Investigations of HP1 and insulator partner protein 1 (HIPP1)

Steve Ehren Glenn

University of Iowa

Copyright © 2018 Steve Ehren Glenn

This dissertation is available at Iowa Research Online: https://ir.uiowa.edu/etd/6583

Recommended Citation
https://doi.org/10.17077/etd.wwur-gtke

Follow this and additional works at: https://ir.uiowa.edu/etd

Part of the Cell Biology Commons
INVESTIGATIONS OF HP1 AND INSULATOR PARTNER PROTEIN 1 (HIPPI)

by

Steve Ehren Glenn

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Molecular and Cellular Biology in the Graduate College of The University of Iowa

December 2018

Thesis Supervisor: Professor Pamela K Geyer
Drosophila HP1 and Insulator Partner Protein 1 (HIPP1) is the homologue of the human co-repressor Chromodomain Y family of proteins that repress neuronal gene expression in mammals. HIPP1 was identified by its extensive co-localization with Heterochromatin Protein 1a (HP1a) in heterochromatic regions of the genome and insulator binding proteins in euchromatic regions. The majority of HIPP1 binding to euchromatin is at binding sites for Drosophila Suppressor of Hairy-wing [Su(Hw)]. Su(Hw) is a zinc finger DNA binding protein that functions as an insulator, activator, and repressor. Transcriptional regulation by Su(Hw) is essential in the ovary and testis, where Su(Hw) functions primarily as a repressor of neuronal genes. However, the mechanism of Su(Hw) dependent repression is not clear. The focus of my thesis work has been defining the role of HIPP1 in development and its contribution to Su(Hw) function and heterochromatin formation. As part of this work, CRISPR was used to generate multiple Hipp1 null alleles and a tagged derivative of the endogenous gene (Hipp1\textsuperscript{GFP}). Hipp1 null flies were found to be viable. Study of HIPP1 expression revealed it is present in most tissues and restricted to the nucleus. HIPP1 showed limited colocalization with HP1a, and tests of repression of transgenes in heterochromatin suggested that HIPP1 is not required for heterochromatin formation. Investigations of HIPP1 binding revealed that Su(Hw) is responsible for the majority of HIPP1 recruitment to euchromatin. Despite this, HIPP1 was found to be dispensable for the transcriptional and insulator functions of Su(Hw) as well as for female and male fertility. These data indicate that HIPP1 is not a critical Su(Hw) cofactor. Further studies are needed to clarify the role of HIPP1 in Drosophila development.
Most cells in the human body have the same genes. The diversity of cell types and tissues arises from controlling the level of expression of genes in different cells. This control is achieved in part through proteins called transcription factors that bind DNA within the genome and either increase or decrease the expression of nearby genes. Many transcription factors have gene specific effects, increasing expression of some genes and decreasing expression of others. The goal of this thesis work is to better understand the underlying principles that determine how transcription factors function to regulate the expression of genes. To accomplish this goal, I studied the fruit fly protein Suppressor of Hairy-wing [Su(Hw)] as a model transcription factor. The primary function of Su(Hw) is to prevent genes that are normally expressed in the brain from being expressed in other tissues. Su(Hw) was shown to interact with an uncharacterized protein, HP1 and Insulator Partner Protein 1 (HIPP1). HIPP1 is the fly version of a human protein, CDYL, that also regulates the expression of neuronal genes. The conservation of HIPP1 from humans to flies suggests that HIPP1 has an important function. This thesis work tested the requirement of HIPP1 in Su(Hw) function and fruit fly development. These studies revealed that Su(Hw) recruits HIPP1 to DNA and that HIPP1 is not required for any identified functions of Su(Hw) or fly survival. By ruling out many predicted functions of HIPP1, these studies may direct future investigations of HIPP1 and transcription factor function.
## TABLE OF CONTENTS

| LIST OF TABLES | ..................................................................................................................... | v |
| LIST OF FIGURES | ....................................................................................................................... | vi |
| CHAPTER 1: INTRODUCTION | .................................................................................................................. | 1 |
| The role of insulators in genome architecture | ................................................................................................. | 3 |
| Mammalian CTCF as an insulator and regulator of genome organization | .................................................. | 6 |
| The role of Drosophila insulators in genome organization and transcription | ........................................ | 8 |
| The role of Su(Hw) in gene expression | ........................................................................................................... | 10 |
| Thesis outline | .................................................................................................................. | 18 |
| CHAPTER 2: MATERIALS AND METHODS | .................................................................................................................. | 23 |
| Drosophila stocks and culture conditions | ......................................................................................................... | 23 |
| Generation of Hipp1 alleles | .................................................................................................................. | 23 |
| Generation of GFP-Rbp9 | .................................................................................................................... | 25 |
| Generation of HIPPI antibody | .................................................................................................................. | 27 |
| Western blot analysis | .................................................................................................................. | 29 |
| Immunohistochemical analysis | .................................................................................................................. | 30 |
| Fertility assays | .................................................................................................................. | 31 |
| Position Effect Variegation (PEV) | .................................................................................................................. | 31 |
| Quantitative PCR analysis of gene expression | ............................................................................................... | 32 |
| Chromatin Immunoprecipitation (ChIP) | .................................................................................................................. | 33 |
| Rapid Amplification of cDNA Ends (RACE) procedure | .............................................................................................. | 34 |
| Northern blotting | .................................................................................................................. | 35 |
| Polytene staining protocol | .................................................................................................................. | 35 |
| CHAPTER 3: INVESTIGATIONS OF HP1 AND INSULATOR PARTNER PROTEIN 1, HIPPI | .................................................................................................................. | 48 |
# LIST OF TABLES

Table 2. 1 CRISPR generated *Hipp1* mutants ................................................................. 43
Table 2. 2 Primers used for qPCR analysis of ovary gene expression .................................. 44
Table 2. 3 Primers used for ChIP-qPCR .............................................................................. 46
Table 3. 1 Primers for verification of *Hipp1* deletion alleles ............................................... 90
Table 3. 2 Percent viability of *Hipp1* mutants .................................................................... 91
Table 3. 3 Effects of *Hipp1* mutants on mid-oogenesis egg chamber death ....................... 92
Table 3. 4 Summary of insulator proteins bound to ChIP sites ............................................ 93
### LIST OF FIGURES

| Figure 1. 1 Genome organization and localization of Drosophila insulator proteins | 21 |
| Figure 1. 2 Su(Hw) binding site logo | 22 |
| Figure 2. 1 Generation of Hipp1<sub>GFP</sub> | 38 |
| Figure 2. 2 Generation of Rbp<sub>GFP</sub> | 40 |
| Figure 2. 3 Generation and validation of peptide antibody | 42 |
| Figure 3. 1 HIPPI is the major Su(Hw) cofactor | 65 |
| Figure 3. 2 HIPPI is the fly homologue of human CDYL | 66 |
| Figure 3. 3 Structure of the Hipp1 locus | 67 |
| Figure 3. 4 Strategy for analysis of CRISPR generated large deletion Hipp1 alleles | 69 |
| Figure 3. 5 Western blot analysis of proteins produced from generated Hipp1 alleles | 70 |
| Figure 3. 6 HIPPI is a nuclear protein expressed in most cell types | 72 |
| Figure 3. 7 HIPPI localizes to foci containing CP190 and Mod67.2 in late stage oocyte nuclei | 73 |
| Figure 3. 8 Loss of Su(Hw) does not alter HIPPI expression pattern in testes | 74 |
| Figure 3. 9 HIPPI has a novel transcript in the testis | 76 |
| Figure 3. 10 HIPPI and HP1a show limited co-localization in interphase cells | 77 |
| Figure 3. 11 Hipp1 is not a modifier of PEV | 79 |
| Figure 3. 12 Hipp1 mutant females are fertile and do not phenocopy Su(Hw) loss | 81 |
| Figure 3. 13 HIPPI is not essential for spermatogenesis | 83 |
| Figure 3. 14 HIPPI is not required for regulation of histone crotonylation at SBSs | 85 |
| Figure 3. 15 HIPPI occupancy at euchromatic sites depends on Su(Hw) | 87 |
| Figure 3. 16 Su(Hw) and HIPPI binding correlate at most sites | 88 |
| Figure 3. 17 HIPPI is not required for Su(Hw) insulator function at gypsy | 89 |
| Figure 4. 1 Alignment of CDYL1b to Drosophila Crotonase-like domain family members | 110 |
Figure 4. 2 HIPP1 has pockets of strong conservation outside the crotonase-like domain ...... 111

Figure 4. 3 HIPP1 associates with chromosomes in the absence of Su(Hw) ......................... 112
CHAPTER 1: INTRODUCTION

Development requires the coordinated regulation of thousands of genes within the genome. This regulation is achieved primarily through the action of transcription factors, proteins that bind DNA and act through multiple mechanisms to ultimately influence PolII recruitment, retention, and productive elongation. Many transcription factors bind specific DNA sequences that serve as regulatory elements to modulate expression of genes. DNA sequence elements that lead to gene activation are known as enhancers, whereas those that lead to gene repression are called silencers. Enhancers can interact with and activate promoters 100s of kilobases away and these contacts are often dynamic, changing as genes become active or inactive during development. (DENG et al. 2014; GHAVI-HELM et al. 2014; SCHOENFELDER et al. 2015; BONEV et al. 2017). Further, promoters often contact multiple enhancers, not strictly based on distance, with promoters often skipping over some enhancers to preferentially interact with others (SCHOENFELDER et al. 2015). Similarly, silencers are able to influence transcription at long distances (LANZUOLO et al. 2007). The potential range of enhancer and silencer action often includes genes that are not regulated by these elements. How promiscuous regulatory elements are constrained to give rise to the observed patterns of gene expression is a long-standing field of research. This field has been advanced by techniques that look at chromatin organization. The technique of chromosome conformation capture (3C) maps contacts between genomic sequences that are distant in the linear genome to get a representation of the three-dimensional organization of chromatin within the nucleus (DEKKER et al. 2002; EAGEN 2018). These studies have revealed underlying principles of genome organization that contribute to transcriptional control.

Genome organization and transcription are tightly linked. Chromatin is broadly separated into two compartments, A and B, that are associated with high and low transcriptional activity.
respectively (LIEBERMAN-AIDEN et al. 2009). Within compartments, chromatin is further divided into domains in which interactions between sequences within the domain occur more frequently than with sequences outside of the domain. These are referred to as topologically associating domains (TADs). The transition point of one TAD to another is referred to as a TAD boundary. A TAD boundary corresponds with a switch in the direction of physical contacts at the transition point, with sequences on each side of the boundary preferentially interacting with distal sequences on the same side of the boundary rather than on the opposite side. A majority of TAD boundaries are conserved across cell types and to a limited degree across species (DIXON et al. 2012; RAO et al. 2014). These physical boundaries often correspond to sharp transitions in transcriptional state and associated histone marks, suggesting that they play a role in transcriptional regulation. Active and repressed chromatin regions tend to interact with themselves over each other (SCHOENFELDER et al. 2015; BONEV et al. 2017). Notably, active promoters establish long distance contacts with each other, perhaps reflecting the formation of transcription factories (SCHOENFELDER et al. 2015; BONEV et al. 2017). TADs are thought to contribute to gene regulation by increasing the chance of interactions between regulatory elements and promoters within the domain and reducing the chance of potentially deleterious interactions across boundaries. In agreement with this, enhancers tend to only contact promoters within the same TAD, with few cross-boundary enhancer-promoter interactions (SCHOENFELDER et al. 2015; BONEV et al. 2017). Further, the expression levels of genes within TADs are more strongly correlated than between neighboring TADs and genes that are coordinately up or downregulated during development are likely to reside within the same TAD (BONEV et al. 2017; RAMIREZ et al. 2018). In line with this, contacts within TADs are dynamic during development, with changes in contact levels between putative enhancers and promoters corresponding to
changes in transcriptional activity as well as the presence of cell-type specific transcription factors (SCHOENFELDER et al. 2015; BONEV et al. 2017). Taken together, these findings suggest that TADs play an integral part in regulating constitutive and developmental transcription.

Along with physical maps of the genome, mapping of histone post-translational modifications as well as non-histone chromatin associated proteins has allowed for classifying chromatin into different states linked to transcriptional activity. Fillion et al. classified the Drosophila genome into five chromatin types. Two chromatin types (RED and YELLOW) are enriched for actively transcribed genes and active histone marks (H3K4me2, H3K79me3). The three remaining chromatin types (GREEN, BLUE, and BLACK) contain primarily transcriptionally inactive genes, and include Heterochromatin Protein 1 associated constitutive heterochromatin marked by H3K9me3 (GREEN), Polycomb group repressed chromatin marked by H3K27me3 (BLUE), and a third class of chromatin that is devoid of specific histone marks (BLACK) (FILION et al. 2010; KHARCHENKO et al. 2011). Transitions in chromatin types often overlap physical TAD boundaries (SEXTON et al. 2012), as illustrated in Figure 1.1. These boundaries are enriched for binding a class of proteins called insulator proteins. Insulator proteins have been extensively studied in Drosophila for their ability to block enhancer-promoter communication and prevent the spread of repressive chromatin. Thus, insulator proteins may mediate boundary formation and establish divergent transcriptional states at boundaries.

The role of insulators in genome architecture

Insulator proteins bind insulators, a class of DNA regulatory element that facilitates independent gene function by constraining enhancers and silencers action. Insulators function in a position dependent manner, blocking enhancer or silencer action only when placed between promoter of a gene and the regulatory element, termed enhancer blocking function (GEYER AND
CORCES 1992). They also block the spread of repressive chromatin, protecting inserted
transgenes from a repressive chromatin environment, termed barrier function (KELLUM AND
SCHEDL 1991; ROSEMAN et al. 1993). Insulators can block enhancers and silencers without
inactivating them, leaving them free to act on genes in the opposite direction (SCOTT AND GEYER
1995). Together, these two functions have resulted in the idea that insulators regulate
transcription by bounding regions of transcriptional activation and repression.

There are two non-mutually exclusive models for how insulator function is achieved. One is
architecturally based, through the formation of chromatin loops. This model posits that the range
of enhancers and silencers action on promoters is limited by the formation of looped domains of
chromatin (FUDENBERG et al. 2016). These loops are proposed to allow interaction of promoters
and regulatory elements within a loop but limit interactions between loops. Insulator loops are
also proposed to facilitate long range contacts by bringing together sequences at the base of the
loop, bypassing the intervening sequence (COMET et al. 2006; KYRCHANOVA et al. 2008).
Evidence for this comes from studies of insulator bypass. Insertion of a single insulator can block
enhancer-promoter communication or prevent the spread of repressive chromatin, however
insertion of a second insulator between the regulatory element and promoter can restore this
communication (CAI AND SHEN 2001). Notably, spread of the repressive H3K27 trimethylation
from a silencer can bypass paired insulators, resulting in repression of a distant gene but leaving
the region between the insulators unmarked and transcriptionally active, in agreement with a
model of looped intervening sequences (COMET et al. 2006; COMET et al. 2011). Similarly,
multiple insulator proteins have been shown to facilitate transgene activation from a distant
enhancer element, suggesting that insulator proteins facilitate long distance contacts between
insulator sites (KYRCHANOVA et al. 2008; ZOLOTAREV et al. 2016; ZOLOTAREV et al. 2017). This
model is further supported by extensive studies in mammals that reveal strong contacts between sequences at TAD boundaries that are enriched for insulator binding sequences in convergent orientation. Deletion or altered orientation of insulator sites can result in local loss of physical insulation and novel deleterious enhancer-promoter contacts (Guo et al. 2015; Lupianez et al. 2015; Narendra et al. 2015). The link between orientation of underlying insulator binding sites and mapped physical contacts strongly suggests that the insulator proteins bound at these positions drives loop formation (Rao et al. 2014; De Wit et al. 2015; Fudenberg et al. 2016).

The second model of insulator function is that insulators function as a decoy (Geyer 1997). In this model, insulators act as promoter-like decoys that compete for interaction with enhancers to inhibit enhancer-promoter communication. Similarly, the barrier function of insulators may be explained by recruitment of transcription factors promoting an open chromatin state to prevent the spread of repressive chromatin marks through nucleosome depletion, nucleosome turnover, or deposition of competing active marks. In fact, it has been suggested that insulators may have evolved from promoters, with recruitment of transcription factors to open and remodel chromatin taking on new roles in insulator function (Raab and Kamakaka 2010). In agreement with this, insulator sites are largely nucleosome depleted. This is thought to be due in part to the action of the NURF chromatin remodeling complex, which is thought to be recruited through the shared insulator cofactor centrosomal protein 190kd (CP190) (Kwon et al. 2016). Further support for this model comes from the observation that some promoters demonstrate insulator-like properties. Promoters with paused PolII have enhancer blocking activity (Chopra et al. 2009) and TFIIC binding sites (tRNA promoters) possess barrier activity (Kirkland et al. 2013). The decoy model is also in agreement with the observation that repressive domains of H3K27me3 are almost entirely bounded by either active genes, insulators, or a combination
thereof. Notably, only domains of H3K27me3 bounded by insulators are sensitive to insulator loss, suggesting that promoters themselves may act as barrier elements in some contexts (SCHWARTZ et al. 2012). In agreement with this, the presence of housekeeping genes and open chromatin (DHS) are strong predictors of TAD boundaries (EL-SHARNOUBY et al. 2017; RAMIREZ et al. 2018). While insulator bypass and orientation dependence of some insulators suggests a looping model, the resemblance between insulators and promoters should not be ignored. It remains to be determined which model more clearly reflects the mechanism of endogenous insulator function.

**Mammalian CTCF as an insulator and regulator of genome organization**

CCCTC-binding factor (CTCF) is an eleven zinc finger DNA binding protein. It was originally identified as a repressor of the chicken c-myc gene (KLENOVA et al. 1993; FILIPOVA et al. 1996), though subsequent studies identified an insulator function as well (BELL et al. 1999). CTCF is required for regulating chromatin contacts at a number of gene clusters in mammals including the hox gene cluster (NARENDRA et al. 2015) and the chicken beta globin locus (BELL et al. 1999; SPLINTER et al. 2006).

CTCF is bound at the majority (~60%) of mammalian TAD boundaries (ZUIN et al. 2014; BONEV et al. 2017). Strong contact between a subset of these boundaries are revealed by Hi-C, a variant of 3C that uses high throughput sequencing to generate genome wide maps of chromatin contacts. Points of strong contact between the boundaries at the edges of a TAD are interpreted as a chromatin loop. Loops enriched for CTCF binding require CTCF to form, indicating its essential role in facilitating long distance contacts (NORA et al. 2017). The loss of CTCF mediated looping increases inter-TAD contacts and decreases intra-TAD contacts (NORA et al. 2017). It has been noted that boundaries are enriched for both CTCF and cohesin, known for its
role in sister chromatin cohesion (NASMYTH AND HAERING 2009). CTCF/cohesin sites are located within TADs as well as at TAD boundaries (SOFUEVA et al. 2013). CTCF has been shown to recruit cohesin, and cohesin is required for CTCF insulator function (PARELHO et al. 2008), suggesting that cohesin may bridge CTCF sites to allow for loop formation. Notably however, loss of CTCF and cohesin have contrasting effects on chromatin organization. Loss of cohesin has more profound effects on chromatin organization, causing a drastic decrease in intra-domain contacts both within and between TADs (ZUIN et al. 2014; SCHWARZER et al. 2017). There is also little overlap between differentially expressed genes in CTCF and cohesin knockdown (ZUIN et al. 2014). These data indicate CTCF and cohesin have different functions. Whereas CTCF is necessary for maintenance of most topological domain boundaries as an insulating factor (NORA et al. 2017), cohesin plays a key role in facilitating chromatin interactions at all levels.

The functional consequences of CTCF mediated looping and physical insulation at the genome wide level are unclear. Despite large-scale changes in organization, there are only a few hundred genes with altered expression upon initial loss of CTCF (ZUIN et al. 2014; NORA et al. 2017). Notably, the clear majority (>80%) of these genes bind CTCF near their promoters (ZUIN et al. 2014; NORA et al. 2017). CTCF predominantly binds slightly upstream of the TSS of regulated genes at the beginning of the nucleosome-depleted region, suggesting it may play a role in preventing promoter occlusion by nucleosomes (NORA et al. 2017). These data suggest that while CTCF does mediate long distance chromatin contacts corresponding to chromatin loops, the primary mechanism of CTCF regulation of gene expression is likely not through facilitating the establishment of TADs, but rather through direct regulation of genes through binding at their promoter.
The role of Drosophila insulators in genome organization and transcription

The role of insulator proteins in chromatin organization and gene expression has been extensively explored in Drosophila. In contrast to vertebrate CTCF, multiple Drosophila proteins with insulator function have been identified including Suppressor of Hairy-wing [Su(Hw)], Drosophila CTCF (dCTCF), Boundary Element-Associated Factor (BEAF32), GAGA factor (GAF), Deformed wings/zeste-white 5 (Dwg/Zw5), Pita, Zinc finger protein interacting with CP190 (ZIPIC), Insulator binding factors 1 and 2 (IBF1 and IBF2), Optix binding protein (Opbp), and the hetero-tripartite early boundary activity (Elba) complex consisting of elba 1, 2, and 3. (Parkhurst et al. 1988; Zhao et al. 1995; Gaszner et al. 1999; Moon et al. 2005; Aoki et al. 2012; Cuartero et al. 2014; Maksimenko et al. 2015; Zoletarev et al. 2017). These insulator proteins facilitate enhancer blocking and in some cases barrier function in reporter constructs and at specific endogenous sites. Some of the most studied endogenous loci are at the DNAse hypersensitivity sites within the Drosophila bithorax complex. The bithorax complex is a group of homeotic genes that controls the development of posterior thoracic and abdominal segments of the adult fly. The bithorax complex contains nine distinct regulatory regions that important for specifying homeotic gene expression in individual parasegments (Kyrchanova et al. 2015). The independent action of each regulatory element is conferred by flanking boundary elements. In an analogous function to its vertebrate orthologue, dCTCF along with GAF, Pita, and Elba possess insulator functions at boundaries within the bithorax complex (Schweinsberg et al. 2004; Moon et al. 2005; Holohan et al. 2007; Maksimenko et al. 2015; Kyrchanova et al. 2017). Loss of CTCF results in aberrant enhancer-promoter contacts that result in mis-regulation of bithorax complex genes and homeotic defects (Holohan et al. 2007). Zw5 and BEAF32 have well characterized insulator function at scs (specialized chromatin structure) and scs’ respectively (Gaszner et al. 1999; Blanton et al. 2003; Kyrchanova et al. 2008). Scs
and scs’ flank a region that becomes highly transcribed upon heat shock, resulting in a puffed-out appearance on polytene chromosomes in larval salivary glands. Scs and scs’ were studied for their ability to demarcate the structural bounds of this transcriptional unit. Scs and scs’ are able to protect a transgene from chromosomal position effects (KELLUM AND SCHEDL 1991). Notably, Scs and scs’ interact with each other and form chromatin loops (BLANTON et al. 2003), supporting an architectural model for their insulator function. However, despite their insulator function in transgene assays, scs and scs’ were later discovered to not be responsible for bounding the heat shock puff (KUHN et al. 2004).

The majority of Drosophila insulator proteins localize to promoters of actively transcribed genes (see Figure 1.1) (JIANG et al. 2009; NEGRE et al. 2010; SEXTON et al. 2012; MAKSIMENKO et al. 2015; ZOLOTAREV et al. 2016). Though each insulator protein has insulator function in transgene assays and in some cases limited genomic contexts, insulator proteins, with the exception of Opbp, bind thousands of sites in the Drosophila genome. For all but a few of these sites, it is unclear if these sites are functional insulators. Studies of BEAF32 found differentially expressed gene pairs that are separated by BEAF32 binding typically have decreased expression when BEAF32 is lost (JIANG et al. 2009). This result is consistent with BEAF32 functioning as a transcriptional activator rather than an insulator at these sites, as loss of an insulator would be expected to result in a convergence of the expression patterns of each gene due to novel shared interactions with regulatory elements. In line with a transcriptional rather than architectural role for BEAF32, despite extensive overlap of BEAF32 at TAD boundaries, BEAF32 knockdown has minimal effects on chromosome conformation (RAMIREZ et al. 2018). GAF, though enriched at promoters, is depleted from TAD boundaries likewise suggesting it does not play a critical role in genome architecture (RAMIREZ et al. 2018). GAF has numerous
transcriptional regulatory roles, primarily as an activator through recruitment of chromatin remodelers to facilitate promoter accessibility and polymerase binding (Okada and Hirose 1998; Fuda et al. 2015). These findings suggest that the term “insulator protein” is inaccurate for these insulator proteins, as they primarily function as direct transcriptional regulators.

In contrast to vertebrate CTCF, dCTCF is not enriched at TAD boundaries (Ramirez et al. 2018), suggesting a functional divergence between CTCF and dCTCF in mediating genome wide looping. Instead, most (77%) of boundaries in Drosophila are located at gene promoters (Ramirez et al. 2018). More specifically, housekeeping genes are enriched at TAD boundaries and depleted from the interior of TADs (Hug et al. 2017). Promoter containing boundaries are enriched the insulator proteins BEAF32 and ZIPIC as well as the insulator cofactor CP190 (Bartkuhn et al. 2009; Jiang et al. 2009; Ramirez et al. 2018). It is unclear what role insulator proteins play in boundary formation at these sites (Ramirez et al. 2018). This is consistent with studies in both mammals and Drosophila that looked at the predictive power of insulator proteins binding sites versus RNA polII recruitment and found that RNA polII is a stronger predictor of chromatin organization (Bonev et al. 2017; Hug et al. 2017; Rowley et al. 2017). Despite this, RNA polII recruitment is not sufficient to mediate boundary formation and active transcription is not required, suggesting that some aspect of promoters results in function as boundaries in agreement with a decoy-model for insulator function. (Bonev et al. 2017; Hug et al. 2017).

The role of Su(Hw) in gene expression

Su(Hw) is one of the best studied insulator proteins in Drosophila. Su(Hw) functions primarily as a transcriptional regulator, and thus serves as a model multifunctional insulator protein to understand basic mechanism of this class of proteins (Soshnev et al. 2013). Su(Hw) possesses enhancer blocking and barrier functions when bound at the gypsy retrotransposon
Su(Hw) binds a 350 bp region just 3’ of the 5’ long terminal repeat of gypsy. This region contains twelve clustered Su(Hw) binding sites (SBSs) that are necessary and sufficient to mediate the insulator function of gypsy (GEYER AND CORCES 1992). Su(Hw) insulator function at gypsy depends on binding three cofactors, Centrosomal protein 190kD (CP190) and Modifier of mdg4 67.2kD isoform (Mod67.2) for enhancer blocking function (GERASIMOVA et al. 1995; GHOSH et al. 2001; PAI et al. 2004) and ENY2 to block the spread of repressive chromatin (barrier function) (KURSHAKOVA et al. 2007). CP190 is a shared insulator cofactor, interacting with multiple insulators including BEAF32 (LIANG et al. 2014), dCTCF (GERASIMOVA et al. 2007), IBF1, IBF2 (CUARTERO et al. 2014), PITA, and ZIPIC (MAKSMENKO et al. 2015). CP190 forms obligate homodimers using its BTB domain (BONCHUK et al. 2011; VOGELMANN et al. 2014). This ability to dimerize is proposed to allow the formation of chromatin loops between different insulator sites (VOGELMANN et al. 2014). Notably, loop formation is predicted to allow BEAF32 to transcriptionally activate long range targets that do not directly bind BEAF32 (LIANG et al. 2014). Similarly, interaction of dCTCF and GAGA with distant sites that do not contain their binding consensus showed significant overlap with CP190 positive BEAF sites, suggesting that CP190 may bridge long range contacts between different insulator proteins (LIANG et al. 2014). Similarly, Mod67.2 can also dimerize and is proposed to help facilitate long distance contacts between insulators (GOLOVNIN et al. 2007).

In addition to gypsy, Su(Hw) is bound at ~3000 non-gypsy sites in the Drosophila genome. In contrast to the twelve Su(Hw) binding sites at gypsy, most non-gypsy sites contain only a single match to the Su(Hw) binding site consensus sequence (PARNELL et al. 2006). This has important implications for the potential function of non-gypsy Su(Hw) binding sites, as four
copies of the Su(Hw) binding site sequence of the third SBS in *gypsy* are needed for insulator function in transgene assays (SCOTT et al. 1999). Loss of Su(Hw) results in the upregulation of predominantly neuronal genes in the ovary and testis, indicating a role as transcriptional repressor (SOSHNEV et al. 2013; DUAN AND GEYER 2018). Su(Hw) binding sites at Su(Hw) regulated genes represent only a small fraction total Su(Hw) binding sites in the genome, with the majority of genes with nearby Su(Hw) binding sites remaining unchanged. Therefore, the function of single binding sites for Su(Hw) within the genome are unclear. Understanding the basic principles that direct Su(Hw) function at these sites might provide insights into shared mechanisms of multifunctional transcription factors.

Su(Hw) might function as a canonical insulator protein at some endogenous binding sites. Studies of some non-*gypsy* binding sites revealed that they were capable of mediating enhancer blocking in transgene assays (PARNELL et al. 2006; RAMOS et al. 2006). Notably these assays also revealed that whether or not a site is capable of acting as an insulator depended upon where they were inserted into the genome, suggesting that genomic context is an important factor in directing Su(Hw) regulatory output. Similar variability in insulator function has been observed for BEAF-32 and dCTCF sites, suggesting that this may be a shared feature of insulators (PARNELL et al. 2006; RAMOS et al. 2006; KUHN-PARNELL et al. 2008; SCHWARTZ et al. 2012). Notably, insulator function can depend on enhancer and promoter strength, suggesting that single SBSs may function as insulators for weaker enhancers or promoters in different contexts (SCOTT et al. 1999). In fact, studies of the bithorax complex have shown that insulators are not interchangeable, and thus a precise calibration of insulator to promoter strength may be critical for proper regulation (HOGGA et al. 2001; IAMPIETRO et al. 2008; KYRCHANOVA et al. 2017). The requirement for at least four Su(Hw) binding sites from *gypsy* for insulator function does not
rule out that individual Su(Hw) binding sites may function as insulators in some contexts (SCOTT et al. 1999). As mentioned previously, genomic context influences insulator activity in transgene contexts. Additionally, 21% of Su(Hw) binding sites overlap binding sites for other insulator proteins. It is possible that Su(Hw) may function cooperatively with other insulator proteins at these sites for insulator function. Such cooperativity between direct DNA binding insulator proteins is observed for dCTCF and Pita in the bithorax complex, where binding sites for both proteins are required for insulator function (MAKSIMENKO et al. 2015). Additionally, cooperativity is observed between CTCF and Su(Hw) at the Su(Hw) dependent enhancer blocking sequence at 1A2 (SOSHNEV et al. 2008). Three copies of the gypsy third Su(Hw) (3R:3) binding site are not enough for insulator function (SCOTT et al. 1999). However, addition of the 1A2 facilitator CTCF binding site, though not possessing insulator function itself, to the 3R:3 construct resulted in insulator function, suggesting that CTCF may facilitate Su(Hw) insulator function at the ~10% of Su(Hw) sites that bind CTCF (SOSHNEV et al. 2008). Similarly, sequences adjacent to CTCF, a subset of which are likely Pita or ZIPIC, are essential for CTCF insulator function at some sites (WETH AND RENKAWITZ 2011; CUARTERO et al. 2014; MAKSIMENKO et al. 2015). Despite functioning as an insulator in transgene assays, 1A2 acts as an enhancer of yar expression, with Su(Hw) functioning as an activator (SOSHNEV et al. 2008). In fact, there is no definitive evidence of a non-gypsy Su(Hw) binding site functioning as an insulator in an endogenous context. Together these data illustrate that the functions of individual Su(Hw) binding sites vary depending on context. Multiple factors may contribute to regulatory output and identifying these factors is the subject of ongoing investigations.

Su(Hw) acts as a transcriptional repressor in both the ovary and testis (PARNELL et al. 2006; SOSHNEV et al. 2013; DUAN AND GEYER 2018). Though most Su(Hw) repressed genes are
neuronal genes, the neuronal genes that are regulated differ between the ovary and testis. These findings suggesting that there might be cell type specific functions for individual SBSs (SOSHNEV et al. 2013; DUAN AND GEYER 2018). Notably, the majority of target genes repressed by Su(Hw) remain unchanged in a Mod67.2 and CP190 mutant background that eliminates gypsy insulator function, suggesting that Su(Hw) functions as a direct transcriptional repressor of these genes rather than functioning as an insulator blocking an enhancer (SOSHNEV et al. 2013). However, in the same Mod67.2 and CP190 mutant background, four out six Su(Hw) activated genes had increased expression, suggesting that insulator function may play some role in regulation of Su(Hw) activated genes (SOSHNEV et al. 2013). A critical Su(Hw) regulated neuronal gene is Rbp9. This gene is a member of the ELAV/Hu family (KIM AND BAKER 1993). ELAV/Hu genes encode proteins containing three RNA recognition motifs and bind poly U tracks in mRNA to regulate mRNA stability, splicing, and translation (KIM-HA et al. 1999). Of the Drosophila ELAV/Hu family genes, Rbp9 is unique in that, in addition to the brain, Rbp9 is also expressed in a limited number of cells in the ovary. Over or under-expression of Rbp9 causes infertility (KIM-HA et al. 1999; JEONG AND KIM-HA 2003). Loss of Su(Hw) causes a three-fold upregulation of Rbp9, resulting in ectopic Rbp9 accumulation (SOSHNEV et al. 2013). Both Su(Hw) loss and Rbp9 overexpression share a common phenotype of oocyte apoptosis, and loss of a single copy of Rbp9 is able to partially rescue su(Hw)/− fertility defects, indicating that regulation of Rbp9 by Su(Hw) is critical for fertility (SOSHNEV et al. 2013). As is often seen with transcription factors, genes that change expression upon Su(Hw) loss seem to be largely indirectly regulated, as only 20-30% have Su(Hw) binding sites nearby (ADRYAN et al. 2007; SOSHNEV et al. 2013) No clear predictors of Su(Hw) insulator versus transcriptional regulatory functions at endogenous sites have been identified.
Reflective of its role as a repressor, Su(Hw) is rarely found at promoters. Instead, Su(Hw) binding is enriched in BLACK chromatin (Figure 1.1) (Filion et al. 2010). BLACK chromatin, or “Void” chromatin, lacks histone marks associated with activation or repression (Filion et al. 2010; Kharchenko et al. 2011; Sexton et al. 2012; Ramirez et al. 2018). Most genes in BLACK chromatin produce no detectable transcripts, and reporter genes inserted into BLACK chromatin are repressed at a frequency even greater than that of insertions into BLUE or GREEN chromatin, suggesting that there is repression of genes in BLACK chromatin (Filion et al. 2010). This chromatin type is also separated into TADs (Filion et al. 2010; Ramirez et al. 2018). Notably, boundaries within repressed chromatin are not associated with a promoters and rarely contain BEAF, ZIPIC, or other promoter associated elements, and are instead enriched for Su(Hw), IBF1, and CTCF (Ramirez et al. 2018). Despite this, boundaries account for only a small percentage of total Su(Hw) and CTCF binding sites (Figure 1.1), supporting the idea that these proteins have functions outside of genome organization. It is currently unclear what function Su(Hw) plays at boundary and non-boundary sites genome wide.

Based on the differences between gypsy and endogenous sequences, a model has been proposed whereby the sequence of a Su(Hw) binding site might direct regulatory output of Su(Hw) (Baxley et al. 2017). Su(Hw) binds a modular consensus sequence consisting of an AT rich upstream region, a central region containing a GCATA motif, and a GC rich downstream sequence (Figure 1.2) (Spana et al. 1988; Soshnev et al. 2012; Baxley et al. 2017). The upstream and downstream modules are not both present in all binding sites, although at least two modules need to be present for Su(Hw) binding. The central module is required for all Su(Hw) binding, as disruption of the core disrupts Su(Hw) independent on whether other modules are present (Baxley et al. 2017). The presence of either the upstream, downstream, or both modules
defines different classes of Su(Hw) binding sites that are associated with different functions of Su(Hw). Su(Hw) binding sites within the gypsy insulator carry the central and upstream module but lack the downstream module. Sites resembling those at gypsy represent a minority of endogenous sites. The SBSs at Rbp9 contain the central and downstream modules, representing the majority of endogenous sites. Mapping of Su(Hw) binding to these modules reveals that different sets of zinc fingers (ZFs) bind each module, and that mutation of individual ZFs leads to separation of Su(Hw) insulator and transcriptional regulatory functions (BAXLEY et al. 2017). This has led to the “Su(Hw) code” hypothesis of function, which posits that differential engagement of Su(Hw) ZFs with DNA at different classes of SBS leads to different regulatory outputs of Su(Hw). The putative “insulator class” consists of sites that contain the upstream and central modules, and the “repressor class” consists of sites that contain the central and downstream modules. Binding to a modular consensus sequence with distinct groups of ZFs is not a feature unique to Su(Hw), the same has been found for CTCF (NAKAHASHI et al. 2013).

Notably, ZFs 10 to 12 are required for binding the upstream module. Disruption of ZFs 10 to 12 also disrupts Su(Hw) association with CP190 and ENY2, concomitant with loss of Su(Hw) binding to the insulator subclass of SBSs and its insulator function (MELNIKOVA et al. 2018). These observations support that DNA binding and ZF involvement may direct alternative recruitment of cofactors to influence regulatory output of Su(Hw).

Transcription factors contribute to the negative or positive transcriptional output of a gene generally through the recruitment of cofactors [for a review see (REITER et al. 2017)]. Cofactors can function in several ways to regulate gene expression, including acting as modifiers of post translational modifications or through chromatin remodeling that ultimately affects PolII recruitment and activation. In the case of Su(Hw), as previously mentioned, CP190, Mod67.2,
and ENY2 are required for the insulator function of Su(Hw) at gypsy (Ghosh et al. 2001; Pai et al. 2004). For each of these cofactors, the mechanism of their contribution to Su(Hw) insulator function is not clear. Both CP190 and Mod67.2 can dimerize, and thus potentially contribute to insulation through an architectural model of insulator function (Golovnin et al. 2007; Bonchuk et al. 2011). However, effects of loss of Mod67.2 on gypsy insulator function are variable depending on the insertion point of gypsy (Georgiev and Gerasimova 1989). Additionally, loss of Mod67.2 results in a repressive function for Su(Hw) bound to gypsy 5’ of the yellow gene, suggesting that Mod67.2 binding inhibits Su(Hw) repressor function (Gerasimova et al. 1995).

Similarly, CP190 is largely enriched at promoters and seems to be more closely tied to gene activation (Bartkuhn et al. 2009; Ramirez et al. 2018). In line with function as an activator, CP190 has been shown to recruit the histone acetyltransferase Gcn5. Recruitment of either Gcn5 or CP190 to silent chromatin can mediate gene activation (Ali et al. 2017). Similarly, CP190 binding is correlated with NURF301/E(bx) binding at Su(Hw) and other insulator sites and is associated with nucleosome depletion (Bartkuhn et al. 2009; Kwon et al. 2016). NURF301 and other NURF complex components co-purify with CP190 and are linked to insulator function (Bohla et al. 2014). Notably, a study found that SBSs lacking CP190 and Mod67.2 binding at the endogenous location of the SBS tended to function as silencers in transgene constructs (Schwartz et al. 2012), suggesting that these cofactors may inhibit Su(Hw) repressor function.

Another protein, dTopors, is an E3 ubiquitin ligase that interacts with Mod67.2 and facilitates Su(Hw) insulator function at gypsy through an unknown mechanism (Capelson and Corces 2005). Su(Hw) cofactors required for its activator and repressor functions are unknown.

One potential candidate cofactor for mediating Su(Hw) repressor function is the newly identified HP1 and Insulator Partner protein (HIP1, CG3680) (Alekseyenko et al. 2014; Rhee
et al. 2014). In Drosophila S2 cells, BioTAP-XL mass spectrometry demonstrated that HIPP1 associates with multiple DNA binding insulator proteins (ALEKSEYENKO et al. 2014), as well as Heterochromatin Protein 1a (HP1a). Of the insulator binding proteins (IBPs) studied, Su(Hw) shows the strongest overlap with HIPP1 by ChIP-seq (56% of HIPP1 sites), with CCCTC-Binding factor (CTCF) representing the next most common HIPP1 partner (19% of HIPP1 sites) (ALEKSEYENKO et al. 2014). HIPP1 is also the major Su(Hw) cofactor, associating with most (86%) SBSs, encompassing all sequence subclasses. This high degree of colocalization suggests that HIPP1 might contribute to Su(Hw) function.

**Thesis outline**

Su(Hw) is a multi-functional transcription factor, possessing insulator, activator, and repressor functions (LOBANENKOV et al. 1990; VOSTROV AND QUITSCHE 1997; BELL et al. 1999; VOSTROV et al. 2002; CIAVATTA et al. 2007; SOSHNEV et al. 2008; BAXLEY et al. 2011; SOSHNEV et al. 2013). Understanding how individual binding sites specify the function of multi-functional transcription factors is important for understanding basic mechanisms of transcriptional regulation. Su(Hw) serves as an ideal candidate for studying these basic mechanisms. Importantly, the identification of a novel Su(Hw) interacting protein, HIPP1, that has homology to human CDYL transcriptional repressor protein, has the potential to elucidate part of this mechanism. The goal of this thesis work is to understand the role of HIPP1 in development and its potential roles in Su(Hw) transcriptional regulatory functions. These studies combine genetic and immunohistochemical approaches to address this question. Chapter two details the techniques and methods used for generation of stocks and reagents. Chapter three details the experiments that were undertaken to address the role of HIPP1 in development and Su(Hw) function. These experiments involved characterizing the effects of novel Hipp1 alleles
on viability as well as a combination of gene expression, genetic interaction, and immunohistochemical analyses to examine the role of HIPPI in Su(Hw) function. Data from these studies reveals that HIPPI is a globally expressed nuclear protein that is not required for viability or fertility. Further, these studies argue against HIPPI involvement in either HP1a or Su(Hw) functions. Despite this, these studies revealed that Su(Hw) is required for the majority of HIPPI binding to SBSs, suggesting an as of yet undiscovered function that requires further investigation. Chapter four discusses these results and puts them into the context of what is known about the mechanisms of Su(Hw), HP1a, and CDYL function.
A.

B.  

<table>
<thead>
<tr>
<th></th>
<th>Genome</th>
<th></th>
<th>Percent of Sites at:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Promoters</td>
</tr>
<tr>
<td>Su(Hw)</td>
<td></td>
<td></td>
<td>19</td>
</tr>
<tr>
<td>CTCF</td>
<td></td>
<td></td>
<td>45</td>
</tr>
<tr>
<td>BEAF32</td>
<td></td>
<td></td>
<td>61</td>
</tr>
<tr>
<td>GAF</td>
<td></td>
<td></td>
<td>40</td>
</tr>
<tr>
<td>Zw5</td>
<td></td>
<td></td>
<td>43</td>
</tr>
<tr>
<td>IBF1/2</td>
<td></td>
<td></td>
<td>30</td>
</tr>
<tr>
<td>PITA</td>
<td></td>
<td></td>
<td>55</td>
</tr>
<tr>
<td>ZIPIC</td>
<td></td>
<td></td>
<td>58</td>
</tr>
<tr>
<td>HIPP1</td>
<td></td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>CP190</td>
<td></td>
<td></td>
<td>50</td>
</tr>
<tr>
<td>Mod67.2</td>
<td></td>
<td></td>
<td>42</td>
</tr>
</tbody>
</table>

- Sites overlap multiple colors
- Sites do not have a defined color
Figure 1. 1 Genome organization and localization of Drosophila insulator proteins

A. Illustration of a generic chromatin conformation capture (Hi-C) contact map. The black line represents part of a chromosome, with each position along the chromosome contacting each other position at some frequency represented by the color of the diamond at the diagonal intersection of those two positions, where darker red indicates higher contact frequency. Underneath is a representative illustration of the distribution of chromatin states from (Filion et al. 2010) in this chromatin, with boundaries of physical association corresponding to transitions in chromatin state.

B. The fraction of total mapped chromatin that corresponds to each state is shown on top (genome) as a cumulative bar graph. The fraction of binding sites for each insulator within each chromatin type is indicated, with each color representing the proportion of sites in that corresponding color. Grey represents the proportion of binding sites that are at the transition between two colors. White represents sites that map to regions that do not have a defined color. The percentage of binding sites for each insulator that are within 200 bps of an annotated transcription start site (refseq genes) or 500 base pairs of a TAD boundary (Ramirez et al. 2018) are indicated on the right.
Figure 1. 2 Su(Hw) binding site logo

MEME generated sequence logo generated from SBSs at Su(Hw) regulated genes (CanB, Sh, CG3104, Nlg2, Or35a, dimm, CyP6a, Jhe, SSE, CG5282, and CG14459). The positions of the Up, Central, and Down modules of the Su(Hw) binding consensus are indicated in green, red, and blue respectively (BAXLEY et al. 2017).
Drosophila stocks and culture conditions

All Drosophila stocks were raised on standard cornmeal/agar medium at 22°C unless otherwise specified. Two su(Hw)<sup>+/+</sup> strains were used in this study, including (1) y<sup>1</sup>w<sup>1118</sup> and (2) Canton S (Bloomington Stock Center, BL1). Three su(Hw) null alleles were used, including (1) su(Hw)<sup>v</sup> that carries a ~1.7 kb deletion encompassing the su(Hw) and RpII15 promoters (HARRISON et al. 1992), (2) su(Hw)<sup>2</sup> that carries an insertion of an ~1.3 kb element into the first intron of the su(Hw) gene (PARKHURST AND CORCES 1986), and (3) su(Hw)<sup>Pb</sup> that carries an insertion of a white marked piggyBac element into the second exon of the su(Hw) gene [su(Hw)<sup>e04061</sup> in Flybase]. Other stocks used include Su(var)2-5<sup>04</sup> and Su(var)3-9<sup>06</sup> provided by Lori Wallrath (U of Iowa) and T(2;3)Sb<sup>V</sup> (BL 878). Su(var)2-5<sup>04</sup> results from a point mutation changing lysine 169 to a stop codon (EISSENBERG et al. 1992). Su(var)3-9<sup>06</sup> results from a ~6 kb DNA insertion that blocks transcript accumulation (WESTPHAL AND REUTER 2002). T(2;3)Sb<sup>V</sup> results from an inversion plus translocation of the Sb<sup>l</sup> mutation, positioning Sb<sup>l</sup> adjacent to the centric heterochromatin in the right arm of chromosome 2 (BEATON et al. 1988).

Generation of Hipp1 alleles

CRISPR was used to generate multiple Hipp1 alleles, using methods outlined in (BIER et al. 2018). Small deletions were generated by embryo injection of single guide RNA expression plasmids, made from pcFD3 (Addgene plasmid 49410). Injected embryos expressed germline Cas9 (yw; nos-Cas9[II-attP40], Bestgene). Putative mutants were screened using a PCR-based restriction enzyme assay, with candidates confirmed using genomic sequencing. Five small deletion alleles were generated, 1G3, 1G5, 2G4, 3G6, and 3G10. A large deletion allele, Hipp1<sup>Δ37</sup>, was generated by injection of a pair of guide RNA plasmids, with putative mutants
screened using PCR (Figure 3.4) Finally, a Hipp1 replacement allele was generated that swapped sequences -314 and +3913 of HIPPI with DsRed (Hipp1\textsuperscript{ΔDsR}). In this mutagenesis, a pair of guide RNA plasmids (SG375, SG378) were co-injected with a pDsRed-attP (Addgene 51019) derivative (SG451) that carried 1 kb of upstream and downstream of Hipp1 relative to the guide RNA cutting sites. A summary of the molecular details of Hipp1 alleles can be found in Table 2.1.

\textit{HIPP1\textsubscript{GFP}} was generated using the scarless tagging method described in (Bier \textit{et al.} 2018) and illustrated in Figure 2.1A. gRNA expression plasmids targeting positions +3179 and +3240 of HIPPI were co-injected with a template plasmid containing 1kb homology arms flanking the GFP coding sequence adjacent to a \textit{piggyBac} transposon containing a DsRed expression construct (pHD-sfGFP-ScarlessDsRed, DGRC #1365). The region around the Hipp1 translation stop site was PCR amplified from the genomic DNA isolated from the fly line used for injection (Stock name: \textit{yw} ; \textit{nos-cas9 (attP40 y+) /CyO}). The PCR product was sequenced to define the genomic sequences around the translation stop site in the injection stock to ensure proper guide RNA design. Candidate gRNA positions were identified by the presence of a protospancer adjacent motif (PAM) sequence. The Zhang lab (MIT) CRISPR design tool was used to ensure there were no candidate off-target cut sites. Template plasmid assembly was done using Gibson assembly (NEB E2611). Primers used to clone the homology arms included a synonymous G to C mutation at +3174 (Ala to Ala) and a G to C change at +3236 in the 3’ untranslated region to eliminate homology to the PAM sequences found in the endogenous Hipp1 gene. The disruption of the PAM sequences in the template and therefore the edited gene was done to prevent Cas9 from cutting either the template plasmid or re-cutting the edited genomic sequence. The final template plasmid (SG1053) and gRNA expression plasmids (SG771 and SG991, targeting G5
and G10 positions respectively in Figure 2.1) were sequenced to confirm proper assembly. Template plasmid and gRNA expression plasmids were co-injected by Bestgene (Stock name: yw;nos-Cas9[II-attP40]). DsRed positive flies were crossed to a piggyBac transposase expressing line (Bloomington stock #8285) to excise DsRed, resulting in an in-frame fusion of the Hipp1 coding and GFP coding sequences. Successful generation of Hipp1GFP was confirmed by sequencing. HIPPI-GFP localization to chromatin was analyzed by staining polytene chromosomes from wandering third instar Hipp1GFP larva with antibodies against GFP and Su(Hw). HIPPI and GFP show nearly 100% colocalization on polytene chromosomes in agreement with the >80% colocalization of Su(Hw) and HIPPI determined by ChIP-seq in S2 cells (Figure 2.1B) (ALEKSEYENKO et al. 2014). These data suggest that modification of HIPPI with a GFP tag does not significantly interfere with its ability to bind to Su(Hw) binding sites.

**Generation of GFP-Rbp9**

The mechanism by which Su(Hw) insulator, activator, and repressor functions are defined at individual Su(Hw) binding sites is unclear. As discussed in the introduction chapter, regulation of Rbp9 by Su(Hw) is critical for female fertility. Thus, Rbp9 represents a model target for testing mechanisms of Su(Hw) transcriptional regulatory function. To address the question of the role of binding site sequence in directing Su(Hw) regulatory function, CRISPR was used to introduce sequence encoding a GFP tag at the translation start site of Rbp9. The tagged Rbp9 allele allows for subsequent study of effects of CRISPR induced alterations in the SBSs on Rbp9GFP expression (see Discussion Chapter). To generate an Rbp9 allele encoding a GFP tag, scarless tagging was performed as described in [86]. The strategy is outlined in Figure 2.2A. The region around the Rbp9 translation start site was sequenced from the y-w-;nos-cas9(attP2 y+)/TM6C (Sb,Tb) injection line to ensure guide recognition. A guide targeting Cas9 to cut just
5’ of the translation start site was identified, and forward and reverse primers corresponding to this sequence were designed with overhangs allowing for annealing into pCFD3 (Addgene plasmid 49410). Separately, a template plasmid was prepared by amplifying the 1 kb regions 5’ and 3’ of the Rbp9 translation start site and assembling them with PCR amplified fragments of pHΔ-sfGFP-ScarlessDsRed (DGRC #1365) corresponding to the plasmid backbone and GFP-DsRed insert using Gibson assembly (NEB E2611S). The resulting plasmid contained the 1kb homology arms flanking the GFP coding sequence adjacent to a piggyBac transposon containing a DsRed expression construct (see template in Figure 2.2). The final template plasmid (SG1032) and gRNA expression plasmid (SG1001, targeting G4 in Figure 2.2) were sequenced to confirm proper assembly. The gRNA expression plasmid and template plasmid were grown up in E.coli and purified using QIAGEN maxiprep kit (cat: 12163) for injection. Template plasmid and gRNA expression plasmids were co-injected by Bestgene. DsRed positive flies were crossed to a piggyBac transposase expressing line (Bloomington stock #8285) to excise DsRed, resulting in an in-frame fusion of the GFP coding sequence with the original Rbp9 translation start site. Successful generation of a fly stock containing the Rbp9GFP allele was confirmed by sequencing. Immunohistochemical analysis was used to determine the expression pattern of Rbp9GFP. GFP-Rbp9 is cytoplasmic and starts to be detected in region 2 of the germaria and subsequently decreases expression until it is no longer detectable by stage four of egg chamber development. This expression pattern closely matches was published for Rbp9 (Figure 2.2B) (SOSHNEV et al. 2013). To determine if GFP-Rbp9 is functional, the Rbp9GFP line was crossed to a line containing Rbp92690, a mutant allele caused by a P element insertion (KIM-HA et al. 1999). Rbp92690/2690 females are infertile, showing an accumulation of undifferentiated germline cells. In contrast Rbp92690/GFP flies had normal oocyte development, with strings of developing egg chambers
These observations demonstrate that $Rbp^{GFP}$ can rescue $Rbp9$ loss. Finally, regulation of $Rbp9^{GFP}$ by Su(Hw) was determined by crossing $Rbp9^{GFP}$ into a $su(Hw)^{-}$ background. Immunohistochemical analysis revealed that in $su(Hw)^{-}$ females expression of $Rbp9^{GFP}$ persists up until $su(Hw)^{-}$ egg chambers die (Figure 2.2D). This expression patterns matches previous reports, though with increased accumulation in the oocyte (SOSHNEV et al. 2013). These data indicate that the $Rbp9^{GFP}$ allele is suitable for future studies of Su(Hw) function.

**Generation of HIPPI antibody**

Peptide Specialty Laboratories (Heidelberg Germany) generated two polyclonal guinea pig HIPPI antibodies. Two peptide antigens consisting of amino acids 570 to 585 (TSARKPRASDSWDYVY) and 599 to 620 (RSNSSYSSNASVSRNSLDNRPG) of HIPPI (Figure 2.3A) were injected together into two guinea pigs (GP1 and GP2). Serum was tested for specificity against HIPPI by western blot of protein extracts from Drosophila ovaries. Unpurified serum antibodies recognized many proteins indicating a need to affinity purify the HIPPI antibody. To purify α-HIPPI antibody, cDNA encoding amino acids 454 to 630 of HIPPI was PCR amplified from RE15416 containing the Hipp1 cDNA (DGRC stock number 1088000, SG 857) into pSC-A-amp/kan StrataClone PCR Cloning Vector (Agilent #240205). It should be noted that the single base pair deletion at position +1966 of the Hipp1 cDNA in the RE15416 flybase sequence is not correct, sequencing confirmed that this base pair is present in the plasmid and the cDNA encodes the full length HIPPI protein. The HIPPI cDNA was subcloned into a protein expression vector (EF1, courtesy of Dr. Ernesto Fuentes at University of Iowa) using NcoI and XhoI restriction enzymes. This expression vector contains the lac operator plus T7 promoter controlling expression of the inserted sequence in frame with sequence encoding a
6xHis tag. *Hipp1*-EF1 (EC93) was transformed into DE3 competent *E. coli* expressing T7 polymerase and induced with 0.1M IPTG for 4 hours at 37 °C. HIPP1:His was purified under denaturing conditions using a Ni-NTA superflow column and eluted with imidazole. Dialysis of HIPP1:His purified protein in elution buffer (20 mM NaH₂PO₄, 500 mM NaCl, 8M Urea) was performed in steps (8, 6, 0.6, 0.1, 0 molar urea) to a final concentration of (25mM HEPES, pH7.6, 15% glycerol, 100mM KCl). Protein concentration was estimated by a Coomassie stained polyacrylamide gel run of a dilution series of dialyzed protein compared to BSA of known concentration. Approximately 2 mgs of HIPP1:His was bound to Actigel ADL beads using ALD coupling solution (Sterogene Bioseparations). Guinea pig serum was heat inactivated at 55 °C for 30 min and placed on ice. Serum was filtered through a .22 µm syringe-driven filter (Millex-GS SLG033SS). HIPP1-Beads were transferred to 10 ml Poly-prep Chromatography Column (BioRad) and washed with filtered 1X PBS. Serum was passed through the column six times. The column was washed twice with 10 mls of PBS. Antibody was eluted in 10 fractions by sequential addition to the column of 500 µls of 0.1M Glycine, pH 2.5 collected into 1.5 ml Eppendorf tubes containing 1M Tris pH8.0 (150ul). Antibody-antigen interactions are usually efficient at physiological pH, lowering the pH can disrupt those interactions without denaturing the proteins. Elution fractions are collected into tubes containing 1M Tris pH8.0 to neutralize the pH and prevent damage to antibodies. Neutral pH of resulting fractions was confirmed by pH paper and fractions were analyzed for protein content by silver stain. The majority of the antibody eluted in fractions 2 and 3. Fraction 3 from both GP1 and GP2 serum was tested for specificity by pre-incubation with ~100 fold excess of purified peptide 1 or 2 supplied by Peptide Specialty Laboratories. Protein extracts from either *Hipp1*+/− (CS) or *Hipp1*GFP/GFP 1-3 day old female ovaries (5 pairs per lane) were run on a western and blotted with 1:100 dilution of each
corresponding GP fraction 3 incubated with either no peptide, or combinations of peptide 1 and 2 (Figure 2.3B). These experiments revealed that serum from both GP1 and GP2 recognized HIPPI, however they recognized different sequences. GP1 serum contains antibodies against peptide 2 and GP2 serum contains antibodies against peptide 1. The GP1 serum showed higher levels of background and thus was not used in future experiments. GP2 serum had a nonspecific band at approximately 150 kDa that was recognized by antibodies against peptide 2. Because HIPPI is primarily recognized by affinity purified GP2 serum antibodies against peptide 1, all future experiments included pre-incubation of this antibody for 1 hour at room temperature with an excess of peptide 2. To test whether this antibody was suitable for use in immunohistochemistry, Hipp1+/+(CS) and Hipp137/DsR2 ovaries were stained. As a comparison, Hipp1GFP/+ and Hipp1+/+ ovaries were stained with antibodies against GFP. Staining ovaries with the anti-peptide antibody revealed nuclear signal in Hipp1+/+ which was absent in Hipp137/DsR2 ovaries. These observations indicate that this antibody detected HIPPI in vivo. However, there was a high level of cytoplasmic background signal making this antibody unusable for immunofluorescence studies (Figure 2.3C).

**Western blot analysis**

Western analysis of HIPPI protein was performed using ovary extracts obtained from 1 to 3-day-old HIPPI females (5 pairs per lane) and testes extracts from <1 day-old males (10 pairs per lane). Western blots were incubated with GP2 affinity purified HIPPI antibody (1:100) that was pre-incubated with >100-fold excess of peptide 2 for one hour to reduce background. HIPPI was detected using horse radish peroxidase (HRP) conjugated second antibodies (1:20,000; Jackson 706-035-148) and analyzed using Advansta WesternBright™ Quantum chemiluminescent kit (K-12042-D20). Blots were re-probed with anti-alpha-tubulin antibody.
Western analysis of histone crotonylation was performed using ovary extracts obtained from 1- to 3-day old females and testis extracts obtained from <1-day-old males. Membranes were incubated with rabbit pan α-crotonyl-lysine (panKcr) antibody (1:2,000; PTM-501) or Rabbit α-H3 antibody (1:2,000; Abcam 1791-100). Proteins were detected using secondary HRP antibodies (1:20,000; BioRad 172-1019) and detected using ECL detection reagents (GE Healthcare, RPN2106). The standard western protocol was modified to optimize for small proteins by excluding SDS from the transfer buffer and by reducing transfer voltage and time from 70V for 90 min to 30V for 70 min.

**Immunohistochemical analysis**

Larval imaginal discs and the central nervous system, as well as adult ovaries (1< day-old) and testes (three-day-old) were dissected into PBS and fixed in 3% EM grade paraformaldehyde in PBT (PBS with 0.3% v/v Triton X-100) for 20 min. Samples were washed in PBT three times for 5 minutes each, with larval tissues being washed an additional 1 hour in 1% PBT (PBS with 1.0% Triton X100). Primary antibodies were diluted in 5% BSA, 0.3% TritonX-100 in PBS and incubated with tissues overnight at 4 °C. Following washes, tissues were stained with 1 µg/ml DAPI (ThermoFisher Scientific) and mounted in ProLong Diamond Antifade Mountant (Invitrogen P36961). Antibodies include mouse α-Lamin Dm0 (DSHB ADL84.12) at 1:200, Rabbit α-CP190 at 1:1000 (WHITFIELD et al. 1988), Chicken α-Mod67.2 at 1:2000 (PARNELL et al. 2006), polyclonal goat α-Su(Hw) at 1:300 dilution (BAXLEY et al. 2011), rabbit polyclonal α-GFP (Life Technologies A11122) at 1:1000, mouse monoclonal α-HP1a (DSHB C1A9) at 1:200 and mouse α-pan polyglycylated tubulin (Millipore AXO 49) at 1:500. Secondary antibodies include donkey α-rabbit AF488 (Invitrogen A21206), donkey α-goat AF568 (Life Technologies A11057), donkey α-rabbit AF568 (Invitrogen A10042), donkey α-goat 488 (Life Technologies
A11055), and donkey α-mouse 647 (Invitrogen A31571). Secondary antibodies were used at a 1:500 dilution. Actin was stained with Texas Red-X phalloidin (Life Technologies T7471) at 1:500. All images were collected on a Zeiss LSM 710 Confocal Microscope and assembled with Adobe Illustrator.

**Fertility assays**

Female fecundity was measured by mating eight females of each genotype to four *Canton S* males per replicate in bottles. Bottles were capped with orange juice plates spread with yeast paste. Every 24 hours, orange juice plates were replaced, and the eggs laid were counted. Fecundity was defined as the number of eggs laid divided by the number of females in the bottle. If a female died during the assay, it was assumed to have died just before screening and the change in number of females was incorporated into calculations of eggs per female per day for subsequent screenings.

Male fertility was assayed using a sperm exhaustion assay as described in (BARTON *et al.* 2016; DUAN AND GEYER 2018). One-day-old males were mated to three virgin females for 3 days. At the end of this period, males were transferred to new vials with three fresh virgin females. Females were discarded from the original vial and the vial was kept and screened for progeny. This mating scheme was repeated for 15 days. Males were scored as fertile if they produced at least five progeny from a three day mating period.

**Position Effect Variegation (PEV)**

To determine effects of loss of HIPPI on heterochromatin formation, we tested the ability of *Hipp1* mutants to modify variegation of a *Stubble* allele (*Sb*) that is sensitive to levels of heterochromatin proteins (DIETZ *et al.* 2015). In these studies, wild type (*yw*), *Su(var)*2-5*04/+*, *Su(var)*3-9*06/06* and *Hipp1*/* females were crossed to *T(2;3)Sb/Y*/*TM3 [Ser]* males. The length of
six bristles on the thorax of adult females was examined to identify the number of non-stubble (long) bristles. Increased levels of heterochromatin inactivate the dominant $Sb^I$ mutation, restoring bristle length to wild type. As a result, higher levels of heterochromatin are associated with higher frequencies of long bristles.

As an additional test of PEV, multiple lines containing SUPor-P elements (ROSEMAN et al. 1995; BELLEN et al. 2004) were obtained as generous gifts from Keith Maggert and David Acevedo. These lines include KV00590, insertion site at Y:3472914, KV108 insertion site on the Y (SWENSON et al. 2016), and KV135, insertion site at Chr2R:1224899. $Hipp1^{1G3/1G3}$, $Hipp1^{ADsR/ADsR}$, and $Hipp1^{A37/A37}$ virgin females were crossed to males of each PEV reporter line and up to eight randomly selected newly eclosed males were imaged depending on the variability of expression. Images were rank ordered from least to most yellow and white expression, the median images are shown in Figure 3.11. To determine if $Hipp1$ is a recessive modifier of PEV, males from each PEV reporter line were crossed to $Hipp1^{ADsR/ADsR}$ virgin females. DsRed positive male progeny were crossed to $Hipp1^{A37/A37}$ virgin females. DsRed positive ($Hipp1^{A37/ADsR}$) and DsRed negative ($Hipp1^{A37/+}$) male progeny were imaged as previously described. As a positive control modifier of PEV, $Su(var)2-5^{CyO}$ virgin females were crossed to males of each PEV reporter line, non-Cy male progeny were imaged as previously described.

**Quantitative PCR analysis of gene expression**

Analysis of gene expression of Su(Hw) regulated targets in the ovary was completed using quantitative PCR. Ovaries were dissected from 6-8 hour old virgin females into ice-cold phosphate buffered saline (PBS) pH 7.4 and frozen at -80°C until needed. Total RNA was isolated from 75 pairs of ovaries per biological replicate using TRIzol reagent (Invitrogen). Genomic DNA was removed using Invitrogen DNase treatment and removal kit (Cat# AM1906).
Generation of cDNA was done using Applied Biosystems Reverse transcription kit (Cat# 4368814). Expression levels were normalized to RpL32 and to one of the replicates of Canton S RNA. Primer sequences are listed in Table 2.2.

**Chromatin Immunoprecipitation (ChIP)**

Chromosome association of HIPP1-GFP and Su(Hw) was tested using ChIP from 100-150 ovaries from 6-8 hour old females per biological replicate. ChIP was performed as described in (BAXLEY et al. 2011). Briefly samples were cross linked with 1.8% formaldehyde for 10 min, washed, and then sonicated on ice with a Fisher sonic dismembrator using 1/8” microtip at 45% amplitude with 8 pulses of 30s, resting 60s between pulses. ChIP of HIPP1-GFP was performed using a modification of a standard protocol, wherein chromatin was cross-linked with 3.0% formaldehyde for 30 min (ALEKSEYENKO et al. 2014), and 16 instead of 8 pulses of sonication were performed. Chromatin was sheared to an average size of 150-300 base pairs as determined by agarose gel electrophoresis. Ten percent of this volume was kept as input and stored at -80 °C for later processing. The remaining 90% was used for pulldown. Antibodies used in ChIP experiments include 5 µls of rabbit polyclonal α-GFP (Abcam Ab290), 7 µls of guinea pig polyclonal α-Su(Hw) (BAXLEY et al. 2011), and 6 µls of rabbit polyclonal α-pan crotonyllysine (PTM-501). As a negative control for HIPP1-GFP ChIP, GFP antibodies were used to ChIP chromatin from wild type (Canton S ovary chromatin). In these studies, all sites showed less than 0.3% input (data not shown), demonstrating the specificity of the GFP ChIP. Immunoprecipitated DNA was quantified using quantitative real time PCR (qPCR) with SYBR
green (Bio-Rad Cat# 170-8882). Analyses were performed on at least two biological replicates. Statistical analysis was performed using PRISM. Primer sequences are listed in Table 2.3.

**Rapid Amplification of cDNA Ends (RACE) procedure**

The lack of a phenotype in flies containing early frameshift inducing mutations in \textit{Hipp1} and the testis specific isoform of CDYL in mammals (Franz \textit{et al.} 2009) led to the question of whether \textit{Hipp1} may have an alternative promoter that may produce a protein despite early frameshift mutations. To address this question, 5’ RACE was carried out to map the transcription start site(s) of \textit{Hipp1} in the testis. RNA was isolated from less than 3 day old \textit{yw} Drosophila testes using Trizol as described previously (Soshnev \textit{et al.} 2008). 5’ RACE was performed using a 5’ RACE system kit (Invitrogen Cat# 18374-041) according to manufacturer’s protocol. Briefly, cDNA/RNA hybrid was created from RNA with a \textit{Hipp1} specific antisense primer (Gene specific primer/GSP) termed GSP1 (GTGCCACCTCCTTCTTAGTCGTG). Reverse transcriptase extends the cDNA from the position of the primer to the 5’ end of the mRNA. Hybrid cDNA:RNA was digested by RNase to remove the RNA strand and the cDNA was poly-C tailed. A second nested \textit{Hipp1} primer (GSP2, TGGGCGAAATGCTTTTGACAGTTGTT) was then paired with a primer binding to the poly-C tail to PCR amplify the cDNA. One microliter of a 1:100 dilution of the initial PCR was used for a second nested PCR amplification using another \textit{Hipp1} primer (GSP3, GGTGTGGTGGGTGCTGGACAT) paired with a primer that binds a unique sequence within the original poly-C binding primer used in the previous step. PCR product was cloned using Strataclone PCR cloning kit (Agilent Cat# 240205). Individual colonies were picked and sequenced.
**Northern blotting**

A 936 base pair fragment of DNA encoding the HIPPI crotonase-like fold domain (+1980 to +2915 in the cDNA, encoding a.a. 666-926) was PCR amplified from Hipp1 cDNA clone RE15416 (DGRC stock number 1088000) to serve as a template for probe generation. Random primers were used to generate $^{32}$P-ATP labelled probe using the Prime-It II Random Primer Labeling Kit (Agilent #300385). RNA was isolated using TRIzol from ~1000 testes and ~200 carcasses from <1 day old yw male flies, and from 200 ovaries from <3 day yw female flies. Poly-A+ RNA was then isolated from between 200-500 ugs of total RNA using the Qiagen mRNA Midi Kit according to manufacturer’s instructions (Cat. 70042). 2.5 ugs of poly-A RNA per sample was run on a formaldehyde-agarose gel and transferred to a Nytran N membrane (Whatman). The membrane was hybridized with the $^{32}$P-ATP labeled probe. The washed blot was used to expose X-ray film.

**Polytene staining protocol**

Salivary glands were dissected from wandering third instar larva into Cohen Buffer (10 mM MgCl$_2$, 25 mM Na$_2$GlyceroP0$_4$, 3 mM CaCl$_2$, 10 mM KH$_2$P0$_4$, 0.5% NP40, 30 mM KCl, 160 mM Sucrose) and incubated for 9 min. Glands were then fixed for 3 min with formaldehyde (0.1 M NaCl, 0.002M KCl, 2% Triton-X 100, 2% formaldehyde, 0.01 M NaPO$_4$, pH 7.0) . Glands were transferred to squish buffer (45% acetic acid, 2% formaldehyde in water) for 2 min. Glands were squashed between a coverslip and slide and flash frozen in liquid nitrogen. Coverslips were removed, and slides were washed in cold 1X TBST 3X for 5 minutes each (10 mM Tris-HCL pH 8.0, 150 mM NaCl, 0.05% Tween 20). Antibodies were prepared in block (10% donkey serum and 10 mg/mL non-fat dry milk in 1X PBS:137 mM NaCl, 3 mM KCl, 10 mM Na$_2$HP0$_4$, 9 mM KH$_2$P0$_4$). Slides were incubated with antibody in a humidity chamber for 1 hour for both primary
and secondary antibodies, washing 3 times with 1X TBST between staining. Antibodies include rabbit α-GFP (TP401) used at 1:1000, goat α-Su(Hw) used at 1:300, and mouse α-HP1α (C1A9) used at 1:200. Secondary antibodies include donkey α-rabbit AF488 (Invitrogen A21206), donkey α-goat AF568 (Life Technologies A11057) and donkey α-mouse 647 (Invitrogen A31571). Slides were stained with 0.2 μg/ml DAPI in PBS for 5 seconds, washed 2 times with 1X TBST for 5 min each, VectaShield was used as a mounting agent and a coverslip was applied.
Figure 2. 1 Generation of Hipp1\textsuperscript{GFP}

A. Schematic of the scarless tagging strategy used to GFP tag HIP1. The Hipp1 gene region is shown with the sequences encoding the crotonase-like fold domain indicated in red. Neighboring genes are indicated in gray. Dashed lines indicate positions targeted by guides to direct CRISPR mediated double strand breaks and homology directed repair (HDR), with G5 and G10 referring to the specific sites targeted by each gRNA expression plasmid. The structure of the template plasmid is shown above with plasmid sequence as the thin line, the blue trapezoids indicating homology arms, green indicating sequence encoding GFP, dark red indicating the DsRed gene and its regulatory sequences, and the arrows indicating the inverted repeats of piggyBac transposon sequence. Following isolation of DsRed positive flies, males were crossed to a line expressing piggyBac transposase to remove the piggyBac transposon containing DsRed. B. Confocal images of polytene chromosomes isolated from wandering third instar Hipp1\textsuperscript{GFP/GFP} larva stained with $\alpha$-GFP (HIPP1, Green) and $\alpha$-Su(Hw) (Red) antibodies. Split indicates the view of the GFP and Su(Hw) channels from the white boxed region split down the middle of the chromosome. Scale bars= 20 $\mu$m.
Figure 2. 2 Generation of $Rbp^{GFP}$

A. Schematic of the strategy used to generate $Rbp9^{GFP}$ expressing a GFP tagged Rbp9 protein. The $Rbp9$ gene region is shown with bent arrows indicating promoters, thin bars indicating untranslated region, thick bars indicating coding sequence, thick black lines indicating intronic sequences, thin lines indicating intergenic regions, and the positions of the three SBSs within the $Rbp9$ genomic region indicted by red inverted triangles. Dashed lines indicate the position targeted by the gRNA (see Methods) to direct a CRISPR mediated double strand break. The structure of the template plasmid is shown above the gene, with plasmid sequence as the thin line, the blue trapezoids indicate homology arms with sequence encoding GFP (green), DsRed and its regulatory sequences (dark red), and the inverted repeats of piggyBac transposon sequence (thick black arrows) are indicated. Following isolation of DsRed positive flies, males were crossed to a line expressing piggyBac transposase to remove the piggyBac transposon containing DsRed.

B. Confocal images of early oogenesis stages from newly eclosed $Rbp9^{GFP/+}$ ovaries stained with DAPI (white), α-Vasa (Red), α-Su(Hw) (Blue) and α-GFP (Green). Scale bars= 20 μm R1-R3 indicate regions one through three of the germaria, S2 indicates a stage two egg chamber.

C. Confocal images of $Rbp9^{2690/2690}$ and $Rbp9^{2690/GFP}$ ovaries stained with DAPI. Scale bars= 200 μm

D. Confocal images of $Rbp9^{GFP/+}; su(Hw)^{+/+}$ and $Rbp9^{GFP/+}; su(Hw)^{2/Pb}$ ovarioles stained with DAPI (left, white, single slice) and α-GFP (right, green, projection). Egg chamber stages are labeled in white, with S2 indicating stage 2, and S3 indicating stage 3 etc. Scale bars= 50 μm.
Figure 2. 3 Generation and validation of peptide antibody

A. Schematic of Drosophila HIPPI protein. Colored regions correspond to the crotonase-like fold domain (Red, amino acids 666-926), the positions of the amino acids in the peptides (1:570-585, 2:599-620) used for antibody generation (Green), and the region used to affinity purify the antibodies (Cyan, amino acids 454-630). The amino acid sequences of the peptides are shown below. B. Western blot of ovary protein extracts using a 1:100 dilution of affinity purified guinea pig 1 or guinea pig 2 antibody. Antibodies were incubated with an excess of peptide 1, peptide 2, both, or neither prior to use as indicated. + indicates CS and GFP indicates HIPPI<sub>GFP/GFP</sub>. SE=short exposure and LE=long exposure. Size markers are in kDa. Dashed lines indicate the expected positions of HIPPI-GFP (Green) and HIPPI (Black), bands not at these sizes are predicted to be non-specific and do not change size between genotypes. C. Left: Confocal images of late stage egg chambers from newly eclosed HIPPI<sup>+/+(CS)</sup> and HIPPI<sup>Δ37/ΔDsR</sup> females stained with DAPI and GP2 antibody (1:100 dilution) incubated with an excess of peptide 2 prior to use. Right: Confocal images of ovaries from a newly eclosed HIPPI<sub>GFP/+</sub> and HIPPI<sup>+/+(CS)</sup> females stained with DAPI and α-GFP antibody. Scale bars= 20 μm.
Table 2. 1 CRISPR generated *Hipp1* mutants

<table>
<thead>
<tr>
<th>Allele</th>
<th>Position of deletion&lt;sup&gt;1&lt;/sup&gt;</th>
<th>Number of inserted nucleotides</th>
<th>Net indel size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1G3</td>
<td>596 to 601</td>
<td>0</td>
<td>-6</td>
</tr>
<tr>
<td>1G5</td>
<td>587 to 609</td>
<td>0</td>
<td>-23</td>
</tr>
<tr>
<td>2G4</td>
<td>766 to 767</td>
<td>0</td>
<td>-2</td>
</tr>
<tr>
<td>3G6</td>
<td>2354</td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>3G10</td>
<td>2351 to 2364</td>
<td>4</td>
<td>-10</td>
</tr>
<tr>
<td>Δ37</td>
<td>593 to 3260</td>
<td>6</td>
<td>-2662</td>
</tr>
<tr>
<td>ΔDsR</td>
<td>-314 to 3912</td>
<td>1350</td>
<td>-2877</td>
</tr>
</tbody>
</table>

1: Position is relative to the *Hipp1* transcription start site.
Table 2. 2 Primers used for qPCR analysis of ovary gene expression

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward Primer</th>
<th>Reverse Primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAPDH</td>
<td>CGCTGCCCAAGACATCATTC</td>
<td>GAAACGACCTCCTCATCGGT</td>
</tr>
<tr>
<td>Ras2</td>
<td>ATTACCAATCGTGTCGCA</td>
<td>TGCCGGTGTTGTAATACAGA</td>
</tr>
<tr>
<td>B-Tubulin</td>
<td>GACACCAGATCGTTGTTGTT</td>
<td>TCCTACTTCGTGGAATGGAT</td>
</tr>
<tr>
<td>RpL32</td>
<td>AATCTCCTTGCCTTCTTG</td>
<td>AAGGGATACGACACAGGAT</td>
</tr>
<tr>
<td>su(Hw)</td>
<td>GCCCGGCTGATAATTCTTA</td>
<td>CAGGGATTTACGCTGCTCTT</td>
</tr>
<tr>
<td>Hipp1</td>
<td>AAGAGTTTCCAGTGAAGTG</td>
<td>CGGAGTGTTGAGTTTCTTT</td>
</tr>
<tr>
<td>CG7582</td>
<td>AAACCTGGGCGGCCTAGTA</td>
<td>GCCCATCTACTGGTTTCTGA</td>
</tr>
<tr>
<td>CG33099</td>
<td>GGAGATGACCTAAACATGACC</td>
<td>CCCGCGTTTATAAGTATGG</td>
</tr>
<tr>
<td>CG6293</td>
<td>GTCAATAGCTCCAGGGTTCT</td>
<td>CTGCCTTGATAATGGTCCTC</td>
</tr>
<tr>
<td>l(1)G0148</td>
<td>CGATCATGCTCAACGGTGAT</td>
<td>TGGCATCACTAAAACGGA</td>
</tr>
<tr>
<td>tipE</td>
<td>CTTTCCGAAAGTAAAGGCT</td>
<td>GAACGAGGGTATTGGAATGG</td>
</tr>
<tr>
<td>mAcR-60C</td>
<td>GTCGAGTTGATGACACAG</td>
<td>TGCTGCCAAAGAAAGAAGA</td>
</tr>
<tr>
<td>Mob2</td>
<td>AAATCCGAAGCTAAATCTGTTTC</td>
<td>CCAACACAAACCACAAATCGAC</td>
</tr>
<tr>
<td>Syn2</td>
<td>ATGGAAAAAGTGGTGACTC</td>
<td>AGCTCCTTGTGTCTTATTC</td>
</tr>
<tr>
<td>CG15760</td>
<td>GTGCCAGCTTTCTCTGTGA</td>
<td>CCTCCTGTCTCTCATCTTCA</td>
</tr>
<tr>
<td>Rbp9</td>
<td>TGTTGCATTCTCGTCATAA</td>
<td>TGCTTGTTCTGGTTTGCTTT</td>
</tr>
<tr>
<td>CG6282</td>
<td>GCATCCGAACCTACCTTGGCG</td>
<td>CTCCACGTAAGACAGGAG</td>
</tr>
<tr>
<td>Rph</td>
<td>GATGCTTCTCTCGCTCTGT</td>
<td>CGCTGGTGCTTGCTTCTTG</td>
</tr>
<tr>
<td>CG9813</td>
<td>GTTAATTTGGCTCAATCGC</td>
<td>TTGGAAGCCTGTAAATGGG</td>
</tr>
<tr>
<td></td>
<td>dpr19</td>
<td>AGAGTTACGCCGCTAGTTCC</td>
</tr>
<tr>
<td>----------------</td>
<td>-----------</td>
<td>----------------------</td>
</tr>
<tr>
<td>dpr19</td>
<td>TGAATCCCTCGTCAGGTGTG</td>
<td></td>
</tr>
<tr>
<td>CG17778</td>
<td>GATGGACGAGATGGAGGTC</td>
<td></td>
</tr>
<tr>
<td>nrv3</td>
<td>GCGACTACTATCCCGAATG</td>
<td></td>
</tr>
<tr>
<td>neuroligin 2</td>
<td>ACATCTGAGTCCAAGTTTAGCC</td>
<td></td>
</tr>
<tr>
<td>fas</td>
<td>ACTTTATTGGTTCCACGCCA</td>
<td></td>
</tr>
<tr>
<td>CG32017</td>
<td>CATTTAATGCGACGACGTGAA</td>
<td></td>
</tr>
<tr>
<td>Site</td>
<td>Forward Primer</td>
<td>Reverse Primer</td>
</tr>
<tr>
<td>------</td>
<td>---------------</td>
<td>---------------</td>
</tr>
<tr>
<td>1A</td>
<td>GTATATCCACATCACCAGACCTCAGG</td>
<td>ACATCCTCGAATCACTATGCAAGTGC</td>
</tr>
<tr>
<td>64A</td>
<td>AGGAAATGGCTGCCTGATGTAAGAG</td>
<td>TTATATGTTGGCTCCCTGCTTCCGC</td>
</tr>
<tr>
<td>gypsy</td>
<td>CAGGCGCACAACACGGTAGT</td>
<td>TTTGCCGAAAAATATGCAATG</td>
</tr>
<tr>
<td>5C</td>
<td>GCGTTAGATTTAGAAATAGGATAATGCC</td>
<td>AATTGCCCAGTCCCAACTCTG</td>
</tr>
<tr>
<td>23A</td>
<td>GGAAATCCAAGCCGAAACCGAA</td>
<td>AAGTGCCATAAAAACGGATGCT</td>
</tr>
<tr>
<td>61F</td>
<td>GCAGCCTTTGTCAGAGGTGAT</td>
<td>CTTTCAGGCTAATATCAGCAACGT</td>
</tr>
<tr>
<td>69D</td>
<td>GATCAGGCATACGCAACCATTG</td>
<td>GCAGCACTGACCTAACCACCATAAGTA</td>
</tr>
<tr>
<td>70C</td>
<td>TTACCCACTGCGATGAATATTAGT</td>
<td>GCAAAAATGATGGCAGCAGT</td>
</tr>
<tr>
<td>100B</td>
<td>GCGTCAAAAGGTTTTCTTAGCAT</td>
<td>AGCTGCGATCAAGCATATT</td>
</tr>
<tr>
<td>CG6282</td>
<td>TCGGAGGACTAGAAAAGCATT</td>
<td>CGGACACACAAACCAAGCC</td>
</tr>
<tr>
<td>CG9813</td>
<td>AAAGCAGGCACAGTGAAAGCCA</td>
<td>GCTCATGCGGCGCAATT</td>
</tr>
<tr>
<td>CG15760</td>
<td>GTGTGAGTCCGACCACAACA</td>
<td>CCTGCGTGTGGCATTTGGT</td>
</tr>
<tr>
<td>dimm</td>
<td>TGTAACCTAGCGGTTGGACC</td>
<td>CCGCAGCTTCTTGGTGT</td>
</tr>
<tr>
<td>Or35a</td>
<td>CCACGATGGCGACAAAAGTATT</td>
<td>TGATGTGGGTGTTGGAGG</td>
</tr>
<tr>
<td>SSE</td>
<td>AGCCATTAGAAAGCAGGAAAAGA</td>
<td>CGGAATGTTGGGCTGGATT</td>
</tr>
<tr>
<td>tipE</td>
<td>TGGAATAATCAATGGCCGCAG</td>
<td>ACACCTCGGAGCTAAAACATCCT</td>
</tr>
</tbody>
</table>
### Table 2.3 - Continued

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>3D</strong></td>
<td>AATCAAGAGTATGCACTAGATTTGCTG</td>
<td>CCGTCTACACACCTACGCATG</td>
<td></td>
</tr>
<tr>
<td><strong>16C1</strong></td>
<td>ATTCCCCTTTAAATGCAGACTCG</td>
<td>ATCATGTTAATGCACTATCGAGTTG</td>
<td></td>
</tr>
<tr>
<td><strong>35B</strong></td>
<td>AGGCAAATATAATCAAGGCTAAAGT</td>
<td>CACGTCTAAGAATACGTCGGCTAGA</td>
<td></td>
</tr>
<tr>
<td><strong>59F</strong></td>
<td>GGAGCTGAAATTTCTCATATAACCAC</td>
<td>GGGCGATGGATAATTTTACATTTGACAT</td>
<td></td>
</tr>
<tr>
<td><strong>62D</strong></td>
<td>TTTGGGCTGCACGTATCT</td>
<td>GCCCGGGCAAAACTTTATTT</td>
<td></td>
</tr>
<tr>
<td><strong>82E</strong></td>
<td>TGCAGTGCGAACATGATTTGCTTA</td>
<td>ACTGCATGCTAAACTGTAGTGGGC</td>
<td></td>
</tr>
<tr>
<td><strong>84B</strong></td>
<td>CTTCGTAGATTGGACAGCCTG</td>
<td>ACTGACAGTTGCCTGACATTTCT</td>
<td></td>
</tr>
<tr>
<td><strong>Jhe</strong></td>
<td>ACGCATTGAAATCATTTAGAGACG</td>
<td>TGCTAAAGAAAGTTATTATTTCTTTC</td>
<td></td>
</tr>
<tr>
<td><strong>CG3104</strong></td>
<td>CAGTTGGGAGCTTGCAGATAATATA</td>
<td>GCAGCGACTGGTCACACTT</td>
<td></td>
</tr>
<tr>
<td><strong>Mob2</strong></td>
<td>CGATTGTGTTGGTGTGTGGTG</td>
<td>GCCCTGAAAGTATGCAACGGAAAA</td>
<td></td>
</tr>
<tr>
<td><strong>Rbp9</strong></td>
<td>TTTTTTTCGAGTCAGTGGGGA</td>
<td>GGTGTAACGTGAACCTGGAACCTACATAATCCG</td>
<td></td>
</tr>
<tr>
<td><strong>16C8</strong></td>
<td>ACCTAGCGGTACAGCTTACATTC</td>
<td>GGCTCCATCTACATTAGGCTTATCA</td>
<td></td>
</tr>
<tr>
<td><strong>16D</strong></td>
<td>CATGAAATATCTAATCAACAGGGATT</td>
<td>GTGAAGTAGATGGCTCTATATGTTCGATT</td>
<td></td>
</tr>
<tr>
<td><strong>44C</strong></td>
<td>ACATCTTACAGACAGTTACCCTTCT</td>
<td>TAGATGACGCTTTTAGCAGTCTATG</td>
<td></td>
</tr>
<tr>
<td><strong>63E</strong></td>
<td>ACCTTTGACTATTAGTATAAGAATAATGAG</td>
<td>GTGGCGGCATCTACATTTCTGAAGA</td>
<td></td>
</tr>
<tr>
<td><strong>78E</strong></td>
<td>AACCTGAGAAGTTACGTGGATTGTA</td>
<td>CACCTATC CGCATAGTCGGTTCA</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 3: INVESTIGATIONS OF HP1 AND INSULATOR PARTNER PROTEIN 1, HIPPI

Introduction

Development requires the precise temporal and spatial regulation of transcription. Central to these processes are DNA binding transcription factors (TF) that read the genome and execute chromatin changes to alter transcription. Multiple classes of DNA binding TFs exist, with Cys$_2$His$_2$ zinc finger (ZF) TFs representing by far the most abundant class in metazoans (ENUAMEH et al. 2013; LAMBERT et al. 2018). Once bound, TFs impact transcription in multiple ways, including transcriptional activation and repression through targeted effects on promoters, as well as transcriptional insulation promoted by the formation of topological domains that shield promoters from inappropriate regulatory inputs. There is an abundance of evidence for individual TFs having multiple functions and demonstrating context-specific regulation (REITER et al. 2017; LAMBERT et al. 2018). How single TFs achieve such multiplicity of effector function remains poorly understood.

Drosophila Suppressor of Hairy-wing [Su(Hw)] represents an exemplar multifunctional TF with insulator, activator and repressor functions (GEYER AND CORCES 1992; ROSEMAN et al. 1993; SOSHNEV et al. 2008; SOSHNEV et al. 2013). Su(Hw) imparts transcriptional regulation using a twelve zinc finger domain to direct DNA binding (SPANA et al. 1988). Insulator function of Su(Hw) depends upon binding to clusters of closely spaced binding sites, exemplified by binding to the cluster of twelve sites in the gypsy retrotransposon (GEYER et al. 1986; GEYER et al. 1988; DORSETT et al. 1989; SCOTT et al. 1999). In contrast, the activator and repressor functions of Su(Hw) are largely associated with standalone non-gypsy Su(Hw) binding sites [SBSs](SOSHNEV et al. 2013). Of these transcriptional contributions, the Su(Hw) repressor function is the most prominent, based on findings that SBSs primarily localize within repressive
‘black’ chromatin (Filion et al. 2010) and nearby genes are generally derepressed upon Su(Hw) loss (Roy et al. 2010; Soshnev et al. 2013; Duan and Geyer 2018). The multiplicity of the Su(Hw) regulatory function has been linked to a “Su(Hw) code” (Baxley et al. 2017), wherein different combinations of Su(Hw) ZFs direct binding to SBSs carrying one of three sequence subclasses, each of which displays a distinct chromatin feature. Notably ZFs 10-12 are involved in binding cofactors ENY2 and CP190 (Melnikova et al. 2018). These observations suggest that Su(Hw) DNA binding may impact cofactor recruitment, leading to context-specific transcriptional regulation.

Insights into possible Su(Hw) function come from identification of a novel interactor, HP1 and Insulator partner protein 1 (HIPP1). HIPP1 was identified for its interaction with Heterochromatin Protein 1 (HP1a) and multiple insulators proteins including Su(Hw), CTCF, Ibf1, Ibf2, PITA, and cofactors CP190 and Mod67.2 (Alekseyenko et al. 2014). HIPP1 was also identified in a separate study that coaffinity purified HA-Flag-tagged Su(Hw) from S2R+ cells and analyzed Su(Hw) interactors by mass spectrometry (Rhee et al. 2014). The detection of HIPP1 with Su(Hw) as the “bait” gives further credence to this interaction (Rhee et al. 2014). HIPP1 is the fly homologue of vertebrate Chromodomain Y-like (CDYL) proteins, a family of transcriptional co-repressors (Caron et al. 2003; Mulligan et al. 2008). CDYL proteins carry an amino-terminal chromodomain that binds methylated H3K9 and H3K27 (Franz et al. 2009) and a carboxyl-terminal domain that displays homology with lipid-metabolizing enzymes of the crotonase superfamily (Lahn et al. 2002; Caron et al. 2003; Wu et al. 2009; Zhang et al. 2011). Biochemical analyses established that the crotonase domain functions both as a protein-protein interaction platform to recruit co-repressors such as histone deacetylases (Caron et al. 2003; Escamilla-Del-Arenal et al. 2013)], as well as a crotonyl-CoA hydratase that
negatively regulates histone lysine crotonylation (Liu et al. 2017a), a histone modification associated with active transcription. These findings implicated HIPPI as a candidate Su(Hw)-dependent co-repressor.

Here, we study the function of HIPPI. To this end, we made multiple Hipp1 alleles, including tagging the endogenous gene to generate Hipp1\textsuperscript{GFP}. Using these tools, we investigated the developmental expression and functional contributions of HIPPI. We find that HIPPI is a nuclear protein expressed in most cell types. Surprisingly, our immunohistochemical analyses uncovered limited co-localization between HIPPI and HP1a, and our genetic studies revealed that Hipp1 mutations do not suppress position effect variegation, indicating that HIPPI is not required for HP1a-dependent heterochromatin formation. Further, we demonstrate that Hipp1 null mutants are viable, as well as female and male fertile. Despite Su(Hw)-dependent HIPPI localization at SBSs, we found that HIPPI loss does not compromise the transcriptional or insulator functions of Su(Hw), indicating that HIPPI is a non-essential Su(Hw) cofactor. Further studies are needed to resolve the role of HIPPI in Drosophila development.

Results

HIPPI carries a conserved crotonase domain

Reciprocal affinity purifications identified the fly homologue of human CDYL as a Su(Hw) cofactor (Alekseyenko et al. 2014; Rhee et al. 2014). Three variants of human CDYL have been identified (Franz et al. 2009), of which only CDYL1b carries a functional chromodomain. For the other variants, one has an N-terminal extension that inactivates the chromodomain (CDYL1a) and the second lacks the chromodomain (CDYL1c). To understand the relationship of HIPPI to human CDYL, we defined the structural conservation of HIPPI and CDYLb (Fig. 2). Alignment of the amino termini of these proteins provided evidence of an extended amino
terminal region that showed signs of a degenerated chromodomain in HIPPI, wherein only a short region of homology was found that included two of the three aromatic cage residues essential for binding methylated lysine [data not shown, (JACOBS AND KHORASANIZADEH 2002)]. Based on these observations, we conclude that fly HIPPI lacks a functional chromodomain. These data imply that HIPPI is most likely to be functionally similar to the CDYL variant lacking a functional chromodomain detectable in testis, CDYL1a (FRANZ et al. 2009). The amino acid sequence of the carboxy terminal region of HIPPI has strong homology with proteins containing a crotonase-like fold domain (CLD; Fig. 3.2). In the crotonase superfamily, this domain carries the active site of the enzyme, wherein conserved structural elements preserve the formation of an “oxyanion hole” that is needed for stabilization of an enolate anion intermediate derived from an acyl-CoA substrate. The crystal structure of CDYL identified residues L403, L452 and D483 as the critical residues forming the oxyanion hole (WU et al. 2009). These residues are conserved in HIPPI (732I, 780L and 812E) and are invariant in other drosophilid HIPPI proteins (data not shown). Our findings suggest that HIPPI carries a functional crotonase domain.

**Hipp1 is a non-essential gene**

To understand the HIPPI function, we generated multiple mutant Hipp1 alleles using CRISPR technology. These included small and large deletions within the Hipp1 coding region, which were confirmed by PCR and sequence analysis (Figures 3.3 and 3.4, Table 2.1). Western blots of ovary extracts assessed effects of these CRISPR-induced indels on protein production. Of the Hipp1 alleles with small deletions, four (1G5, 2G4, 3G6, 3G10) were predicted to cause premature termination of the encoded protein, whereas the fifth mutant (1G3) was predicted to remove two amino acids (Table 2.1). Indeed, we found that the four putative premature
termination mutants failed to accumulate any HIPP1 protein, while the fifth generated a full-length protein (Figure 3.5). We did not detect HIPP1 protein produced in flies heteroallelic for large deletions (Δ37 and ΔDsR). In total, six Hipp1 null alleles were generated.

Once Hipp1 mutants were available, we defined effects of HIPP1 loss on viability. To this end, we crossed Hipp1+/TM6c males and females and determined the number of Hipp1+/+ (TM6c, Sb) versus Hipp1−/− (non-TM6c, Sb) progeny. We found that Hipp1 adults were obtained at or near the expected number, with these adults displaying normal morphology (Table 1). We conclude HIPP1 is not essential for Drosophila development. Strikingly, these findings contrast those for CDYL, wherein the knockout mouse is inviable (Wan et al. 2013).

**Hipp1 is broadly expressed during development**

To understand possible developmental roles, we defined the tissue distribution of HIPP1. As our peptide antibodies did not work in immunohistochemistry, we used CRISPR technology to engineer the endogenous gene to encode HIPP1-GFP (Figure 3.3), choosing a carboxy-terminal tag based on previous studies of CDYL (Escamilla-Del-Arenal et al. 2013). Western analysis of Hipp1GFP/+ demonstrated that HIPP1-GFP was stably produced at wild type levels (Figure 3.5). Furthermore, genetic analyses demonstrated that Hipp1GFP adults are produced at wild type levels (Table 3.2). We conclude that HIPP1-GFP serves as a faithful reporter of HIPP1.

We used Hipp1GFP individuals to determine whether HIPP1 and Su(Hw) are commonly co-expressed. Larval tissues were examined first. These experiments revealed that HIPP1 is nuclear enriched (Figure 3.6A), a finding that contrasts with the localization of other members of the crotonase family that are found in the cytoplasm, specifically in peroxisomes and mitochondria (Furuta et al. 1980; Geisbrecht et al. 1999). Expression of HIPP1 extensively overlaps with that of Su(Hw) in larval tissues, with the exception of brain and the ventral nerve cord (Figure
In these neuronal tissues, HIPP1-GFP is largely present, whereas Su(Hw) is largely absent. Even so, the optic lobe and central brain carry clusters of cells that express only Su(Hw), and not HIPP1 (Figure 3.6A). These studies show that HIPP1 and Su(Hw) extensively co-localize but are not obligate partners in all cell types.

We determined the spatial localization of HIPP1-GFP in the ovary, an adult tissue that requires Su(Hw) (BAXLEY et al. 2011; DUAN AND GEYER 2018). Drosophila ovaries are divided into ovarioles that carry an organized developmental program of advancing stages of oocyte maturation (BASTOCK AND ST JOHNSTON 2008). At the anterior end of an ovariole is the germarium, a specialized structure that contains somatic cells that comprise the stem cell niche and two to three germline stem cells (GSCs). Upon GSC division, differentiating germ cells undergo four incomplete mitotic divisions to form a sixteen-cell cyst called an egg chamber. Continued germ cell differentiation produces an egg chamber with fifteen polyploid nurse cells, one diploid oocyte and a surrounding layer of somatic follicle cells. In the absence of Su(Hw), oogenesis is blocked, due to complete apoptosis of mid-stage egg chambers. This defect results from loss of transcriptional regulation in both somatic and germ cells (SOSHNEV et al. 2013). We co-stained ovaries dissected from Hipp1GFP females with GFP and Su(Hw) antibodies to define their extent of co-localization. These studies revealed that HIPP1-GFP is present in all somatic and germ cells (Figure 3.6B), even in the mitotically active region of the germarium that lacks Su(Hw). In later stages of oogenesis, HIPP1-GFP localization parallels that of Su(Hw), being found on nurse cell chromosomes but excluded from nuclear regions that contain the nucleolus (Figure 3.6B). In the oocyte nucleus, Su(Hw) staining is absent, however insulator cofactors Mod67.2 and CP190 are diffuse up until stage 6, at which point they begin to colocalize into foci that are then maintained (BAXLEY et al. 2011). These foci resemble insulators bodies, aggregates
of insulators proteins that form in response to osmotic stress or cell death that are proposed to serve as temporary storage compartments for these proteins (Golovnin et al. 2008; Schoborg et al. 2013). HIPPI colocalizes with these foci (Figure 3.7) supporting the idea that HIPPI can interact with at least one of the foci components in the absence of Su(Hw) (Alekseyenko et al. 2014). These studies reveal extensive co-localization of HIPPI and Su(Hw) in the ovary.

Su(Hw) is also required in testes for sperm development (Duan and Geyer 2018). For this reason, we examined HIPPI expression in the testis. Drosophila testes carry a single stem cell niche, called the hub, that supports two stem cell populations, GSCs and cyst stem cells. Spermatogenesis begins upon asymmetric division of both stem cell populations to form a differentiation unit of somatic cyst cells and germ cells. Subsequent mitotic and meiotic divisions of the germ cells produce 64 spermatids that differentiate into sperm. In the absence of Su(Hw), males demonstrate an age-dependent block in late stages of spermatogenesis, resulting in a failure to produce sperm. These defects result from a loss of Su(Hw) in somatic cells of the testis (Duan and Geyer 2018). Expression of HIPPI and Su(Hw) was determined by co-staining testes isolated from Hipp1GFP males with GFP and Su(Hw) antibodies. We found low levels of HIPPI-GFP in all somatic cells (Figure 3.6C, data not shown). In contrast, HIPPI-GFP and Su(Hw) are produced in a complementary pattern in germ cells. Early spermatocytes have high levels of Su(Hw) that diminish as they develop, while HIPPI-GFP levels start out low and then increase at the time Su(Hw) declines (Fig. 4C). We tested whether the low expression of HIPPI in early spermatocytes was due to Su(Hw) and found that HIPPI expression pattern was unchanged in a su(Hw) mutant (Figure 3.8). Although Su(Hw) and HIPPI overlap in testis is less extensive as compared to other non-neuronal tissues, HIPPI is present in somatic cells where Su(Hw) function is essential for male fertility (Duan and Geyer 2018). These
immunohistochemical studies reveal that HIPPI and Su(Hw) co-localize in the testis where Su(Hw) regulation is required.

**Hipp1 produces a novel transcript in the testis**

Given the transcriptional complexity of genes in testis (Batut et al. 2013; Brown et al. 2014) and the testis specific isoform of CDYL (Franz et al. 2009), we asked whether HIPPI may have similar transcriptional complexity in the Drosophila testis. To address this, we looked at a RAMPAGE (RNA Annotation and Mapping of Promoters for the Analysis of Gene Expression) data set (Batut et al. 2013) from multiple tissues and found evidence for a testis specific promoter at Hipp1 +1079 (Figure 3.9A). This potential transcript is predicted to encode a HIPPI short isoform consisting of a.a. 338-926 of HIPPI. To examine whether this potential promoter is conserved, and thus likely under selection pressure suggesting function, we examined the second transcription start site region across drosophilids and looked for evidence of core promoter elements (Figure 3.9B) (Sloutskin et al. 2015). We found that there were putative initiator sites (5/6 match to consensus) around the RAMPAGE identified start site in *melanogaster* and *simulans*, as well as at a site nearby in *ananassae*, however there were no corresponding sequences in *persimilis* or *virils*. This suggests that this promoter may have only recently been gained and may not be active in other drosophilids. To further test for the presence of this promoter, we performed Rapid Amplification of cDNA Ends (RACE) from RNA isolated from testis. In addition to clones mapping to the promoter of Hipp1 at +3 relative to the annotated transcription start site, we obtained clones mapping to +1088 and +1099 of HIPPI from two different experiments (shown as green arrow heads in Figure 3.9B). These data are consistent with the often broad transcription start sites associated with lack of strong positioning core promoter elements (Rach et al. 2009). To determine whether this transcript was unique to
the testis, and whether it may be alternatively spliced, we examined RNA from ovary, testis, and male carcass by Northern blot using a cDNA probe consisting of the sequences encoding the crotonase domain (Figure 3.9C). While carcass and ovary only had a single apparent band at the expected size (3.2 kb), we saw three apparent bands in the testis at approximately 3.2 kb, 2.9 kb, and 1.9 kb. The expected size of the novel promoter transcript is 2.2 kb. The difference in size may reflect an alternative transcription termination site, as the full length HIPPI transcript has a 455 bp UTR. Interestingly, this would also explain the 2.9 kb transcript if this alternative termination were to occur for the full length transcript. Alternatively, this transcript may arise from a third promoter, however based on the positioning of possible codons encoding a translation start site we would not predict such a transcript to give rise to a detectable protein as these codons are not in frame with the normal Hipp1 coding sequence. We tested whether this short isoform produced a protein by western and observed a band of the expected molecular weight for a protein containing a.a. 338-926 of full length HIPPI (HIPPI short isoform 64 kD + 27 kD GFP) (Figure 3.9D). These data support the usage of a second promoter giving rise to a short HIPPI isoform in the testis.

**HIPPI shows limited co-localization with HP1a**

HIPPI was biochemically identified as a high confidence HP1a interacting protein in S2 cells (Alekseyenko et al. 2014). Based on these findings, we predicted that we would detect foci of HIPPI overlapping those of HP1a in our immunohistochemical analyses, but this was not observed (Figure 3.10). To directly assess the HIPPI and HP1a partnership, we co-stained Hipp1GFP derived tissues with antibodies against GFP and HP1a. These two proteins show limited co-localization in cells that carry discrete HP1a loci (Figure 3.10). Despite this finding, genome-wide mapping studies in S2 cells had determined that HIPPI broadly associated with
heterochromatic regions (ALEKSEYENKO et al. 2014). In S2 cells, HIPPI is enriched in second, 
third, and Y chromosome heterochromatin, suggesting that HIPPI might be required for HP1a 
function in these regions. To evaluate this postulate, we tested the ability of Hipp1 mutants to 
suppress variegation of the Sb^l allele in the context of T(2;3)Sb^v. This translocation places the 
dominant Sb^l mutation adjacent to centric heterochromatin of the second chromosome. Flies 
carrying T(2;3)Sb^v display a mosaic thoracic bristle phenotype, wherein bristles are both short 
(Sb) and long (Sb^+; Figure 3.11). This phenotype reflects the variable spread of heterochromatin 
to the Sb^l gene, such that when heterochromatin reaches Sb^l, it inactivates expression of the 
dominant mutant allele and restores bristle length. Genetic backgrounds that decrease levels of 
heterochromatin, such as mutations in the gene encoding HP1a [Su(var)2-5] or the histone H3K9 
methyltransferase Su(var)3-9, dominantly decrease the number of wild type bristles (Fig. S2B). 
We crossed T(2;3)Sb^v into multiple Hipp1 mutant backgrounds and quantified the level of Sb^+ 
(long, wild type) and Sb (short, mutant) bristles. We found that all Hipp1 null mutants had 
numbers of wild type bristles that were similar to each other and to control lines (Figure 3.11A). 
These data suggest that HIPPI does not contribute to HP1a-dependent heterochromatin 
formation on the second chromosome, consistent with its limited co-localization with HP1a. As a 
second test, we determined whether Hipp1 mutants modified position effect variegation of the 
yellow and white genes within SUPor P transposons integrated into second or Y heterochromatin 
(Fig. 6C, D). We analyzed three SUPor P transgenic lines, chosen because these lines displayed 
moderate to severe repression of yellow and white gene expression (unmodified phenotype 
Hipp1^1G3+/*) and were sensitive to HP1a loss (suppressed phenotype Su(var)2-5^04). Whereas 
heterozygous loss of HP1a partially restored yellow and white expression, phenotypic 
improvement was absent in both heterozygous and homozygous Hipp1 null mutants (Fig.
These data indicate that HIPPI is not required for HP1a-dependent heterochromatin formation on the second or Y chromosomes, consistent with its limited co-localization with HP1a. Taken together, our findings indicate a limited partnership between HIPPI and HP1a, a surprising result given their strong association in S2 cells. We suggest that this protein partnership might be regulated. We note that the composition of heterochromatin changes during the cell cycle, demonstrated by the movement of TFs such as GAGA factor or Proliferation disrupter (Prod) between euchromatin and satellite sequences in heterochromatin (PATERO et al. 1998). In our studies, we primarily examined non-dividing cells, whereas S2 cells are actively dividing. It remains possible that differences in protein composition in heterochromatin during different stages of the cell cycle influences HIPPI recruitment or stabilization at these genomic regions.

**HIPPI is not required for fertility**

The extensive co-localization of Su(Hw) and HIPPI in the ovary and testes suggested that HIPPI might contribute to the function of Su(Hw) in these tissues. For this reason, we determined effects of HIPPI loss on oogenesis and spermatogenesis, as well as tested genetic interactions between *Hipp1* and *su(Hw)* mutants. In all cases, we examined effects in mutant animals that carried heteroallelic combinations to avoid complications associated with homozygous chromosomes that carry second site mutations.

We determined whether HIPPI was required for oogenesis. First, we measured the fecundity of *Hipp1* females and found that HIPPI loss had no effect on egg production (Figure 3.12). Further, *Hipp1* ovaries carried all stages of oogenesis with only low levels of apoptosis of mid-stage egg chambers (Table 3.3), as defined by staining with antibodies against Vasa and DAPI that reveals apoptotic egg chambers. Second, we tested for genetic interactions between
Hipp1 and su(Hw) null alleles. Trans-heterozygotes (Hipp1Δ37/+, su(Hw)Pb/+ and Hipp1ΔDsR/+, su(Hw)2/+), as well as Hipp1 mutants that were heterozygous for a su(Hw) mutation (HIPPIΔ37ΔDsR, su(Hw)2/+), all showed normal oogenesis and egg production (Table 3.3, Figure 3.12A). Third, we determined whether Su(Hw) regulated genes were mis-regulated in Hipp1 mutant ovaries. We reasoned that loss of HIPPI might alter transcription of Su(Hw) regulated genes without affecting oogenesis, as our prior studies showed that up-regulation of Rbp9 was primarily responsible for su(Hw)/- infertility (SOSHNEV et al. 2013). To this end, we isolated RNAs from Hipp1 and su(Hw) ovaries and measured RNA levels using qPCR (Figure 3.12B). In total, the transcriptional output of four Su(Hw) activated and fifteen Su(Hw) repressed genes was defined. We reasoned that if HIPPI was required for Su(Hw) regulation, then both heteroallelic mutant backgrounds should show changes in gene expression similar to those found in su(Hw) mutants. Notably, only one gene met these criteria. Mob2 was up-regulated in both Hipp1 mutant backgrounds, although the degree of up-regulation was reduced relative to that found in su(Hw) mutants (Figure 3.12B). Based on these data, we conclude that HIPPI does not contribute to Su(Hw) regulated transcription in the ovary.

We also determined the requirement for HIPPI in spermatogenesis. Notably, mouse CDYL has been implicated in spermatogenesis (LAHN et al. 2002; CARON et al. 2003; LIU et al. 2017a). Immunohistochemical studies found that CDYL is expressed mainly in round spermatids and spermatocytes (LIU et al. 2017a), a pattern similar to HIPPI. Investigation of the effects of CDYL loss on spermatogenesis is challenging because of the lethality of the Cdyl knockout mouse (WAN et al. 2013). However, tests of over-expression of CDYL were completed, which showed compromised sperm production, with these defects linked to hypo-crotonylation of histones (LIU et al. 2017a). First, we measured fertility of Hipp1/- males. These studies used a
sperm depletion assay to monitor fertility as males age (Figure 3.13A). In contrast to *su(Hw)* mutants, we found that *Hipp1*−/− males retained wild type levels of fertility over the two-week period, indicating sustained spermatogenesis. Second, we tested for genetic interactions between *Hipp1* and *su(Hw)* mutants. These studies uncovered a genetic interaction only in *Hipp1*Δ37/ΔDsR, *su(Hw)*2/Pb males, wherein fertility of these males declined faster than that of *su(Hw)*2/Pb males (Figure 3.13A). These findings suggest that HIPPI has modest contributions to male fertility in a *su(Hw)* mutant background. Third, we examined the testis phenotype in *Hipp1* mutant males. We focused on post-meiotic stages of spermiogenesis, because Su(Hw) loss affects these stages (Duan and Geyer 2018). A critical component of these stages is sperm individualization, which is characterized by the formation of actin rich individualization complexes (ICs) around sperm nuclei that travel from the posterior to anterior tip of the testis. ICs promote encasement of each sperm in its own plasma membrane [Figure 3.13; (Fabian and Brill 2012)]. As *su(Hw)*−/− testes age, ICs become disorganized and a bulge appears at the end of the testes, defects that coincide with reduced sperm production and small seminal vesicles (Figure 3.13 B+C). To evaluate *Hipp1*−/− testis phenotypes, we stained three-day-old testes with antibodies against polyglycylated Tubulin (PolyG Tub) that identifies sperm tails undergoing individualization, as well as phalloidin and antibodies against cleaved Caspase 3 to identify ICs. These analyses reveal a wild type testis phenotype (Figure 3.13 B+C), with evidence of continued IC formation, the absence of a posterior bulge and large seminal vesicles. Our data suggest that HIPPI is not required for regulation of Su(Hw) target genes involved in spermiogenesis. Fourth, we examined histone crotonylation in the *Hipp1*−/− testes, as CDYL has been implicated in the negative regulation of histone lysine crotonylation (Kcr) in spermatogenesis (Liu et al. 2017a). If HIPPI functions similarly to CDYL in the testes, then levels of histone Kcr should increase. To this end, western
analysis was completed with the α-crotonyl-lysine (panKcr) antibody used previously (Liu et al. 2017a). Only low levels of histones were recovered in testes extracts, with these histones showing low levels of Kcr that was unchanged upon HIPP1 loss (Figure 3.14A). Strikingly, the major Kcr protein is ~50 kD protein and no change in its modification was observed in Hipp1 mutants. We also analyzed protein extracts obtained from wild type and Hipp1−/− ovaries. In this case, histone Kcr is abundant, and is unchanged upon HIPPI loss (Figure 3.14A). While global levels of crotonylation appear unchanged, we asked whether levels of crotonylation were altered around HIPPI binding sites. We performed ChIP of histone H3 and crotonylated lysine from both HIPP1+/+ (CS) and HIPP1−/− (Δ37/ΔDsR) backgrounds (Figure 3.14C). We found no correlation between changes in crotonylated lysine levels and the presence or absence of HIPPI. These observations suggest that loss of HIPPI does not affect levels of crotonyl-lysine in germ cells. Taken together, we conclude that HIPPI is not essential for spermatogenesis.

**Su(Hw) recruits HIPPI to euchromatic regions**

Chromosome association of Su(Hw) is influenced by its cofactors Mod67.2 and CP190 (Schwartz et al. 2012; Soshnev et al. 2012). We wondered whether HIPPI also facilitated Su(Hw) chromosome association. To answer this question, we completed ChIP-qPCR of Su(Hw) in a Hipp1 null background. These studies found Su(Hw) binding to SBSs was unchanged upon HIPPI loss (Figure 3.15A)). We conclude that HIPPI is not required for Su(Hw) binding in the genome.

Although our data suggest that HIPPI is not essential for regulation of Su(Hw) target genes, over half of HIPPI associated regions are bound by Su(Hw) (Alekseyenko et al. 2014). This prompted our investigation of whether Su(Hw) is required for recruitment of HIPPI to euchromatin. First, we identified HIPPI occupied sites in the ovary. We dissected ovaries from
Hipp1-GFP females and immunoprecipitated HIPP1 using GFP antibodies, analyzing DNA enrichment using qPCR (ChIP-qPCR). Parallel studies were conducted using Canton S ovaries as a negative control, confirming the specificity of the GFP antibodies (data not shown). In total, 30 genomic regions were assayed. We chose these regions based on two criteria. First, HIPP1 occupied these sites in S2 cells. Second, these regions were occupied by different IBPs. The categories included Su(Hw) and HIPP1 absent regions (negative controls), as well as HIPP1 regions that were SBSs bound only by Su(Hw) [Su(Hw) only], SBSs bound by other IBPs [Su(Hw)+ IBPs] and non-Su(Hw) IBPs [IBPs-Su(Hw); Table 3.4]. We found HIPP1-GFP associated with 73% (22/30) of the predicted HIPP1 regions, including 72% of all SBSs and 80% of non-Su(Hw) IBPs regions (Fig. 3.15B). These data suggest that the overlap of HIPP1 occupancy in S2 cells and the ovary is strong. Strikingly, HIPP1-GFP occupancy at SBSs is higher than at non-Su(Hw) IBP regions. Even so, many SBSs did not bind HIPP1. Notably, the majority of these ovary lost regions (88%, 7/8) correspond to SBSs in Su(Hw) target genes. We asked whether this loss correlated with Su(Hw) occupancy at these sites by ChIP for Su(Hw) in the same background (Hipp1GFP/GFP) (Figure 3.16). This analysis revealed that Su(Hw) had a similar level of binding independent of HIPP1 presence or absence (23A and CG6282). These findings are consistent with our observations that HIPP1 is not required for Su(Hw) regulation of target genes in the ovary.

Having identified ovarian HIPP1 occupied regions, we asked whether Su(Hw) plays a role in HIPP1 chromosome association. We reasoned that if Su(Hw) were required for HIPP1 recruitment and/or retention, then loss of Su(Hw) would reduce HIPP1 association at SBSs and not at other genomic regions. To this end, we conducted ChIP-qPCR analysis of chromatin isolated from Hipp1GFP/GFP, su(Hw)+ females. Indeed, loss of Su(Hw) decreased HIPP1-GFP
occupancy at all (17/18) SBSs, whereas retention at all non-Su(Hw) regions was unaffected (4/4, Figure 3.15B). Strikingly, HIPPI occupancy was significantly decreased at regions bound by both Su(Hw) and other IBPs. As SBSs represent a large component (56%) of HIPPI bound regions, these findings imply that Su(Hw) has a major role in determining HIPPI euchromatic occupancy.

HIPPI binds the *gypsy* insulator in the ovary (Figure 3.15A). These observations suggested that HIPPI might have a role in establishing *gypsy* insulator function. To test this possibility, we defined effects of HIPPI loss on enhancer blocking, using the classic *gypsy*-induced *yellow* and *cut* mutations (CORCES AND GEYER 1991). The *gypsy* insulator in the *y* gene blocks the action of the wing enhancer, producing a yellow wing blade and the *gypsy* insulator in the *cut* gene blocks the wing margin enhancer, causing cuts or notches in the wing margin. Both of these mutant phenotypes are reversed by loss of Su(Hw) [Figure 3.17, *su(Hw)*2/*pb*], such wings are dark and smooth. In contrast, *y* *cut*, *Hipp1*Δ37/ΔDsR males had yellow and notched wings (Figure 3.17). Further, *Hipp1* mutants failed to dominantly enhance phenotypes found in heterozygous *su(Hw)* mutants. These studies reveal that HIPPI is not required for enhancer blocking by the *gypsy* insulator.

**Concluding remarks**

HIPPI is a biochemically identified partner protein of HP1a and Su(Hw) (ALEKSEYENKO *et al.* 2014; RHEE *et al.* 2014). Here, we investigated the functional contributions of HIPPI in Drosophila development. We generated multiple *Hipp1* null alleles (Figure 3.3), finding that *Hipp1* is a non-essential gene that is dispensable for female and male fertility (Figures 3.12 and 3.13). In interphase cells, we show that HIPPI is a globally expressed nuclear protein that largely localizes to chromosomes outside of heterochromatic domains (Figure 3.10). These observations
prompted investigation of HIPPI contributions to heterochromatin formation, showing that HIPPI is not a dominant modifier of position effect variegation (Figure 3.11). Together, our data suggest that HIPPI has a non-essential partnership with HP1a in heterochromatin. We also investigated a euchromatic role for HIPPI. Our data support early findings that Su(Hw) and HIPPI are protein partners (ALEKSEYENKO et al. 2014). We find that HIPPI chromosome association in the ovary strongly overlaps that defined in S2 cells (Figure 3.15). Further, HIPPI occupancy is strongest at SBSs relative to other IBP sites and depends upon Su(Hw) (Figure 3.15). Even so, HIPPI is not required for Su(Hw) regulatory functions, including its repressor, activator or insulator functions (Figs. 3.12, 3.13, 3.17). Notably, the absence of a regulatory contribution correlates with low HIPPI occupancy at Su(Hw) target gene SBSs and the gypsy insulator (Figure 3.15B). Our studies leave the developmental role of Drosophila HIPPI unresolved.

HIPPI represents the fly homologue of CDYL, a corepressor with crotonyl CoA hydratase function. This conclusion is based on the conservation of the crotonase domain and presence of the N terminal region of HIPPI containing putative remnants of a chromo-domain absent from other crotonase proteins in Drosophila. Further, the exon-intron structure of the Hipp1 gene within the region encoding the crotonase domain supports ancestry with Cdyl. Even so, fly and human HIPPI/CDYL appear to have evolved different developmental roles. Although HIPPI is not essential for Drosophila development, mouse Cdyl is required for viability and is linked to male fertility (WAN et al. 2013; LIU et al. 2017a). More studies are needed to determine whether other crotonase domain proteins compensate for HIPPI loss during development.
Figure 3. 1 HIPP1 is the major Su(Hw) cofactor

A. Shown is a UCSC Genome Browser view of a representative 418 kb region of chromosome 3R. Chip-seq tracks (top) and called peaks (bottom) are shown for Su(Hw), HIPP1, CP190, Mod67.2, dCTCF using datasets from S2 cells (OSHNEV et al. 2012; ONG et al. 2013; ALEKSEYENKO et al. 2014). Su(Hw)+HIPP1-CTCF sites are highlighted in green. CTCF+HIPP1-Su(Hw) sites are highlighted in red. B. Summary of frequency of co-localization of insulator proteins at individual genomic regions using data shown in A., with the number at the right corresponding to number of occurrences of each combination of insulator proteins indicated (black boxes).
Figure 3. 2 HIPP1 is the fly homologue of human CDYL

Shown is a diagram of the HIPP1 protein, which has a C-terminal crotonase-like domain (CLD, red) and an uncharacterized N-terminus (NT) with the corresponding amino acid positions indicated. Shown below is the alignment of the HIPP1 CLD with human CDYLb, Enoyl-coA hydratase, and bacterial Crotonase. Red boxes show the location of the three structural amino acid residues (red boxes) that are predicted to form the oxyanion hole in CDYL.
Figure 3. 3 Structure of the Hipp1 locus

A. Shown is the structure of the Hipp1 gene, with exons (large rectangles) colored to indicate the positions of the conserved crotonase-like fold domain (CLD, red) and bent arrows show directions of transcription. Structures of the neighboring CG3634 and CG3618 genes are indicated in grey. Inverted triangles above the Hipp1 gene indicate the locations of the small CRISPR generated deletions (1G3, 1G5, 2G4, 3G6 and 3G10), whereas the locations of the large deletions (∆37 and ∆DsR) are shown below the gene. The position of insertion of the GFP coding region is shown (raised green rectangle). Asterisks indicate the location of the epitopes recognized by the HIPPI antibody.
Figure 3. 4 Strategy for analysis of CRISPR generated large deletion Hipp1 alleles

*Top:* Structure of the Hipp1 gene and three newly generated alleles, including Hipp1-GFP (inverted green triangle) and two deletion alleles (Δ37 and ΔDsR). Locations of primers used in PCR analyses are indicated as arrowheads. Black arrowheads correspond to endogenous sequences and green arrowheads to GFP sequences. PCR reactions using five primer pairs were completed, using the indicated primer set (Table 3.1). Expected sizes (in kilobases) of the corresponding fragments amplified are shown at the right. *Bottom:* Ethidium stained agarose gel showing the PCR products generated from indicated PCR reactions. Genomic DNA analyzed was isolated from homozygous animals of the indicated genotype. + corresponds to Canton S.
Figure 3. 5 Western blot analysis of proteins produced from generated Hippl alleles

Western blot of protein extracts obtained from ovaries dissected from wild type (+/+, Canton S) or Hippl mutant females of the indicated genotype. Blots were probed with α-HIPP1 and α-αTubulin as a loading control. Positions of full-length proteins are shown by black and green arrowheads, indicating HIPP1 and HIPP1-GFP, respectively. Asterisks indicate positions of degradation products.
Figure 3. 6 HIPP1 is a nuclear protein expressed in most cell types

A-C. Confocal images of tissues dissected from Hipp1<sup>GFP/+</sup> and stained with antibodies against GFP (HIPP1, green) and Su(Hw) (red), with the merge image at the right. A. Top panels: Representative images of tissues dissected from third instar larvae, showing neuronal tissues of the central brain (CB), optic lobe (OL), ventral nerve cord (VNC), as well as non-neuronal tissues (eye disc, ED; leg disc, LD). Bottom panels: Magnification of boxed region of the central brain isolated from Hipp1<sup>GFP/+</sup> wandering third instar larva. This section reveals that some cell types express Su(Hw), but not HIPP1. Scale bars, 50 μm. B. Top panels: Image of a germarium, with the position of the somatic niche shown as a dashed line. Bottom panels: an early stage egg chamber (EC, bottom) that contains differentiated germ cells (nurse cells, NC) surrounded by somatic follicle cells (FCs). Scale bars, 20 μm. C. Top panels: Image of a testis that shows the somatic niche (hub, asterisk) and developing germ cell cysts. Bottom panels: Magnification of the boxed region to highlight the transition between Su(Hw) positive spermatocytes (stages S1 to S2) and Su(Hw) absent spermatocytes (S3). HIPP1 expression is stronger in mid-to-late stage spermatocytes. Scale bars, 50 μm.
Figure 3. HIPP1 localizes to foci containing CP190 and Mod67.2 in late stage oocyte nuclei

Confocal image of a segment of a stage 9 egg chamber from Hipp1\(^{GFP/\ast}\) stained with DAPI (white) and antibodies against GFP (green), Mod67.2 (red), and CP190 (yellow). Scale bars=10 \(\mu\)m.
Figure 3. Loss of Su(Hw) does not alter HIPPI expression pattern in testes

Representative image of a testis from a <1 day old Hipp1<sup>GFP/+</sup>, su(Hw)+/+ and Hipp1<sup>GFP/+</sup>, su(Hw)<sup>2/Pb</sup> males stained with antibody against lamin Dm0 (blue), GFP [HIPPI] (Green), and Su(Hw) (red). Dashed lines indicate transitions from S1 (left), S2 (middle), S3+ (right) spermatocytes as defined by nuclear size (lamin Dm0) and DAPI morphology (not shown).
Figure 3. 9 HIPPI has a novel transcript in the testis

A. UCSC browser view of RNA Annotation and Mapping of Promoters for the Analysis of Gene Expression (RAMPAGE) data from [Batut et al. 2013] for multiple tissues indicated. 1: mix of female and male tissue was used for the analysis. All data shown is for the plus strand. R1 and R2 are two separate replicates performed with testis tissue. B. Alignment of the Hipp1 gene region across drosophilids (million years evolutionary distance from melanogaster). Red arrowhead, site corresponding to RAMPAGE novel peak in the testis. Green arrowheads, transcription start sites identified by 5’ RACE. Red boxed region indicate putative initiator binding sites with the consensus shown below [Ohler et al. 2002]. C. Northern blot of RNA from carcass (male), ovary and testis using 32P-dATP- labeled probe corresponding to the crotonase like domain cDNA sequence of Hipp1. D. Western blot of protein extract from ovary and testis from Hipp1<sup>GFP/GFP</sup> and Hipp1<sup>+/+</sup> (CS) flies probed with antibodies against GFP and αTubulin.
Figure 3. 10 HIPPI and HP1a show limited co-localization in interphase cells

*Top:* Representative confocal image of an early stage egg chamber (stage 4) in an ovary dissected from a 1< day-old *Hipp1*<sup>GFP/+</sup> female stained with DAPI, α-GFP (green), and α-HP1a (red). *Middle:* Representative confocal image of the anterior portion of a 1< day-old testis dissected from a *Hipp1*<sup>GFP/+</sup> male and stained the same as (Top). Anterior is to the left. In testes, HP1a localizes diffusely in nuclei. *Bottom:* Representative confocal image of a third instar larval wing disc stained the same as (Top). Scale bars, 20 µm.
Figure 3. 11 Hipp1 is not a modifier of PEV

A. **Left:** Images of the thorax in a *Sb*+ (*Canton S*) and a *Sb*V animal, illustrating the six thoracic bristles that were quantified as either long (black arrowhead) or short (red arrowhead). **Right:** Box plots of quantification of the number of wild type (long) bristles per thorax of wild type flies (*Sb*+/+, CS) or *Sb*V/+ flies carrying either no known modifiers of PEV (*Sb*V crossed to 1: CS or 2: yw), or mutations of known suppressors of variegation *Su(var)2-5* and *Su(var)3-9*, or mutations of *Hipp1* (see Table 2.1) with the number of individuals analyzed shown above each bar. Each box represents the 25th to 75th percentile interval, the middle line represents the median, and the whiskers represent range. **B.** Representative dissecting scope image of abdomen and eyes of flies containing a SUPor-P transgene construct (ROSEMAN *et al.* 1995; BELLÉN *et al.* 2004) containing *yellow* and *mini-white* genes with the insertion site chromosome of the SUPor-P element listed below (see Methods Chapter). *Hipp1*¹G3/+ flies were used a negative control as the closest matched genetic background to the mutant alleles. *Su(var)2-5*⁰⁴/+ flies were used as a positive control of the effects of a known modifier of PEV.
Figure 3. 12 Hipp1 mutant females are fertile and do not phenocopy Su(Hw) loss

A. Fecundity (eggs laid per female per day) of five-day-old Hipp1<sup>−/−</sup>, su(Hw)<sup>−/−</sup> and heterozygous Hipp1<sup>+/−</sup>, su(Hw)<sup>+/−</sup> mutant females of the indicated genotypes, crossed to wild type males. The two wild type (+/+) reference strains were Canton S (1) and yw (2). Genotypes are noted under the graph. Error bars indicate the standard deviation from a minimum of three independent experiments. Fecundity was compared between genotypes using a one-way ANOVA followed by Tukey post hoc analysis. Asterisks indicate genotypes that were significantly different from control lines (p value < 0.01). Only su(Hw) null backgrounds showed a significant difference in egg laying.

B. Heat map of fold changes of gene expression defined by quantitative reverse transcription PCR (RT-qPCR) of Su(Hw) target genes, measuring gene expression levels in RNA isolated from 1<sup>st</sup> day-old su(Hw)<sup>+/+</sup> (Canton S), two su(Hw)<sup>−/−</sup>, two Hipp1<sup>−/−</sup> and one Hipp1<sup>−/−</sup>, su(Hw)<sup>−/−</sup> double mutant backgrounds. Fold change in expression was determined by normalizing levels to the housekeeping gene RpL32 and is relative to RNA levels in one of the three su(Hw)<sup>+/+</sup> (Canton S) RNA samples. The color key corresponding to fold change is shown below. Asterisks indicate gene expression changes relative to Canton S, * p<0.05, ** p<0.01, *** p<0.001 (Student's t-test).
Figure 3. 13 HIPPI is not essential for spermatogenesis

A. Quantification of the male fertility in wild type (1, Canton S and 2, yw), two Hipp1\textsuperscript{+/-}, two heteroallelic Hipp1\textsuperscript{+/+}, su(Hw)\textsuperscript{+/+} mutants, and one Hipp1\textsuperscript{+/-}, su(Hw)\textsuperscript{-/-} double mutant background. The number of males tested is shown above each data set. Bars indicate standard deviation from a minimum of three replicates. Significant changes in fertility between groups and over time were determined using repeated-measures ANOVA (#, not significant; ** p< 0.01; *** p<0.001).

B. Representative confocal images of 3-day-old wild type (Canton S), su(Hw)\textsuperscript{2/Pb} and Hipp1\textsuperscript{A37/ADsR} testis stained with antibodies against polyglycylated Tubulin (red, marks sperm tails), cleaved Caspase 3 (yellow, marks ICs) and phalloidin (blue, marks actin in ICs and elsewhere in the testis). Scale bars: 200 μm. Asterisk marks anterior of testis. S.V. denotes the seminal vesicle.

C. Quantification of maximum width at the end of the testis (just before seminal vesicle) from 7 day old (Canton S), su(Hw)\textsuperscript{2/Pb}, Hipp1\textsuperscript{A37/ADsR} and Hipp1\textsuperscript{A37/ADsR} su(Hw)\textsuperscript{2/Pb} males.
Figure 3. 14 HIPP1 is not required for regulation of histone crotonylation at SBSs

A. Western blot of proteins extracted from 1-to 3-day old ovaries and <1 day-old testes from Hipp1+/+ and Hipp1Δ37/ΔDsR animals probed with a pan α-crotonyl-lysine (panKcr) antibody and an antibody against Histone H3 (α-H3). A nuclear HeLa cell extract was run as a positive control. B. Western blot of proteins extracted from ovary and testes as described in (A), with transfer conditions optimized for small proteins (see Chapter 2-Methods). C. ChIP-qPCR analysis of crotonylated proteins at SBSs using antibodies against H3 and crotonylated lysine (panKcr) in Hipp1+/+(CS) and Hipp1Δ37/ΔDsR backgrounds at sites that either lack Su(Hw) and HIPP1, bind Su(Hw) and HIPP1, or bind Su(Hw) and HIPP1 plus other insulator proteins. Bars represent the standard deviation from three independent replicates. Asterisks indicate a significant change (t-test, p-value<.05) in binding between the Hipp1+/+(CS) and Hipp1Δ37/ΔDsR.
Figure 3. 15 HIPP1 occupancy at euchromatic sites depends on Su(Hw)

A. ChIP-qPCR analysis of Su(Hw) binding in Hipp1−/− (orange) and wild type (Canton S, red) ovaries. Three classes of sites were tested, including sites that lack Su(Hw) and HIPP1 (negative controls), SBSs that bind HIPP1 (HIPP1 present), and SBSs that lack HIPP1 (HIPP1 absent). Bars represent standard deviation of at least two biological replicates. Asterisks indicate a significant change in Su(Hw) binding between genotypes (t-test, p-value <0.05).

B. ChIP-qPCR analysis of HIPP1-GFP occupancy in the ovary, using chromatin isolated from Hipp1GFP/GFP females that carry either su(Hw) wild type (green) or su(Hw) null (su(Hw)2/Pb) alleles (blue). Four classes of sites were tested based on HIPP1 occupancy in S2 cells (ALEKSEYENKO et al. 2014), including 1) negative controls that lack HIPP1, Su(Hw) and other IBPs, 2) Su(Hw) only SBSs, 3) Su(Hw) SBSs bound by other IBPs, 4) IBP that lack Su(Hw). Table 3.4 lists IBPs found at each site, using data from modEncode and (CUARERO et al. 2014; MAKSIMENKO et al. 2015; BAXLEY et al. 2017). Tested sites were named for their cytological position or for the Su(Hw) target gene. These genes represent targets that are mis-regulated in su(Hw)−/− ovaries (blue) or testes (red). Asterisks indicate a significant enrichment of HIPP1-GFP in su(Hw)+/+ relative to the negative controls, p-value <0.05 (t-test). Bars represent the standard deviation from three independent replicates.
Figure 3. 16 Su(Hw) and HIPP1 binding correlate at most sites

Shown is ChIP-qPCR analysis of Su(Hw) and HIPP1-GFP binding at a number of sites lacking Su(Hw) (negative controls) and Su(Hw) binding sites. ChIP is from ovaries of newly eclosed Hipp1<sup>GFP/GFP</sup> females. Bars represent the standard deviation from three independent replicates.
Figure 3. 17 HIPPI is not required for Su(Hw) insulator function at gypsy

Shown are representative wings dissected from 1< day-old males that carry an X chromosome with two classic mutations caused by the gypsy insulator, yellow$^2$ ($y^2$) and cut$^6$ (ct$^6$). Males were generated that carried the $y^2$ ct$^6$ chromosome and were Hipp1$^{-/-}$, heterozygous Hipp1$^{-/+}$, su(Hw)$^{-/-}$ or Hipp1$^{-/-}$, su(Hw)$^{-/+}$, as indicated. When insulator function is lost, males carrying the $y^2$ ct$^6$ chromosome have a dark wing blade and smooth wing margin.
Table 3. Primers for verification of Hipp1 deletion alleles

| Primers for PCR analysis of large deletion alleles |
|---------------------------------|---------------------------------|
| A | GGTAGTTGATCTCGGCCGACCTTTAG | CTTGCCAGCTCAACATGAATCTC |
| B | GATACTAGTCACTGATCGCCTAC | TCCTCATCCAGACCGGGGATTT |
| C | GCACTCTGGGAGGGGAGAT | CGTGTAAAGCGATAATTTAATTGCAAACTTTCC |
| D | CTGTTCGCAACGGAACTTTTGGG | CGCTTCTCGTTGGCTCTT** |
| E | GGTGGATCTGGAGGTTCCCG** | CGTGTAAAGCGATAATTTAATTGCAAACTTTCC |

**: Primers bind within the GFP CDS
Table 3. 2 Percent viability of Hipp1 mutants

<table>
<thead>
<tr>
<th>Allele transmitted by female</th>
<th>Allele transmitted by male</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Δ37/TM6c</td>
<td>Δ DsR/ TM6c</td>
<td>Df3Lb/TM6c</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δ37/ TM6c</td>
<td>% viable</td>
<td># of TM6c</td>
<td>% viable</td>
<td># of TM6c</td>
<td>% viable</td>
</tr>
<tr>
<td>ΔDsR/ TM6c</td>
<td>84</td>
<td>319</td>
<td>142</td>
<td>257</td>
<td>92</td>
</tr>
<tr>
<td>Df3L/ TM6c</td>
<td>84</td>
<td>119</td>
<td>112</td>
<td>251</td>
<td>111</td>
</tr>
<tr>
<td>GFP/TM6c</td>
<td>88</td>
<td>57</td>
<td>80</td>
<td>383</td>
<td>116</td>
</tr>
</tbody>
</table>

*Percent viability was as determined by dividing the total number of non-balancer progeny obtained by half of the total number of TM6c progeny, multiplied by 100.

bDf(3L) refers to Df(3L)BSC452 that carries a 196 kb deletion that includes HIPPI and 28 other genes. Viability of homozygous mutants was not defined (gray boxes).
Table 3. Effects of *Hipp1* mutants on mid-oogenesis egg chamber death

<table>
<thead>
<tr>
<th>Genotype</th>
<th>% Apoptosis</th>
<th># ovarioles scored</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Hipp1</em>+/+, <em>su(Hw)</em>+/+ (CS)</td>
<td>4.7</td>
<td>107</td>
</tr>
<tr>
<td><em>Hipp1</em>+/+, <em>su(Hw)</em>+/+ (yw)</td>
<td>1.7</td>
<td>230</td>
</tr>
<tr>
<td><em>Hipp1</em>+/+, *su(Hw)*Pb/+</td>
<td>5.9</td>
<td>85</td>
</tr>
<tr>
<td><em>Hipp1</em>+/+, *su(Hw)*2/+</td>
<td>5.7</td>
<td>140</td>
</tr>
<tr>
<td><em>Hipp1</em>+/+, *su(Hw)*2/Pb</td>
<td>100.0</td>
<td>71</td>
</tr>
<tr>
<td><em>Hipp1</em>ΔDsR/+, *su(Hw)*2/+</td>
<td>0.0</td>
<td>204</td>
</tr>
<tr>
<td><em>Hipp1</em>Δ37/+, *su(Hw)*Pb/+</td>
<td>0.9</td>
<td>217</td>
</tr>
<tr>
<td><em>Hipp1</em>ΔDsR/Δ37, <em>su(Hw)</em>+/+</td>
<td>2.0</td>
<td>344</td>
</tr>
<tr>
<td><em>Hipp1</em>ΔDsR/Δ37, *su(Hw)*2/+</td>
<td>6.6</td>
<td>91</td>
</tr>
<tr>
<td><em>Hipp1</em>ΔDsR/Δ37, *su(Hw)*2/Pb</td>
<td>100.0</td>
<td>42</td>
</tr>
</tbody>
</table>
Table 3. 4 Summary of insulator proteins bound to ChIP sites

<table>
<thead>
<tr>
<th>Class</th>
<th>Site</th>
<th>IBPs bound</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Su(Hw)</td>
</tr>
<tr>
<td>Negative Controls</td>
<td>1A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>13F</td>
<td></td>
</tr>
<tr>
<td></td>
<td>64A</td>
<td></td>
</tr>
<tr>
<td>Su(Hw) only</td>
<td>5C</td>
<td></td>
</tr>
<tr>
<td></td>
<td>23A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>61F</td>
<td></td>
</tr>
<tr>
<td></td>
<td>69D</td>
<td></td>
</tr>
<tr>
<td></td>
<td>70C</td>
<td></td>
</tr>
<tr>
<td></td>
<td>100B</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CG6282</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CG9813</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CG15760</td>
<td></td>
</tr>
<tr>
<td></td>
<td>dimm</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Or35a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SSE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TipE</td>
<td></td>
</tr>
<tr>
<td>IBPs + Su(Hw)</td>
<td>3D</td>
<td></td>
</tr>
<tr>
<td></td>
<td>16C1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>35B</td>
<td></td>
</tr>
<tr>
<td></td>
<td>59F</td>
<td></td>
</tr>
<tr>
<td></td>
<td>62D</td>
<td></td>
</tr>
<tr>
<td></td>
<td>82E</td>
<td></td>
</tr>
<tr>
<td></td>
<td>84B</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Jhe</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CG3104</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mob2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rbp9</td>
<td></td>
</tr>
<tr>
<td>IBPs - Su(Hw)</td>
<td>16C8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>16D</td>
<td></td>
</tr>
<tr>
<td></td>
<td>44C</td>
<td></td>
</tr>
<tr>
<td></td>
<td>63E</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 4: DISCUSSION AND FUTURE DIRECTIONS

Insulator proteins are a diverse group of proteins that are often multifunctional. They possess the ability to block enhancers and the spread of repressive chromatin. Insulator protein binding sites are enriched at physical boundaries as well points of transition in chromatin state. However, recent studies of chromatin organization have found a limited role for the insulator proteins tested in global chromatin organization (Zuin et al. 2014; Schwarzer et al. 2017; Ramirez et al. 2018). Instead, most insulators seem to function primarily as local insulators or as transcriptional activators or repressors (Jiang et al. 2009; Soshnev et al. 2013; Nora et al. 2017). In particular, my investigations focused on using Su(Hw) as a model insulator protein to study the multifunctional nature shared by many of these proteins. Su(Hw) has well characterized insulator functions at gypsy, however the primary function of Su(Hw) in the Drosophila genome is as a transcriptional repressor of neuronal genes in the ovaries and testes (Soshnev et al. 2013; Duan and Geyer 2018). Despite the role of Su(Hw) as a repressor in the ovary, of the thousands of binding sites for Su(Hw) near genes, only about 3% are putative direct regulators of expression (Soshnev 2012). It is therefore unclear how the regulatory output of Su(Hw) at individual Su(Hw) binding sites in the genome is specified. In my studies I used Su(Hw) as a model multifunctional transcription factor to study mechanisms specifying regulatory output.

The investigations described in this thesis were aimed at elucidating the mechanisms of the non-insulator function of Su(Hw) through study of a novel interacting protein HIPPI1. Here, I investigated the role of HIPPI1 in development and in multiple functions of Su(Hw). These studies revealed that that Hipp1 is a non-essential gene, and that it was dispensable for both female and male fertility. Investigations of HIPPI1 expression revealed that it is more broadly
expressed than Su(Hw), being detectable in several cell types within the ovary and testis in which Su(Hw) is undetectable. Interestingly, it has a complementary expression pattern to Su(Hw) in the larval central brain. Further, these studies revealed a previously undescribed *Hipp1* transcript in the testis that gives rise to a truncated protein with undetermined function. HIPPI recruitment to chromatin is primarily Su(Hw) dependent, even when other insulator proteins are present. Despite this, the function of HIPPI at these sites remains unclear, as loss of HIPPI did not result in changes of Su(Hw) regulated genes in the ovary. Further studies are needed to elucidate the function of HIPPI and the mechanism of Su(Hw) repressor activity.

**HIPPI is a non-essential gene**

Loss of CDYL is early post-natal lethal in mice (WAN *et al.* 2013). To test whether HIPPI was required in development, *Hipp1* null alleles were generated and the viability of *Hipp1*−/− flies was tested. Somewhat surprisingly, loss of HIPPI had no effect on viability and the flies appeared morphologically normal. This may reflect loss of the essential function of CDYL in HIPPI, or there might be functional redundancy with other proteins in Drosophila. There are 17 predicted crotonase-like superfamily members in Drosophila (Uniprot database). Putative paralogs with significant homology to the CDYL crotonase-like fold domain are shown in Figure 4.1. These include several proteins with putative human orthologs other than CDYL, including CG13890 (34% identity to human PECI) and CG9577 (45% identity to human ECH1.), CG6543 (59% identity to human ECH1.), CG6984 (42% identity to human ECHDC3.), CG8778 (60% identity to human AUH.), and CG4594 (40% identity to human ECI1.) (Hu *et al.* 2011; Faust *et al.* 2012). The human proteins ECH1 and PECI primarily localize to the mitochondria and peroxisomes respectively, the localization of the other human proteins is unclear (FitzPatrick *et al.* 1995; Geisbrecht *et al.* 1999). Another putative crotonase-like fold domain protein,
CG5844, had no clear human orthologue. A study investigating the network of transcription factor interacting proteins in Drosophila examined Su(Hw) interactors by immunoprecipitation followed by mass spec. That study found only HIPPI among crotonase-like fold domain superfamily members as a significant interactor of Su(hw) (Rhee et al. 2014). It is therefore unlikely that any other crotonase-like fold domain proteins compensate for loss of HIPPI at Su(Hw) sites. However, it cannot be ruled out that HIPPI may compete with other crotonase proteins or other non-crotonase Su(Hw) cofactors whose interaction with Su(Hw) may be affected in a Hipp1−/− background. Further, the structures of the genes encoding other Drosophila crotonase-like fold domain proteins suggests that they are more distantly related to Cdyl than Hipp1. Hipp1 and Cdyl share the same relative splice junction positions within the region encoding the crotonase-like domain, suggesting these genes are orthologs. CG13890 is the only other gene encoding a crotonase-like domain protein that shares a splice junction position with Cdyl, however the protein is more strongly conserved with human PECI, suggesting that it is a more distant relative of Cdyl. However, we also cannot rule out that some other developmentally required function of HIPPI is compensated for by these proteins. There are a number of RNAi and putative mutation lines caused by P-element insertions for the non-HIPPI crotonase-like domain proteins. However, they have not been validated. To address the question of redundancy in the future, I would use the scarless tagging technique that was employed for Rbp9 and Hipp1. By inserting GFP-piggyBac[DsRed] at the translation start site, the initial line should be a putative null that would be validated by qPCR (See Figures 2.1 and 2.2). Subsequent jumping of the piggyBac element to bring the GFP tag in frame with the coding sequence would then generate alleles that could be used to look at the cellular localization of each protein. Alternatively, a double cut method could be used to delete the coding sequence. Such null alleles
would be tested for genetic interaction with HIPPI in affecting viability to determine whether they may compensate for HIPPI loss during development.

**HIPPI is not localized with HP1 in most cells**

As part of these studies, I examined whether HIPPI and HP1a colocalized in multiple tissues and found that in the cells in which HP1a forms foci, HIPPI was largely absent from HP1a foci. (Figure 3.10). This was surprising given the extensive colocalization of HIPPI with HP1a defined by ChIP-seq in S2 cells. However, it should be noted that in a different study of HP1a interactors, a majority (80%) of the proteins identified as HP1a interactors by IP-mass spec failed to colocalize with HP1a. Further, those that did colocalize showed a diverse range of colocalization patterns, with many showing a diffuse staining pattern like that seen for HIPPI (SWENSON et al. 2016). HP1a interacting proteins also show variations in their binding in different stages of the cell cycle (PLATERO et al. 1998; SWENSON et al. 2016). The tissues examined for colocalization are primarily in interphase. Thus, it remains possible that HIPPI only extensively colocalizes with HP1 during certain stages of the cell cycle, or that the weak localization of HIPPI to heterochromatin implied by the diffuse staining of HIPPI observed during interphase is sufficient to give rise to the signal observed by ChIP in S2 cells (ALEKSEYENKO et al. 2014). Examining HIPPI-GFP localization relative to HP1a in cells positive for H3S10p, or doing live imaging of embryos comparing HIPPI-GFP and RFP-H3 would determine whether HIPPI localizes to chromatin during mitosis. The interaction of HIPPI with multiple insulator proteins, some which do and some that do not localize to chromatin during mitosis, makes the localization of HIPPI during this stage of the cell cycle hard to predict. While CTCF persists on chromosomes through mitosis, Su(Hw) is absent, and HP1a is removed during early prophase and reloaded during anaphase/telophase (SWENSON et al. 2016).
In the case of CDYL, binding of methylated lysine by the chromodomain is required for CDYL localization to heterochromatin. Examination of the HIPP1 protein sequence revealed that HIPP1 does not have any regions of significant homology to chromodomains. The chromodomains of Human and Drosophila HP1 show significant conservation (47 of their 59 residues align at 68% identical, 82% similarity), suggesting that the lack of identifiable homology is not merely due to evolutionary distance, as functional chromodomains are strongly conserved among orthologs. The lack of a chromodomain suggests that HIPP1 has an alternative mechanism for interaction with heterochromatin. The chromo-shadow domain of HP1a functions as a protein-protein interaction domain (LECHNER et al. 2005). This domain has specificity for binding proteins with a PxVxL motif, where X is any amino acid (SMOTHERS AND HENIKOFF 2000). Notably, the chromo-shadow domain also interacts variant motifs including VxVxL and LxVxL of some HP1 interacting proteins (LECHNER et al. 2005; MENDEZ et al. 2011). While CDYLb does not contain these motifs, HIPP1 possesses two candidate variant interaction motifs, VKVML (a.a. 335-339) and LAVAL (a.a. 752-756). Examination of the conservation of these sequences across drosophilids revealed that while the second site was not conserved, the first motif was conserved as either the VxVxL variant or the LxVxL variant in eleven out of twelve Drosophila species examined, suggesting that it might represent a conserved HP1a interaction motif. The absence of these motifs in CDYL and their presence in HIPP1 suggests that there might have been a switch in the mechanism by which HIPPI/CDYL localizes to heterochromatin. Testing what regions of HIPPI are required for HP1a interaction in vitro could allow this hypothesis to be tested.

HIPP1 localizes primarily to autosomal heterochromatin, with ChIP-seq reads being distributed across heterochromatic regions (ALEKSEYENKO et al. 2014). Given the association
between HP1a and HIP1, the effect of HIP1 loss on repression of genes inserted into heterochromatin was determined by examining position effect variegation. Loss of a single copy of Hipp1 did not result in changes in either Sb', white, or yellow genes that had been inserted into heterochromatin. Further, loss of both copies of Hipp1 did not result in changes in white expression and seemed to result in a mild enhancement of yellow variegation for the Y chromosome lines. (Figure 3.11) The lack suppression of variegation with loss of Hipp1 is consistent with the minimal co-localization of HIP1 and HP1a observed in multiple cell types (Figure 3.10) and suggests that HIP1 is not required for repression in centric heterochromatin or on the Y chromosome. Notably, many genes within heterochromatin require HP1 for expression, with their expression decreasing upon HP1 loss (Wakimoto and Hearn 1990; Howe et al. 1995; Lu et al. 2000). Similarly, HP1a is required for the expression of a subset of euchromatic genes that bind HP1a (Piacentini et al. 2003; Cryderman et al. 2005; Cryderman et al. 2011). While HIP1 and HP1a euchromatic ChIP peaks do not significantly overlap, one proposed mechanism for HP1 activator function is through binding of transcripts, with localization being defined by RNA immunoprecipitation (Piacentini et al. 2009). It is unclear if HIP1 may play a role in the activator function HP1, but this could be tested by looking at expression of HP1a regulated genes in Hipp1 null flies.

Finally, it should be noted that other studies of HP1a interactors have not identified HIP1 (Rhee et al. 2014; Ryu et al. 2014; Swenson et al. 2016). This may be due to differences in pulldown procedure, as the Kuroda group fixed their cells before pulldown (Alekseyenko et al. 2014). It should also be noted that when HIP1 was pulled down in Alekseyenko et al., peptide counts for HP1a were only ~2 fold enriched over input, indicating that pull down of HIP1 only co-precipitated a small fraction of total HP1a in the cell. This suggests that HIP1-
HP1a interactions occur more rarely or are weaker affinity than compared to the interaction with Su(Hw), which had a lower bound enrichment of 4.9 fold. Together these findings support the view that HIPP1-HP1a interactions are either rare or weak, in agreement with our immunohistochemical analysis showing minimal overlap.

**Su(Hw) is required for HIPPI localization to SBSs**

Our analysis of HIPPI binding in the ovary revealed that the majority of HIPPI binding to SBSs requires Su(Hw), even at sites bound by multiple insulator proteins. However, there was a distinct lack of HIPPI binding at the Su(Hw) only SBSs of target genes (Figure 3.15B). Whether or not this is functionally important is still unclear. Since HIPPI binding was mapped in S2 cells, it is possible that HIPPI has tissue specific binding patterns. To test this, HIPPI versus Su(Hw) binding was analyzed in *Hipp1GFP/GFP* flies and these studies revealed that the levels of HIPPI and Su(Hw) binding correlated in most cases with a couple exceptions (Figure 3.16). The SBSs at *CG6282* and 23A were devoid of HIPPI binding, despite having a similar level of Su(Hw) binding as the SBS at 69D. It has been noted that Su(Hw) has two distinct populations of sites, those that bind CP190 and Mod67.2, and those that bind Su(Hw) only (SOSHNEV et al. 2012). Su(Hw) binding sites that do and do not bind CP190 were examined for DNA sequences enriched in either class and none were able to be identified (SCHWARTZ et al. 2012). Notably, preliminary analysis of CP190 versus HIPPI binding in the ovary revealed that sites absent for HIPPI binding also lacked CP190 binding, despite the SBS at CG6282 binding HIPPI and SBS at 23A binding both HIPPI and CP190 in other cell types (ALEKSEYENKO et al. 2014). This limited number of sites is not enough to determine whether CP190 and HIPPI recruitment are linked in the ovary, further sites need to be tested. If CP190 and HIPPI binding correlate in the ovary, the relationship between CP190 and HIPPI could be further explored. *CP190*– flies are
pupal lethal (PAl et al. 2004). However, ChIP for HIPPI-GFP from \textit{CP190} wildtype and null larva could be used to determine the effect of CP190 loss on HIPPI recruitment. This would need to be paired with ChIP for Su(Hw) to ensure that Su(Hw) is not being lost in this background, as CP190 has been reported to be contribute to Su(Hw) binding (Schwartz et al. 2012). If CP190 facilitates HIPPI recruitment, then loss of CP190 is expected to decrease HIPPI binding at SBSs. It should be noted that around 29% of Su(Hw) sites positive for HIPPI in S2 cells also lack CP190 binding in S2 cells (Figure 3.1), suggesting that HIPPI is able to interact with Su(Hw) independent of CP190. It is therefore not expected that HIPPI binding would be lost from all SBSs in a \textit{CP190} mutant background.

While these studies indicate a stronger recruitment of HIPPI to non-target gene sites as opposed to target sites by ChIP (Figure 3.15), the level of HIPPI binding at both SBSs and non-SBSs in Figure 3.15 should be interpreted on a site by site basis. The selection of non-target sites was intentionally biased towards strong Su(Hw) binding sites, which tend to be sites with all three modules of the Su(Hw) consensus and show increased enrichment for CP190 and Mod67.2 (Baxley et al. 2017). In contrast, the target gene SBSs analyzed by ChIP were selected based on whether the regulated gene had previously been validated to change by qPCR and whether there was only a single candidate SBS within 2kb for testing. Mob2 was an exception to this rule with two SBSs due to its changes in gene expression. The second Mob2 SBS does not recruit HIPPI in S2 cells or the ovary (Alekseyenko et al. 2014) (data not shown). Based on the possible link between CP190 and HIPPI binding in the ovary, and the link between Su(Hw) and HIPPI binding in S2 cells, the selection of high Su(Hw) occupancy non-target gene SBSs likely resulted in a bias towards stronger HIPPI binding than at target gene SBSs. Due to the biased selection, it cannot be said for certain whether there is a significant depletion of HIPPI binding at target gene
SBSs as compared to SBSs on average. Performing ChIP-seq of HIPPI in the ovary and comparing HIPPI occupancy at SBSs around Su(Hw) regulated genes versus non-target SBSs could be used to determine if there is a link between Su(Hw) repression of genes and the absence of HIPPI. It is interesting to note that a previous study of a limited number of sites found that SBSs that did not bind CP190 and Mod67.2 sites tended to have repressor function in a transgene assay whereas CP190 and Mod67.2 binding SBSs did not (SCHWARTZ et al. 2012). It is possible that HIPPI may inhibit Su(Hw) repressor function, and that it is lost from target gene SBSs in the ovary to allow for Su(Hw) dependent repression of these genes. Though HIPPI seems unlikely based on homology to CDYL to be an inhibitor of Su(Hw) repressor function, if HIPPI does act as an inhibitor, then loss of HIPPI may result in downregulation of non-target genes with nearby Su(Hw) binding sites. This relationship could be examined by RNA-seq analysis of gene expression in a Hipp1−/− background. Expression of downregulated genes with a nearby Su(Hw) binding site could then be examined in a su(Hw)−/−, Hipp1−/− background to determine if this repression was due to Su(Hw).

As part of these studies we tested the expression of a variety of genes that were previously validated to have differential expression in the ovary upon loss of Su(Hw). However, as our ChIP studies and the S2 cell studies indicate, HIPPI is not restricted to binding Su(Hw) sites. HIPPI was maintained on the 4/4 non-Su(Hw) insulator binding protein sites tested in the absence of Su(Hw). In fact, there was even a slight but significant increase at 16C8 (Figure 3.15). Further, preliminary polytene staining data reveals that HIPPI remains bound to many sites in the genome in su(Hw)−/− (Figure 4.3). Though much of the banding pattern is lost in su(Hw)−/− and there seems to be an increase in diffuse HIPPI staining and clear banding is at some sites. The resolution of polytene staining prevents identification of these sites, however
they presumably correspond to the hundreds of euchromatic HIPPI sites that lack Su(Hw) (Figure 3.1). Notably, the diffuse staining also overlaps the HP1a positive chromocenter, suggesting that HIPPI may have increased binding at heterochromatin upon Su(Hw) loss. These data suggest that Su(Hw) might sequester HIPPI away from non-Su(Hw) binding sites. Aberrant HIPPI recruitment to sites could potentially contribute to the mis-regulation of the ~100 upregulated or the ~100 downregulated genes that change in a su(Hw) mutant background but lack a nearby Su(Hw) binding sites (Soshnev et al. 2013). This hypothesis could be tested by looking for candidate non Su(Hw) insulator sites at non-target genes and testing for changes in HIPPI occupancy in a su(Hw) mutant background. This analysis would be paired with gene expression analysis in a su(Hw)/, HIPPI/ mutant background. If mis-localization of HIPPI in a su(Hw) mutant contributes to the mis-regulation of non-target genes, then loss of both might restore gene expression to wildtype levels.

HIPPI is recruited by Su(Hw) through a yet unknown mechanism. The studies completed so far do not determine whether the interaction between Su(Hw) is direct or indirect. The retention of HIPPI at non-Su(Hw) insulator protein binding sites (Figure 3.15B), as well as the interaction of HIPPI with multiple insulator proteins suggest that HIPPI likely directly interacts with at least one other insulator binding protein. HIPPI localization to foci containing CP190 and Mod67.2 in late stage oocytes, where Su(Hw) is undetectable, further argues for HIPPI interaction with a non-Su(Hw) insulator protein (Figure 3.7). The most parsimonious answer would be for HIPPI to interact with CP190, as it was found as one of the top HIPPI interactors and CP190 interacts with all the other insulator proteins pulled down with HIPPI. (Gerasimova et al. 2007; Alekseyenko et al. 2014; Cuartero et al. 2014; Maksimenko et al. 2015).

Notably, there are regions of high conservation throughout the N-terminus of HIPPI across
different drosophilid species (Figure 4.2). This suggests that there is a strong selection pressure for them to remain unchanged over 40 million years of evolution (D. melanogaster to D. virilis).

There is also nothing resembling these sequences in CDYL or the rest of the human or Drosophila proteome suggesting they recently arose since the divergence of HIPPI and the CDY family ancestral protein. One of these regions may be a Su(Hw) or CP190 interaction domain. Mapping these putative interaction domains may be carried out in vitro using bacterially expressed fragments of HIPPI and looking for interaction with Su(Hw) and CP190. I have generated expression plasmids that express different fragments of HIPPI tagged with GST to look for potential direct interactions. These studies have yet to be performed and remain the subject of future investigations.

**HIPPI is not required for Su(Hw) repressor function**

The experiments performed in these studies demonstrated that HIPPI is not required for Su(Hw) repression of target genes involved in su(Hw)^− induced loss of fertility. Thus, the function of HIPPI remains unclear. Because HIPPI is not recruited to the Su(Hw) binding sites near the target genes that were tested for changes in gene expression, it remains possible that HIPPI is required for the regulation of only a limited subset of genes that were not analyzed. To better answer this question, it is prudent to determine whether recruitment of HIPPI itself can mediate repression of a gene. While CDYL was shown to be capable of directing transcriptional repression (CARON et al. 2003), HIPPI has not. Creating a gal4(DBD)- HIPPI fusion construct and expressing it in UAS-lacZ reporter line could allow for the determination of whether HIPPI may direct repression of gene expression. If HIPPI is capable of mediating repression, then the role of HIPPI in Su(Hw) dependent repression could be further explored by looking at gene expression changes in Hipp1 null flies in an unbiased approach by RNA-seq. Despite the lack of
phenotype for \(Hipp1^{-}\) in ovary, this tissue would be used to enable matching gene expression and ChIP data for Su(Hw). RNA-seq would also allow exploration of potential alternative functions of HIPPI, such as regulation of HP1a activated or repressed genes.

The lack of expression changes in Su(Hw) target genes in a \(Hipp1\) null ovaries indicates that HIPPI is not required for repression of Su(Hw) target genes, thus the mechanism of Su(Hw) repressor function remains unclear. Drosophila Lethal (3) malignant brain tumor [L(3)mbt] was identified as a Su(Hw) and HIPPI interacting protein by immunoprecipitation followed by mass spectrometry experiments (GuruHarsha et al. 2011; Rhee et al. 2014) L(3)mbt is a 1477 amino acid nuclear protein that is ubiquitously expressed in Drosophila. L(3)mbt contains three MBT domains that bind mono or dimethylated histone tails (Trojer et al. 2007). L(3)mbt is part of two repressive chromatin complexes: The RBF, E2F2, and Myb-interaction proteins (dREAM complex), as well as the L(3)mbt-interaction complex (LINT complex) consisting of L(3)Mbt, Lin-1, and the co-repressor CoREST. Both complexes have been shown to repress developmental gene expression (Georlette et al. 2007; Meier et al. 2012). L(3)mbt, as part of the LINT complex, represses the expression of neuronal and testis specific genes in female germline cells (Coux et al. 2018). Given the similarity of this function to that of Su(Hw) in the ovary (Soshnev et al. 2013), L(3)mbt represents another candidate partner of Su(Hw) that might facilitate repression. L(3)mbt binding has extensive overlap with CP190 and other promoter associated insulators, however it only binds a small portion (13%) of Su(Hw) sites (Richter et al. 2011). However, one of these sites is at \(Rbp9\), the critical neuronal gene regulated by Su(Hw), specifically at the SBS next to the second promoter of \(Rbp9\) that becomes active upon loss of Su(Hw). In line with this, pilot studies undertaken by Alexey Soshnev found that \(Rbp9\), along with a number of other Su(Hw) repressed targets, becomes upregulated in a \(l(3)mbt\) temperature
sensitive mutant line (SOSHNEV 2012). Similar to Su(Hw), loss of l(3)mbt specifically in the germline leads to increased egg chamber apoptosis (COUX et al. 2018). Together, these studies suggest that L(3)mbt may be another candidate for facilitating Su(Hw) function.

**Alternative functions of Su(Hw) that may be mediated by HIPPI1**

Although these investigations find that HIPPI1 is not required for Su(Hw) repressor or insulator functions, there is evidence Su(Hw) has other functions. One alternative function of Su(Hw) is in directing origin recognition complex (ORC) assembly. Su(Hw) is hypothesized to mediate ORC assembly at up to 6% of replication origin sites in Drosophila (VOROBYEVA et al. 2013), primarily in repressed BLACK chromatin. This in agreement with another recent study that showed Su(Hw) was important for the localization of ORC components at genomic regions amplified in follicular cells in early oogenesis (KRASNOV et al. 2018). The model put forth is that Su(Hw) recruits an acetyltransferase (SAGA) to acetylate histones, which are then bound by bromodomain containing proteins in the SWI/SNF chromatin remodeling complex, leading to chromatin remodeling and nucleosome depletion allowing for assembly of ORC components (VOROBYEVA et al. 2013). My investigations did not directly address whether HIPPI1 may a play a role in ORC component recruitment. However, H3 binding at HIPPI1 sites was analyzed as part of the investigation of the role of HIPPI1 in regulating crotonyl-lysine levels (Figure 3.14). The finding that all of the Su(Hw) binding sites tested (7/7) had unchanged H3 binding in a Hipp1 mutant is inconsistent with HIPPI1 regulating nucleosome localization at Su(Hw) binding sites. Notably, while there was a 2-fold increase in H3 binding upon loss of Su(Hw) at 62D (VOROBYEVA et al. 2013), no changes in H3 were observed at this site in a Hipp1−/− background (Fig. 3.14). It should be noted that control sites showed significant decreases in H3 upon Hipp1 loss, which makes interpretation of the data difficult. While these preliminary data and
observations trend against HIPPI playing a role in origin functions of Su(Hw), more studies with a direct comparison of H3 and replication complex components at SBSs in Hipp1−/−, su(Hw)−/−, and wild type backgrounds are needed.

Interestingly CDYL was also recently linked to replication as a bridge between chromatin assembly factor complex 1 (CAF-1), a histone chaperone complex that facilitates assembly of histones onto newly replicated DNA (Smith and Stillman 1989; Kamakaka et al. 1996), and the replicative helicase microchromosome maintenance (MCM) complex responsible for unwinding double stranded DNA at the replication fork (Bochman and Schwacha 2009; Liu et al. 2017b). MCM is enriched at Su(Hw) and other insulator sites (Vorobyeva et al. 2013). Analysis of HIPPI binding in S2 cells (Alekseyenko et al. 2014) as well as CTCF and Su(Hw) binding from modEncode data (Celniker et al. 2009) reveals a slight enrichment for MCM components at HIPPI positive Su(Hw) sites (14% of SBSs) versus HIPPI negative Su(Hw) sites (6% of SBSs). Similarly, at non-Su(Hw) CTCF sites, MCM components are found at 49% of HIPPI positive CTCF sites, compared to 35% of HIPPI negative CTCF sites. Notably, CDYL has been found to interact with MCM components however the region of CDYL found to associate with MCM4 (a.a. 211-309) only partially overlaps the region of CDYl conserved in HIPPI (a.a. 290-544) (Liu et al. 2017b). It remains to be tested whether HIPPI can facilitate MCM recruitment. This could be tested by examining MCM recruitment to Su(Hw) sites in a Hipp1−/− mutant background.

**Exploring complementary mechanisms of directing Su(Hw) regulatory output**

The “Su(Hw) code” model posits that binding site sequence might direct alternative regulatory functions of Su(Hw) (Baxley et al. 2017). In order to test this model a GFP tagged Rbp9 allele was generated as outlined in Chapter 2. Regulation of Rbp9 is independent of known
insulator cofactors and is an essential target of regulation for fertility (Soshnev et al. 2013). The Rbp9 gene has three promoters, of which transcription start site 3 (TSS3) is the major promoter used in the ovary with low expression from TSS1 and TSS2 (see Figure 2.2 for a structure of Rbp9). In su(Hw) mutants, TSS2 becomes active increasing TSS2 expression eleven fold, with combined TSS1, TSS2, and TSS3 expression increasing three fold (Soshnev et al. 2013). This allows for comparison of the effects of any manipulations of Su(Hw) binding sites on multiple nearby promoters, one Su(Hw) regulated and two Su(Hw) non-regulated promoters. The Rbp9 5’ untranslated region contains three SBSs positioned in-between TSS1 and 2, at TSS2, and between TSS2 and 3. By anchoring a qPCR primer within the GFP coding sequence and using CRISPR to manipulate SBS sequence within the context of the GFP allele, Rbp9GFP provides a means to monitor the effect of binding site sequence on the regulatory output of Su(Hw). It also allows for allele specific immunohistochemical analysis as well. Each binding site within the Rbp9 promoter region could be eliminated by CRISPR to determine which one(s) are required for regulation, and subsequently swapping in “insulator class” and “transcriptional regulatory” class sites of equal affinity for Su(Hw) could be performed to determine what effect sequence has on function.

The combination of sites at Rbp9 also provides the opportunity to explore the possibility of binding site orientation in Su(Hw) regulation of Rbp9. It has been repeatedly shown that CTCF loop formation is tightly linked to binding site orientation, with the majority of loops containing CTCF binding sites in convergent orientation (Rao et al. 2014; Bonev et al. 2017). Further, changing binding site orientation disrupts loop formation and can lead to aberrant enhancer-promoter interactions (De Wit et al. 2015). It is unlikely that Su(Hw) sites at Rbp9 form loops that are required for regulation, due to the lack of requirement for CP190 or
Mod67.2. However, CTCF activator function has been linked to binding site orientation, with 90% of promoters with single CTCF sites possessing a binding site in direct orientation with the direction of transcription (Nora et al. 2017). There may be a similar orientation requirement for Su(Hw) repressor function. Studies in flies have also revealed a dependence on orientation for some insulator interactions, suggesting this is a conserved property of insulators (Kyrchanova et al. 2008). By changing the orientation of individual binding sites within the Rbp9 promoter region, the contribution of this aspect of Su(Hw) binding could also be explored. Together these experiments would clarify the role of SBS sequence in directing Su(Hw) regulatory output.
Figure 4. 1 Alignment of CDYL1b to Drosophila Crotonase-like domain family members

Shown is an alignment of Human CDYL1b and fly proteins of the Crotonase-like superfamily identified as putative orthologs by BLAST. The bounds of homology (275-544) were defined as the region of CDYL1b with significant alignment (Blast E value < 0.05) to any of the fly proteins listed. Fly and protein sequences were aligned by Clustal and this alignment generated the above graphic using an in-house python script. The positions of identical (blue) and similar (red) residues to the corresponding position in CDYL1b within the bounds of homology are indicated. Sequences unique to each protein are shown in black. Green arrows indicate the position of the amino acids encoded by nucleotides at the splice junction of each gene.
Figure 4. 2 HIPP1 has pockets of strong conservation outside the crotonase-like domain

Shown is an alignment of HIPP1 from different Drosophila species of different evolutionary distances from *D. melanogaster* (millions of years of evolution). The position of the crotonase-like domain is indicated above, and the alignment is color coded based on identical (blue), similar (red), or unique (black) residues relative to the corresponding position in HIPP1 from *melanogaster*. Asterisks indicate regions of strong homology from *D. melanogaster* to *D. virilis.*
Figure 4. 3 HIPPI associates with chromosomes in the absence of Su(Hw)

A. Shown are confocal images of polytene chromosomes isolated wandering third instar larva from Hipp1\textsuperscript{GFP/+}, su(Hw)\textsuperscript{+/-} stained with DAPI (white) and antibodies against HP1a (blue), Su(Hw) (red), and GFP (HIPPI, green). B. Confocal polytene images from Hipp1\textsuperscript{GFP/+}, su(Hw)\textsuperscript{2/Pb} stained with DAPI and antibodies against HP1a (blue) and GFP (HIPPI, green).
REFERENCES


Bartkuhn, M., T. Straub, M. Herold, M. Herrmann, C. Rathke et al., 2009 Active promoters and insulators are marked by the centrosomal protein 190. Embo j 28: 877-888.


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


Geyer, P. K., M. M. Green and V. G. Corces, 1988 Reversion of a gypsy-induced mutation at the yellow (y) locus of Drosophila melanogaster is associated with the insertion of a newly defined


Golovnin, A., A. Mazur, M. Kopantseva, M. Kurshakova, P. V. Gulak et al., 2007 Integrity of the Mod(mdg4)-67.2 BTB domain is critical to insulator function in Drosophila melanogaster. Mol Cell Biol 27: 963-974.

Golovnin, A., L. Melnikova, I. Volkov, M. Kostuchenko, A. V. Galkin et al., 2008 'Insulator bodies' are aggregates of proteins but not of insulators. EMBO Rep 9: 440-444.


Kuhn-Parnell, E. J., C. Helou, D. J. Marion, B. L. Gilmore, T. J. Parnell et al., 2008 Investigation of the properties of non-gypsy suppressor of hairy-wing-binding sites. Genetics 179: 1263-1273.


Kurshakova, M., O. Maksmenko, A. Golovnin, M. Pulina, S. Georgieva et al., 2007 Evolutionarily conserved E(y)2/Sus1 protein is essential for the barrier activity of Su(Hw)-dependent insulators in Drosophila. Mol Cell 27: 332-338.


Lobanenkov, V. V., R. H. Nicolas, V. V. Adler, H. Paterson, E. M. Klenova et al., 1990 A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene. Oncogene 5: 1743-1753.


Melnikova, L., M. Kostyuchenko, A. Parshikov, P. Georgiev and A. Golovnin, 2018 Role of Su(Hw) zinc finger 10 and interaction with CP190 and Mod(mdg4) proteins in recruiting the Su(Hw) complex to chromatin sites in Drosophila. PLoS One 13: e0193497.


Moon, H., G. Filippova, D. Loukinov, E. Pugacheva, Q. Chen et al., 2005 CTCF is conserved from Drosophila to humans and confers enhancer blocking of the Fab-8 insulator. EMBO Rep 6: 165-170.


