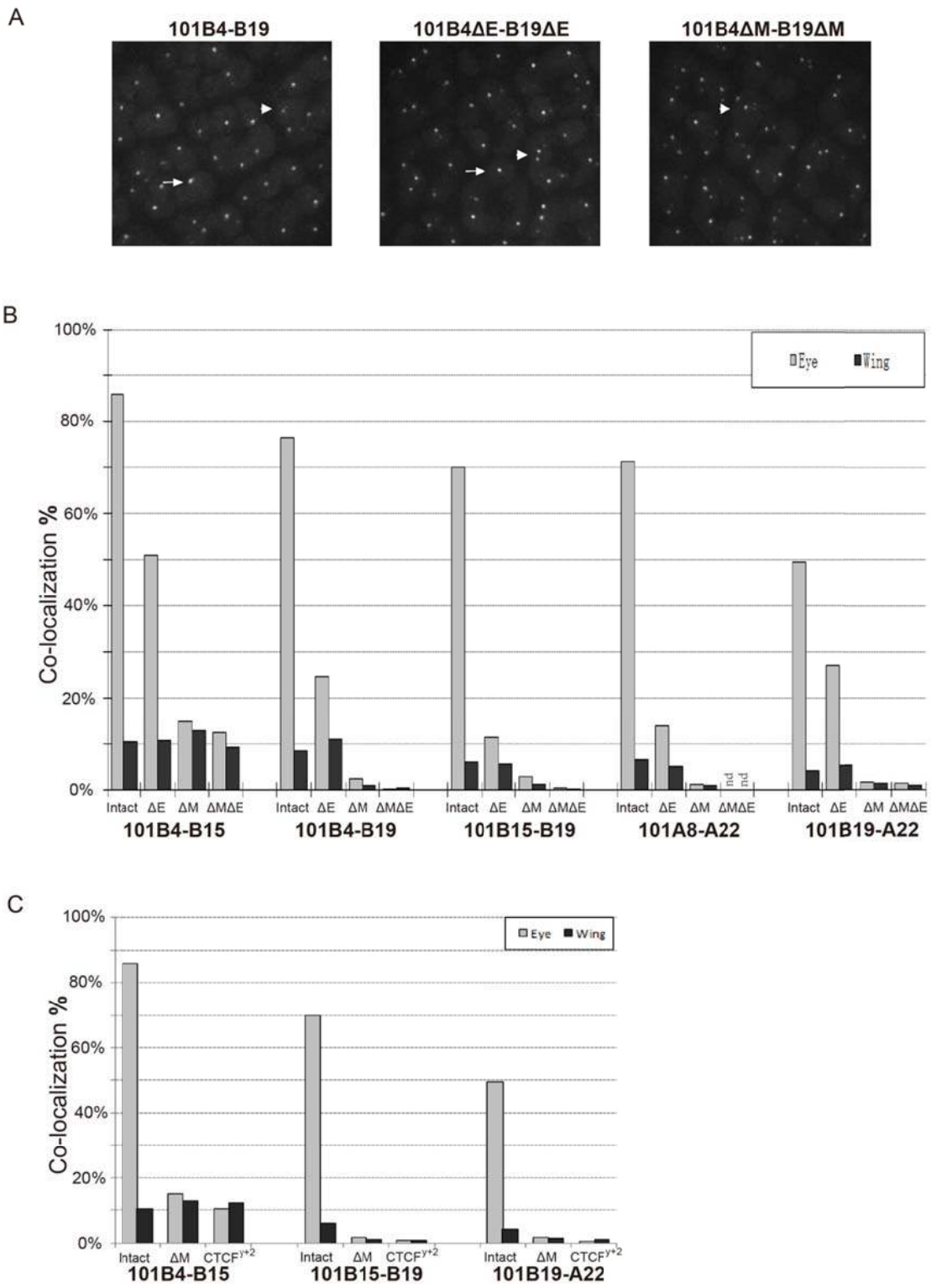


Figure 3.2 Frequency of long-range interaction of Mcp-enhancer transgenes.

(A) Typical GFP images of eye imaginal disc nuclei showing one-dot (arrows) and two-dot (arrowheads) nuclei from Mcp-101B4-B19 (left), Mcp-101B4ΔE-B19ΔE (middle), Mcp-101B4ΔM (right). ΔE, deletion of eye enhancer; ΔM, deletion of *Mcp*.

Figure legend of Figure 3.2 continued.

(B) Interaction between Mcp-enhancer transgenes. For each of the pairwise combinations between five lines bearing the *Mcp*-enhancer construct, the histogram shows frequencies of co-localization (one dot nuclei) for the starting lines (intact), for the lines after eye enhancer deletion (ΔE), for the lines after *Mcp* deletion (ΔM), or for the lines after double deletion (ΔMΔE). Gray bars show the ratio of colocalization in the eye disc cells, while the dark bar show the ratio in the wing and membrane cells. (C) Interactions in dCTCF mutant background. The frequency of co-localization after dCTCF mutation drop down to the background level, similar to the *Mcp* deletion (ΔM).

3.4.3 Insulators are required for the long-range interaction

Next, we deleted the *Mcp* part of the transgene (ΔM) to see if this is still necessary to mediate the long-range interaction. The results (Fig. 3.2B) show that deletion of *Mcp* almost totally abolished the interaction both in eye and wing imaginal disc cells (varies from 70% down to less than 2%). Double deletion of both *Mcp* and Enhancer shows similar results (Fig. 3.2B, ΔMΔE). The only exception is the 101B4-B15 line which still displays around 10% co-localization after *Mcp* deletion or double deletion, probably because the two transgenes are on the same chromosome arm and fairly close to each other (~5Mb away), and other insulator elements surround the insertion sites.

The Mcp fragment contains two separable functions: the insulator and the PRE (Busturia et al. 2001; Gruzdeva et al. 2005; Kyrchanova et al. 2007). Previously, we have shown that the insulator, not the PRE function is essential for co-localization at the low level observed in the absence of the eye enhancer (Li et al. 2011). To determine if the insulator is still required for the high level co-localization, we specifically abrogated this function. The *Mcp* insulator has been shown to bind the insulator protein dCTCF, which can then recruit other insulator proteins such as CP190 (Bushey et al., 2009; Holohan et al., 2007; Mohan et al. 2007; Nègre et al., 2010; Li et al., 2011). To inactivate insulator function, we introduced a homozygous dCTCF loss of function mutation in our experiment. The results show that loss of dCTCF knocks down the interaction of two remote transgenes to a level similar to that

seen when the *Mcp* is deleted (Fig. 3.2C). Therefore, we conclude that insulator function is still essential for long-range interaction and, in its absence, the enhancer alone cannot interact with its distant partners.

3.4.4 TRX, not PC, is also required for high interaction

Since PREs generally repress gene expression and antagonize enhancer action, we reasoned that the high interaction we observed might only need the insulator part of *Mcp* and the enhancer activity.

To test this hypothesis, we made another construct similar to Mcp-101A except that only the insulator part of *Mcp* was used instead of the 820-bp *Mcp* fragment (Fig. 3.3A). Several independent 100A lines were generated, and combined to test the frequency of co-localization. Unexpectedly, these lines only show the basic 5~7% insulator-dependent co-localization frequency both in the eye and wing imaginal disc cells (Fig. 3.3B), similar to that obtained with the insulator alone, with no enhancer and no PRE (H.-bing Li et al. 2011). These results indicate that the PRE part is in fact important for the enhancer to mediate high frequency long-range interaction.

In order to understand the function of the PRE, we returned to the 820-bp Mcp-101 lines, and tested them in a *Polycomb* mutant background. Since homozygous Pc^{-} flies die at the embryonic stage, we could only test whether reducing the *Pc* dosage to half the normal level affects co-localization. As shown in figure 3.3C, the *Pc* heterozygous mutation does not have any effect on the long-distance interaction. So the high interaction is not sensitive to PC levels although the remaining Pc activity might provide sufficient function. To see if the transgenes associate with Polycomb bodies, we repeated the experiment in the presence of a PC-GFP transgene that would enable

us to see the PcG bodies in live cells (Dietzel et al., 1999; Ficz et al. 2005), in addition to the LacI-RFP transgene to visualize the *Mcp* transgenes (Thakar et al. 2005; 2006). The live-imaging results showed that the *Mcp* transgenes do not co-localize with PcG bodies (Fig. 3.3D), indicating that PcG proteins do not bind to PRE in the eye disc.

Trithorax protein (TRX) binds constitutively to all known or putative PREs (therefore also called TREs) regardless of whether these sites also bind PcG proteins and regardless of the transcriptional activities of the target genes. PcG target genes have been shown to be positively regulated by TRX, a histone methyltranferase that is known to methylate H3K4 and to antagonize PcG repression (Schwartz & Pirrotta, 2007; 2008). To test whether TRX might be the PRE-binding factor required for high level co-localization, we crossed our Mcp-101 lines into a TRX-deficient genetic background. Homozygous *trx* loss of function mutations are embryonic lethal, therefore we can only test the co-localization frequency under heterozygous *trx* conditions. The results showed that reducing the level of TRX by half brings the high interaction in the eye disc cells down to the basic level, the same as that found in the membrane and wing disc cells (Fig. 3.3E). Together, those data demonstrate that the high frequency co-localization is hightly dependent on TRX concentration but is not very sensitive to the PC concentration. This argues that the TRX/TRE, but not the PRE function, is required for the high interaction, and indicates that epigenetic regulation is involved in the long-range interaction.

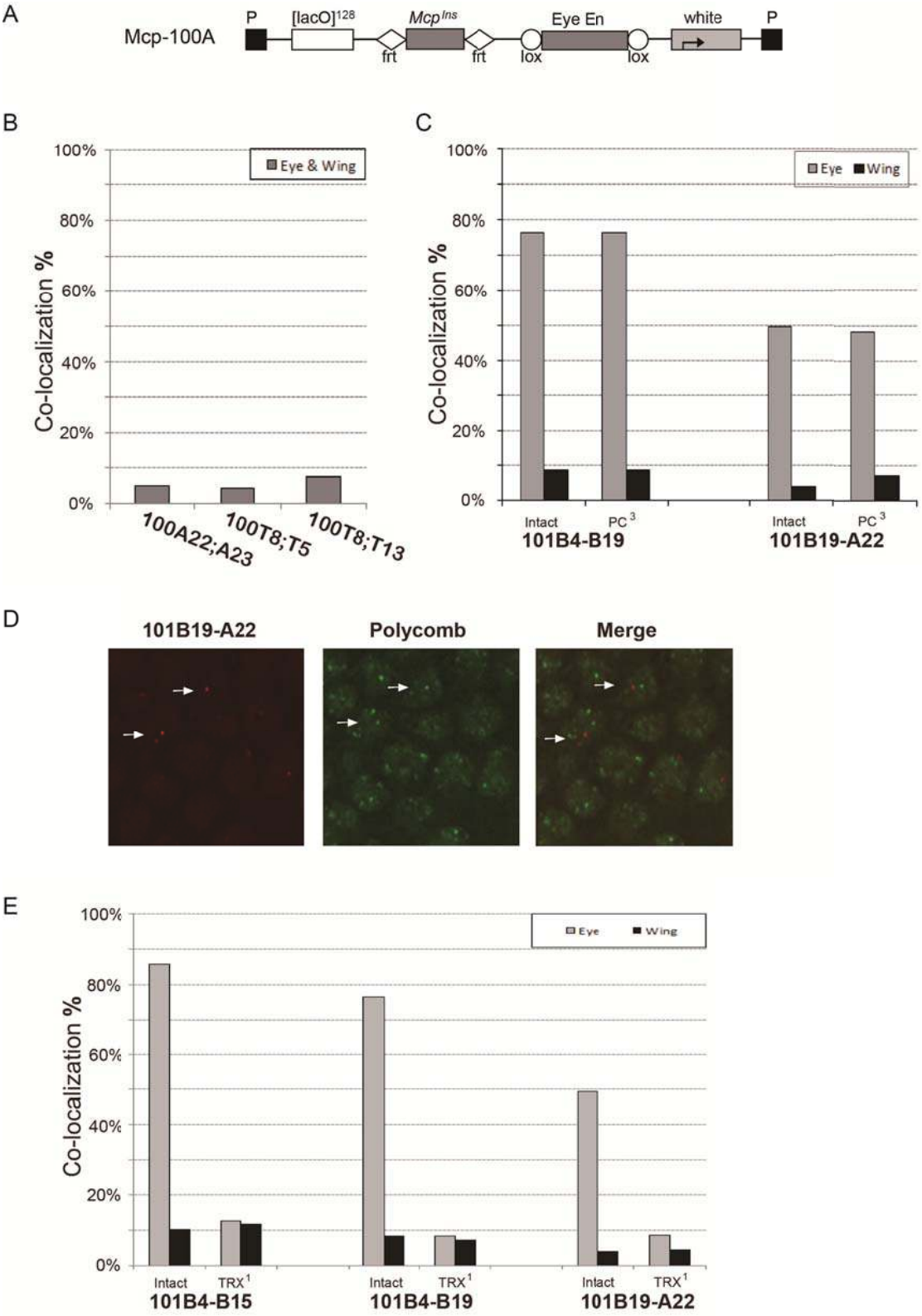


Figure 3.3 The high interaction ratio needs TRE/TRX, but no PC.

(A) Structure of new Mcp insulator–white enhancer construct. Only the insulator part of *Mcp* is used here instead of whole *Mcp*. (B) Interactions between Mcp

Figure legend of Figure 3.3 continued.

insulator-enhancer transgenes. The histogram shows the frequencies of co-localization of three lines both in eye and wing disc cells, which have similar interaction ratio. (C) Interactions between Mcp-enhancer transgenes after PC mutation. The histogram shows the frequencies of co-localization of Mcp-101 transgenes after PC mutation. (D) In-vivo imaging of mRFP labeled 101B19-A22 and GFP labeled Polycomb. Figures correspond to several eye membrane cell nuclei of deconvolved single slices from 3D stacks. (E) Interactions between Mcp-enhancer transgenes after TRX mutation. The histogram shows the frequencies of co-localization of Mcp-101 transgenes after TRX mutation.

3.4.5 A Different enhancer also promotes co-localization

If the frequency of co-localization is related to the level of transcription activity rather than some unknown factor that binds specifically to the eye enhancer, a different enhancer should also promote the long-range interaction mediated by the insulator. To test this, we constructed Mcp-126A, similar to Mcp-101A except that the enhancer was replaced by five copies of the GAL4 binding site (5×UAS) (Fig. 3.4A). Three independent insertion sites, all on the third chromosome, were combined pairwise and tested in the presence or absence of the *Arm*-Gal4 driver, which can activate UAS enhancer both in eye and wing disc cells. The frequency of co-localization is consistently higher in eye disc cells than in wing disc and membrane cells with or without Arm-Gal4 driver, perhaps because that the mini-*white* gene used in these constructs has a residual eye-specific activity even in the absence of the eye enhancer. More interestingly, the co-localization increases substantially both in eye and wing disc cells in all three lines, from ~12% to 25% in eye disc cells and ~8% to ~18% in wing disc cells in the presence of *Arm*-Gal4 (Fig. 3.4B). The UAS/GAL4 activation therefore increases co-localization, although the frequency is not as high as that obtained in eye enhancer lines, probably because the *Arm*-GAL4 activation is much weaker than that due to the eye enhancer.

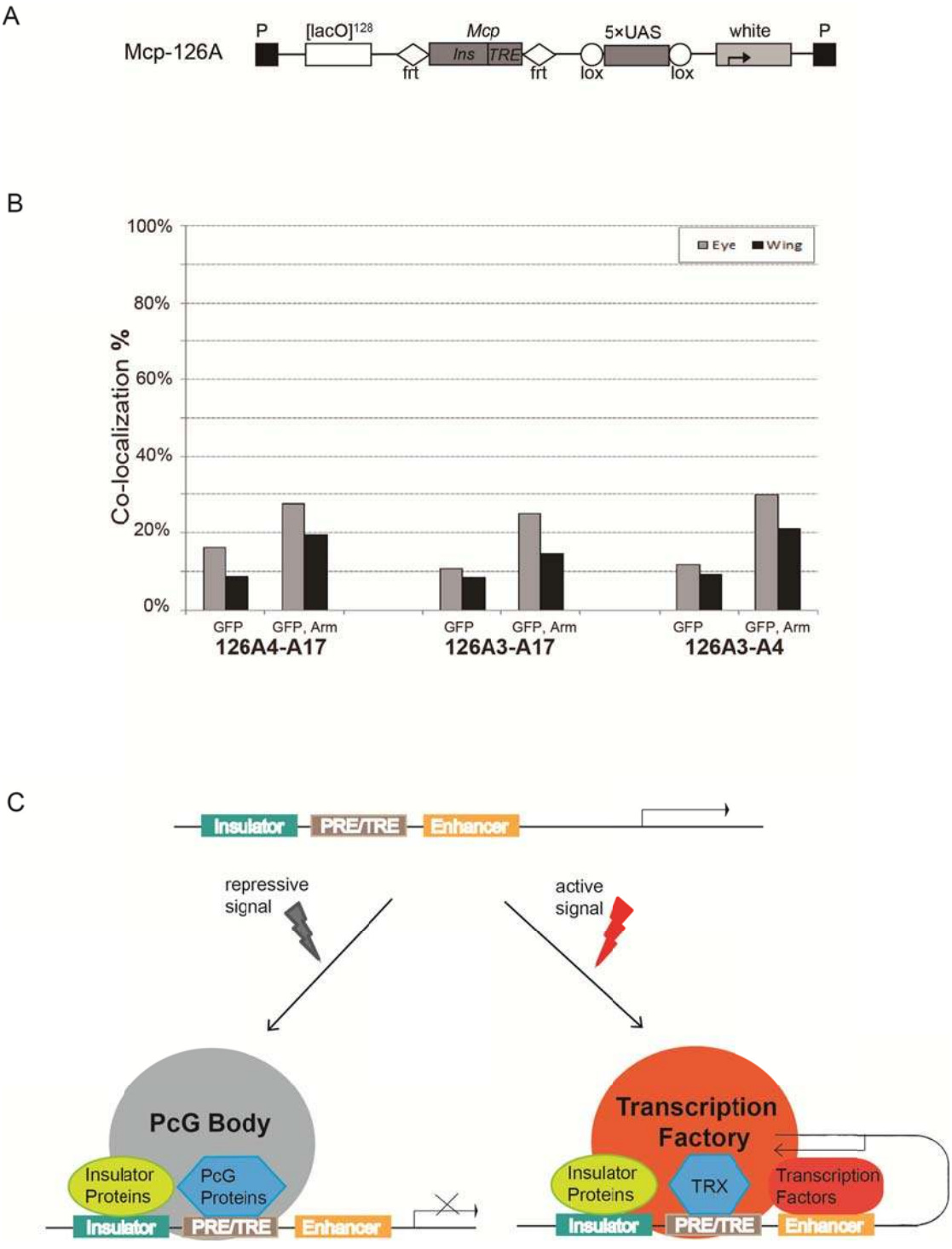


Figure 3.4 UAS enhancer could also promote the long-range interaction mediate by *Mcp*.

(A) Structure of new Mcp-UAS construct. The 5×UAS sequence is used here instead of white enhancer. (B) Interaction of Mcp-UAS transgenes after Arm-Gal4 driver

Figure legend of Figure 3.4 continued.

activation. The histogram shows the frequencies of co-localization with and without Arm-Gal4 activation both in eye (Gray bar) and in wing (Black bar) imaginal disc cells. (C) Model for insulator dual action. Insulator alone mediates basic interactions. Once the gene receives repressive signals, the PRE recruit PcG proteins, and insulator brings the gene into a pre-existing cognate PcG body. If the gene receives active signals, the TRX bind to TRE/PRE, transcription factors bind to the enhancer, then the insulator bring the gene into a cognate transcription factories.

3.5 Discussion

In this study, we demonstrate that transcriptional enhancers can promote long-range interactions mediated by *Mcp* element. While insulator part of *Mcp* and insulator proteins are required and sufficient to mediate the basic low frequency co-localization, the TRE/TRX is needed to help the enhancer to stabilize the long-distance interaction. Our observation also suggested that the interaction frequency/ratio is correlated with the transcription activity levels, the stronger of the transcription, the higher of the interaction.

3.5.1 Insulators and Transcription factories

The traditional view of transcription describes RNA polymerase as being loaded at the promoter and proceeding like a locomotive tracking on the DNA template and synthesizing RNA. The observation that nascent RNA is found at a relatively small set of discrete foci in the nuclei where it co-localizes with RNA pol Ⅱ led to the current model for transcription (Jackson et al., 1993; Wensink, 1993; Osborne et al. 2004), according to which polymerases concentrate in discrete ‘factories’, immobilized by attachment to a sub-nuclear structure, where multiple templates undergo transcription by being reeled into the factory while the newly synthesized RNAs are extruded (Cook, 1999; Cook, 2010; Sutherland et al. 2009; Chakalova et al. 2010). It has been shown by 4C that co-regulated genes cluster in the same transcription factories (Schoenfelder et al. 2009). However, one main gap in the model is an account of how genes get into the transcription factories. In this study, we provide evidence that insulator elements can

possibly bring active enhancers (genes?) into the transcription factories. Active enhancers could stabilize long-distance interaction inside the transcription factories between two distant transgenes mediated by *Mcp* insulators, and the strength of the interaction depends on the activities of the enhancer.

3.5.2 Insulators and gene regulation

Insulators have been suggested to play important roles in the maintenance of independent gene domains and in the organization of genomes (West et al., 2002). Currently, the popular model for insulator function, termed the 'structural model', propose that insulators organize the chromatin fiber within the nuclear space and create transcriptionally independent looped domains by anchoring chromatin to fixed nuclear substrate, such as the nuclear lamin and matrix, thus isolate the signals generated in a domain (Valenzuela & Kamakaka, 2006; Mongelard et al. 2001; Geyer et al. 2002). Our results would be consistent with the structural model, if we envision that the *Mcp* insulator serves as the link that connects the associated genes into pre-existing nuclear bodies, depending on the epigenetic state of the associated genes.

Previous studies have extensively addressed the looping properties and associated proteins of the insulator elements, while only few publications correlate insulator functions with the gene regulation. Cavalli and colleagues showed that *Fab-7* transgenes could co-localize with endogenous *Fab-7* inside PcG bodies, which contributed to the reporter gene silencing. They also showed that *Antp* and *Abd-B* genes

could have ~20% interaction inside PcG bodies when both were repressed, but co-localization was reduced to 6% when one gene was active and the other was silenced, and they believed that PC is the factor that bring distanct DNA together to form the PcG bodies (Bantignies et al. 2011). We previously showed that it is the insulator part of the *Fab-7* or *Mcp* that mediates the low ~7% basic interaction (Li et al. 2011), and in this research, we showed that the distant interaction was abrogated after CTCF mutation, which proved that the insulator function is essential for the long-distance interaction. So, it is possible that the PcG proteins could promote and stabilize the long-range interaction mediated by insulators inside the PcG bodies (from 6% to 20%), thus enhancing the repression. In this study, for the first time, we found that insulator elements are also involved in co-localization of transcriptionally active genes, perhaps by bringing active genes into transcription factories, and the association is stabilized by the enhancer activities.

3.5.3 Insulator dual structural model

We provide direct evidence that insulator part of *Mcp* and *Fab-7* bring the PREs together from long distance, althoght at low frequency (Li et al. 2011). Together with the results from Cavalli and colleagues, we assume that insulators bring PREs into the PcG bodies. And in the current research, we showed that two transgenes could co-localize at high frequency in the presence of both enhancer and TRE within the eye disc cells where the transgene is at active state, while only interact at basal level in the wing disc cells where the transgene is silenced. Presumabably, the transgene will

associate with the transcription factories, since the reporter gene is highly transcribed and give rise to dark red eye of the fly line. Similarly, the transgene may go to the PcG bodies where it is silenced in the wing disc cells, which is also showed by live-imaing results with Pc-GFP lines.

So, it seems that insulators not only partner with silencers to bring genes into silence bodies, but may also partner with enhancers to bring active genes into transcription factories. Our current study, together with our previous results, supports the dual structural role model of insulators in gene regulation (Fig. 4C): insulator alone only mediate basic low interaction, provide all kinds of interaction possibilities for other DNA elements; Once the gene receives repressive signals, the PcG proteins are recruited to the PRE, then the insulator will bring this gene into one of the cognate pre-existing PcG bodies; When the gene receives active signals, the transcription factors will be recruited to the enhancers, the TRX will bind to the PRE and antagonize the PcG and possibly modify the chromatin, then the insulators will bring this poised gene into a nearby transcription factory. It is possible that different enhancers will determine which factory to go. We may envision that the PcG bodies and transcription factories may be located in the interchromatin compartment, a continuous space between chromosomal territories. The spaces where the PcG bodies reside represent the silenced zone and memorize the genes inside the bodies even after cell cycles, while the transcription factories reside in the active zone. The insulator elements may participate

in the formation of those bodies/zones and help genes to arrive those different bodies/zones, depending on the epigenetic marks and the signals received.

This dual functional model of insulator need more evidence, and the current results just provide some clues that lead to this attractive model. And there is some minor confictions between this model and some of results in this research. First, according to this model, the transcription of both transgenes should be increased in the presence of insulator which bring two distanct transgenes into the same Pol Ⅱ factories. However, I do not see any significant transcription decrease (judged by the no changes of the eye color) after *Mcp* deletion both in Mcp-101A and Mcp-101B lines, and the results showed that two transgenes do not interact with each other without Mcp. What this means is that the two transgenes still could go to their own transcription factories and get highly transcribed in the presence of enhancer only, and they do not need *Mcp* insulator to bring the genes into transcription factories. This confliction could be reconciled by the possibilities that there are other other types of unknown insulators in the transgene which could bring the transgene into their closest neighbouring transcription factory, while two transgene would be more likely to share the same factory in the presence of *Mcp* because some specific factors bind on *Mcp*. The second confliction arising from this model is that, the *Mcp* insulator seems not functional in the Mcp-101B lines because those lines show strong transcription even with the *Mcp* insulator placed in between the eye enhancer and the *white* gene in heterozygous state. This could be due to the facts that the *Mcp* insulator is a weak insulator and the *white*

eye enhancer is a very strong enhancer that could overpass the insulation effects of a weak insulator. The additional paradox is that the PRE part of *Mcp* is functional as evidenced by the PRE phenotype of pairing sensitive silencing (PSS) effects shown in the Mcp-101B fly lines, which means that the transgenes may bind PcG proteins, while the transgenes supposed to be in highly active state in the presence of enhancer and strongly colocalized with each other inside the transcription factories. Althogh Vazquez observed the PSS-like phenotype of two remote transgenes in heterozygous fly lines, I do not see this phenotype with Mcp-101B lines after recombining two transgenes into one fly line. So, the quantification of the reporter *white* gene expression is needed to explain whether or not the insulator increases the transcription of the transgene. The last problem is that the PcG/PRE should increase the frequency of co-localization mediated by the insulator according to this model, but I do not see the PRE effects on the insulator mediated basal colocalization in the results of previous chapter. I do see the PRE effects on the reporter gene expression in those lines, because the eye color increase a little bit after PRE deletion (Appendix table 5). The possible explanation for this is that the PREs used there is weak PRE, and the minimal PRE used in those fly lines may lose some important sequence that is needed for the PcG enhanced co-localization.

One important thing that must be borne in mind is that this model is all based on transgene assays, and we need evidence from in-vivo genes. Recently, Bantignies et al. (2011) showed that *Antp* gene of ANT-C and the *Abd-B* gene of BX-C colocalize in the

same PcG bodies when both genes are repressed in the head of embryos in ~20% of the cells, while only 6% if one gene is active and the other one is silenced. If my model is correct, the colocalization frequency of both genes could also increase when both genes is in active states, and in both cases mutation of CTCF should abrogate the interaction of the two genes. Those possibilities could be tested by using the FISH or 3C techniques.

3.5.4 Where the transgene go?

Prelimiary data from live-imaging and immune-staining showed that the transgene will join the PcG bodies in silenced state in some of the wing disc cells, and join in the transcription factories in active state in some of the eye disc cells. However, firstly, the data is hard to quantify due to the nature of the techniques and the definition of the bodies. Second, in preliminary experiments, I do not observe the colocalization between the transgenes and the PcG bodies in the majority of the wing disc cells, neither the transgenes and the Pol Ⅱ foci in most of the eye disc cells. So, the questions here are where those genes go. According to the image data, the transgene seems do not bind PC in repressed states in most of the wing disc cells, and the transgenes do not associate with strong Pol Ⅱ foci in most of the eye disc cells when in active states.

It is also interesting to know the relationship between the active transgenes with endogenous Mcp. My results in the previous chapter has shown that the Mcp

transgene could always interact with endogenous Mcp detected by 3C techniques. It will provide very useful information by quantifying the 3C interactions between the Mcp-101 transgenes and the endogenous Mcp, to see the 3C interaction changes when the transgenes is in active states and in high frequency colocalization, and when the transgene is in silence states. Another experiment that could be done to see where the transgene go, is using one transgene with enhancer (active state) and one transgene without enhancer (possibley silenced state), and then quantify the interaction frequency and immune-stain the cells to see each transgene location.

Based on this insulator structure model, the gene will join one of the existing bodies based on the signal they received. Then what happened to the genes if they either not receive active nor repressive signals? It is possible that the genes are likely to interact with their neighouring genes and join one of the closet bodies nearby. So what kind of signals that direct the insulators into specific bodies inside the nuclei? How the insulator drags the genes into the bodies? Considering the involvement of PcG proteins and TRX protein, it is possible that the epigenetic marks are parts of the signaling system. The gene may then be brought to one of its cognate body by the actin-motor system. This might explain the previously observed phenomena that the gene will join one of transcription factories when it is activated and get out of its chromosome territories.

The more interesting questions is what are genes inside a specific body and what the insulator elements that bring genes into one special body. Are the genes inside one transcription factoy all transcriptionally active and co-regualted? Are the genes inside one PcG body all silenced and Pc targets? Do the insulators associated with those genes inside the same body bind to the same set of insulator proteins? In other words, do they belong to the same class of insulators? All those interesting questions could be possibly answered with the 4C techniques that apply on proper tissues with various controls.

3.5.5 The role of TRE/TRX

One of very interesting conclusion in the current research is that TRX/TRE is required for the high frequency co-localization in the eye disc cells. The high frequency interaction between two transgenes is very sensitive to the dosage of TRX. But how TRX contribute to the interaction is elusive. The first questions that is interesting to ask is whether the higher interaction between the remote *Antp* and *Abd-B* genes in the embryo head cells also depend on the TRX function. Or, is the TRX only required for the higher frequency interaction only in the case where both genes are active as in the case of transgenes? TRX constituively bind to the PRE/TRE irrespective of the transcription states, and ASH1 is associated with TRX and TRE when the target gene is active, and both proteins are methyltransferase that modify chromatin and deposit active epigenetic marks. So, it is also interesting to test the effect of *ash1* mutation on the high frequency colocalization between two Mcp-101 lines in the eye disc cells.

There are could be several ways that TRX contribute to the high-frequency colocalization by interacting with other proteins/pathways. First, RNAi machinery proteins, such as AGO1, PIWI1, PIWI2, are reported to be required for the long-distance interaction between Fab-7 transgenes and between Gypsy transgenes (Grimaud et al., 2006; Lei et al. 2006). It is possible that TRX interacts with the RANi proteins, or TRX is involved in the small RNA production process. Second, Cohesin proteins has been reported to associate with insulator protein CTCF, and both may contribute to the formation of enhancer-promoter loops. So it is possible that TRX may interact with Cohesin proteins and involve in the long distance DNA interaction. This possibility become increasingly interesting with the recent publication from Paro and colleages (Strubbe et al., 2011), which showed that PcG proteins could be co-purified with Cohesin and TrxG proteins, and Cohesin is important for the Pairing Sensitive Silencing phenotype. They do not explain the correlation between the Cohesin an TrxG proteins, which could fit in my insulator model, Cohesin could contribute together with the insulator and TRX to the formation of both silence body and transcription factories. Another experiment that could be done to see the TRX interactors is pulling down the transgenes and associated proteins using the antibody to GFP-LacI, which specifically bind to the tagging 128 copies of LacO repeats.

3.5.6 An alternative model

While the insulator structural model all base on the *Mcp* transgenes, it is possible that this model may not apply on other insulators other than *Mcp*. There are could be some

unknown factors that bind to *Mcp* sequence, which contribute the specific properties of mediating long-range DNA co-localization. The first indentified insulators, *scs* and *scs'*, were shown to have no such long-range interaction properties, but they could form local loops (Blanton et al., 2003). 4C results using insulator bind sites showed mainly local interactions (unpublished observations). So, some, if not most, insulators mainly involved in local loop formation, while for *Mcp* and *Fab-7*, they acquired some specific elements that enable them to reach far.

In conclusion, the data described here implicate that insulator elements could bring active enhancer (genes) into transcription factories, and regulate gene expression by organizing chromatin into active zones and silenced zones.

Chapter 4. Summary and Conclusions

Insulator is a relatively new class of DNA elements comparing to enhancer, promoter and silencer, and little is known regarding its properties and functions. *Fab-7* and *Mcp*, the boundarires of cis-regulatory domains of the bithorax complex in *Drosophila*, each contain two functional parts: the silencer (Polycomb Response Element) and the insulator. By using genetics functional assay, previous data in the lab showed that the *Mcp* insulator is essential for trans interaction while neither the powerful silencer *bxd* PRE nor the *Mcp* PRE can mediate such interactions. To further prove that it is the insulator part, not the PRE, of *Mcp* and *Fab-7* mediate the long-range interactions between Polycomb targets, I made new transgenes with the insulator part and PRE part of *Mcp* and *Fab-7* each flanked by LOX and FRT elements, and employed the technologies of Chromosome Conformation Capture (3C) and in-vivo live-imaging. The results showed that the *Mcp* and *Fab-7* transgenes could interact with their endogenous partners, and interact with each other at low frequency (~6%). Deletion of the PRE part does not influence the frequency of the co-localization, while deletion of the insulator part abrogates the interaction. I continue to show that the co-localization between two transgenes actually do not depend on their endogenous partners, since the deletion of endogenous partner does not affect the frequency of long-range interactions. The ChIP results showed that *Mcp* and *Fab-7* both bind the dCTCF and dCP190 insulator proteins, and in-vivo imaging results showed that *Mcp* transgene could also interact with *Fab-7* transgene at long distance, which indicate that *Mcp* and *Fab-7* may belong to the same class of insulator

elements. I also showed that Polycomb protein binds to the transgene, and this binding is lost after the deletion of the PRE fragments. Overall, I proved that it is insulator, but not PRE, that have the ability to mediate long-distance interactions between Polycomb targets, and we propose that insulator bring the Polycomb target genes into the PcG bodies, thus enhance silencing.

Vazquez et al. (2006) used similar in-vivo imaging systems and showed that *Mcp* transgenes could trans interact at a frequency of up to 90% of the eye disc cells, which is in conflict with our observation of 6% interaction. I repeat his experiments and found that the high frequency interaction is limited in the eye disc cells, while the wing disc and membrane cells have only 10% interaction. I compare the constructs used in their experiments and our previous experiments, and found it is likely because tha they used an additional *white* eye enhancer element, which is active only in the eye disc cells. Therefore, I hypothesize that the enhancer could promote the long-range interaction mediated by insulator, which means insulator may involve in transcription. To prove this hypothesis, I made similar transgenes with *Mcp* and *white* eye enhancer. The live-imaging results show that the transgenes interact with each other at high frequency (from 50% to 80%) in the eye disc cells, but low (4~10%) at the wing disc and membrane disc cells, which is in agreement with the results of Vazquez et al. (2006). Deletion of the eye enhancer brings the frequency of interaction down to 20% in the eye disc cells, but do not influence the frequency in the membrane and wing disc cells. The frequency in eye disc cells is still higher than

that in wing cells, which could be due to the fact that *white* gene is still transcribed in eye disc cells but not in wing disc cells. After deletion of the *Mcp* part, or both *Mcp* and eye enhancer in the transgene, the frequency of the interaction just drop down to background level, which is due to the absence of insulator, since the mutation of insulator protein dCTCF also abrogate the trans interactions between the intact transgenes. Since the trans interaction is mediated by the insulator part, not the PRE part, of *Mcp*, I reason that the PRE part of *Mcp* may not be required for the high frequency interaction in the eye disc cells. To test this hypothesis, I construct new transgenes with only the insulator part of *Mcp* and *white* eye enhancer. Surprisingly, the results showed only low frequency (~6%) interaction both in eye and wing disc cells, which suggest that the PRE part of the *Mcp* is also required for the high frequency interaction in the eye disc cells. In order to understand the function of the PRE, I returned to the 820-bp *Mcp* lines, and tested them in a *Polycomb* mutant background. Reducing the Pc dosage to half the normal level does not affect colocalization. Since Trithorax protein (TRX) binds constitutively to all known or putative PREs (therefore also called TREs) and antagonize PcG repression (Schwartz & Pirrotta, 2007; 2008), I then test the co-localization frequency under heterozygous trx condition, and found that reducing the level of TRX brings the high interaction in the eye disc cells down to basic level, the same as that found in the membrane and wing disc cells. To test if the high frequency interaction is eye enhancer specific, I generated new transgenic fly lines with the UAS-Gal enhancer and 820-bp *Mcp*. After activation by the *Arm*-Gal4 driver, the interaction frequency in both the eye and wing

disc cells significantly increased up to 25%, which is not as high as eye enhancer transgene, may be due to the fact that Arm-Gal4 is not as strong enhancer as the white eye enhancer. So the frequency of co-localization is related to the level of transcription activity rather than some unknown factor that binds specifically to the eye enhancer.

In conclusion, my results showed insulator not only partner with silencers to bring genes into silence bodies, but also partner with enhancers to possibly bring active genes into transcription factories. Our works supports the dual structural role model of insulators in gene regulation (Fig. 3.4C): insulator alone only mediate basic low interaction, provide all kinds of interaction possibilities for other DNA elements; Once the gene receive repressive signals, the PcG proteins are recruited to the PRE, then the insulator will bring this gene into one of the cognate pre-exist PcG bodies; While when the gene receive active signals, the transcription factors will be recruited to the enhancers, the TRX will bind to the PRE and antagonize the PcG and possibly modify the chromatins, then the insulators will bring this poised gene into a nearby transcription factory. I envision that the PcG bodies and transcription factories may be located in the interchromatin compartment, a continuous space between chromosomal territories. The spaces where the PcG bodies reside represent the silenced zone and memorize the genes inside the bodies even after cell cycles, while the transcription factories reside in the active zone. The insulator elements may participate in the

formation of those bodies/zones and help genes to arrive those different bodies/zones, depending on the epigenetic marks and the signals received.

Appendices

Appendix 1. The *Mcp* and *Fab-7* Sequence used in various transgenic flies

***Mcp* sequence**

Chr3R:

Pst I *Sal* I

12694579 ctgcagacttaaattgatttaaagtttgctgccttttcaacgacagttcaaatgcaaattggctggtcgaccggccgttttccgttttattgcgaatatt 12694678

12694679 aaatgaaattaatgaaattttctgcgccataatcctttgcaaaacgcataaatttgctcattaagtgtgcgcaaatattgtatgtatccgctccgctaaa 12694778

Mcp^{340}

12694779 aggtctatatactttatatacttgtattgatttttaagctcagataaataagctcagagtacataagcgacgcccaaaaagcccaaatgtagagcttttt 12694878

Mcp^{210}

12694879 cgaaattaaacagaaagtcgggtctgcaaataagggcttttctggggaagaaataaattatatcttaataaatatatttt***aaacttaactcagacttaga*** 12694978

12694979 ***tttattttatcacttatttttaagtgatttaaataatttaaaaatttatttgttacataaatttagccaatatccaaacctttgcgctggcgccccctа*** 12695078

Mcp PRE

12695079 ***ttgtttttctttgcagcttatgctttgctgacaacccaccagaggacgctcgctgttggaaacgcattacgcacacttacaacgcttggg***ttttctcatg 12695178

12695179 tgttagtgcgtgagagtaagtgagacaacaggcttattgatgtagtcttcctccttacacataatacatggccgcgcgacaaagatggcaacattgatgg 12695278

Pst I

12695279 ctgcctctgaaaacatggcctctttttccgacattgtatctgtgtgacgtttgactgcagatgcgtttgtgtgggtagtaaatgtatcttctgcgtttaa 12695378

Xba I

12695379 gtcgattttgtcaactaaatttgcgctttgttacccctgaaaatgggagctcatgcgcagtatgcagctggtgcggaattttctctaga 12695466

***Fab*-7 Sequence**

Chr3R:

Fab-7 Insulator

12724265 actgcagtgaagacacgaaccccaaggacgcatttccaattgggaaagaaacccattggtgcagactttgttcaacattgttgttgagccgtgcgattgc 12724364

12724365 cccaatcattcttatcagcaaaaagcagagctgtgccattgtttgatattttgccacccacaatgcatccaactttgttgccaagtgagcgaaaaactta 12724464

12724465 ttatatttcgcccgcacaatcccctcaaaaaatgaatgcaagcccaaaaaaacaaaaaaaaaaaaaagacgagaaaaagaacaggacgagtggcaaaagct 12724564

12724565 ggcaaagcagcaaaaatcgtaaaaaagaaaattgcatttccccaaagcagcgaaacttgcgcaggacttttgagattctattaaattctaacaagatttc 12724664

12724665 aagctgtgtggcgggggggaagaggaagagagcggaaagtgcagcgcccaataagcaaatggcagctgtcacggggaagcacagagagtgcagaaaggggga 12724764

12724765 aaaaacattggggcatatcaacgcgccaaaaagaaaaacaaaaagagcgaggtagaatgtcgctcaaagagcgacacgtgaacaggtgcagtagtaaata 12724864

12724865 taagcaaagagagttggaaagagtattggctaagagcgaccgctcactaacacatagataaattaagagagacgtgataagagaaccgcacgcacaccac 12724964

12724965 cgcaaaatccaattggaagagagcgactgcttgaatgtattggtaagcaagagagcggctaggtttgatggtttgattggaattcagttgccgttcgaaa 12725064

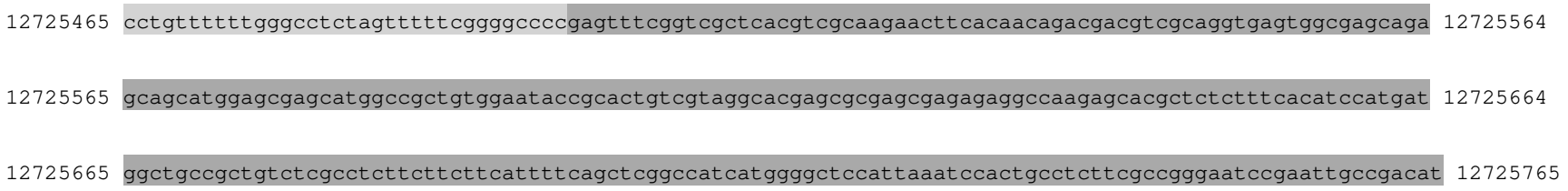
12725065 tatttttgataaaatataaaaaattaattcagaggctgaggcaagtctaaaaacaatgctttgcctaagaattcgtagcttttataatttttttttttaat 12725164

12725165 tttgaaattagcattttatttttaaatgattctccaattaagccaactggtttccaactctagcggtgaccctcacctttggtttgcgtaccgactaa 12725264

12725265 gtccgagcagtgctgcgcatccctttgagccttagtatacccatctcgctcttagccacccctaaataccgttacttaccctgggcaacttccttcgtc 12725364

12725365 cgtcggcctttgtttctgcatttttttttgtttttgtctgggcgacgacgcagtcgcagaaagtccctcgaaattcctccgctccctcgctcgctcacaat 12725464

Fab-7 PRE


12725465 cctgttttttgggcctctagttttcggggccccgagtttcggtcgctcacgtcgcaagaacttcacaacagacgacgtcgcaggtgagtggcgagcaga 12725564
12725565 gcagcatggagcgagcatggccgctgtggaataccgcactgtcgtaggcacgagcgcgagcgagagaggccaagagcacgctctctttcacatccatgat 12725664
12725665 ggctgccgctgtctcgcctcttcttcttcattttcagctcggccatcatggggctccattaaatccactgcctcttcgccgggaatccgaattgccgacat 12725765

Appendix Table 2. Co-localization of *Mcp* transgenes

Fly Cross	One-Dot Cells / Total Cells	Percentage	Comparison	P-value
a: M16-M25	20 / 319	6.30%		
b: M16ΔP-M25ΔP	20 / 346	5.78%	b vs. a	0.7909
c: M16ΔI -M25ΔI	2 / 1381	0.14%	c vs. a	<.0001
a: M16; M31	21 / 269	7.81%		
b: M16ΔP; M31ΔP	104 / 1533	6. 80%	b vs. a	0.5426
c: M16ΔI; M31ΔI	1 / 765	0.13%	c vs. a	<.0001
a: M25; M31	32 / 412	7.77%		
b: M25ΔP; M31ΔP	108 / 1663	6.50%	b vs. a	0.3565
c: M25ΔI; M31ΔI	0 / 1179	0%	c vs. a	<.0001

For each transgene combination, the fraction and percentage of nuclei showing co-localisation is listed. The significance of the colocalisation is calculated by a Chi square test of the comparison indicated and the corresponding P-value is given.

Appendix Table 3. Co-localization of *Fab-7* transgenes

Fly Cross	One-Dot Cells / Total Cells	Percentage	Comparison	P-value
a: F4; F9	29 / 499	5.81%		
b: F4ΔP; F9ΔP -Fab7[1]	99 / 1502	6.60%	b vs. a	0.5375
c: F4ΔI; F9ΔI	4 / 1693	0.24%	c vs. a	<.0001
a: F12; F4	151 / 1978	7.63%		
b: F12ΔP; F4ΔP	53 / 813	6.52%	b vs. a	0.3039
c: F12ΔI; F4ΔI	3 / 1295	0.23%	c vs. a	<.0001
a: F12; F9	61 / 882	6.92%		
b: F12ΔP;F9ΔP	39 / 562	6.94%	b vs. a	0.9864
c: F12ΔP; F9ΔP-Fab7[1]	124 / 1878	6.60%	c vs. a	0.7589
d: F12ΔI; F9ΔI	5 / 1289	0.39%	d vs. a	<.0001

For each transgene combination, the fraction and percentage of nuclei showing co-localisation is listed. The significance of the colocalisation is calculated by a Chi square test of the comparison indicated and the corresponding P-value is given. Two of the transgene combinations were tested also in a genetic background deleted for the endogenous *Fab-7*. The results show that the endogenous element is not required for interactions between transgenes.

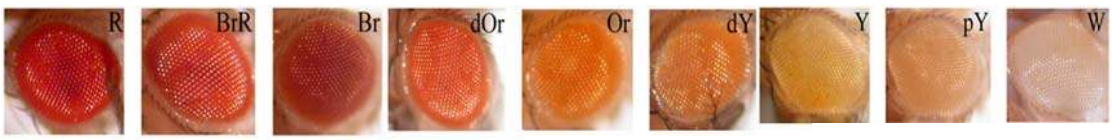
Appendix Table 4. Co-localization of *Mcp* with *Fab-7*.

Fly Cross	One-Dot Cells / Total Cells	Percentage	Comparison	P-value
a: F9-M31	51/991	5.15%	a vs. c	<0.0001
b: F9ΔP-M31ΔP	66/1519	4.34%	b vs. c	<0.0001
c: F9ΔI-M31ΔI	3/1027	0.29%		
a: F9-M31	51/991	5.15%	a vs. d	<0.0001
b: F9ΔP-M31ΔP	66/1519	4.34%	b vs. d	<0.0001
d: F9ΔPΔI-M31ΔPΔI	4/894	0.45%		

For each transgene combination, the fraction and percentage of nuclei showing co-localisation is listed. The significance of the colocalisation is calculated by a Chi square test of the comparison indicated and the corresponding P-value is given.

Appendix Table 5. Score the eye color of the *LacO-Mcp* and *LacO-Fab-7* lines and their derivatives.

	R	BrR	Br	dOr	Or	dY	Y	pY	W
M16						×			
M16△P				×					
M16△I							×		
M16△P△I							×		
M25									×
M25△P								×	
M25△I									×
M25△P△I									×
M31							×		
M31△P				×					
M31△I						×			
M31△P△I						×			
F4					×				
F4△P			×						
F4△I						×			
F4△P△I							×		
F9							×		
F9△P					×				
F9△I									×
F9△P△I								×	
F12				×					
F12△P				×					
F12△I							×		
F12△P△I						×			



The color gradation from white to red is indicated by W, pY, Y, dY, Or, dOr, Br, BrR, R (Gruzdeva, 2005). The score is based on the Homozygous male, and no paring sensitive silencing was observed for all lines.

Appendix Table 6. Co-localization of all the Mcp-101 and drivatives.

	Intact		ΔE		ΔM		ΔMΔE	
	Eye	Wing	Eye	Wing	Eye	Wing	Eye	Wing
101B4-B15	85.9%	10.3%	51.1%	10.7%	14.9%	12.8%	12.4%	9.1%
101B4-B19	76.4%	8.5%	24.6%	10.9%	1.4%	0.9%	0.19%	0.35%
101B15-B19	70%	6.1%	11.3%	5.4%	1.8%	1.1%	0.34%	0.21%
101A8-A22	71.3%	6.5%	13.8%	5%	2.1%	1.9%	nd	nd
101B19-A22	49.4%	4.1%	26.9%	5.2%	1.7%	1.4%	1.24%	0.78%

For each transgene combination, the fraction and percentage of nuclei showing co-localization is listed. ΔE denote eye enhancer deleted, ΔM denote *Mcp* deleted, while ΔMΔE denote both *Mcp* and enhancer deleted. 'Eye' denote that the the numbers were counted in the eye disc cells which were highlighted in red, while 'Wing' denote the numbers were counted in the wing and membrane cells. 'nd' denote not done.

Reference

BANERJI, J. , Rusconi, S., Schaffner, W. (1981). Expression of a β -globin gene is enhanced by remote SV40 DNA sequences. Cell 27, 299-308.

Bantignies, F., Grimaud, C., Lavrov, S., Gabut, M., and Cavalli, G. (2003). Inheritance of Polycomb-dependent chromosomal interactions in Drosophila. Genes & development 17, 2406-20.

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, 214-26.

Barges, S., Mihaly, J., Galloni, M, Hagstrom, K, Müller, M, Shanower, G., Schedl, P., Gyurkovics, H, and Karch, F. (2000). The Fab-8 boundary defines the distal limit of the bithorax complex iab-7 domain and insulates iab-7 from initiation elements and a PRE in the adjacent iab-8 domain. Development (Cambridge, England) 127, 779-90.

Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell 129, 823-37.

Becker, M., Baumann, C., John, S., Walker, D.A., Vigneron, M., McNally, J.G., and Hager, G.L. (2002). Dynamic behavior of transcription factors on a natural promoter in living cells. EMBO reports 3, 1188-94.

Bell, A.C., West, A.G., and Felsenfeld, G. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. Cell 98, 387-396.

Bellés, X. (2010). Beyond Drosophila: RNAi in vivo and functional genomics in insects. Annual review of entomology 55, 111-28.

Belmont, A.S., and Straight, A.F. (1998). In vivo visualization of chromosomes using lac operator-repressor binding. Trends in Cell Biology 8, 121-124.

Blackwood, E.M. ; Kadonaga, J.T. (1998). Going the Distance: A Current View of Enhancer Action. Science 281, 60-63.

Blanton, J., Gaszner, M., and Schedl, P. (2003). Protein:protein interactions and the pairing of boundary elements in vivo. Genes & development 17, 664-75.

Blow, M.J., McCulley, D.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., Afzal, V., Bristow, J., Ren, B., Black,

B.L., Rubin, E.M., Visel, A., and Pennacchio, L.A. (2010). ChIP-Seq identification of weakly conserved heart enhancers. Nature Genetics 42, 806-10.

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

Brown, J.L., Mucci, D., Whiteley, M., Dirksen, M.L., and 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, 1057-64.

Brown, J.L., Fritsch, C., Mueller, J., and Kassis, J.A. (2003). The Drosophila pho-like gene encodes a YY1-related DNA binding protein that is redundant with pleiohomeotic in homeotic gene silencing. Development (Cambridge, England) 130, 285-94.

Buchenau, P., Hodgson, J., Strutt, H., and Arndt-Jovin, D. J. (1998). The Distribution of Polycomb-Group Proteins During Cell Division and Development in Drosophila Embryos: Impact on Models for Silencing. J. Cell Biol. 141, 469-481.

Bushey, A.M., Ramos, E., and Corces, V.G. (2009). Three subclasses of a Drosophila insulator show distinct and cell type-specific genomic distributions. Genes & development 23, 1338-50.

Busturia, A., Lloyd, A., Bejarano, F., Zavortink, M., Xin, H., and Sakonju, S. (2001). The MCP silencer of the Drosophila Abd-B gene requires both Pleiohomeotic and GAGA factor for the maintenance of repression. Development (Cambridge, England) 128, 2163-73.

Byrd, K., and Corces, V.G. (2003). Visualization of chromatin domains created by the gypsy insulator of Drosophila. The Journal of cell biology 162, 565-74.

Cai, H.N., and Shen, P. (2001). Effects of cis arrangement of chromatin insulators on enhancer-blocking activity. Science (New York, N.Y.) 291, 493-5.

Cao, R., Tsukada, Y.-I., and Zhang, Y. (2005). Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. Molecular cell 20, 845-54.

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 (New York, N.Y.) 298, 1039-43.

Capelson, M., and Corces, V.G. (2004). Boundary elements and nuclear organization. Biol. Cell 96, 617-629.

Capelson, M., and Corces, V.G. (2005). The ubiquitin ligase dTopors directs the nuclear organization of a chromatin insulator. Mol. Cell 20, 105-116.

Capelson, M., and Corces, V.G. (2006). SUMO conjugation attenuates the activity of the gypsy chromatin insulator. The EMBO journal 25, 1906-14.

Carter, D.R.F., Eskiw, C., and Cook, P.R. (2008). Transcription factories. Biochemical Society transactions 36, 585-9.

Chakalova, L., Debrand, E., Mitchell, J.A., Osborne, C.S., and Fraser, P. (2005). Replication and transcription: shaping the landscape of the genome. Nat. Rev. Genet. 6, 669-677.

Chakalova, L., and Fraser, P. (2010). Organization of Transcription. Cold Spring Harbor perspectives in biology 2, a000729.

Chambeyron, S., and Bickmore, W.A. (2004). Does looping and clustering in the nucleus regulate gene expression. Curr. Opin. Cell Biol. 16, 256-262.

Chambeyron, S., and Bickmore, W.A. (2004). Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. Genes & development 18, 1119-30.

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, 2553-64.

Cheutin, T., McNairn, A.J., Jenuwein, T., Gilbert, D.M., Singh, P.B., and Misteli, T. (2003). Maintenance of stable heterochromatin domains by dynamic HP1 binding. Science (New York, N.Y.) 299, 721-5.

Chinwalla, V., Jane, E.P., and Harte, P.J. (1995). The Drosophila trithorax protein binds to specific chromosomal sites and is co-localized with Polycomb at many sites. The EMBO journal 14, 2056-65.

Chung, J.H., Whiteley, M., and Felsenfeld, G. (1993). A 5’ element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila. Cell 74, 505-14.

Chung, J.H. (1997). Characterization of the chicken beta -globin insulator. Proceedings of the National Academy of Sciences 94, 575-580.

Cook, P.R. (1994). RNA polymerase: structural determinant of the chromatin loop and the chromosome. BioEssays : news and reviews in molecular, cellular and developmental biology 16, 425-30.

Cook, P.R. (1999). The Organization of Replication and Transcription. Science 284, 1790-1795.

Cook, P.R. (2010). A model for all genomes: the role of transcription factories. Journal of molecular biology 395, 1-10.

Cremer, T., and Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat. Rev. Genet 2, 292-301.

Crickmore, M.A., Ranade, V., and Mann, R.S. (2009). Regulation of Ubx expression by epigenetic enhancer silencing in response to Ubx levels and genetic variation. PLoS genetics 5, e1000633.

Czermin, B., Melfi, R., McCabe, Donna, Seitz, V., Imhof, A., and Pirrotta, V. (2002). Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. Cell 111, 185-96.

Dekker, J. (2006). The three “C” s of chromosome conformation capture: controls, controls, controls. Nature methods 3, 17-21.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science (New York, N.Y.) 295, 1306-11.

Dietzel, S., Niemann, H., Brückner, B., Maurange, C., and Paro, R. (1999). The nuclear distribution of Polycomb during Drosophila melanogaster development shown with a GFP fusion protein. Chromosoma 108, 83-94.

Dostie, J., Richmond, T.A., Arnaout, R.A., Selzer, R.R., Lee, W.L., Honan, T.A., Rubio, E.D., Krumm, A., Lamb, J., Nusbaum, C., Green, R.D., and Dekker, J. (2006). Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. Genome research 16, 1299-309.

Duncan, I.W. (2002). Transvection effects in Drosophila. Annual review of genetics 36, 521-56.

Déjardin, J., and Cavalli, G. (2004). Chromatin inheritance upon Zeste-mediated Brahma recruitment at a minimal cellular memory module. The EMBO journal 23, 857-68.

Déjardin, J., Rappailles, A., Cuvier, O., Grimaud, C., Decoville, M., Locker, D., and Cavalli, G. (2005). Recruitment of Drosophila Polycomb group proteins to chromatin by DSP1. Nature 434, 533-8.

Faro-Trindade, I., and Cook, P. (2006). Mechanisms of Gene Regulation Mechanisms of Gene Regulation Transcription factories : structures conserved during differentiation and evolution. Society, 23-27.

Festenstein, R., Pagakis, S. N., Hiragami, K., Lyon, D., Verreault, A., Sekkali, B., and Kioussis, D. (2003). Modulation of heterochromatin protein 1 dynamics in primary Mammalian cells. Science (New York, N.Y.) 299, 719-21.

Ficz, G., Heintzmann, R., and Arndt-Jovin, D.J. (2005). Polycomb group protein complexes exchange rapidly in living Drosophila. Development (Cambridge, England) 132, 3963-76.

Fischer, J.A., Giniger, E., Maniatis, T., and Ptashne, M. (1988). GAL4 activates transcription in Drosophila. Nature 332, 853-6.

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 & development 17, 1870-81.

Francis, N.J., Kingston, R.E., and Woodcock, C.L. (2004). Chromatin compaction by a polycomb group protein complex. Science (New York, N.Y.) 306, 1574-7.

Gall, J.G., and Pardue, M.L. (1969). Formation and Detection of RNA-DNA Hybrid Molecules in Cytological Preparations. Proceedings of the National Academy of Sciences 63, 378-383.

Gao, X., Yang, L., Petros, J.A., Marshall, F.F., Simons, J.W., and Nie, S. (2005). In vivo molecular and cellular imaging with quantum dots. Current opinion in biotechnology 16, 63-72.

Gaszner, M., and Felsenfeld, G. (2006). Insulators: exploiting transcriptional and epigenetic mechanisms. Nat. Rev. Genet 7, 703-713.

Gaszner, M., Vazquez, J., and Schedl, P. (1999). The Zw5 protein, a component of the scs chromatin domain boundary, is able to block enhancer-promoter interaction. Genes & Development 13, 2098-2107.

Gerasimova, T.I., Byrd, K., and Corces, V.G. (2000). A chromatin insulator determines the nuclear localization of DNA. Mol. Cell 6, 1025-1035.

Gerasimova, T.I., and Corces, V.G. (1998). Polycomb and trithorax group proteins mediate the function of a chromatin insulator. Cell 92, 511-21.

Gerasimova, T.I, Gdula, D.A., Gerasimov, D.V., Simonova, O., and Corces, V.G. (1995). A Drosophila protein that imparts directionality on a chromatin insulator is an enhancer of position-effect variegation. Cell 82, 587-97.

Gerasimova, T.I., Lei, E.P., Bushey, A.M., and Corces, V.G. (2007). Coordinated control of dCTCF and gypsy chromatin insulators in Drosophila. Molecular cell 28, 761-72.

Geyer, P.K., and Clark, I. (2002). Protecting against promiscuity: the regulatory role of insulators. Cellular and Molecular Life Sciences (CMLS) 59, 2112-2127.

Ghosh, D., Gerasimova, T.I., and Corces, V.G. (2001). Interactions between the Su(Hw) and Mod(mdg4) proteins required for gypsy insulator function. EMBO J 20, 2518-2527.

Gohl, D., Müller, M., Pirrotta, V., Affolter, M., and Schedl, P. (2008). Enhancer blocking and transvection at the Drosophila apterous locus. Genetics 178, 127-43.

Grewal, S.I.S., and Elgin, S.C.R. (2007). Transcription and RNA interference in the formation of heterochromatin. Nature 447, 399-406.

Grimaud, C., Bantignies, F., Pal-Bhadra, M., Ghana, P., Bhadra, U., and Cavalli, G. (2006). RNAi components are required for nuclear clustering of Polycomb group response elements. Cell 124, 957-71.

Gruzdeva, N., Kyrchanova, O., Parshikov, A., Kullyev, A., and Georgiev, P. (2005). The Mcp element from the bithorax complex contains an insulator that is capable of pairwise interactions and can facilitate enhancer-promoter communication. Molecular and cellular biology 25, 3682-9.

Gurudatta, B.V., and Corces, V.G. (2009). Chromatin insulators: lessons from the fly. Briefings in functional genomics & proteomics 8, 276-82.

Gyurkovics, H., Gausz, J., Kummer, J., and Karch, F. (1990). A new homeotic mutation in the Drosophila bithorax complex removes a boundary separating two domains of regulation. The EMBO journal 9, 2579-85.

Hagstrom, K., Muller, M., and Schedl, P. (1996). Fab-7 functions as a chromatin domain boundary to ensure proper segment specification by the Drosophila bithorax complex. Genes & development 10, 3202-15.

Hagstrom, K., Muller, M., and Schedl, P. (1997). A Polycomb and GAGA Dependent Silencer Adjoins the Fab-7 Boundary in the Drosophila Bithorax Complex. Genetics 146, 1365-1380.

Hagège, H., Klous, P., Braem, C., Splinter, E., Dekker, J., Cathala, G., de Laat, W., and Forné, T. (2007). Quantitative analysis of chromosome conformation capture assays (3C-qPCR). Nature protocols 2, 1722-33.

Hart, C.M., Zhao, K., and Laemmli, U.K. (1997). The scs' boundary element: characterization of boundary element-associated factors. Molecular and cellular biology 17, 999-1009.

Heard, E., and Bickmore, W. (2007). The ins and outs of gene regulation and chromosome territory organisation. Current Opinion in Cell Biology, 2-7.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., Wang, W., Weng, Z., Green, R.D., Crawford, G.E., Ren, B. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nature genetics 39, 311-8.

Holohan, E.E., Kwong, C., Adryan, B., Bartkuhn, M., Herold, M., Renkawitz, R., Russell, S., and White, R. (2007). CTCF genomic binding sites in Drosophila and the organisation of the bithorax complex. PLoS genetics 3, e112.

Horard, B., Tatout, C., Poux, S., and Pirrotta, V. (2000). Structure of a Polycomb Response Element and In Vitro Binding of Polycomb Group Complexes Containing GAGA Factor. Molecular and Cellular Biology 20, 3187-3197.

Horike, S., Cai, S., Miyano, M., Cheng, J.-F., and Kohwi-Shigematsu, T. (2005). Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. Nature genetics 37, 31-40.

Iborra, F.J., Pombo, A., Jackson, D.A., and Cook, P.R. (1996). Active RNA polymerases are localized within discrete transcription "factories' in human nuclei. Journal of cell science 109 (Pt 6, 1427-36.

Jack, J., Dorsett, D., Delotto, Y., and Liu, S. (1991). Expression of the cut locus in the Drosophila wing margin is required for cell type specification and is regulated by a distant enhancer. Development (Cambridge, England) 113, 735-47.

Jackson, D.A., Hassan, A.B., Errington, R.J., and Cook, P.R. (1993). Visualization of focal sites of transcription within human nuclei. The EMBO journal 12, 1059-65.

Kahn, T.G., Schwartz, Y.B., Dellino, G.I., and Pirrotta, V. (2006). Polycomb complexes and the propagation of the methylation mark at the Drosophila ubx gene. The Journal of biological chemistry 281, 29064-75.

Karch, F, Weiffenbach, B., Peifer, M., Bender, W., Duncan, I., Celniker, S., Crosby, M., and Lewis, E B (1985). The abdominal region of the bithorax complex. Cell 43, 81-96.

Karch, F., Galloni, M., Sipos, L., Gausz, J., Gyurkovics, H., and Sched, P. (1994). Mcp and Fab -7: molecular analysis of putative boundaries of cis-regulatory domains in the bithorax complex of Drosophila melanogaster. Nucleic Acids Research 22, 3138-3146.

Kassis, J.A. (2002). Pairing-sensitive silencing, polycomb group response elements, and transposon homing in Drosophila. Advances in genetics 46, 421-38.

Kellum, R., and Schedl, P. (1992). A group of scs elements function as domain boundaries in an enhancer-blocking assay. Molecular and cellular biology 12, 2424-31.

Kellum, R., and Schedl, P. (1991). A position-effect assay for boundaries of higher order chromosomal domains. Cell 64, 941-950.

Khoury, G., and Gruss, P. (1983). Enhancer elements. Cell 33, 313-4.

Kim, T.H., Abdullaev, Z.K., Smith, A.D., Ching, K.A., Loukinov, D.I., Green, R.D., Zhang, M.Q., Lobanenkov, V.V., and Ren, B. (2007). Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell 128, 1231-45.

Kimura, H., Tao, Y., Roeder, R.G., and Cook, P.R. (1999). Quantitation of RNA polymerase II and its transcription factors in an HeLa cell: little soluble holoenzyme but significant amounts of polymerases attached to the nuclear substructure. Molecular and cellular biology 19, 5383-92.

Kimura, H., Sugaya, K., and Cook, P.R. (2002). The transcription cycle of RNA polymerase II in living cells. The Journal of cell biology 159, 777-82.

Klymenko, T., and Müller, J. (2004). The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. EMBO reports 5, 373-7.

Klymenko, T., Papp, B., Fischle, W., Köcher, T., Schelder, M., Fritsch, C., Wild, B., Wilm, M., and Müller, J. (2006). A Polycomb group protein complex with sequence-specific DNA-binding and selective methyl-lysine-binding activities. Genes & development 20, 1110-22.

Kravchenko, E., Savitskaya, E., Kravchuk, O., Parshikov, A., Georgiev, P., and Savitsky, M. (2005). Pairing between gypsy insulators facilitates the enhancer

action in trans throughout the Drosophila genome. Molecular and cellular biology 25, 9283-91.

Kuhn, E., and Geyer, P.K. (2003). Genomic insulators: connecting properties to mechanism. Current Opinion in Cell Biology 15, 259-265.

Kuhn, E.J., Viering, M.M., Rhodes, K.M., and Geyer, P.K. (2003). A test of insulator interactions in Drosophila. The EMBO journal 22, 2463-71.

Kumaran, R.I., Thakar, R., and Spector, D.L. (2008). Chromatin dynamics and gene positioning. Cell 132, 929-34.

Kurukuti, S., Tiwari, V.K., Tavoosidana, G., Pugacheva, E., Murrell, A., Zhao, Z., Lobanenkov, V., Reik, W., and Ohlsson, R. (2006). CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. Proceedings of the National Academy of Sciences of the United States of America 103, 10684-9.

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 & development 16, 2893-905.

Kyrchanova, O., Toshchakov, S., Parshikov, A., and Georgiev, P. (2007). Study of the functional interaction between Mcp insulators from the Drosophila bithorax complex: effects of insulator pairing on enhancer-promoter communication. Molecular and cellular biology 27, 3035-43.

Lei, E.P., and Corces, V.G. (2006). RNA interference machinery influences the nuclear organization of a chromatin insulator. Nature genetics 38, 936-41.

Levine, M. (2010). Transcriptional enhancers in animal development and evolution. Current biology : CB 20, R754-63.

Levine, S.S., King, I.F.G., and Kingston, R.E. (2004). Division of labor in polycomb group repression. Trends in biochemical sciences 29, 478-85.

Lewis, E.B. (1998). The bithorax complex: the first fifty years. The International journal of developmental biology 42, 403-15.

Lewis, E.B. (1954). The Theory and Application of a New Method of Detecting Chromosomal Rearrangements in Drosophila melanogaster. The American Naturalist 88, 225-239.

Li, H.-B., Muller, M., Bahechar, I., Kyrchanova, O., Ohno, K., Georgiev, P., and

Pirrotta, V. (2011) Insulators, not Polycomb complexes, are required for long-range interactions between Polycomb targets in Drosophila melanogaster. Mol. Cell. Biol., 31(4):616-25. Epub 2010 Dec 6.

Ling, J.Q., Li, T., Hu, J.F., Vu, T.H., Chen, H.L., Qiu, X.W., Cherry, A.M., and Hoffman, A.R. (2006). CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. Science (New York, N.Y.) 312, 269-72.

Lomvardas, S., Barnea, G., Pisapia, D.J., Mendelsohn, M., Kirkland, J., and Axel, R. (2006). Interchromosomal interactions and olfactory receptor choice. Cell 126, 403-13.

Lunyak, V.V. (2008). Boundaries. Boundaries...Boundaries??? Current opinion in cell biology 20, 281-7.

Maeda, R.K., and Karch, F. (2006). The ABC of the BX-C: the bithorax complex explained. Development (Cambridge, England) 133, 1413-22.

Majumder, P., and Cai, H.N. (2003). The functional analysis of insulator interactions in the Drosophila embryo. Proceedings of the National Academy of Sciences of the United States of America 100, 5223-8.

Melnikova, L., Juge, F., Gruzdeva, N., Mazur, A., Cavalli, G., and Georgiev, P. (2004). Interaction between the GAGA factor and Mod(mdg4) proteins promotes insulator bypass in Drosophila. Proceedings of the National Academy of Sciences of the United States of America 101, 14806-11.

Michalet, X., Pinaud, F.F., Bentolila, L.A., Tsay, J.M., Doose, S., Li, J.J., Sundaresan, G., Wu, A.M., Gambhir, S.S., and Weiss, S. (2005). Quantum dots for live cells, in vivo imaging, and diagnostics. Science (New York, N.Y.) 307, 538-44.

Mihaly, J., Hogga, I., Gausz, J., Gyurkovics, H., and 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 (Cambridge, England) 124, 1809-20.

Mishra, R.K., Mihaly, J., Barges, S., Spierer, A., Karch, F., Hagstrom, K., Schweinsberg, S.E., and Schedl, P. (2001a). The iab-7 polycomb response element maps to a nucleosome-free region of chromatin and requires both GAGA and pleiohomeotic for silencing activity. Molecular and cellular biology 21, 1311-8.

Mitchell, J.A., and Fraser, P. (2008). Transcription factories are nuclear subcompartments that remain in the absence of transcription. Genes & development 22, 20-5.

Mohan, M., Bartkuhn, M., Herold, M., Philippen, A., Heinl, N., Bardenhagen, I., Leers, J., White, R.A., Renkawitz-Pohl, R., Saumweber, H., and Renkawitz, R. (2007). The Drosophila insulator proteins CTCF and CP190 link enhancer blocking to body patterning. The EMBO journal 26, 4203-14.

Mongelard, F., and Corces, V.G. (2001). Two insulators are not better than one. Nat. Structure Biol. 8, 192-194.

Moon, H., Filippova, G., Loukinov, D., Pugacheva, E., Chen, Q., Smith, S.T., Munhall, A., Grewe, B., Bartkuhn, M., Arnold, R., Burke, L.J., Renkawitz-Pohl, R., Ohlsson, R., Zhou, J., Renkawitz, R., Lobanenkov, V. (2005). CTCF is conserved from Drosophila to humans and confers enhancer blocking of the Fab-8 insulator. EMBO reports 6, 165-70.

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

Muravyova, E., Golovnin, A., Gracheva, E., Parshikov, A., Belenkaya, T., Pirrotta, V., and Georgiev, P. (2001). Loss of insulator activity by paired Su(Hw) chromatin insulators. Science (New York, N.Y.) 291, 495-8.

Murrell, A., Heeson, S., and Reik, W. (2004). Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. Nature genetics 36, 889-93.

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., and Simon, J.A. (2002). Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. Cell 111, 197-208.

de Napoles, M., Mermoud, J.E., Wakao, R., Tang, Y.A., Endoh, M., Appanah, R., Nesterova, T.B., Silva, J., Otte, A.P., Vidal, M., Koseki, H., and Brockdorff, N. (2004). Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. Developmental cell 7, 663-76.

Nègre, N., Brown, C.D., Shah, P.K., Kheradpour, P., Morrison, C.A., Henikoff, J.G., Feng, X., Ahmad, K., Russell, S., White, R.A., Stein, L., Henikoff, S., Kellis, M., and White, K.P. (2010). A comprehensive map of insulator elements for the Drosophila genome. PLoS genetics 6, e1000814.

Nègre, N., Hennetin, J., Sun, L.V., Lavrov, S., Bellis, M., White, K.P., and Cavalli, G. (2006). Chromosomal distribution of PcG proteins during Drosophila development. PLoS biology 4, e170.

Osborne, C.S., Chakalova, L., Brown, K.E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J.A., Lopes, S., Reik, W., and Fraser, P. (2004). Active genes dynamically colocalize to shared sites of ongoing transcription. Nature genetics 36, 1065-71.

Osborne, C.S., Chakalova, L., Mitchell, J.A., Horton, A., Wood, A.L., Bolland, D.J., Corcoran, A.E., and Fraser, P. (2007). Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. PLoS biology 5, e192.

Pai, C.-Y., Lei, E.P., Ghosh, D., and Corces, V.G. (2004). The centrosomal protein CP190 is a component of the gypsy chromatin insulator. Molecular cell 16, 737-48.

Pal-Bhadra, M., Leibovitch, B.A., Gandhi, S.G., Rao, M., Bhadra, U., Birchler, J.A., and Elgin, S.C.R. (2004). Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. Science (New York, N.Y.) 303, 669-72.

Papantonis, A., Larkin, J.D., Wada, Y., Ohta, Y., Ihara, S., Kodama, T., and Cook, P.R. (2010). Active RNA Polymerases: Mobile or Immobile Molecular Machines? PLoS Biology 8, e1000419.

Papp, B., and Müller, J. (2006). Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins. Genes & development 20, 2041-54.

Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H.C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T., Cobb, B.S., Yokomori, K., Dillon, N., Aragon, L., Fisher, A.G., and Merkenschlager, M. (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. Cell 132, 422-33.

Pennacchio, L.A., Ahituv, N., Moses, A.M., Prabhakar, S., Nobrega, M.A., Shoukry, M., Minovitsky, S., Dubchak, I., Holt, A., Lewis, K.D., Plajzer-Frick, I., Akiyama, J., De Val, S., Afzal, V., Black, B.L., Couronne, O., Eisen, M.B., Visel, A., and Rubin, E.M. (2006). In vivo enhancer analysis of human conserved non-coding sequences. Nature 444, 499-502.

Pirrotta, V. (1998). Polycombing the genome: PcG, trxG, and chromatin silencing. Cell 93, 333-6.

Pirrotta, V., and Rastelli, L. (1994). White gene expression, repressive chromatin domains and homeotic gene regulation in Drosophila. BioEssays : news and reviews in molecular, cellular and developmental biology 16, 549-56.

Pirrotta, V., Chan, C.S., McCabe, D., and Qian, S. (1995). Distinct Parasegmental and Imaginal Enhancers and the Establishment of the Expression Pattern of the Ubx Gene. Genetics 141, 1439-1450.

Pombo, A., Cuello, P., Schul, W., Yoon, J.B., Roeder, R.G., Cook, P.R., and Murphy, S. (1998). Regional and temporal specialization in the nucleus: a transcriptionally-active nuclear domain rich in PTF, Oct1 and PIKA antigens associates with specific chromosomes early in the cell cycle. The EMBO journal 17, 1768-78.

Pombo, A., Jackson, D.A., Hollinshead, M., Wang, Z., Roeder, R.G., and Cook, P.R. (1999). Regional specialization in human nuclei: visualization of discrete sites of transcription by RNA polymerase III. The EMBO journal 18, 2241-53.

Poux, S., Kostic, C., and Pirrotta, V. (1996). Hunchback-independent silencing of late Ubx enhancers by a Polycomb Group Response Element. The EMBO journal 15, 4713-22.

Poux, S., Horard, B., Sigrist, C.J.A., and Pirrotta, V. (2002). The Drosophila trithorax protein is a coactivator required to prevent re-establishment of polycomb silencing. Development (Cambridge, England) 129, 2483-93.

Qian, S., Capovilla, M., and Pirrotta, V. (1993). Molecular mechanisms of pattern formation by the BRE enhancer of the Ubx gene. The EMBO journal 12, 3865-77.

Qian, S., Capovilla, M., and Pirrotta, V. (1991). The bx region enhancer, a distant cis-control element of the Drosophila Ubx gene and its regulation by hunchback and other segmentation genes. The EMBO journal 10, 1415-25.

Qian, S., Varjavand, B., and Pirrotta, V. (1992). Molecular Analysis of the zeste-white Interaction Reveals a Promoter-Proximal Element Essential for Distant Enhancer-Promoter Communication. Genetics 131, 79-90.

Ramos, E., Ghosh, D., Baxter, E., and Corces, V.G. (2006). Genomic organization of gypsy chromatin insulators in Drosophila melanogaster. Genetics 172, 2337-49.

Robinett, C.C., Straight, A., Li, G., Willhelm, C., Sudlow, G., Murray, A., and Belmont, A.S. (1996). In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. The Journal of cell biology 135, 1685-700.

Robinett, C., O'Connor, A., and Dunaway, M. (1997). The repeat organizer, a specialized insulator element within the intergenic spacer of the Xenopus rRNA genes. Mol. Cell. Biol. 17, 2866-2875.

Rodin, S., Kyrchanova, O., Pomerantseva, E., Parshikov, A., and Georgiev, P. (2007). New properties of Drosophila fab-7 insulator. Genetics 177, 113-21.

Saurin, A.J., Shao, Z., Erdjument-Bromage, H., Tempst, P., and Kingston, R.E. (2001). A Drosophila Polycomb group complex includes Zeste and dTAFII proteins. Nature 412, 655-60.

Saurin, A.J., Shiels, C., Williamson, J., Satijn, D.P.E., Otte, A.P., Sheer, D., and Freemont, P.S. (1998). The Human Polycomb Group Complex Associates with Pericentromeric Heterochromatin to Form a Novel Nuclear Domain. J. Cell Biol. 142, 887-898.

Schoborg, T.A., and Labrador, M. (2010). The phylogenetic distribution of non-CTCF insulator proteins is limited to insects and reveals that BEAF-32 is Drosophila lineage specific. Journal of molecular evolution 70, 74-84.

Schoenfelder, S., Sexton, T., Chakalova, L., Cope, N.F., Horton, A., Andrews, S., Kurukuti, S., Mitchell, J.A., Umlauf, D., Dimitrova, D.S., Eskiw, C.H., Luo, Y., Wei, C.L., Ruan, Y., Bieker, J.J., and Fraser, P. (2009). Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. Nature Genetics 42, 53-61.

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, 700-5.

Schwartz, Y.B., and Pirrotta, V. (2007). Polycomb silencing mechanisms and the management of genomic programmes. Nature reviews. Genetics 8, 9-22.

Schwartz, Y.B., and Pirrotta, V. (2008). Polycomb complexes and epigenetic states. Current opinion in cell biology 20, 266-73.

Schweinsberg, S., Hagstrom, K., Gohl, D., Schedl, P., Kumar, R.P., Mishra, R., and Karch, F. (2004). The enhancer-blocking activity of the Fab-7 boundary from the Drosophila bithorax complex requires GAGA-factor-binding sites. Genetics 168, 1371-84.

Shao, Z., Raible, F., Mollaaghababa, R., Guyon, J.R., Wu, C.T., Bender, W., and Kingston, R.E. (1999). Stabilization of chromatin structure by PRC1, a Polycomb complex. Cell 98, 37-46.

Siegal, M. L., and Hartl, D.L. (1996). Transgene Coplacement and High Efficiency Site-Specific Recombination With the Cre/loxP System in Drosophila. Genetics 144, 715-726.

Sigrist, C.J., and Pirrotta, V. (1997). Chromatin insulator elements block the silencing of a target gene by the Drosophila polycomb response element (PRE) but allow trans interactions between PREs on different chromosomes. Genetics 147, 209-21.

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., Wit, E. de, Steensel, Bas van, and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nature genetics 38, 1348-54.

Simonis, M., Kooren, J., and de Laat, W. (2007). An evaluation of 3C-based methods to capture DNA interactions. Nature Methods 4, 895-901.

Spana, C., Harrison, D.A., and Corces, V.G. (1988). The Drosophila melanogaster suppressor of Hairy-wing protein binds to specific sequences of the gypsy retrotransposon. Genes & Development 2, 1414-1423.

Spector, D.L. (2001). Nuclear domains. J. Cell Sci. 114, 2891-2893.

Splinter, E., Grosveld, F., and de Laat, W. (2004). 3C technology: Analyzing the spatial organization of genomic loci in vivo. Methods Enzymol 375, 493-507.

Spradling, A., and Rubin, G. (1982). Transposition of cloned P elements into Drosophila germ line chromosomes. Science 218, 341-347.

Straight, A.F., Belmont, A.S., Robinett, C.C., and Murray, A.W. (1996). GFP tagging of budding yeast chromosomes reveals that protein-protein interactions can mediate sister chromatid cohesion. Curr. Biol 6, 1599-1608.

Strubbe, G., Popp, C., Schmidt, A., Pauli, A., Ringrose, L., Beisel, C., and Paro, R. (2011). Polycomb purification by in vivo biotinylation tagging reveals cohesion and Trithorax group proteins as interaction partners. Proc Natl Acad Sci USA Mar 17, Epub ahead of print.

Surface, L.E., Thornton, S.R., and Boyer, L.A. (2010). Polycomb Group Proteins Set the Stage for Early Lineage Commitment. Cell stem cell 7, 288-298.

Sutherland, H., and Bickmore, W.A. (2009). Transcription factories: gene expression in unions? Nature reviews. Genetics 10, 457-66.

Taddei, A., Hediger, F., Neumann, F.R., and Gasser, S.M. (2004). The function of nuclear architecture: a genetic approach. Annual review of genetics 38, 305-45.

Thakar, R., and Csink, A.K. (2005). Changing chromatin dynamics and nuclear organization during differentiation in Drosophila larval tissue. Journal of cell science 118, 951-60.

Thakar, R., Gordon, G., and Csink, A.K. (2006). Dynamics and anchoring of heterochromatic loci during development. Journal of cell science 119, 4165-75.

Tolhuis, B., Palstra, R.-J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and Interaction between Hypersensitive Sites in the Active β-globin Locus. Molecular Cell 10, 1453-1465.

Tolhuis, B., de Wit, E., Muijrers, I., Teunissen, H., Talhout, W., van Steensel, B., and van Lohuizen, M. (2006). Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in Drosophila melanogaster. Nature genetics 38, 694-9.

Udvardy, A., Maine, E., and Schedl, P. (1985). The 87A7 chromomere: Identification of novel chromatin structures flanking the heat shock locus that may define the boundaries of higher order domains. Journal of Molecular Biology 185, 341-358.

Valenzuela, L., and Kamakaka, R.T. (2006). Chromatin insulators. Annual review of genetics 40, 107-38.

Vazquez, J., Mu, M., Pirrotta, V., and Sedat, J.W. (2006). The Mcp Element Mediates Stable Long-Range Chromosome – Chromosome Interactions in Drosophila □. Molecular Biology of the Cell 17, 2158 -2165.

Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., Afzal, V., Ren, B., Rubin, E.M., and Pennacchio, L.A. (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature 457, 854-8.

Visel, A., Prabhakar, S., Akiyama, J.A., Shoukry, M., Lewis, K.D., Holt, A., Plajzer-Frick, I., Afzal, V., Rubin, E.M., and Pennacchio, L.A. (2008). Ultraconservation identifies a small subset of extremely constrained developmental enhancers. Nature genetics 40, 158-60.

Volpi, E.V., Chevret, E., Jones, T., Vatcheva, R., Williamson, J., Beck, S., Campbell, R.D., Goldsworthy, M., Powis, S.H., Ragoussis, J., Trowsdale, J., and Sheer, D. (2000). Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei. J. Cell Sci. 113, 1565-1576.

Wallace, H.A., Plata, M.P., Kang, H.-J., Ross, M., and Labrador, M. (2010). Chromatin insulators specifically associate with different levels of higher-order chromatin organization in Drosophila. Chromosoma 119, 177-94.

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, 873-8.

Wansink, D.G, Schul, W., Kraan, I. van der, Steensel, B van, Driel, R. van, and Jong, L. de (1993). Fluorescent labeling of nascent RNA reveals transcription by RNA polymerase II in domains scattered throughout the nucleus. The Journal of cell biology 122, 283-93.

Wansink, D.G. (1993). Fluorescent labeling of nascent RNA reveals transcription by RNA polymerase II in domains scattered throughout the nucleus. The Journal of Cell Biology 122, 283-293.

Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T., Yahata, K., Imamoto, F., Aburatani, H., Nakao, M., Imamoto, N., Maeshima, K., Shirahige, K., and Peters, J.M. (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature 451, 796-801.

West, A.G, Gaszner, M., and Felsenfeld, G. (2002). Insulators: many functions, many mechanisms. Genes & development 16, 271-88.

Wiblin, A.E., Cui, W., Clark, A.J., and Bickmore, W.A. (2005). Distinctive nuclear organisation of centromeres and regions involved in pluripotency in human embryonic stem cells. Journal of cell science 118, 3861-8.

Winoto, A., and Baltimore, D. (1989). A novel, inducible and T cell-specific enhancer located at the 3' end of the T cell receptor alpha locus. The EMBO journal 8, 729-33.

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

Würtele, H., and Chartrand, P. (2006). Genome-wide scanning of HoxB1-associated loci in mouse ES cells using an open-ended Chromosome Conformation Capture methodology. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology 14, 477-95.

Xi, H., Shulha, H.P., Lin, J.M., Vales, T.R., Fu, Y., Bodine, D.M., McKay, R.D., Chenoweth, J.G., Tesar, P.J., Furey, T.S., Ren, B., Weng, Z., and Crawford, G.E. (2007). Identification and characterization of cell type-specific and ubiquitous chromatin regulatory structures in the human genome. PLoS genetics 3, e136.

Xie, X., Mikkelsen, T.S., Gnirke, A., Lindblad-Toh, K., Kellis, M., and Lander, E.S. (2007). Systematic discovery of regulatory motifs in conserved regions of the human genome, including thousands of CTCF insulator sites. Proceedings of the National Academy of Sciences of the United States of America 104, 7145-50.

Zeitlinger, J., Zinzen, R.P., Stark, A., Kellis, M., Zhang, H., Young, R.A., and Levine, M. (2007). Whole-genome ChIP-chip analysis of Dorsal, Twist, and Snail suggests integration of diverse patterning processes in the Drosophila embryo. Genes & development 21, 385-90.

Zhao, K. (1995). Visualization of chromosomal domains with boundary element-associated factor BEAF-32. Cell 81, 879-889.

Zhao, Z., Tavoosidana, G., Sjölinder, M., Göndör, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Sandhu, K.S., Singh, U., Pant, V., Tiwari, V., Kurukuti, S., and Ohlsson, R. (2006). Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nature genetics 38, 1341-7.

Zhong, X.-P. (1997). An enhancer-blocking element between alpha and delta gene segments within the human T cell receptor alpha /delta locus. Proceedings of the National Academy of Sciences 94, 5219-5224.

Zhou, G.-L., Xin, L., Song, W., Di, L.-J., Liu, G., Wu, X.-S., Liu, D.-P., and Liang, C.-C. (2006). Active chromatin hub of the mouse alpha-globin locus forms in a transcription factory of clustered housekeeping genes. Molecular and cellular biology 26, 5096-105.

Zhou, J, Barolo, S., Szymanski, P., and Levine, M (1996). The Fab-7 element of the bithorax complex attenuates enhancer-promoter interactions in the Drosophila embryo. Genes & Development 10, 3195-3201.

Zink, D., and Paro, R. (1995). Drosophila Polycomb-group regulated chromatin inhibits the accessibility of a trans-activator to its target DNA. The EMBO journal 14, 5660-71.

Curriculum Vitae

Huabing Li

EDUCATION

2005 – 2011 Doctor of Philosophy, Biochemistry
Rutgers, the State University of New Jersey and University of Medicine and Dentistry of New Jersey, New Brunswick/Piscataway, NJ, USA

2002 – 2005 Master of Science, Genetics
Nankai University, Tianjin, China

1999 – 2002 Bachelor of Science, Biological Sciences
Nankai University, Tianjin, China

PUBLICATIONS

Li, H-B., Muller, M., Bahechar, I., Kyrchanova, O., Ohno, K., Georgiev, P., and Pirrotta, V. (2011) Insulators, not Polycomb complexes, are required for long-range interactions between Polycomb targets in Drosophila melanogaster. Mol. Cell. Biol., 31(4):616-25. Epub 2010 Dec 6.