ABSTRACT

Title of dissertation: RNA SILENCING AND HIGHER ORDER CHROMATIN ORGANIZATION IN DROSOPHILA

Nellie Moshkovich, Doctor of Philosophy, 2011

Dissertation directed by: Dr. Elissa P. Lei
NIH, NIDDK, LCDB
David A. O'Brochta
University of Maryland, Department of Entomology

Higher order chromatin organization influences gene expression, but mechanisms by which this phenomenon occurs are not well understood. RNA silencing, a conserved mechanism that involves small RNAs bound to an Argonaute protein, mediates gene expression via transcriptional or post-transcriptional regulation. Recently, a role for RNA silencing in chromatin has been emerging. In fission yeast, a major role of RNA

interference (RNAi) is to establish pericentromeric heterochromatin. However, whether this mechanism is conserved throughout evolution is unclear. In *Drosophila*, a powerful model organism, there are multiple functionally distinct RNA silencing pathways. Previous studies have suggested the involvement of the Piwi-interacting RNA (piRNA) and endogenous small interfering RNA (endo-siRNA) pathways in heterochromatin formation in order to silence transposable elements in germline and somatic tissues, respectively, but direct evidence is lacking. We addressed whether the genomic locations generating these small RNAs may act as AGO-dependent platforms for heterochromatin recruitment. Our genetic and biochemical analyses revealed that heterochromatin is nucleated independently of endo-siRNA and piRNA pathways suggesting that RNAi-dependent heterochromatin assembly may not be conserved in metazoans.

Chromatin insulators are regulatory elements characterized by enhancer blocking and barrier activity. Insulators form large nuclear foci termed insulator bodies that are tethered to the nuclear matrix and have been proposed to organize the genome into distinct transcriptional domains by looping out intervening DNA. In *Drosophila*, RNA silencing has been reported to affect nuclear organization of *gypsy* insulator complexes and formation of Polycomb repression bodies. Our studies revealed that AGO2 is required for CTCF/CP190-dependent *Fab-8* insulator function independent of its catalytic activity or Dicer-2. Moreover, AGO2 associates with euchromatin but not heterochromatin genome-wide. Also, AGO2 associates physically with CP190 and CTCF, and mutation of *CTCF*, *CP190*, or *AGO2* decreases chromosomal looping interactions and alters gene expression. We propose a novel RNAi-independent role for AGO2 in the nucleus. We postulate that insulator proteins recruit AGO2 to chromatin to

promote or stabilize chromosomal interactions crucial for proper gene expression. Overall, our findings demonstrate novel mechanisms by which RNA silencing affects gene expression on the level of higher order chromatin organization.

RNA SILENCING AND HIGHER ORDER CHROMATIN ORGANIZATION IN DROSOPHILA

by

Nellie Moshkovich

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2011

Advisory Committee:

Professor David A. O'Brochta, Chair
Dr. Elissa P. Lei, Co-chair
Professor Leslie Pick
Professor Stephen M. Mount
Professor Jonathan D. Dinman
Professor Judd O. Nelson

DEDICATION

I would like to dedicate this work to my family; to my mother, Galina, for teaching me the value of hard work, determination and perseverance, and for instilling in me the inspiration to set high goals and the confidence to achieve them; to my father, Yuriy, for emphasizing to me the importance of science and encouraging me in my academic pursuits; to my husband, Igor, my emotional anchor, for his cool head and warm heart; and to my daughter, Nikita, my little bookworm, so that she may be proud of me as much as I am proud of her.

ACKNOWLEDGEMENTS

My deepest gratitude goes to my advisor, Elissa P. Lei. I have been fortunate to have an advisor who not only provided me with rigorous scientific training but also cared for all my successes and failures as if they were her own. Thank you for sharing your enthusiasm and dedication to science with me, it was really inspiring. Thank you also for leading by example, holding me to a high research standard, and for taking a chance on a rather unconventional graduate student.

I would also like to acknowledge the members of my dissertation committee, David A. O'Brochta, Leslie Pick, Stephen M. Mount and Jonathan D. Dinman for their guidance, encouragement and constructive criticism throughout the years. I would like to thank the members of Lei lab, past and present, for their indispensable assistance in my research and writing efforts over the last four years. I would like in particular to acknowledge the colleagues I collaborated with - Patrick Boyle, for his many empirical and intellectual contributions, the many valuable discussions, and his friendship; Parul Nisha, for laying the foundation for some of my research, her support and invaluable advice on graduate school and science; Ryan Dale, for unlimited computational support and sharing the magic of computational biology with me; Leah Matzat, for insightful discussions, reading of numerous drafts, her infectious enthusiasm and camaraderie. I would also like to thank Brandi Thompson, Matthew Emmett, Patrick Murphy and Madoka Chinen. I extend my appreciation to Erik Baehrecke for jump starting my graduate career and for teaching me the basics of fly husbandry. Many thanks go to my friends and colleagues from Baehrecke lab, who made the first two years of graduate

school more bearable, Christina Kary McPhee, Michelle Beaucher, Deb Berry, Jahda Hill, Yakup Batlevi and Sudeshna Dutta.

Finally, this journey would not have been possible if it were not for my parents. It is through their personal sacrifice that I had the luxury to strive for academic excellence and I am indebted to them. I also would like to thank my husband for keeping everything in perspective for me and for his love and care. Last but not least, I acknowledge my daughter who has grown into a wonderful human being while her mom was occasionally absent chasing laurels.

TABLE OF CONTENTS

LIST OF TABLES ..... **viii**

LIST OF FIGURES ..... **ix**

LIST OF ABBREVIATIONS ..... **xi**

CHAPTER 1. Introduction **1**

RNA silencing **1**
Transposable element silencing by the piRNA and endo-siRNA pathways **3**
A role for RNA silencing in the nucleus **7**

Heterochromatin and RNA silencing **8**
RNAi-mediated heterochromatic silencing in *S. pombe* **9**
RNA silencing involvement in heterochromatin recruitment in metazoans **11**

Polycomb-mediated gene repression and RNA silencing **13**

Chromatin insulators, nuclear organization and RNA silencing **14**

Conclusion **20**

CHAPTER 2. HP1 recruitment in the absence of Argonaute proteins in *Drosophila* **23**

Abstract **23**

Introduction **25**

Results **29**
Heterochromatin dependent transcriptional silencing at piRNA clusters **29**
piRNA and endo-siRNA pathway mutants decrease transcription at piRNA clusters **37**
HP1 chromatin association is increased at piRNA clusters in somatic tissues of RNA silencing mutants **40**
HP1 also associates with piRNA clusters in ovaries **48**
HP1 chromatin association is not affected greatly by depletion of Piwi in somatic ovarian follicle cells **50**
Loss of piRNA production from a single cluster results in global HP1 mislocalization **53**
Mutation of the *flam* piRNA cluster suppresses heterochromatic silencing at a distant site **58**

Discussion **59**

AGO2 and Piwi are not required for HP1 association at piRNA clusters 59
Additional candidate platforms for Piwi-dependent HP1 recruitment 63
Functions for *piwi* outside of the gonad 63
HP1 redistribution in piRNA pathway mutants 64
Conclusions 65

Acknowledgements 67
Author contributions 68

Materials and Methods 69

CHAPTER 3. RNAi-independent role for Argonaute2 in CTCF/CP190 chromatin insulator function 77

Abstract 77

Introduction 78

Results 82
AGO2 associates with euchromatin and not repetitive sequences 82
AGO2 colocalizes with chromatin insulator sites throughout the genome 94
AGO2 associates with active promoters 95
AGO2 chromatin association does not correspond to regions of the genome that produce endo-siRNA 96
AGO2 binds to PREs and overlaps extensively with TrxG and PcG proteins 100
AGO2 opposes Polycomb function 104
AGO2 but not its catalytic activity is specifically required for *Fab-8* insulator activity 109
AGO2 interacts physically with CTCF and CP190 114
AGO2 chromatin association requires CP190 and CTCF 118
AGO2 associates with chromatin downstream of CTCF and CP190 122
CP190, CTCF and AGO2 are required for looping at the *Abd-B* locus 125
AGO2 is required for proper expression of *Abd-B* similar to CTCF 126

Discussion 127
AGO2 localizes predominantly to euchromatin and not heterochromatin 127
RNAi-independent function for AGO2 at chromatin 128
Role of AGO2 in *Fab-8* insulator function 128
Role of AGO2 in long range chromosomal interactions 130
Conclusions 131

Acknowledgements 134

Author contributions 135

Materials and Methods 136

CHAPTER 4. Discussion and future directions ……. 154

RNA silencing and its role in heterochromatin assembly …… **154**
piRNA- and endo-siRNA-mediated TE silencing ……... **154**
Alternative mechanisms for heterochromatin nucleation …… **156**

RNA silencing and its effects on chromatin insulators ……. **157**
AGO2: a multifunctional protein ……. **157**
Role for AGO2 in transcriptional regulation ……... **159**
A connection between two functions of AGO2 ……... **161**
AGO2 role in nuclear organization …….. **163**

Conclusion…….. **164**

REFERENCES …… 166

LIST OF TABLES

Table 2-1. Expression of mini-*white* in fly lines harboring P element insertions in four top piRNA clusters **…… 32**

Table 2-2. Primer set sequences used for ChIP at the *flam* piRNA cluster **……. 75**

Table 3-1. Sources of data for tiling arrays, endo-siRNA, and genome features **… 150**

Table 3-2. List of primers **…… 152**

LIST OF FIGURES

Figure 1-1. *Drosophila* piRNA and siRNA pathways. **……... 6**

Figure 1-2. Model for RNAi-mediated centromeric silencing in *S. pombe*. **…….. 10**

Figure 1-3. Functional properties of chromatin insulators. **…… 15**

Figure 1-4. Nuclear organization of a chromatin insulator. **…… 17**

Figure 1-5. *Abd-B cis*-regulatory region of the bithorax complex (BX-C). **…… 19**

Figure 2-1. Schematic representation of four top piRNA clusters. **……. 31**

Figure 2-2. piRNA and endo-siRNA pathway mutants display increased silencing of transcriptional reporters at or near piRNA clusters. **…….. 35**

Figure 2-3. *Dcr-2* mutants display increased HP1 chromatin association and increased silencing at piRNA clusters. **……. 38**

Figure 2-4. HP1 associates with chromatin at piRNA clusters, and its levels increase in RNA silencing mutants. **……..…… 42**

Figure 2-5. Su(Hw) does not associate with chromatin at piRNA clusters in heads. **……. 43**

Figure 2-6. HP1 chromatin association levels are increased in *piwi* mutants at piRNA clusters. **……..… 44**

Figure 2-7. HP1 chromatin association levels are increased in *AGO2* mutants at piRNA clusters. **……..…… 45**

Figure 2-8. HP1 protein levels in wildtype, $flam^{1}$, $AGO2^{51B}$ and $piwi^{1}/piwi^{2}$ fly heads. **……..….. 46**

Figure 2-9. HP1 associates with chromatin at piRNA clusters in ovaries. **……..… 48**

Figure 2-10. Depletion of Piwi from ovarian somatic follicle cells does not affect HP1 recruitment to piRNA clusters. **……..… 51**

Figure 2-11. Mutation of the *flam* piRNA cluster results in global HP1 redistribution. **…… 55**

Figure 2-12. $spn\text{-}E^{hls\text{-}E1}/spn\text{-}E^{hls\text{-}E616}$ mutants display accumulation of HP1 at the chromocenter. **…… 56**

Figure 2-13. ChIP primer efficiency and specificity. **……..…** **74**

Figure 3-1. ChIP-seq profiles of AGO2 in S2 and S3 cells at BX-C. **……** **85**

Figure 3-2. Overlap between genome-wide binding sites of AGO2, insulator, TrxG/PcG, transcription related factors, and promoters. **…….** **92**

Figure 3-3. Distribution of endo-siRNA densitities for AGO2 binding sites compared to 3' *cis*-NATs and transcriptionally active or inactive regions. **…….** **99**

Figure 3-4. AGO2 behaves as a TrxG protein. **……...** **103**

Figure 3-5. Effects of *Pc* and *TRX* knock-down on AGO2 association with chromatin. **……..…** **108**

Figure 3-6. AGO2 but not its catalytic activity is required for *Fab-8* insulator function. **……** **113**

Figure 3-7. AGO2 associates physically with CP190. **…….** **117**

Figure 3-8. AGO2 is required for looping interactions throughout the *Abd-B* locus and proper gene expression. **…….** **124**

Figure 3-9. Model for AGO2 function with respect to CTCF/CP190 chromatin insulator activity. **……...** **133**

LIST OF ABBREVIATIONS

Abd-B	*Abdominal-B*
AGO	Argonaute
Aub	Aubergine
BX-C	*Bithorax* complex
CP190	Centrosomal protein 190
3C	Chromosome conformation capture
ChIP	Chromatin immunoprecipitation
CTCF	CCCTC-binding factor
Dcr	Dicer
dsRNA	Double stranded RNA
endo-siRNA	Endogenous siRNA
flam	*flamenco*
Fab-8	*Frontabdominal-8*
GAF	GAGA factor
HP1	Heterochromatin protein 1
IP	Immunoprecipitation
KD	Knock down
Mod(mdg4)2.2	Modifier of mdg4 2.2
NAT	Natural antisense transcript
OSC	Ovarian somatic cell
piRNA	Piwi-interacting RNA
PEV	Position-effect variegation
PcG	Polycomb group
PRE	Polycomb response element
RNAi	RNA interference
RITS	RNA-induced transcriptional silencing complex
siRNA	Short interfering RNA
Su(Hw)	Suppressor of Hairy wing
Su(var)	*Suppressor of variegation*
TSS	Transcription start site
TE	Transposable element
TrxG	Trithorax group
WT	wildtype

CHAPTER 1

INTRODUCTION

Eukaryotic genomes are organized into a complex system where physical interactions between genes and regulatory elements ensure proper regulation of gene expression. It has become apparent that nuclear organization and chromatin folding are vital for spatial and temporal control necessary for proper gene expression during development and differentiation. The mechanisms regulating nuclear organization still need to be addressed. Gaining mechanistic insight into the function of key regulators mediating higher order chromatin organization as well as elucidating new contributors may aid in understanding abnormal biological processes such as tumorigenesis in which gene expression is altered. Recently, RNA silencing has been emerging as one such contributing factor.

RNA SILENCING

RNA silencing pathways are evolutionarily conserved mechanisms that control gene expression via sequence-specific interactions mediated by a small RNA bound to an Argonaute (AGO) effector protein and are involved in numerous biological processes such as post-transcriptional gene regulation, defense against transposable elements (TEs) and pathogens, and chromatin organization. RNA silencing pathways are characterized by the activity of an Argonaute effector protein that binds small RNAs directly. Ranging

from one AGO gene in *Schizosaccharomyces pombe* (*S. pombe*) to 27 in *Caenorhabditis elegans* (*C. elegans*), the number varies greatly among species. The five Argonautes in *Drosophila melanogaster* can be divided into two families based on homology. The AGO subfamily includes AGO1 and AGO2, and the Piwi subfamily consists of Piwi, Aubergine (Aub), and AGO3 (reviewed in Hutvagner and Simard 2008). *Drosophila* AGO1 and AGO2 are expressed ubiquitously (Williams and Rubin 2002; Rehwinkel et al. 2006). AGO1 is primarily required for the microRNA pathway, which regulates mRNA expression and functions chiefly through mRNA destabilization as well as translational repression (reviewed in Czech and Hannon 2011). Existing as a mechanism to protect against exogenous double stranded RNA (dsRNA), AGO2 associates with 21-22 nt short interfering RNA (siRNA) produced by Dicer-2 (Dcr-2) in a pathway that is required for viral immunity and a robust RNAi response (Hammond et al. 2001; Wang and Ligoxygakis 2006). In addition, AGO2 binds endogenous siRNAs (endo-siRNAs) that mediate TE silencing in somatic tissues (Chung et al. 2008; Czech et al. 2008; Ghildiyal et al. 2008; Kawamura et al. 2008).

The Piwi clade of AGO proteins mediates TE silencing in the gonad. The expression of *piwi*, *aub*, and *AGO3* is mainly, although not exclusively, in the gonad (Williams and Rubin 2002; Saito et al. 2006; Brennecke et al. 2007; Brower-Toland et al. 2007). Whereas Piwi is a nuclear protein that is detected in both germ and somatic cells in *Drosophila* ovaries, Aub and AGO3 localize to the cytoplasm and accumulate in the nuage, an electron-dense perinuclear structure associated with germline RNA processing (Cox et al. 2000; Saito et al. 2006; Brennecke et al. 2007; Gunawardane et al. 2007; Nishida et al. 2007). The Piwi proteins associate with a class of small RNAs termed

Piwi-interacting RNAs (piRNAs) involved in the repression of TEs in the germline (Saito et al. 2006; Brennecke et al. 2007; Nishida et al. 2007; Yin and Lin 2007).

All AGO proteins are characterized by three functional domains crucial for their interaction with small RNAs: the PAZ, Mid and PIWI domains. The PAZ and Mid domains bind small RNAs (Song et al. 2003; Ma et al. 2004). The PIWI domain is a catalytically active RNase H-like domain that cleaves the targeted RNA molecule (Song et al. 2004; Ma et al. 2005). The endonuclease activity of AGO proteins is often referred to as Slicer activity. Slicer activity of certain *Drosophila* AGO proteins has been shown to be vital for biological functions such as small RNA-based TE and pathogen defense. The role of AGO proteins in post-transcriptional gene silencing is rather well-characterized. Other aspects of AGO function such as its roles in transcription and chromatin organization or yet unknown functions will be elucidated in the future.

Transposable element silencing by the piRNA and endo-siRNA pathways

Eukaryotic genomes are beset with TEs, mobile genetic elements, transposition and recombination of which can cause genetic instability leading to deleterious mutations. Suppression of TEs is especially imperative in the gonad in order to limit the propagation of unwanted mutations and is achieved principally by the activity of the Piwi subfamily of AGO proteins. In *Drosophila*, Piwi, Aub, and AGO3 bind to 23-30 nt piRNAs that are predominantly derived from genomic locations termed piRNA clusters (Brennecke et al. 2007; Yin and Lin 2007). These piRNA producing loci are mainly pericentromeric and enriched in transposon sequences. Based on comparative sequence

analysis of piRNAs immunopurified from the ovary, the “ping-pong” or “amplification loop” model for germline piRNA biogenesis was proposed (Figure1-1; Brennecke et al. 2007; Gunawardane et al. 2007). Precursor transcripts from piRNA clusters, derived from either one or both strands, give rise to piRNAs bound by Piwi, Aub, or AGO3. Those piRNAs antisense to a homologous TE transcript can result in its cleavage, and this event defines the 5’ end of a secondary piRNA that can then bind and cleave an antisense piRNA cluster transcript, and the cycle can continue. Piwi appears to play a minor role in ping-pong piRNA amplification (Li et al. 2009a; Malone et al. 2009), which is thought to occur primarily in the cytoplasmic nuage driven by Aub and AGO3 (Harris and Macdonald 2001; Brennecke et al. 2007; Gunawardane et al. 2007).

Piwi independently serves an additional role in the silencing of certain TEs expressed in somatic follicle cells surrounding the ovary. This somatic piRNA pathway depends on Piwi alone and therefore does not undergo ping-pong amplification (Li et al. 2009a; Malone et al. 2009; Saito et al. 2009). The *flamenco (flam)* piRNA cluster, which controls the *gypsy*, *ZAM*, and *Idefix* retrotransposons (Prud'homme et al. 1995; Desset et al. 2003), is one of the major sites of primary piRNA production (Lau et al. 2009; Li et al. 2009a; Malone et al. 2009; Saito et al. 2009). Piwi associates with piRNAs generated by *flam* and other piRNA clusters, and has been proposed to cleave homologous TE transcripts using its Slicer activity (Saito et al. 2006).

TE silencing in somatic tissues is mediated by AGO2. AGO2 binds endo-siRNAs, the majority of which silence the expression of TEs outside of the gonad (Chung et al. 2008; Czech et al. 2008; Ghildiyal et al. 2008; Kawamura et al. 2008). Silencing is achieved by Dcr-2-mediated cleavage of dsRNAs into 21-22 nt siRNA that are loaded

into AGO2, which cleaves the target TE mRNA using its Slicer activity. Interestingly, many endo-siRNAs map to the same genomic regions that generate piRNAs. Although some redundancy in TE targeting between the two pathways is plausible, some TEs were shown to be silenced by only one of these pathways (Watanabe et al. 2008).

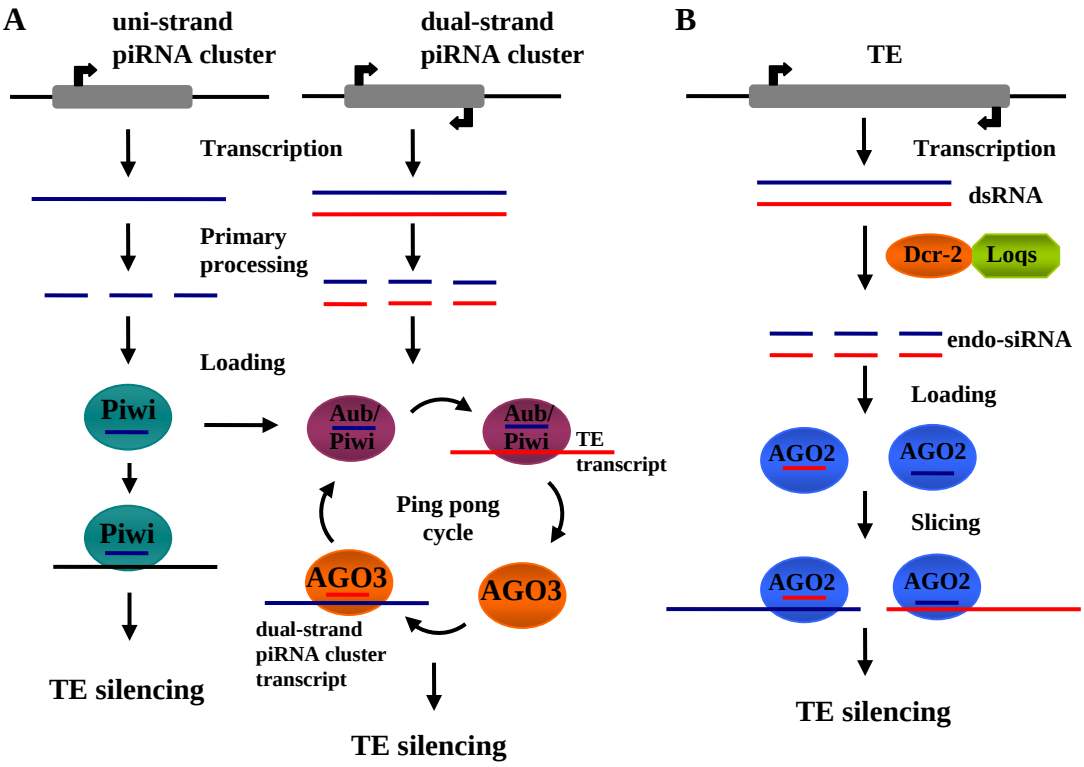


Figure 1-1. *Drosophila* piRNA and siRNA pathways. A. piRNA pathway. Single-stranded piRNA precursors are generated from either uni-strand piRNA clusters that are transcribed in one direction or dual-strand clusters that produce piRNAs from both strands. After primary processing, piRNAs derived from uni-strand clusters are loaded into Piwi to target TEs in ovarian somatic cells. Aub and AGO3-loaded piRNAs originating from the dual-strand clusters are further amplified in the "ping-pong" cycle targeting TEs in ovarian germline cells. B. Endo-siRNA pathway. Most endo-siRNAs are derived from TE transcripts. DsRNA precursors are processed by Dcr-2 and Loqs. Endo-siRNAs are loaded onto AGO2. AGO2-endo-siRNA complexes target TE transcripts cleaving them by AGO2 Slicer activity. Modified from Siomi et al., *Nat Rev Gen* 2011.

A role for RNA silencing in the nucleus

Although RNA silencing was originally believed to be a cytoplasmic mechanism, there is a sufficient amount of empirical evidence indicating that it can also function in the nucleus in various organisms. In *S. pombe*, RNAi functions primarily to nucleate and maintain heterochromatin required for centromere function but also has a function in euchromatin. In the unicellular ciliate *Tetrahymena thermophilica (T. thermophilica*), the RNAi machinery triggers programmed genome elimination that protects against potentially harmful TEs, which are also silenced by heterochromatin recruitment (reviewed in Mochizuki 2010). In *C. elegans*, NRDE-3, an AGO protein, can localize to the nucleus upon siRNA loading (Guang et al. 2008). NRDE-3 along with NRDE-2 has also been shown to inhibit RNA Polymerase II (Pol II) elongation (Guang et al. 2010). In addition, another AGO protein, CSR-1, mediates a pathway that affects meiotic silencing of unpaired chromatin and chromosome segregation during early embryogenesis (Claycomb et al. 2009; van Wolfswinkel et al. 2009). Recently, a study reported human Ago1 and Ago2 localization to the nucleus in HeLa cells (Weinmann et al. 2009). Interestingly, in mammals, AGO1 and AGO2 guide synthetic siRNAs to gene promoters (Janowski et al. 2006; Kim et al. 2006). However, the specific mechanisms of AGO-directed RNA silencing in the nucleus and its function on chromatin remains vague in metazoans.

HETEROCHROMATIN AND RNA SILENCING

Given the vast size and complexity of eukaryotic genomes, the genetic material must be efficiently packaged into the nucleus. The basic structural unit is a nucleosome, which consists of an octamer of core histones (H2A, H2B, H3, and H4) with 147 bp of DNA wrapped around it (Kornberg 1974). The individual nucleosomes are further arranged into 11 nm "beads on a string" arrays that make up higher order structure known as chromatin. Microscopically, chromatin can be classified into two distinct forms, euchromatin and heterochromatin (reviewed in Richards and Elgin 2002). During interphase, euchromatin is visualized as decondensed chromatin whereas heterochromatin is the compacted form. In *Drosophila*, an estimated one-third of the genome is composed of repetitive and noncoding sequences associated with heterochromatin. This form of chromatin is characterized by repeat-rich sequences, hypoacetylation of histone tails, and dimethylation of histone H3 on lysine 9 (H3K9me2) (reviewed in (Grewal and Elgin 2007)). A conserved nonhistone Heterochromatin Protein 1 (HP1) is a critical component of heterochromatin, localizing predominantly at and near centromeres but also residing at telomeres, the Y, and fourth chromosomes. These regions tend to be rich in TEs, which must be suppressed in order to maintain genomic stability but can serve a cellular function, particularly in the case of *Het-A* and *TART* at the telomeres (reviewed in (Mason et al. 2008)).

RNAi-mediated heterochromatic silencing in *S. pombe*

The paradigm for how RNA silencing controls gene expression at the chromatin level comes from pioneering genetic and biochemical studies in *S. pombe*, which have shed considerable light on mechanisms of heterochromatin assembly. The RNA interference (RNAi) machinery was found to play a key role in heterochromatin formation by detecting the transcription of specific DNA repeats located at the mating type locus and the centromere and subsequently nucleating heterochromatin. For example, double-stranded RNAs (dsRNA) transcribed from pericentromeric repeats are processed by the RNase III endonuclease Dicer1 into siRNAs (Figure1-2; Volpe et al. 2002). The single Argonaute, Argonaute1 binds these siRNAs as part of the RNA-induced transcriptional silencing complex (RITS) (Motamedi et al. 2004). Loading of RITS with siRNA and recruitment of the complex to the site of dsRNA transcription requires the Clr4 histone methyltransferase, which methylates H3K9 (Noma et al. 2004). This methylation mark serves as a binding site for Swi6, a fission yeast homolog of HP1, leading to heterochromatin establishment and spreading.

Importantly, heterochromatin can also be nucleated independently of RNAi by other mechanisms. In fission yeast, for example, in the absence of RNAi the ATF/CREB stress-activated proteins promote heterochromatin formation at the mating type locus (Jia et al. 2004), and the Taz1 protein can establish HP1 recruitment to telomeres (Kanoh et al. 2005). These studies exemplify the redundancy of RNAi and additional mechanisms with respect to the formation of heterochromatin.

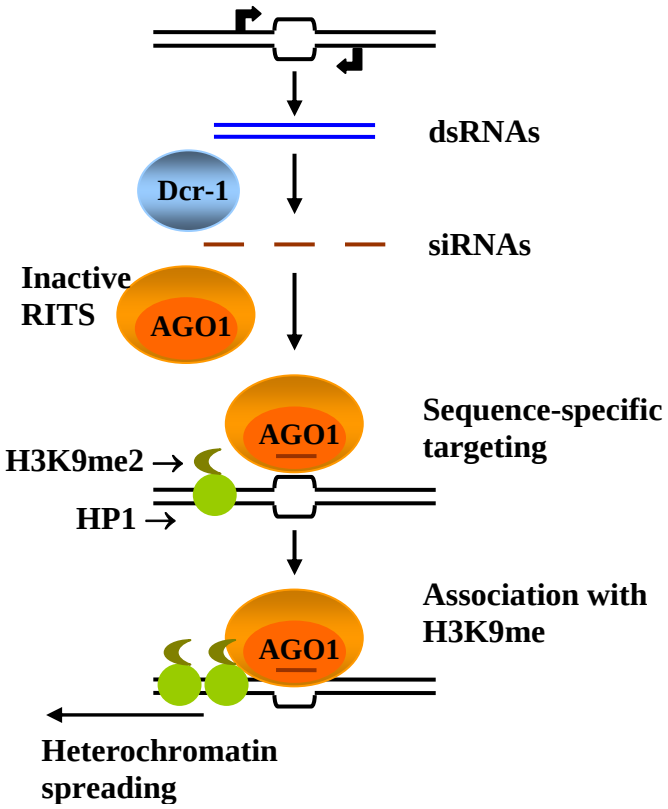


Figure 1-2. Model for RNAi-mediated centromeric silencing in *S. pombe*. In this pathway, dsRNAs transcribed from pericentromeric repeat-rich regions are processed by Dcr-1 into siRNAs that are loaded into the RITS complex containing Ago1. Association of the RITS complex with the pericentric repeats via sequence-specific pairing guides H3K9 methylation leading to heterochromatin nucleation and spreading. Modified from Verdel et al., *Science*, 2004.

RNA silencing involvement in heterochromatin recruitment in metazoans

Based on the model in *S. pombe*, it has been hypothesized that the mechanism of RNAi-dependent heterochromatin assembly is evolutionarily conserved between unicellular eukaryotes and metazoans. Multiple studies suggested that one or more *Drosophila* RNA silencing pathways may participate in transcriptional TE silencing by inducing heterochromatin formation, but direct evidence is lacking. Initial evidence that implicated piRNA pathways, and Piwi in particular, in establishment or maintenance of heterochromatin in the germline came from an observation that mutation of *piwi*, *aub*, or *spn-E*, encoding an RNA helicase required for the germline piRNA pathway (Vagin et al. 2006; Malone et al. 2009), results in defects in heterochromatic silencing and visible changes in heterochromatin localization. These mutants display reduced silencing of pericentromeric transcriptional reporters and exhibit mislocalization of HP1 and H3K9me2 in salivary gland polytene chromosomes (Pal-Bhadra et al. 2004) suggesting a role for the piRNA silencing factors in nucleating or maintaining heterochromatin. These results, however, contrast with a study that investigated the role of Piwi on *3R-TAS* subtelomeric region, which is a site of Piwi chromatin association and piRNA production. The study demonstrated that a mutation of *piwi* results in an increase of transcriptional silencing at *3R-TAS* region and an increase of HP1 association suggesting an activating role for Piwi on chromatin (Yin and Lin 2007). Moreover, HP1 was identified as an interactor of Piwi in yeast two-hybrid screens (Brower-Toland et al. 2007). Furthermore, both proteins associate specifically with the chromatin of two transposable elements, *1360* and the *F element*. Based on their findings, the authors propose that Piwi could

serve as a recruitment platform for HP1 binding and heterochromatin-mediated silencing. However, this model appears not to be applicable to the *3R-TAS* subtelomeric region. Thus, it remains an open question whether other sites in the genome could serve as Piwi-dependent HP1 recruitment sites.

Two studies suggest that the endo-siRNA silencing pathway may participate in transcriptional TE silencing by inducing heterochromatin formation. First, mutation of *AGO2* results in pleiotropic cellular defects in early embryos including mislocalization of HP1 and the histone H3 variant CID, which binds specifically the centromere (Deshpande et al. 2005). Later in development, *AGO2* mutants display mislocalization of HP1 on polytene chromosomes of the larval salivary gland (Fagegaltier et al. 2009). Additionally, silencing of a pericentromeric transcriptional reporter is relieved when the maternally derived pool of AGO2 is reduced. Despite these defects, *AGO2* mutant flies develop normally and are fertile, suggesting that these defects are mild and can be compensated by other mechanisms.

In plants and other metazoans, it is similarly unclear whether RNA silencing can establish heterochromatin directly. In *Arabidopsis thaliana*, RNA-directed RNA polymerase, RDR2, a dicer protein, DCL3 and AGO4 mediate RNA-directed DNA methylation (RdDM), which has been implicated in centromeric repeat and TE silencing. However, H3K9me2 present in these regions is not lost in *dcl3* and *rdr2* mutants (Qi et al. 2006), a situation reminiscent of *S. pombe* siRNA-dependent heterochromatin assembly at the mating-type locus where a redundant pathway maintains heterochromatin formation in a manner independent of RNAi. A study in chicken demonstrated that Ago2 associate at very low levels with a constitutive heterochromatic domain that separates the

folate receptor gene and the β-globin locus (Giles et al. 2010). The authors proposed that this heterochromatic region is maintained by a Dicer and Argonaute 2 dependent mechanism, however, it is not known whether this represents a general mechanism to maintain heterochromatin. Furthermore, a recent study in mouse cells proposed a model for a *de novo* HP1 targeting to pericentric heterochromatin by long non-coding RNAs corresponding to major satellites (Maison et al. 2011). Importantly, the authors did not detect any small dsRNA corresponding to these heterochromatic regions.

POLYCOMB-MEDIATED GENE REPRESSION AND RNA SILENCING

The highly conserved proteins of the two antagonizing complexes Polycomb group (PcG) and trithorax group (TrxG) maintain gene-expression profiles of crucial developmental regulators through either repression or activation, respectively. PcG/TrxG target genes possess *cis*-regulatory dual response elements termed Polycomb Response Elements (PREs), to which both PcG and TrxG proteins can be independently recruited by DNA-binding proteins (Simon and Kingston 2009). In *Drosophila*, in the silenced state, the DNA-binding protein Pleiohomeotic (PHO) (Klymenko et al. 2006) binds PREs and recruits Polycomb repressive complex 1 (PRC1) (Mohd-Sarip et al. 2005). PRC1-mediated ubiquitination in concert with PRC2-directed methylation of H3 at lysine 27(Wang et al. 2004) lead to Polycomb-induced gene silencing. Other DNA-binding proteins such as GAGA factor (GAF) have been suggested to function at PREs (Muller and Kassis 2006). During late embryogenesis, PcG and TrxG proteins maintain the

repressed or activated state of the genes in homeotic clusters respectively by binding to PREs and modifying chromatin structure.

RNA silencing has been suggested to affect the PcG response in a regulatory manner by influencing nuclear organization. A phenomenon, known as cosuppression, where multiple transgenic copies silence an endogenous gene, is sensitive to both PcG and RNA silencing components (Pal-Bhadra et al. 2002). Additionally, a transgenic reporter line that contains the *Drosophila Fab-7* region, a PRE-containing boundary element that controls *Abd-B* expression, can induce PcG-dependent silencing of a reporter gene and of the endogenous *scalloped* (*sd*) gene (Bantignies et al. 2003). Argonaute mutants, *ago1*, *piwi* and *aub* were shown to disrupt *Fab-7* PRE pairing-dependent silencing (Grimaud et al. 2006). Furthermore, RNA silencing has been suggested to play a role in PcG-dependent long-distance interaction between PRE-containing loci. Termed PcG bodies, these PRE-containing loci form approximately 50 nuclear foci that have been proposed to be sites of transcriptional repression that contain multiple PcG complexes bound to different PREs. Two remote PRE-containing loci have been shown to colocalize to the same PcG body and these long-range interactions were decreased in RNA silencing mutants.

CHROMATIN INSULATORS, NUCLEAR ORGANIZATION AND RNA SILENCING

It has become increasingly apparent that DNA topology is a critical determinant of gene regulation. While enhancers activate their target promoters over long distances,

insulators act to restrict these communications (reviewed in Wallace and Felsenfeld 2007). Chromatin insulators are defined as DNA-protein complexes that are characterized by two functional properties, enhancer blocking and barrier activity (Figure 1-3). In the former case, a chromatin insulator, positioned between an enhancer and a promoter, can interfere with enhancer-promoter interactions in a directional manner. In the latter case, an insulator can buffer transgenes against the spread of silent chromatin.

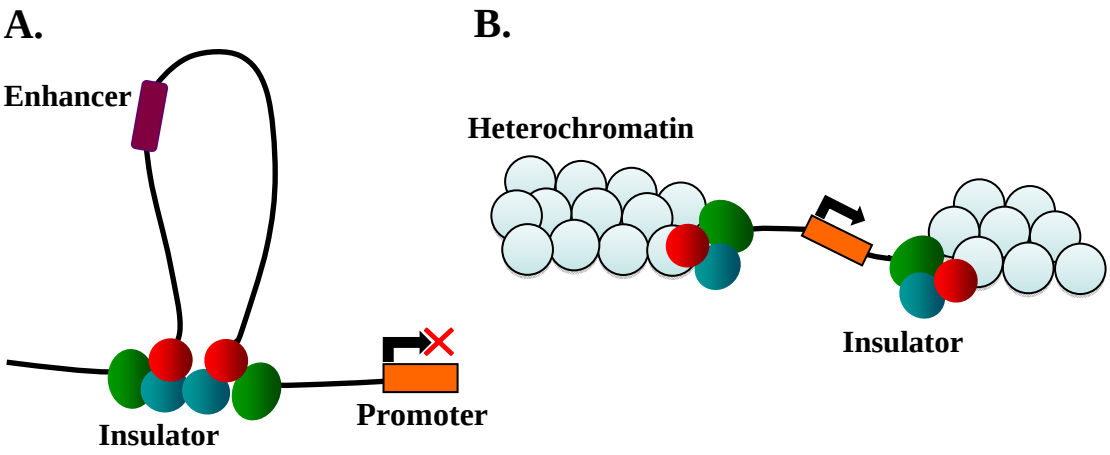


Figure 1-3. Functional properties of chromatin insulators. A) Enhancer blocking activity of chromatin insulators restricts enhancer-promoter communications while (B) barrier function demarcates silent and active chromatin domains.

Unlike vertebrates, which possess only one known insulator protein, the CCCTC-binding factor (CTCF) (Phillips and Corces 2009), *Drosophila* employs at least five known insulators defined by their DNA binding proteins. Two well-characterized insulators are the *gypsy* insulator and *Frontabdominal-8* (*Fab-8*) (Moon et al. 2005; Gerasimova et al. 2007; Mohan et al. 2007) each containing binding sites for the zinc-finger DNA binding proteins Suppressor of Hairy wing (Su(Hw)) and CTCF, respectively

(Figure 1-4). The *gypsy* insulator contains an additional component, Mod(mdg4)2.2, a BTB domain protein that directly interacts with Su(Hw) but not DNA. The two insulators share a common component, Centrosomal protein 190 (CP190) (Pai et al. 2004; Gerasimova et al. 2007), which interacts with Su(Hw), Mod(mdg4)2.2 and CTCF. Despite thousands of insulator binding sites throughout the genome, insulator complexes localize to a few large nuclear foci termed insulator bodies. These bodies are proposed to be insulator sites clustered together to organize chromatin into distinct transcriptional domains, and the integrity of insulator bodies is highly correlated with *gypsy* insulator function (Figure 1-4; reviewed in Bushey et al. 2008). It has been proposed that chromatin insulators interact with each other or with other *cis*-regulatory elements to form chromatin loops. Chromosome conformation capture (3C) studies have demonstrated an interaction between two insulator sites (Blanton et al. 2003). Moreover, *gypsy* insulator loops have been visualized in salt-extracted nuclei by fluorescence in situ hybridization (Byrd and Corces 2003).

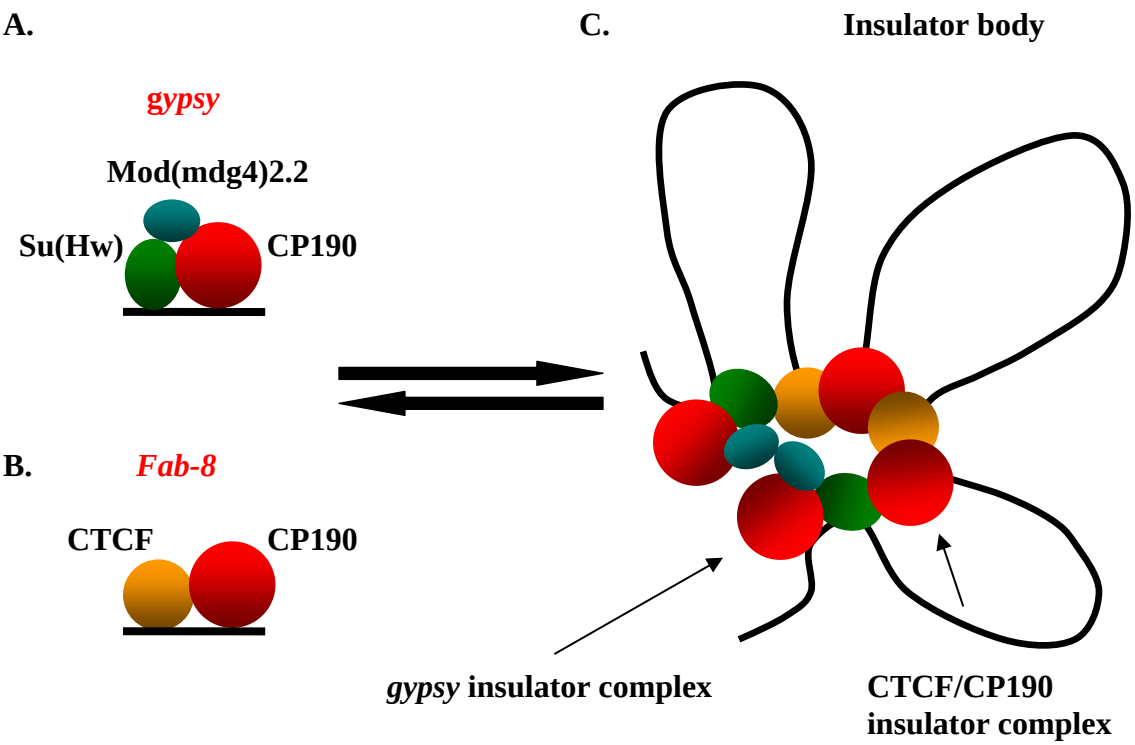


Figure 1-4. Nuclear organization of a chromatin insulator.

A. The *gypsy* insulator is defined by its binding protein Su(Hw), which also interacts with Mod(mdg4)2.2 and CP190. B. The *Fab-8* insulator, located in the *Abd-B* locus of the BX-C complex, is bound by CTCF and CP190. C. Insulator proteins colocalize in large nuclear structures termed insulator bodies. It has been proposed that insulator complexes come together to establish transcriptional domains by looping of the DNA.

The *Fab-8* insulator is a part of a large homeotic gene (*Hox*) cluster, known as the *bithorax* complex (BX-C), that controls the identity of nine parasegments (PS5-14) in the posterior two-thirds of the fly by regulating the expression of three homeotic genes, *Ultrabithorax* (*Ubx*), *abdominal-A* (*abd-A*) and *Abdominal-B* (*Abd-B*) (Figure 1-5). Each parasegment-specific *infraabdominal* (*iab*) domain contains an enhancer that is kept

autonomous by boundary elements, *Mcp*, *Fab-6*, *Fab-7* and *Fab-8* (reviewed in Maeda and Karch 2009). For example, in early embryogenesis, *Abd-B* expression is controlled by enhancers in *iab-5-8 cis*-regulatory domains in PS10-13. CTCF is present at *Mcp*, *Fab-6*, and *Fab-8* boundary elements and has been shown to be required for *Fab-8* insulator function (Moon et al. 2005; Holohan et al. 2007). Moreover, *Fab-8* interacts with a region bound by CTCF near the *Abd-B* promoter (Kyrchanova et al. 2008), suggesting that boundary elements regulate proper communication between enhancers and the *Abd-B* promoter via CTCF organizing chromatin domains at BX-C. Insulators and other *cis*-regulatory regions in the *Abd-B* locus engage in numerous long-range interactions, and the precise topology of the locus has been postulated to be a central mechanism of tissue-specific *Abd-B* regulation (Cleard et al. 2006; Lanzuolo et al. 2007; Kyrchanova et al. 2008; Bantignies et al. 2011). However, the mechanism by which chromosome looping is achieved at this locus has not been elucidated. Vertebrate CTCF has been demonstrated to mediate chromosomal looping at several developmentally regulated loci in concert with cohesin (reviewed in Merkenschlager 2010), but it is not known whether *Drosophila* CTCF, which only shares homology in the zinc-finger DNA binding domain, retains the capacity to promote looping. Lastly, a recent genome-wide study in mammalian cells revealed that CTCF mediates numerous promoter-enhancer communications suggesting a role for CTCF that is diverse from its enhancer-blocking function (Handoko et al. 2011).

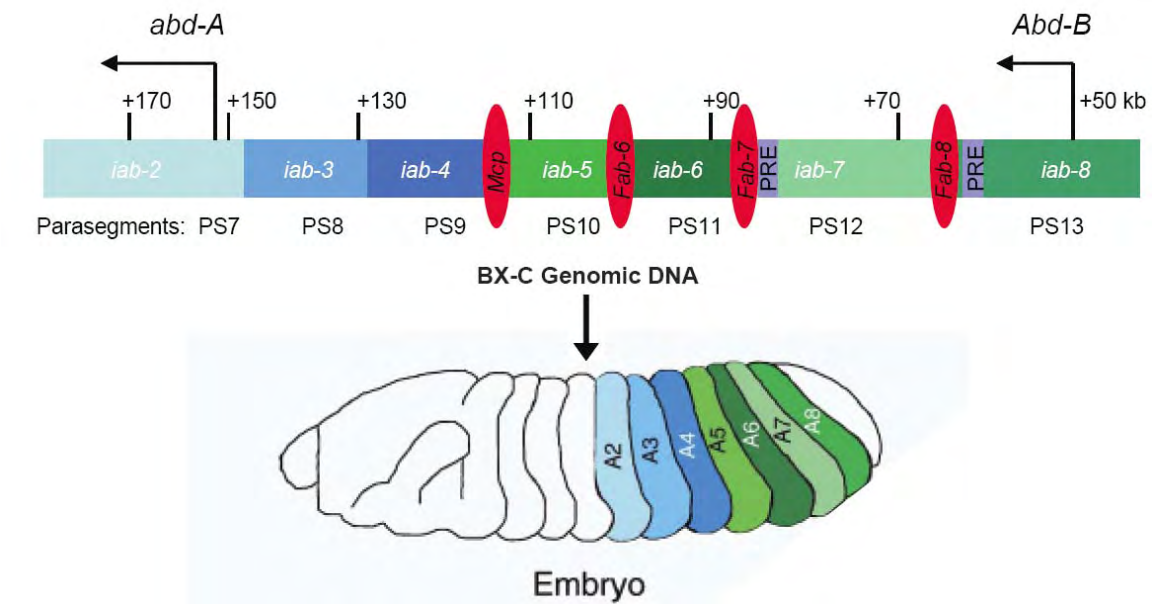


Figure 1-5. *Abd-B cis*-regulatory region of the bithorax complex (BX-C). A schematic representation of *iab-2* through *iab*-8 *cis*-regulatory domains that encompass two transcription units: *abd-A* and *Abd-B*, arrows indicate the direction of transcription. Boundary elements, *Mcp*, *Fab-6*, *Fab-7*, and *Fab-8* and PREs are also indicated. The corresponding abdominal segments specified by the *iab* domains are shown in the *Drosophila* embryo. Modified from Akbari et al., *Dev. Biol.*, 2006.

Several observations implicate RNA silencing in insulator function. First, Rm62, a DEAD-box putative RNA helicase required for dsRNA-mediated silencing copurifes with CP190 insulator complexes from *Drosophila* nuclear embryonic extract in an RNase A-sensitive manner (Lei and Corces 2006). Second, genetic analysis of *Rm62* mutants revealed improved activity of the *gypsy* insulator while mutations in Argonaute genes, *piwi* and *aub*, cause decreased *gypsy* insulator function. Third, microscopic examination of insulator bodies in larval imaginal discs of *Rm62*, *piwi* and *aub* mutants revealed

disruption of insulator body organization. Insulator body phenotypes correlated with *gypsy* insulator activity, suggesting that RNA silencing plays a role in nuclear organization of *gypsy* insulator complexes. Whether RNA silencing can affect chromatin insulator activity of an insulator other than *gypsy* has not been determined.

CONCLUSION

It has become increasingly apparent that long-range chromosomal interactions driven by *cis*-regulatory elements are critical for proper control of gene expression. Chromatin insulators disrupt enhancer-promoter interactions and can protect transgenes from the effects of silent chromatin exerted by heterochromatin or PcG-induced repression. In *Drosophila*, both the *gypsy* chromatin insulator and PREs have been postulated to mediate long-range interactions to promote higher order chromatin organization. These long-range interactions are, however, perturbed in RNA silencing mutants. Furthermore, a functional link has been reported between the *gypsy* insulator, acting to restrict PRE-mediated chromatin looping, and PcG-induced silencing (Comet et al. 2011). The plethora of evidence suggests that RNA silencing has a functional role in higher order chromatin organization. The mechanistic details regarding the interaction between RNA silencing components and insulator proteins along with PREs are, however, missing.

Here, I investigate whether RNA silencing affects gene expression at the level of higher order chromatin organization. For my studies I utilize *Drosophila melanogaster*, an outstanding model organism that has a short life cycle, a large array of available

genetic tools, a sizable pool of characterized mutants and cell lines. Specifically, I investigate two aspects of RNA involvement in chromatin function; (1) whether a role for RNA silencing in heterochromatin nucleation and maintenance is conserved in *Drosophila*, and (2) elucidate how RNA silencing affects the function of the CTCF class of chromatin insulators. First, I examine the effects of piRNA and endo-siRNA silencing mutants on heterochromatin recruitment to the sites of piRNA and endo-siRNA production utilizing genetic and biochemical approaches. Second, I investigate the role of AGO2 on CTCF/CP190-dependent *Fab-8* insulator function, PREs and active promoters, and its involvement in nuclear organization. Overall, these studies will shed more light on mechanistic details of RNA silencing function in higher order chromatin organization.

CHAPTER 2

HP1 RECRUITMENT IN THE ABSENCE OF ARGONAUTE PROTEINS IN DROSOPHILA

ABSTRACT

Highly repetitive and transposable element rich regions of the genome must be stabilized by the presence of heterochromatin. A direct role for RNA interference in the establishment of heterochromatin has been demonstrated in fission yeast. In metazoans, which possess multiple RNA silencing pathways that are both functionally distinct and spatially restricted, whether RNA silencing contributes directly to heterochromatin formation is not clear. Previous studies in *Drosophila melanogaster* have suggested the involvement of both the *AGO2*-dependent endogenous small interfering RNA (endo-siRNA) as well as Piwi-interacting RNA (piRNA) silencing pathways. In order to determine if these Argonaute genes are required for heterochromatin formation, we utilized transcriptional reporters and chromatin immunoprecipitation of the critical factor Heterochromatin Protein 1 (HP1) to monitor the heterochromatic state of piRNA clusters, which generate both endo-siRNAs and the bulk of piRNAs. Contrary to expectation, we find that mutation of *AGO2* or *piwi* increases silencing at piRNA clusters corresponding to an increase of HP1 association. Furthermore, loss of piRNA production from a single piRNA cluster results in genome-wide redistribution of HP1 and reduction of silencing at a distant heterochromatic site suggesting indirect effects on HP1 recruitment. Taken

together, these results indicate that heterochromatin forms independently of endo-siRNA and piRNA pathways.

INTRODUCTION

Heterochromatin, characterized by scarcity of genes, low levels of transcription, late replication, and low recombination rates, maintains genomic stability by hindering propagation of potentially deleterious transposable elements (TEs). One model for heterochromatin assembly comes from studies in fission yeast where a role for RNA interference (RNAi), a process by which dsRNAs silence the expression of target genes, was demonstrated (reviewed in (Grewal and Elgin 2007)). Based on the model in *S. pombe*, where transcription from pericentromeric repeat-rich regions initiates heterochromatin formation via interaction with the RNA silencing machinery, it has been suggested that the mechanism may be evolutionarily conserved between unicellular eukaryotes and metazoans.

In the *Drosophila* genome, there is an extensive overlap between TE-rich regions and pericentromeric and telomeric heterochromatin marked by non-histone Heterochromatin Protein 1 (HP1) and histone 3 dimethylated at lysine 9 (H3K9me2). The phenomenon of position-effect variegation (PEV) provided the first glimpse into the role of heterochromatin in gene silencing in *Drosophila*. When a normally euchromatic gene is relocated near heterochromatin, variegated expression results from variable levels of heterochromatin spreading over the gene in each cell. Screens for dominant mutations that either suppress or enhance {*Enhancer of variegation* [*E(var)*]} PEV were performed to identify key components of heterochromatin. For example, mutation of *Su(var)3-9*, which encodes an H3K9 methyltransferase, was identified in a large screen for modifiers of PEV (Tschiersch et al. 1994). Accordingly, loss of HP1, encoded by *Su(var)2-5*,

causes increased expression of a gene subject to PEV, while an extra copy has the reverse effect (Eissenberg et al. 1990).

Sequencing of the small RNA population associated with the Piwi claude of AGO proteins, Piwi, Aub and AGO3 from *Drosophila* ovaries identified a subclass of ~24-30 nt long RNAs, termed Piwi-interacting RNAs (piRNAs) (Brennecke et al. 2007; Yin and Lin 2007). Interestingly, piRNAs originate from discrete loci, the majority of which are located at pericentromeric and telomeric regions containing TEs and other repetitive elements. Termed piRNA clusters, these loci range in length between a few of to hundreds Kb. Two prominent clusters include the *42AB* piRNA locus on chromosome 2R and *flamenco* (*flam*) piRNA producing locus on the X chromosome. The *flam* locus was originally discovered as a region involved in the control of retrotransposons: *gypsy, Idefix,* and *ZAM* (Prud'homme et al. 1995; Desset et al. 2003). Examination of piRNAs associated with Piwi, Aub and AGO3 revealed nucleotide signatures indicative of the amplification "ping-pong" cycle (Brennecke et al. 2007; Gunawardane et al. 2007). Aub and AGO3, which both possess the Slicer endonuclease activity, are cytoplasmic proteins that most likely silence TEs post-transcriptionally in the germ cells. Interestingly, in the ovarian somatic cell (OSC) line, where Piwi is highly expressed but not Aub or AGO3, Piwi-associated piRNAs that are mostly derived from *flam* locus are derived by the primary processing pathway that does not involve the amplification step (Saito et al. 2009). Piwi, a nuclear protein that is present in both somatic and germ cells of the ovary, is required for the TE silencing in gonadal somatic cells. Thus, it is possible that Piwi can recruit heterochromatin machinery to silence the repetitive elements in the nucleus.

The role of Piwi and the somatic primary piRNAs as epigenetic regulators remains controversial since Piwi seems to exert the opposite effects at different genomic sites. On one hand, *piwi* mutants exhibit defects in heterochromatic silencing and heterochromatin localization, and Piwi and HP1 interact physically and associate with heterochromatic *1360* and *F element* transposable elements, suggesting that Piwi and the somatic primary piRNAs may silence TEs by recruiting heterochromatin (Pal-Bhadra et al. 2004; Vagin et al. 2006; Brower-Toland et al. 2007; Malone et al. 2009). On the other hand, *piwi* mutants show both increased transcriptional silencing and HP1 association with the *3R-TAS* subtelomeric region, indicating an epigenetic activation role. Therefore, it remains an open question whether the piRNA pathway can recruit heterochromatin to its target sites.

Sequencing of small RNAs associated with AGO2 revealed a population of ~21-22 nt long species in *Drosophila* embryos, ovaries and S2 embryonic cells termed endogenous siRNAs (endo-siRNAs) (Chung et al. 2008; Czech et al. 2008; Ghildiyal et al. 2008; Kawamura et al. 2008). The majority of endo-siRNAs exhibit homology to TEs and other repetitive sequences and overlap considerably with piRNA clusters. Studies have reported HP1 mislocalization defects in early embryos and on larval polytene chromosomes in *ago2* mutants suggesting endo-siRNA pathway involvement in heterochromatin formation (Deshpande et al. 2005; Fagegaltier et al. 2009). Whether the endo-siRNA pathway can directly recruit HP1 to heterochromatin remains unclear.

In this study, I investigated whether HP1 association with heterochromatin in *Drosophila* is mediated by either the *piwi* dependent piRNA pathways or by the *AGO2* dependent endo-siRNA pathway. I hypothesized that similarly to the RNAi-dependent

heterochromatin recruitment in fission yeast, heterochromatin assembly depends on piRNA and endo-siRNA pathways that silence TEs in the germline and soma respectively. In order to address this hypothesis, I examined heterochromatin localization as marked by HP1 to piRNA producing loci in piRNA and endo-siRNA pathway mutants genetically, utilizing transcriptional reporters, and biochemically by Chromatin Immunoprecipitation (ChIP). Transcriptional reporters bearing mini-*white* transgene and positioned inside or in close proximity to top piRNA producing loci, *flam*, *42AB*, and *80EF* on chromosome 3L were used to measure mini-*white* expression in the adult eye. ChIP, an immunoprecipitation technique used to determine whether specific proteins are associated with specific genomic regions *in vivo*, was utilized to assay HP1 localization to piRNA clusters. Specifically, I hypothesized that if heterochromatin recruitment depends on piRNA and/or endo-siRNA silencing pathways then a disruption in Piwi and AGO2 proteins would result in (a) decreased silencing as measured by transcriptional reporters at the piRNA producing loci and (b) a decrease of HP1 association with piRNA clusters.

Here, I show that piRNA clusters are subject to heterochromatic silencing and bound by HP1. Contrary to expectation, mutation of *AGO2*, *piwi* or *aub* results in increased silencing at piRNA clusters and an increase in HP1 association with these loci. Furthermore, loss of piRNA production at a single piRNA locus results in global redistribution of HP1 and a reduction of silencing at a distant heterochromatic site. Therefore, our results indicate that HP1 can associate with chromatin independently of both endo-siRNA and piRNA silencing pathways in the soma.

RESULTS

Heterochromatin dependent transcriptional silencing at piRNA clusters

We sought to determine if HP1 is recruited to heterochromatin by AGO2 or Piwi. The majority of genomic regions that produce the bulk of piRNA, termed piRNA clusters, are pericentromeric and rich in transposable elements (Brennecke et al. 2007; Yin and Lin 2007). These regions also produce endo-siRNA (Chung et al. 2008; Czech et al. 2008; Ghildiyal et al. 2008; Kawamura et al. 2008), and due to their proximity to the centromere, may be heterochromatic and serve as platforms for Argonaute mediated HP1 recruitment.

In order to test genetically whether pericentromeric piRNA clusters are heterochromatic, we examined a collection of fly lines bearing P element transgene insertions inside or in close proximity to four piRNA producing loci, *flam*, *80EF*, *42AB*, and *38C*. The P elements contain a mini-*white* transcriptional reporter that was assayed for expression in the adult eye. Genomic locations of these transgene insertions are indicated in relation to previously identified small RNAs immunoprecipitated with Piwi, Aub/AGO3, and AGO2 respectively from various cell types (Figure 2-1) (Brennecke et al. 2007; Yin and Lin 2007; Czech et al. 2008; Kawamura et al. 2008). Lines harboring P elements inside or in the vicinity of a piRNA cluster exhibit variegating coloration of distinct eye facets similar to PEV, suggesting the presence of variably spreading heterochromatin at their sites of insertion (Figure 2-2, Table 2-1). Interestingly, insertions within a piRNA cluster that display high *mini-white* expression without

variegation harbor *SUPor-P* constructs, which contain Suppressor of Hairy wing [Su(Hw)] insulator sequences that flank and likely protect the *mini-white* reporter from the effects of surrounding heterochromatin (Roseman et al. 1993).

Figure 2-1. Schematic representation of four top piRNA clusters.

Genomic locations of small RNAs, primer sets used for ChIP, and P element insertions at (A) *flam* piRNA cluster on chromosome X, (B) *80EF* piRNA cluster on chromosome 3L, (C) *42AB* piRNA cluster on chromosome 2R and (D) *38C* piRNA cluster on chromosome 2L. Sequence datasets derived from previous studies were mapped to the genome using Bowtie software allowing two or zero mismatches (Langmead et al. 2009). Piwi-immunoprecipitated (Brennecke et al. 2007; Yin and Lin 2007), Aub or AGO3-immunoprecipitated (Brennecke et al. 2007) and AGO2-immunoprecipitated (Czech et al. 2008; Kawamura et al. 2008) reads mapping to multiple locations in the genome are indicated in red (with two mismatches allowed) and pink (with zero mismatches allowed) while uniquely mapping reads are in dark blue (with two mismatches allowed) and light blue (with zero mismatches allowed). Primer sets used for ChIP analysis are indicated by yellow arrowheads. Strongly variegating (dark green triangle), weakly variegating (light green triangle), and non-variegating P elements with high expression levels (white triangle) are indicated. (A) *P{EPgy2}DIP1*^{EY02625}, (B) *PBac(PB) c06482*, and (C) *P{EPgy2}EY08366* P element insertions are marked by an asterisk. *SUPor-P* P elements containing insulator sequences are marked by an "I". Centromere proximal end is marked by a hollow C. RepeatMasker detected sequences are represented in black.

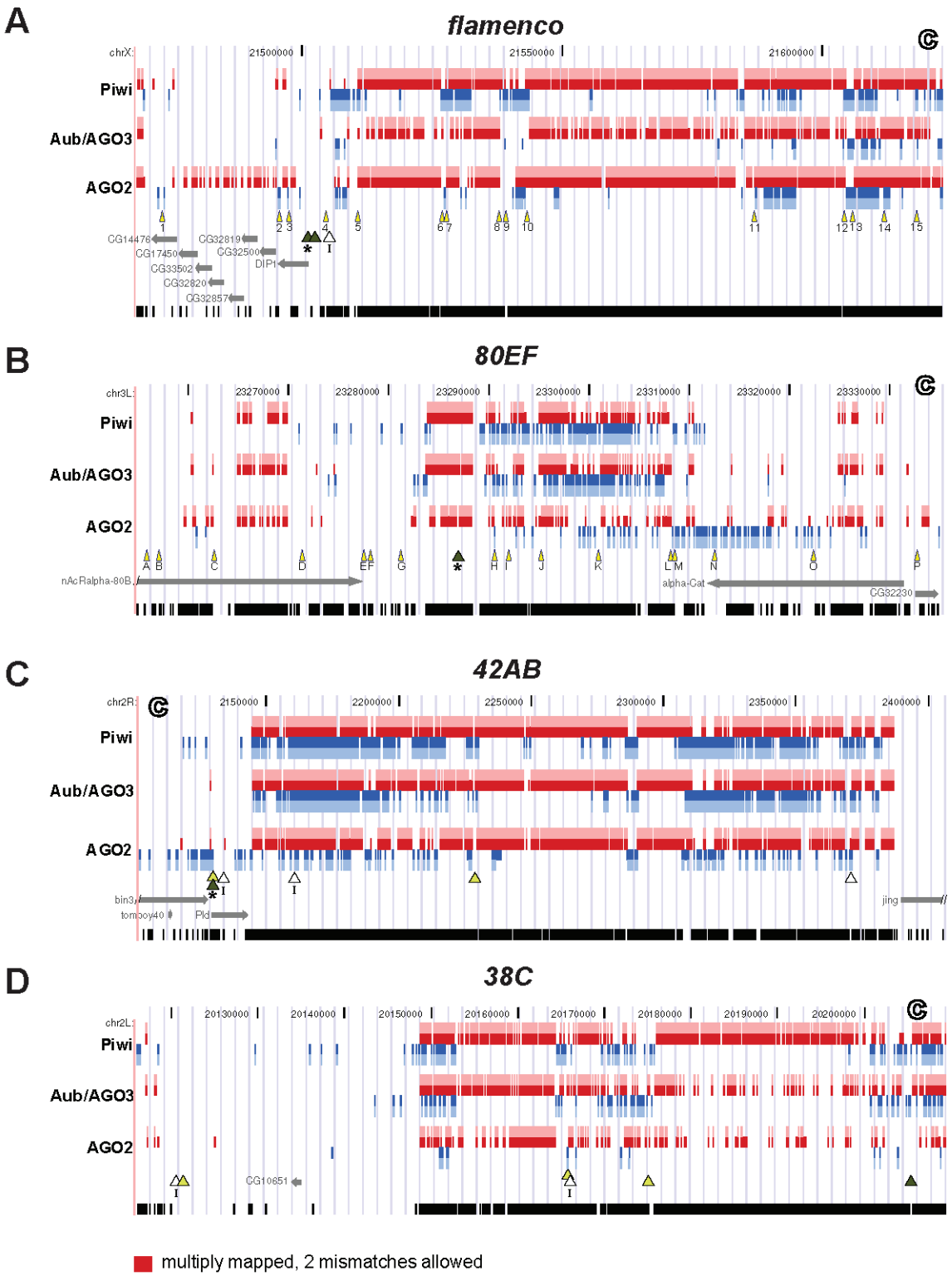

A
flamenco
chrX:
21500000
21550000
21600000
Piwi
Aub/AGO3
AGO2
CG14476
CG32819
CG17450
CG32500
DIP1
CG33502
CG32820
CG32857
B
80EF
chr3L:
23270000
23280000
23290000
23300000
23310000
23320000
23330000
nAcRalpha-80B
alpha-Cat
CG32230
C
42AB
chr2R:
2150000
2200000
2250000
2300000
2350000
2400000
bin3
tomboy40
Pld
jing
D
38C
chr2L:
20130000
20140000
20150000
20160000
20170000
20180000
20190000
20200000
CG10651
multiply mapped, 2 mismatches allowed
multiply mapped, 0 mismatches allowed
uniquely mapped, 2 mismatches allowed
uniquely mapped, 0 mismatches allowed
RepeatMasker
primer set
reporter
insulated
Variegation
yes
weak
no
P-element insertion
centromere
proximal

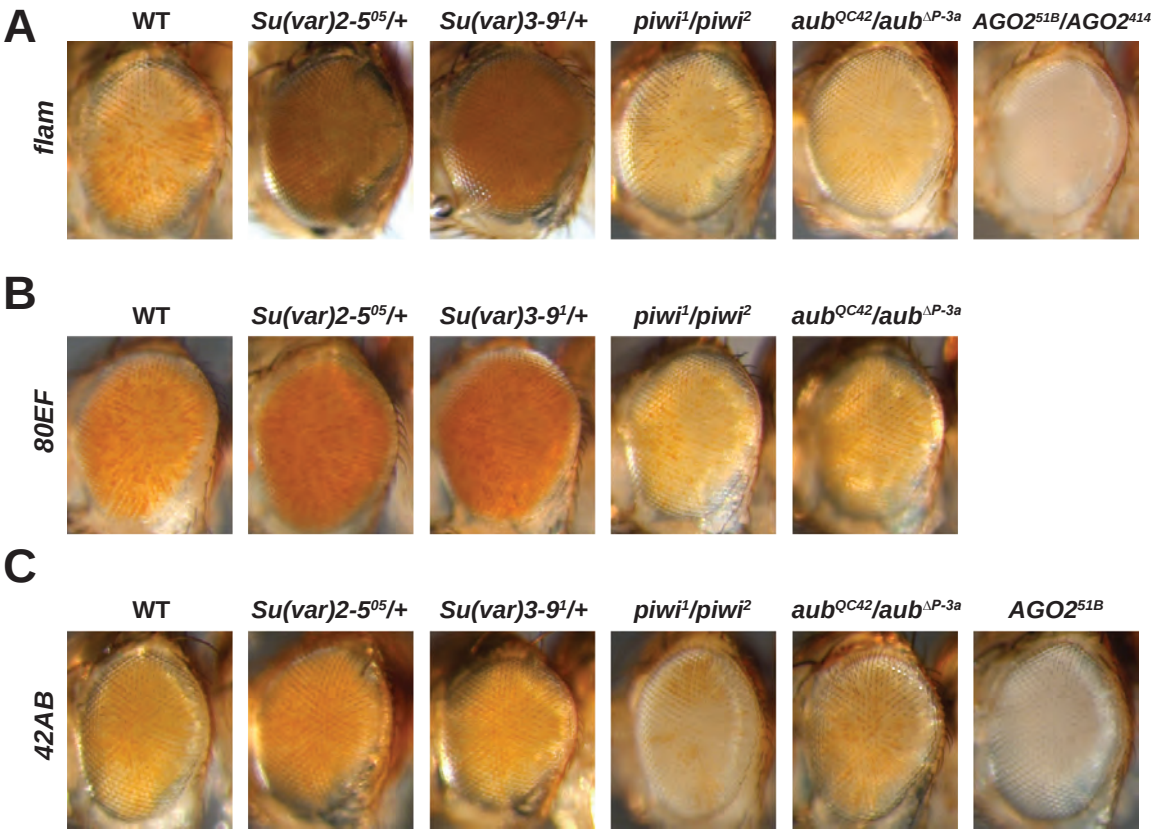
Table 2-1. Expression of mini-*white* in fly lines harboring P element insertions in four top piRNA clusters.

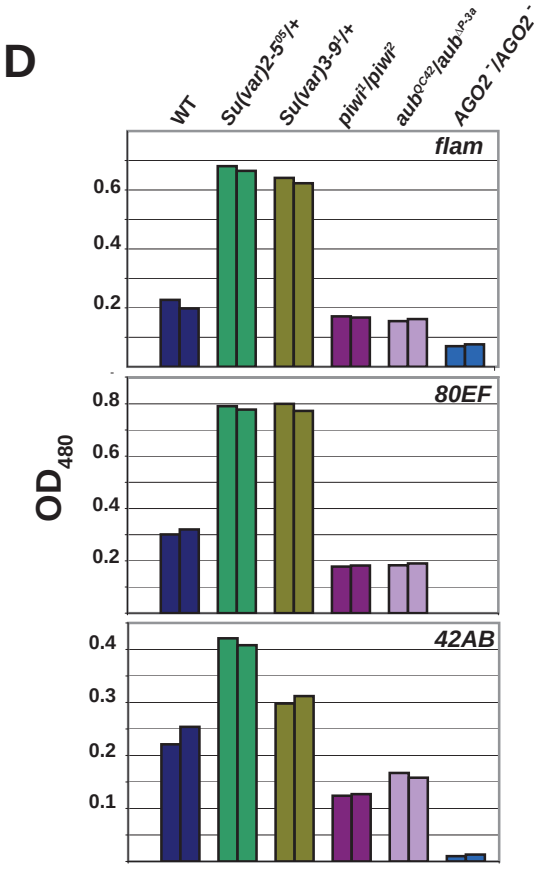
Insertion	piRNA cluster	Genomic coordinates of insertion	Variegation	Inside piRNA cluster
P{EPgy2}DIP1^{EY02625}	*flam*	X:21,501,171 [-]	Yes	No
P{SUPor-P}flam^{KG00476}	*flam*	X:21,505,285 [-]	No	No
P{GT1}flam^{BG02658}	*flam*	X:21,502,538 [-]	Yes	No
PBac(WH)CG32230^{f00651}	*80EF*	3L:23,237,018 [+]	Weak	No
PBac(PB) c06482	*80EF*	3L:23,286,922 [-]	Yes	Yes
PBac(PB)CG40470^{c06318}	*80EF*	3L:23,849,420 [+]	No	No
P{GT1}BG01672^{BG01672}	*42AB*	2R:2,370,529 [-]	No	Yes
P{EPgy2}EY08366	*42AB*	2R:2,129,510 [+]	Yes	No
P{XP}d02126	*42AB*	2R:2,129,452 [+]	Weak	No
P{SUPor-P}Pld^{KG02714}	*42AB*	2R:2,133,438 [+]	No	No
P{SUPor-P}KG09351	*42AB*	2R:2,160,357 [-]	No	Yes
PBacf(WH)04291	*42AB*	2R:2,228,280 [-]	Weak	Yes
P{EPgy2}EY01034	*38C*	2L:20,205,306	Yes	Yes
P{XP}d02757	*38C*	2L:20,174,988 [+]	Weak	Yes
PBac(WH)f03348	*38C*	2L:20,165,746	Weak	Yes
P{SUPor-P}KG05288	*38C*	2L:20,166,034 [+]	No	Yes
PBac{RB}e03575	*38C*	2L:20,121,359 [+]	Weak	No
P{SUPor-P}KG02342	*38C*	2L:20,120,504 [-]	No	No

The genomic coordinates for four top piRNA clusters were defined as previously determined by Brennecke et al., 2007. The genomic coordinates of the P-element insertions were confirmed by PCR with primers specific to the P-elements and flanking genomic sequences.

Expression analysis of these transcriptional reporter insertions indicates that piRNA clusters and their immediate vicinity are subject to HP1 dependent silencing. Reporter expression levels of three lines harboring an insertion at *flam*, *80EF*, or *42AB* with the most apparent variegation were tested for dependence on heterochromatin. *P{EPgy2}DIP1*^{EY02625} is inserted in a gene located on the centromere distal side of the *flam* piRNA producing locus on the X chromosome (Figure 2-1A), *PBac(PB)c06482* resides within the *80EF* cluster on chromosome 3L (Figure 2-1B), and *P{EPgy2}EY08366* borders the centromere proximal edge of the *42AB* piRNA locus on chromosome 2R (Figure 2-1C). In order to test whether these reporters are sensitive to perturbation of heterochromatin, the expression of mini-*white* was examined in *Su(var)2-5*^{05}/+ and *Su(var)3-9*^{1}/+ dominant loss-of-function mutants, which are compromised for HP1 and H3K9 methyltransferase activity respectively. As expected, decreased silencing of mini-*white* expression resulting in increased pigmentation was observed for all three insertions in the heterochromatin mutants compared to wild type (Figure 2-2), suggesting that the vicinity of P element insertion are indeed heterochromatic.

Figure 2-2. piRNA and endo-siRNA pathway mutants display increased silencing of transcriptional reporters at or near piRNA clusters. Adult eyes of wild type, *Su(var)2-5*^{05}/+, *Su(var)3-9*^{1}/+, *piwi*^{1}/*piwi*^{2}, *aub*^{QC42}/*aub*$^{\Delta P\text{-}3a}$, and/or *AGO2*$^{-}$/*AGO2*$^{-}$ mutants carrying a mini-*white* transgene inserted inside or in close proximity to the (A) *flam*, (B) *80EF*, and (C) *42AB* piRNA clusters. *AGO2*^{51B}/*AGO2*^{414} mutants are examined in (A) while *AGO2*^{51B} mutants are examined in (C). (D) Levels of eye pigment measured at 480 nm extracted from male heads of the indicated genotypes (A-C).

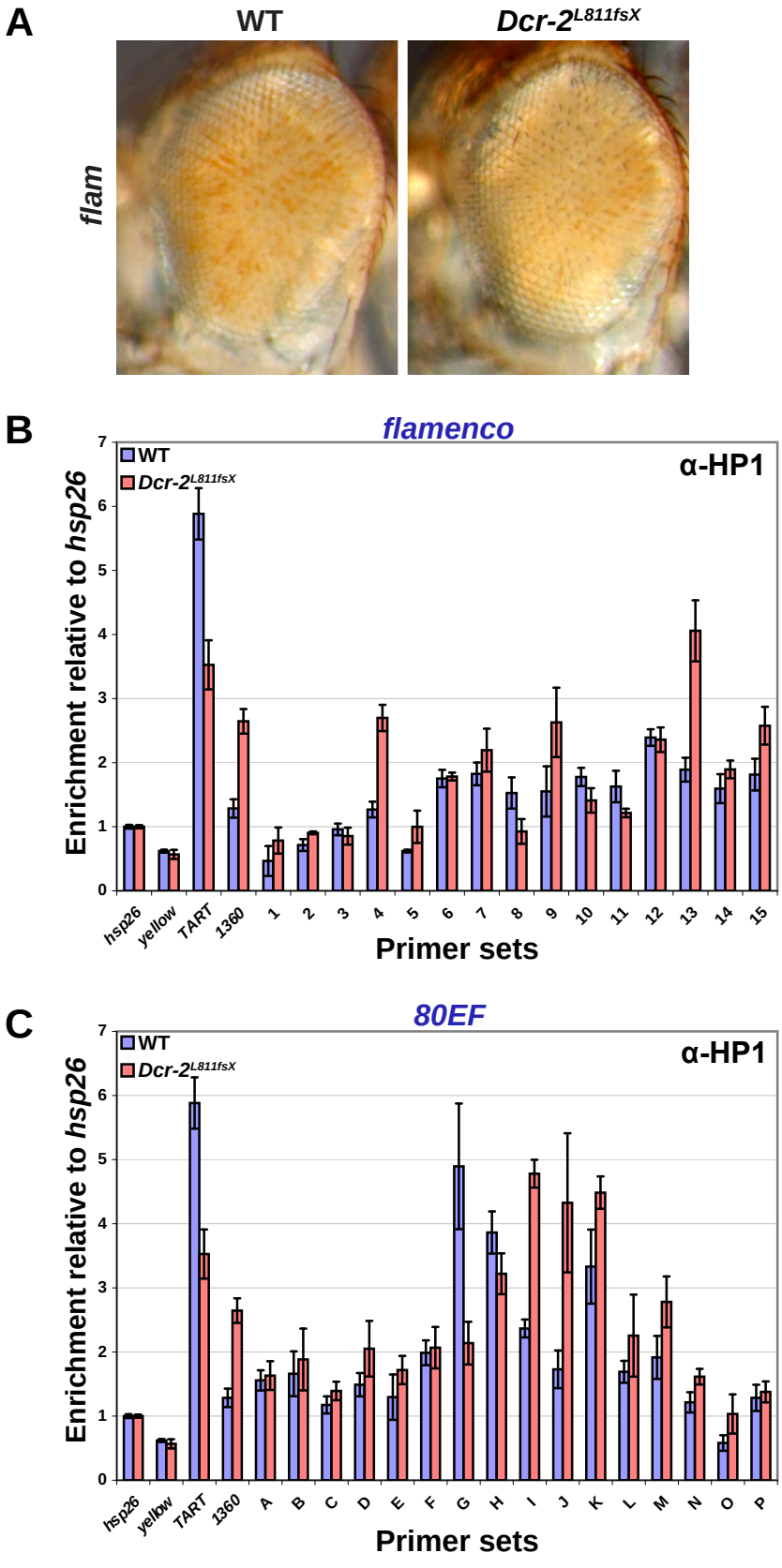

A
WT
Su(var)2-5^05/+
Su(var)3-9^1/+
piwi^1/piwi^2
aub^QC42/aub^ΔP-3a
AGO2^51B/AGO2^414
flam
B
WT
Su(var)2-5^05/+
Su(var)3-9^1/+
piwi^1/piwi^2
aub^QC42/aub^ΔP-3a
80EF
C
WT
Su(var)2-5^05/+
Su(var)3-9^1/+
piwi^1/piwi^2
aub^QC42/aub^ΔP-3a
AGO2^51B
42AB


D
WT
Su(var)2-5^05/+
Su(var)3-9^1/+
piwi^1/piwi^2
aub^QC42/aub^ΔP-3a
AGO2^-/AGO2^-
flam
80EF
42AB
OD480
0.6
0.4
0.2
0.8
0.6
0.4
0.2
0.4
0.3
0.2
0.1

piRNA and endo-siRNA pathway mutants decrease transcription at piRNA clusters

We next tested whether the transcriptional reporters at piRNA clusters are sensitive to perturbations in the piRNA and endo-siRNA silencing pathways. If Piwi were responsible for direct recruitment of HP1 to piRNA clusters, mutation of *piwi* should increase mini-*white* expression similarly to disruption of heterochromatin. Surprisingly, $piwi^{1}/piwi^{2}$ loss-of-function mutants exhibit a substantial loss of reporter expression indicating increased silencing when compared to wild type (Figure 2-2). Furthermore, $aub^{QC42}/aub^{\Delta P\text{-}3a}$ loss-of-function piRNA pathway mutants result in a similar reduction of mini-*white* expression. Strikingly, the *flam* transcriptional reporter expression level was decreased dramatically in the transheterozygous endo-siRNA pathway mutant, $AGO2^{51B}/AGO2^{414}$ compared to wild type (Figure 2-2A). Similarly, in the $AGO2^{51B}$ null mutant, the *42AB* transcriptional reporter displays almost complete silencing (Figure 2-2C). Spectroscopic analysis of extracted eye pigment verifies the overall changes in mini-*white* expression levels for each genotype compared to wild type (Figure 2-2D). Additionally, examination of $Dcr\text{-}2^{L811fsX}$ mutants shows a similar mild increase in silencing for the transcriptional reporter inserted near *flam* (Figure 2-3A). The opposite effects of piRNA and endo-siRNA pathway mutations compared to heterochromatin mutations suggest that these RNA silencing pathways may actually oppose heterochromatin formation at piRNA clusters.

Figure 2-3. *Dcr-2* mutants display increased HP1 chromatin association and increased silencing at piRNA clusters. ChIP at (A) *flam* and (B) *80EF* piRNA clusters in wild type (blue) and *Dcr-2*^{L811fsX} /+ (orange) from adult heads with antibodies specific to HP1. Values shown are percent input immunoprecipitated for each primer set normalized to *hsp26*. Error bars indicate standard deviation of quadruplicate PCR measurements. (C) Adult eyes of wild type and *Dcr-2*^{L811fsX} mutants carrying a mini-*white* transgene inserted in close proximity to the *flam* piRNA cluster.


A
WT
Dcr-2L811fsX
flam
B
flamenco
WT
Dcr-2L811fsX
α-HP1
Enrichment relative to hsp26
0
1
2
3
4
5
6
7
hsp26
yellow
TART
1360
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Primer sets
C
80EF
WT
Dcr-2L811fsX
α-HP1
Enrichment relative to hsp26
0
1
2
3
4
5
6
7
hsp26
yellow
TART
1360
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Primer sets

HP1 chromatin association is increased at piRNA clusters in somatic tissues of RNA silencing mutants

In order to further examine the heterochromatic nature of piRNA clusters at higher resolution, ChIP assays were performed in adult heads to assess HP1 association with two piRNA clusters, *flam* and *80EF*, in the soma. Genomic locations of primer sets that uniquely amplify regions spanning these piRNA clusters are indicated in Figure 2-1A-B. As positive controls, primers for two transposable elements known to recruit HP1, *TART*, a telomere-specific non-LTR retrotransposon, and *1360*, a DNA transposon, were also tested (Fanti et al. 1998; Sun et al. 2004). Euchromatic genes *hsp26* and *yellow* were also included in the analysis as negative controls for HP1 association.

In wildtype fly heads, HP1 is observed at or near locations that give rise to piRNAs and endo-siRNAs at both *flam* and *80EF* loci. ChIP was performed using α-HP1 antibodies in chromatin prepared from wildtype heads, and the amount of DNA associated was determined by quantitative PCR using specific primer sets. As expected, low levels of *hsp26* and *yellow* are immunoprecipitated with HP1, while *TART* and *1360* levels are enriched above the euchromatic genes by over six-fold (Figure 2-4). At *flam*, HP1 associates with the majority of regions that produce high levels of piRNAs or endo-siRNAs approximately two to three-fold over the euchromatic sites (Figure 2-4A, primer sets 1-15). Similarly, at *80EF*, HP1 immunoprecipitates piRNA and endo-siRNA producing regions two to three-fold higher than the negative controls indicating the presence of heterochromatic marks at these loci (Figure 2-4B, primer sets G-M). Regions flanking these areas display approximately one to two-fold enrichment over euchromatic

sites, which may be due to tapering of HP1 spreading (Figure 2-4B, primer sets A-F and N-P). ChIP using antibodies directed against the chromatin insulator protein Su(Hw) verified its presence at known insulator sequences *gypsy* and *1A-2* (Parnell et al. 2003) but only background levels at *TART*, *1360*, and piRNA clusters, indicating the specificity of HP1 association at these sites (Figure 2-5). Rabbit IgG negative control immunoprecipitations yielded negligible amounts of DNA for all sites tested (<0.3% input).

Consistent with the transcriptional reporter assay, RNA silencing mutants display elevated levels of HP1 at piRNA clusters. ChIP of HP1 was performed in $piwi^1/piwi^2$ mutant heads, and similar levels at positive and negative controls were obtained compared to wild type (Figure 2-4). In contrast, at the *flam* locus, a two to five-fold increase in HP1 levels is observed at the centromere proximal side of the locus compared to wild type (Figure 2-4A, primer sets 6-15). Little change in HP1 recruitment is observed at the centromere distal end of *flam* in $piwi^1/piwi^2$ mutants (Figure 2-4A, primer sets 1-5). At *80EF*, HP1 levels increase two to three-fold in $piwi^1/piwi^2$ mutants compared to wild type across all primer sets examined (Figure 2-4B, primer sets A-P). In order to address differences in strain background and potential accumulation of TEs in *piwi* mutant strains, we performed ChIP assays comparing $piwi^1/piwi^2$ mutants to a $piwi^1/+$ heterozygous strain and obtained similar results (Figure 2-6).

ChIP experiments performed in $AGO2^{51B}$ mutant heads show a similar overall increase of HP1 at piRNA clusters compared to $piwi^1/piwi^2$ mutants. Levels of HP1 at *hsp26*, *yellow*, *TART*, and *1360* are similar in $AGO2^{51B}$ mutants and wild type while differences are apparent at piRNA clusters (Figure 2-4). At *flam*, $AGO2^{51B}$ mutants

display a two to seven-fold increase of HP1 association with the centromere proximal side compared to wild type (Figure 2-4A, primer sets 6-15). At the centromere distal end, no significant changes in HP1 levels are detected (Figure 2-4A, primer sets 1-5). For *80EF, AGO2*^{51B} mutants show similar levels of HP1 to wild type at the centromere distal end (Figure 2-4B, primer sets A-D) while an approximately two to five-fold increase of HP1 is detected in the remainder of the regions tested (Figure 2-4B, primer sets E-P). Moreover, ChIP assays in *AGO2*^{51B} homozygous mutants compared to an *AGO2*^{51B}/+ heterozygous strain produced similar results (Figure 2-7). Similar to *AGO2*^{51B} mutants, *Dcr-2*^{L811fsX} mutants show an increase of HP1 at regions that produce small RNAs compared to wild type (Figure 2-3 (B-C)). HP1 protein levels in wildtype, *piwi*^{1}*/piwi*^{2}, and *AGO2*^{51B} fly heads are similar indicating that the increased chromatin association observed is not due to an increased amount of HP1 (Figure 2-8). The increased HP1 chromatin association with piRNA clusters in RNA silencing mutants compared to wild type is consistent with increased silencing of P element insertions, and these results suggest that at least some of the observed effects on reporter gene expression in RNA silencing mutants are due to chromatin related events. Taken together, these data suggest an antagonistic effect of Piwi, Aub, and AGO2 on HP1 recruitment to chromatin in somatic tissue.

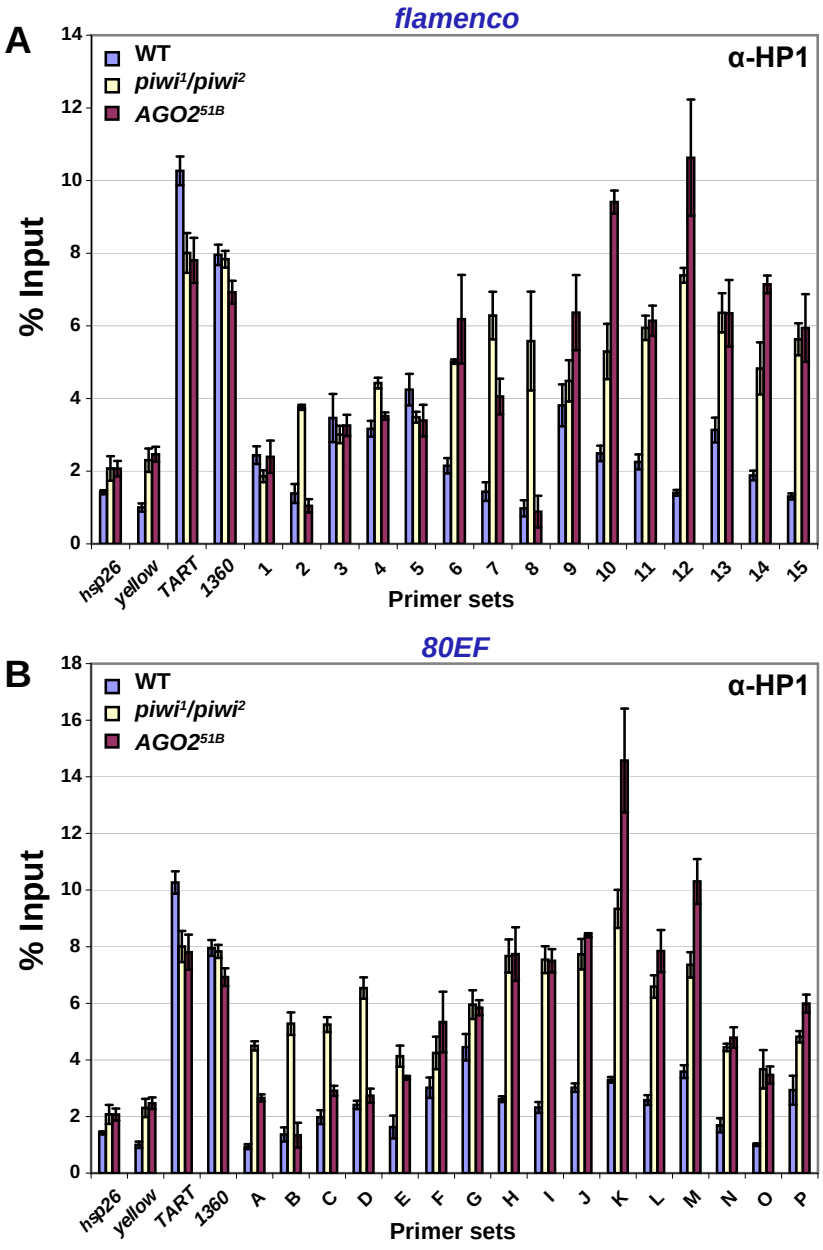


Figure 2-4. HP1 associates with chromatin at piRNA clusters, and its levels increase in RNA silencing mutants.

ChIP at (A) *flam* and (B) *80EF* piRNA clusters in wild type (blue), *piwi*^{1}/*piwi*^{2} (yellow), and *AGO2*^{51B} (red) mutants from adult heads with antibodies specific to HP1. Percent input immunoprecipitated is shown for each primer set, and error bars indicate standard deviation of quadruplicate PCR measurements.

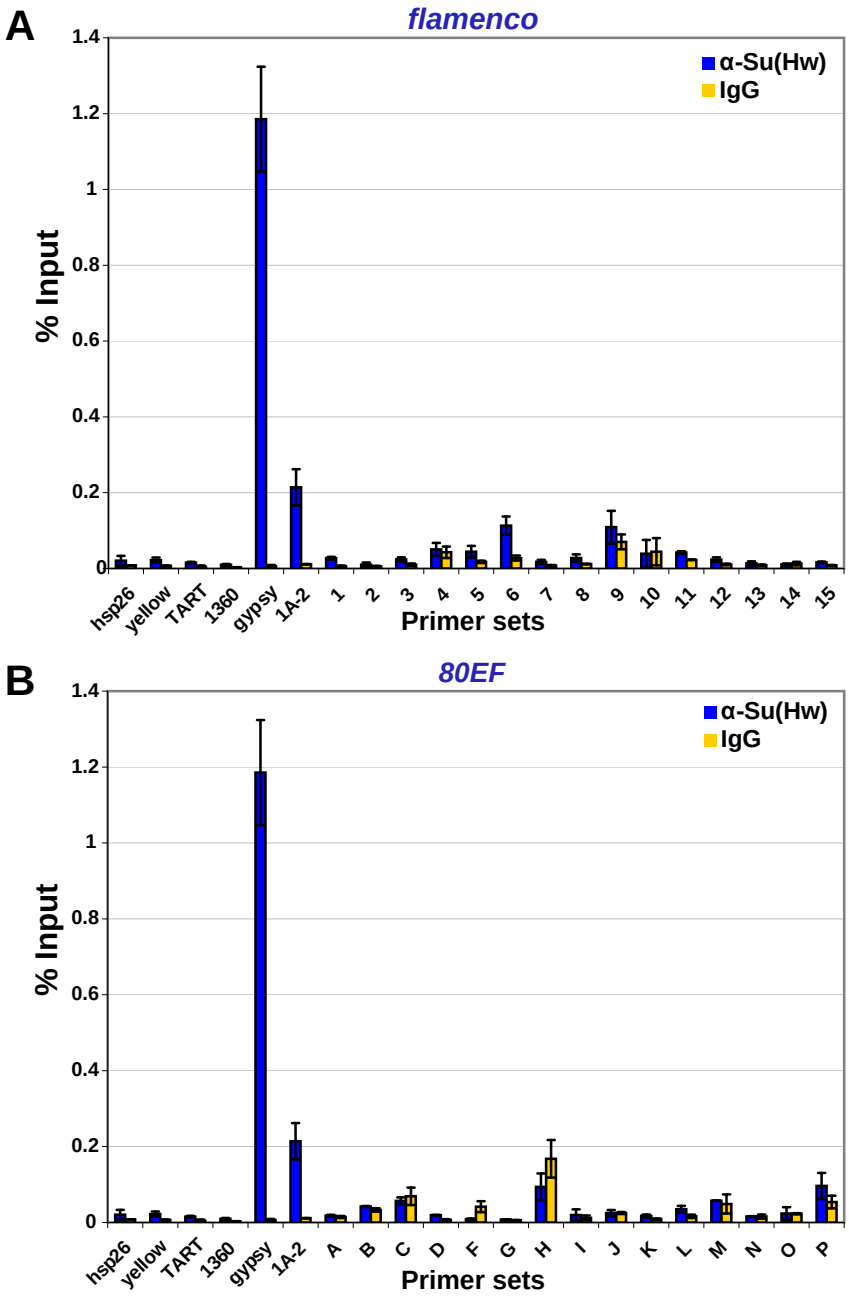


Figure 2-5. Su(Hw) does not associate with chromatin at piRNA clusters in heads. ChIP at (A) flam and (B) 80EF piRNA clusters in wild type with antibodies specific to Su(Hw) (blue) and rabbit normal serum (yellow). Percent input immunoprecipitated is shown for each primer set, and error bars indicate standard deviation of quadruplicate PCR measurements.

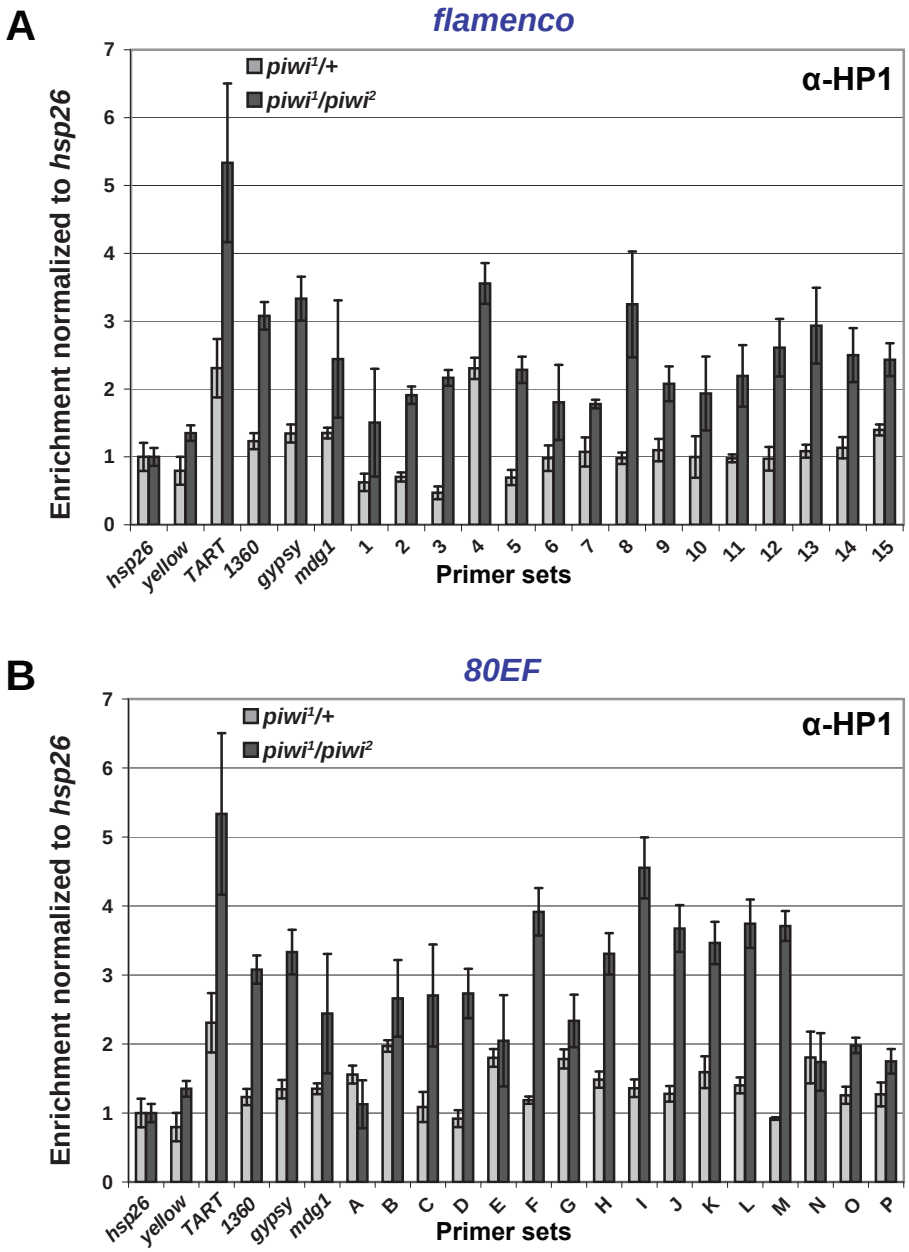


Figure 2-6. HP1 chromatin association levels are increased in piwi mutants at piRNA clusters.

ChIP at (A) *flam* and (B) *80EF* piRNA clusters in *piwi*^{1}/+ (light grey) and *piwi*^{1}/*piwi*^{2} (dark grey) from adult heads with antibodies specific to HP1. Values shown are percent input immunoprecipitated for each primer set normalized to *hsp26*. Error bars indicate standard deviation of quadruplicate PCR measurements.

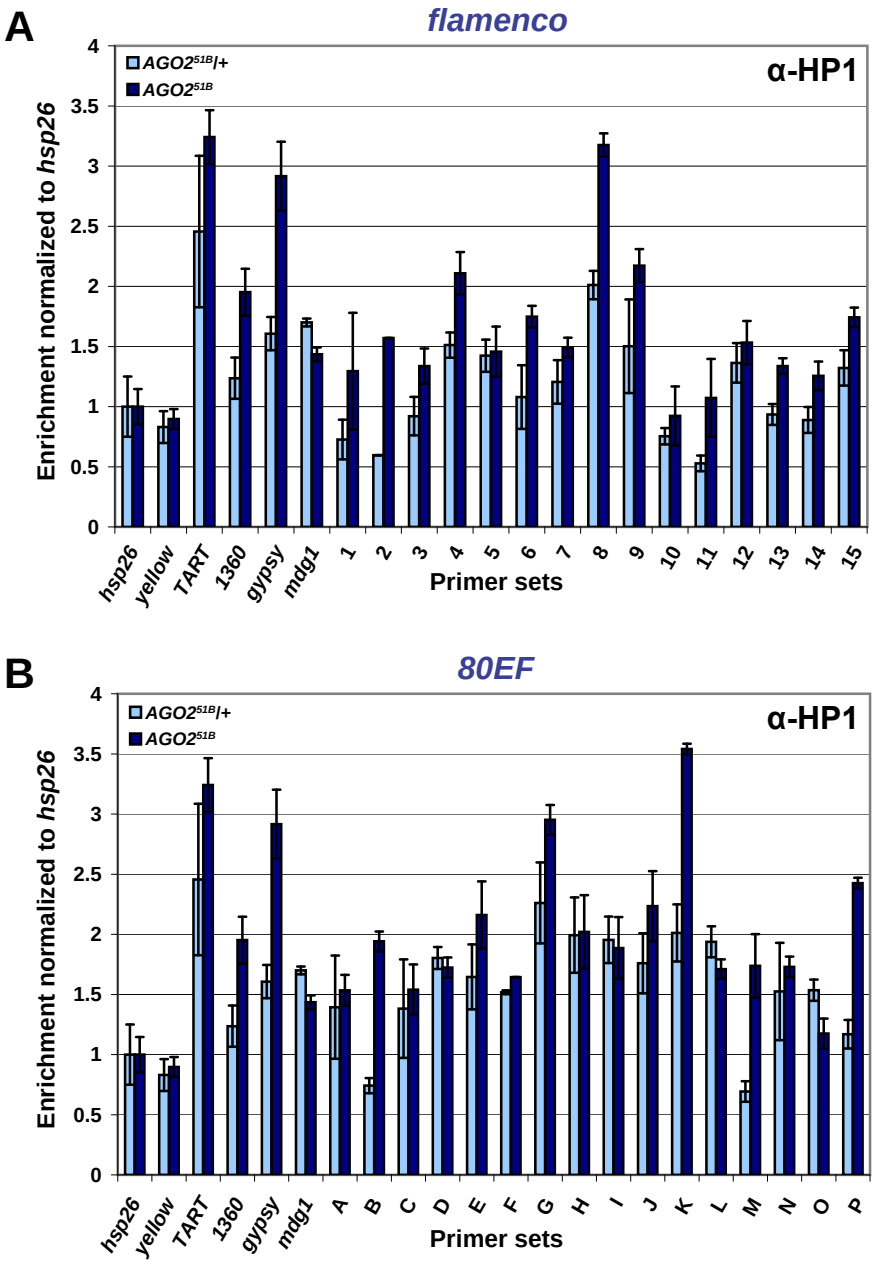


Figure 2-7. HP1 chromatin association levels are increased in AGO2 mutants at piRNA clusters.

ChIP at (A) *flam* and (B) *80EF* piRNA clusters in $AGO2^{51B}$/+ (light blue) and $AGO2^{51B}$ (dark blue) from adult heads with antibodies specific to HP1. Values shown are percent input immunoprecipitated for each primer set normalized to *hsp26*. Error bars indicate standard deviation of quadruplicate PCR measurements.

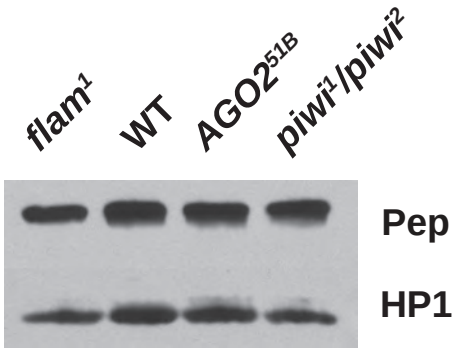


Figure 2-8. HP1 protein levels in wildtype, *flam*^{1}, *AGO2*^{51B} and *piwi*^{1}/*piwi*^{2} fly heads. Total protein was extracted from twenty adult heads by homogenization in RIPA buffer and separated by SDS-PAGE. Immunoblotting of HP1 and Protein on Ecdysone Puffs (Pep), a nuclear protein serving as a loading control, is shown.

HP1 also associates with piRNA clusters in ovaries

Given the evidence that transposable elements are mainly silenced in the gonad via piRNA pathways and in the soma via the endo-siRNA pathway, we wanted to determine whether HP1 also associates with piRNA clusters in gonadal tissues. Therefore, we investigated HP1 recruitment to piRNA clusters in wildtype ovaries by ChIP. As in heads, low levels of *hsp26* and *yellow* are immunoprecipitated with HP1, whereas *TART* and *1360* levels are enriched above the euchromatic genes by over ten-fold (Figure 2-9). At the *flam* locus, a four to fifteen-fold increase over the euchromatic sites in HP1 levels is observed at most sites at the centromere proximal side of the locus (Figure 2-9A, primer sets 4-15). Similarly, at *80EF*, HP1 immunoprecipitates small RNA producing regions two to twenty-fold higher than euchromatic sites indicating the presence of heterochromatic marks at these loci (Figure 2-9B, primer sets A-P). Rabbit IgG negative control immunoprecipitations yielded negligible amounts of DNA for all sites tested. We were unable to immunoprecipitate DNA at levels above background from either heads or whole ovaries using multiple antibodies to Piwi, Aub, AGO3, and AGO2 that have been used in previous studies for immunoprecipitation or immunofluorescence (data not shown; Miyoshi et al. 2005; Saito et al. 2006; Brower-Toland et al. 2007; Gunawardane et al. 2007).

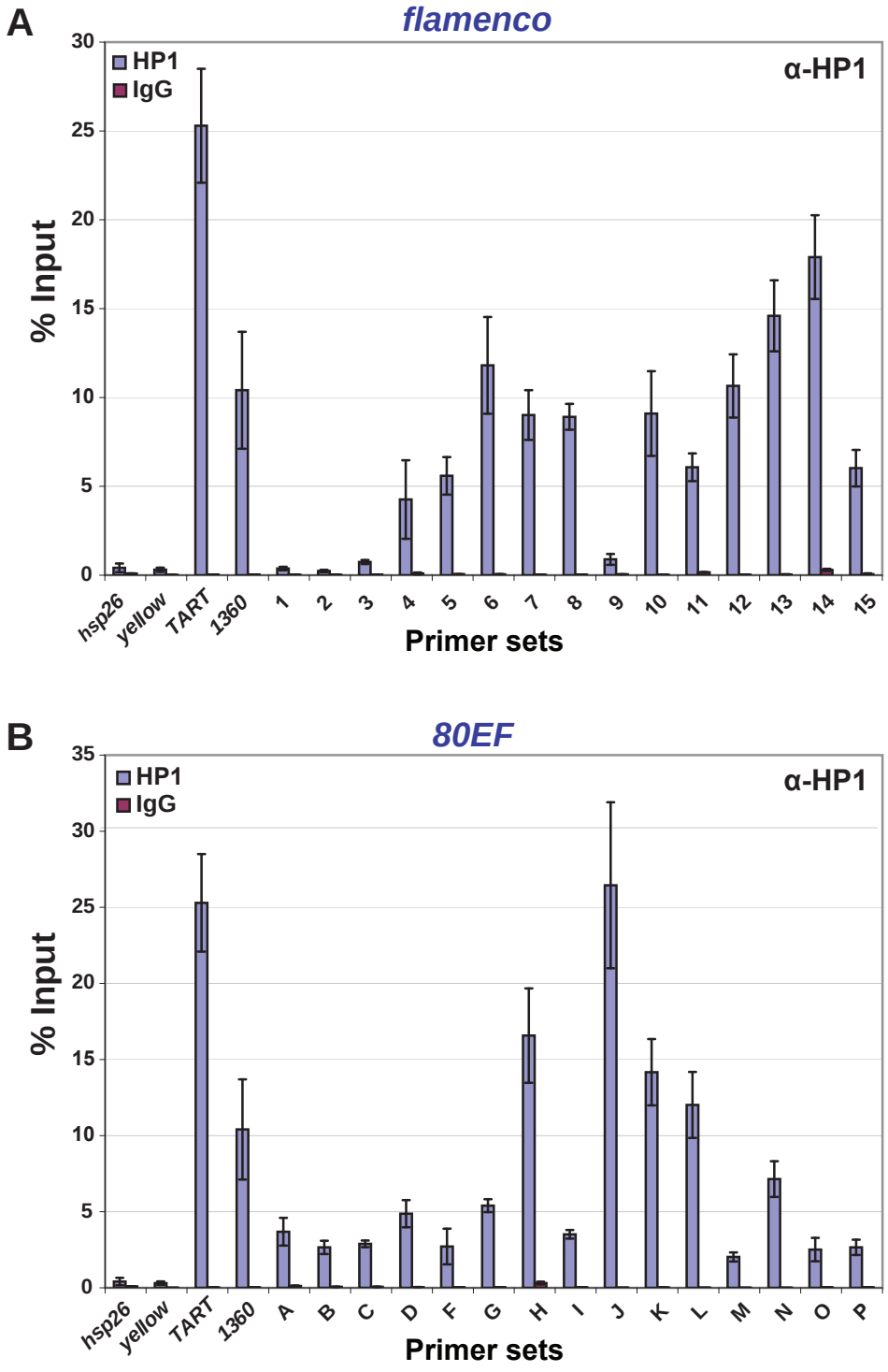


Figure 2-9. HP1 associates with chromatin at piRNA clusters in ovaries.

ChIP at (A) *flam* and (B) *80EF* piRNA clusters in wildtype ovaries with antibodies specific to HP1 (blue) and normal rabbit IgG (red).

HP1 chromatin association is not affected greatly by depletion of Piwi in somatic ovarian follicle cells

We wished to address whether HP1 association with piRNA clusters is dependent on Piwi in the gonad, which express high levels of both proteins. Due to a complete loss of germ cells and the severe underdevelopment of ovary tissue in *piwi* mutants, it was not possible to obtain enough mutant material to perform ChIP. Therefore, we examined the recruitment of HP1 to chromatin in an ovarian somatic follicle cell line (OSC) that expresses Piwi but not Aub or AGO3 and produces only primary piRNAs, a large proportion of which derive from the *flam* locus (Saito et al. 2009). The majority of Piwi was depleted from OSC cells by siRNA-mediated knockdown, and depletion of Piwi does not affect HP1 or Lamin protein levels compared to mock transfected cells (Figure 2-10A).

Subsequently, we investigated HP1 recruitment to piRNA clusters by ChIP in OSC cells. In mock treated cells, low levels of *hsp26* and *yellow* are immunoprecipitated with HP1, while *TART* and *1360* levels are enriched above the euchromatic genes by 1.5- to over two-fold (Figure 2-10(B-C)). Two additional TEs tested, *gypsy* and *mdg1*, are immunoprecipitated at similar levels to *TART* with HP1 (Figure 2-10(B-E)). At *flam*, HP1 associates with the piRNA cluster similar to TE levels (Figure 2-10(B-C)). Despite much lower piRNA production from the *80EF* cluster in OSC compared to *flam* (Saito et al. 2009), HP1 associates with piRNA producing regions of *80EF* at similar levels to *flam* and TEs (Figure 2-10C, primer sets A-P). Overall, the HP1 recruitment profile in OSC is similar to that of heads and whole ovaries albeit at lower relative levels. In Piwi

knockdown cells, no significant differences are seen for HP1 recruitment to all sites compared to mock treated cells except a two-fold decrease at the *1360* element. Rabbit IgG negative control immunoprecipitations yielded low amounts of DNA for all sites tested (<0.06% and <0.07% input for mock and Piwi knockdown cells, respectively).

Importantly, Piwi association with chromatin is detectable in OSC cells, but its profile differs from that of HP1. In mock treated cells, antibodies directed against Piwi (Saito et al. 2006) immunoprecipitate euchromatic sites at levels similar to that of TEs (Figure 2-10(D-E)). Furthermore, the majority of regions producing piRNA at *flam* is also immunoprecipitated at comparable levels to both euchromatic sites and TEs (Figure 2-10D). Moreover, levels of Piwi association with *80EF* is akin to that of *flam*, while several sites in both *flam* and *80EF* clusters show particular enrichment of Piwi up to three-fold compared to the average association with other sites tested (Figure 2-10(D-E)). In Piwi knockdown cells, Piwi chromatin association drops two to five-fold, down to background levels at all sites except for some residual association with two sites in or near the *flam* locus. Mouse IgG negative control immunoprecipitations yielded low amounts of DNA in comparison to α-Piwi immunoprecipitations in mock treated cells for all sites tested (<0.04% and <0.02% input for mock and Piwi knockdown cells, respectively). We conclude that in ovarian somatic follicle cells, reduction of the total pool of Piwi as well as the chromatin bound fraction does not affect HP1 association with piRNA clusters and has a minimal effect on HP1 association with TE chromatin association.

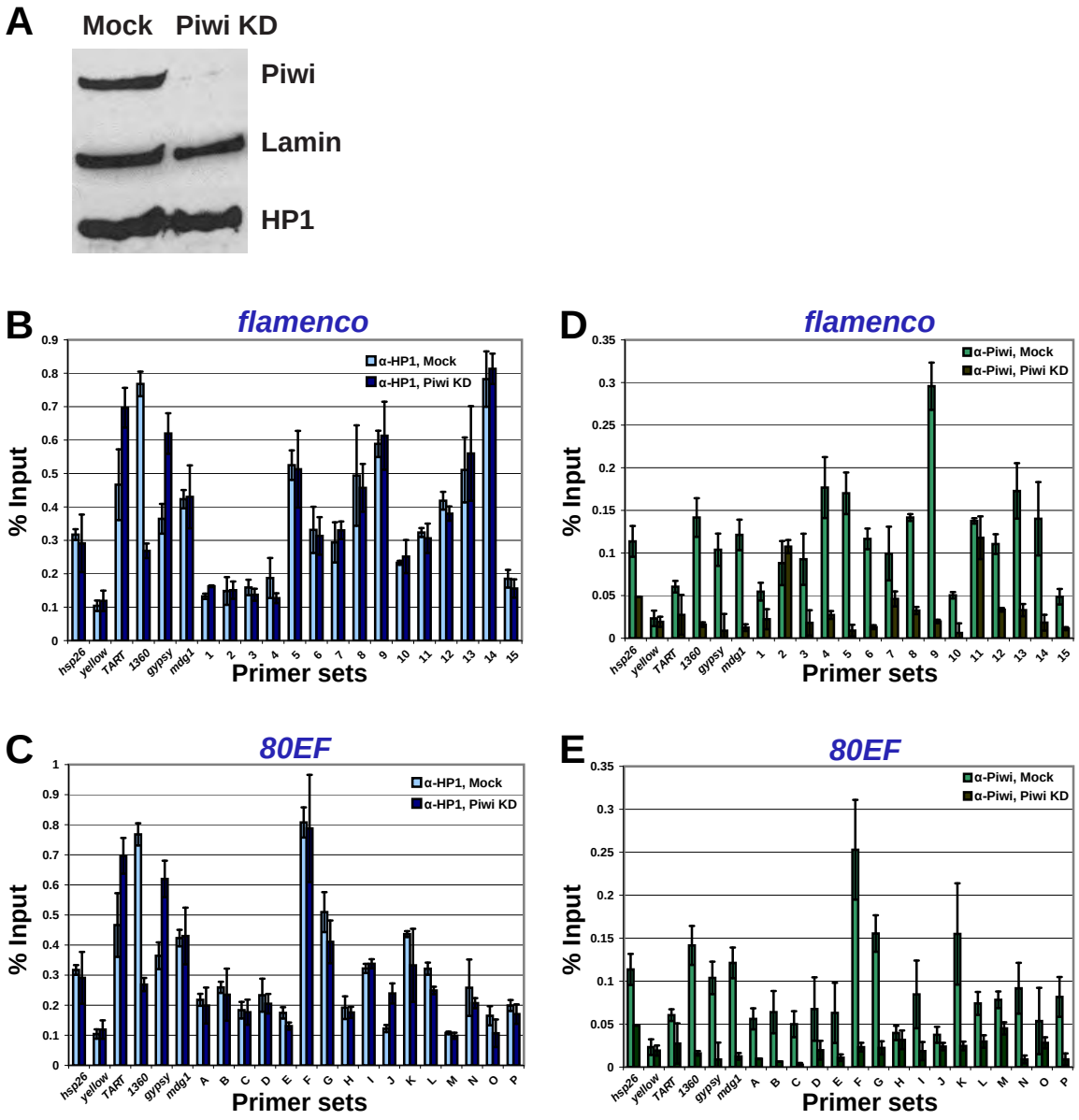


Figure 2-10. Depletion of Piwi from ovarian somatic follicle cells does not affect HP1 recruitment to piRNA clusters.

(A) Western blotting of Piwi, HP1 and Lamin in OSC cells that were either mock treated (left lane) or treated with siRNA directed against *piwi* (right lane, Piwi KD).

ChIP at *flam* (B, D) and *80EF* (C, E) piRNA clusters in mock treated and Piwi KD OSC cells with antibodies specific to HP1 (B-C) or Piwi (D-E).

Loss of piRNA production from a single cluster results in global HP1 mislocalization

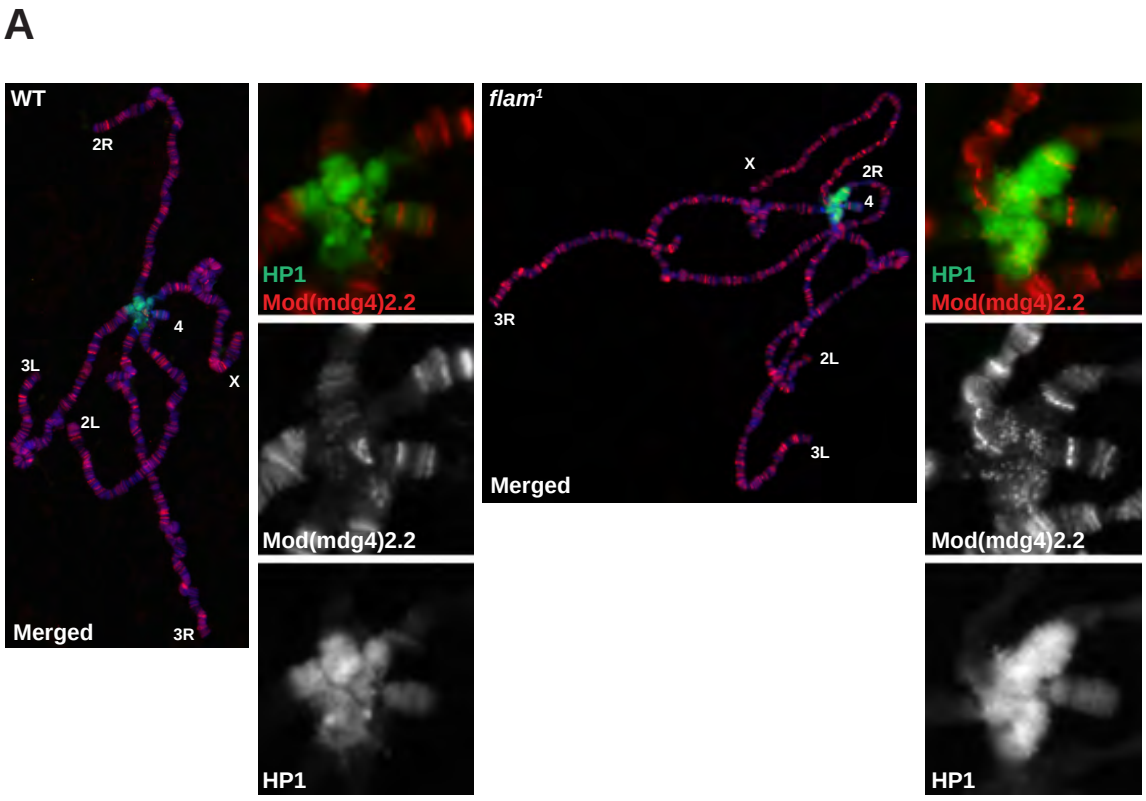
We next sought to determine whether loss of piRNA production at a single piRNA cluster would affect HP1 recruitment to chromatin. Previous studies have shown that mutation of various RNA silencing components results in global mislocalization of HP1 on polytene chromosomes (Pal-Bhadra et al. 2004; Fagegaltier et al. 2009). Mutation of *flam* has been previously shown to result in loss of piRNA production (Brennecke et al. 2007) and upregulation of the *gypsy* retroelement (Prud'homme et al. 1995). In order to obtain a genome-wide view of HP1 chromatin association in *flam* mutants, we examined the localization of HP1 to highly replicated salivary gland polytene chromosomes from either wild type or $flam^{1}$ mutant third instar larvae by indirect immunofluorescence using α-HP1 antibodies. In wild type, HP1 localizes predominantly to a concentration of heterochromatin where the centromeres of each chromosome coalesce, termed the chromocenter (Figure 2-11A, green). In contrast, $flam^{1}$ mutants display expansion of HP1 at the chromocenter. Spreading of HP1 is apparent on the second and third chromosomes, but not on the X chromosome, where *flam* is located. As a reference, we also examined the localization of the chromatin insulator protein Mod(mdg4)2.2, which is unchanged in localization between wild type and $flam^{1}$ (Figure 2-11A, red). The extent of HP1 chromocenter expansion is comparable to the level of HP1 expansion that we observe in $spn\text{-}E^{hlsE1}/spn\text{-}E^{hlsE616}$ mutants (Figure 2-12). A lesser degree of HP1 expansion was also observed in $flam^{BG02658}/flam^{KG00476}$ mutants (data not shown). Finally, total HP1 levels are unchanged in $flam^{1}$ whole flies compared to wild

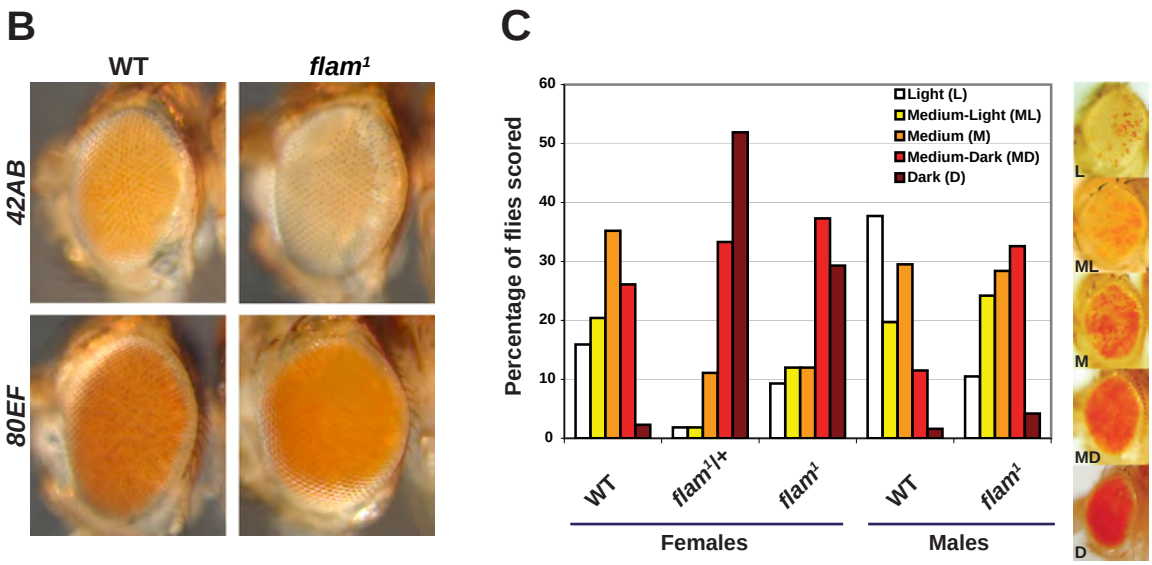
type (Figure 2-8). These results indicate a global change in HP1 localization resulting from inactivation of a single piRNA cluster.

We reasoned that accumulation of HP1 at the chromocenter of *flam*^{1} mutants may result in an increase in silencing at pericentromeric sites. Therefore, the expression of transcriptional reporters at *42AB* or *80EF* piRNA clusters, which are located on different chromosomes from the *flam* locus, was examined in *flam*^{1} mutants. Compared to wild type, *flam*^{1} mutants harboring a P element insertion at either *42AB* or *80EF* piRNA clusters display mildly decreased pigmentation suggesting increased silencing at these distinct pericentromeric loci (Figure 2-11B).

Figure 2-11. Mutation of the *flam* piRNA cluster results in global HP1 redistribution.

(A) Wild type (left) and *flam*^{1} (right) polytene chromosomes stained with antibodies directed against HP1 (green) and a reference protein Mod(mdg4)2.2 (red). DNA is stained with DAPI (blue). Chromosome arms are labeled, and insets of the enlarged chromocenter are shown. (B) Adult eyes of wild type and *flam*^{1} mutants carrying a *mini-white* transgene inserted in *42AB* (top row) and *80EF* (bottom) piRNA clusters. (C) Degree of eye pigmentation due to expression of the *DX1* transgene array at 50C on chromosome 2L, which undergoes repeat-induced heterochromatic silencing, in wildtype, *flam*^{1}/+, and *flam*^{1} female flies and wildtype and *flam*^{1} male flies. Scoring of variegation in the eye is categorized into five groups that range between light (few pigmented facets) to dark (almost all pigmented facets). Percentage of flies falling into each category was graphed. Representative eyes are shown on right.


A
WT
2R
4
3L
X
2L
Merged
3R
HP1
Mod(mdg4)2.2
Mod(mdg4)2.2
HP1
flam1
X
2R
4
3R
2L
3L
Merged
HP1
Mod(mdg4)2.2
Mod(mdg4)2.2
HP1


B
WT
flam1
42AB
80EF
C
Percentage of flies scored
60
50
40
30
20
10
0
Light (L)
Medium-Light (ML)
Medium (M)
Medium-Dark (MD)
Dark (D)
WT
flam1/+
flam1
WT
flam1
Females
Males
L
ML
M
MD
D

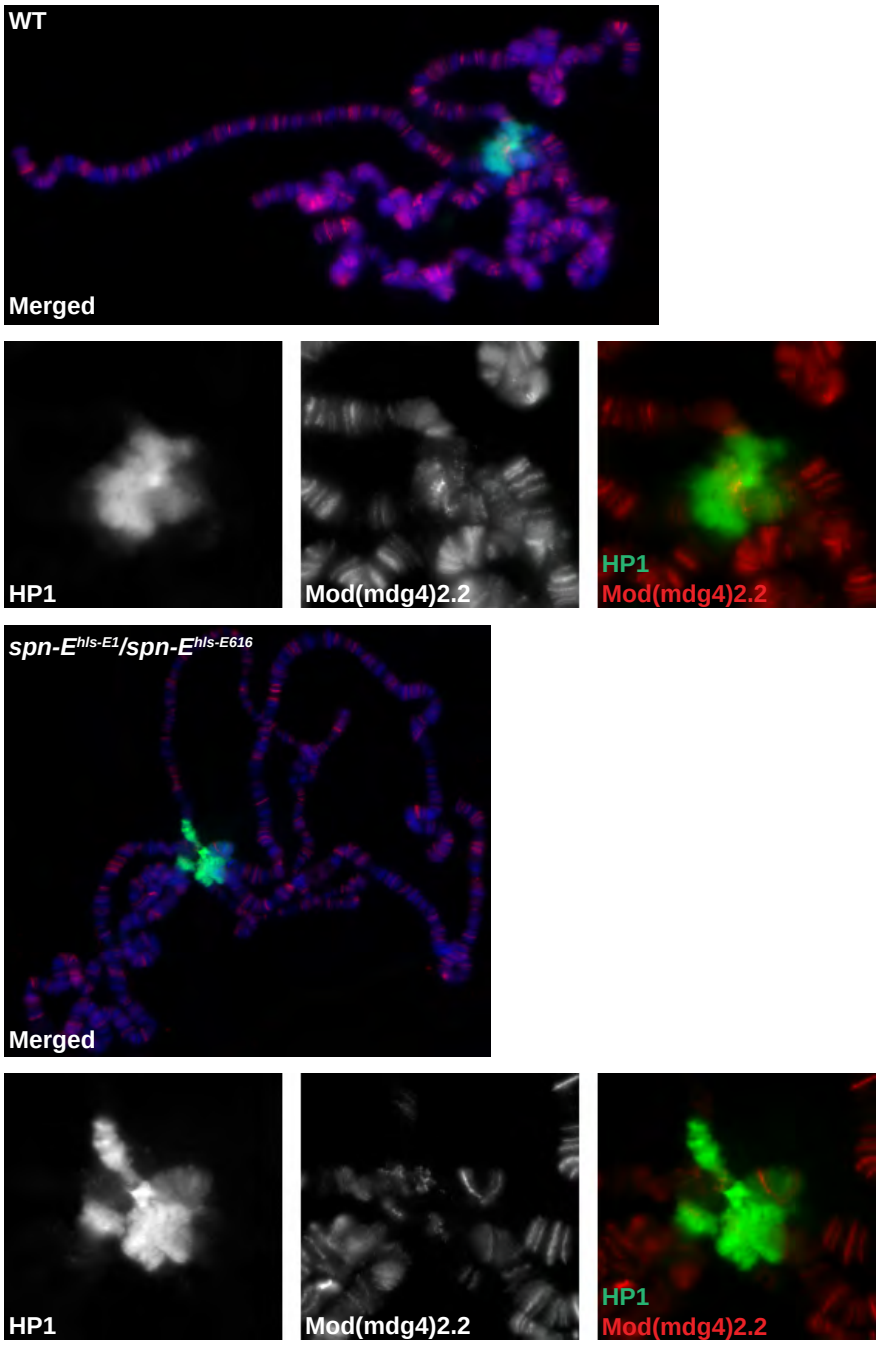


Figure 2-12. *spn-E*$^{hls-E1}$*/spn-E*$^{hls-E616}$ mutants display accumulation of HP1 at the chromocenter.

Wild type and *spn-E*$^{hls-E1}$*/spn-E*$^{hls-E616}$ polytene chromosomes stained with C1A9 antibody directed against HP1 (green) or a reference protein Mod(mdg4)2.2 (red). DNA is stained with DAPI (blue).

Mutation of the *flam* piRNA cluster suppresses heterochromatic silencing at a distant site

Finally, to verify HP1 genome-wide redistribution in *flam*^{1}mutants, we examined the effect of *flam*^{1} on the silencing of a centromere distal heterochromatic site on a different chromosome. The *DX1* transgene array consists of seven *mini-white* P elements with one inverted copy at a normally euchromatic site at 50C on chromosome 2R (Dorer and Henikoff 1994). Due to this configuration, the array forms ectopic repeat induced heterochromatin and displays a variegated phenotype similar to PEV that is dependent on HP1. Expression of the *DX1* array was assessed based on variegation of eye pigmentation in wild type, heterozygous *flam*^{1}/+, and homozygous or hemizygous *flam*^{1} mutants (Figure 2-11C). Due to a wide range of eye coloration, variegation was scored by categorization into five groups that ranged between Light (few pigmented facets) to Dark (almost all facets pigmented). For females, 3% of wild type was classified as Dark, while 29% of *flam*^{1}/+ and 52% of *flam*^{1} mutants displayed the same high level of pigmentation. In males, 15% of wild type was scored as Medium-Dark or Dark while 40% of *flam*^{1} males fell into these categories. These results indicate that mutation of *flam* can suppress heterochromatic silencing in *trans*. Taken together with the HP1 centromeric expansion in polytene chromosomes and increased pericentromeric silencing in *flam*^{1} mutants, there appears to be a global redistribution of HP1 resulting from the loss of piRNA production from a single locus.

DISCUSSION

In this study, we tested directly whether the Argonautes AGO2 or Piwi recruit HP1 to chromatin. As candidate sites for Argonaute/HP1 interaction, we examined whether piRNA clusters may be heterochromatic using both genetic and molecular approaches. First, P elements inserted at or near pericentromeric piRNA clusters were assayed as transcriptional reporters, and these transgenes were found to display variegated expression that is increased in heterochromatin mutants. Next, ChIP with α-HP1 antibodies showed that HP1 associates with piRNA clusters at levels significantly above euchromatic sites. However, mutation of *piwi*, *aub*, or *AGO2* leads to a modest increase in silencing of transcriptional reporters as well as an increase of HP1 association at piRNA clusters in heads. In ovarian somatic follicle cells, in which both Piwi and HP1 are highly expressed, depletion of Piwi results in little or no change in HP1 recruitment to piRNA clusters and TEs. Furthermore, loss of piRNA production at a single locus results in expansion of HP1 at the centromere. In these $flam^{1}$ mutants, silencing of a distant heterochromatic transgene array is reduced, further indicating a global redistribution of HP1 and suggesting indirect effects. Taken together, the results argue against direct recruitment of HP1 or maintenance of its association by AGO2 or Piwi in the soma.

AGO2 and Piwi are not required for HP1 association at piRNA clusters

Several reasons dictated the choice of piRNA clusters as the focus of our analyses. First, both endo-siRNAs and piRNAs are generated from these loci (Brennecke

et al. 2007; Yin and Lin 2007; Chung et al. 2008; Czech et al. 2008; Ghildiyal et al. 2008; Kawamura et al. 2008). Next, we reasoned that at least some piRNA clusters are likely to be heterochromatic because of their strong bias toward TE-rich pericentromeric positions in the genome (Brennecke et al. 2007; Yin and Lin 2007), in close proximity to the vast majority of HP1 localization. In fact, early cloning attempts determined that the *flam* locus is located in a repetitive, TE rich heterochromatic region (Robert et al. 2001). Furthermore, the pericentromeric position of these clusters likely coincides with the transition between euchromatin and heterochromatin, corresponding to the borders of HP1 spreading. This characteristic allows variegation assays, which monitor the variable spreading of HP1 and heterochromatin, to be extremely sensitive. ChIP assays at the borders of HP1 spreading would also likely be optimally sensitive to both local and overall changes in HP1 chromatin association. Finally, piRNA clusters contain enough unique sequence for specific primer design and monitoring by directed ChIP analysis.

Given that *AGO2* is the predominant Argonaute expressed outside the gonad that participates in the silencing of TEs in the soma, we tested whether AGO2 could recruit HP1 to chromatin in somatic tissue. Moreover, it has been shown that *AGO2* mutants exhibit mislocalization of HP1 (Deshpande et al. 2005; Fagegaltier et al. 2009). However, our results show that mutation of *AGO2* results in a strong increase of silencing of transcriptional reporters at or near piRNA clusters and a mild increase of HP1 chromatin association in heads. Given the extent of increased silencing in the *AGO2* mutant compared to *piwi* or *aub* mutants, which accumulate HP1 on chromatin to a similar degree, a posttranscriptional step of silencing likely contributes to the negative effects observed on transcriptional reporters. *AGO2* mutants show a plethora of cellular

defects during early nuclear divisions but develop normally and are fertile suggesting that effects on these various processes as well as HP1 localization are mild or otherwise compensated for (Deshpande et al. 2005). Therefore, *AGO2* is unlikely to be required for HP1 recruitment in this tissue.

Additionally, we find that HP1 association at piRNA clusters does not depend on the presence of Piwi. Our analysis of piRNA clusters included *flam*, a primary piRNA cluster, and *80EF*, a germline piRNA producing locus. We examined both *flam* and *80EF* clusters in somatic head tissue and ovaries, which are a mixed population of somatic follicle and germline derived cells. In heads, there is no apparent requirement for *piwi* with respect to HP1 recruitment to the piRNA clusters or to TEs that were examined.

In our study, Piwi chromatin association was detected only in OSC cells, and its presence is dispensable for HP1 chromatin association. The *flam* piRNA cluster produces high levels of primary piRNA in OSC while *80EF* is active for piRNA production in germ cells but not in OSC (Li et al. 2009a; Malone et al. 2009; Saito et al. 2009). Nonetheless, Piwi associates with both the *flam* and *80EF* clusters at comparable levels, suggesting that the amount of piRNA production from a particular locus does not correlate with Piwi chromatin association. Furthermore, the pattern of Piwi chromatin association in OSC differs from that of HP1 in that there is no particular enrichment of Piwi at TEs above euchromatic sites and only a minor accumulation at a few sites in the *flam* and *80EF* piRNA clusters. When Piwi levels were reduced by siRNA knockdown, Piwi chromatin association was essentially abolished but HP1 recruitment was not affected except for a two-fold decrease over the *1360* element. Previous studies suggested that the *1360* element may be responsible for nucleating heterochromatin on

the largely heterochromatic fourth chromosome and further showed that mutation of factors representing all RNA silencing pathways, *piwi*, *aub*, *spn-E*, *Dcr-1*, and *Dcr-2*, affect *1360* dependent heterochromatic silencing (Sun et al. 2004; Haynes et al. 2006). Unlike the results in adult heads, no accumulation of HP1 over piRNA clusters was detected as a result of Piwi knockdown in OSC cells. This discrepancy may reflect differential effects in distinct cell types or the length of the Piwi knockdown in OSC cells, which was at least adequate to essentially eliminate Piwi chromatin association. In a related but independently derived ovarian somatic follicle cell line (OSS), Piwi and HP1 do not colocalize in the nucleus (Lau et al. 2009), and this finding supports the conclusion that Piwi does not direct HP1 recruitment in this cell type. Also consistent with our results, HP1 remains localized to the chromocenter in salivary gland polytene chromosomes in *piwi* null mutants (Pal-Bhadra et al. 2004; Brower-Toland et al. 2007). We conclude that association of HP1 with chromatin can occur independently of *AGO2* and *piwi* in somatic tissue.

A previous study addressed the role of the germline piRNA pathway in HP1 association with transposable elements. The *spn-E* gene controls predominantly germline piRNA production but does not affect the somatic piRNA pathway (Malone et al. 2009). ChIP was used to show that *spn-E* mutants display significantly decreased levels of H3K9me3 and HP1 at telomeric *Het-A* but similar to wildtype HP1 levels at the *I-element* and *copia* TEs, which are distributed throughout the genome (Klenov et al. 2007). This modest reduction of HP1 at *Het-A* was apparent in ovaries but not in carcasses, which contain only somatic tissue. One caveat to this study is that ChIP was performed using primers that detect all TEs matching a particular sequence, thus measuring average HP1

and H3K9me levels on TEs across the genome. Nonetheless, this work suggests a limited role for the germline piRNA pathway in HP1 recruitment at the telomere.

Additional candidate platforms for Piwi-dependent HP1 recruitment

Several studies have shown that Piwi associates with at least some heterochromatic sites in the genome, but direct evidence that any of these sites serve as recruitment platforms for HP1 and subsequent spreading is lacking. The best characterized Piwi-associated site is the heterochromatic 3R-TAS subtelomeric region, which generates the abundant Piwi bound 20nt 3R-TAS piRNA. Surprisingly, the role of *piwi* at this location is transcriptional activation, as *piwi* mutants display increased transcriptional silencing of a nearby reporter transgene as well as an increase of HP1 association at 3R-TAS (Yin and Lin 2007). Likewise, we observe a mild corresponding increase in HP1 association and silencing at piRNA clusters in *piwi* mutants suggesting that *piwi* function could in fact oppose HP1 recruitment at multiple sites in the genome. Our results are consistent with the possibility that piRNA clusters act as boundaries to the spread of pericentromeric heterochromatin. The mechanism of Piwi dependent transcriptional activation has not been determined, but considering that Piwi interacts with the chromoshadow domain of HP1 (Brower-Toland et al. 2007), Piwi may compete for binding with other HP1 interactors such as Su(var)3-9 that promote heterochromatic silencing.

Functions for *piwi* outside of the gonad

The majority of Piwi protein is found in both somatic and germline tissues of the gonad, yet *piwi* clearly exerts an effect on non-gonadal somatic tissues as well. RT-PCR analysis shows that the *piwi* transcript is readily detectable outside the gonad and in somatic cell lines (Rehwinkel et al. 2006; Brower-Toland et al. 2007), but the Piwi protein is difficult to detect (Brower-Toland et al. 2007). Nevertheless, mutation of *piwi* suggests important functions for this gene outside of the gonad. For example, *piwi* is essential for viability, and loss-of-function mutants display a variety of phenotypes manifest in various non-gonadal somatic tissues such as demonstrated in this study and others, which show a requirement for *piwi* in pairing-dependent silencing, nucleolar integrity, and chromatin insulator function (Pal-Bhadra et al. 2002; Grimaud et al. 2006; Lei and Corces 2006; Peng and Karpen 2007). For each of these chromatin related studies, it remains a possibility that even a small amount of maternally deposited Piwi could trigger early events in the oocyte or embryo that persist throughout development, manifesting phenotypes visible in adult somatic tissues.

HP1 redistribution in piRNA pathway mutants

Our results along with previous studies have demonstrated that HP1 mislocalizes from the chromocenter in a subset of piRNA pathway mutants. We found that polytene chromosomes of $flam^{1}$ mutants exhibit expanded HP1 chromocenter distribution. This result is intriguing because the $flam^{1}$ mutation affects a single piRNA cluster on the X chromosome but HP1 spreading to other chromosomes is apparent. A previous study

detected spreading of HP1 to euchromatic arms especially in *spn-E* mutants (Pal-Bhadra et al. 2004), and we confirmed this result albeit to a lesser degree, with spreading being comparable to the extent seen in *flam*^{1} mutants. Perhaps the increase of TE expression in RNA silencing mutants can stimulate HP1 recruitment and spreading from the centromere, which contains the highest concentration of TEs. In fact, transcription of pericentromeric repeats stimulates RNAi-dependent heterochromatin formation in fission yeast (Zofall and Grewal 2006; Chen et al. 2008; Kloc and Martienssen 2008).

Redistribution of HP1 in RNA silencing mutants may indirectly affect silencing at various heterochromatic locations in the genome. Seemingly inconsistent with HP1 spreading, *spn-E*, *aub*, and *piwi* mutants display decreased silencing of P element transgene arrays such as *DX1* and single insertions at pericentromeric regions on chromosomes 2 and 4 (Pal-Bhadra et al. 2004). In our study, we found that mutation of *flam* also results in loss of silencing at *DX1*, which is distant from the *flam* locus. This reduced silencing in *trans* could not be due to posttranscriptional events as there are no shared sequences between *DX1* and the *flam* locus. Therefore, we consider the possibility that there exists a finite pool of HP1 that accumulates at the centromere in *flam* and other RNA silencing mutants at the cost of reduced density and reduced silencing at other heterochromatic regions such as the transgene array, the fourth chromosome, and the telomere. The concept of a limited population of HP1 was suggested previously to explain the finding that the Y chromosome behaves as a suppressor of variegation by acting as a sink for HP1 (Dorer and Henikoff 1994).

Conclusions

Studies in multiple organisms have identified or suggested alternative mechanisms to RNA silencing for the recruitment of HP1 to chromatin. In fission yeast, overlapping and redundant RNAi-dependent and independent mechanisms of heterochromatin formation have been elucidated. In mouse cells, HP1 localization to pericentromeric heterochromatin was found to be RNase A sensitive suggesting that an RNA moiety may be involved in HP1 recruitment (Maison et al. 2002). Our data indicate that heterochromatin can form independently of RNA silencing in *Drosophila*. It will be interesting to determine if any of these alternative mechanisms of heterochromatin formation are conserved throughout evolution.

ACKNOWLEDGEMENTS

I would like to thank Ryan Dale for computational support and the generation of Figure 2-1 and C. Berg, A. Beyer, J. Birchler, V. Corces, J. Eissenberg, S. Elgin, G. Hannon, S. Henikoff, H. Lin, P. Macdonald, U. Schäfer, P. Schedl, and M. Siomi for fly strains and antibodies. I am indebted to K. Saito and M. Siomi for OSC cells and detailed culture protocols, Y. Zhang for ChIP protocols, and members of the Lei laboratory for discussions.

AUTHOR CONTRIBUTIONS

All the experiments were performed by N.M., and were designed by N.M. and E.P.L. E.P.L. contributed to the writing of the manuscript presented in this chapter.

MATERIALS AND METHODS

Drosophila strains

Fly stocks were maintained at 25°C on standard cornmeal medium. Lines containing *P{EPgy2}DIP1*^{EY02625} and *P{EPgy2}EY08366* were obtained from the Bloomington Drosophila Stock Center, and a line harboring *PBac(PB) c06482* was obtained from the Exelixis Collection at Harvard Medical School. Genomic coordinates of these P-element insertions were confirmed by PCR with primers specific to the P-elements and flanking genomic sequences followed by sequencing. For transcriptional reporter assays, transgenes were crossed or recombined into mutant backgrounds and scored against crosses to *yw*^{67c23} as a reference. For ChIP and immunofluorescence, Oregon-R was used as a wildtype control. The *y v f mal flam*^{1}*/FM3* stock was selected for heterozygous females each generation to prevent mobilization and accumulation of TEs. For the *DX1* variegation assay, *DX1/CyO* was crossed to *y w v f mal flam*^{1}*/FM7c; CyO/Sp* flies or *yw*^{67c23}*; CyO/Sp* as a reference.

Transcriptional reporter and eye pigmentation assays

Eye pigmentation of 40 to 60 adult males six days of age was examined, and representative eye photos were taken. To quantify overall levels of eye pigmentation, the heads of 25 male flies of each genotype were dissected, and eye pigmentation was measured as previously described (Pal-Bhadra et al. 2004). Briefly, heads were

homogenized in 0.8 ml of methanol, acidified with 0.1% HCl and centrifuged. The absorbance of the supernatant was measured at 480 nm.

Chromatin immunoprecipitation

Crosslinking and sonication

Wildtype (Oregon-R) heads or ovaries were dissected on dry ice. Fly heads/ovaries were washed in 5 ml PBS containing 0.01% Triton X-100 and centrifuged for 1 min at 500 rcf to pellet heads. Supernatant was discarded, and 1 ml of crosslinking solution (50 mM HEPES, pH 8.0, 1 mM EDTA, 0.5 mM EGTA, 100 mM NaCl, and 1.8% formaldehyde) and 3 mL n−heptane were added. The mixture was shaken vigorously for 20 min at room temperature. Supernatant was discarded, and heads were resuspended in 5 ml PBS containing 125 mM glycine and 0.01% Triton X-100. The mixture was shaken for 5 min at room temperature. Supernatant was discarded after centrifugation, and 5 ml of ice cold PBS containing 0.01% Triton X-100 was added. Supernatant was removed and heads were resuspended in 5 ml ice cold PBS containing 0.01%Triton X-100 and protease inhibitors (Roche). Heads were Dounce homogenized with pestle A (Kontes) for tissue disaggregation and complete homogenization. The mixture was centrifuged at 400 rcf for 1 min, and supernatant was transferred to a fresh tube. Supernatant was centrifuged at 9190 rcf for 5 min at 4°C and supernatant discarded afterwards. The pellet was resuspended in 5 ml ice cold Cell Lysis Buffer (5 mM PIPES, pH 8, 85 mM potassium chloride, 0.5% Nonidet P40 (NP40) and protease inhibitors).

The mixture was Dounce homogenized with pestle B to release the cell nuclei and centrifuged at 9190 rcf for 5 min at 4°C. Supernatant was removed, and the pellet was resuspended in 1 ml of ice cold Nuclear Lysis Buffer (50 mM Tris HCl, pH 8.0, 10 mM EDTA, 1% SDS and protease inhibitors) and incubated for 20 min at 4°C. Then 0.5 ml ice cold IP dilution buffer (0.01% SDS, 1% Triton X-100, 1.2 mM EDTA pH 8.0, 16.7 mM Tris.HCl, pH 8.0, 167 mM NaCl and protease inhibitors) and 0.3 g of acid washed 212−300 micron glass beads (Sigma) was added. The mixture was sonicated in ice water 8 times for 30 s with 30 s intervals, transferred to microfuge tubes and centrifuged at 18407 rcf for 10 min at 4°C.

Quality control of input chromatin

100 µl of chromatin was adjusted to 200 µl with IP dilution buffer and decrosslinked at 65^{o}C overnight. 2µl of proteinase K at 20 mg/ml (Invitrogen) was added, and the mixture was incubated at 55^{o}C for 2 h. Chromatin was extracted twice with equal volume of phenol:chloroform (Sigma) and once with chloroform (Sigma). 2 µl of glycogen and one-tenth volume of 3M NaOAc, pH5.2 and 2.5 volumes of ethanol were added. The mixture was incubated for 30 min at -20°C, centrifuged and washed twice with 70% ethanol. The pellet was dissolved in 50 µl water. Chromatin size was checked on a 1% agarose gel. DNA was quantified using Nanodrop ND1000 spectrophotometer.

Preparation of Protein A beads

rProtein A agarose beads (GE Healthcare) were washed with IP dilution buffer 3 times and blocked in IP dilution buffer containing 1% BSA rotating at 4°C overnight. Beads were washed three times in IP dilution buffer and resuspended in IP dilution buffer for later usage.

Chromatin preclear

Chromatin was diluted three to five times with IP dilution buffer (depending on the DNA concentration). 30 µl of washed Protein A agarose beads were added to each fraction of diluted chromatin to be used in later immunoprecipitation and incubated at 4°C for 2 h.

Immunoprecipitation

Same volume of precleared chromatin was aliqouted for each IP sample. 3-5 µl of antibodies were added to aliquoted chromatin and incubated rotating at 4°C overnight. 30 µl of washed beads previously blocked with IP dilution buffer containing 1% BSA were added to each IP sample. The mixture was incubated at 4°C for 4 h. Chromatin-bound beads were centrifuged at 587 rcf for 2 min, and supernatant was discarded. Beads were washed three times with 1 mL Low salt wash buffer (0.1% SDS, 1.0% Triton X-100, 2.0 mM EDTA pH 8.0, 20 mM Tris.HCl, pH 8.0, 150 mM NaCl), three times with High salt wash buffer (0.1% SDS, 1.0% Triton X-100, 2.0 mM EDTA, pH 8.0, 20 mM Tris.HCl, pH 8.0, 500 mM NaCl), and two times with LiCl buffer (0.25 M LiCl, 1 mM EDTA, pH 8.0, 10 mM Tris HCl, pH 8.0, 1% NP40, 1% SDC) for 5 min with rotation at RT and pelleted at 587 rcf for 2 min. Chromatin was eluted with 200 µl of freshly prepared IP Elution Buffer (0.1 M $NaHCO_3$, 1% SDS) at 65°C for 30 min. This step was repeated one more time. NaCl (5 M, 20µl), EDTA (0.5 M, 8µl), Tris (1 M, pH 8.0, 16µl) were added to 400 µl of eluted chromatin and incubated overnight at 65^{o}C. 4 µl of Proteinase K at 20

mg/ml were added for 3 h at 55°C. Chromatin was extracted with phenol:chloroform and precipitated as described before. IP samples were dissolved in 50 μl water.

ChIP quantification

The quantities of target genomic regions precipitated by different antibodies were calculated as percent input based on four-point standard curves constructed from input DNA for each primer set. Standard deviation of each PCR performed in quadruplicate was calculated to determine the error of measurement. Two independent ChIP samples were analyzed, and similar results were obtained. ChIP primers were designed to be unique, detecting only sequences present in the *flam* and *80EF* piRNA loci and verified by *in silico* PCR. All primers (Table 2-2 and 2-3) were checked for both specificity and efficiency by standard agarose gel electrophoresis and real time PCR respectively. Primers to piRNA clusters amplify in the same DNA dilution range as primers specific to *hsp26* and *yellow* single copy genes compared to high copy TE elements (Figure 2-13).

Culture of OSC cell line and siRNA knockdowns

The OSC line was maintained and Piwi siRNA knockdown was performed as previously described (Saito et al. 2009). Briefly, 3×10^6 trypsinized cells were resuspended in 0.1 mL of Solution V of the Cell Line Nucleofector Kit V (Amaxa Biosystems) and mixed with 200 pmol of siRNA duplex. Transfection was conducted according to the manufacturer's protocol using the nucleofector program T-029, and the transfected cells were incubated at 25°C for 48 hrs. Protein knockdowns were verified by

Western blotting, and ChIP assays were performed on mock and *piwi* siRNA transfected cells ($5x10^6$ cells per IP).

Immunostaining of polytene chromosomes

Preparation and immunostaining of salivary gland polytene chromosomes was performed as described previously (Gerasimova et al. 2000). Primary antibodies directed against HP1 (Covance) and Mod(mdg4)2.2 (generated similarly as in Mongelard et al. 2002) and Alexa Fluor 488 labeled anti-guinea pig or Alexa Fluor 594 labeled anti-rabbit secondary antibodies (Invitrogen-Molecular Probes) were used. The chromosomes were viewed using a Leica epifluorescence microscope and photographed using a Hamamatsu digital camera.

DX1 variegation assay

Eye pigmentation of 100 to 200 flies was scored. The scoring of variegation was categorized into five groups: Light, Medium-Light, Medium, Medium-Dark and Dark corresponding to the percentage of pigmented facets. Percentage of flies falling into each category was graphed. Representative eye photos were taken.

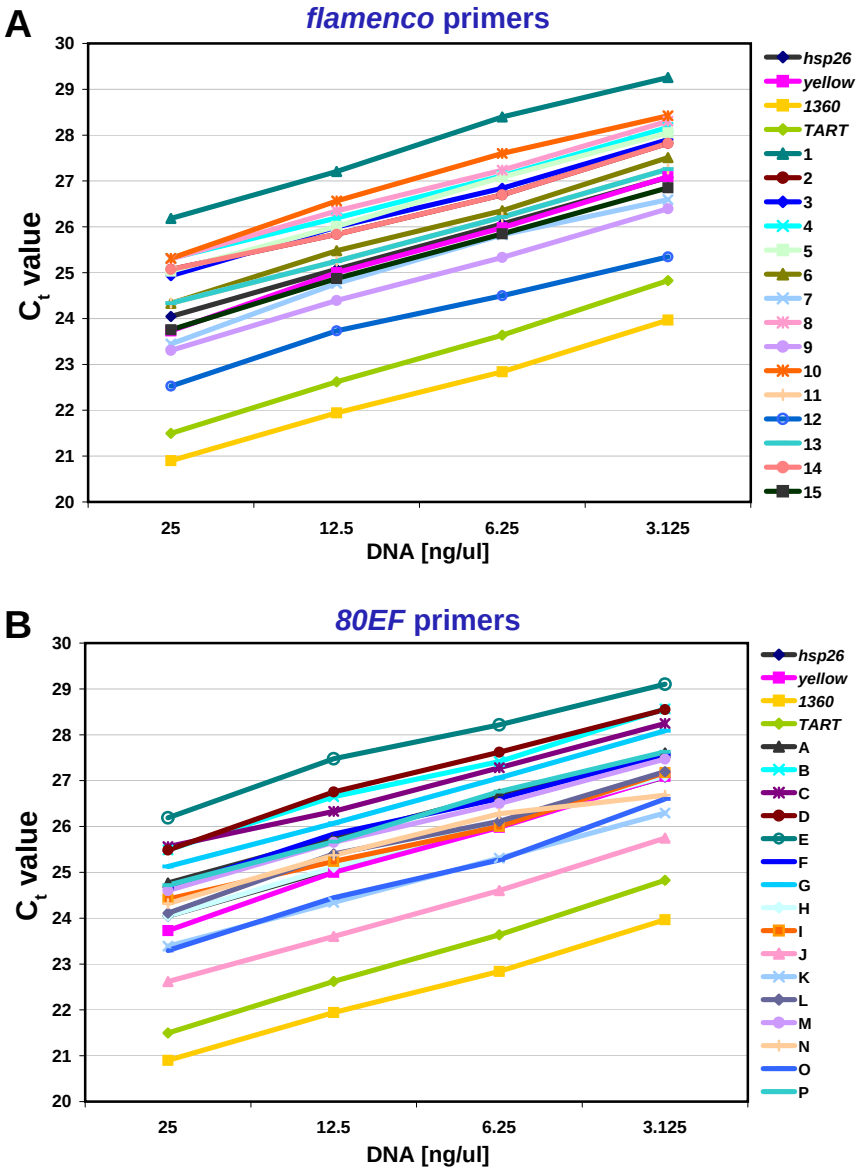


Figure 2-13. ChIP primer efficiency and specificity.

PCR amplification efficiency and specificity of ChIP primers at (A) *flam* and (B) *80EF* piRNA loci are graphed as a function of cycle threshold (Ct) values over DNA concentration. Ct values of standard curves of input from a representative experiment were graphed to show that primers to piRNA clusters amplify in the same DNA dilution range as primers specific to single copy genes *hsp26* and *yellow* compared to high copy TE elements.

Table 2-2. Primer set sequences used for ChIP at the *flam* piRNA cluster.

	Sequence 5' to 3'	Genomic coordinates
1	cgttcatgtcgttccacaac tgcacggatcgtggttatta	chrX:21473147+21473333 187bp
2	aaaccacttcgcggatttc tgcattttgatttcttgtgctc	chrX:21495641+21495810 170bp
3	aacgaggccagattcaacat gaatcagtacgagggcaagg	chrX:21497545+21497751 207bp
4	caagttggggtttcgtgttt attgaaccttaccccgacaa	chrX:21504581+21504730 150bp
5	ggagtgggatggatagacga cctggacacaggaccaaagt	chrX:21510544+21510729 186bp
6	ctcgggattttgcgttacat ggcagctaaccgtggataaa	chrX:21526571+21527094 231bp
7	gtggcttcacaaaacacgac cgaaggcttacacgcaagat	chrX:21527157+21527376 220bp
8	cctaccaacccagcgaataa tgctcttaagcctgcgaaat	chrX:21537802+21538037 236bp
9	cgatccgtttatgcaggtct ctgccaacaaatccatttcc	chrX:21539214+21539437 224bp
10	tgcctgtcgtactttgcttg ccaatgaattgccgctagtt	chrX:21543255+21543448 194bp
11	cgcgactgattggaagaact tctaagcccaacgtacacga	chrX:21586797+21586976 180bp
12	tcaggattcctccagaggtg ggccgctatgagtttcatgt	chrX:21604099+21604347 249bp
13	tgcgtgacgtaagcaaactc ttttatcggtggtgggaaag	chrX:21605733+21605922 190bp
14	cgggtgtaggtcacttggtt cagttaccaacgcaatcacg	chrX:21611603+21611784 182bp
15	tgcgtgccttttaaggagtc cgctgaatgcgatagtgaca	chrX:21618001+21618210 210bp

Table 2-3. Primer set sequences used for ChIP at the *80EF* piRNA cluster.

	Sequence 5' to 3'	Genomic coordinates
A	aggcacatggatgaacaaca gtttggttaacgggcaacat	chr3L:23255817+23255966 150bp
B	accgtgcatcccaatatcat ccaccaaaagaaagaacacg	chr3L:23257001+23257188 188bp
C	aggacacacatgcttgctttt cgataaatcttcttttggcaga	chr3L:23262518+23262618 101bp
D	tagcattacggcgaatggac ctctgcaataaagcgcacac	chr3L:23271313+23271532 220bp
E	gcttcgaagaagtgcaatca ttttgagcgggttttattcg	chr3L:23277432+23277632 201bp
F	ggacggtttgtttgtcttcg gactcgatgtggccatgata	chr3L:23278084+23278273 190bp
G	ttttgcatgtggcaataatca cgcatcggatattgtctgtg	chr3L:23281146+23281329 184bp
H	cgaggcatgtcgtagctgta gccctagtggcctcttctct	chr3L:23290484+23290709 226bp
I	cctcattttcgcctcgatta aaaagaaccgcaagagagca	chr3L:23291884+23292130 247bp
J	tcgatgagcaagatgtgagg aaacgagatggccaacaaag	chr3L:23295139+23295322 184bp
K	agggtccggttctcttctgt aaaacttggttgccctgatg	chr3L:23300821+23301000 180bp
L	tcgtggtgcagttgagagtc aagagcggcagagagtcaag	chr3L:23307902+23308093 192bp
M	aaatcaaacggagtttctgtttct caagctcaaagtgccatcaa	chr3L:23308478+23308657 180bp
N	tttcggaagctggtacaaag cgccgcttatattttgaacg	chr3L:23312351+23312521 171bp
O	ctagtttttcagcgtgcttgg ctaagaaggcaattgcgaaag	chr3L:23322270+23322429 160bp
P	ggagctattggagccgtcta tgtactcttgccatggttcg	chr3L:23332664+23332763 100bp

CHAPTER 3

RNAi-INDEPENDENT ROLE FOR ARGONAUTE2 IN CTCF/CP190 CHROMATIN INSULATOR FUNCTION

ABSTRACT

A major role of the RNAi pathway in *S. pombe* is to nucleate heterochromatin, but it remains unclear whether this mechanism is conserved. To address this question in *Drosophila*, genome-wide localization of Argonaute2 (AGO2) by ChIP-seq in two different embryonic cell lines was performed revealing that AGO2 localizes to euchromatin but not heterochromatin. This localization pattern is further supported by immunofluorescence staining of polytene chromosomes and cell lines, and these studies also indicate that a substantial fraction of AGO2 resides in the nucleus. Intriguingly, AGO2 colocalizes extensively with CTCF/CP190 chromatin insulators but not with genomic regions corresponding to endogenous siRNA production. Moreover, AGO2, but not its catalytic activity or Dicer-2, is required for CTCF/CP190-dependent *Fab-8* insulator function. AGO2 interacts physically with CTCF and CP190, and depletion of either CTCF or CP190 results in genome-wide loss of AGO2 chromatin association. Finally, mutation of *CTCF*, *CP190*, or *AGO2* leads to reduction of chromosomal looping interactions, thereby altering gene expression. I propose that RNAi-independent recruitment of AGO2 to chromatin by insulator proteins promotes the definition of transcriptional domains throughout the genome.

INTRODUCTION

RNA silencing is an evolutionary conserved mechanism that involves small RNAs bound to an Argonaute (AGO) protein that act as transcriptional or post-transcriptional regulators of gene expression. The paradigm for how RNA silencing controls gene expression at the chromatin level comes from studies in fission yeast, in which the RNA interference (RNAi) machinery establishes heterochromatin at the centromere and mating type locus to ensure proper chromosome segregation and to promote stability of repetitive regions. At the centromere, RNAs transcribed from pericentromeric repeats are processed by the Dcr1 endonuclease and Ago1 Argonaute protein, which leads to the recruitment of the histone H3K9 methyltransferase and Swi6/HP1 binding (reviewed in (Grewal and Elgin 2007)).

In *Drosophila*, it remains unclear whether the RNAi pathway is involved directly in heterochromatin formation. The primary endogenous function of the RNAi/siRNA pathway is to silence the expression of transposable elements (TEs) in the soma (reviewed in (Okamura and Lai 2008)). Silencing is achieved by Dcr-2-mediated cleavage of double-stranded RNAs (dsRNAs) into 21-22 nt siRNA that are loaded into AGO2, which cleaves the target TE mRNA using its Slicer activity. Less well understood is the function of non-TE endo-siRNAs also produced by Dcr-2 activity and loaded into AGO2, which are generated from hairpin transcripts and regions of 3' overlap of convergent transcripts (3' *cis*-NATs). Two studies implicated AGO2 in heterochromatin formation based on mislocalization of HP1 and desilencing of pericentromeric transcriptional reporters in *AGO2* mutants (Deshpande et al. 2005;

Fagegaltier et al. 2009). However, direct analysis of HP1 recruitment by chromatin immunoprecipitation (ChIP) and HP1-dependent silencing at small RNA generating loci led to the suggestion that *AGO2* and other Argonaute genes may not be required for heterochromatin formation in the soma (Moshkovich and Lei 2010).

Chromatin insulators are DNA-protein complexes defined functionally either as barriers that prevent the spread of silent chromatin or enhancer blockers that constrain enhancer-promoter communication. Unlike vertebrates, which possess only one known insulator protein, CTCF (reviewed in (Phillips and Corces 2009)), *Drosophila* employs at least five different insulator complexes. Two well-characterized insulators are the *gypsy* (also known as Su(Hw)) insulator and the *Fab-8* insulator of the *Abd-B* locus in the bithorax complex (BX-C) (reviewed in (Bushey et al. 2008)). The *gypsy* and *Fab-8* insulators harbor binding sites for the zinc-finger DNA-binding proteins Su(Hw) and CTCF, respectively, and both insulator complexes share a common component, CP190. Genome-wide insulator proteins are present at thousands of distinct DNA-binding sites but in diploid cells they concentrate at a small number of nuclear foci termed insulator bodies, which are dependent on CP190 for their integrity. Highly correlated at least with *gypsy* insulator function, insulator bodies have been proposed to serve as tethering sites for large chromosomal loops or other higher order chromatin structures.

It has become increasingly apparent that DNA topology is a critical determinant of gene regulation. While enhancers activate their target promoters over long distances, insulators act to restrict these communications (reviewed in (Wallace and Felsenfeld 2007)). Insulators and other *cis*-regulatory regions in the *Abd-B* locus engage in numerous interactions, and the precise topology of the locus has been postulated to be a

central mechanism of tissue-specific *Abd-B* regulation (Cleard et al. 2006; Lanzuolo et al. 2007; Kyrchanova et al. 2008; Bantignies et al. 2011). However, the mechanism by which chromosome looping is achieved at this locus has not been elucidated. Vertebrate CTCF has been shown to mediate chromosomal looping at several developmentally regulated loci jointly with cohesin (reviewed in (Merkenschlager 2010)), and a recent study reported that CTCF promotes promoter-enhancer interactions genome-wide. However, it is unknown whether *Drosophila* CTCF can promote looping.

In *Drosophila*, AGO2 or other RNA silencing factors appear to play important roles in chromatin and nuclear organization, such as formation of Polycomb Group (PcG) repression bodies (Grimaud et al. 2006) and *gypsy* chromatin insulator bodies (Lei and Corces 2006). Furthermore, AGO proteins have been detected in the nucleus and are thought to have a functional role in that compartment. Overall, these studies suggest novel mechanisms by which RNA silencing affects gene expression on the level of higher order chromatin organization.

Here, I hypothesized that similarly to the *gypsy* insulator, RNA silencing can affect the *Fab-8* insulator. A comprehensive genetic analysis of diverse RNA silencing mutants on *Fab-8* insulator function revealed that AGO2 was the only RNA silencing component to exert an effect. In order to test whether AGO2 may have a function on chromatin, ChIP-seq analysis of AGO2 in two *Drosophila* cell lines was performed. Instead of repetitive sequence, AGO2 associates primarily with euchromatic sites, the majority of which correspond to chromatin insulators. Intriguingly, AGO2 chromatin association does not correspond to regions of the genome that produce endo-siRNAs. I demonstrate that AGO2, but not its catalytic activity, is required for CTCF/CP190-

dependent *Fab-8* insulator function. Additionally, AGO2 interacts physically with CP190 and CTCF. Chromosome conformation capture (3C) experiments demonstrate that CTCF/CP190-dependent looping interactions may regulate AGO2 recruitment to chromatin. Also, depletion of AGO2 leads to a decrease in chromosomal looping and, thus, altered gene expression. Therefore, I propose a novel RNAi-independent role for AGO2 on chromatin to promote or stabilize insulator-dependent looping interactions to define transcriptional domains throughout the genome.

RESULTS

AGO2 associates with euchromatin and not repetitive sequences

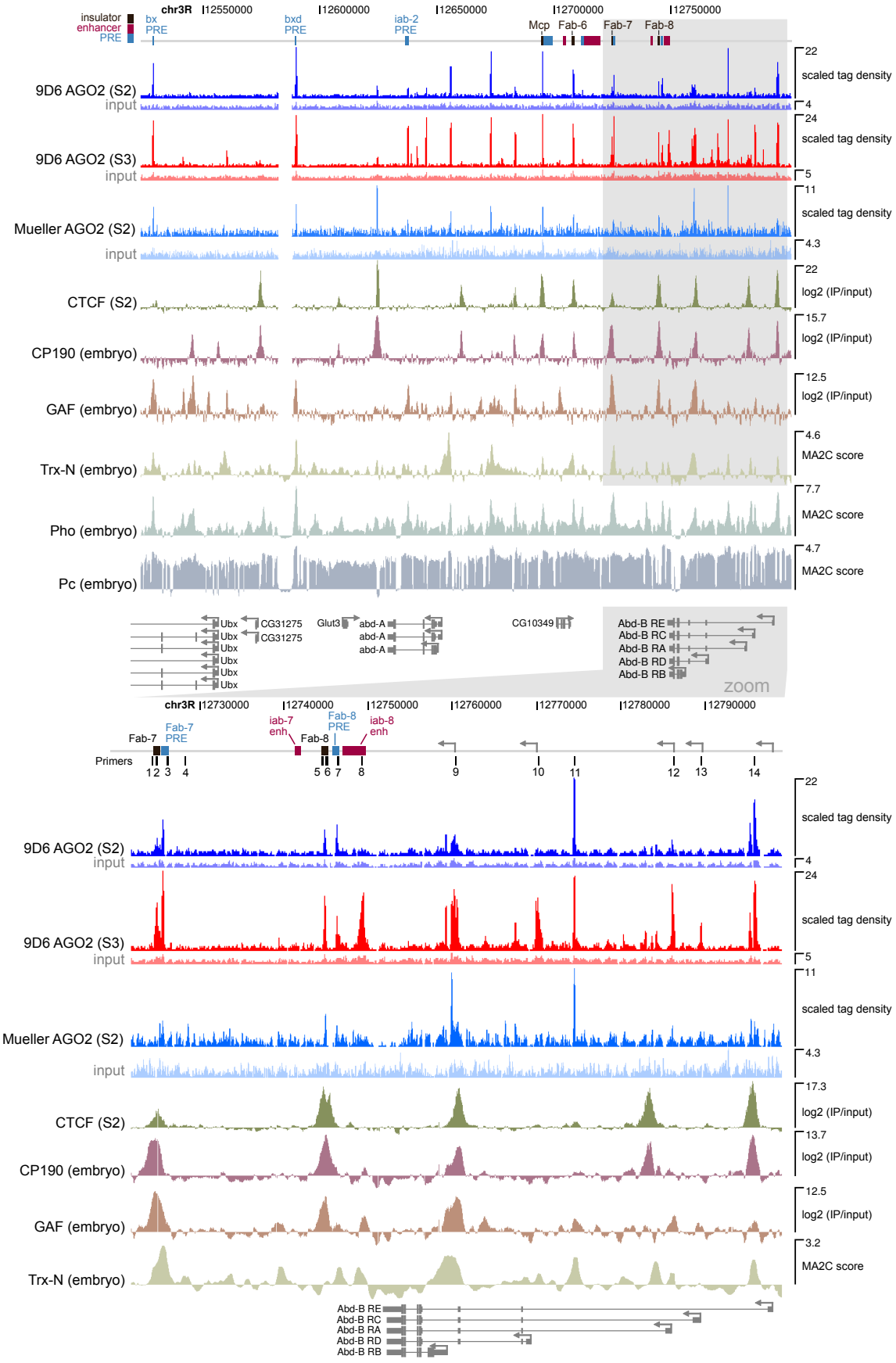
In order to obtain high-resolution information about the genome-wide chromatin association profile of AGO2, ChIP-seq analysis of AGO2 in S2 and S3 *Drosophila* embryonic cell lines was performed. ChIP was carried out using a previously characterized monoclonal antibody, 9D6, capable of isolating AGO2 and associated small RNAs (Kawamura et al. 2008). Greater than 9M reads per input or IP sample were obtained, leading to the identification of 3367 AGO2 bound sites between both cell types using a 5% false discovery rate threshold with the MACS algorithm (Zhang et al. 2008). Approximately 86% of AGO2 sites in S2 overlap with those found in S3 suggesting that AGO2 genome-wide localization is mainly consistent between cell types. Comparing the fraction of total reads mapping to repetitive sequences indicates no enrichment of repetitive sequences in the IP versus input (chi-square test, $p < 2e^{-16}$); therefore, I conclude that AGO2 localizes predominantly to euchromatic regions.

Strikingly, the majority of AGO2 sites overlap with known chromatin insulator sites throughout the genome. As a model region, I inspected the 300 kb BX-C Hox gene cluster and observed association of AGO2 with all known *cis*-regulatory domain boundaries in both cell lines (Figure 3-1). These insulators include the *Abd-B* locus boundary elements *Mcp*, *Fab-6*, *Fab-7*, and *Fab-8*. We obtained a similar ChIP-seq profile with lower signal using an independent α-AGO2 polyclonal antibody (Meyer et al. 2006). Moreover, three independent antibodies capable of immunoprecipitating

AGO2 (Jiang et al. 2005; Meyer et al. 2006; Czech et al. 2008) show similar enrichment profiles at the *Abd-B* locus as determined by ChIP followed by quantitative PCR (data not shown; Moshkovich et al. 2011). For subsequent genome-wide binding site analyses, 9D6 data was used exclusively because of its high signal-to-noise ratio and well-characterized specificity (Kawamura et al. 2008) (Figure 3-1, data not shown).

Figure 3-1. **ChIP-seq profiles of AGO2 in S2 and S3 cells at BX-C.**

AGO2 ChIP-seq profiles of input DNA and IPs in S2 and S3 cells compared with tiling array ChIP data for CTCF, CP190, GAF (Negre et al. 2010), Trx-N, Pho, and Pc (Schuettengruber et al. 2009) in indicated cell types or embryos over the BX-C region (top) and *Abd-B* locus (bottom). Coding sequences, promoters, and *cis*-regulatory regions are shown. ChIP-seq scales are in reads per million unique mapped reads. Input samples are shown on the same scale relative to respective IP and are therefore directly comparable. ChIP-chip data are expressed either as $\log_2$ (IP/input) or as MA2C score. The bottom of each scale bar indicates zero. AGO2 ChIP-seq analysis was performed by Ryan K. Dale.


insulator
enhancer
PRE
chr3R
12550000
12600000
12650000
12700000
12750000
bx PRE
bxd PRE
iab-2 PRE
Mcp
Fab-6
Fab-7
Fab-8
9D6 AGO2 (S2)
input
9D6 AGO2 (S3)
input
Mueller AGO2 (S2)
input
CTCF (S2)
CP190 (embryo)
GAF (embryo)
Trx-N (embryo)
Pho (embryo)
Pc (embryo)
22
scaled tag density
4
24
scaled tag density
5
11
scaled tag density
4.3
22
log2 (IP/input)
15.7
log2 (IP/input)
12.5
log2 (IP/input)
4.6
MA2C score
7.7
MA2C score
4.7
MA2C score
Ubx
CG31275
Glut3
abd-A
CG10349
Abd-B RE
Abd-B RC
Abd-B RA
Abd-B RD
Abd-B RB
zoom
chr3R
12730000
12740000
12750000
12760000
12770000
12780000
12790000
Fab-7
Fab-7 PRE
iab-7 enh
Fab-8
Fab-8 PRE
iab-8 enh
Primers
1 2 3 4
5 6 7 8
9
10
11
12
13
14
9D6 AGO2 (S2)
input
9D6 AGO2 (S3)
input
Mueller AGO2 (S2)
input
CTCF (S2)
CP190 (embryo)
GAF (embryo)
Trx-N (embryo)
22
scaled tag density
4
24
scaled tag density
5
11
scaled tag density
4.3
17.3
log2 (IP/input)
13.7
log2 (IP/input)
12.5
log2 (IP/input)
3.2
MA2C score
Abd-B RE
Abd-B RC
Abd-B RA
Abd-B RD
Abd-B RB

AGO2 colocalizes with chromatin insulator sites throughout the genome

Consistent with binding at BX-C boundary sites, approximately 62% of AGO2 sites overlap with known chromatin insulator proteins. Comparison of AGO2 ChIP-seq profiles with previously determined genome-wide ChIP tiling array analyses indicates extensive overlap with the insulator proteins CP190, CTCF, BEAF-32, and modest similarity to Mod(mdg4)2.2 compared to random expectation (Figure 3-2 (A-B); Moshkovich et al. 2011). In contrast, AGO2 sites display no statistically significant overlap with the *gypsy* insulator protein Su(Hw), indicating specificity of the AGO2 correspondence with CTCF/CP190 insulators.

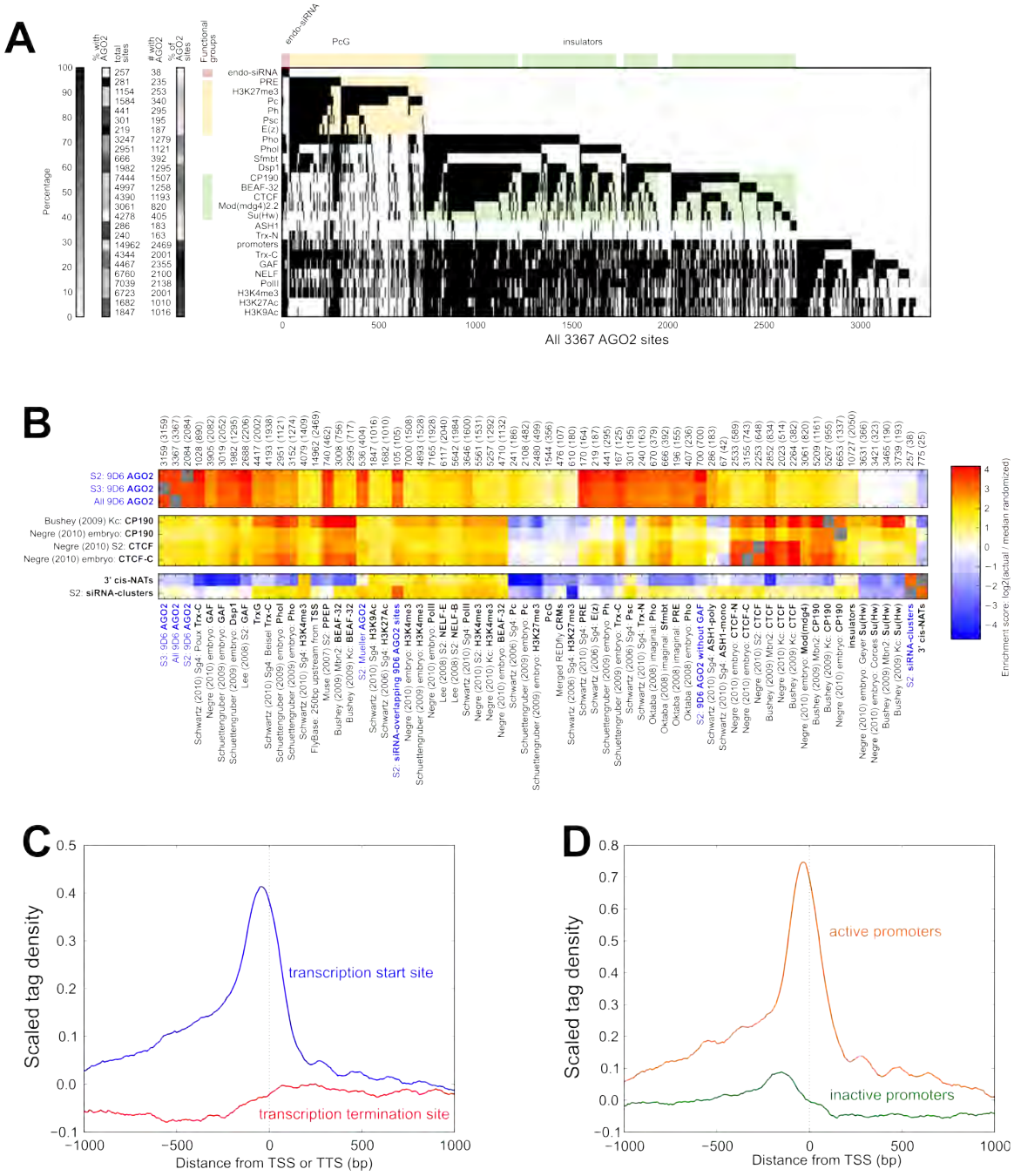
In order to confirm the genome-wide colocalization of AGO2 with insulator proteins and specific association with euchromatin, highly replicated polytene chromosomes of third instar larvae were stained by indirect immunofluorescence. AGO2 staining is mainly observed at euchromatic DAPI interbands, which correspond to decondensed regions of the genome bearing the majority of transcribed genes (data not shown; Moshkovich et al. 2011). In contrast, AGO2 is not visible at the heterochromatic chromocenter, at which the centromere of each chromosome coalesces. In $AGO2^{51B}$ null mutants (Xu et al. 2004), this staining pattern is dramatically reduced, verifying the specificity of the antibody (data not shown). In wild type, modest genome-wide colocalization is observed between AGO2 and CTCF while more extensive overlap is seen between AGO2 and CP190, consistent with the ChIP-seq results (data not shown).

Figure 3-2. Overlap between genome-wide binding sites of AGO2, insulator, TrxG/PcG, transcription related factors, and promoters.

(A) Binary heat map of AGO2 binding sites ordered by supervised hierarchical clustering. Each column represents one of the 3367 AGO2 binding sites across both S2 and S3 cell types, and each row represents overlapping binding sites for a particular factor across all available data sets. A mark in a row indicates that the indicated protein colocalizes with AGO2 at that site. AGO2 sites are classified into functional groups (endo-siRNA, PcG, and insulators). Feature counts for each factor and the number of features that intersect with the set of all AGO2 sites are shown. Corresponding percentages of overlap for each factor or for AGO2 are represented as grayscale values (left).

(B) Heat map of $\log_2$ enrichment scores for pairwise comparisons of binding sites for AGO2, CP190, CTCF, 3' *cis*-NATs, and endo-siRNA clusters with additional data sets. Enrichment score was calculated by dividing the actual overlapping feature count by the median overlapping feature count from 1000 random shufflings of features. Empirical p values reported in the text are the percentile of the actual overlapping feature count in this null distribution. Color scale corresponding to enrichment value is indicated (right). Positive values indicate significant enrichment while negative values indicate significant negative correlation of enrichment. Self-self comparisons are indicated in grey, and pairwise comparisons that are not statistically significant ($p > 0.001$) are indicated in white. Numbers along top of each column indicate the total number of features in each data set, and the number of sites that interact with all AGO2 sites are indicated in parentheses.

(C) Half of AGO2 binding sites correspond to promoters. Profile of S2 AGO2 ChIP-seq tag density subtracted by input density around transcription start sites (blue) or termination sites (red) from coding genes (FlyBase R5.23) generated using CEAS. (D) AGO2 associates preferentially with active promoters. Profile of S2 AGO2 ChIP-seq tag density subtracted by input density around TSSs associated (orange) or not associated (green) with H3K4me2 and Pol II 250 bp upstream or 750 bp downstream. AGO2 binding site comparisons with other datasets was performed by Ryan K. Dale.


A
B
C
D
Percentage
All 3367 AGO2 sites
Scaled tag density
transcription start site
transcription termination site
Distance from TSS or TTS (bp)
active promoters
inactive promoters
Distance from TSS (bp)

AGO2 associates with active promoters

Like insulator proteins, over half of AGO2 binding sites are located at promoters. Extensive promoter association has been reported for the insulator proteins CP190, CTCF, Mod(mdg4)2.2, and BEAF-32 but not Su(Hw) (Bushey et al. 2009; Jiang et al. 2009; Smith et al. 2009), with a preference for active promoters (Negre et al. 2010). Genome-wide, 61% of AGO2 sites in S2 cells are found within 250 bp upstream of a transcription start site (TSS), with a slight bias upstream of the TSS (Figure 3-2C). In contrast, no enrichment of binding is seen proximal to transcription termination sites (as defined by the presence of polyA sites). Next, AGO2 binding sites at TSSs associated with RNA Pol II and H3K4me3 in the body of the gene were compared to those lacking these active marks of transcription revealing that AGO2 associates preferentially with active promoters (Figure 3-2D), corresponding to approximately 14% of all active promoters. Consistent with this finding, AGO2 associates with all five active *Abd-B* promoters in S3 cells but with only the RB and RE (also known as m and γ respectively) inactive promoters in S2 cells (Figure 3-1). Moreover, AGO2 associates with the *iab-8* enhancer in S3 but not in S2 cells, suggesting that its chromatin association with certain enhancers and promoters may be dependent on active transcription.

AGO2 chromatin association does not correspond to regions of the genome that produce endo-siRNA

In contrast to its association with insulator proteins, AGO2 genome-wide localization does not coincide with regions of the genome that produce Dicer-dependent endo-siRNA. First, the genome-wide distribution of AGO2 binding sites was compared to a set of 257 clusters of high-density AGO2-bound unique endo-siRNAs in S2 cells (Czech et al. 2008; Ghildiyal et al. 2008; Kawamura et al. 2008; Okamura et al. 2008) demonstrating overlap with less than 1% of AGO2 sites, which is not statistically significant compared to random expectation (Figure 3-2 (A-B), p=0.35). As an additional test, the densities of unique endo-siRNA matching AGO2 chromatin binding sites in comparison to regions of the genome known to produce endo-siRNAs were calculated. Only thirty percent of 3' *cis*-natural antisense transcripts (*cis*-NATs) have been shown to produce Dicer-dependent endo-siRNAs (Okamura et al. 2008; Roy et al. 2010). The endo-siRNA densities of all known 3' *cis*-NATs were used to perform a conservative comparison to AGO2 chromatin-associated sites. Then, the endo-siRNA densities of sets of AGO2 binding regions and 3' *cis*-NATs shuffled throughout the genome in order to randomize their positions were calculated. Normalized to relative random expectations, substantially more 3' *cis*-NATs produce >8 endo-siRNAs per kb compared to AGO2 sites, which produce much lower levels of endo-siRNA (Figure 3-3 (A-B)). Production of such a low level of endo-siRNA at AGO2 sites may be due to the fact that AGO2 bound sites are associated with active transcription. These observations suggest that actively transcribed regions may produce more endo-siRNA than transcriptionally silent

regions. In fact, the top 5% highest endo-siRNA density AGO2 sites cluster with marks of active transcription (dat not shown, Moshkovich et al. 2011). Therefore, the endo-siRNA density analysis using Pol II and H3K27me3-bound regions, which represent transcriptionally active and inactive sites respectively, was repeated. Overall, Pol II-bound regions produce moderately higher levels of endo-siRNA than AGO2 sites while H3K27me3-bound regions produce lower levels of endo-siRNA compared to respective random expectations (Figure 3-3 (A-B)). Similar results were obtained using a nuclear library of S2 endo-siRNA (Figure 3-3 (C-D); Fagegaltier et al. 2009). These results suggest that regions of active transcription tend to produce low levels of endo-siRNA, and the majority of AGO2 binding sites correspond to little or no endo-siRNA production.

Figure 3-3. Distribution of endo-siRNA densitities for AGO2 binding sites compared to 3' *cis*-NATs and transcriptionally active or inactive regions.

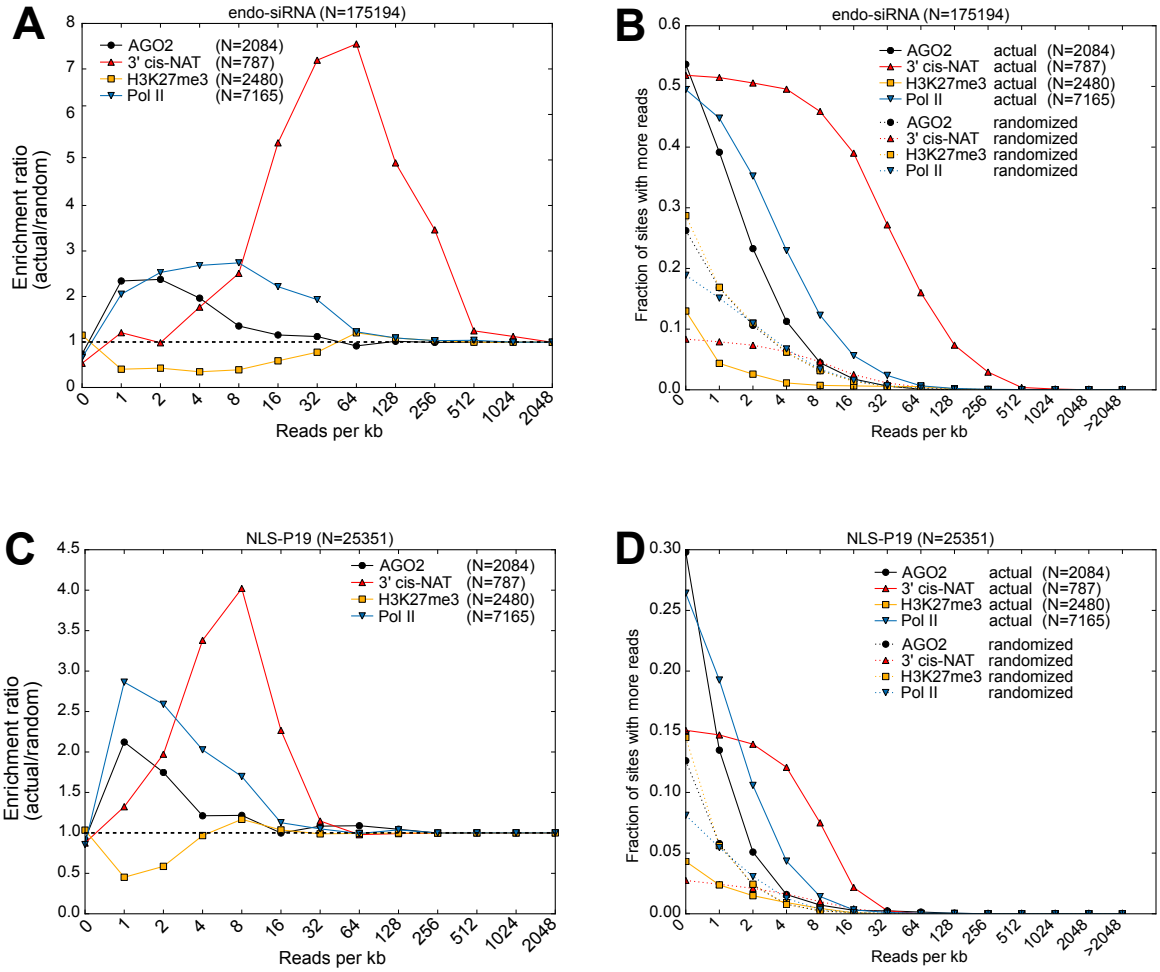
(A) Enrichment scores of endo-siRNA read densities in S2 AGO2 sites (black), 3' *cis*-NATs (Okamura et al. 2008) (red), H3K27me3 domains (Schuettengruber et al. 2009) (yellow), and S2 Pol II domains (Negre et al. 2010) (blue) expressed as a ratio of actual over random. For each bin, the enrichment score is calculated as (actual + 0.01) / (randomized + 0.01) to avoid dividing by zero. No enrichment with respect to random is indicated by a horizontal dashed line at 1.

(B) Complementary cumulative distributions of the density of endo-siRNA reads in S2 AGO2 sites (black), 3' *cis*-NATs (red), H3K27me3 domains (yellow), and S2 Pol II domains (blue). Solid lines show actual distribution while dotted lines show the average distribution over 1000 random intrachromosomal shufflings of binding sites.

(C) As in A, but using siRNAs bound to NLS-P19, a viral suppressor protein directed to the nucleus that acts upstream of AGO2 by binding double stranded siRNAs produced by Dcr-2 cleavage.

(D) As in (B) but using NLS-P19 bound siRNAs.

SiRNA analyses were performed by Ryan K. Dale.


A
endo-siRNA (N=175194)
AGO2 (N=2084)
3' cis-NAT (N=787)
H3K27me3 (N=2480)
Pol II (N=7165)
Enrichment ratio (actual/random)
Reads per kb
B
endo-siRNA (N=175194)
AGO2 actual (N=2084)
3' cis-NAT actual (N=787)
H3K27me3 actual (N=2480)
Pol II actual (N=7165)
AGO2 randomized
3' cis-NAT randomized
H3K27me3 randomized
Pol II randomized
Fraction of sites with more reads
Reads per kb
C
NLS-P19 (N=25351)
AGO2 (N=2084)
3' cis-NAT (N=787)
H3K27me3 (N=2480)
Pol II (N=7165)
Enrichment ratio (actual/random)
Reads per kb
D
NLS-P19 (N=25351)
AGO2 actual (N=2084)
3' cis-NAT actual (N=787)
H3K27me3 actual (N=2480)
Pol II actual (N=7165)
AGO2 randomized
3' cis-NAT randomized
H3K27me3 randomized
Pol II randomized
Fraction of sites with more reads
Reads per kb

AGO2 binds to PREs and overlaps extensively with TrxG and PcG proteins

Approximately 15% of AGO2 sites correspond to regions that can be regulated by both TrxG and Polycomb Group (PcG) proteins. The TrxG and PcG complexes maintain transcriptional activation or repression, respectively, of critical developmental regulators and are recruited by DNA-binding proteins that recognize Polycomb Response Elements (PREs), which are frequently juxtaposed to chromatin insulators (reviewed in Simon and Kingston 2009). This close configuration is particularly evident in the BX-C locus, in which insulators act as barriers to constrain PRE activity directionally. The high resolution afforded by ChIP-seq allows AGO2 detection specifically at all known PREs in the BX-C (*bx*, *bxd*, *iab-2*, *Fab-6*, *Fab-7*, and *Fab-8*) despite their close proximity to insulators in this locus (Figure 3-1, 3-2A). Additionally, AGO2 associates with 84% of PREs across the genome, as previously defined (Oktaba et al. 2008; Schwartz et al. 2010) (Figure 3-2A, left panel). It should be noted that the probability-based enrichment values calculated for the AGO2 overlap with PREs and associated factors are higher than that with insulator proteins; this result is influenced by the small number of sites bound by PcG proteins genome-wide compared to insulator proteins (Figure 3-2B). Finally, mild but statistically significant overlap is also detected between AGO2 and annotated *cis*-regulatory modules (CRMs) in the REDFly database (Gallo et al. 2011), which is biased towards extensively studied TrxG and PcG regulated genes.

AGO2 chromatin localization at PREs resembles that of TrxG proteins more closely than that of PcG proteins. Genome-wide, AGO2 overlaps extensively with the TrxG proteins Trx-N, Trx-C, and Ash1 as well as with the recruiter proteins Pho, Phol,

Sfmbt, Dsp1, and GAF, which also associate with non-PRE sites in the genome (Figure 3-2 (A-B)). Furthermore, AGO2 colocalizes substantially with the sharply peaking PRE-associated PcG proteins E(z), Ph, and Psc, as well as the broadly spreading Pc and H3K27me3; however, AGO2 itself does not bind chromatin in extended domains (Figure 3-1, 3-2A-B). Furthermore, AGO2 binds both *Fab-7* and *Fab-8* PREs in S2 cells, in which *Abd-B* is silent, as well as in S3 cells, in which *Abd-B* is expressed (Figure 3-1). Likewise, recruiter and TrxG proteins bind at *Abd-B* and its PREs irrespective of transcriptional expression state (Beisel et al. 2007), whereas PcG recruitment at *Abd-B* is only apparent in S2 cells (Breiling et al. 2004). This observation suggests that AGO2 does not require PcG proteins in order to associate with PREs.

In order to obtain further insight into the specificity of AGO2 chromatin association, *de novo* motif analysis of AGO2 binding sites was performed. The analysis of the central 500 bp of 500 random AGO2 binding sites using the MEME algorithm (Bailey and Elkan 1995) identified a GA-rich consensus binding sequence reminiscent of the binding motif for the TrxG and insulator-associated GAGA-factor (GAF) (Figure 3-4A) (Farkas et al. 1994; Belozerov et al. 2003; Schweinsberg et al. 2004). Similar results were obtained using all or non-GAF occupied AGO2 binding sites with the GADEM (Li 2009) and Weeder (Pavesi and Pesole 2006) algorithms (data not shown).

Figure 3-4. AGO2 behaves as a TrxG protein.

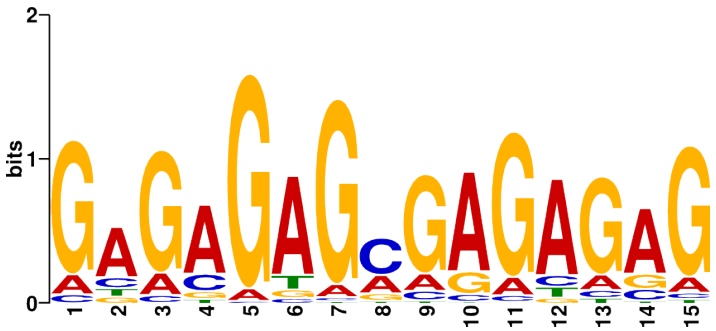
(A) Highest-scoring *de novo* AGO2 binding site motif found by MEME using the center 500 bp of a random subset of 500 AGO2 binding site sequences.

(B) Percentage of adult male flies displaying second and/or third legs with at least one ectopic sex comb tooth as an indication of posterior to anterior transformation was scored in the indicated genotypes, and number of flies (n) scored is shown.

(C) Western blotting of AGO2, Pc, and Mod(mdg4)2.2 in wildtype and $AGO2^{51B}$ adult male extracts.

De novo AGO2 motif analysis was performed by Ryan K. Dale.

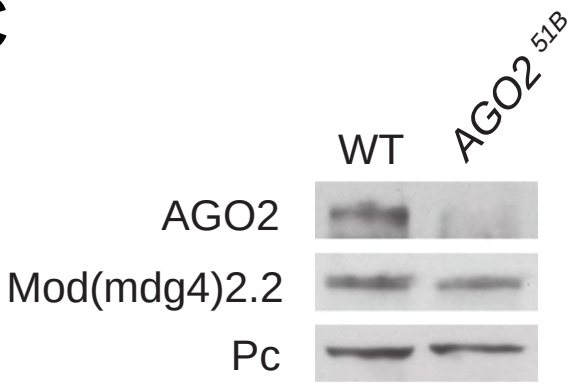
A



B

Genotype	n	% transformation
Pc^{4} / +	302	61.6
AGO2^{414} / +	357	0
AGO2^{51B} / +	337	0
+, *Pc*^{4} / *AGO2*^{414}, +	376	44.1
+, *Pc*^{4} / *AGO2*^{51B}, +	449	36.6
+, *Pc*^{4} / *AGO2*^{V966M}, +	278	61.9
AGO2^{V966M}, *Pc*^{4} / *AGO2*^{V966M}, +	339	62.0

C



AGO2 opposes Polycomb function

Given the high overlap of AGO2 with TrxG proteins, we tested whether AGO2 affects either TrxG or PcG function. I anticipated that *AGO2* may function as a *trxG* gene since the genes that encode GAF and Mod(mdg4)2.2 chromatin insulator proteins have been shown to behave as *trxG* genes (Farkas et al. 1994; Gerasimova and Corces 1998). I examined the classic posterior to anterior transformation phenotype of Pc^{4}/+ mutants and determined that 62% of adult males exhibit ectopic sex combs on second and/or third legs (Figure 3-4B). The $AGO2^{414}$/+ mutation results in a mild suppression of the Pc^{4}/+ phenotype such that a reduced number of double mutant males, 44%, display transformation. Interestingly, the partial loss-of-function $AGO2^{414}$/+ mutation is not defective for RNAi-dependent silencing in the heterozygous state (Okamura et al. 2004). Furthermore, heterozygous null $AGO2^{51B}$/+ mutants display stronger suppression of the Pc^{4}/+ phenotype in that only 37% of flies exhibit transformation. Neither $AGO2^{414}$/+ nor $AGO2^{51B}$/+ mutants, which both harbor deletions of the first two exons of *AGO2* (Okamura et al. 2004; Xu et al. 2004), exhibit developmental delays compared to wild type (data not shown). The $AGO2^{V966M}$ point mutation results in production of wildtype levels of catalytically inactive protein incompetent for RNAi-dependent silencing (Kim et al. 2007) but capable of associating with polytene chromosomes (data not shown). Importantly, the heterozygous $AGO2^{V966M}$/+ or homozygous $AGO2^{V966M}$ mutations do not affect the Pc^{4}/+ phenotype, indicating that Slicer catalytic activity of AGO2 is not required for the suppression of the Pc^{4}/+ phenotype. This suppression is not due to an indirect effect on *Pc* gene expression as Pc protein levels are equivalent in wild type and

AGO2^{51B} mutants (Figure 3-4C). These results indicate that *AGO2* behaves as a *trxG* gene and can counteract PcG function.

In order to determine whether PcG affects AGO2 recruitment on chromatin, I performed ChIP analysis of the *Abd-B* locus in S2 cells depleted of Pc. Transfection of dsRNA corresponding to *Pc* results in reduction of the target protein by over 90% (Figure 3-5A). I first analyzed Trithorax (TRX) and H3K27ac association with chromatin in mock-treated cells. In S2 cells, where *Abd-B* is PcG-repressed, TRX is present at S2 cells at baseline levels equivalent to *Rpl32* (Figure 3-5B). H3K27ac histone mark, which is associated with TrxG activity, is detected at IgG negative control levels. These observations are consistent with published data (Schwartz et al. 2010). AGO2 ChIP results in an approximate 12-fold enrichment of *Fab-7* (set 2), eight-fold enrichment of *Fab-8* PRE (set 7), four-fold enrichment of the *Abd-B* RE promoter (set 14) and six-fold enrichment of an intronic site (set 11) compared to *RpL32*, which shows low AGO2 association. In Pc-depleted cells, TRX and H3K27ac association with chromatin is increased. AGO2 recruitment to most sites is increased approximately 1.5 to three-fold across the entire *Abd-B* locus. Therefore, similarly to TrxG proteins, increased AGO2 association with chromatin is correlated with derepression of target genes caused by the reduction in Pc levels.

I also examined whether TrxG proteins affect AGO2 association with chromatin by performing AGO ChIP of the *Abd-B* locus in S3 cells depleted of TRX. TRX protein levels were reduced using dsRNA RNAi by over 90% (Figure 3-5C). In mock-treated S3 cells, where *Abd-B* is transcribed, ChIP of Pc and H3K27ac revealed baseline levels of Pc and strong H3K27ac signal over the entire locus (Figure 3-5D). AGO2 association with

chromatin is similar to S2 profile. TRX KD correlated with increased Pc association and loss H3K27ac over the *Abd-B* locus. Interestingly, no change in AGO2 association with chromatin was detected upon TRX depletion. Therefore, TrxG proteins do not affect AGO2 recruitment to chromatin. Taken together with AGO2 ChIP in Pc-depleted S2 cells, these results suggest that AGO2 may associate with open chromatin possibly to promote transcription.

Figure 3-5. Effects of *Pc* and *TRX* knockdown on AGO2 association with chromatin.

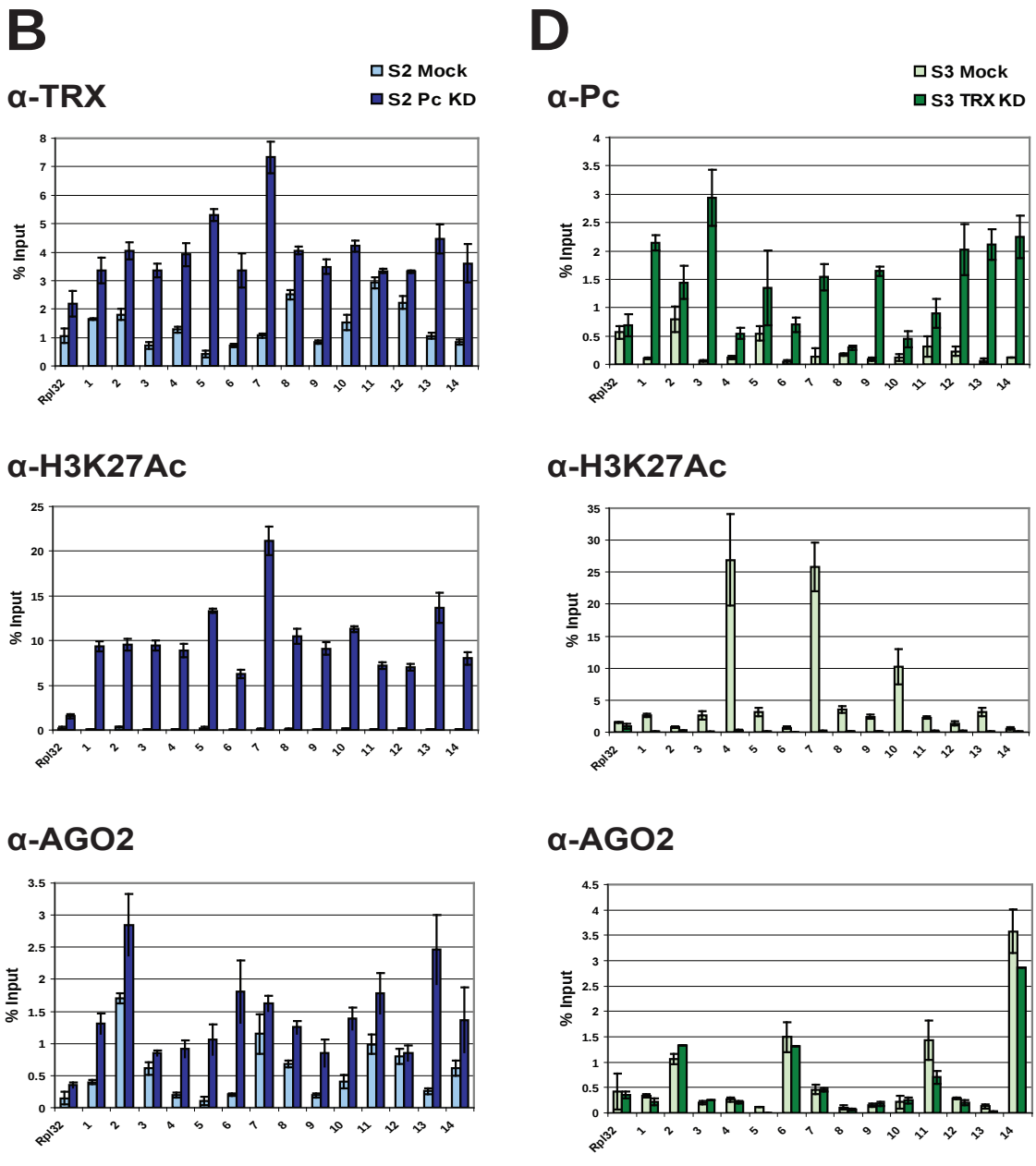
(A) Western blotting of lysates from S2 cells mock treated (lane 1) or transfected with *Pc* dsRNA (lane 2).

(B) S2 cells mock treated (light blue) or transfected with *Pc* dsRNA (dark blue) were subjected to ChIP using α-TRX-N, α-H3K27ac and α-AGO2 antibodies. Locations of primer sets are indicated in Figure 3-1. Percent input DNA immunoprecipitated is shown for each primer set, and error bars indicate standard deviation of quadruplicate PCR measurements. IgG negative control IPs for all sites yielded <0.06% input.

(C) Western blotting of lysates from S3 cells mock treated (lanes 1) or transfected with *TRX* dsRNA (lane 2).

(D) S3 cells mock treated (light green) or transfected with *Pc* dsRNA (dark green) were subjected to ChIP using α-Pc, α-H3K27ac and α-AGO2 antibodies. Percent input DNA immunoprecipitated is shown for each primer set, and error bars indicate standard deviation of quadruplicate PCR measurements. IgG negative control IPs for all sites yielded <0.05% input.


A
Mock S2
Pc KD
CP190
AGO2
Pc
1 2
C
Mock S3
TRX KD
TRX
AGO2
Lamin
1 2


B
S2 Mock
S2 Pc KD
α-TRX
% Input
Rpl32
α-H3K27Ac
α-AGO2
D
S3 Mock
S3 TRX KD
α-Pc
α-H3K27Ac
α-AGO2

AGO2 but not its catalytic activity is specifically required for *Fab-8* insulator activity

Given the high overlap of AGO2 with insulator sites throughout the genome, particularly of the CP190 class, I wished to determine whether *AGO2* is required for activity of the well-characterized CTCF/CP190 dependent insulator *Fab-8* of the *Abd-B* locus. I utilized a transgenic enhancer blocking assay in which a genomic fragment containing the *Fab-8* insulator and PRE positioned between a mini-*white* (mini-w^+) reporter and w^+ enhancer reduces reporter expression, resulting in intermediate levels of pigmentation in the adult eye (Barges et al. 2000). Compared to wild type, $AGO2^{414}$/+, $AGO2^{51B}$/+, $AGO2^{414}$, and $AGO2^{51B}$ mutants carrying the *Fab-8* insulator transgene display increased eye pigmentation corresponding to the strength of *AGO2* loss-of-function mutation, indicating a positive role for *AGO2* in *Fab-8* insulator function (Figure 3-6A). Importantly, the $AGO2^{V966M}$ catalytic activity mutant remains fully competent for *Fab-8* insulator activity. In comparison, loss-of-function $CP190^{4-1}$/$CP190^{H31-2}$ mutants (Pai et al. 2004) that reduce *Fab-8* insulator function (Gerasimova et al. 2007) display a more modest increase of mini-w^+ expression than $AGO2^{51B}$/+ mutants (data not shown). No differences compared to wild type are detected in *AGO2* mutant flies carrying a transgene containing only the *Fab-8* PRE or no *cis*-regulatory sequence (Figure 3-6A), indicating that the effects on the *Fab-8* insulator reporter are likely specific to the insulator. Importantly, comprehensive genetic analysis of RNA silencing mutants revealed that *AGO2*, but not other RNA silencing factors, is required for *Fab-8* insulator activity (data not shown, Moshkovich et al. 2011).

In order to obtain mechanistic insight into the possible function of AGO2 with respect to *Fab-8* insulator activity, the *in vivo* localization of insulator proteins in *AGO2* mutants was examined. Previously it was shown that positive or negative effects of certain RNA silencing mutants on *gypsy* insulator activity correlate with the integrity of insulator bodies (Lei and Corces 2006). The $AGO2^{51B}$ null mutation does not appear to reduce *Fab-8* function by disrupting the integrity of insulator bodies (Figure 3-6B). Furthermore, no overall differences in the ability of CTCF and CP190 to associate with chromatin or specifically with the BX-C on polytene chromosomes of wild type compared to $AGO2^{51B}$ mutants were observed (Figure 3-6C). Finally, Western blotting of wild type and $AGO2^{51B}$ mutants indicates no effect on CTCF or CP190 protein levels (Figure 3-6D).

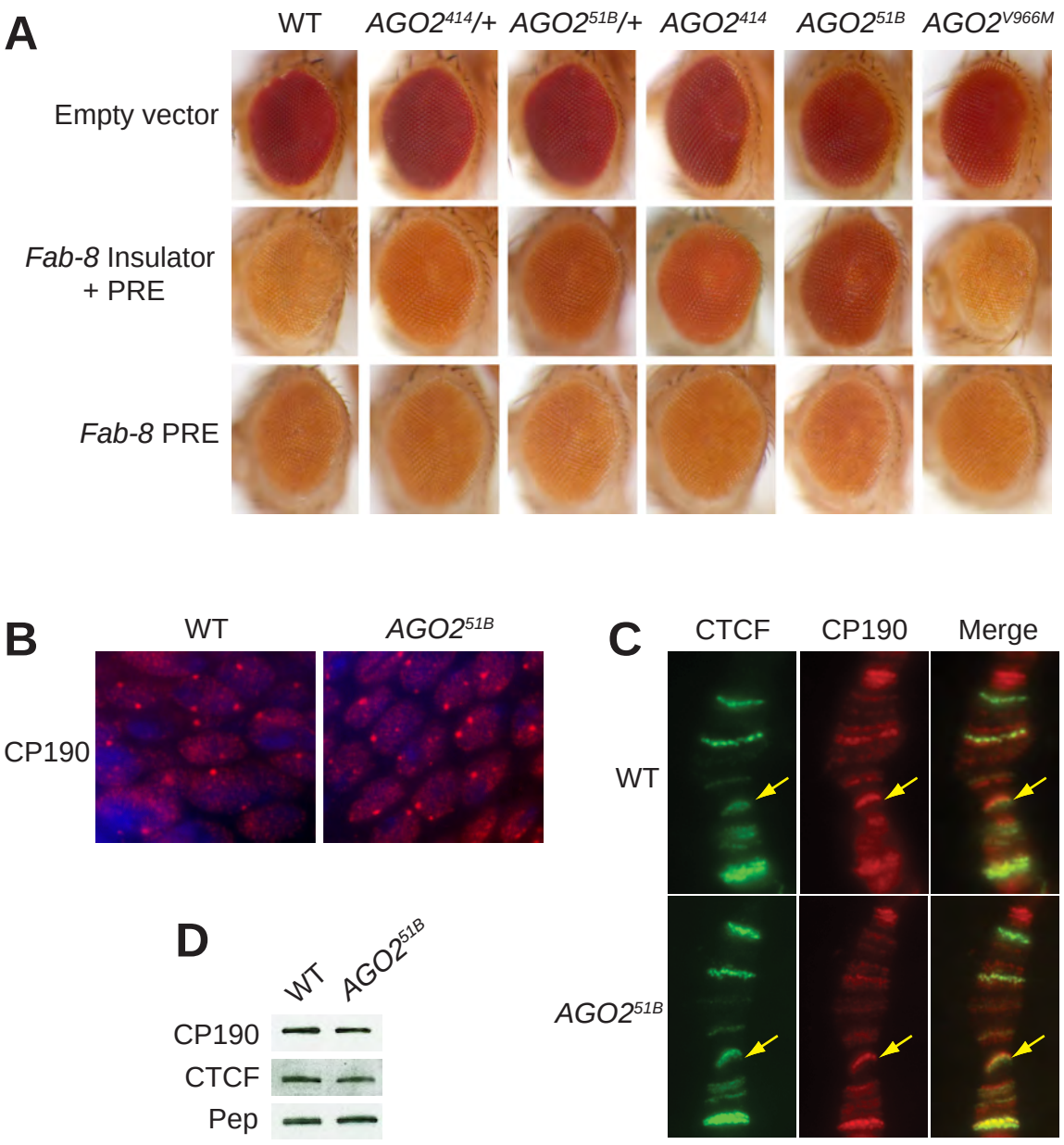
Figure 3-6. AGO2 but not its catalytic activity is required for *Fab-8* insulator function.

(A) Eye color due to expression of a transgenic construct carrying no regulatory element (top row), *Fab-8* insulator and PRE (middle row) or *Fab-8* PRE (bottom row) between the mini-*white* enhancer and its coding sequence in wildtype, $AGO2^{414}$/+, $AGO2^{51B}$/+ $AGO2^{414}$, $AGO2^{51B}$, and $AGO2^{V966M}$ flies.

(B) Visualization of insulator bodies by indirect immunofluorescence of whole mount larval imaginal discs using α-CP190 antibodies (red) merged with DAPI staining (blue) in wild type and $AGO2^{51B}$ mutants.

(C) Polytene chromosome staining of α-CTCF (green), α-CP190 (red), and merged images in wildtype and $AGO2^{51B}$ mutants. Arrows point to the BX-C locus.

(D) Western blotting of CP190, CTCF, and Pep (loading control) in wildtype and $AGO2^{51B}$ pupal extracts.


A
WT
AGO2414/+
AGO251B/+
AGO2414
AGO251B
AGO2V966M
Empty vector
Fab-8 Insulator + PRE
Fab-8 PRE
B
WT
AGO251B
CP190
C
CTCF
CP190
Merge
WT
AGO251B
D
WT
AGO251B
CP190
CTCF
Pep

AGO2 interacts physically with CTCF and CP190

In order to address whether AGO2 influences chromatin insulator activity in a direct manner, its subcellular localization compared to that of CP190 was examined. In S2 cells, CP190 localization is mainly diffuse within the nucleus whereas in S3 cells, CP190 is nuclear but also concentrates into insulator bodies reminiscent of those seen in larval imaginal disc cells (data not shown; Moshkovich et al. 2011). In both S2 and S3 cells, AGO2 localizes throughout the cell but concentrates preferentially in the nucleoplasm in the majority of cells. Nuclear signal is reduced upon siRNA knockdown of AGO2 (data not shown). Importantly, AGO2 staining is excluded from the heterochromatic DAPI dot and is mainly nonoverlapping with the heterochromatin protein HP1 (data not shown).

Next, physical interactions between insulator complexes and RNA silencing components were probed for. Immunoprecipitation of AGO2 from embryonic nuclear extracts at high monovalent salt concentrations results in copurification of CTCF and CP190 but not the *gypsy* insulator protein Mod(mdg4)2.2 (Figure 3-7A). In addition, column-based immunoaffinity purification of CP190-associated complexes from nuclear extracts verifies the presence of core *gypsy* and *Fab-8* insulator components CP190, Su(Hw), Mod(mdg4)2.2, and CTCF, and reveals association of the RNA silencing components Rm62, Piwi, and AGO2 (Figure 3-7B). Interactions between insulator proteins, Piwi, or AGO2 with CP190 complexes are not affected by RNaseA treatment under conditions that disassociate Rm62 (Figure 3-7C), suggesting that RNA does not mediate physical associations between Piwi or AGO2 and CP190. Physical interactions

between these RNA silencing components and CP190, either direct or in the context of larger complexes, are consistent with the direct involvement of Piwi and Rm62 in *gypsy* insulator activity and that of AGO2 in CTCF/CP190 insulator activity.

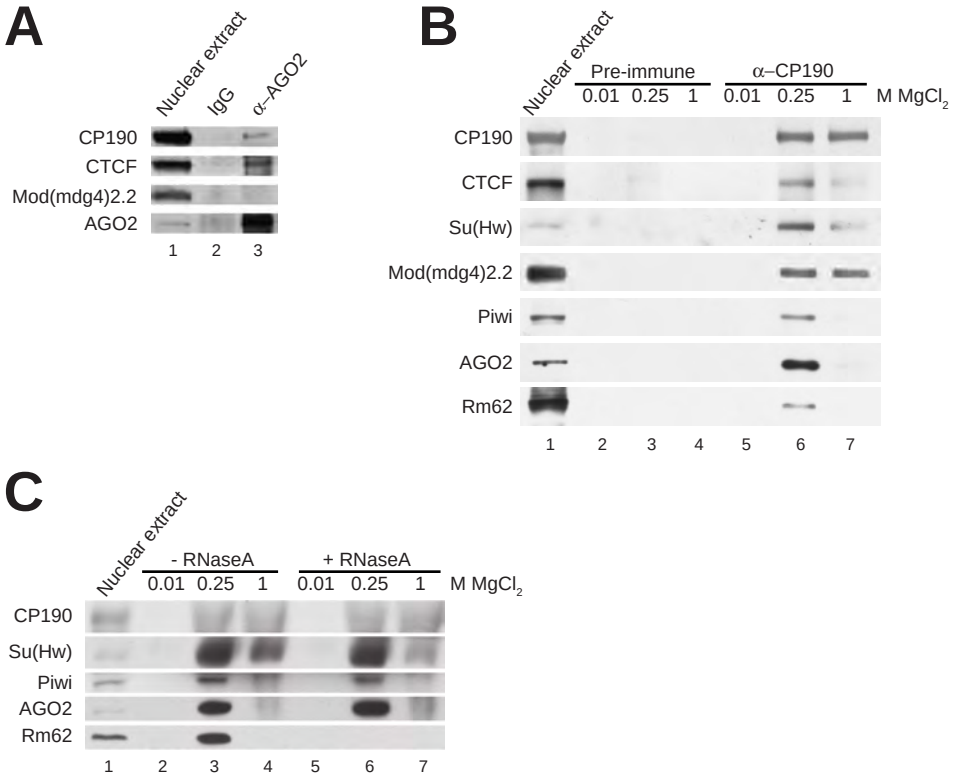
Figure 3-7. AGO2 associates physically with CP190 and CTCF.

(A) Western blotting of embryonic nuclear extracts immunoprecipitated with α-AGO2 antibodies. Nuclear extract (lane 1) bound to control IgG (lane 2) or α-AGO2 immobilized on ProtA-sepharose (lane 3) at > 1.1 M monovalent salt concentration.

(B) Western blotting of embryonic nuclear extracts (lane 1) bound to a control preimmune column (lanes 2-4) or α-CP190 column (lanes 5-7) and step eluted with increasing $MgCl_2$ concentrations as indicated.

(C) Western blotting of embryonic nuclear extracts (lane 1) bound to α-CP190 columns either untreated (lanes 2-4) or treated (lanes 5-7) with RNaseA and step eluted with increasing $MgCl_2$ concentrations as indicated.

Immunoaffinity purification with α-CP190, RNaseA treatment, and immunoprecipitation with α-AGO2 was performed by Elissa P. Lei.


A
Nuclear extract
IgG
α–AGO2
CP190
CTCF
Mod(mdg4)2.2
AGO2
1 2 3
B
Nuclear extract
Pre-immune
α–CP190
0.01 0.25 1 0.01 0.25 1 M MgCl2
CP190
CTCF
Su(Hw)
Mod(mdg4)2.2
Piwi
AGO2
Rm62
1 2 3 4 5 6 7
C
Nuclear extract
- RNaseA
+ RNaseA
0.01 0.25 1 0.01 0.25 1 M MgCl2
CP190
Su(Hw)
Piwi
AGO2
Rm62
1 2 3 4 5 6 7

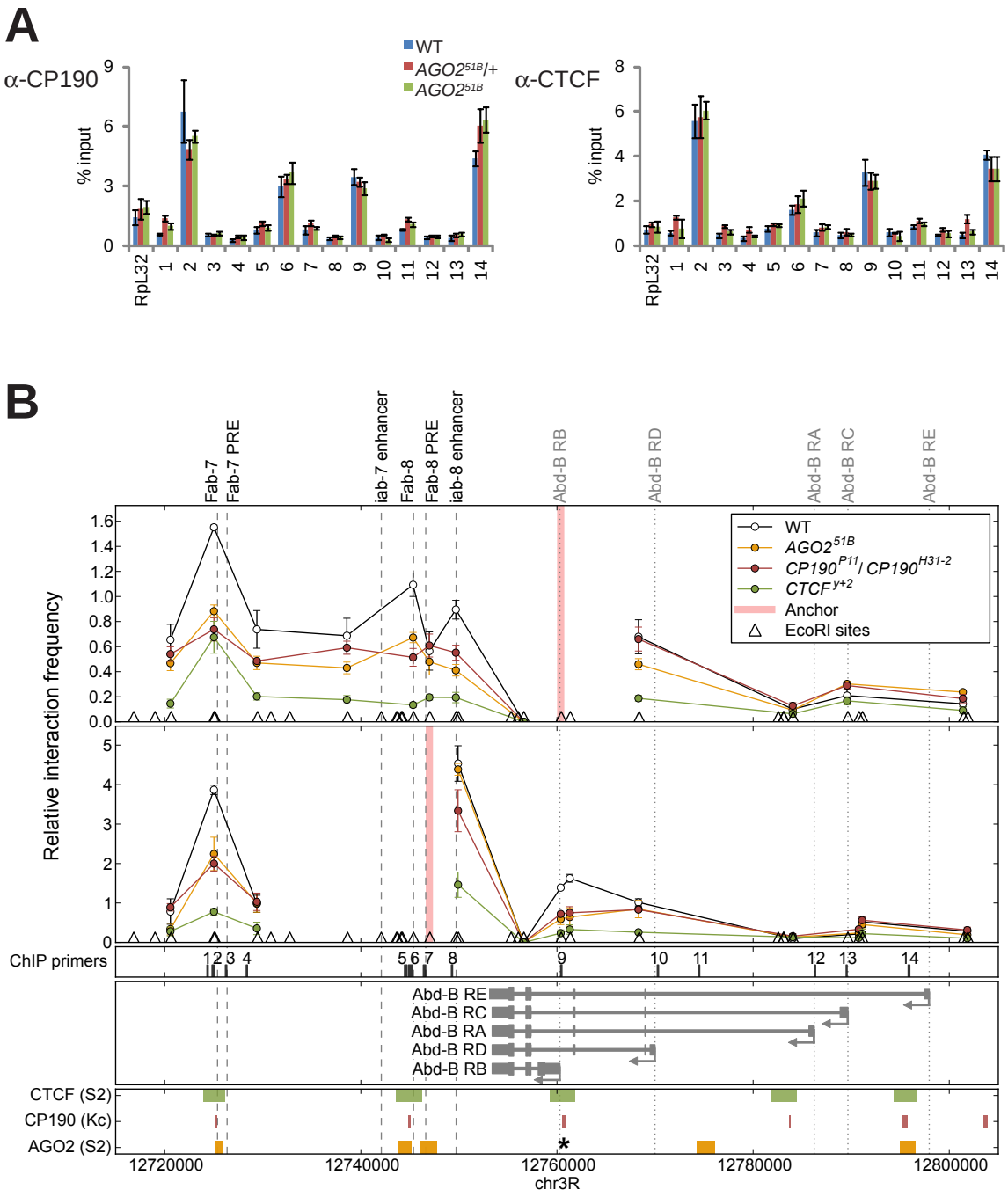
AGO2 associates with chromatin downstream of CTCF and CP190

In order to examine whether AGO2 recruitment to chromatin is downstream to that of CTCF and CP190, chromatin association of these insulator proteins in the absence of AGO2 was examined. No changes in CP190 or CTCF recruitment in AGO2 knockdowns were observed; however, a significant amount of residual AGO2 remains on chromatin despite at least 90% depletion of total AGO2 (data not shown). As a more rigorous test, I examined $AGO2^{51B}$ null mutants derived from mothers with $AGO2^{51B}$ ovaries by deriving germline clones; these mutants contain no maternal or zygotic protein. ChIP was performed on adult heads of $AGO2^{51B}$/+ or $AGO2^{51B}$ mutant siblings derived from the germline clones as well as from wild type flies. ChIP profiles of CTCF and CP190 in adult head tissue are similar to that observed in S2 and S3 cells but with considerable enrichment at the *Fab-7* insulator (Figure 3-8A, primer set 2). Importantly, no changes were observed in $AGO2^{51B}$ null mutants compared to heterozygous siblings or to wild type. Pc chromatin association is also unchanged in $AGO2^{51B}$ null mutants (data not shown). These results in combination with the finding that CP190 and CTCF localization is unchanged in polytene chromosomes of $AGO2^{51B}$ null mutants (Figure 3-6B), suggest that AGO2 is not required for CTCF or CP190 recruitment.

Figure 3-8. AGO2 is required for looping at the *Abd-B* locus.

(A) CTCF and CP190 chromatin association is unaffected in $AGO2^{51B}$ null mutants. Adult heads of wild type (blue) as well as $AGO2^{51B}$/+ (red) or $AGO2^{51B}$ (green) derived from $AGO2^{51B}$ germline clones were subjected to ChIP using α-CP190, α-CTCF, and α-Pc antibodies. Locations of primer sets are indicated in Figure 3-11B. Percent input DNA immunoprecipitated is shown for each primer set, and error bars indicate standard deviation of quadruplicate PCR measurements.

(B) 3C looping interactions between *cis*-regulatory elements of the *Abd-B* locus are dependent on *AGO2*. Relative interaction frequencies between *EcoRI* restriction fragments (triangles) and anchor regions (red vertical lines) are shown for wild type (open circles), $CP190^{P11}/CP190^{H31-2}$ (filled red circles), $CTCF^{y+2}$ (filled green circles) and $AGO2^{51B}$ (filled orange circles) mutant larval brains and imaginal discs.


A
α-CP190
α-CTCF
WT
AGO2^51B/+
AGO2^51B
% input
RpL32
1
2
3
4
5
6
7
8
9
10
11
12
13
14
B
Fab-7
Fab-7 PRE
iab-7 enhancer
Fab-8
Fab-8 PRE
iab-8 enhancer
Abd-B RB
Abd-B RD
Abd-B RA
Abd-B RC
Abd-B RE
WT
AGO2^51B
CP190^P11/ CP190^H31-2
CTCF^y+2
Anchor
EcoRI sites
Relative interaction frequency
ChIP primers
Abd-B RE
Abd-B RC
Abd-B RA
Abd-B RD
Abd-B RB
CTCF (S2)
CP190 (Kc)
AGO2 (S2)
12720000
12740000
12760000
12780000
12800000
chr3R

CP190, CTCF and AGO2 are required for looping at the *Abd-B* locus

AGO2 association with non-insulator regions of *Abd-B* in CP190 and CTCF knockdowns is lost and may be the result of changes in looping interactions at this locus (data not shown, Moshkovich et al. 2011). In order to determine whether CP190 or CTCF insulator proteins mediate these or other long-range interactions in this locus, I examined locus-wide interactions by 3C in diploid larval brains and imaginal discs of wild type compared to *CTCF* (Gerasimova et al. 2007) and *CP190* null mutants. These tissues represent a mixed population with a minority of cells expressing *Abd-B*. We scanned pair wise interactions using available EcoRI restriction sites in an 80 kb region encompassing the *Fab-7* insulator to the most distal *Abd-B* RE promoter. Using an anchor at the *Abd-B* RB promoter, high levels of interaction are observed between the *Abd-B* RB promoter anchor and *Fab-7*, *Fab-8*, and *iab-8* enhancer in wild type (Figure 3-8B). These interactions are decreased 1.5 to two-fold in both $CP190^{P11}/CP190^{H31-2}$ mutants and further decreased in $CTCF^{y+2}$ null mutants. In addition, using an anchor at *Fab-8*, interactions with *Fab-7* and the *Abd-B* RB promoter are decreased approximately two-fold in *CP190* mutants with a greater decrease in *CTCF* mutants compared to wild type. These results indicate a requirement for both CP190 and CTCF for looping interactions between insulators, PREs, enhancers, and promoters of the *Abd-B* locus.

I also addressed the possibility that *AGO2* is required for insulator-dependent looping interactions at *Abd-B*. Similar to *CP190* and *CTCF* mutants, 3C in diploid larval brains and imaginal discs of *AGO2* null mutants from germline clones demonstrated high levels of interaction between the *Abd-B* RB promoter anchor and *Fab-7*, *Fab-8*, and *iab-8*

enhancer in wild type (Figure 3-8B). These interactions are decreased 1.5 to two-fold in *AGO2*^{51B} mutants. High frequency interactions between the *Fab-8* anchor and *Fab-7* and the *Abd-B* RB promoter are decreased approximately two-fold in *AGO2* compared to wild type, suggesting that *AGO2* is required for CTCF/CP190 insulator-dependent looping interactions at *Abd-B*. The looping interactions, mediated by both AGO2 and the insulator proteins, may be required for proper expression of *Abd-B* (data not shown; Moshkovich et al. 2011).

DISCUSSION

Here, I provide the first evidence for an Argonaute protein functioning directly on euchromatin to effect changes in gene expression. The genome-wide binding profile of AGO2 displays striking overlap with insulator proteins. Genetic analysis revealed that AGO2, independent of its catalytic activity, promotes *Fab-8* insulator activity. Like known insulator proteins, AGO2 also associates with promoters and can oppose PcG function. Genome-wide AGO2 recruitment to chromatin is downstream of CTCF and CP190 binding and may be partially achieved via looping interactions among *cis*-regulatory regions and promoters. I propose that AGO2 may act to facilitate or stabilize looping that is needed to partition the genome into independent transcriptional domains (Figure 3-9).

AGO2 localizes predominantly to euchromatin and not heterochromatin

The presented results suggest that the main function of AGO2 on chromatin resides in euchromatin and not in heterochromatin. Immunofluorescence localization of AGO2 in polytene chromosomes and cell lines indicates exclusion from heterochromatic and HP1-enriched regions. Furthermore, the majority of chromatin-associated AGO2 resides in non-repetitive euchromatic but not repeat-rich regions as determined by genome-wide ChIP-seq. I suggest that the role of AGO2 in RNAi-dependent silencing of TEs occurs primarily at the posttranscriptional level and that AGO2 harbors a second RNAi-independent activity to promote chromatin insulator function.

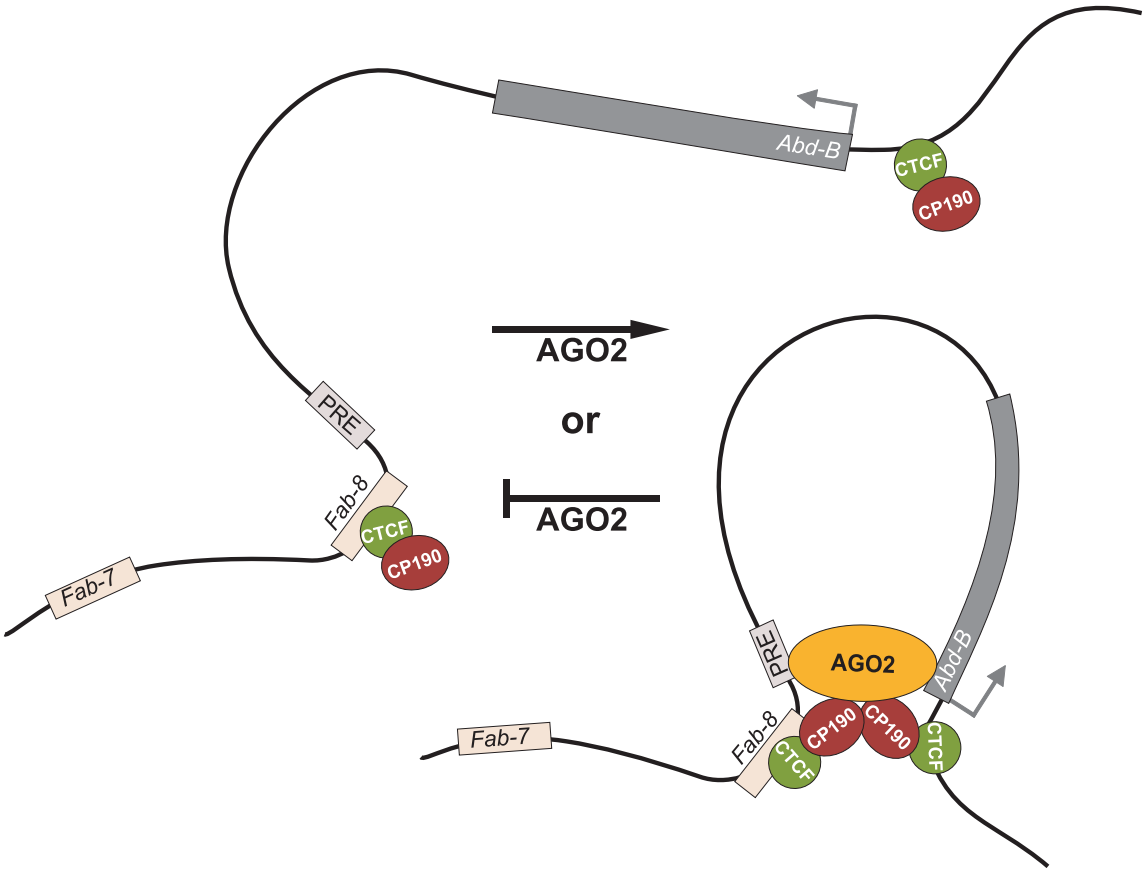


Figure 3-9. Model for AGO2 function with respect to CTCF/CP190 chromatin insulator activity.

Looping at the *Abd-B* locus between the *Fab-8* insulator and *Abd-B* promoter is dependent on CTCF/CP190 insulator interactions. This specialized configuration promotes interactions between *Fab-8* associated *cis*-regulatory elements and the promoters to facilitate proper gene expression. AGO2 is recruited downstream of CTCF/CP190 chromatin association and acts to either promote or stabilize looping interactions. Transfer of AGO2 to non-insulator sites may be achieved through CTCF/CP190-dependent looping interactions.

RNAi-independent function for AGO2 at chromatin

Several observations suggest that AGO2 chromatin association is mainly, if not exclusively, independent of the RNAi pathway. First, AGO2 chromatin association does not correspond to regions of the genome that produce high levels of endo-siRNAs, which are dependent on Dcr-2 and AGO2 (Chung et al. 2008; Czech et al. 2008; Ghildiyal et al. 2008; Kawamura et al. 2008; Okamura et al. 2008). Second, AGO2 but not Dcr-2 is required for *Fab-8* insulator function. Finally, a catalytically inactive AGO2 protein, which is defective for RNAi, retains the ability to associate with chromatin and is functional with respect to both TrxG function and *Fab-8* insulator activity.

An intriguing question raised by these findings is whether or not the functions of AGO2 in RNAi and chromatin insulator activity are completely distinct. It was determined that *CP190* mutants remain competent for silencing using a *GMR-wIR* hairpin transgene (Lee et al. 2004), suggesting that AGO2 chromatin association is not required for RNAi (data not shown). Nevertheless, it remains possible that chromatin-associated AGO2 is loaded with siRNA. Future work will address how AGO2 subcellular localization and seemingly disparate functions in RNAi and chromatin insulator activities are regulated.

Role of AGO2 in *Fab-8* insulator function

AGO2 but not other RNA silencing factors exerts a unique positive role in *Fab-8* insulator function. Importantly, a catalytically inactive mutant form of AGO2 expressed at wildtype levels retains insulator activity, further suggesting that the RNAi pathway is dispensable for *Fab-8* insulator function. A significant fraction of AGO2 resides in the nucleus, and physical interaction is observed between AGO2 and CP190. This interaction is insensitive to RNaseA, suggesting that RNA does not mediate the interaction between AGO2 and CP190. It remains possible that AGO2 can interact with siRNA or other RNA while associated with the insulator complex, although there is no evidence to support this hypothesis.

I show for the first time that chromosomal looping in the *Abd-B* locus is dependent on CTCF, CP190, and AGO2. Confirming and extending previous studies, we find that the *Abd-B* RB promoter interacts frequently with *Fab-7, Fab-8,* and the *iab-8* enhancer and moreover that the *Fab-8* region also contacts *Fab-7* as well as multiple *Abd-B* promoters. Currently, the significance of insulator protein promoter association is unclear, but insulators may be thus situated to control looping interactions between promoters and *cis*-regulatory elements. Depletion of CP190 or CTCF reduces these high frequency looping interactions, and loss of this specialized chromatin configuration could result in disassociation of AGO2. Given this possibility, AGO2 may act to detect the insulator-dependent conformation of this locus.

AGO2 is recruited to chromatin insulator sites as well as non-insulator sites in a CTCF/CP190-dependent manner. I speculate that AGO2 chromatin association with insulator sites could result from physical interactions with CP190 complexes, while AGO2 recruitment to other sites may be achieved at least in part by chromatin looping

mediated by CP190 and CTCF. In fact, it was recently shown that PcG proteins can be transferred from a PRE to a promoter as a result of intervening insulator-insulator interactions (Comet et al. 2011). Once recruited to chromatin, AGO2 could perform a primarily structural function to promote or stabilize the frequency of CTCF/CP190-dependent looping interactions.

Role of AGO2 in long range chromosomal interactions

AGO2 appears to promote *Fab-8* insulator activity independently of an effect on *gypsy* insulator body localization. Previous work showed that both the *gypsy* class and CTCF/CP190 insulators colocalize to insulator bodies, suggesting that these subnuclear structures may be important for both *gypsy* and *Fab-8* activities (Gerasimova et al. 2007). However, since *Fab-8* activity is not affected by RNA silencing components that disrupt *gypsy* insulator body localization, this subnuclear structure appears to be dispensable for *Fab-8* function. Recent work indicates that the BX-C harbors multiple redundant *cis*-regulatory elements that can maintain looping interactions of this locus (Bantignies et al. 2011), suggesting that the configuration of the BX-C may not require a nuclear scaffold such as the *gypsy* insulator body.

AGO2 mutations suppress the *Polycomb* phenotype, indicating that AGO2 behaves similarly to *trxG* genes and opposes *PcG* function. A previous study proposed that RNA silencing factors promote long-range PRE-dependent chromosomal pairing as well as PcG body formation but did not examine *AGO2* (Grimaud et al. 2006). I found that the $AGO2^{51B}$ null mutation has no effect on *Fab-X* PRE pairing-dependent silencing

on *sd* as assayed in that study (data not shown), and the genetic results suggest that AGO2 is unlikely to promote PRE-dependent interactions or PcG body formation, which are both positively correlated with PcG function. Interestingly, it has recently been shown in the case of AGO2-associated *Fab-7* and *Mcp* boundary elements, that long-range interactions are dependent on insulator sequences and not PREs (Li et al. 2011). Future studies will elucidate the complex interplay between PcG and insulator organization as well as the role of AGO2 in the regulation of these structures.

Conclusions

It remains to be seen whether *Drosophila* AGO2 euchromatin association and function may be conserved in other organisms. In *C. elegans*, the nuclear NRDE RNAi pathway can block transcriptional elongation of Pol II on a target transcript when treated with exogenous complementary dsRNA (Guang et al. 2010). Interestingly, this negative transcriptional effect is contemporaneous with an increase in H3K9me3. Whether the Argonaute protein NRDE-3/WAGO-12, which lacks Slicer activity, associates with euchromatin to effect this repression is not yet known. Furthermore, the *C. elegans* Argonaute Csr-1, loaded with 22G endo-siRNAs antisense to mRNAs of holocentric chromosomes, may serve as chromosomal attachment points to promote efficient chromosome segregation (Claycomb et al. 2009; van Wolfswinkel et al. 2009). Recently, it has been shown that *S. pombe* Ago1 participates in surveillance mechanisms to prevent read-through transcription of mRNA (Gullerova and Proudfoot 2008; Zofall et al. 2009; Halic and Moazed 2010). However, the majority of Ago1 associates with

heterochromatic regions (Noma et al. 2004), and it is not clear thus far if Ago1 directly associates with euchromatin or acts posttranscriptionally. An emerging theme from studies of RNAi in various model systems is that genome integrity and control of gene expression may be achieved by multiple yet overlapping mechanisms.

ACKNOWLEDGEMENTS

We thank A. Beyer for α-Pep, F. Fuller-Pace for α-p68, J. Kassis for α-Pho, Q. Liu, A. Mueller, and M. Siomi for α-AGO2, P. O'Farrell and D. Moazed for α-Pc; and C. Berg, J. Birchler, R. Carthew, V. Corces, G. Hannon, S. Hou, F. Karch, J. Kassis, G. Shanower, P. Schedl, and P. Zamore for strains. We are indebted to P. Murphy for antibody characterization and M. Emmett for primers; A. Dean for 3C protocols; S. Grewal, F. Karch, and B. Oliver for discussions; and J. Kassis and L. Matzat for critical reading of the manuscript.

AUTHOR CONTRIBUTIONS

Elissa P. Lei	Performed immunoaffinity purification with α-CP190 and RNaseA treatment. Performed immunoprecipitation with α-AGO2. Performed immunostaining of salivary gland polytene chromosomes and S2 and S3 cell staining. Contributed to the writing of the manuscript presented in this chapter.
Ryan K. Dale	Performed AGO2 ChIP-seq analysis, AGO2 binding site comparisons with other datasets, d*e novo* AGO2 motif analysis and siRNA analyses. Generated Figures 3-1, 3-2, 3-3 and 3-4A.
Parul Nisha	Performed whole mount immunofluorescence staining of insulator bodies.
Patrick J. Boyle	Designed primers for ChIP.
Brandi A. Thompson	Designed primers for 3C and prepared 3C control template.

MATERIALS AND METHODS

Drosophila strains

Flies were maintained on standard cornmeal medium at R.T. or 25°C. Newly eclosed flies were collected and aged for 24-27 h and examined for eye pigmentation. Larvae for immunostaining of imaginal discs were raised at 25°C. Larvae for immunostaining of polytene chromosomes were raised at 18°C. The *Fab-8* Insulator + PRE transgene contains a HindIII-EcoRI fragment, and the *Fab-8* PRE transgene contains an EcoRI-AflII fragment (Barges et al. 2000). Transgenes were scored as single copy. The $AGO2^{51B}$/+ mutation was tested on five independent *Fab-8* Insulator + PRE insertion lines, and similar results were observed.

Homozygous $AGO2^{51B}$ flies exhibit a high degree of male and female sterility, but these phenotypes appear to be caused by second site mutations unlinked to the *AGO2* mutation. Furthermore, $AGO2^{51B}$ mutants exhibit a low, variable level of protein likely maternally deposited. Consequently, homozygous mutant germline clones were produced by recombining the $AGO2^{51B}$ mutation with FRT2A and inducing recombination with a ovo^{D1} marked FRT2A chromosome using a hs-FLP recombinase induced for 1h in larvae at 5 d and 6 d of age as described previously (Selva and Stronach 2007). These flies were then crossed to $AGO2^{51B}$/+ males to obtain the desired progeny. The progeny were verified by Western blotting, PCR, and ChIP, and the same results were obtained with $AGO2^{321}/AGO2^{454}$ null mutants (Hain et al. 2010) from $AGO2^{321}$ germline clones (data not shown).

Indirect immunofluorescence

Preparation and immunostaining of salivary gland polytene chromosomes was performed as described previously (Lei and Corces 2006). Cell staining and whole mount staining are detailed in Supplemental Experimental Procedures. Rabbit α-Su(Hw) (Moshkovich and Lei 2010) and guinea pig α-CP190 (generated similarly as in (Pai et al. 2004)), rabbit α-CP190 (Pai et al. 2004), rat α-CTCF (Gerasimova et al. 2007), rabbit α-CTCF (Gerasimova et al. 2007), mouse α-AGO2 (9D6) (Kawamura et al. 2008), and rabbit α-Pc antibodies (a kind gift from D. Moazed and P. O'Farrell, UCSF) were used for staining. Samples were mounted with Vectashield (Vector Laboratories), and images were acquired on a Leica DM5000 epifluorescence microscope with Openlab (Perkin Elmer) software.

Western blotting

Lysates from whole pupae, the anterior third of third instar larvae, whole flies, or cell lines were prepared as described previously (Lei and Corces 2006). Guinea pig α-CP190, rabbit α-CP190, (generated similarly as in (Pai et al. 2004)), guinea pig α-Su(Hw) (generated similarly as in (Moshkovich and Lei 2010)), guinea pig α-Mod(mdg4)2.2 (Moshkovich and Lei 2010), rabbit α-CTCF, rat α-CTCF, mouse α-Pep (Amero et al. 1991), MAD1 mouse α-p68 (Ishizuka et al. 2002), mouse α-AGO2 (4D2)

(Okamura et al. 2004), mouse α-AGO2 (9D6), rabbit α-Piwi (Abcam ab-5207), rabbit α-AGO2 (Abcam ab-5072), mouse α-Lamin (ADL67.10) (Stuurman et al. 1996), rabbit α-Pc and rabbit α-TRX-C (Schuettengruber et al. 2009) were used for Western blotting. Specificity of generated and commercial antibodies was verified by blotting mutant fly lysates and/or cells knocked down with the corresponding dsRNA.

Immunoaffinity purification

Immunoaffinity purification with α-CP190 and RNaseA treatment was carried out as described previously (Lei and Corces 2006). Immunoprecipitation with α-AGO2 (9D6) was performed using nuclei isolated from 20 g of 0-24 h embryos as described previously (Lei and Corces 2006). Nuclei were lysed by sonication in 5 mL HBSMT-0.3% + 1 M KCl (50 mM HEPES, 150 mM NaCl, 1 M KCl, 3 mM $MgCl_2$, 0.3% Triton X-100 (vol/vol), pH 7) including 1 mM PMSF and Complete protease inhibitor cocktail (Roche) and extracts prepared as described previously (Lei and Corces 2006). 1.2 mL of extract was bound overnight at 4˚C to 1 mL α-AGO2 (9D6) tissue culture supernatant or 1.4 µg mouse IgG (Santa Cruz) prebound to rProtA-sepharose for 1 h at 4˚C. Beads were washed three times with HBSMT-0.3% + 1 M KCl then once with HBSM (50 mM HEPES, 150 mM NaCl, 5 mM KCl, 3 mM $MgCl_2$), and eluted with denaturing sample buffer by boiling for 5 min. Samples were Western blotted as described previously (Pai et al. 2004).

Chromatin immunoprecipitation and ChIP-seq

Preparation of ChIP samples and analysis was performed essentially as described previously (Moshkovich and Lei 2010). S2 and S3 cells were grown at 25°C in Shield and Sangs M3 Insect medium (Sigma) supplemented with 0.1% yeast extract, 0.25% bactopeptone, and 10% fetal bovine serum (HyClone). Immunoprecipitations were performed with α-AGO2 (9D6), rabbit α-CP190 (this study), and rabbit α-CTCF, rabbit α-Pho (Fritsch et al. 1999), rabbit α-Pc, rabbit α-TRX N-terminal (Schwartz et al. 2010), rabbit α-H3K9ac (Abcam, ab4729), mouse IgG, and rabbit IgG (Santa Cruz Biotechnology) coupled to rProtein A agarose beads (GE Healthcare). Similar results but with lower signal were obtained with rabbit α-AGO2 (Jiang et al. 2005), rabbit α-AGO2 (Meyer et al. 2006) or with α-Flag (M2, Sigma) or α-HA (12CA5, Santa Cruz) using chromatin prepared from HA/Flag-AGO2 transgenic flies also expressing wildtype AGO2 (Czech et al. 2008). Primers used are indicated in Supplemental Table S3.

Samples for ChIP-seq from input DNA and AGO2 ChIP were prepared according to the manufacturer's protocol (Illumina). DNA was sequenced on an Illumina Genome Analyzer at the NIDDK Genomics Core. Computational methods are detailed in Supplemental Materials and Methods. AGO2 ChIP-seq data are available at GEO (GSE22623).

AGO2 ChIP-seq analysis

36 bp sequenced tags were mapped to the *D. melanogaster* genome (dm3) with Bowtie (Langmead et al. 2009), retaining uniquely mapping reads with up to 2

mismatches in the first 28 bp (-m1 -n2 --best). Binding sites were identified using MACS 1.3.7.1 (Zhang et al. 2008) (--tsize 36 --bw 150). The resulting peaks were filtered to retain only those peaks with a false discovery rate of 5% or below. To assess the possible enrichment of AGO2 binding in heterochromatin, we separated the reads into uniquely mapping (i.e., euchromatic) and multiply mapping (i.e., repetitive sequence or heterochromatin) and asked whether the ratio of unique to multiply mapping reads was different in IP versus input using a chi-square test.

AGO2 binding site comparisons with other datasets

ChIP-chip tiling array data were downloaded from GEO, ArrayExpress, or supplemental material as available (Schwartz et al. 2006; Czech et al. 2008; Kawamura et al. 2008; Lee et al. 2008; Oktaba et al. 2008; Bushey et al. 2009; Schuettengruber et al. 2009; Negre et al. 2010; Schwartz et al. 2010) as described in Table S2 and stored as a collection of BED-format files. Where possible, called peaks from the original study were used to retain consistency with published work. Otherwise the peak-calling algorithm described by the authors was implemented to obtain binding sites. BED files were mapped from dm2 to dm3 if necessary with the liftOver tool from UCSC (http://genome.ucsc.edu/cgi-bin/hgLiftOver) and were filtered to remove features from the heterochromatic chromosomes since some data sets only considered the euchromatic chromosomes. Promoters were defined as 250 bp upstream from the transcription start site (TSS) of transcripts (mRNA, tRNA, rRNA, snoRNA, snRNA, miRNA) annotated in FlyBase r5.29. The "AGO2 minus GAF" BED file was created by removing all sites in

the AGO2 S2 BED file that overlapped by at least 1 bp with the GAF S2 sites. The TrxG (all Ash1, Trx-C, and Trx-N sites), PcG (all E(z), H3K27me3, Pc, Ph, PRE, and Psc sites), and insulators (all BEAF-32, CTCF, CP190, and Su(Hw) sites) BED files were created by concatenating all sites together and merging into non-redundant sites.

7747 active promoters were defined in S2 cells as having both H3K4me3 and PolII 250 bp upstream or 750 bp downstream from the TSS. 7281 inactive promoters were defined as having neither factor in this same window. H3K4me3 data from S2 cells were obtained from (Negre et al. 2010) (GEO accession GSM409457). PolII data from S2 cells were obtained from modENCODE consortium (modMine.org). Unique transcription start sites for all annotated RNAs were retrieved from FlyBase r5.29 and provided to CEAS (Shin et al. 2009) to calculate a profile for IP and input separately. The resulting profiles were scaled by library size, and the input profile was then subtracted from the IP profile to obtain the final scaled, input-subtracted profile. Profiles for transcription termination, active promoters, and inactive promoters were calculated similarly.

For the binary heatmap, supervised hierarchical clustering of overlap by at least 1 bp was performed as in (Kim et al. 2008). For the colocalization heatmap, permutation tests were performed to assess the degree of overlap between binding sites of each pair of factors similar to a previous study (Negre et al. 2010). Specifically, for each pair of factors A and B, the number of features in A that overlapped with at least one feature in B was calculated using BEDtools v2.6.1 (Quinlan and Hall 2010). Then features in A were randomly shuffled within each chromosome, and intersections with B were again calculated. This was repeated 1000 times for each pair of factors, resulting in a null

distribution of intersections. An empirical p-value was calculated as the percentile of the original intersection count within this distribution, and the enrichment score is defined as original intersection count divided by the median of the null distribution. The matrix of pairwise enrichment scores was clustered using complete linkage using correlation as the distance metric (as implemented in scipy.cluster in the SciPy package for Python).

De novo motif analysis

MEME (Bailey and Elkan 1994) v4.5.0 was used for *de novo* motif finding. The center 500 bp were extracted from a randomly selected set of 500 of the 2084 S2 AGO2 peaks. DNA mode was used with a maximum number of 3 iterations, a maximum width of 15 bp, up to 3 motifs, allowing zero or one motif per sequence, and in reverse-complement mode (parameters -dna -maxiter 3 -maxw 15 -nmotifs 3 -revcomp -mod zoops). Similar motifs were found with GADEM using all 2084 sites and the center-weighted option and Weeder using the center 500 bp of 500 random AGO2 sites and default parameters.

siRNA analyses

Data were obtained from GSM266765 (Kawamura et al. 2008), GSM280087 (Czech et al. 2008), GSM272652 (Okamura et al. 2008) and GSM239051 (Ghildiyal et al. 2008) and pooled. NLS-P19 data (Fagegaltier et al. 2009) were processed separately. Adapters were removed from reads before being mapped to the dm3 assembly with

Bowtie, allowing only uniquely-mapping reads and allowing up to one mismatch in the first 28 bp (parameters -m 1 -k 1 -n 1 --best --strata). Reads were then size-filtered, retaining only those 19-22 nt in length, and further filtered by removing those overlapping by 1 or more bp with annotated mirBase miRNA sites (mirbase.org) (Griffiths-Jones et al. 2008). Reads falling in the *white* gene were also removed. Finally, PCR bias was removed with Picard's MarkDuplicates program (Li et al. 2009b), collapsing duplicate reads into a single read. For the heatmaps, siRNA reads were then clustered according to a previously described algorithm (Czech et al. 2008).

For the cumulative histograms, all siRNA reads retained after the PCR-filtering step were used, and their overlap was computed with all possible 3' *cis*-NATs, as previously reported (Okamura et al. 2008), and divided by the length of each 3' *cis*-NAT to obtain an siRNA density for each feature. Random backgrounds were calculated by averaging the histograms of 1000 randomizations, where each randomization consisted of intrachromosomal shuffling of sites bound by a particular factor while keeping the siRNA reads fixed. siRNA densities were similarly calculated for PolII, H3K27me3, and AGO2.

Double stranded RNA and siRNA knockdowns

Amplicons used for dsRNA knockdowns were designed based on recommendations from the *Drosophila* RNAi Screening Center. Templates were PCR amplified from genomic DNA using primers containing the T7 promoter sequence. dsRNAs were produced by *in vitro* transcription of PCR templates using the MEGAscript T7 kit (Ambion) and purified using NucAway Spin Columns (Ambion). Transfections

using 200 ng-1.25 μg of dsRNA or 100 pmol of siRNA per million cells, or no dsRNA/siRNA for mock treatment were performed using Cellfectin (Invitrogen), Effectene (Qiagen), or Cell Line Nucleofector Kit V (Amaxa Biosystems) transfection reagent using the recommended protocol. Four to six days after transfection, cells were collected, and knockdown efficiency was confirmed by Western blotting. Highest knockdown efficiencies were generally obtained using the Amaxa system. No differences were seen with mock treatment, *GFP* dsRNA, or *luciferase* dsRNAs. Primers used are indicated in Supplemental Table S3, and 3C methods are detailed in Supplemental Materials and Methods.

Cell staining

S2 and S3 cells were prepared for staining by washing in PBS and affixed to poly-L-lysine coated slides for 10 min at R.T. Cells were fixed in 4% paraformaldehyde + 0.1% Triton X-100 in PBS for 30 min and then blocked in 10% normal goat serum + 0.1% Triton X-100 in PBS for at least 30 min. Cells were incubated with rabbit α-CP190 (Pai et al. 2004) and mouse α-AGO2 (9D6) including 0.1% Triton X-100 overnight at 4°C, washed 3 times in PBS + 0.1% Triton X-100 (PBS-X-0.1%) for 15 min at R.T., and incubated with Alexa594 anti-rabbit and Alexa488 anti-mouse antibodies (Invitrogen) in PBS-X-0.1% + 10% normal goat serum for 2 h at 37°C. Samples were washed 3 times in PBS-X for 15 min at R.T., and stained with 100 ng/mL DAPI in PBS at R.T. for 1 min.

Whole mount staining

For whole mount immunofluorescence staining of insulator bodies, brain and imaginal disc complexes from third instar larvae were dissected in PBS and fixed in 4% paraformaldehyde in PBS + 0.1% Tween (PBS-Tw) for 20 min on a rotating wheel. Samples were rinsed three times in PBS-Tw and blocked for 2 h in PBS + 0.3% Triton X-100 (PBS-X-0.3%) + 10% normal goat serum then incubated with rabbit α-CP190 (Pai et al. 2004) with blocking solution overnight at 4°C on a rotating wheel. Samples were washed 3 times for 5 min each with PBS-X-0.3% followed by 3 times for 20 min in the same buffer on a rotating wheel at R.T. Alexa594 anti-rabbit antibody in PBS-X-0.3% + 10% normal goat serum was added for 2 h at 37°C on a rotating wheel. Samples were washed 3 times for 5 min with PBS-X followed by 3 times for 20 min, followed by one PBS-Tw wash for 5 min. Samples were stained with 100 ng/mL DAPI in PBS at R.T. for 5 min, followed by a PBS-Tw wash for 5 min.

Chromosome Conformation Capture (3C)

Drosophila S2 cells used for the 3C assays were grown in M3+BYPE (10% FBS). Before crosslinking, 4×10^6 cells were resuspended in 5 mL of fresh media. Crosslinking was performed by adding formaldehyde directly to the media at final concentration of 1% and incubating for 10 min at R.T. Reactions were quenched by adding glycine to a final concentration of 0.125 M and incubating for 5 min at RT. Reactions were incubated on ice 5 min followed by centrifugation at 1200 rpm at 4°C for 5 min. Cells were then washed with 5 mL of cold PBS and centrifuged at 1200 rpm at 4°C for 5 min. Lysis was

performed by incubating cells in 1 mL of Lysis buffer [10 mM NaCl, 0.2% NP-40, 10 mM Tris pH=8, protease inhibitors [1 Mini Complete tablet (Roche) per 10 mL Lysis buffer)] at 37°C for 20 min. Samples were then centrifuged at 4200 rpm at 4°C for 5 min, and the lysis step was repeated once.

Brains and imaginal discs were dissected from ten male and ten female third instar larvae in Schneider's S2 medium without serum and immediately centrifuged for 30 s at 6000 rpm. Pellets were resuspended in 700 µl of Fixing buffer (50 mM HEPES pH 7.6, 100 mM NaCl, 0.1 mM EDTA, 0.5 mM EGTA) and 100 µl 16% paraformaldehyde and rocked for 15 min at R.T. Reactions were quenched by adding 1 mL Stop solution (PBS, 0.01% TritonX-100, 0.125 M glycine) and rocked for 10 min at R.T. Reactions were washed with 1 mL Wash solution (50 mM Tris, 10 mM EDTA, 0.5 mM EGTA, 0.25% TritonX-100) twice rocking for 10 min. Pellets were homogenized with 200 µl of Lysis buffer using a motorized pellet pestle. 800 µl of Lysis buffer was added, and samples were incubated at 37° C for 20 min. Samples were then centrifuged at 4200 rpm at 4°C for 5 min, and resuspended in 1 mL Lysis buffer and incubated another 20 min at 37°C.

Pellets from cell or larval samples were then washed with 0.8 mL of digestion buffer [0.2% NP-40, 1X NEBuffer 3 (NEB)] and centrifuged at 4200 rpm at 4°C for 5 min. Nuclei were resuspended in 1.6 mL of digestion buffer with SDS added to a final concentration of 0.1% and incubated at 65°C for 30 min. Triton X-100 was added to a final concentration of 1% and incubated at 37°C for 15 min. A 40 µL aliquot of the sample was taken and used as the undigested control. The remaining sample was digested with 1600 U of EcoRI (NEB) at 37°C O/N. Samples were incubated 20 min at

65°C to inactivate EcoRI. A 40 µL aliquot of the sample was taken here and used as the digested control. The remaining sample was then diluted to 4 mL with ligation buffer [final concentrations were 1% Triton X-100, 1X T4 DNA Ligase Reaction Buffer (NEB)] and incubated at 37°C for 30 min. After the addition of 4800 U of T4 DNA Ligase (NEB), each sample was incubated at 16°C O/N. Proteinase K was added to all samples including the controls at a final concentration of 65 ng/mL. Samples were then incubated at 65°C O/N to reverse the crosslinking. After de-crosslinking, samples were combined with 1 vol of phenol/chloroform/isoamyl alcohol (25:24:1), vortexed 15 s, and centrifuged for 4 min at 13,000 rpm. The top layer was transferred to a new tube, and the procedure was repeated using 1 vol chloroform. The top layer was collected and subsequently diluted with 1 vol of dH_2O. The sample was combined with 0.1 vol of 3M NaOAc pH=5.2 and 2.5 vol of 100% ethanol. After incubating 1 hr at -80°C, samples were centrifuged 20 min at 4°C at13,000 rpm. Pellets were washed with 75% ethanol and centrifuged 5 min at 4° at13,000 rpm. Pellets were air dried at R.T. prior to resuspension in an appropriate volume of TE. Loading adjustment and digestion efficiency tests were performed using previously described methods (Hagege et al. 2007). Loading adjustment was performed by SYBR green quantitative PCR to the *yellow* locus, and samples were adjusted accordingly before TaqMan quantitative PCR for 3C.

Preparation of the Control Template for 3C

To prepare a control template containing all possible ligation products, equimolar amounts of bacterial artificial chromosomes [RP48-36F20 & CH321-96A10 (CHORI)]

spanning the loci of interest were digested in 1X NEBuffer 3 with EcoRI (NEB) at a concentration of 12 U/μg DNA. Digested DNA was purified by phenol:chloroform extraction and ethanol precipitation. DNA was then ligated with T4 DNA ligase in 1X T4 DNA Ligase Reaction Buffer (NEB) at 16°C O/N. Ligated DNA was purified by phenol:chloroform extraction and ethanol precipitation. A second digest was performed using HindIII (NEB) to linearize any DNA circles. Digested DNA was purified by phenol:chloroform extraction and ethanol precipitation.

Quantitative PCR for 3C

Primers were designed to flank all EcoRI restriction sites within the region of interest. Custom TaqMan TAMRA Probes (Applied Biosystems) were designed with 5'FAM reporter dye and 3'TAMRA quencher dye. Real-time PCR reactions were prepared using the TaqMan Universal PCR Master Mix (Applied Biosystems) and the recommended protocol. Each reaction was performed in quadruplicate. To normalize for the PCR efficiency of each primer pair/probe combination, the BAC control template was used to generate standard curves for each combination. Interaction frequencies were calculated based on the C_t values of each sample relative to the standard curve for the given primer pair/probe combination. Primers and probes used are listed in Table 3-3.

Table 3-1. Sources of data for tiling arrays, endo-siRNA, and genome features.

Antibody or feature type	Lab	Year	Cell Type	Label in heatmap	reference	Source	Peak calling[1]
Active promoters	Multiple	2010	S2	S2: active promoters	Negre et al. 2010, modMine	GSM409457, http://intermine.modencode.org	7
AGO2 9D6	all	2011	all	All 9D6 AGO2	this study	GSE22623	6
AGO2 9D6	Lei	2011	S2	S2: 9D6 AGO2	this study	GSE22623	2
AGO2 9D6	Lei	2011	S3	S3: 9D6 AGO2	this study	GSE22623	2
AGO2 Mueller	Lei	2011	S2	S2: Mueller AGO2	this study	GSE22623	2
AGO2 9D6 without GAF	Lei	2011	S2	S2: 9D6 AGO2 without GAF	this study	GSE22623	1
all-*cis*NAT	Lai	2010	all	3' *cis*-NATs	Okamura et al. 2008	Table S1	1
ASH1-mono	Pirrotta	2010	Sg4	Schwartz (2010) Sg4: ASH1-mono	Schwartz et al. 2010	GSM454525	3
ASH1-poly	Pirrotta	2010	Sg4	Schwartz (2010) Sg4: ASH1-poly	Schwartz et al. 2010	GSM454524	3
BEAF-32	Corces	2009	Kc	Bushey (2009) Kc: BEAF-32	Bushey et al. 2009	Supplemental files	1
BEAF-32	Corces	2009	Mbn2	Bushey (2009) Mbn2: BEAF-32	Bushey et al. 2009	Supplemental files	1
BEAF-32	White	2010	embryo	Negre (2010) embryo: BEAF-32	Negre et al. 2010	GSM409067	1
CP190	Corces	2009	Kc	Bushey (2009) Kc: CP190	Bushey et al. 2009	Supplemental files	1
CP190	Corces	2009	Mbn2	Bushey (2009) Mbn2: CP190	Bushey et al. 2009	Supplemental files	1
CP190	White	2010	embryo	Negre (2010) embryo: CP190	Negre et al. 2010	GSM409068	1
CTCF	Corces	2009	Kc	Bushey (2009) Kc: CTCF	Bushey et al. 2009	Supplemental files	1
CTCF	Corces	2009	Mbn2	Bushey (2009) Mbn2: CTCF	Bushey et al. 2009	Supplemental files	1
CTCF	White	2010	Kc	Negre (2010) Kc: CTCF	Negre et al. 2010	GSM409079	1
CTCF	White	2010	S2	Negre (2010) S2: CTCF	Negre et al. 2010	GSM409078	1
CTCF-C	White	2010	embryo	Negre (2010) embryo: CTCF-C	Negre et al. 2010	GSM409069	1
CTCF-N	White	2010	embryo	Negre (2010) embryo: CTCF-N	Negre et al. 2010	GSM409070	1
dSfmbt	Mueller	2008	imaginal	Oktaba (2008) imaginal: Sfmbt	Oktaba et al. 2008	Table S3	3
DSP1	Cavalli	2009	embryo	Schuettengruber (2009) embryo: Dsp1	Schuettengruber et al. 2009	Table S17.xls	1
E(z)	Pirrotta	2006	Sg4	Schwartz (2006) Sg4: E(z)	Schwartz et al. 2006	MEXP-535	3
GAF	Cavalli	2009	embryo	Schuettengruber (2009) embryo: GAF	Schuettengruber et al. 2009	Table S17.xls	1
GAF	Gilmour	2008	S2	Lee (2008) S2: GAF	Lee Mol Cell Biol 2008	Supplemental files	4
GAF	White	2010	embryo	Negre (2010) embryo: GAF	Negre et al. 2010	GSM409071	1
H3K27Ac-F	Pirrotta	2010	Sg4	Schwartz (2006) Sg4: H3K27Ac	Schwartz et al. 2010	GSM454533	3
H3K27me3	Cavalli	2009	embryo	Schuettengruber (2009) embryo: H3K27me3	Schuettengruber et al. 2009	Table S17.xls	1
H3K27me3	Pirrotta	2006	Sg4	Schwartz (2006) Sg4: H3K27me3	Schwartz et al. 2006	MEXP-535	3
H3K4me3	Cavalli	2009	embryo	Schuettengruber (2009) embryo: H3K4me3	Schuettengruber et al. 2009	Table S17.xls	1
H3K4me3	Pirrotta	2010	Sg4	Schwartz (2010) Sg4:	Schwartz et al.	GSM454526	3

				H3K4me3	2010		
H3K4me3	White	2010	embryo	Negre (2010) embryo: H3K4me3	Negre et al. 2010	GSM409075	1
H3K4me3	White	2010	Kc	Negre (2010) Kc: H3K4me3	Negre et al. 2010	GSM409458	1
H3K4me3	White	2010	S2	Negre (2010) S2: H3K4me3	Negre et al. 2010	GSM409457	1
H3K9Ac	Pirrotta	2010	Sg4	Schwartz (2010) Sg4: H3K9Ac	Schwartz et al. 2010	GSM454529	3
Inactive promoters	Multiple	2010	S2	S2: inactive promoters	Negre et al. 2010, modMine	GSM409457, http://intermine.modencode.org	7
insulator	all	all	all	insulators	multiple, see methods	see methods	6
mod(mdg4)	White	2010	embryo	Negre (2010) embryo: Mod(mdg4)	Negre et al. 2010	GSM409072	1
NELF-B	Gilmour	2008	S2	Lee (2008) S2: NELF-B	Lee et al. 2008	Supplemental files	4
NELF-E	Gilmour	2008	S2	Lee (2008) S2: NELF-E	Lee et al. 2008	Supplemental files	4
PC	Cavalli	2009	embryo	Schuettengruber (2009) embryo: Pc	Schuettengruber et al. 2009	Table S17.xls	1
PC	Pirrotta	2006	Sg4	Schwartz (2006) Sg4: Pc	Schwartz et al. 2006	MEXP-535	3
pcg	all	all	all	PcG	multiple, see methods	see methods	6
PH	Cavalli	2009	embryo	Schuettengruber (2009) embryo: Ph	Schuettengruber et al. 2009	Table S17.xls	1
PHO	Cavalli	2009	embryo	Schuettengruber (2009) embryo: Pho	Schuettengruber et al. 2009	Table S17.xls	1
PHO	Mueller	2008	embryo	Oktaba (2008) embryo: Pho	Oktaba et al. 2008	Table S1	1
PHO	Mueller	2008	imaginal	Oktaba (2008) imaginal: Pho	Oktaba et al. 2008	Table S2	1
PHOL	Cavalli	2009	embryo	Schuettengruber (2009) embryo: Phol	Schuettengruber et al. 2009	Table S17.xls	1
PolII	Pirrotta	2010	Sg4	Schwartz (2010) Sg4: PolII	Schwartz et al. 2010	GSM454527	3
PolII	White	2010	embryo	Negre (2010) embryo: PolII	Negre et al. 2010	GSM409077	1
PPEP	Adelman	2007	S2	Muse (2007) S2: PPEP	Muse et al. 2007	Table S1	1
PRE	Muller	2008	imaginal	Oktaba (2008) imaginal: PRE	Oktaba et al. 2008	Table S4	1
PRE	Pirrotta	2010	Sg4	Schwartz (2010) Sg4: PRE	Schwartz et al. 2010	Table S6	1
promoters	generated	2010	all	FlyBase: 250bp upstream from TSS	FlyBase	FlyBase r5.29	1
PSC	Pirrotta	2006	Sg4	Schwartz (2006) Sg4: Psc	Schwartz et al. 2006	MEXP-535	3
redfly	Gallo	2011	all	Merged REDfly CRMs	Gallo et al. 2011	http://redfly.ccr.buffalo.edu/	1
siRNA-cluster	Hannon-Siomi	2008	S2	S2: siRNA-clusters	multiple, see methods	see Methods	5
su(Hw)	Corces	2009	Kc	Bushey (2009) Kc: Su(Hw)	Bushey et al. 2009	Supplemental files	1
su(Hw)	Corces	2009	Mbn2	Bushey (2009) Mbn2: Su(Hw)	Bushey et al. 2009	Supplemental files	1
su(Hw)-1	White	2010	embryo	Negre (2010) embryo: Su(Hw)-Corces	Negre et al. 2010	GSM409073	1
su(Hw)-2	White	2010	embryo	Negre (2010) embryo: Su(Hw)-Geyer	Negre et al. 2010	GSM409074	1
siRNA-overlapping 9D6 AGO2 sites	This study	2011	S2	S2: siRNA-overlapping AGO2 9D6 sites	Multiple, see methods	See Methods	5
TRX-C	Cavalli	2009	embryo	Schuettengruber (2009) embryo: Trx-C	Schuettengruber et al. 2009	Table S17.xls	1

TRX-C-Beisel	Pirrotta	2010	Sg4	Schwartz (2010) Sg4: Trx-C-Beisel	Schwartz et al. 2010	GSM454521	3
TRX-C-Poux	Pirrotta	2010	Sg4	Schwartz (2010) Sg4: Trx-C-Poux	Schwartz et al. 2010	GSM454521	3
TrxG	all	all	all	TrxG	multiple, see methods	see methods	6
TRX-N	Pirrotta	2010	Sg4	Schwartz (2010) Sg4: Trx-N	Schwartz et al. 2010	GSM454523	3

Description of original data sources used in this study including GEO or ArrayExpress accessions as available.

[1]Peak-calling methods performed in order to obtain BED-format files for this study

1 discrete features (peaks, domains, genes) available directly

2 MACS v1.3.7.1 FDR 5%

3 Distance-based clustering algorithm from (Negre et al., 2010)

4 WIG data segmented into peaks

5 see Supplemental Materials and Methods for siRNA analysis

6 Concatenated and merged from other BED files as described in Supplementary Materials and Methods

7 Identified using H3K4me3 and PolII

Table 3-2. List of primers.

Primer	Sequence	Coordinates	Application[1]
1F	AGAAACCCATTGGTGCAGAC	chr3R:12724311-12724330	CH
1R	CAAAGTTGGATGCATTGTGG	chr3R:12724422-12724441	CH
2F	TCAAAGAGCGACACGTGAAC	chr3R:12724828-12724847	CH
2R	CATCAAACCTAGCCGCTCTC	chr3R:12725015-12725034	CH
3F	TCTTCGGGATGGCAATAAAC	chr3R:12726187-12726206	CH
3R	ACGATGTCGGATTCCTGAAC	chr3R:12726318-12726337	CH
4F	GGTTCTATTCCTAAAATTCTGTATGC	chr3R:12728314-12728339	CH
4R	GCATAACTCAAGGCCCGTTA	chr3R:12728417-12728436	CH
5F	CACGTGTTCGGTTTTCCTTT	chr3R:12744456-12744475	CH
5R	TTCCCTCCAATATGCAGACC	chr3R:12744686-12744705	CH
6F	GCAAGCGAAGAGTTCCATTC	chr3R:12744868-12744887	CH
6R	ACTGTCGGAGAGCGACATCT	chr3R:12744997-12745016	CH
7F	TGGTGGAAGGAGAAAACTGG	chr3R:12746370-12746389	CH
7R	TGCAGCGAGACAATAAAACG	chr3R:12746602-12746621	CH
8F	GCCAACCAGAAGGTCGTAAA	chr3R:12749233-12749252	CH
8R	GCTTCTCTTGGCGTTTCATC	chr3R:12749340-12749359	CH
9F	GCACTGTTTCAACTAGCGCCTTCA	chr3R:12760371-12760394	CH
9R	TTGAAGAGCGGATGCCTTCACACGTA	chr3R:12760508-12760533	CH
10F	CAAAGACGCGAACAAGTGAA	chr3R:12770226-12770245	CH
10R	TTGAACTTTGGCGGTACGAT	chr3R:12770343-12770362	CH
11F	CAAGTGAGTAGGCGATACGG	chr3R:12774451-12774470	CH
11R	CTCACGCTCTCGCAAAGTG	chr3R:12774537-12774555	CH
12F	CTCGTCTGCCTTCCATTCTC	chr3R:12786264-12786283	CH
12R	TTCTTTTGTCCGGGTAGTGG	chr3R:12786350-12786369	CH
13F	GGCGAATACGAAATCACCAC	chr3R:12789485-12789504	CH
13R	GTCAGGAGGAAACCACGAGA	chr3R:12789567-12789586	CH
14F	CACGCATTCTGCTGGTACAT	chr3R:12795809-12795828	CH
14R	CGGGCTCGTATCTGTGTCTC	chr3R:12795961-12795980	CH
RpL32F	CTGCATGAGCAGGA	chr3R:25871170-25871184	CH
RpL32R	ATGACCATCCGCCC	chr3R:25871488-25871501	CH
17	CTATGGGCCGCCAAATACCGTTTCAA	chrX:254644-254669	DE,LA
18	TGGAGACTACATTGCCTGAATTGGCG	chrX:254797-254822	DE,LA
81	GGAATTTCTTTGTCATTTCCACTGTGCC	chr3R:12743700-12743727	DE,SP
82	CGCTCTCTATCTGATCATCACACACGGT	chr3R:12743796-12743823	DE
83	ACAGAGTGCACTTGAGAAATCGGC	chr3R:12744035-12744058	DE,SP
84	TTGAAATTCCCGGTTCAAGTGCGG	chr3R:12744181-12744204	DE
85	CCGCACTTGAACCGGGAATTTCAA	chr3R:12744181-12744204	DE,SP
86	TGAGACATCAGGAAGAGGTTCGTTGG	chr3R:12744291-12744316	DE
87	TGTACATGAAGCATATGTAGACG	chr3R:12745316-12745338	DE,SP,TRTP
88	CTTCTACTTGCAGAACTTGTACTC	chr3R:12745394-12745417	DE
89	TTTCGTGGGTGGATGACTTTCCC	chr3R:12746979-12747001	DE,SP,TRTP
90	GCAACATATGGTACTTCCTGCGCT	chr3R:12747098-12747121	DE
91	TTTGAGCCAGAGGCTGACGACAT	chr3R:12749676-12749698	DE,SP,TRTP
92	TCAGATGAGCAGAATGCCGAAGGA	chr3R:12749770-12749793	DE
93	CATCCTCGGGCACTTTGAGACAACTT	chr3R:12749899-12749924	DE,SP,TRTP
94	GATCTGCTGTTGGTTAGAACACCTTT	chr3R:12749975-12750000	DE
103	AGTACGAAACACGAACTATGTGGGCG	chr3R:12720571-12720596	DE,SP,TRTP
104	AAATCACGTTCGTCGGAAGTGGAGA	chr3R:12720671-12720695	DE
105	AAGAGAGCGGCTAGGTTTGATGGT	chr3R:12725013-12725036	DE,SP,TRTP
106	AGACTTGCCTCAGCCTCTGAAT	chr3R:12725091-12725112	DE
107	ATTCAGAGGCTGAGGCAAGTCT	chr3R:12725091-12725112	DE,SP
108	TCACCGCTAGAGTTGGAAACCAGT	chr3R:12725211-12725234	DE
109	CCAACCATGCACACATCCAGGTAA	chr3R:12729389-12729412	DE,SP,TRTP
110	TGAAGAGAAGGCGGTTGGTCTGTT	chr3R:12729553-12729576	DE
111	TCATGTCGATTTCAGTCCGTAGCCAG	chr3R:12730755-12730780	DE,SP
112	TTAGCCCTGCCATAAAGTTCGGTTCC	chr3R:12730894-12730919	DE
113	TATGCAGTTGACGTCGGTTGATGC	chr3R:12732700-12732723	DE,SP
114	ACGAGATGGTGCGTCCATAAAGGT	chr3R:12732797-12732820	DE
115	TTGTTGTTCTGCTGATTGGCCTGG	chr3R:12755226-12755249	DE,SP
116	TTGGCCAGAAATTTGCAGCTGACC	chr3R:12755369-12755392	DE
117	GTCTTGGTAGCATTGAACAGTTAGGACAG	chr3R:12755990-12756018	DE,SP
118	AGTTAAGTGACCTCGCCAGCCAAT	chr3R:12756076-12756099	DE
119	ACATTTAGGTGGAATTTGAACGCCTCT	chr3R:12756585-12756611	DE,SP,TRTP
120	GCATCTTGCAACTCTAGTTTGGGAGG	chr3R:12756742-12756767	DE

121	GCACTGTTTCAACTAGCGCCTTCA	chr3R:12760371-12760394	DE,SP,TRTP
122	TTGAAGAGCGGATGCCTTCACACGTA	chr3R:12760508-12760533	DE
123	CGCATTTAGTTGAAGAGTCCAACTGCT	chr3R:12761292-12761318	DE,SP,TRTP
124	GGAAATAGATTGCGGCAGTTAATTACAAGT	chr3R:12761385-12761414	DE
125	GAATGGGAAAAGTTTCCGGCCTAAC	chr3R:12738565-12738589	DE,SP,TRTP
126	GGAAACATATTTTGGGATGGGCTTT	chr3R:12738732-12738756	DE
127	TTCGCCGCCATTTGCCGAAGG	chr3R:12746942-12746962	DE,SP,TRTP
129	AGAGGTAGTTAGACGATCGTGGGT	chr3R:12768301-12768324	DE,SP,TRTP
130	TGAGTGGATTTGACCACTTGGGTG	chr3R:12768412-12768435	DE
131	ATTCTGGCGATTCTGTCCCTTCCA	chr3R:12782513-12782536	DE,SP,TRTP
132	CTGGCATAGCAACGTAACAACTATGGG	chr3R:12782635-12782661	DE
133	ACCGACATCTTCATATCTGCCTTGC	chr3R:12782989-12783013	DE,SP
134	CGAAATTAAAGCATGTTCTCATTTAGG	chr3R:12783106-12783132	DE
135	CCTAAATGAGAACATGCTTTAATTTCGG	chr3R:12783106-12783133	DE,SP
136	GCTCGAAAAATCCAAGATAATTGACTGACC	chr3R:12783215-12783244	DE
137	GCACTCTCATATTTCCAAGAGCACACC	chr3R:12784025-12784051	DE,SP,TRTP
138	GGGTGTGTCCATACTTGCACTGT	chr3R:12784180-12784202	DE
139	CTCTCGTGGTTTCCTCCTGACC	chr3R:12789566-12789587	DE,SP,TRTP
140	TGTGTGTGTCAGGTGTTGTTACCC	chr3R:12789734-12789757	DE
141	AGAAGGACAAGGGAATGGGTGTGA	chr3R:12790763-12790786	DE,SP,TRTP
142	AGCAACTGCGGAGGCCATAAATTG	chr3R:12790879-12790902	DE
143	TCAATTGAAGCGCATCGCAACCGT	chr3R:12791097-12791120	DE,SP,TRTP
144	AGAAGCATGCTCCAGTTGACCCAA	chr3R:12791184-12791207	DE
145	CAAATATGCCGCCGGCTTTGGAAT	chr3R:12801377-12801400	DE,SP,TRTP
146	GGAGCTGCAAGGCTATCTTGATATGTATG	chr3R:12801639-12801667	DE
147	AAGATAGGAGTGGATGATGGCGCA	chr3R:12801832-12801855	DE,SP,TRTP
148	TTCGGAGATCGACGTTTAAGCCTG	chr3R:12801940-12801963	DE
121Probe	TCCAGAGCCAGTCCCAGTCGAAGTG	chr3R:12760415-12760439	TRTP
89Probe	TCTCAGCACGCGCTTTTCGTGG	chr3R:127469650-12746986	TRTP
125Probe	TGCGAGTTTATTAACCGCAAGTAATTTCACCAAA	chr3R:12738596-12738629	TRTP
AGO2-S	caaccacagcagcugcaacdTdT	chr3L:15547468-15547490	SI
AGO2-AS	guugcagcugcugugguugdTdT	chr3L:15547571-15547593	SI
Pc Fwd	TAATACGACTCACTATAGGGAGAGTACCGTGTCAAGTGGAAG	chr3L:21309750-21310571	DS
Pc Rev	TAATACGACTCACTATAGGGAGATTTGGTATGTTATTGTTCTCGG		DS
TRX Fwd	TAATACGACTCACTATAGGGAGACAAACGTCTACCACCACCC	chr3R:10098104-10098949	DS
TRX Rev	TAATACGACTCACTATAGGGAGAGCCCATTAGCGTGGTATCC		DS

[1]Legend

CH	Chromatin Immunoprecipitation
DS	dsRNA Amplicons
DE	Digestion Efficiency Test
LA	Loading Adjustment Test
SP	Standard PCR Detection of Interactions
TRTP	TaqMan Real-time PCR Detection of Interactions
SI	siRNA knockdowns

CHAPTER 4

DISCUSSION AND FUTURE DIRECTIONS

RNA SILENCING AND ITS ROLE IN HETEROCHROMATIN ASSEMBLY

piRNA- and endo-siRNA-mediated TE silencing

Our results indicate that heterochromatin forms independently of endo-siRNA and piRNA pathways. These findings suggest two possibilities for the heterochromatic silencing of TEs; sites other than piRNA producing loci serve as Piwi-dependent HP1 recruitment sites in *Drosophila* genome, or the existence of mechanisms alternative to RNA silencing that recruit HP1 to chromatin. We considered the first possibility and, in order to gain insight into the genome-wide chromatin association of Piwi, performed ChIP-seq of Piwi in *Drosophila* ovarian somatic cell (OSC) line. Our preliminary analysis revealed Piwi association with repeat-rich sequences but not with euchromatin (data not shown) indicating that no other Piwi-dependent HP1 binding platforms exist in euchromatin. Interestingly, our directed HP1 ChIP detected a two-fold decrease in HP1 recruitment at the *1360* element in OSC cells depleted of Piwi when compared to mock-treated cells. Coincidently, the *1360* and *F* elements, two TEs known to be preferentially bound by HP1, were shown to bind Piwi (reviewed in Brower-Toland et al. 2007). Whether HP1 recruitment to the *F* element is affected in Piwi-depleted OSC cells is

unknown. The same experiment, however, revealed that an additional target of HP1, *TART*, a telomere-specific non-LTR retrotransposon, was immunoprecipitated with HP1 at levels similar between mock-treated and Piwi-depleted OSC cells. Therefore, it is possible that a differential mechanism of TE heterochromatic silencing that may depend on Piwi exists at specific genomic sites. Further analysis will reveal the identity of these Piwi-associated repeat-rich sequences.

It is likely that different Piwi clade proteins silence TEs by mechanisms that may involve additional proteins. A recent study reported that one HP1 variant, Rhino, may play a role in piRNA-mediated TE silencing (Klattenhoff et al. 2009). Rhino, which is specifically expressed in the female germline, is required for TE silencing, and its localization to nuclear foci is independent of piRNA production. A model, which is mechanistically distinct from centromeric heterochromatin silencing in yeast, was proposed where Rhino does not appear to be involved in TE silencing. According to this model, Rhino binds to *42AB* piRNA cluster to promote transcription of piRNAs, which associate with Aub and AGO3, that most likely direct TE silencing through posttranscriptional target cleavage. The authors speculated that *rhino*, which is a rapidly evolving gene, may be involved in a battle between TE propagation and maintenance of germline DNA integrity. The TE integration machinery is constantly evolving to escape silencing. Rhino's rapid evolution may be due to interaction with TE integration proteins in order to promote transposition into the piRNA clusters and generate *trans*-silencing piRNAs. It has been hypothesized that piRNA clusters serve as hot spots for TE entrapment but the mechanistic details of this phenomenon and the role that HP1 variants may play have not been elucidated.

Our results also demonstrate no requirement for AGO2 in HP1 recruitment in somatic tissues. Furthermore, high-resolution genome-wide chromatin association profile of AGO2 in S2 and S3 *Drosophila* embryonic cell lines revealed that chromatin-associated AGO2 localizes in euchromatic but not repeat-rich regions. When compared to regions of the genome that produce Dicer-dependent endo-siRNAs, AGO2 genome-wide localization revealed no overlap. We also performed AGO2 ChIP-seq in OSC line which expresses AGO2 at high levels. Our preliminary results revealed that the majority of chromatin-associated AGO2 localizes in euchromatic regions. A more in depth computational analysis will reveal whether there is any AGO2 enrichment for repetitive sequences in ovarian somatic cells.

Alternative mechanisms for heterochromatin nucleation

Our results show that piRNA and endo-siRNA pathways do not recruit heterochromatin to the piRNA producing loci in *Drosophila* somatic tissues. We also show that HP1 recruitment to chromatin is independent of Piwi in the ovarian somatic cells. Studies have shown that heterochromatin formation can be achieved independently of RNAi. In fission yeast, the ATF/CREB stress-activated proteins nucleate heterochromatin at the silent mating-type locus in an RNAi-independent manner (Jia et al. 2004). Also, fission yeast telomere binding protein Taz1 can establish HP1 recruitment to telomeres independent of RNAi (Kanoh et al. 2005). In *Drosophila*, DDP1 dodeca-satellite binding protein, a single-stranded nucleic acid binding protein that associates with pericentric heterochromatin, has been suggested to contribute to

heterochromatin organization and function (Cortes and Azorin 2000). Furthermore, a recent study in mouse cells demonstrated that long nuclear non-coding transcripts that correspond to major satellite repeats at the pericentric heterochromatin associate with SUMO-modified HP1, a modification that promotes the initial targeting of HP1 to these regions (Maison et al. 2011). The same study did not detect any small dsRNA corresponding to major satellites suggesting that RNAi-mediated heterochromatin assembly, as pertains to *S. pombe*, may be a pathway that is not evolutionary conserved in metazoans.

RNA SILENCING AND ITS EFFECTS ON CHROMATIN INSULATORS

AGO2: a multifunctional protein

Our findings suggest two distinct roles for AGO2. In addition to its RNAi-dependent posttranscriptional silencing of TEs, we show that AGO2 functions in euchromatin in a Dicer-independent manner to promote or stabilize CTCF/CP190-dependent looping interactions that define transcriptional domains throughout the genome. Although, we show that the Slicer activity of AGO2 is not required for *Fab-8* activity and *CP190* mutants, which lose AGO2 association with chromatin, are still functional for RNAi, whether the functions of AGO2 in RNAi and chromatin insulator activity are completely separate remains to be elucidated.

Functionally multifaceted proteins are not a rare phenomenon in eukaryotic systems. AGO2 localizes to insulators, PREs, and promoters. One common feature of these sites is that they correspond to high nucleosome turnover, presumably needed to permit access to the macromolecular machinery that defines their activities (Mito et al. 2007; Deal et al. 2010). Intriguingly, *de novo* motif analysis of AGO2 binding sites resulted in a GA-rich motif similar to the GAF binding sequence. As multifunctional as AGO2, GAF is required for *Fab-7* and SF1 insulator activities (Belozerov et al. 2003; Schweinsberg et al. 2004), has been classified as a trxG protein (Farkas et al. 1994), and is associated with NELF-dependent paused polymerases (Lee et al. 2008). GAF associates with DNA directly through a zinc-finger DNA-binding domain, which is not present in AGO2 (reviewed in Adkins et al. 2006). However, both proteins harbor a polyglutamine-rich region of unknown function, which could mediate interactions with common proteins. Since AGO2 binding sites that do not overlap with GAF still contain the GA-rich motif, GAF binding does not appear to be prerequisite for AGO2 binding. Furthermore, in CP190 mutants, AGO2 but not GAF chromatin association is lost, suggesting that GAF and AGO2 recruitment are achieved by independent mechanisms.

Over twenty years of genetic and biochemical studies suggests that GAF promotes open chromatin, although, the mechanism is not well understood. We propose that once recruited to chromatin by CP190 and CTCF, AGO2 could serve to open chromatin and promote insulator activity by maintaining a nucleosome-free state and stabilizing or increasing the frequency of looping interactions between insulators, promoters, and PREs. It has been proposed that GAF may recruit to nucleosome remodeling factors to assist in local nucleosome turnover, which depending on gene

context, could lead to transcriptional activation or suppression (Xiao et al. 2001). Identification of additional proteins interacting with AGO2 such as components of the chromatin remodeling machinery may shed more light on the mechanism of AGO2 associated with chromatin.

Role for AGO2 in transcriptional regulation

The observation that more than half of AGO2 sites are associated with promoters suggests that AGO2 may regulate transcription directly. Given that AGO2 behaves similarly to a trxG protein, we anticipated that AGO2 may activate transcription. Transcription *per se*, however, may not be required for AGO2 chromatin association since TrxG proteins, depletion of which reduces transcription of *Abd-B* (Schwartz et al. 2010), are not required for AGO2 recruitment to chromatin. Interestingly, AGO2 depletion leads to either upregulation or downregulation of hundreds of transcripts (Rehwinkel et al. 2006). Upregulation would be consistent with direct posttranscriptional regulation via the RNAi pathway, but these effects could alternatively be due to direct transcriptional repression. Despite potentially complex relationships between AGO2 and gene expression levels, we found that the promoters of a statistically significant number of upregulated transcripts are bound by AGO2 (data not shown), suggesting that AGO2 may negatively regulate some of the transcripts to which it binds. It is unclear why some transcripts, whether bound by AGO2 or not, are downregulated upon AGO2 depletion. These gene expression changes could indicate an additional function for AGO2 in transcriptional activation that will be elucidated by future research. Alternatively, these

events could result from secondary effects of increased expression of RNAi-dependent or transcriptional targets of AGO2.

We also observed a considerable genome-wide overlap between AGO2, paused Pol II, and NELF, which is required for promoter proximal pausing of Pol II. Despite the overlap, AGO2-dependent gene expression does not correlate with NELF-dependent transcriptional effects. It was shown that depletion of NELF leads to both positive and negative effects on gene expression via two independent mechanisms (Muse et al. 2007; Gilchrist et al. 2008). On the one hand, NELF can attenuate gene expression by maintaining Pol II in a paused state in cases where transcription elongation is rate-limiting. On the other hand, NELF also promotes gene expression by stabilizing paused Pol II and maintaining an open chromatin structure at the promoter. Because promoters with paused polymerase, including *Abd-B* RB, have recently been shown to possess NELF-dependent enhancer blocking activity in transgene assays (Chopra et al. 2009), we compared gene expression profiles of AGO2 or NELF depletion in S2 cells but did not identify any statistically significant correlation (data not shown).

We also compared the AGO2-dependent transcription profile with that of CP190 or CTCF but did not detect any resemblance (data not shown). Like AGO2, extensive promoter association has been reported for the insulator proteins CP190, CTCF, Mod(mdg4)2.2, and BEAF-32 but not Su(Hw) (Bushey et al. 2009; Jiang et al. 2009; Smith et al. 2009; Negre et al. 2010). In CP190 and CTCF depleted S2 cells, gene expression changes are also observed in both positive and negative directions (Bartkuhn et al. 2009). Correlation between binding and change in gene expression was identified for both positively and negatively regulated transcripts, with CP190 more frequently

found at the promoter than CTCF. Our correlation analysis of gene expression changes among AGO2, CP190, or CTCF depleted cells indicates that there is no statistically significant correlation between AGO2 and either insulator protein. Currently, the significance of AGO2 or insulator protein promoter association is unclear.

A connection between two functions of AGO2

AGO2 may be an example of how the cell uses the same protein for two distinct purposes since the differential effects of the *AGO2* loss-of-function mutant and *AGO2* catalytic activity mutant on *Fab-8* insulator activity and *Pc* suppression suggests that AGO2 function in chromatin organization may be uncoupled from its role in RNAi. In order to understand whether the functions of AGO2 in RNAi and chromatin can be separated a systematic dissection of different aspects of AGO2 structure need to be performed. We show that AGO2, present in the nuclear pool, interacts with CP190 and CTCF in an RNA-independent manner. However, whether AGO2 can interact with siRNAs while associated with the insulator complex is not clear. Although, no small RNAs associated with the insulator proteins could be identified (data not shown), examining *AGO2* mutants defective in the PAZ and Mid domains, which would lack the ability to bind small RNAs, on *Fab-8* insulator activity and *Pc* suppression may address that question.

The regulation of RNAi and chromatin-related activities may be directed by AGO2 present in the cytoplasmic and nuclear fractions, respectively. An emerging theme from studies of post-translational modifications of mammalian AGO proteins is

that these modifications are crucial for the function of AGO proteins. One study identified hydroxylation of the human endogenous Ago2, a modification that appears to be crucial for stability of Ago2 and proper function of RISC activity (Qi et al. 2008). Another study identified phosphorylation of human Ago2, which contributes to Ago2 localization (Zeng et al. 2008). Also, a study in mice revealed that ubiquitylation of Ago2 affects its turnover (Rybak et al. 2009). It would be interesting to see whether the functions of AGO2 in RNAi and chromatin may be modulated by distinct post-translational modifications. Future research will examine *Drosophila* AGO2 modifications in the context of transcriptional regulation by two different mechanisms.

Another possibility is that distinct AGO2 isoforms may influence RNAi and chromatin organization in a different manner. The *AGO2* locus is predicted to give rise to two transcripts, *AGO2-RB* and *AGO2-RC* (FlyBase, 6_2010). The two isoforms are mostly identical except for distinct transcription start sites due to alternative splicing. The two transcripts encode two protein isoforms that differ only by 6-9 amino acids at their amino termini and have a 1208 amino acid region in common. A recent study in *Drosophila* reported that the *AGO2* locus produces an alternative transcript, which is predicted to encode a putative short isoform of AGO2 that is lacking the amino-terminal domain (Hain et al. 2010). However, the authors were not able to verify protein production from this transcript *in vivo*. Future studies will address whether different AGO2 isoforms vary in abundance in cytoplasmic and nuclear pools and whether they play distinct roles in these cellular compartments. Lastly, *AGO2* loss-of-function mutant and *AGO2* catalytic activity mutant may regulate transcription of different target genes.

Therefore, performing a gene expression profiling of these mutants in order to identify their targets may also be of interest.

AGO2 role in nuclear organization

Nuclear bodies such as insulator and PcG bodies have been proposed to be hubs of nuclear proteins interacting with regulatory elements in order to organize chromatin. Previously it was shown that changes in *gypsy* insulator activity in *piwi* and *aub* mutants corresponds to an improvement in the overall organization of insulator bodies (Lei and Corces, 2006), a structure that may be established in the early embryo and persist throughout development. Since no changes in insulator protein expression or recruitment to chromatin have been detected in these mutants, these RNA silencing pathways appear to specifically affect nuclear organization of the *gypsy* insulator. We examined *gypsy* insulator activity in *AGO2* mutants and found it to be similarly improved (data not shown). An increase of *gypsy* insulator activity in *AGO2* mutants corresponds to an improvement in the overall organization of insulator bodies. Moreover, localization of *gypsy* insulator proteins Su(Hw) and CP190 in *Drosophila* polytene chromosomes is not altered in *AGO2* mutants (data not shown). Interestingly, neither the ability of CTCF and CP190 to associate with chromatin or specifically with the BX-C nor localization of insulator bodies is affected in *AGO2* mutants. Therefore, it appears that *Fab-8* function is uncoupled from insulator body formation since *Fab-8* activity is not affected by RNA silencing mutants that alter insulator body localization. These observations suggest a differential effect of AGO2 on different classes of insulators. It would be interesting to

examine AGO2 effects on other classes of *Drosophila* insulators such BEAF-bound *scs'* insulator and ZW5-defined *scs* insulator. Given that the two insulators have been shown to interact by looping out the intervening DNA (Blanton et al. 2003), investigating the effects of AGO2 on this communication may also be worth of attention.

Unlike RNA silencing mutants, *ago1*, *piwi* and *aub*, which disrupt PcG-mediated long-range interactions between PREs, AGO2 does not appear to function in this phenomenon. Although, we did not determine whether AGO2 localizes to PcG bodies, a mutation in *AGO2* has no effect on *Fab-X*. Future studies will elucidate the role of AGO2 in *Mcp* PRE-mediated long-range interactions. Overall, however, our findings indicate that AGO2 function in general chromatin organization may not necessarily be connected to the formation of nuclear bodies.

CONCLUSION

Here, we demonstrate that RNA silencing affects gene expression at the level of higher order chromatin organization in *Drosophila melanogaster*. Specifically, our findings reveal that small RNA silencing pathways do not mediate heterochromatin formation in fly somatic tissues suggesting that alternative mechanisms may be involved. We also uncover a novel Dicer-independent function of AGO2 in CTCF/CP190 insulator activity.

The role of AGO2 in chromatin organization is intriguing but whether it is conserved in mammals remains an open question. Having a highly divergent sequence,

Drosophila AGO2 is very much distinct from AGO1, as well as yeast, plant, and mammalian Argonautes. Its carboxy-terminal region is conserved from plants to vertebrates, and the PAZ and PIWI domains are well characterized for RNA binding and cleavage. The AGO2 amino-terminal region is uniquely long in comparison to other Argonautes. This domain includes long stretches of glutamine-rich repeats, and its function is not well understood (Meyer et al. 2006). Future studies will address which domain of *Drosophila* AGO2 is responsible for its association with chromatin.

It remains an open question whether Argonaute function in chromatin is conserved in the mammalian CTCF activity as well as long-range chromosomal interactions mediated by this insulator protein. Additionally, two recent studies identified an endogenous PRE in the human homeotic genes HOXD12 and HOXD11 and a mouse PRE (Sing et al. 2009; Woo et al. 2010). Interestingly, the mouse PRE contains GAGA-binding motif. Future research will address whether AGO2 binding to PREs is a conserved phenomenon.

Our findings provide a first glimpse into understanding the complex interplay between the RNA silencing machinery and higher order chromatin structure to affect changes in gene expression. Future studies will decipher whether the two distinct functions of AGO2 are fully separable or are interconnected. Future research will also address the mechanisms that govern interactions between insulators, PREs and AGO2 necessary for correct nuclear organization and chromatin folding that are fundamental for proper gene expression during development and differentiation.

REFERENCES

Adkins, N.L., Hagerman, T.A., and Georgel, P. 2006. GAGA protein: a multi-faceted transcription factor. *Biochem Cell Biol* **84**(4): 559-567.

Amero, S.A., Elgin, S.C., and Beyer, A.L. 1991. A unique zinc finger protein is associated preferentially with active ecdysone-responsive loci in Drosophila. *Genes Dev* **5**(2): 188-200.

Bailey, T.L. and Elkan, C. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol* **2**: 28-36.

-. 1995. The value of prior knowledge in discovering motifs with MEME. *Proc Int Conf Intell Syst Mol Biol* **3**: 21-29.

Bantignies, F., Grimaud, C., Lavrov, S., Gabut, M., and Cavalli, G. 2003. Inheritance of Polycomb-dependent chromosomal interactions in Drosophila. *Genes Dev* **17**(19): 2406-2420.

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

Barges, S., Mihaly, J., Galloni, M., Hagstrom, K., Muller, 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* **127**(4): 779-790.

Bartkuhn, M., Straub, T., Herold, M., Herrmann, M., Rathke, C., Saumweber, H., Gilfillan, G.D., Becker, P.B., and Renkawitz, R. 2009. Active promoters and insulators are marked by the centrosomal protein 190. *EMBO J* **28**(7): 877-888.

Beisel, C., Buness, A., Roustan-Espinosa, I.M., Koch, B., Schmitt, S., Haas, S.A., Hild, M., Katsuyama, T., and Paro, R. 2007. Comparing active and repressed expression states of genes controlled by the Polycomb/Trithorax group proteins. *Proc Natl Acad Sci U S A* **104**(42): 16615-16620.

Belozerov, V.E., Majumder, P., Shen, P., and Cai, H.N. 2003. A novel boundary element may facilitate independent gene regulation in the Antennapedia complex of Drosophila. *EMBO J* **22**(12): 3113-3121.

Blanton, J., Gaszner, M., and Schedl, P. 2003. Protein:protein interactions and the pairing of boundary elements in vivo. *Genes Dev* **17**(5): 664-675.

Breiling, A., O'Neill, L.P., D'Eliseo, D., Turner, B.M., and Orlando, V. 2004. Epigenome changes in active and inactive polycomb-group-controlled regions. *EMBO Rep* **5**(10): 976-982.

Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G.J. 2007. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* **128**(6): 1089-1103.

Brower-Toland, B., Findley, S.D., Jiang, L., Liu, L., Yin, H., Dus, M., Zhou, P., Elgin, S.C., and Lin, H. 2007. Drosophila PIWI associates with chromatin and interacts directly with HP1a. *Genes Dev* **21**(18): 2300-2311.

Bushey, A.M., Dorman, E.R., and Corces, V.G. 2008. Chromatin insulators: regulatory mechanisms and epigenetic inheritance. *Mol Cell* **32**(1): 1-9.

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 Dev* **23**(11): 1338-1350.

Byrd, K. and Corces, V.G. 2003. Visualization of chromatin domains created by the gypsy insulator of Drosophila. *J Cell Biol* **162**(4): 565-574.

Chen, E.S., Zhang, K., Nicolas, E., Cam, H.P., Zofall, M., and Grewal, S.I. 2008. Cell cycle control of centromeric repeat transcription and heterochromatin assembly. *Nature* **451**(7179): 734-737.

Chopra, V.S., Cande, J., Hong, J.W., and Levine, M. 2009. Stalled Hox promoters as chromosomal boundaries. *Genes Dev* **23**(13): 1505-1509.

Chung, W.J., Okamura, K., Martin, R., and Lai, E.C. 2008. Endogenous RNA interference provides a somatic defense against Drosophila transposons. *Curr Biol* **18**(11): 795-802.

Claycomb, J.M., Batista, P.J., Pang, K.M., Gu, W., Vasale, J.J., van Wolfswinkel, J.C., Chaves, D.A., Shirayama, M., Mitani, S., Ketting, R.F., Conte, D., Jr., and Mello, C.C. 2009. The Argonaute CSR-1 and its 22G-RNA cofactors are required for holocentric chromosome segregation. *Cell* **139**(1): 123-134.

Cleard, F., Moshkin, Y., Karch, F., and Maeda, R.K. 2006. Probing long-distance regulatory interactions in the Drosophila melanogaster bithorax complex using Dam identification. *Nat Genet* **38**(8): 931-935.

Comet, I., Schuettengruber, B., Sexton, T., and Cavalli, G. 2011. A chromatin insulator driving three-dimensional Polycomb response element (PRE) contacts and Polycomb association with the chromatin fiber. *Proc Natl Acad Sci U S A* **108**(6): 2294-2299.

Cortes, A. and Azorin, F. 2000. DDP1, a heterochromatin-associated multi-KH-domain protein of Drosophila melanogaster, interacts specifically with centromeric satellite DNA sequences. *Mol Cell Biol* **20**(11): 3860-3869.

Cox, D.N., Chao, A., and Lin, H. 2000. piwi encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. *Development* **127**(3): 503-514.

Czech, B. and Hannon, G.J. 2011. Small RNA sorting: matchmaking for Argonautes. *Nat Rev Genet* **12**(1): 19-31.

Czech, B., Malone, C.D., Zhou, R., Stark, A., Schlingeheyde, C., Dus, M., Perrimon, N., Kellis, M., Wohlschlegel, J.A., Sachidanandam, R., Hannon, G.J., and Brennecke, J. 2008. An endogenous small interfering RNA pathway in Drosophila. *Nature* **453**(7196): 798-802.

Deal, R.B., Henikoff, J.G., and Henikoff, S. 2010. Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones. *Science* **328**(5982): 1161-1164.

Deshpande, G., Calhoun, G., and Schedl, P. 2005. Drosophila argonaute-2 is required early in embryogenesis for the assembly of centric/centromeric heterochromatin, nuclear division, nuclear migration, and germ-cell formation. *Genes Dev* **19**(14): 1680-1685.

Desset, S., Meignin, C., Dastugue, B., and Vaury, C. 2003. COM, a heterochromatic locus governing the control of independent endogenous retroviruses from Drosophila melanogaster. *Genetics* **164**(2): 501-509.

Dorer, D.R. and Henikoff, S. 1994. Expansions of transgene repeats cause heterochromatin formation and gene silencing in Drosophila. *Cell* **77**(7): 993-1002.

Eissenberg, J.C., James, T.C., Foster-Hartnett, D.M., Hartnett, T., Ngan, V., and Elgin, S.C. 1990. Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in Drosophila melanogaster. *Proc Natl Acad Sci U S A* **87**(24): 9923-9927.

Fagegaltier, D., Bouge, A.L., Berry, B., Poisot, E., Sismeiro, O., Coppee, J.Y., Theodore, L., Voinnet, O., and Antoniewski, C. 2009. The endogenous siRNA pathway is involved in heterochromatin formation in Drosophila. *Proc Natl Acad Sci U S A* **106**(50): 21258-21263.

Fanti, L., Dorer, D.R., Berloco, M., Henikoff, S., and Pimpinelli, S. 1998. Heterochromatin protein 1 binds transgene arrays. *Chromosoma* **107**(5): 286-292.

Farkas, G., Gausz, J., Galloni, M., Reuter, G., Gyurkovics, H., and Karch, F. 1994. The Trithorax-like gene encodes the Drosophila GAGA factor. *Nature* **371**(6500): 806-808.

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

Gallo, S.M., Gerrard, D.T., Miner, D., Simich, M., Des Soye, B., Bergman, C.M., and Halfon, M.S. 2011. REDfly v3.0: toward a comprehensive database of transcriptional regulatory elements in Drosophila. *Nucleic Acids Res* **39**(Database issue): D118-123.

Gerasimova, T.I., Byrd, K., and Corces, V.G. 2000. A chromatin insulator determines the nuclear localization of DNA. *Mol Cell* **6**(5): 1025-1035.

Gerasimova, T.I. and Corces, V.G. 1998. Polycomb and trithorax group proteins mediate the function of a chromatin insulator. *Cell* **92**(4): 511-521.

Gerasimova, T.I., Lei, E.P., Bushey, A.M., and Corces, V.G. 2007. Coordinated control of dCTCF and gypsy chromatin insulators in Drosophila. *Mol Cell* **28**(5): 761-772.

Ghildiyal, M., Seitz, H., Horwich, M.D., Li, C., Du, T., Lee, S., Xu, J., Kittler, E.L., Zapp, M.L., Weng, Z., and Zamore, P.D. 2008. Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. *Science* **320**(5879): 1077-1081.

Gilchrist, D.A., Nechaev, S., Lee, C., Ghosh, S.K., Collins, J.B., Li, L., Gilmour, D.S., and Adelman, K. 2008. NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. *Genes Dev* **22**(14): 1921-1933.

Giles, K.E., Ghirlando, R., and Felsenfeld, G. 2010. Maintenance of a constitutive heterochromatin domain in vertebrates by a Dicer-dependent mechanism. *Nat Cell Biol* **12**(1): 94-99; sup pp 91-96.

Grewal, S.I. and Elgin, S.C. 2007. Transcription and RNA interference in the formation of heterochromatin. *Nature* **447**(7143): 399-406.

Griffiths-Jones, S., Saini, H.K., van Dongen, S., and Enright, A.J. 2008. miRBase: tools for microRNA genomics. *Nucleic Acids Res* **36**(Database issue): D154-158.

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**(5): 957-971.

Guang, S., Bochner, A.F., Burkhart, K.B., Burton, N., Pavelec, D.M., and Kennedy, S. 2010. Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription. *Nature* **465**(7301): 1097-1101.

Guang, S., Bochner, A.F., Pavelec, D.M., Burkhart, K.B., Harding, S., Lachowiec, J., and Kennedy, S. 2008. An Argonaute transports siRNAs from the cytoplasm to the nucleus. *Science* **321**(5888): 537-541.

Gullerova, M. and Proudfoot, N.J. 2008. Cohesin complex promotes transcriptional termination between convergent genes in S. pombe. *Cell* **132**(6): 983-995.

Gunawardane, L.S., Saito, K., Nishida, K.M., Miyoshi, K., Kawamura, Y., Nagami, T., Siomi, H., and Siomi, M.C. 2007. A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. *Science* **315**(5818): 1587-1590.

Hagege, H., Klous, P., Braem, C., Splinter, E., Dekker, J., Cathala, G., de Laat, W., and Forne, T. 2007. Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nat Protoc* **2**(7): 1722-1733.

Hain, D., Bettencourt, B.R., Okamura, K., Csorba, T., Meyer, W., Jin, Z., Biggerstaff, J., Siomi, H., Hutvagner, G., Lai, E.C., Welte, M., and Muller, H.A. 2010. Natural variation of the amino-terminal glutamine-rich domain in Drosophila argonaute2 is not associated with developmental defects. *PLoS One* **5**(12): e15264.

Halic, M. and Moazed, D. 2010. Dicer-independent primal RNAs trigger RNAi and heterochromatin formation. *Cell* **140**(4): 504-516.

Hammond, S.M., Boettcher, S., Caudy, A.A., Kobayashi, R., and Hannon, G.J. 2001. Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* **293**(5532): 1146-1150.

Handoko, L., Xu, H., Li, G., Ngan, C.Y., Chew, E., Schnapp, M., Lee, C.W., Ye, C., Ping, J.L., Mulawadi, F., Wong, E., Sheng, J., Zhang, Y., Poh, T., Chan, C.S., Kunarso, G., Shahab, A., Bourque, G., Cacheux-Rataboul, V., Sung, W.K., Ruan, Y., and Wei, C.L. 2011. CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat Genet* **43**(7): 630-638.

Harris, A.N. and Macdonald, P.M. 2001. Aubergine encodes a Drosophila polar granule component required for pole cell formation and related to eIF2C. *Development* **128**(14): 2823-2832.

Haynes, K.A., Caudy, A.A., Collins, L., and Elgin, S.C. 2006. Element 1360 and RNAi components contribute to HP1-dependent silencing of a pericentric reporter. *Curr Biol* **16**(22): 2222-2227.

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 Genet* **3**(7): e112.

Hutvagner, G. and Simard, M.J. 2008. Argonaute proteins: key players in RNA silencing. *Nat Rev Mol Cell Biol* **9**(1): 22-32.

Ishizuka, A., Siomi, M.C., and Siomi, H. 2002. A Drosophila fragile X protein interacts with components of RNAi and ribosomal proteins. *Genes Dev* **16**(19): 2497-2508.

Janowski, B.A., Huffman, K.E., Schwartz, J.C., Ram, R., Nordsell, R., Shames, D.S., Minna, J.D., and Corey, D.R. 2006. Involvement of AGO1 and AGO2 in mammalian transcriptional silencing. *Nat Struct Mol Biol* **13**(9): 787-792.

Jia, S., Noma, K., and Grewal, S.I. 2004. RNAi-independent heterochromatin nucleation by the stress-activated ATF/CREB family proteins. *Science* **304**(5679): 1971-1976.

Jiang, F., Ye, X., Liu, X., Fincher, L., McKearin, D., and Liu, Q. 2005. Dicer-1 and R3D1-L catalyze microRNA maturation in Drosophila. *Genes Dev* **19**(14): 1674-1679.

Jiang, N., Emberly, E., Cuvier, O., and Hart, C.M. 2009. Genome-wide mapping of boundary element-associated factor (BEAF) binding sites in Drosophila melanogaster links BEAF to transcription. *Mol Cell Biol* **29**(13): 3556-3568.

Kanoh, J., Sadaie, M., Urano, T., and Ishikawa, F. 2005. Telomere binding protein Taz1 establishes Swi6 heterochromatin independently of RNAi at telomeres. *Curr Biol* **15**(20): 1808-1819.

Kawamura, Y., Saito, K., Kin, T., Ono, Y., Asai, K., Sunohara, T., Okada, T.N., Siomi, M.C., and Siomi, H. 2008. Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. *Nature* **453**(7196): 793-797.

Kim, D.H., Villeneuve, L.M., Morris, K.V., and Rossi, J.J. 2006. Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. *Nat Struct Mol Biol* **13**(9): 793-797.

Kim, J., Chu, J., Shen, X., Wang, J., and Orkin, S.H. 2008. An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* **132**(6): 1049-1061.

Kim, K., Lee, Y.S., and Carthew, R.W. 2007. Conversion of pre-RISC to holo-RISC by Ago2 during assembly of RNAi complexes. *RNA* **13**(1): 22-29.

Klattenhoff, C., Xi, H., Li, C., Lee, S., Xu, J., Khurana, J.S., Zhang, F., Schultz, N., Koppetsch, B.S., Nowosielska, A., Seitz, H., Zamore, P.D., Weng, Z., and Theurkauf, W.E. 2009. The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell* **138**(6): 1137-1149.

Klenov, M.S., Lavrov, S.A., Stolyarenko, A.D., Ryazansky, S.S., Aravin, A.A., Tuschl, T., and Gvozdev, V.A. 2007. Repeat-associated siRNAs cause chromatin silencing of retrotransposons in the Drosophila melanogaster germline. *Nucleic Acids Res* **35**(16): 5430-5438.

Kloc, A. and Martienssen, R. 2008. RNAi, heterochromatin and the cell cycle. *Trends Genet* **24**(10): 511-517.

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

Kornberg, R.D. 1974. Chromatin structure: a repeating unit of histones and DNA. *Science* **184**(139): 868-871.

Kyrchanova, O., Toshchakov, S., Podstreshnaya, Y., Parshikov, A., and Georgiev, P. 2008. Functional interaction between the Fab-7 and Fab-8 boundaries and the upstream promoter region in the Drosophila Abd-B gene. *Mol Cell Biol* **28**(12): 4188-4195.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**(3): R25.

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

Lau, N.C., Robine, N., Martin, R., Chung, W.J., Niki, Y., Berezikov, E., and Lai, E.C. 2009. Abundant primary piRNAs, endo-siRNAs, and microRNAs in a Drosophila ovary cell line. *Genome Res* **19**(10): 1776-1785.

Lee, C., Li, X., Hechmer, A., Eisen, M., Biggin, M.D., Venters, B.J., Jiang, C., Li, J., Pugh, B.F., and Gilmour, D.S. 2008. NELF and GAGA factor are linked to promoter-proximal pausing at many genes in Drosophila. *Mol Cell Biol* **28**(10): 3290-3300.

Lee, Y.S., Nakahara, K., Pham, J.W., Kim, K., He, Z., Sontheimer, E.J., and Carthew, R.W. 2004. Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell* **117**(1): 69-81.

Lei, E.P. and Corces, V.G. 2006. RNA interference machinery influences the nuclear organization of a chromatin insulator. *Nat Genet* **38**(8): 936-941.

Li, C., Vagin, V.V., Lee, S., Xu, J., Ma, S., Xi, H., Seitz, H., Horwich, M.D., Syrzycka, M., Honda, B.M., Kittler, E.L., Zapp, M.L., Klattenhoff, C., Schulz, N., Theurkauf, W.E., Weng, Z., and Zamore, P.D. 2009a. Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. *Cell* **137**(3): 509-521.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. 2009b. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**(16): 2078-2079.

Li, H.B., Muller, M., Bahechar, I.A., Kyrchanova, O., Ohno, K., Georgiev, P., and Pirrotta, V. 2011. Insulators, not Polycomb response elements, are required for long-range interactions between Polycomb targets in Drosophila melanogaster. *Mol Cell Biol* **31**(4): 616-625.

Li, L. 2009. GADEM: a genetic algorithm guided formation of spaced dyads coupled with an EM algorithm for motif discovery. *J Comput Biol* **16**(2): 317-329.

Ma, J.B., Ye, K., and Patel, D.J. 2004. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* **429**(6989): 318-322.

Ma, J.B., Yuan, Y.R., Meister, G., Pei, Y., Tuschl, T., and Patel, D.J. 2005. Structural basis for 5'-end-specific recognition of guide RNA by the A. fulgidus Piwi protein. *Nature* **434**(7033): 666-670.

Maeda, R.K. and Karch, F. 2009. The bithorax complex of Drosophila an exceptional Hox cluster. *Curr Top Dev Biol* **88**: 1-33.

Maison, C., Bailly, D., Peters, A.H., Quivy, J.P., Roche, D., Taddei, A., Lachner, M., Jenuwein, T., and Almouzni, G. 2002. Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nat Genet* **30**(3): 329-334.

Maison, C., Bailly, D., Roche, D., Montes de Oca, R., Probst, A.V., Vassias, I., Dingli, F., Lombard, B., Loew, D., Quivy, J.P., and Almouzni, G. 2011. SUMOylation promotes de novo targeting of HP1alpha to pericentric heterochromatin. *Nat Genet* **43**(3): 220-227.

Malone, C.D., Brennecke, J., Dus, M., Stark, A., McCombie, W.R., Sachidanandam, R., and Hannon, G.J. 2009. Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. *Cell* **137**(3): 522-535.

Mason, J.M., Frydrychova, R.C., and Biessmann, H. 2008. Drosophila telomeres: an exception providing new insights. *Bioessays* **30**(1): 25-37.

Merkenschlager, M. 2010. Cohesin: a global player in chromosome biology with local ties to gene regulation. *Curr Opin Genet Dev* **20**(5): 555-561.

Meyer, W.J., Schreiber, S., Guo, Y., Volkmann, T., Welte, M.A., and Muller, H.A. 2006. Overlapping functions of argonaute proteins in patterning and morphogenesis of Drosophila embryos. *PLoS Genet* **2**(8): e134.

Mito, Y., Henikoff, J.G., and Henikoff, S. 2007. Histone replacement marks the boundaries of cis-regulatory domains. *Science* **315**(5817): 1408-1411.

Miyoshi, K., Tsukumo, H., Nagami, T., Siomi, H., and Siomi, M.C. 2005. Slicer function of Drosophila Argonautes and its involvement in RISC formation. *Genes Dev* **19**(23): 2837-2848.

Mochizuki, K. 2010. RNA-directed epigenetic regulation of DNA rearrangements. *Essays Biochem* **48**(1): 89-100.

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. *EMBO J* **26**(19): 4203-4214.

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

Mongelard, F., Labrador, M., Baxter, E.M., Gerasimova, T.I., and Corces, V.G. 2002. Trans-splicing as a novel mechanism to explain interallelic complementation in Drosophila. *Genetics* **160**(4): 1481-1487.

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., and Lobanenkov, V. 2005. CTCF is conserved from Drosophila to humans and confers enhancer blocking of the Fab-8 insulator. *EMBO Rep* **6**(2): 165-170.

Moshkovich, N. and Lei, E.P. 2010. HP1 recruitment in the absence of argonaute proteins in Drosophila. *PLoS Genet* **6**(3): e1000880.

Moshkovich, N., Nisha, P., Boyle, P.J., Thompson, B.A., Dale, R.K., and Lei, E.P. 2011. RNAi-independent role for Argonaute2 in CTCF/CP190 chromatin insulator function. *Genes Dev* **25**(16).

Motamedi, M.R., Verdel, A., Colmenares, S.U., Gerber, S.A., Gygi, S.P., and Moazed, D. 2004. Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. *Cell* **119**(6): 789-802.

Muller, J. and Kassis, J.A. 2006. Polycomb response elements and targeting of Polycomb group proteins in Drosophila. *Curr Opin Genet Dev* **16**(5): 476-484.

Muse, G.W., Gilchrist, D.A., Nechaev, S., Shah, R., Parker, J.S., Grissom, S.F., Zeitlinger, J., and Adelman, K. 2007. RNA polymerase is poised for activation across the genome. *Nat Genet* **39**(12): 1507-1511.

Negre, 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 Genet* **6**(1): e1000814.

Nishida, K.M., Saito, K., Mori, T., Kawamura, Y., Nagami-Okada, T., Inagaki, S., Siomi, H., and Siomi, M.C. 2007. Gene silencing mechanisms mediated by Aubergine piRNA complexes in Drosophila male gonad. *RNA* **13**(11): 1911-1922.

Noma, K., Sugiyama, T., Cam, H., Verdel, A., Zofall, M., Jia, S., Moazed, D., and Grewal, S.I. 2004. RITS acts in cis to promote RNA interference-mediated transcriptional and post-transcriptional silencing. *Nat Genet* **36**(11): 1174-1180.

Okamura, K., Balla, S., Martin, R., Liu, N., and Lai, E.C. 2008. Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in Drosophila melanogaster. *Nat Struct Mol Biol* **15**(9): 998.

Okamura, K., Ishizuka, A., Siomi, H., and Siomi, M.C. 2004. Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. *Genes Dev* **18**(14): 1655-1666.

Okamura, K. and Lai, E.C. 2008. Endogenous small interfering RNAs in animals. *Nat Rev Mol Cell Biol* **9**(9): 673-678.

Oktaba, K., Gutierrez, L., Gagneur, J., Girardot, C., Sengupta, A.K., Furlong, E.E., and Muller, J. 2008. Dynamic regulation by polycomb group protein complexes controls pattern formation and the cell cycle in Drosophila. *Dev Cell* **15**(6): 877-889.

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. *Mol Cell* **16**(5): 737-748.

Pal-Bhadra, M., Bhadra, U., and Birchler, J.A. 2002. RNAi related mechanisms affect both transcriptional and posttranscriptional transgene silencing in Drosophila. *Mol Cell* **9**(2): 315-327.

Pal-Bhadra, M., Leibovitch, B.A., Gandhi, S.G., Rao, M., Bhadra, U., Birchler, J.A., and Elgin, S.C. 2004. Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. *Science* **303**(5658): 669-672.

Parnell, T.J., Viering, M.M., Skjesol, A., Helou, C., Kuhn, E.J., and Geyer, P.K. 2003. An endogenous suppressor of hairy-wing insulator separates regulatory domains in Drosophila. *Proc Natl Acad Sci U S A* **100**(23): 13436-13441.

Pavesi, G. and Pesole, G. 2006. Using Weeder for the discovery of conserved transcription factor binding sites. *Curr Protoc Bioinformatics* **Chapter 2**: Unit 2 11.

Peng, J.C. and Karpen, G.H. 2007. H3K9 methylation and RNA interference regulate nucleolar organization and repeated DNA stability. *Nat Cell Biol* **9**(1): 25-35.

Phillips, J.E. and Corces, V.G. 2009. CTCF: master weaver of the genome. *Cell* **137**(7): 1194-1211.

Prud'homme, N., Gans, M., Masson, M., Terzian, C., and Bucheton, A. 1995. Flamenco, a gene controlling the gypsy retrovirus of Drosophila melanogaster. *Genetics* **139**(2): 697-711.

Qi, H.H., Ongusaha, P.P., Myllyharju, J., Cheng, D., Pakkanen, O., Shi, Y., Lee, S.W., and Peng, J. 2008. Prolyl 4-hydroxylation regulates Argonaute 2 stability. *Nature* **455**(7211): 421-424.

Qi, Y., He, X., Wang, X.J., Kohany, O., Jurka, J., and Hannon, G.J. 2006. Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. *Nature* **443**(7114): 1008-1012.

Quinlan, A.R. and Hall, I.M. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**(6): 841-842.

Rehwinkel, J., Natalin, P., Stark, A., Brennecke, J., Cohen, S.M., and Izaurralde, E. 2006. Genome-wide analysis of mRNAs regulated by Drosha and Argonaute proteins in Drosophila melanogaster. *Mol Cell Biol* **26**(8): 2965-2975.

Richards, E.J. and Elgin, S.C. 2002. Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. *Cell* **108**(4): 489-500.

Robert, V., Prud'homme, N., Kim, A., Bucheton, A., and Pelisson, A. 2001. Characterization of the flamenco region of the Drosophila melanogaster genome. *Genetics* **158**(2): 701-713.

Roseman, R.R., Pirrotta, V., and Geyer, P.K. 1993. The su(Hw) protein insulates expression of the Drosophila melanogaster white gene from chromosomal position-effects. *EMBO J* **12**(2): 435-442.

Roy, S., Ernst, J., Kharchenko, P.V., Kheradpour, P., Negre, N., Eaton, M.L., Landolin, J.M., Bristow, C.A., Ma, L., Lin, M.F., Washietl, S., Arshinoff, B.I., Ay, F., Meyer, P.E., Robine, N., Washington, N.L., Di Stefano, L., Berezikov, E., Brown, C.D., Candeias, R., Carlson, J.W., Carr, A., Jungreis, I., Marbach, D., Sealfon, R., Tolstorukov, M.Y., Will, S., Alekseyenko, A.A., Artieri, C., Booth, B.W., Brooks, A.N., Dai, Q., Davis, C.A., Duff, M.O., Feng, X., Gorchakov, A.A., Gu, T., Henikoff, J.G., Kapranov, P., Li, R., Macalpine, H.K., Malone, J., Minoda, A., Nordman, J., Okamura, K., Perry, M., Powell, S.K., Riddle, N.C., Sakai, A., Samsonova, A., Sandler, J.E., Schwartz, Y.B., Sher, N., Spokony, R., Sturgill, D., van Baren, M., Wan, K.H., Yang, L., Yu, C., Feingold, E., Good, P., Guyer, M., Lowdon, R., Ahmad, K., Andrews, J., Berger, B., Brenner, S.E., Brent, M.R., Cherbas, L., Elgin, S.C., Gingeras, T.R., Grossman, R., Hoskins, R.A., Kaufman, T.C., Kent, W., Kuroda, M.I., Orr-Weaver, T., Perrimon, N., Pirrotta, V., Posakony, J.W., Ren, B., Russell, S., Cherbas, P., Graveley, B.R., Lewis, S., Micklem, G., Oliver, B., Park, P.J., Celniker, S.E., Henikoff, S., Karpen, G.H., Lai, E.C., Macalpine, D.M., Stein, L.D., White, K.P., and Kellis, M. 2010. Identification of Functional Elements and Regulatory Circuits by Drosophila modENCODE. *Science*.

Rybak, A., Fuchs, H., Hadian, K., Smirnova, L., Wulczyn, E.A., Michel, G., Nitsch, R., Krappmann, D., and Wulczyn, F.G. 2009. The let-7 target gene mouse lin-41 is a stem cell specific E3 ubiquitin ligase for the miRNA pathway protein Ago2. *Nat Cell Biol* **11**(12): 1411-1420.

Saito, K., Inagaki, S., Mituyama, T., Kawamura, Y., Ono, Y., Sakota, E., Kotani, H., Asai, K., Siomi, H., and Siomi, M.C. 2009. A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila. *Nature* **461**(7268): 1296-1299.

Saito, K., Nishida, K.M., Mori, T., Kawamura, Y., Miyoshi, K., Nagami, T., Siomi, H., and Siomi, M.C. 2006. Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. *Genes Dev* **20**(16): 2214-2222.

Schuettengruber, B., Ganapathi, M., Leblanc, B., Portoso, M., Jaschek, R., Tolhuis, B., van Lohuizen, M., Tanay, A., and Cavalli, G. 2009. Functional anatomy of polycomb and trithorax chromatin landscapes in Drosophila embryos. *PLoS Biol* **7**(1): e13.

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

Schwartz, Y.B., Kahn, T.G., Stenberg, P., Ohno, K., Bourgon, R., and Pirrotta, V. 2010. Alternative epigenetic chromatin states of polycomb target genes. *PLoS Genet* **6**(1): e1000805.

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**(3): 1371-1384.

Selva, E.M. and Stronach, B.E. 2007. Germline clone analysis for maternally acting Drosophila hedgehog components. *Methods Mol Biol* **397**: 129-144.

Shin, H., Liu, T., Manrai, A.K., and Liu, X.S. 2009. CEAS: cis-regulatory element annotation system. *Bioinformatics* **25**(19): 2605-2606.

Simon, J.A. and Kingston, R.E. 2009. Mechanisms of polycomb gene silencing: knowns and unknowns. *Nat Rev Mol Cell Biol* **10**(10): 697-708.

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

Smith, S.T., Wickramasinghe, P., Olson, A., Loukinov, D., Lin, L., Deng, J., Xiong, Y., Rux, J., Sachidanandam, R., Sun, H., Lobanenkov, V., and Zhou, J. 2009.

Genome wide ChIP-chip analyses reveal important roles for CTCF in Drosophila genome organization. *Dev Biol* **328**(2): 518-528.

Song, J.J., Liu, J., Tolia, N.H., Schneiderman, J., Smith, S.K., Martienssen, R.A., Hannon, G.J., and Joshua-Tor, L. 2003. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat Struct Biol* **10**(12): 1026-1032.

Song, J.J., Smith, S.K., Hannon, G.J., and Joshua-Tor, L. 2004. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **305**(5689): 1434-1437.

Stuurman, N., Sasse, B., and Fisher, P.A. 1996. Intermediate filament protein polymerization: molecular analysis of Drosophila nuclear lamin head-to-tail binding. *J Struct Biol* **117**(1): 1-15.

Sun, F.L., Haynes, K., Simpson, C.L., Lee, S.D., Collins, L., Wuller, J., Eissenberg, J.C., and Elgin, S.C. 2004. cis-Acting determinants of heterochromatin formation on Drosophila melanogaster chromosome four. *Mol Cell Biol* **24**(18): 8210-8220.

Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G., and Reuter, G. 1994. The protein encoded by the Drosophila position-effect variegation suppressor gene Su(var)3-9 combines domains of antagonistic regulators of homeotic gene complexes. *EMBO J* **13**(16): 3822-3831.

Vagin, V.V., Sigova, A., Li, C., Seitz, H., Gvozdev, V., and Zamore, P.D. 2006. A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* **313**(5785): 320-324.

van Wolfswinkel, J.C., Claycomb, J.M., Batista, P.J., Mello, C.C., Berezikov, E., and Ketting, R.F. 2009. CDE-1 affects chromosome segregation through uridylation of CSR-1-bound siRNAs. *Cell* **139**(1): 135-148.

Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I., and Martienssen, R.A. 2002. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297**(5588): 1833-1837.

Wallace, J.A. and Felsenfeld, G. 2007. We gather together: insulators and genome organization. *Curr Opin Genet Dev* **17**(5): 400-407.

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

Wang, L. and Ligoxygakis, P. 2006. Pathogen recognition and signalling in the Drosophila innate immune response. *Immunobiology* **211**(4): 251-261.

Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T., Surani, M.A., Sakaki, Y., and Sasaki, H. 2008. Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* **453**(7194): 539-543.

Weinmann, L., Hock, J., Ivacevic, T., Ohrt, T., Mutze, J., Schwille, P., Kremmer, E., Benes, V., Urlaub, H., and Meister, G. 2009. Importin 8 is a gene silencing factor that targets argonaute proteins to distinct mRNAs. *Cell* **136**(3): 496-507.

Williams, R.W. and Rubin, G.M. 2002. ARGONAUTE1 is required for efficient RNA interference in Drosophila embryos. *Proc Natl Acad Sci U S A* **99**(10): 6889-6894.

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**(1): 99-110.

Xiao, H., Sandaltzopoulos, R., Wang, H.M., Hamiche, A., Ranallo, R., Lee, K.M., Fu, D., and Wu, C. 2001. Dual functions of largest NURF subunit NURF301 in nucleosome sliding and transcription factor interactions. *Mol Cell* **8**(3): 531-543.

Xu, K., Bogert, B.A., Li, W., Su, K., Lee, A., and Gao, F.B. 2004. The fragile X-related gene affects the crawling behavior of Drosophila larvae by regulating the mRNA level of the DEG/ENaC protein pickpocket1. *Curr Biol* **14**(12): 1025-1034.

Yin, H. and Lin, H. 2007. An epigenetic activation role of Piwi and a Piwi-associated piRNA in Drosophila melanogaster. *Nature* **450**(7167): 304-308.

Zeng, Y., Sankala, H., Zhang, X., and Graves, P.R. 2008. Phosphorylation of Argonaute 2 at serine-387 facilitates its localization to processing bodies. *Biochem J* **413**(3): 429-436.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nussbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**(9): R137.

Zofall, M., Fischer, T., Zhang, K., Zhou, M., Cui, B., Veenstra, T.D., and Grewal, S.I. 2009. Histone H2A.Z cooperates with RNAi and heterochromatin factors to suppress antisense RNAs. *Nature* **461**(7262): 419-422.

Zofall, M. and Grewal, S.I. 2006. Swi6/HP1 recruits a JmjC domain protein to facilitate transcription of heterochromatic repeats. *Mol Cell* **22**(5): 681-692.