Modifiers of *P*-element-dependent silencing in *Drosophila melanogaster.*

by

Allen Thomas Melbourne McCracken

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Molecular Biology and Genetics

Department of Biological Sciences

Spring 2013
Edmonton, Alberta

Abstract

If all cells in a multicellular organism contain exactly the same genetic information, the question arises of how tissue types with distinct gene expression profiles are formed and maintained over the life of the organism. These different temporal and spatial gene expression patterns are thought to be built by activating and repressing proteins and RNAs to create self-perpetuating chromatin states. Identifying these components is the first and fundamental step in understanding this type of control of gene expression, and is the focus of this thesis. My model system centers on *P{lacW}ci*^{Dplac} , a *white* (*w*$^{+}$) transgene insert on chromosome 4 of *Drosophila melanogaster.* *P{lacW}ci*^{Dplac} is a previously characterized enhancer trap of *ci*^{D} that should be sensitive to many of the proteins that regulate *ci* during development. Normally, the *white* gene within *P{lacW}ci*^{Dplac} is expressed throughout the adult eye and presents a uniform red eye phenotype. However, the presence of other *P* elements results in stochastic silencing of the *w*$^{+}$ of this transgene and a variegated phenotype in a process called *P* element dependent silencing (PDS). A derivative allele of *P{lacW}ciDplac* was isolated, called *E1*, that contained a distal *gypsy* element insertion. This allele variegates in the absence of other *P* elements, and the variegating phenotype can be suppressed by and enhanced by modifiers of *w*^{m4} in a manner similar to heterochromatic Position Effect Variegation (hPEV). I performed a genetic screen for modifiers of *E1* variegation and isolated mutations that fell into

complementation groups on both the second and third chromosomes. I identified 5 of these groups as: TAF4, a general transcription factor; *cg*, an already characterized regulator of *ci*; *ash1* and *trx*, known regulators of homeotic genes not previously shown to act at *ci*; and *CG8878*, a putative protein kinase of unknown specificity. I also isolated a complementation group that was too weak in phenotype to accurately map via recombination and several singles, which were not pursued farther. I chose to investigate the alleles of *ash1*, *trx*, and *CG8878*. This thesis describes their generation, isolation and further characterization.

Acknowledgements

I would like to thank Dr. John Locke and Scott Hanna for their support, criticisms, valuable discussion and expertise during my graduate studies. I thank NSERC for Dr. Locke's Discovery Grant that supported this research; my wife, Tanya Lee-McCracken for assistance in stock keeping; MBSU for technical support, and the University of Alberta Department of BioSci for TA Support.

Table of Contents

List of Figures

List of Tables

List of Symbols and Abbreviations

+: wild type allele

Δ: delta (deletion)

(*2R*): chromosome 2 right arm

(*2Rh*): chromosome 2 right arm heterochromatin

A: Adenine or Alanine depending on context

aa: amino acid(s)

ADP: Adenosine-diphosphate

Amp^R: *Ampicillin resistance* gene

ap^{Xa}: *Xasta* allele of *apterous* gene

ash1: *absent, small or homeotic discs 1* gene

AWS: associated with SET

BLAST: Basic Local Alignment Search Tool

bp: base pairs

BSA: bovine serum albumin

Bl: *Bristle* gene

brm: *brahma* gene

bw: *brown* gene

bw^D: Dominant allele of the *brown* gene

bxd: *bithoraxoid* gene

C- carboxy, Cytosine or Cysteine depending on context

^oC: degrees Celcius

Cam^R: *Chloramphenicol resistance* gene

CBP: *CREB binding protein*

CD3; *Cluster of Differentiation 3* gene

CFTR: Cystic fibrosis transmembrane conductance regulator

cg: *combgap* gene

CG8878: Celera Genetics predicted gene 8878

ChIP: Chromatin Immuno recipitation

ci: *cubitus interruptus* gene

ci^{D}: *Dominant* allele of *cubitus interruptus* gene

cM: centi-Morgans

CpG: Cytosine-phosphate-Guanine

Cy: *Curly* gene

CYP3A: Cytochrome P450, family 3, subfamily A gene

CyO: Curly of Oster chromosome

Δ2-3: *P(ry+, Δ2-3)99B*

D: Aspartic acid

DAM: DNA methyltransferase

dE2F: *Drosophila* E2F transcription factor

Df(2): chromosome 2 deficiency

Df(3): chromosome 3 deficiency

DNA: deoxyribonucleic acid

dp: *dumpy* gene

e: *ebony* gene

E: Glutamic acid

E1: *P{lacW}ci*^{E1}

E2: *P{lacW}ci*^{E2}

EMS: ethylmethane sulfonate

en: *engrailed* gene

esc: *extra sex combs* gene

E(var): Enhancer of variegation gene

E(var)3-93-E: *Enhancer of variegation 3-9* gene

E(var)3-9: *Enhancer of variegation 3-9* gene

E(z): *enhancer of zeste* gene

F-: Phenylalanine or Flag tagged

g: grams

G: Guanine or Glycine

GST: Glutathione S-transferase Gene Fusion System

GuHCl: Guanidine Hydrochloride

H: Histidine

HAT: Histone Acetyltransferase

H2a: Histone 2a

H3: Histone 3

H4: Histone 4

HDAC: Histone deacetylase

HMTase: Histone Methyl Transferase

HOX: Homeobox

HOXA9: *Homeobox A9* gene

HOXA10: *Homeobox A10* gene

HP1: heterochromatin protein 1

hPEV: heterochromatic Position Effect Variegation

hs: *heat shock*

hsp70: *heat shock protein 70*

hsp26: *heat shock protein 26*

hsp83: *heat shock protein 83*

Hu: *Humeral* allele of *Antennapedia* gene

I: Isoleucine

IgH: immuno-globulin heavy-chain

ING2: Inhibitor of growth protein 2

INM: inner nuclear membrane

In(1)wm^{4}: *Inversion (1) white-mottled 4*

IPTG: Isopropyl β-D-1-thiogalactopyranoside

K: Lysine

kb: kilobases

kDa: kilodaltons

KP : a *P* derived element caused by internally deleting nucleotides 808-2560

kis: *kismet* gene

L: liter or Leucine

Lsd1: Lysine specific deaminase 1 gene

M: molar or Methionine

me1: mono-methylated

me2 : di-methylated

me3: tri-methylated

mg: milligrams

miRNA: microRNA

ml: milliliters

MLL: *mixed-lineage leukemia* gene

mM: millimolar

mod(mdg4): *modifier of midgut 4* gene

mRNA: messenger RNA

mor: *moira* gene

MYND: (myeloid, Nervy, and DEAF-1) domain

µg: micrograms

µl: microliters

N- amino or Asparagine

NSD2: Nuclear receptor-binding SET domain protein 2

NURF: Nucleosome Remodelling Factor gene

p: phosphate

P: Proline

PABC1: *Polyadenylate-binding protein, cytoplasmic 1* gene

PBS: phosphate buffered saline

Pc: *Polycomb* gene

PcG: *Polycomb* Group

Pol II: Polymerase Two

PRC1,2: *Polycomb* Repressive Complex 1, 2

Pci: *P{lacW}ci*^{DPlac}

PCR: polymerase chain reaction

PDS: *P* element dependent silencing

PHD: Plant Homeo Domain

piwi: *piwi* gene

PRE: Polycomb response element

P-Sal: *P[SalI]89D*

*P-Sal**: *P-Sal* mutants

ptn^{D}: an allele of *mod(mdg4)*

Q: Glutamine

R: Arginine

Rb: *retinoblastoma* gene

RIGS: repeat induced gene silencing

RING: (Really Interesting New Gene) finger domain

RNA: ribonucleic acid

RNAi: ribonucleic acid interference

RpS3a: *Ribosomal protein S3A* gene

RT: reverse transcription

RT-PCR: reverse transcription polymerase chain reaction

ry: *rosy* gene

S: Sulphur, Synthesis or Serine depending on context

Sb: *Stubble* gene

Sce: *Sex combs extra* gene

SCF: Skp, Cullin, F-box containing complex

SETDB1: *SET domain, bifurcated 1*gene

Ser: *Serate* gene

siRNA: short interfering RNA

su(Hw): *suppressor of Hairy wing* gene

Su(var): Suppressor of variegation gene

Su(var)205: *Suppressor of variegation 205* gene

Su(var)3-3: *Suppressor of variegation 3-3* gene

Su(var)3-7: *Suppressor of variegation 3-7* gene

Su(var)3-9: *Suppressor of variegation 3-9* gene

SWI/SNF: SWItch/Sucrose NonFermentable yeast nucleosome remodeling complex

T: Thymine or Threonine depending on context

Taf4: *TBP-associated factor 4* gene

Tb: *Tubby* gene

Tb: Tubby phenotype

TM6: third multiply inverted chromosome 6

TM6B: third multiply inverted chromosome 6B

TRE: *trithorax* response element

trx: *trithorax* gene

TrxG: *trithorax* Group

Ubx: *Ultrabithorax* gene

UTR: untranslated region

V: Valine

vasa-IVS3-beta-geo;: a germline specific reporter construct

vg^{21-3}: an allele of the *vestigial* gene

W: Tryptophan

w: *white* gene

w^{+}: uniform red eye color phenotype

w^{m4}: *Inversion (1) white-mottled 4*

w^{+mC} : *mini-white*, a recombinant *w* gene produced by adding $w^{+}3.3(Xb, Sp)$ to $w^{+}0.8(Sp, H3)$

w^{var}: variegated eye color phenotype

wg^{Sp}: *wingless Sternopleural* gene

Y: Tyrosine

y: *yellow* gene

YNG1: Yeast homolog of mammalian ING1

Zn^{2+}: divalent Zinc cation

Chapter 1 – Introduction

Overview

All multi-cellular organisms start from a single celled zygote, which over the course of development gives rise to a variety of distinct cell types with specialized functions. Even though essentially all of these derived somatic cells have identical genotypes, they express only a subset of all the genes. Thus different cell types are produced due to differential gene expression; certain genes are expressed in some cells yet not in others. Some are regulated by environmental cues, while others are maintained in an active or inactive state after passing certain temporal windows during development. The identity and function of each cell type is determined by its individual gene expression profile, which must be faithfully reproduced after every cell division during development. This epigenetic establishment and maintenance of defined temporal and spatial gene expression patterns largely depends on the activities of activating and repressing proteins and RNAs in creating stable chromatin states. Identifying the components and understanding the mechanism of this control is fundamental to describing gene expression and being able to use these systems to our advantage.

As part of this characterization I have used the study of *P{lacW}ci*^{Dplac}, an enhancer trap of *ci*^{D} that was shown to mimic the expression pattern of *ci* (Eaton & Kornberg 1990) and thus should be sensitive to many of the proteins that normally regulate *ci* during development. Expression of *P{lacW}ci*^{Dplac} was found in this lab to be repressed by the presence of other *P* elements, *KP*s or other *P* derivatives capable of mimicing some of the characteristics of P cytotype such as modifying P- repressor sensitive alleles, but not enabling P element mobilization and transposition (*P*-like cytotype). During investigation of this phenomenon, called *P*-element dependent silencing (PDS), Dan Bushey, a former graduate student in the

lab, isolated two gypsy inserts upstream of *P{lacW}ci*^{Dplac} that resulted in *white* variegation in the absence of *P*-like cytotype. He showed that this variegation was sensitive to global levels of several of the major proteins involved in heterochromatic Position Effect Variegation (hPEV) and therefore in heterochromatin formation. One of these alleles of *P{lacW}ci*^{Dplac}, called *E1* was then used by me to screen for enhancers of PDS, and hopefully, therefore, for hPEV. I isolated several complementation groups consisting of: TAF4, a general transcription factor; *cg*, an already characterized regulator of *ci*; *ash1* and *trx*, known regulators of homeotic genes not previously shown to act at *ci*; and *CG8878*, encoding a putative Serine/Threonine/ Tyrosine kinase of unknown specificity. I also isolated a complementation group that was too weak in phenotype to accurately map via recombination and several mutants, which did not fall into complementation groups; these were not pursued further. I chose to investigate the alleles of *ash1*, *trx*, and *CG8878*; this thesis involves the work in their generation, isolation and further characterization.

Background

Epigenetics

Epigenetics comes from the Greek *epi* (*επ)*, meaning over or above, and *genetics* and is the study of heritable changes in phenotype or gene expression without corresponding changes to the gene itself as determined by the underlying nucleotide sequence. Examples of effectors of epigenetic changes are DNA modification such as CpG methylation in mammals, and histone modifications such as methylation of lysine and arginine residues, acetylation of lysine residues, phosphorylation of serine and threonine residues, biotinylation, sumoylation, and ubiquitination of lysine residues and ADP-ribosylation of lysine and glutamic acid residues. All of these modifications alter gene expression without mutating the DNA

sequence of the affected genes. Maintaining a stable inheritance pattern of gene expression is epigenetics (Allis *et al*. 2007).

Position Effect Variegation (PEV)

One of the first examples of an epigenetic effect was the discovery of heterochromatic position effect variegation (hPEV) of the *white (w)* gene in *Drosophila melanogaster* by H.J. Muller in 1930 (reviewed in Wakimoto 1998). The *white*$^{+}$ protein product is needed for transport of colourless pigment precursors into the eye, thus lack of w^{+} function results in a white eye, presenting a simple phenotypic assay for the expression of the w^{+} gene. In the well studied model system *Inversion(1)white-mottled 4*, (w^{m4}) the *Inversion(1)3C1-2;20F* places the normally euchromatic *white* gene near pericentric heterochromatin resulting in a variegated or mottled eye phenotype. Random silencing of *white* gene expression during tissue development is maintained in a stable state through multiple cell divisions causing clone-like inheritance and a mosaic pattern of gene expression. Since most (all?) *Drosophila* genes located in euchromatin can be silenced through heterochromatin formation, research into hPEV has provided insight into the functional differences between heterochromatin and euchromatin. This model system has identified the factors that help distinguish chromosomal domains, how these domains influence the transcriptional state of a gene, and the inheritance of that state, or epigenesis.

Chromatin can be subdivided into two basic categories, heterochromatin and euchromatin. These were originally determined cytologically by Heitz (1928) in the moss *Pellia epiphylla*. Certain parts of five of *Pellia epiphylla*'s nine chromosomes remained condensed throughout interphase while others become invisible at late telophase. Heitz termed these new, condensed autosomal structures as heterochromatin, and the decondensed regions as euchromatin. Gene

poor heterochromatin stained darkly with 45% carmine acetic acid indicating dense chromatin packing while gene rich euchromatin stained lightly indicating loose packing. Euchromatin replicates earlier in the cell cycle than does heterochromatin, is meiotic recombination active, has high gene density (which corresponds to being transcriptionally active) and polytenizes in *Drosophila.* Heterochromatin is late replicating, meiotic recombination inactive, has low gene density which corresponds to being transcriptionally inactive and does not polytenize in *Drosophila.* Heterochromatin can be further subdivided into several types along a continuum between constitutive and facultative heterochromatin. Both types affect the expression of genes; transcriptionally silent constitutive heterochromatin can affect nearby genes via PEV while facultative heterochromatin is thought to be the result of gene silencing via histone methylation and subsequent repressive complex binding. Constitutive heterochromatin is usually composed of repetitive sequences with structural functions such as centromeres or telomeres and localizes to the nuclear periphery while facultative heterochromatin is not repetitive, is interspersed throughout the nucleus, and can lose its condensed structure and become transcriptionally active in response to specific developmental or environmental cues. Constitutive heterochromatin is often di or tri-methylated at H3K9 and thus stains heavily for HP1 which binds to $H3K9^{me3}$ via its chromodomain while facultative heterochromatin can be methylated at H3K27 which is bound by PRC1 via Pc's chromodomain. Thus, both constitutive and facultative heterochromatins have been thought to achieve compaction and subsequent gene silencing via similar mechanisms (Riddle & Elgin 2006).

There are currently two complementary models that attempt to explain hPEV, the mass-action model and the nuclear organization model. The mass-action model postulates that the primary determinant of hPEV is the

cis-spreading of a condensed, heterochromatic state across the rearrangement breakpoint imposing an altered chromatin conformation onto the adjacent euchromatic region. This state inhibits access of the transcriptional machinery and results in transcriptional repression of the affected genes. Random, stochastic variation in the extent of this linear spreading during certain stages of development explains the variegated phenotype (Locke *et al*. 1988). Evidence for this model includes: cytogenetic studies showing changes in the polytene chromosome banding pattern of euchromatin adjacent to the rearrangement; modification of variegation by altering histone dosage or level of acetylation; transgene studies showing a more ordered nucleosomal array; and decreased restriction endonuclease sensitivity in heterochromatic locations compared to euchromatic insertions (Wakimoto 1998; Wallrath 1998; Weiler & Wakimoto 1995; Wallrath & Elgin 1995). hPEV can be relieved by moving the affected gene away from the breakpoint either by recombination or by the induction of a second rearrangement. hPEV affects genes further along the chromosome with strength of effect decreasing with distance, and more distally affected genes only variegate in cells that also show variegation of genes closer to the breakpoint indicating a progressive effect along the chromosome. Adding an extra Y-chromosome suppresses hPEV, presumably by the titration of heterochromatinizing factors away from the variegating site to the largely heterochromatic Y. Reciprocally, removing the Y-chromosome has the opposite effect. However, the mass-action model cannot easily explain the variegation of euchromatic genes located several megabases away from the breakpoint nor the sensitivity to inter-chromosomal interactions such as homologue pairing (reviewed by Girton & Johansen 2008).

The alternative nuclear compartmentalization model of hPEV hypothesizes that chromosomal rearrangements could disrupt the normal Rabl organization and place a normally euchromatic gene into a nuclear

compartment that lacks the correct concentration of transcription factors, resulting in a lack of expression and PEV. The classic example of this is the bw^D allele of the *brown* (*bw*) eye color gene (Dernburg *et al*. 1996; Csink & Henikoff 1996; Belyaeva *et al*. 1997). This mutation results from the insertion of a 1–2 megabase block of centromeric heterochromatin at 59E, near the distal tip of the right arm of chromosome *2* (*2R*). In bw^+/bw^D heterozygotes this insertion results in the dominant *trans* inactivation of the wild-type homologue. The degree of bw^+ *trans*-inactivation in the adult eye depended upon the degree of bw^D-*2Rh* association. These results are strong indicators that homologous pairing and nuclear sub-localization can contribute to PEV and that the insertion of DNA can alter the expression of adjacent genes (reviewed by Girton & Johansen 2008).

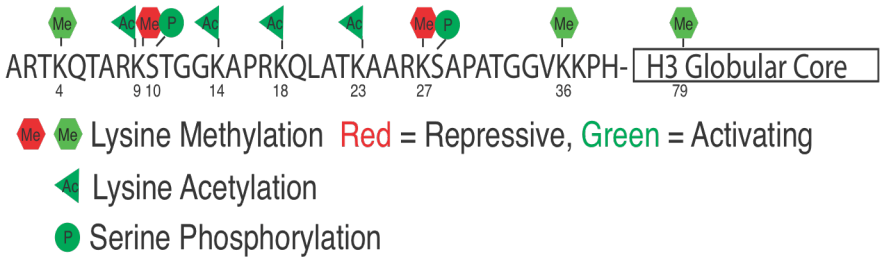
The spatial localization of mammalian chromatin within the nucleus has been shown to be important for transcription (Sutherland and Bickmore, 2009). The inner nuclear membrane and lamina (INM-lamina) comprise a prominent compartment with a unique set of trans-membrane proteins on a basal network of lamin intermediate filaments (Schirmer and Foisner, 2007). This compartment makes extensive chromatin contacts with a large fraction of the mammalian genome segregated into transcriptionally inactive, nuclear lamina associated chromosomal domains. The murine immuno-globulin heavy-chain (IgH) locus is associated with the INM-lamina in fibroblasts and hematopoietic progenitors; however, in B-lineage progenitors it is in the nucleoplasm, coinciding with its transcriptional activation (Kosak *et al*., 2002; Reddy *et al*., 2008). Likewise, activation of the *β-globin* and *CFTR* loci coincide with their dissociation from the lamina during development (Ragoczy *et al*., 2006; Zink *et al*., 2004). This association with the nuclear lamina is also seen for developmentally regulated transgenes during *Caenorhabditis elegans* embryogenesis; with dissociation from the nuclear lamina occurring in activated lineages (Meister *et al*., 2010). Zullo *et.al.* (2012) used genomic repositioning

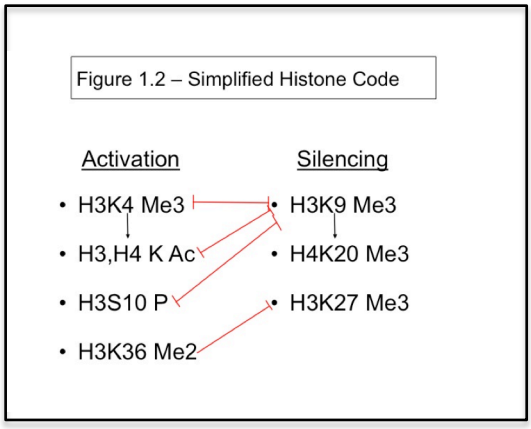
assays to demonstrate that the developmentally regulated *IgH* and *Cyp3a* loci contain discrete DNA regions that associate with the nuclear lamina in fibroblasts with concomitant transcriptional repression. Fine-scale mapping showed these regions were enriched for a GAGA motif that directed lamina association. This repeated motif was bound by cKrox in a complex with HDAC3, and knockdown of either *cKrox* or *HDAC3* resulted in dissociation from the nuclear lamina. Since cKrox is a transcriptional repressor, and histone deacetylation is transcriptionally inactivating, these results provide a model that couples nuclear lamina compartmentalization of chromatin domains with the repression of gene activity.

Histone code summary

Before going further, a review the "histone code" is needed. The histone code hypothesis states that the transcription of DNA is partially regulated by chemical modifications to histone proteins, usually to their amino terminal ends. Two copies each of histone proteins 2a, 2b, 3, and 4 form an octamer around which the DNA wraps itself 1.67 times to form a nucleosome. The four histone proteins consist of large globular cores and flexible N-terminii. These terminii protrude from the nucleosome and their amino acids can be modified, thus making them potentially relevant for both higher order structure and in transcriptional regulation. Histone tail modifications recruit effector proteins via the specific interaction of the modified histone with specialized protein domains such as chromodomains or bromodomains, which recognize methylated and acetylated lysines respectively. Common modifications include the acetylation, methylation or ubiquitination of lysines, phosphorylation of serines, and methylation of arginine residues. Figure 1.1 shows the major sites of modification of H3 relevant to this study that are discussed below, while figure 1.2 shows the relationships between the various modifications.

Figure 1.1 - Common covalent modifications to the amino terminal tail of histone 3.





One of the most basic repressive histone modifications in *Drosophila* is the di- and tri-methylation of H3K9 on chromosomes 1-3 by SU(VAR) 3-9 and by dSETdB on chromosome 4. This methylated lysine is recognized and bound by the chromodomain of HP1. HP1 can then associate with SU(VAR) 3-9 via its chromoshadow domain, providing a mechanism for the spread of heterochromatin, which is presumed to be resistant to access by transcriptional machinery. However, methylation of H3K9 can be blocked by phosphorylation of H3S10 by JIL-1, a serine specific tandem kinase discussed below.

One of the most fundamental activating histone modifications in *Drosophila* is the methylation of H3K4 by TRX. Once methylated, this promotes subsequent acetylation of lysines 9, 14, 18, and 23 on histone 3 by HATs. Acetylation of H3K9 prevents its methylation making activation and inactivation mutually exclusive. In order for H3K9 to be methylated, H3K4 must therefore first be demethylated. This is done by SU(VAR)3-3, the *Drosophila* homolog of human Lsd1, the first histone demethylase to

be identified (Rudolph *et al.* 2007; Shi *et al.* 2004). SU(VAR)3-3, demethylates H3K4me2 and H3K4me1 facilitating the subsequent methylation of H3K9 by SU(VAR)3-9. Loss of *Su(var) 3-3* results not only in elevated global levels of monomethyl and dimethyl-H3-K4 methylation along with elevated expression of a subset of genes (Rudolph *et al*. 2007) but also in the suppression of position effect variegation, suggesting a disruption of the balance between euchromatin and heterochromatin (Di Stefano *et al.* 2007). *Su(var)3-3* mutations also prevent extension of H3K9me2 at pericentric heterochromatin, indicating that *Su(var)3-3*-mediated removal of activating histone marks is a prerequisite for subsequent H3K9 methylation and heterochromatin formation (Rudolph 2007).

Another repressive mark is the di and tri-methylation of H3K27 by E(Z), a component of PRC2. This methylated lysine is recognized and bound by the chromodomain of PC, the eponymous member of PRC1, resulting in genetic silencing. This has been postulated to be due to the spread of facultative heterochromatin, which is presumed to be similar to constitutive heterochromatin in resistance to access by transcriptional machinery. In a manner analogous to the interdependence of H3K4 and K9 methylation status, the methylation status of H3K27 depends on the status of H3K36. Yuan *et al*. (2011) demonstrated that H3K36 di-methylation by ASH1H inhibited the spread of H3K27 di- and tri-methylation by PRC2. Furthermore, the trxG protein CBP associates with ASH1 and acetylates H3K27 blocking its methylation by E(z). As well, ASH1 recruits TRX which methylates H3K4 as mentioned above.

Suppressors of PEV in Drosophila

Mutations in genes whose products are involved in establishing and promoting heterochromatin formation, such as S*u(var)205*, *Su(var)3-7*, and *Su(var)3-9* suppress PEV. For example, *Su(var)205* encodes the major heterochromatin associated protein, HP1, normally found in

chromocenter and telomeric heterochromatin and along the 4th chromosome (Eissenberg *et al*. 1992; Eissenberg & Elgin 2000). HP1 spread into euchromatic regions across a variegating breakpoint correlates with the silencing of adjacent genes. HP1 is recruited to silenced arrays of *P* transposons carrying a mini-*white* gene (such as BX2 and T1) in polytene chromosomes (Fanti *et al*. 1998); however, it is also detectable in lesser abundance at single copy *white* transgenes which are not silenced and at transposon arrays with a different eye color marker that do not exhibit *P* element dependent silencing (PDS). Thus, the presence of HP1 above a certain threshold of concentration may be required (reviewed in Elgin 1996; Eissenberg & Elgin 2000).

Su(var) 3-7 encodes a heterochromatin-associated protein with seven zinc finger domains suggesting DNA binding activity (Cleard *et al*. 1995). SU(VAR)3-7 has a distribution identical to that of HP1 in *Drosophila* larval salivary gland polytene chromosomes; as well, antibodies to SU(VAR)3-7 co-immunoprecipitate HP1 from embryo extracts indicating the association of the *Su(var)3-7* product with HP1 in heterochromatin (Reuter *et al.* 1990; Cleard *et al*. 1997; Delattre *et al*. 2000). Jaquet *et al*. (2002) proposed that the spacing of zinc fingers in the N-terminal half of Su(var)3-7 allows them to contact DNA at a distance. This would facilitate the packing of scattered DNA sequences into a more compact conformation in a manner similar to that described for the 12 zinc finger protein encoded by *Suppressor of Hairy-wing* (Shen *et al*. 1994).

Su(var)3-9 encodes the major Drosophila histone H3K9 methyltransferase; SU(VAR)3-9 is the second major constituent of heterochromatin protein complexes. Green fluorescent protein tagged SU(VAR)3-9 fusion products localize to the chromocenter, telomeres, and all main blocks of translocated pericentric heterochromatin. Functional analysis of SU(VAR)3-9 homologues in fission yeast, *Drosophila* and

mammals demonstrate an evolutionarily conserved SU(VAR)3-9-dependent gene silencing process which has been demonstrated by human SUV39H1 transgene rescue of *Drosophila Su(var)3-9* mutant phenotypes (Schotta *et al*. 2002, 2003a,b). *Su(var)205*, *3-7* and *3-9* are classed together as haplo-insufficient loci with a triplo-dependent enhancer effect as the presence of an extra wild type gene leads to enhancement of variegation, while dominant suppression of PEV results from haploinsufficiency. The evidence thus far indicates that the proteins encoded by these genes are needed for proper heterochromatin structure formation.

Taken together these results indicate that suppression of PEV at the variegating locus results from a decrease in heterochromatin formation and suggest a mechanism for maintaining heterochromatin during DNA replication. HP1 is recruited to DNA via interactions between histone 3 trimethylated at lysine-9 by SUVAR39 methylase and the HP1 chromodomain (Bannister *et al*. 2001). Since SUVAR39 methylase is associated with HP1 (via its chromo-shadow domain), which in turn is bound to the chromatin, it would then direct methylation of new histones after the DNA was replicated, recruiting new HP1. *Su(var) 3-7* over expression results in ectopic H3K9 dimethylation and HP1 localization suggesting it plays a role in SUVAR39 recruitment (Delattre *et al*. 2004) while *retinoblastoma* (*Rb*) has been shown to be necessary to direct SUVAR39H1 methylation of H3K9, and for subsequent binding of HP1 to the *cyclin E* promoter. (Soren *et al*., 2001) This helps to explain the stable inheritance of a heterochromatic state in these regions.

hPEV and chromosome 4

While hPEV was first discovered and examined with a large chromosome inversion, it can also be studied via the use of transgene reporters inserted into or near heterochromatic regions such as the pericentric

heterochromatin, or on chromosome 4, as shown by *in-situ* hybridization. Chromosome 4 is the smallest autosome, being only 4.2 Mb in size with 3 Mb constituting the centromere leaving only ~1.2 Mb in a polytenizing arm with approximately 82 known or predicted genes. Chromosome 4 has an organization reminiscent of mammalian genomes with a repeat density of 30% and therefore exhibits characteristics of both heterochromatin and euchromatin such as late S-phase replication, lack of meiotic recombination, and the ability to induce hPEV while having a high gene density and being transcriptionally active. Most reporter transgenes inserted into chromosome 4 undergo hPEV silencing which is not unexpected as the entire chromosome is highly enriched for H3K9 di- and trimethylation and heterochromatin protein 1a (reviewed in Riddle and Elgin 2008). Transgene reporter inserts into sites that variegate have reduced endonuclease sensitivity and a relatively ordered nucleosome spacing while inserts into sites that do not variegate are more similar to autosomes in nature. Variegating inserts respond to *Su(var)205* and *3-7* in the same manner as w^{m4} and other centromeric hPEV lines but not to *Su(var)3-9* (Riddle & Elgin 2006, Sun *et al*. 2000, Wallrath & Elgin 1995): however, on chromosome 4 H3K9 is methylated by dSETDB1, the Drosophila ortholog of mammalian SETDB1 (Tzeng *et al*. 2007). Recent chromatin immunoprecipitation (ChIP)-array analysis for 18 histone modifications and 17 chromosomal proteins in S2 and BG3 cell lines has demonstrated that chromosome 4 displays a distinct chromatin profile similar to pericentric heterochromatin with interspersed euchromatic regions. Most notably, the centromere proximal region encompassing *ci* displays two distinct chromatin signatures of transcriptionally active genes within large blocks of pericentric type heterochromatin. The left block covering the *ci* coding region is enriched in H3K36me and H3K27ac, marks of states 3 (brown), while the right block covering the *ci* upstream regulatory region is enriched in well-established TSS and elongation

signatures including H3K36me3, which are marks of state 2 (mauve) (Kharchenko *et al*. 2011). Riddle *et al*. (2011), in a similar study, also found the *ci* coding and upstream regulatory regions to be positively associated with the transcription activating marks H3K3me3, H3K36me3, and H3K27ac as well as the proteins ASH1 and *Painting of fourth* with a block of pericentric-like heterochromatin just upstream of the transcription start site seperating the coding and regulatory regions. Intriguingly, they also found a positive correlation with HP1a and *Su(var)3-7*, proteins normally associated with pericentric heterochromatin. Interestingly, Riddle *et al*. (2011) did not find any euchromatic domains as defined by H3K9me2/H3K9me3/HP1a depletion and activation mark association. Instead, active genes on chromosome 4 were characterized by a distinct combination of *Painting of fourth*, H3K36me3, HP1a, and H3K9me di and trimethylation. The authours observed a strong correlation between transcriptionally permissive sites (full expression of an hsp70-white transgene reporter resulting in a red eye phenotype) and *Polycomb* (PcG) regulation. The authours claim that all non-variegating lines and no variegating reporter lines are inserted in regions that contain H3K27me3 and PC and lack HP1a; however, close examination of their figures indicates that the single non-variegating insert near *ci* is not. Their results suggest that: HP1a and PC occupy separate domains on chromosome 4; that PcG regulated domains may be transcriptionally permissive for hsp70-white reporters in the critical cell type; and that the majority of chromosome 4 genes are associated with HP1a, a heterochromatic mark normally correlated with silencing.

Molecular mechanisms

This switching of chromatin states can be blocked by the presence of euchromatic markers. Such markers include methylation of H3K4 and acetylation of H3K9. As well, methylation of H3K9 can be blocked by

phosphorylation of H3S10, which is under the control of JIL-1, a serine specific tandem kinase that functions to maintain euchromatic domains and counteract heterochromatin formation and gene silencing. Loss of JIL-1 results in H3K9 dimethylation (H3K9me2) and HP1 spreading to ectopic locations while genetic interaction assays have shown that *JIL-1* functions antagonistically to *Su(var)3-9* (Schotta *et al*. 2002; Zhang *et al*. 2006).

Enhancers of PEV in Drosophila

E(var) loci encode products that are usually thought of as either constituents of euchromatin or factors that resist epigenetic silencing, such as negative regulators of constituents of heterochromatin formation. Therefore, mutations in genes that encode transcription factors, which presumably confer a more open chromatin state, can enhance hPEV (reviewed in Weiler and Wakimoto 2002). Examples of this are *E(var)3-93E* and members of the *trithorax* group such as *trithorax* (*trx*), *ash1*, *trithorax-like*, *zeste*, *additional sex combs*, and *mod(mdg4)* (Flybase). Null mutations of *zeste* are strong enhancers of PEV affecting *w, roughest* and an allele of *notch* called *split* in decreasing order. It is possible that *zeste* participates redundantly in the opening and stabilization of transcriptionally active chromatin domains; however, this probably reflects its role in transcriptional regulation rather than in partitioning the genome into higher order structure, thus *zeste* is not generally considered a true enhancer of PEV *(E(var))* (Judd 1995; Weiler & Wakimoto 2002).

Transcriptional activators are often enhancers of PEV

trithorax (*trx*) and *ash1* are both SET (***S****u(var)3-9*, ***E****nhancer of zeste* and ***t****rithorax*) domain containing epigenetic activators necessary for maintenance of post-gastrulation homeotic gene expression. They do so via association with cis-regulatory elements called *trithorax* response elements (TREs), which produce non-coding RNAs in a tissue specific manner. The human *trx* homologue *mixed-lineage leukemia* (*MLL*) protein

trimethylates H3K4 at the human *HOXA9* locus (Dou *et al*. 2005); this epigenetic mark is generally associated with transcriptionally active regions of chromatin (Eissenberg & Shilatifard 2010). Thus, while H3K9 methylation serves to anchor HP1 and promote heterochromatin formation, H3K4 methylation blocks HP1 binding and maintains a domain in a transcriptionally active euchromatic state. ASH1 (which is a major gene in this thesis) is targeted to TREs via binding of its SET domain to TRE transcripts and subsequent recruitment to the corresponding TRE template. This process is sensitive to siRNA degradation of TRE transcripts as well as to RNase H or A but not RNase III indicating that single stranded RNA is important for the association and that it is the TRE transcripts themselves that hybridize with the template DNA (Sanchez-Elsner *et al*. 2006). ASH1 recruitment to the *bxd* region of its *Ubx* target gene coincides with its target gene's epigenetic activation (Beisel *et al*. 2002).

E(var)3-93E encodes *dE2F*, a Drosophila transcription enhancer and cell cycle regulator. Deficiency of *dE2F* enhances variegation of w^{m4} and *yellow* while over expression of *dE2F* via the mild heat shock of flies containing an *hs-dE2F* transgene suppresses PEV. Thus, *dE2F* is a haplo-enhancer and triplo-suppressor of classic PEV. This dose dependency seems to indicate *dE2F* has a direct effect on chromatin structure that is surprising for a cell cycle regulator and transcriptional activator. (Seum *et al.* 1996)

Mod(mdg4) protein genetically interacts with *Su(Hw)* protein to confer directionality on *gypsy*'s boundary or insulator function, stopping transcription from upstream promoters; this boundary becomes bidirectional in the absence of *mod(mdg4)*. Its widespread distribution and the *E(var)* phenotype of *mod(mdg4)* mutants suggest a role in either stopping heterochromatin formation or enhancing euchromatin formation

for *mod(mdg4)* separate from its insulator function. Mod(mdg4) co-localizes with *su(Hw)* at about 200 sites that lack *gypsy* on polytene chromosomes. These sites are hypothesized to contain sequences similar to the gypsy insulator sequence and to therefore be functionally equivalent. *Mod(mdg4)* protein is also present at about 300 other loci without SU(HW). Since *mod(mdg4)* lacks DNA binding domains, it must be associating with other unidentified DNA binding proteins to form different types of boundaries or insulators or perhaps to play some as yet unidentified role in gene expression (Gerasimova *et al*. 1995; Gdula *et al*. 1996; Gerasimova and Corces 1996). Some of the proteins found to co-localize with the *mod(mdg4)* protein product are Trithorax and Polycomb group proteins. (Flybase)

Before the Homeotics: the Segmentation Genes

As the Drosophila embryo develops, a sequential order of gene expression partitions the cells into ever decreasing lengthwise domains, beginning with the maternal coordinate gene products. Their concentration gradients are interpreted into repetitive patterns by the gap, pair rule and segment polarity genes. Finally, selector genes known as the homeotic or HOX genes define each parasegment's identity. Homeotic gene products are required throughout development and adult life, long after the segmentation gene products that determined their early expression have disappeared. Domains of homeotic gene activation are maintained in subsequent cell generations by TrxG proteins, while domains of gene repression are maintained by proteins of the PcG. However, PcG and TrxG involvement may occur earlier than HOX gene regulation as genetic studies have demonstrated that gap genes *hunchback*, *knirps*, and *giant* and the pair rule gene *even skipped* are directly or indirectly regulated by the PcG during early embryogenesis (McKeon *et al*. 1994; Pelegri and Lehmann 1994). PREs have been defined or predicted in *hunchback*, *knirps*, *giant*, *even skipped* and several other segmentation genes as well

as in the segment polarity genes *engrailed* (Kassis 1994), *hedgehog* (Maurange & Paro 2002), and *ci* (Schwartz *et al.* 2006; Tolhuis *et al*. 2006).

The finding that so many genes in the segmentation hierarchy may be regulated by PcG and TrxG proteins indicates that the transcription factor cascades may be bolstered or maintained by chromatin mechanisms at every level of the hierarchy (Ringrose & Paro 2004).

PcG & TrxG proteins regulate transcription by covalently modifying chromatin

The current model of how TrxG and PcG proteins regulate transcription to maintain a particular state is through facilitating a series of chromatin modifications that promote either open or closed conformations, similar to euchromatin or heterochromatin (facultative heterochromatin). *E*(*z*), a histone methyltransferase in PRC2 methylates H3K27 in PREs. This epigenetic mark is recognized by the chromodomain of Pc, a subunit of PRC1. THUS, PRC1 promotes the binding of PRC1 to chromatin via a mechanism similar to the recruitment of HP1 by SU(VAR)3-9 in the formation of heterochromatin.

Similarly, covalent modification of histone tails in TREs by trxG methyltransferases such as TRX and ASH1, or histone acetyltransferases such as CBP or ASH1 targets or regulates the activities of trxG members such as BRM (SWI/SNF) or KIS required for ATP-dependent chromatin remodeling. BRM and other SWI/SNF subunits contain bromodomains that interact with acetylated histone tails while KIS contains two chromodomains similar to PRC subunits and HP1. This provides a mechanism by which heritable histone modifications could ensure transmission of either an open conformation that is transcriptionally permissive or a closed one that is transcriptionally repressive (reviewed in Ringrose & Paro 2004; Schuettengruber *et al*. 2011)

Are trxG proteins transcriptional activators or anti-repressors?

Genetic studies showing that removing PcG complexes in the absence of ASH1 and TRX (Klymenko and Muller, 2004) results in transcriptional reactivation suggest that TrxG proteins may act as PcG antagonists rather than transcriptional activators. Yuan et al. (2011) demonstrated that ASH1H is a H3K36 dimethylase, and that H3K36me2 inhibited the spread of H3K27 di- and tri-methylation by PRC2. Substitution of alanine for lysine at H3K36 had the same effect indicating that contact with H3K36 or its neighbours is important for PRC2 activity. Therefore, methylating or mutating H3K36 most likely impairs the enzymatic activity of PRC2 by restricting its binding. Furthermore, the TrxG protein CBP associates with ASH1 and acetylates H3K27 blocking its methylation by E(z). Taken together these results suggest that one of the modes of action of TrxG proteins is as PcG antagonists.

P element Dependent Silencing (PDS) in *Drosophila melanogaster*

P{lacW}ci^{Dplac}

Previous work in the Locke lab involved investigation into *P* element dependent silencing (PDS), which is similar to PEV, but deals with *P{lacW}ci*^{Dplac} , a transgene insert close to the centromere on chromosome 4 between *Ribosomal protein S3A* (*RpS3A*) and *cubitus interruptus (ci)*. *P{lacW}ci*^{Dplac} was originally isolated as an enhancer trap of *ci*^{D} (Eaton & Kornberg 1990) and as such should be sensitive to many of the proteins that normally regulate *ci* during development. The *ci* gene product is required for proper anterior-posterior boundary formation in embryos and imaginal discs with *ci* being expressed only in the anterior compartment and *engrailed* (*en*) in the posterior. In the eye-antennal disc *en* expression is limited to the posterior third of the antennal disc with *ci* being expressed throughout the rest of the eye-antennal disc. The eye-antennal disc rotates ~180° after evaginating from the embryonic ectoderm inverting its

anterior-posterior axis in the adult fly compared to the embryo. Thus the anterior half of the disc becomes the more posterior eye where *ci* and the w^{+mC} gene in the enhancer trap *P{lacW}ci*^{Dplac} are expressed throughout development.

P elements

In M cytotype flies, the w^{+} transgene in *P{lacW}ci*^{Dplac} is expressed in a uniform manner throughout the eye resulting in a wild type (red) eye. However, in flies containing *P* elements (*P* cytotype) or KPs (internally deleted *P* elements capable of producing a putative *P* repressor (*P*-like cytotype) variegation occurs resulting in patches of white ommatidea on a red background. Like hPEV, this indicates random silencing of the w^{+} minigene during development, presumably by similar phenomena such as heterochromatin spreading. This hypothesis is supported by the *P* cytotype dependent transcriptional repression of germ-line-expressed *hsp83* or *vasa*-IVS3-beta-*geo* reporter transgenes; like the *w* mini gene, neither the *hsp83* nor *vasa* promoter contains *P*-element protein binding sites suggesting that *P* cytotype transcriptional repression may occur through a chromatin-based transcriptional silencing mechanism (Roche & Rio 1998). To date, this *P* element dependent silencing (PDS) is limited to *P{lacW}ci*^{Dplac} , plus another w^{+} transgene that is inserted close to *P{lacW}ci*^{Dplac}, between the gene for ribosomal protein subunit 3a and that for *cubitus interruptus*. *P{lacW}* insertions in other locations do not display PDS and chromosomal translocations of this construct away from the centromere suppress this silencing (28). *P{lacW}ci*^{Dplac} variegating inserts respond to *Su(var)205* and *Su(var)3-7* in the same manner as w^{m4} and other centromeric hPEV lines, but not to *Su(var) 3-9*. However, on chromosome 4, H3K9 is methylated by dSETDB1, the *Drosophila* ortholog of mammalian SETDB1 (Wallrath & Elgin 1995; Sun *et al*., 2000, Riddle & Elgin 2006; Tzeng *et al*. 2007). As chromosome 4 is heavily heterochromatic with the centromere being even more so, this presents

another system for examining factors for suppressing and enhancing the formation of heterochromatin and their roles in gene expression, as well as the role of *P* cytotype.

PDS and *P* cytotype

The PDS system requires the presence of other non-autonomous *P* elements capable of mimicking some of the characteristics of *P* cytotype such as modifying *P*- repressor sensitive alleles, but not enabling *P* element mobilization and transposition. Some examples are *P{ry+ SalI}89D*, *KP-U*, and *KP-D*. These *P* elements have mutations that prevent transposase production, but produce a protein product with much of the same function as the putative 66-kDa repressor. PDS is not the same as *P* cytotype as it occurs in somatic tissue, is not dependent on maternal inheritance, and is not affected by the T1 array of *P{lacW}* inserts. In addition, position is important as translocating *P{lacW}ci*^{Dplac} away from the centromere suppresses PDS (Bushey & Locke 2004). The position dependence and variegated phenotype suggest silencing may occur from a similar mechanism as heterochromatic PEV. Further evidence for this is the formation of heterochromatin in arrays of three or more tandem repeats of *P{lacW}* inserts causing variegation of the w^+ transgenes, which is modified by the same mutations as hPEV. Silencing of a single insert at a homologous location in *trans* to these arrays also occurs as previously mentioned for *bw*^{D}, and moving these arrays next to constitutive heterochromatin enhances variegation. The presence of the same non-autonomous P elements that cause PDS at *P{lacW}ci*^{Dplac} enhances variegation at these tandem arrays (Dorer & Henikoff 1997,Josse *et al*. 2002).

Recent work in our lab

PDS and Su(var) mutations

In our lab, Bushey and Locke (2004) have used genetic screens to identify two mutations in *Su(var) 205* and 11 at *Su(var)3-7* that suppress PDS at *P{lacW}ci*^{Dplac} indicating further similarity of silencing mechanism with hPEV. *Su(var)205*^{P4} and *Su(var)205*^{P5} are relatively weak suppressors of PDS at *P{lacW}ci*^{Dplac}. Testing of other extant *Su(var)205* alleles revealed a weaker to non-existent suppression of PDS compared to their much stronger suppression of PEV at *w*^{m4}. The *Su(var) 3-7* mutants were much stronger suppressors of variegation at *P{lacW}ci*^{Dplac} than the *Su(var(205)* alleles, with suppression ranging from strong to nearly complete; however, only two of our *Su(var)3-7* mutants suppressed variegation at *w*^{m4} while the other four tested did not. They dose dependently suppressed silencing by both type I and type II P elements at *P{lacW}ci*^{Dplac} indicating that both *P* element types act in PDS via a common mechanism, and that this is dependent on *Su(var)3-7* dose. Flies with two wild type *Su(var)3-7* alleles showed strong to complete PDS by both types of elements whereas in those with one wild type and one mutant *Su(var)3-7* PDS was nearly completely suppressed. Increasing *Su(var)3-7*$^{+}$ dose to three copies resulted in variegation of *P{lacW}ci*^{Dplac} in the absence of other non-autonomous P-elements indicating *Su(var)3-7* product is already acting at this loci and does not require recruitment by *P* element proteins, this decrease in *w*$^{+}$ expression was confirmed by pigment analysis. Therefore, *Su(var)3-7* is a haploinsufficient triplo enhancer of PDS as well as of hPEV as previously described. The *Su(var)3-7* mutants did not interfere with *P{ry SalI}89D's* ability to modify *vg*$^{21-3}$ or *P{lacZ}* expression indicating repression of repressor sensitive alleles by the *P* repressor protein and PDS silencing of *w*$^{+}$ occur through two different mechanisms (Bushey and Locke 2004).

Moving *P{lacW}ci*^{Dplac} to other locations suppresses PDS; however, increasing *Su(var)3-7*$^{+}$ dosage in these translocations resulted in reduced pigment levels in M strain flies indicating reduced *w*$^{+}$ expression in the absence of variegation. It is already known that mutations in *Su(var)205* suppress hPEV at the BX2 and T1 *P{lacW}* arrays; testing these arrays for dependence on *Su(var)3-7*$^{+}$ dosage revealed enhanced variegation in the absence of other non autonomous *P* elements. Therefore, silencing at these repeats and at *P{lacW}ci*^{Dplac} involves dependence upon a similar set of chromatin proteins, which could explain why *P* elements enhance silencing in both phenomena (Bushey and Locke 2004).

P{lacW}^{ciE1} and *P{lacW}*^{ciE2}

During the investigation of PDS at *P{lacW}ci*^{Dplac} two spontaneous mutants, *P{lacW}*^{ciE1} and *P{lacW}*^{ciE2}, were recovered that showed a variegated eye phenotype in the absence of other *P* elements and a complete white eye phenotype when combined with *P{ry SalI}89D. Su(var)205* mutant alleles weakly suppressed variegation at *P{lacW}ci*^{Dplac} while *Su(var)3-7* mutant alleles had a much stronger effect similar to their patterns of suppression of PDS. Therefore, the same dose dependent modifiers act on both *P{lacW}*^{ciE}s as do on PDS. Analysis of *P{lacW}*^{ciE1} and *P{lacW}*^{ciE2} revealed gypsy element insertions approximately 1 kb upstream from P{lacW}ci^{Dplac} in opposite orientations 547 bp apart. Both *P{lacW}*^{ciE1} and *P{lacW}*^{ciE2} were tested against mutants of *su(Hw)* for suppression and *mod(mdg4)* for enhancement. The four mutants of *su(Hw)* tested either suppressed variegation weakly, or not at all for both *P{lacW}*^{ciE1} and *P{lacW}*^{ciE2}. They did; however, suppress the known *su(Hw)* dependent mutant *lozenge*^{1} male eye phenotype, indicating that the suppression system was functional. The allele of *mod(mdg4)* that was tested weakly enhanced silencing. For all *su(Hw)* and *mod(mdg4)* mutations the effect was too weak to genetically map for segregation with the mutation concerned. Taken together, it was concluded that variegation

by *P{lacW}*^{ciE1} and *P{lacW}*^{ciE2} was not as a result of the gypsy insulator function described earlier (Bushey & Locke 2004).

Both *P{lacW}*^{ciE1} and *P{lacW}*^{ciE2} *trans*-silenced w^+ expression from *P{lacW}ci*^{Dplac} but not from translocations with the exception of T(3;4)DB66C where the effect was weaker than for *P{lacW}ci*^{Dplac} . Therefore, position or pairing contributes to trans-silencing. Combining PDS caused by *P{ry SalI}89D* with trans-silencing caused by *P{lacW}*^{ciE1} and *P{lacW}*^{ciE2} resulted in an increase in w^+ silencing. This was true for *P{lacW}*^{ciE1} and the *hsp70-*w^+ transgene in *P{hsp26-pt-T}*$^{ci2\text{-}M1021.R}$ (Bushey & Locke 2004).

Rationale for this research

The *cis* changes in sequence at *P{lacW}*^{ciE1} (*E1*) and *P{lacW}*^{ciE2} (*E2*) seem to cause changes in heterochromatin functionally similar to that caused by *P* elements in trans. Their effects are modified by some of the same modifiers as PEV and PDS and in the same manner as for PDS. However, unlike PDS at *P{lacW}ci*^{Dplac} they allow for an easier screen for enhancers of PDS (and therefore presumably PEV) due to their intermediate variegated phenotype which can not only be visibly suppressed as shown for *Su(var)205* and *Su(var)3-7* mutants, but also enhanced closer to full white eye phenotype as shown in the presence of other *P* elements. Since *P{lacW}ci*^{Dplac} is an enhancer trap of *ci*^{D}, this system should be capable of revealing epigenetic regulators of *ci* as well as modifiers of heterochromatin formation. This research should help address the question of what aspects of epigenetic gene regulation, facultative heterochromatin formation, and hPEV are separate, and in common.

<u>Modifiers</u>

True *E(vars)* are underrepresented in the literature (Weiler & Wakimoto 2002), presumably due to limitations inherent in the screening systems

used. The system proposed here allows for simple *E(var)* and *Su(var)* screening and is a less complicated genetic system than used previously for PDS as all components are on the same chromosome. Since the w^+ mini-gene is under its own promoter (rather than the *P* promoter) loci specific to *P* element transcription should not isolated. The elimination of *P{ry SalI}89D* prevents the isolation of mutants that enhance or suppress PDS by changing transcription from the *P* promoter of *P{ry SalI}89D.* Chromosome 4 is the most similar to human chromosomes of all *Drosophila* chromosomes as the genes are interspersed with numerous short repeats (Sun *et al*. 2000). Given this and the conserved nature of heterochromatinization as previously discussed, mutants isolated should help identify human homologues and improve our knowledge of heterochromatin formation and chromatin structure in humans as well as in *Drosophila*.

Genetic screen for Enhancers of PDS

I began this work by screening for enhancers of PDS using the eye phenotype of $P\{lacW\}^{ciE1}$ (mutagenize *E1* homozygous males with EMS) in a manner similar to Bushey and Locke (30). Putative recessive lethal mutants would be tested for transmission via transvection with *E1*/*Pci*, stocked, sorted into complementation groups, and mapped via recombination to approximate chromosomal loci. If the recessive lethal phenotype mapped to the same approximate locus, and was inseparable from the enhancer phenotype by recombination, then deficiency mapping would be used to further refine the target location. Complementation analysis against extant alleles of candidate genes would identify mutant loci, which would then be sequenced to identify individual lesions. Pigment analysis would allow confirmation and quantification of the extent of PDS enhancement as well as specificity. Extant alleles were also tested to confirm enhancement; unfortunately, all outside stocks available contained either *P*s, *KP*s or *En(var)*s of w^{m4}. Nonetheless, I have isolated enhancers

of PDS that can then be characterized further and add to our understanding of the complicated phenomena of PDS, PEV, and heterochromatin/euchromatin formation and maintenance in epigenesis. I have also recovered a transcriptional activator to add to our knowledge of gene regulation and gene interactions.

References

Allis, C.D., Jenuwein, T., and Reinberg, D. 2006. *Epigenetics*, pp. 23–61, Cold Spring Harbor Laboratory, New York.

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**: 120-124.

Beisel, C., Imhof, A., Greene, J., Kremmer, E. & Sauer, F. 2002. Histone methylation by the *Drosophila* epigenetic transcriptional regulator Ash1. *Nature* **419**: 857-862.

Belyaeva, E.S., Koryakov, D,E., Pokholkova, G.V., Demakova, O.V., and Zhimulev, I.F., 1997. Cytological study of the brown dominant position effect. *Chromosoma* **106**: 124-132.

Bushey, D. and Locke, J. 2004. Mutations in Su(var)205 and Su(var)3-7 Suppress P-Element-Dependent Silencing in Drosophila melanogaster. *Genetics* **168**: 1395–1411.

Cléard, F., Matsarskaia, M., and Spierer, P. 1995. The modifier of position effect variegation *Suvar(3)7* of *Drosophila*: there are two alternative transcripts and seven scattered zinc fingers, each preceded by a tryptophan box. *Nucleic Acids Res.* **23**: 796-802.

Cleard, F., Delattre, M., and Spierer, P. 1997. SU(VAR)3-7, a *Drosophila* heterochromatin-associated protein and companion of HP1 in the genomic silencing of position-effect variegation. *EMBO J* **16**: 5280–8.

Delattre, M., Spierer, A., Tonka, C.H., and Spierer P. 2000. The genomic silencing of position-effect variegation in *Drosophila melanogaster*: interaction between the heterochromatin-associated proteins Su(var)3-7 and HP1. *J. Cell Sci.* **113**: 4253–61.

Delattre, M., Spierer, A., Jaquet, Y., and Spierer, P. 2004. Increased expression of *Drosophila* Su(var)3-7 triggers Su(var)3-9-dependent heterochromatin formation. *J. Cell Sci.* **117**: 6239–47.

Di Stefano, L., Ji, J.Y., Moon, N.S., Herr, A., and Dyson, N. 2007. Mutation of *Drosophila Lsd1* Disrupts H3-K4 Methylation, Resulting in Tissue-Specific Defects during Development. *Curr. Biol.***17**: 808–812.

Dorer, D. R., and Henikoff, S. 1997. Transgene repeat arrays interact with distant heterochromatin and cause silencing in cis and trans. *Genetics* **147:** 1181-1190.

Dou, Y., Milne, T.A., Tackett, A.J., Smith, E.R., Fukuda, A., Wysocka, J., Allis, C.D., Chait, B.T., Hess, J.L. and Roeder, R.G. 2005. Physical Association and Coordinate Function of the H3 K4 Methyltransferase MLL1 and the H4 K16 Acetyltransferase MOF. *Cell* **121**: 873-85.

Eaton, S., and Kornberg, T.B., 1990. Repression of ci-D in posterior compartments of *Drosophila* by *engrailed*. *Genes Dev*. **4**: 1068–1077.

Elgin, S.C.R. 1996. Heterochromatin and gene regulation in *Drosophila. Curr. Opin. Genet. Dev*. **6**: 193–202.

Eissenberg, J.C., Morris, G.D., Reuter, G., Hartnett, T. 1992. The heterochromatin-associated protein HP-1 is an essential protein in *Drosophila* with dosage-dependent effects on position-effect variegation. *Genetics*, **131**: 345-352.

Eissenberg, J.C. and Elgin, S.C.R. 2000. The HP1 protein family: getting a grip on chromatin. *Curr. Opin. Genet. Dev*. **10**: 204–210.

Fanti, L., Dorer, D.R., Berloco, M., Henikoff, S., and Pimpinelli, S. 1999. Heterochromatin protein 1 binds transgene arrays*. Chromosoma* **107**: 286-292.

Gdula, D.A., Gerasimova, T.I., and Corces, V.G. 1996. Genetic and molecular analysis of the *gypsy* chromatin insulator of *Drosophila. Proc. Natl. Acad. Sci. USA* **93**: 9378-9383.

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

Gerasimova, T.I. and Corces, V. G. 1996. Boundary and insulator elements in chromosomes. *Curr. Opin. Genet. Dev*. **6**: 185-192.

Girton, J.R. and Johansen, K.M. 2008. Chromatin Structure and the Regulation of Gene Expression: The Lessons of PEV in *Drosophila. Advances in Genetics*. **61**: 1-43.

Heitze E. 1928. Das Heterochromatin der Moose. *Jahrb Wiss Bot* **69**: 762-818.

Jaquet, Y., Delattre, M., Spierer, A., and Spierer, P. 2002. Functional dissection of the *Drosophila* modifier of variegation *Su(var)3-7*. *Development* **129:** 3975-3982.

Josse, T., Boivin, A., Anxolabehere, D., and Ronsseray, S. 2002. *P* element-encoded regulatory products enhance Repeat-Induced Gene Silencing (RIGS) of P-lacZ-white clusters in *Drosophila melanogaster. Mol. Genet. Genomics* **268:** 311-320.

Judd, B.H. 1995. Mutations of zeste that mediate transvection are recessive enhancers of position-effect variegation in *Drosophila melanogaster*. *Genetics* **141**: 245–253.

Kassis, J.A. 1994. Unusual properties of regulatory DNA from the *Drosophila engrailed* gene: three “pairing-sensitive” sites within a 1.6-kb region. *Genetics* **136**: 1025–38.

Kharchenko, P.V., Alekseyenko, A.A., Schwartz, Y.B., Minoda, A., Riddle, N.C., Ernst, J., Sabo, P.J., Larschan, E., Gorchakov, A.A., Gu, T., Linder-Basso, D., Plachetka, A., Shanower, G., Tolstorukov, M.Y., Luquette, L.L., Xi, R., Jung, Y.L., Park, R.W., Bishop, E.P., Canfield, T.P., Sandstrom, R., Thurman, R.E., MacAlpine, D.M., Stamatoyannopoulos, J.A., Kellis, M., Elgin, S.C.R., Kuroda, M.I., Pirrotta, V., Karpen, G.H., and Park, P.J. 2011. Comprehensive analysis of the chromatin landscape in *Drosophila*. *Nature* **471**: 480–485.

Klymenko, T. and Muller, J. 2004. The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. *EMBO Rep.* **5**: 373–377.

Kosak, S.T., Skok, J.A., Medina, K.L., Riblet, R., Le Beau, M.M., Fisher, A.G., and Singh, H. 2002. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science* **296**: 158–162.

Locke, J., Kotarski, M.A., and Tartof, K.D. 1988. Dosage-dependent modifiers of position effect variegation in *Drosophila* and a mass action model that explains their effect. *Genetics* **120**: 181-198.

Maurange, C., and Paro, R. 2002. A cellular memory module conveys epigenetic inheritance of *hedgehog* expression during *Drosophila* wing imaginal disc development. *Genes Dev.* **16**: 2672–83.

McKeon, J., Slade, E., Sinclair, D.A., Cheng, N., Couling, M., and Brock, H.W. 1994. Mutations in some *Polycomb* group genes of *Drosophila* interfere with regulation of segmentation genes. *Mol. Gen. Genet.* **244**: 474–83.

Meister, P., Towbin, B.D., Pike, B.L., Ponti, A., and Gasser, S.M. 2010. The spatial dynamics of tissue-specific promoters during *C. elegans* development. *Genes Dev*. **24**: 766–782.

Muller, H.J., 1930. Types of visible variation induced by X-Rays in *Drosophila*. *J. Genet*. **22**: 299–334.

Pelegri, F., and Lehmann, R. 1994. A role of *Polycomb group* genes in the regulation of gap gene expression in *Drosophila*. *Genetics* **136**: 1341–53.

Ragoczy, T., Bender, M.A., Telling, A., Byron, R., and Groudine, M. 2006. The locus control region is required for association of the murine beta-globin locus with engaged transcription factories during erythroid maturation. *Genes Dev*. **20**: 1447–1457.

Reddy, K.L., Zullo, J.M., Bertolino, E., and Singh, H. 2008. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* **452**: 243–247.

Reuter, G., Giarre, M., Farah, J., Gausz, J., Spierer, A., and Spierer P. 1990. Dependence of position-effect variegation in *Drosophila* on dose of a gene encoding an unusual zinc-finger protein. *Nature* **344**: 219–23.

Riddle, N. C. and Elgin, S.C. 2006. The dot chromosome of *Drosophila*: insights into chromatin states and their change over evolutionary time. *Chromosome Res* **14**: 405-16.

Riddle, N.C., Elgin S.C. 2008 A role for RNAi in heterochromatin formation in Drosophila. Curr. Top. Microbio.l Immunol. 320: 185–209.

Riddle, N.C., Jung, Y.L., Gu, T., Alekseyenko, A.A., Asker, D., Gui, H., Kharchenko, P.V., Minoda, A., Plachetka, A., Schwartz, Y.B., Tolstorukov, M.Y., Kuroda, M.I., Pirrotta, V., Karpen, G.H, Park, P.J., and Elgin, S.C.R. 2011. Enrichment of HP1a on *Drosophila* Chromosome 4 Genes Creates an Alternate Chromatin Structure Critical for Regulation in this Heterochromatic Domain. *PLOS Genetics* **8**(9): e1002954. doi:10.1371/journal.pgen.1002954

Ringrose, L., and Paro, R. 2004. Epigenetic Regulation Of Cellular Memory By The Polycomb And Trithorax Group Proteins *Annu. Rev. Genet.* **38**: 413–43.

Roche, S.E., and Rio, D.C. 1998. *Trans*-silencing by P elements inserted in subtelomeric heterochromatin involves the *Drosophila* Polycomb group gene, *Enhancer of zeste*. *Genetics* **149**: 1839-1855.

Rudolph, T., Yonezawa, M., Lein, S., Heidrich, K., Kubicek, S., Schafer, C., Phalke, S., Walther, M., Schmidt, A., Jenuwein, T., Reuter, G., 2007. Heterochromatin formation in *Drosophila* is initiated through active removal of H3K4 methylation by the LSD1 homolog SU(VAR)3-3. *Mol. Cell*. **26**: 103–115.

Sanchez-Elsner, T., Gou, D., Kremmer, E., and Sauer, F. 2006. Noncoding RNAs of Trithorax Response Elements Recruit *Drosophila* Ash1 to Ultrabithorax. *Science* **311:** 1118-1123.

Schirmer, E.C., and Foisner, R. 2007. Proteins that associate with lamins: many faces, many functions. *Exp. Cell Res*. **313**: 2167–2179.

Schuettengruber, B., Martinez, A-M., Iovino, N., and Cavalli, G. 2011. Trithorax group proteins: switching genes on and keeping them active *Nature Reviews Molecular Cell Biology* **12**: 799-814.

Schotta, G., Ebert, A., Krauss, V., Fischer, A., Hoffmann, J., Rea, S., Jenuwein, T., Dorn, R., and Reuter, G. 2002. Central role of *Drosophila* SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. *EMBO J*. **21**: 1121–1131.

Schotta, G., Ebert, A., and Reuter, G. 2003a. SU(VAR)3-9 is a conserved key function in heterochromatic gene silencing. *Genetica* **117**: 149-158.

Schotta, G., Ebert, A., Dorn, R., and Reuter, G. 2003b. Position effect variegation and the genetic dissection of chromatin regulation in *Drosophila*. *Semin. Cell Dev. Biol.* **14:** 67–75.

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

Seum, C., Spierer, A., Pauli, D., Szidonya, J., Reuter, G., and Spierer, P. 1996. Position-effect variegation in *Drosophila* depends on the dose of the gene encoding the E2F transcriptional activator and cell cycle regulator. *Development* **122**: 1949–56.

Shen, B., Kim, J., and Dorsett, D. 1994. The enhancer-blocking suppressor of Hairy-wing zinc finger protein of *Drosophila melanogaster* alters DNA structure. *Mol. Cell Biol.* **14**: 5645-5652.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., and Casero, R.A. 2004. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119:** 941–953.

Soren J.N., Schneider, R., Bauer, U-M., Bannister, A.J., Morrison, A., O'Carroll, D., Firestein, R., Cleary, M., Jenuwein, T., Herrera, R.E., and Kouzarides, T. 2001. Rb targets histone H3 methylation and HP1 to promoters. *Nature* **412**: 561-565.

Sun, F. L., Cuaycong, M. H., Craig, C. A., Wallrath, L. L., Locke, J., and Elgin, S.C.R. 2000. The fourth chromosome of *Drosophila melanogaster*: interspersed euchromatic and heterochromatic domains. *Proc. Natl. Acad. Sci. USA* **97:** 5340–5345.

Sutherland, H., and Bickmore, W.A. 2009. Transcription factories: gene expression in unions? *Nat. Rev. Genet*. **10**: 457–466.

Tolhuis, B., de Wit, E., Muijrers, I., Teunissen, H., Talhout, W., van Steensel, B., and van Lohuizen, M. 2006. Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in *Drosophila melanogaster*. *Nat. Genet.* **38**: 694–699.

Tzeng, T., Lee, C., Chan, L., and Shen, C. 2007. Epigenetic regulation of the *Drosophila* chromosome 4 by the histone H3K9 methyltransferase dSETDB1. *Proc. Natl. Acad. Sci. USA* **104**: 12691-12696.

Wakimoto, B.T. 1998. Beyond the Nucleosome: Epigenetic Aspects of Position–Effect Variegation in *Drosophila. Cell* **93**: 321–324.

Wallrath, L.L. 1998. Unfolding the mysteries of heterochromatin. *Curr. Opin. Genet. Dev*. **8**: 147–153.

Wallrath, L., and Elgin, S.C.R. 1995. Position effect variegation in *Drosophila* is associated with an altered chromatin structure. *Genes Dev*. **9**: 1263–77.

Weiler, K.S. and Wakimoto, B.T. 1995. Heterochromatin and gene expression in *Drosophila*. *Annu. Rev. Genet.* **29:** 577–605.

Weiler, K.S., and Wakimoto, B.T. 2002. Suppression of heterochromatic gene variegation can be used to distinguish and characterize *E(var)* genes potentially important for chromosome structure in *Drosophila melanogaster*. *Mol. Genet. Genomics* **266**: 922–32.

Yuan, W. *Xu, M., Huang, C., Liu, N., Chen, S., and Zhu, B.* 2011. H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J. Biol. Chem.* **286**: 7983–7989 27.

Zhang, W., Deng, H., Bao, X., Lerach, S., Girton, J., Johansen, J., and Johansen, K.M. 2006. The JIL-1 histone H3S10 kinase regulates dimethyl H3K9 modifications and heterochromatic spreading in *Drosophila. Development* **133**: 229-235.

Zink, D., Amaral, M.D., Englmann, A., Lang, S., Clarke, L.A., Rudolph, C., Alt, F., Luther, K., Braz, C., Sadoni, N., et al. 2004. Transcription-dependent spatial arrangements of *CFTR* and adjacent genes in human cell nuclei. *J. Cell Biol.* **166**: 815–825.

Zullo, J.M., Demarco, I.A., Piqué-Regi, R., Gaffney, D.J., Epstein, C.B., Spooner, C.J., Luperchio, T.R., Bernstein, B.E., Pritchard, J.K., Reddy, K.L., Singh, H. 2012. DNA sequence-dependent compartmentalization and silencing of chromatin at the nuclear lamina. Cell. 149:1474-87.

Chapter 2 – Manuscript for *ash1* research

This chapter is a version of a manuscript in preparation for submission for publication.

Mutations in *ash1* enhance *P* element dependent silencing (PDS) in Drosophila melanogaster.

Allen McCracken and John Locke*

Department of Biological Sciences,
University of Alberta,
Edmonton, Alberta, Canada, T6G 2E9

*Corresponding Author:
John Locke
Department of Biological Sciences
CW 405, Biological Sciences Bldg
University of Alberta
Edmonton, Alberta
Canada, T6G 2E9
Email: john.locke@ualberta.ca
Phone: 780-492-2193
Fax: 780-492-9234

Abstract

In *Drosophila melanogaster*, the mini-w^+ transgene in *P{lacW}ci*^{Dplac} *(Pci)* is normally expressed throughout the adult eye; however, when *P* or *KP* elements are also present, a variegated eye phenotype results indicating random w^+ silencing during development. This is *P* element dependent silencing (PDS). In previous work, we recovered a spontaneous mutation of *P{lacW}ci*^{Dplac} called *P{lacW}ci*$^{Dplac\ E1}$ (*E1*) that showed w^+ variegation in the absence of *P* elements, presumably due to the insertion of a *gypsy* element adjacent to the transgene. Heterochromatin modifiers, such as *Su(var)205* and *Su(var)3-7* act dose dependently on *Pci* and *E1* indicating their involvement in the variable gene silencing. In a genetic screen for second-site modifiers of *E1* variegation, we isolated 5 mutations in *ash1* that enhance this variegated phenotype. These mutant alleles enhance silencing of *E1* alone, and in transvection with *E1*/*Pci*, but suppress heterochromatic position effect variegation (hPEV) at w^{m4}. This is consistent with a model whereby ASH1 marks transcriptionally active chromatin domains. If ASH1 function is lost or altered, heterochromatin can spread into these domains creating a sink for heterochromatic proteins, and leaving less heterochromatin elsewhere.

Introduction

The first known example of an epigenetic effect was that of heterochromatic Position Effect Variegation (hPEV) in *Drosophila melanogaster* (Muller 1930). The euchromatic *white* (*w*) gene was translocated next to heterochromatin resulting in a variegated or mottled eye colour phenotype. The initial random silencing of *white* gene expression during tissue development is maintained in a somewhat stable state through multiple cell divisions causing clonal inheritance and a mosaic pattern of gene expression. The primary determinant of hPEV is the *cis*-spreading of a condensed, heterochromatic state from a pericentric initiation site across the rearrangement break point into the adjacent region, which then inhibits access of the transcriptional machinery and results in transcriptional repression (Locke *et al*. 1988). In normal pericentric heterochromatin spreading would be contained by termination sites. In the absence of termination sites, there is random variation in the extent of this cis-spreading during certain stages of development to cause the variegation (reviewed in Eissenberg & Reuter 2009). Homolog position and pairing also contribute to PEV as shown by the bw^D allele of the *brown* (*bw*) eye color gene, which results from the insertion of a 1–2 megabase block of centromeric heterochromatin at 59E, near the distal tip of the right arm of chromosome *2* (*2R*). In bw^+/bw^D heterozygotes this insertion results in the dominant *trans*-inactivation of the wild-type homologue. The degree of bw^+ *trans*-inactivation in the adult eye depends upon the degree of bw^D-*2Rh* association indicating that homologous pairing and nuclear localization contribute to PEV (Dernburg *et al*. 1996; Csink & Henikoff 1996; Belyaeva *et al*. 1997).

In *Drosophila*, expression of the *white* gene is cell autonomous, and necessary for the import of pigment precursors into the adult eye. The white-eye phenotype allows for identification of flies transformed with w^+ containing *P*-element constructs, and in some cases also allows for the

identification of regulatory regions near the construct's insertion site. We are using *P{lacW}ci*^{Dplac}, a transgene insert close to the centromere on chromosome 4 between *Ribosomal protein S3A* (*RpS3A*) and *cubitus interruptus (ci)* which was originally isolated as an enhancer trap of *ci* (Eaton & Kornberg 1990), but we have discovered that the w^+ expression is also sensitive to the presence of P elements (Bushey & Locke 2004; Sameny & Locke 2011). This is *P* element dependent silencing (PDS) and appears similar to PEV. In flies lacking *P* elements (M strains) *Pci* is expressed in a uniform manner (red eye phenotype), while in flies containing *P* elements (P strains) or *KP* s (derivative elements capable of mimicking some of the characteristics of *P* cytotype such as modifying *P*-repressor sensitive alleles, but not enabling *P* element mobilization and transposition) variegation occurs resulting in patches of white ommatidia on a red background in the eye. An example of such an element used in this lab is *P{ry+ SalI}89D*. Another w^+ transgene, *P{hsp26- pt-T}ci*$^{2\text{-}M1021.R}$, present near this location also undergoes PDS (Bushey & Locke 2004). Like PEV, this indicates random silencing of the w^+ minigene during development, presumably by similar phenomena such as heterochromatin spreading as variegating inserts respond to *Su(var)205* and *Su(var)3-7* in the same manner as w^{m4} and other centromeric PEV lines (Sun *et al*. 2000, Wallrath & Elgin 1995). As with PEV, position is important as insertions of the same transgene in other locations do not display PDS and translocating *P{lacW}ci*^{Dplac} away from the centromere suppresses PDS (Bushey & Locke 2004). The position dependence and variegated phenotype suggest silencing may occur from a similar mechanism as heterochromatic PEV.

During investigation of PDS at *Pci* two spontaneous mutants, *P{lacW}ci*^{DplacE1} and *P{lacW}ci*^{DplacE2} (hereafter *E1* and *E2*), were recovered that showed a variegated eye phenotype in the absence of other *P* elements and a complete white eye phenotype when combined

with *P{ry+ SalI}89D.* Both *Su(var)205* mutant alleles and *Su(var)3-7* mutant alleles suppressed silencing of *E1* and *E2*. Analysis of *E1* and *E2* revealed *gypsy* element insertions approximately 1 kb upstream from *P{lacW}ci*^{Dplac} in opposite orientations 547 bp apart. Testing of *E1* and *E2* against mutants of *su(Hw)* and *mod(mdg4)* showed that variegation by *E1* and *E2* was not as a result of the *gypsy* insulator function *per se* (Bushey & Locke 2004); however, this does not rule out some interaction between the *gypsy* insulator and the *wari* element at the 3' end of the w^+ transgene (Chetverina *et al.* 2008). Both *E1* and *E2 trans*-silenced w^+ expression from *P{lacW}ci*^{Dplac}, but not from translocations, with the exception of *T(3;4)DB66C* where the effect was weaker than for *P{lacW}ci*^{Dplac}. Therefore position or pairing contributes to trans-silencing similar to the dominant *trans*-inactivation of the wild-type homologue in bw^+/bw^D heterozygotes (Bushey & Locke 2004).

PDS at *E1* allows for an easy visual screen for second-site dominant enhancers of silencing due to its intermediate variegated phenotype. This should identify genes involved in modifying chromatin changes at *E1* as well as epigenetic regulators of *ci*. From such a screen we describe here the isolation and characterization of 5 mutants in *ash1*, a putative *ci* regulator.

Materials and Methods

Drosophila stocks and mutations: Unless otherwise cited, *D. melanogaster* mutations were described previously (Lindsley & Zimm 1992). The *P{lacW}ci*^{Dplac} allele (Eaton & Kornberg 1990) is a *P{lacZ*$^{P\T.W}$*w*$^{+mC}$ *amp*^{R} *ori = lacW}* construct inserted ~3 kbp upstream from the *ci* locus on chromosome *4*. *P{lacW}ci* ^{DplacE1} has a gypsy element insertion ~ 1kb further upstream as previously described (Bushey & Locke 2004). *y*^{1} *w* P{lacW}3-76a* and *ash1*^{B1}were originally isolated by Y. N. Jan and J. Kennison respectively, and were provided by the Bloomington stock center.

Fly stocks (Table A1) were maintained at room temperature on standard yeast/cornmeal medium. Mutagenesis involved *w; dp; e; P{lacW}ci* ^{DplacE1} males treated with 25 mM EMS (Ashburner 1989) mated to *y w* virgin females and screened for an enhanced eye phenotype in the progeny. Potential mutants were mated to *w*$^{-}$*; dp*$^{-}$*; e*$^{-}$*; P{lacW}ci* ^{Dplac} flies to confirm transmission and segregation and to determine chromosomal location. Mutant *ash1* alleles were kept as balanced stocks with *TM6B, Hu e*.

Genetic Mapping: The enhancer and recessive lethal phenotypes were separately mapped by recombination relative to *Gl Sb H e* markers. During the enhancer mapping recombinants to both left and right of the enhancer were collected. These were tested for retention of the enhancer phenotype by crossing males to *w*$^{-}$*; dp*$^{-}$*; e*$^{-}$*; P{lacW}ci* ^{DplacE1} virgin females, and for retention of the recessive lethal phenotype by crossing to other members of the same complementation group. The recessive lethal location was refined by complementation analysis against deficiencies in the region. At least 100 progeny were scored and if the heterozygous mutant combination was not recovered the combination was considered lethal.

DNA sequencing: Mutants were crossed back to the parental chromosome and overlapping gene segments amplified by PCR (Tables A4, A5) and sequenced. Point mutations were identified as half-height, double peaks. All polymorphisms and mutations were confirmed by reading both strands. The region containing the deletion in mutant *3a42a* was purified as a separate smaller band (reduced by ~ 230 nt) after PCR amplification and sequenced in both directions. In mutant *4a89a* the deletion was suggested by double peaks on the sequencing chromatogram. The hypothetical deletion removed a Pst I site, so after digestion with Pst I the remaining high molecular weight band was sequenced from both ends. Homozygous *4a89a* mutant dead embryos were sequenced on both strands to confirm the deletion end points.

Eye pigment assays: The amount of w^+ gene activity was determined by measuring the amount of brown eye pigment using a modification of the method of Ephrussi and Herold (1944). Heads from 6-8 day old adult flies were stored at -20°C, until extracted. For each genotype, three replicate samples of 10 heads were extracted in 200 μL of acidified ethyl alcohol (1% HCl in ethanol) with shaking for 48 hours. Absorbance at 470 nm was then measured using a 96 well Costar flat bottom plate in a Bio-Tek PowerWave XS spectrophotometer. Photographs of adult flies eyes were taken under mineral oil using a Zeiss stereo-microscope and a Nikon Coolpix 995 digital camera. For both the eye pigment assay and the adult eye photographs, the parental chromosome was used as a control.

Attempted Rescue of mutant loci: A construct was made with *ash1* under its own promoter fused in frame to a 6X Myc tag using TOPO cloning and the Gateway system. This was then sub cloned into a vector containing *attB* bacterial attachment sites and *PhiC31* integrase-mediated transgenesis used to transform flies containing the *attP* phage attachment

sites (*attP* docking sites) by BestGene Inc. Unfortunately, two attempts at transformation by BestGene Inc were unsuccessful.

Results

Screen for enhancers of E1

We screened ~44,000 flies for phenotypic enhancement of w^+ silencing in *E1* and recovered 58 mutations, which fell into five simple and three complex recessive lethal complementation groups, as well as many single alleles that complemented all other mutants. A simple group with 5 alleles (*3a1a'*, *3a31a*, *3a42a*, *4a5a* and *4a89a*) was examined further and the enhancer of *E1* phenotype genetically mapped to between 3-46.9 and 3-48.2 (Table 2.1). Recombinants exchanging either side of the enhancer were tested and shown to retain the recessive lethality common to the complementation group. The recessive lethality common to *4a5a* and *4a89a* mapped to 3-47.0 (n=466) and that common to *3a1a'* and *3a31a* mapped to 3-49.3 (n=353). From this we concluded that (1) all 5 enhancers mapped to the same locus, (2) that the common lethality also mapped to the same locus, and (3) that they were tightly linked. Deficiency mapping refined this position to 76B4;77A1, within Df(3L)XS705 (76B4:76D3) but not Df(3L)Exel6135 (76B9:76C5), which includes the *ash1 locus*. All 5 alleles failed to complement the recessive lethality of $ash1^{22}$ and $ash1^{B1}$; therefore we concluded they were mutant alleles of *ash1* and were temporarily designated *3a1a'*, *3a31a*, *3a42a*, *4a5a* and *4a89a*.

We wished to test extant alleles, $ash1^{22}$ and $ash1^{B1}$, for enhancement of the variegating eye phenotype of *E1*. However, PCR tests using internally directed *P* terminal repeat primers (Tables A4, A8) amplified an approximately 1150 bp band in both $ash1^{22}$ and $ash1^{B1}$ stocks and a 2.9 kb band in $ash1^{22}$. These bands are indicative of a *KP* and full length *P* element, respectively (O'Hare & Rubin 1983, Black *et al*. 1987). Attempts to segregate the $ash1^{22}$ and $ash1^{B1}$ mutations from interfering elements were unsuccessful and thus we could not assay their effect on *E1* modification.

Characterization of our *ash1* mutants

DNA sequencing of the entire coding region of each of the mutant alleles, in heterozygotes with the parental unmutagenized chromosome, and of $ash1^{B1}$ confirmed that each allele had a change that predicted an altered amino acid sequence (Table 2.1). Four of the alleles had premature stop codons, as does $ash1^{B1}$. The fifth has an amino acid substitution (H1873W) in ASH1's PHD finger indicating that this domain is necessary for ASH1 function. Figure 2.1 depicts these changes along the protein with respect to its predicted domains. The five alleles are now designated $ash1^{W790*}$, $ash1^{Q893*}$, $ash1^{N303*}$, $ash1^{H1873W}$, and $ash1^{W770*}$.

The lethal stage of each of the five mutants was determined by crossing $y^{-}\ w^{-};\ ash1^{*}/+$ virgin females of each allele to $y^{-}\ w^{-};\ Df(3L)Exel9007/+$ males (76B3;76B9, 3L:19417791;19628895) (Figure 2.2). Eggs, collected at 4 to 6 hour intervals, were scored after 12, 24, 48 and 72 hours for viability. Larvae were reared in the same containers and scored post-eclosion. Tripoulas *et al*. (1996) found that the null allele $ash1^{22}$ (Gln47→ amber) was lethal in the 3rd instar to pupariation stages. Hypomorphs of *ash1* survived to pharate adults and antimorphs died in the first to third instars. Of our alleles, $ash1^{N303*}$, $ash1^{W770*}$, $ash1^{W790*}$ and $ash1^{Q893}$ can be classified as larval recessive lethals. The embryonic lethality exhibited by the $ash1^{N303*}$ and the $ash1^{Q893*}$ bearing chromosomes appears be due to the presence of second site or synthetic lethal(s) as overall mortality exceeds 25%. $ash1^{H1873W}$ shows an intermediate degree of 12.5% larval lethality and close to 75 % overall mortality post pupariation. By the criteria of Tripoulas *et al.* (1996) we have classed $ash1^{H1873W}$ as a hypomorph based on the reduced larval lethality. Importantly, four alleles show no lethality after pupariation consistent with them being null alleles or stronger according to the criteria set out by Tripoulas *et al.* (1996).

Variegation of w^+ in *E1* heterozygotes was visibly enhanced for all alleles in both sexes, frequently producing flies indistinguishable from w^-. Representative examples are given in Figure 2.3. The extent of *E1* enhancement was quantified by pigment analysis of our *ash1* mutants, and *ash1B1*. Male *ash1* mutants had less pigment than non-mutant internal control flies from the same cross. Female *ash1* mutants also showed enhancement compared to non-mutant internal controls from the same cross, with all but $ash1^{W790*}$ being significant at the 95% confidence limit (Figure 2.4). This quantitatively confirms that all our *ash1* alleles enhance silencing of *E1*.

We also assayed pigment amounts in *E1/Pci* flies because the amount of eye pigment was very low in the progeny of the above crosses, and silencing so strong in *E1* heterozygous flies. The introduction of *Pci* in trans with *E1* normally produces an eye with substantially more pigment, making it easier to assay for enhancement. All alleles demonstrated a dose dependent relationship with *ash1* in that flies with *ash1* mutants had half the pigment of non-mutant internal control flies from the same cross in both sexes (Figure 2.5). Variegation was visibly enhanced for all alleles in both sexes. A pattern of silencing was also present starting at the posterior edge of the eye and weakening as it progressed anteriorly as shown in Figure 2.6.

To test if w^+ silencing was dependent upon the presence of *E1* we examined the effect of *ash1* mutants on *Pci*/+ flies, which lack *E1*. *Pci*/+ flies with wild type *ash1* have approximately half the pigment as the corresponding *E1/Pci* flies from the previous analysis. All *ash1* mutant *Pci*/+ flies were significantly reduced in pigment compared to non-mutant internal controls from the same cross, with $ash1^{N303*}$ males being significant at >95% and the rest > 99% confidence limit (Figure 2.7). Variegation was visually much harder to detect in this cross, but had a

pattern of weak silencing starting at the posterior edge of the eye and decaying rapidly as it progressed anteriorly (Figure 2.8).

To address the possibility that *ash1* product was interacting directly with the w^+ transgene within the *P{lacW}ci*^{Dplac} insert, we tested *y*^{1} *w* P{lacW}3-76a* /+ flies that lacked *E1* or *Pci* with *ash1* mutants. The *P{lacW}3-76a* transgene is the same as in *P{lacW}ci*^{Dplac} but is located at 18A1 (60.7 cM) on the X chromosome. It has a full red eye. None of our *ash1* mutant alleles showed any significant difference from non-mutant control flies of the same cross indicating that *ash1*'s silencing effect is dependent upon the insert location and not on the w^+ transgene *P{lacW}ci*^{Dplac} construct itself (Figures 2.9,2.10).

The *ash1 mutants affect hPEV**. Given that *ash1* mutants enhanced PDS in *P{lacW}ci*^{Dplac}, but had no effect on the same construct in *P{lacW}3-76a*, we asked whether or not *ash1* had any effect on hPEV at w^{m4} by crossing y^- *w*; *ash1** e^-/+ males to w^{m4}; e^- virgin females. Variegation of w^{m4} in female flies was visibly suppressed by all five mutants isolated in this study while male w^{m4} flies were also suppressed by four of the five mutants (Figure 2.11). Mutation *ash1*^{W790*} was originally isolated from a female, and we did not replace the Y-chromosome to match the other four mutants. *ash1*^{W790*} has the same Y-chromosome as the y^- w^-*; +; +;* + flies used as controls, and shows similar enhancement of w^{m4} variegation (Figure 2.12). This Y-chromosome enhancer effect appears to be much stronger than the suppression of variegation exhibited by *ash1*. This pushes w^{m4} variegation almost to a w^- phenotype, and overwhelms any Su(var) effect of *ash1*^{W790*} in males. Due to the high variability of the w^{m4} phenotype, statistical significance is problematic; however, *ash1*^{N303*} which most likely represents a null allele due to its being a stop codon early on in the polypeptide sequence and *ash1*^{H1873W} which is a mutation predicting a substitution of the middle conserved

histidine of *ash1*'s PHD finger both show significant suppression of w^{m4} variegation at the 95% confidence limit.

An interaction with *trx* mutations

One of the other complementation groups found to influence the transcriptional status of the *ci* region in this screen was *trx*. This group of 3 *trx* alleles silenced expression of both *E1*/+ and *Pci* /+ in a manner indistinguishable from *ash1* (McCracken & Locke 2012). Since ASH1 is known to interact with, and presumably recruit, TRX at other loci (Kuzin *et al*. 1994, Rozovskaia *et al*. 1999, Tillib *et al*. 1999), we asked whether or not they interacted at *Pci.* To answer this question, we crossed y^{-} w^{-}*:* $ash1^{H1873W*}$ e^{-}/+;*Pci* males to y^{-} w^{-}; trx^{R1583*} e^{-}/*TM3, Sb Ser* e^{-} females. The progeny of the appropriate combination of genotypes were collected and the eye pigment measured. Both *ash1* and *trx* single mutants had about half the pigment of non-mutant internal balancer control flies while the double mutant flies had about one quarter the pigment of non-mutant internal control flies indicating a cumulative, dosage dependent effect on reporter silencing for both *ash1* and *trx* (Figure 2.13).

Discussion

We have shown that the 5 En(PDS) mutations described in this study are alleles of *ash1:* (1) The enhancement phenotype of all 5 alleles map to the *ash1* locus. (2) The lethal phenotype maps to the *ash1* locus. (3) The two phenotypes could not be separated by recombination. (4) All 5 alleles failed to complement all deficiencies tested that uncover *ash1* and all 5 alleles failed to complement both $ash1^{22}$ and $ash1^{B1}$. (5) Four alleles contain mutations resulting in stop codons before ASH1's SET domain and therefore likely represent null alleles. The fifth, $ash1^{H1873W}$, is a hypomorph with reduced larval lethality, however, it is phenotypically indistinguishable from our other 4 alleles in its enhancement of *E1*, *E1*/*Pci*, *Pci* and suppression of w^{m4}. We therefore consider $ash1^{H1873W}$ to be a strong hypomorph. This allele has a mutation predicting a substitution of the middle conserved histidine residue in ASH1's PHD finger. Together with 2 conserved cysteine residues they coordinate 2 Zn^{2+} ions; loss of the histidine residue should abolish this function. The ING2, YNG1 and NURF PHD fingers have been reported to bind to histone H3 tri-methylated on lysine 4 (H3K4me3), this may therefore be a common property among PHD fingers (Shi *et al.* 2006; Martin *et al.* 2006; Wysocka *et al.* 2006). The modification H3K4me3 is associated with the transcription start site of active genes, and therefore may be how part of how ASH1 recognizes a transcriptionally active gene sequence that is required to be "locked on" by TrxG proteins.

ASH1 acts at the *ci* locus

Pci was isolated as an enhancer trap of *ci* and is an allele of *ci.* While *Pci* (and the *E1* gypsy element) are inserted in the *ci* distal regulatory region, both ci^{57g}, a deletion upstream of *Pci*, and ci^{1}, a gypsy insert upstream of *Pci,* exhibit *ci* phenotypes. This implies that the *ci* regulatory region extends past *Pci* and *E1* and that changes in their expression mimics *ci* regulatory region changes. Expression of the *Pci* enhancer-trap

reporter in imaginal discs accurately mimicked that of ci^D RNA with both being expressed specifically in anterior compartment cells (Eaton & Kornberg 1990). All alleles isolated of *ash1* are *En(PDS)* with *E1*, *E1/Pci* and act at *Pci* to lower *w* expression without *E1*. The lack of ASH1 function results in increased silencing of the *white* reporter gene, which should be mimicking *ci* expression. Since these mutants do not affect *P{lacW}3-76a*, this effect is not construct dependent and is not due to a direct interaction with the *white* promoter, but with the *ci* regulatory region itself. Since *Pci* reporter expression is halved when ASH1 dose is halved, and does not depend on the presence of *E1*, we infer that ASH1 normally acts at the *ci* regulatory region, likely in a dose dependent manner. Since Polycomb Response Elements (PREs) and Trithorax Response Elements (TREs) share similar components and distribution (Tillib *et al*. 1999; Schuettengruber *et al*. 2007; Ringrose & Paro 2007), and Pc(G) proteins have been shown to bind at *ci* (Schwartz *et al.* 2006; Tolhuis *et al*. 2006) it is likely that TRE(s) exist there as well for ASH1 to interact with. Transcription of *bxd* TREs have been shown to mediate transcriptional activation of *Ultrabithorax* (*Ubx*) by recruiting ASH1 to the template TREs (Sanchez-Elsner *et al*. 2006). One of the other complementation groups isolated in this screen has been identified as *trx*, which would be expected if a TRE were present at *ci*. Members of this *trx* complementation group showed a degree of synthetic lethality with those of the *ash1* complementation group (Supplemental Table A3). Genetic evidence has indicated that the activities of ASH1 and TRX are functionally related. Mutants in *ash1* and *trx* may exhibit synthetic lethality; ASH1 co-immunoprecipitates with TRX from embryonic nuclear extracts; and ASH1 and Trx co-localize at multiple sites on polytene chromosomes with TRX accumulation reduced in an *ash1* mutant background (Kuzin *et al*. 1994; Rozovskaia *et al*. 1999). These results suggest a model in which TRE transcription recruits ASH1. ASH1 then binds to and methylates H3 via its

SET domain allowing recruitment of TRX, and explains the loss of TRX on polytene chromosomes from an *ash1* mutant, and the synthetic lethality. The existence of a PRE/TRE at *ci*, the isolation of *ash1* and *trx* as enhancers of reporter gene silencing at *ci*, and their equivalence and additive effects at *ci* are consistent with this model. If ASH1 recruits TRX to *ci*, then loss of one dose of *ash1* should result in approximately half of the amount of TRX protein recruited to regulate the *ci* region, and be functionally similar to the loss of one dose of *trx* itself. If both *ash1* and *trx* doses are halved, then half the amount of ASH1 recruiting from a pool with half the amount of TRX, should result in recruitment of approximately one quarter of the normal amount of TRX protein to the *ci* region. Thus the double heterozygote mutant should produce one quarter of the normal amount of *white* reporter as seen in Figure 2.13.

ASH1 is a histone methyl-transferase (HMTase) that was first reported to methylate lysine residues 4 and 9 in Histone H3 and 20 in Histone H4 in *Drosophila* (Beisel *et al.* 2002), and Histone H3 lysine 4 in humans (Gregory *et al*. 2007). Histone methylation in the promoter of ASH1 target genes in *Drosophila* correlates with their transcriptional activation and is thought to serve as a binding surface for a chromatin remodelling complex containing the epigenetic activator Brahma (Brm). ChIP analysis of Ultrabithorax transcription in Drosophila indicates that transcriptional activation, trivalent methylation by ASH1 and recruitment of BRM coincide (Beisel *et al*. 2002). We therefore tested a null allele of *brahma*, brm^{2}, against *E1* to see if it enhanced variegation and for complementation against all third chromosome recessive lethal mutants recovered in this screen. Not only did it fail to have any visible effect on *E1* variegation, it complemented all 29 third chromosome recessive lethal mutants recovered in this screen (data not shown). We take these as a strong indication that *brahma* is not involved in the epigenetic regulation of *ci* by *ash1* and *trx*.

More recent work has shown that the actual target of ASH1 methylation is Histone H3K36 in both *Drosophila* and humans (Tanaka *et al*. 2007). Yuan *et al*. (2011) also demonstrated that ASH1H is a H3K36 dimethylase, and that H3K36me2 inhibited the spread of H3K27 di- and tri-methylation by PRC2. Substitution of alanine for lysine at H3K36 had the same effect indicating that contact with H3K36 or its neighbours is important for PRC2 activity. Therefore, methylating or mutating H3K36 most likely impairs the enzymatic activity of PRC2 by restricting its binding. Furthermore, the trxG protein CBP associates with ASH1 (Bantignies *et al*. 2000) and acetylates H3K27 in a TRX dependent manner, blocking its methylation by *Enhancer of zeste E(z)*, a member of *Polycomb* Repressive Complex Two (PRC2) (Feng Tie *et al*. 2009). PRC2 trimethylation of H3K27 is essential for *Polycomb* Repressive Complex One (PRC1) binding and transcriptional silencing of Polycomb target genes, whereas acetylation of H3K27 (H3K27ac) has been shown to be associated with active genes. Taken together these results suggest that one of the modes of action of ASH1 is as a PcG antagonist as well as a recruiter of *trx*, which is itself a H3K4 histone methyltransferase.

ASH1 protein localizes at over 100 sites on polytene chromosomes of larval salivary glands implying that it functions to maintain the expression pattern of multiple genes (Tripoulas *et al*. 1996). This action of ASH1 at many sites explains our mutants' Su(var) effect on w^{m4}. This opposite effect on w^{m4} is consistent with a model whereby loss of *ash1* results in the titration of heterochromatinizing factors away from pericentric heterochromatin. Loss of ASH1 at *ci* would not cause enough facultative heterochromatin formation to titre any measureable amount of *Su(var)* activity away from pericentric heterochromatin, but if this ectopic heterochromatin formation was to occur at many sites the loss of heterochromatin forming factors would add up to a measurable amount as seen at w^{m4} in this study. This also implies that the facultative

heterochromatin promoted by *Pc*G proteins and antagonized by *trx*G proteins shares some components such as HP1 with pericentric heterochromatin, and is consistent with the previous finding of *Su(var)205* acting to suppress expression of our reporter construct (Bushey & Locke 2004).

The fact that we isolated $ash1^{H1873W}$ as a recessive lethal member of the *ash1* complementation group indicates that ASH1's PHD finger is necessary for ASH1 function. The dose responsive loss of reporter gene activity in a $ash1^{H1873W}$ heterozygote indicates a failure of $ash1^{H1873W}$ mutant ASH1 to act at *ci*; most likely due to ASH1 being unable to localize or bind to nucleic acids properly.

For all of our alleles the pattern of loss of *w* expression follows progression of the morphogenetic furrow in *E1/+* and *E1/Pci* flies with ommatidea specified first having the least pigment. This suggests that a lower dose of *ash1* results in a later locking of *Pci* into a transcriptionally active state.

References

Ashburner, M. 1989. *Drosophila*, A laboratory manual. Cold Spring Harbor Press, Cold Spring Harbor, New York.

Bantignies, F., Goodman, R.H. & Smolik, S.M. 2000. Functional interaction between the coactivator *Drosophila* CREB-binding protein and ASH1, a member of the trithorax group of chromatin modifiers. *Mol. Cell. Biol*. **20**: 9317–9330.

Beisel, C., Imhof, A., Greene, J., Kremmer, E., & Sauer, F. 2002. Histone methylation by the *Drosophila* epigenetic transcriptional regulator *Ash1*. *Nature* **419**: 857-862.

Belyaeva, E.S., Koryakov,D,E., Pokholkova,G.V., Demakova,O.V., and Zhimulev, I.F. 1997. Cytological study of the brown dominant position effect. *Chromosoma* **106**:124-132.

Black, D. M., M.S. Jackson, M.G. Kidwell and G.A. Dover. 1987. KP elements repress P-induced hybrid dysgenesis in *D. melanogaster*. *EMBO J*. **6**: 4125-4135.

Bushey, D., and Locke, J. 2004. Mutations in *Su(var)205* and *Su(var)3-7* Suppress *P*-Element-Dependent Silencing in *Drosophila melanogaster*. *Genetics* **168**: 1395–1411.

Chetverina, D., Savitskaya, E., Maksimenko, O., Melnikova, L., Zaytseva, O., Parshikov, A., Galkin, A.V., and Georgiev, P. 2008. Red flag on the white reporter: a versatile insulator abuts the white gene in *Drosophila* and is omnipresent in mini-white constructs. *Nucleic Acids Research* **36**: 929–937.

Csink, A. K., and Henikoff, S. 1996. Genetic modification of heterochromatic association and nuclear organization in *Drosophila*. *Nature* **381**: 529–531.

Dernburg, A.F., Broman, K.W., Fung, J.C., Marshall, W.F., Philips, J., Agard D.A., and Sedat, J.W. 1996. Perturbation of Nuclear Architecture by Long-Distance Chromosome Interactions. *Cell* **85**: 745–759.

Eaton, S., and Kornberg, T.B. 1990. Repression of ci-D in posterior compartments of *Drosophila* by *engrailed*. *Genes Dev*. **4**: 1068–1077.

Eissenberg, J. C, and Reuter, G. 2009. Cellular mechanism for targeting heterochromatin formation in *Drosophila*. *Int. Rev. Cell Mol. Biol.* **273**:1–47.

Ephrussi, B., and Herold, J.L. 1944. Studies of eye pigments of *Drosophila*. I. Methods of extraction and quantitative estimation of the pigment components. *Genetics* **29**: 148–175.

Feng Tie, F., Banerjee, R., Stratton C.A., Prasad-Sinha. J., Stepanik, V., Zlobin, A., Diaz, M.O., Scacheri, P.C. and Harte, P.J. 2009. CBP-mediated acetylation of histone H3 lysine 27 antagonizes *Drosophila* Polycomb silencing. *Development* **136**: 3131-3141.

Gregory, G.D., Vakoc, C.R., Rozovskaia, T., Zheng, X., Patel, S., Nakamura, T., Canaani, E., & Blobel, G.A. 2007. Mammalian ASH1L Is a Histone Methyltransferase That Occupies the Transcribed Region of Active Genes. *Molecular and Cellular Biology* **27**: 8466–8479

Kuzin, B., S. Tillib, Y. Sedkov, L. Mizrokhi, and Mazo, A. 1994. The *Drosophila trithorax* gene encodes a chromosomal protein and directly regulates the region specific homeotic gene fork head. *Genes Dev*. **8**: 2478–2490.

Lindsley, D. L., and Zimm, G.G. 1992. The Genome of *Drosophila melanogaster*. Academic Press, New York.

Locke, J., Kotarski, M.A., and Tartof, K.D., 1988. Dosage dependent modifiers of position effect variegation in *Drosophila* and a mass action model that explains their effect. *Genetics* **120**:181-198.

Martin, D. G., Baetz, K., Shi, X., Walter, K.L., MacDonald, V.E., Wlodarski, M.J., Gozani, O., Hieter, P., and Howe, L. 2006. The Yng1p plant homeodomain finger is a methyl-histone binding module that recognizes lysine 4-methylated histone H3. *Mol. Cell. Biol*. **26**: 7871–7879.

Muller, H. J., 1930. Types of visible variation induced by X-Rays in *Drosophila. J. Genet.* **22**: 299–334.

O'Hare, K. and G. M. Rubin. 1983. Structure of *P* transposable elements and their sites of insertion and excision in the *Drosophila melanogaster* genome. *Cell* **34**: 25-35.

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

Rozovskaia, T., Tillib, S., Smith, S., Sedkov, Y., Rozenblatt-Rosen, O., Petruk, S., Yano, T., Nakamura, T., Ben-Simchon, L., Gildea, J., Croce, C.M. Shearn, A., Canaani, E., and Mazo, A. 1999. Trithorax and ASH1 Interact Directly and Associate with the Trithorax Group-Responsive bxd Region of the *Ultrabithorax* Promoter. *Mol. Cell. Biol.***19**: 6441–6447.

Sameny, A. and Locke, J. 2011. The *P*-element-induced silencing effect of *KP* transposons is dose dependent in *Drosophila melanogaster*. *Genome* **54**: 752-762.

Sanchez-Elsner, T., Gou, D., Kremmer, E., and Sauer, F. 2006. Noncoding RNAs of Trithorax Response Elements Recruit *Drosophila* Ash1 to Ultrabithorax. *Science* **311**: 1118-1123.

Schuettengruber ,B., Chourrout, D., Vervoort, M., Leblanc,B., and Cavalli. G., 2007. Genome Regulation by Polycomb and Trithorax Proteins. *Cell* **128**: 735–745.

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

Shi, X., Hong, T., Walter, K.L., Ewalt, M., Michishita, E., Hung, T., Carney, D., Peña, P., Lan, F., Kaadige, M.R., Lacoste, N., Cayrou, C., Davrazou, F., Saha, A., Cairns, B.R., Ayer, D.E., Kutateladze, T.G., Shi, Y., Côté, J., Chua, K.F., and Gozani, O. 2006. ING2 PHD domain

links histone H3 lysine 4 methylation to active gene repression. *Nature* **442**: 96–99 .

Sun, F. L., Cuaycong, M. H., Craig, C. A., Wallrath, L. L., Locke, J., and Elgin, S.C.R. 2000. The fourth chromosome of *Drosophila melanogaster*: interspersed euchromatic and heterochromatic domains. *Proc.Natl. Acad. Sci. USA* **97**: 5340–5345.

Tanaka, Y., Katagiri, Z., Kawahashi, K., Kioussis, D., and Kitajima, S. 2007. Trithorax-group protein ASH1 methylates histone H3 lysine 36. *Gene* **397**: 161–8.

Tillib, S., Petruk, S., Sedkov, Y., Kuzin, A., Fujioka, M., Goto, T., and Mazo. A., 1999. Trithorax- and Polycomb-group response elements within an *Ultrabithorax* transcription maintenance unit consist of closely situated but separable sequences. *Molecular and Cellular Biology* **19**: 5189–5202.

Tolhuis, B., de Wit, E., Muijrers, I., Teunissen, H., Talhout, W., van Steensel, B., and van Lohuizen, M. 2006 Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in *Drosophila melanogaster. Nat. Genet.* **38**: 694–699.

Tripoulas, N., LaJeunesse, D., Gildea, J., and Shearn, A. 1996. The *Drosophila ash1* Gene Product, Which Is Localized at Specific Sites on Polytene Chromosomes, Contains a SET Domain and a PHD Finger. *Genetics* **143**: 913-928.

Wallrath, L.L., and Elgin, S.C.R. 1995. Position effect variegation in *Drosophila* is associated with an altered chromatin structure. *Genes & Development* **9**: 1263-1277.

Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tacket,t A.J., Chait, B.T., Badenhorst, P., Wu, C., Allis, C.D. 2006. A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442**: 86–90.

Yuan, W., Xu, M., Huang, C., Liu, N., Chen, S., and Zhu, B. 2011. H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J. Biol. Chem*. **286:** 7983–7989.

Tables and Figures

Table 2.1 - En (PDS) mutants

Mutant	Enhancer Position	Base pair change	Predicted amino acid change	Stage of Lethality
3a1a'	47.2 (n=876)	G2369A	W790Amber	Larval
3a31a	48.2 (n=896)	C2677T	Q893Ochre	Larval
3a42a	46.9 (n=787)	Δ907-1139	N303-1 frameshift	Embryonic/ Larval
4a5a	47.0 (n=994)	C5617T	H1873W	Larval/ Pupal
4a89a	46.9 (n=920)	Δ2310-2334	W770Opal	Embryonic/ Larval
ash1^{B1}	Not Determined	C4978T	Q1660Amber	Not Determined

Table 2.1 Table summarizes *ash1* mutations showing the mutagen used, coding sequence change, amino acid alteration, stage of lethality, and source of the mutant. The extant mutation, *ash1*^{B1}, had not previously been sequenced.

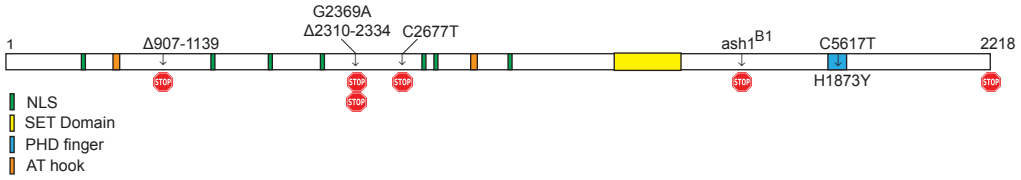


Figure 2.1 Schematic representation of ASH1. Important domains predicted by SMART (http://smart.embl-heidelberg.de/) and the location of mutants used in this study are indicated. Mutant designations are above the protein backbone while the nature of the corresponding mutation is below.

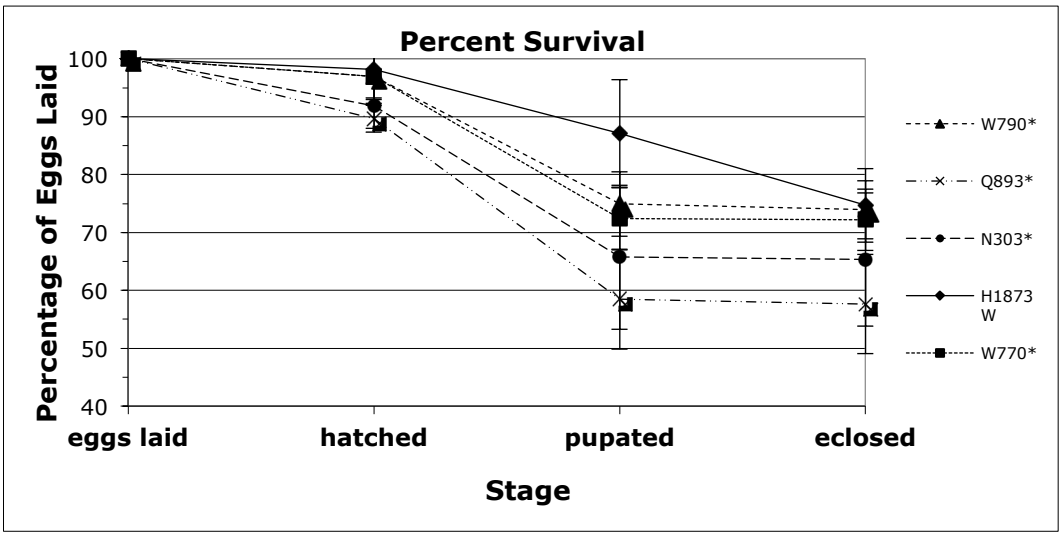


Figure 2.2 Effect of *ash1* alleles recovered in this screen on viability. Results expressed as a percentage of the number (N) of eggs laid. *ash1*^{W790*} N = 492, *ash1*^{Q893*}, N = 583, *ash1*^{N303*} N = 331, *ash1*^{H1873W} N = 578, *ash1*^{W770*} N = 562. All values normalized to wild type (*y*$^{-}$ *w*$^{-}$; +, +, +).

(*y*$^{-}$ *w*$^{-}$*; ash1* e*$^{-}$/+ ♀s X *y*$^{-}$ *w*$^{-}$*; Df(3L)Exel9007/+* ♂s (76B3;76B9, 3L:19417791;19628895))

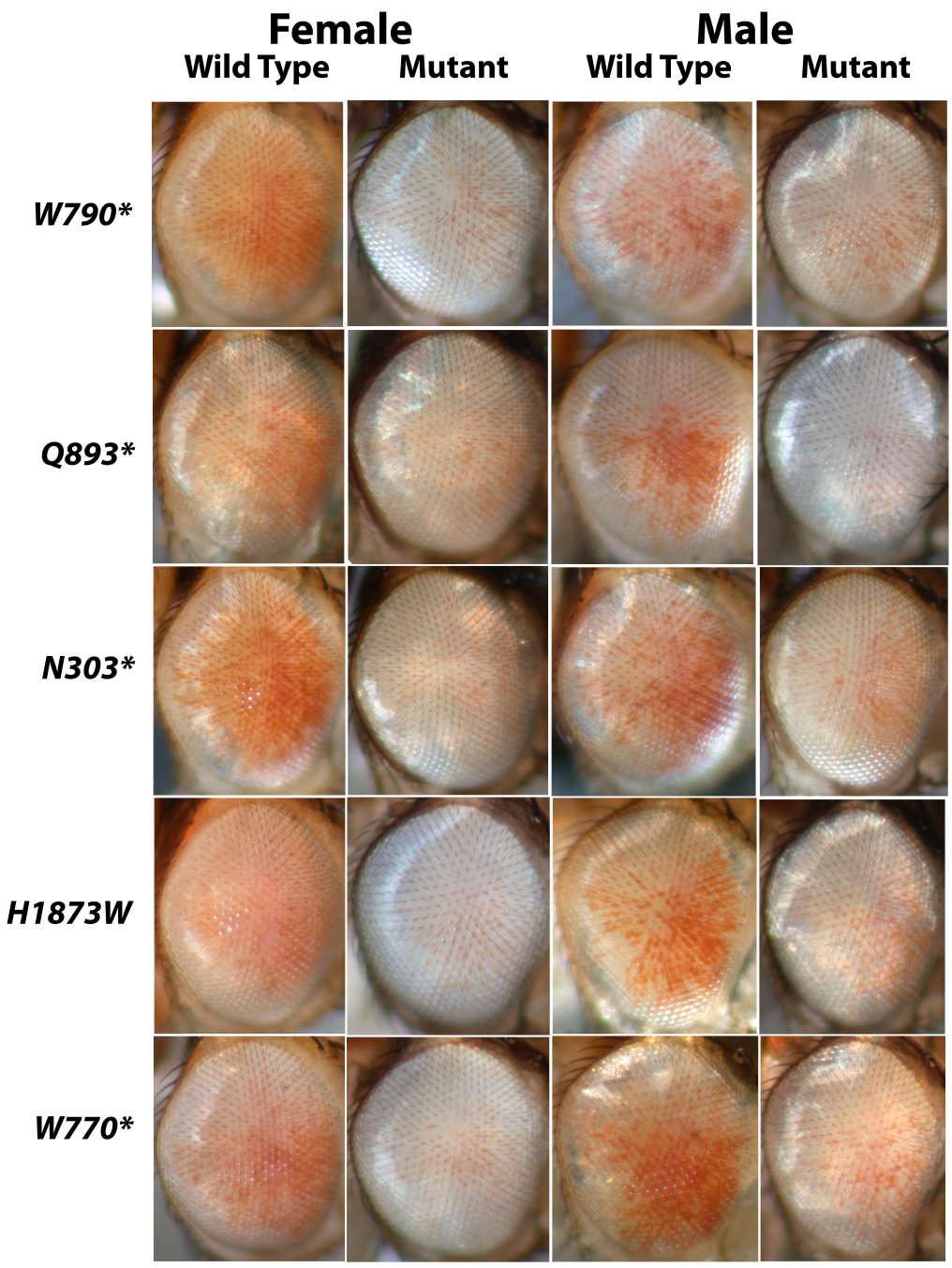


Figure 2.3 *ash1* mutants' enhance *E1*. Photographs of representative examples from each class of progeny from heterozygous *ash1* mutants crossed back to the parental *E1* stock used in this mutagenesis. *ebony* versus *ebony*$^{+}$ flies were compared for each sex. Flies are posed facing right.

(y^{-} w^{-}; *ash1* e^{-}/+ ♂ X w^{-}; dp^{-}; e^{-}; *E1*♀)

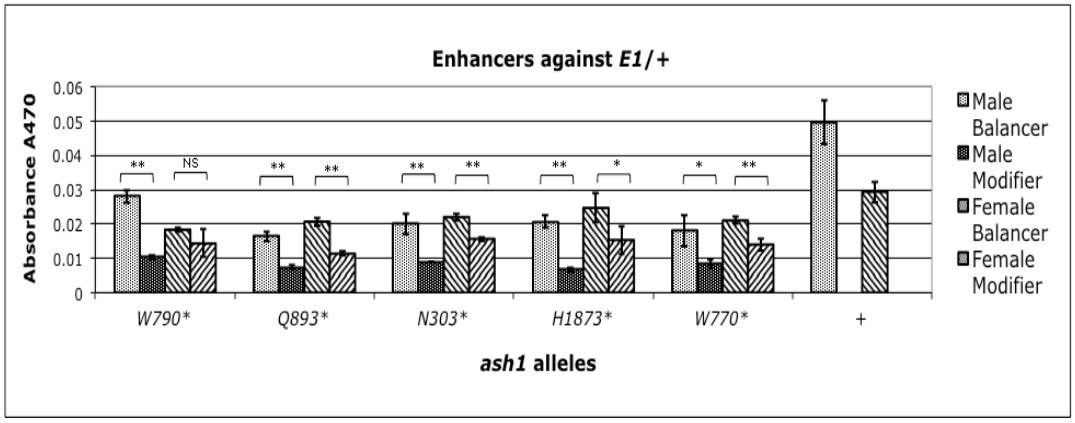


Figure 2.4 Pigment analysis of *ash1* mutants' effect on *E1*. Heterozygous *ash1* mutants were crossed back to the parental *E1* stock used in this mutagenesis. *ebony* versus *ebony*$^{+}$ flies were compared for each sex. (y^{-} w^{-}; *ash1* e^{-}/+ ♂ X w^{-}; dp^{-}; e^{-}; *E1*♀) Statistical significance between pairs (T-Test = 1-tailed, independent (unpaired, unequal variance)) is given above each mutant pair tested, * = $p< 0.05$, ** = $p<0.01$, NS= not significant.

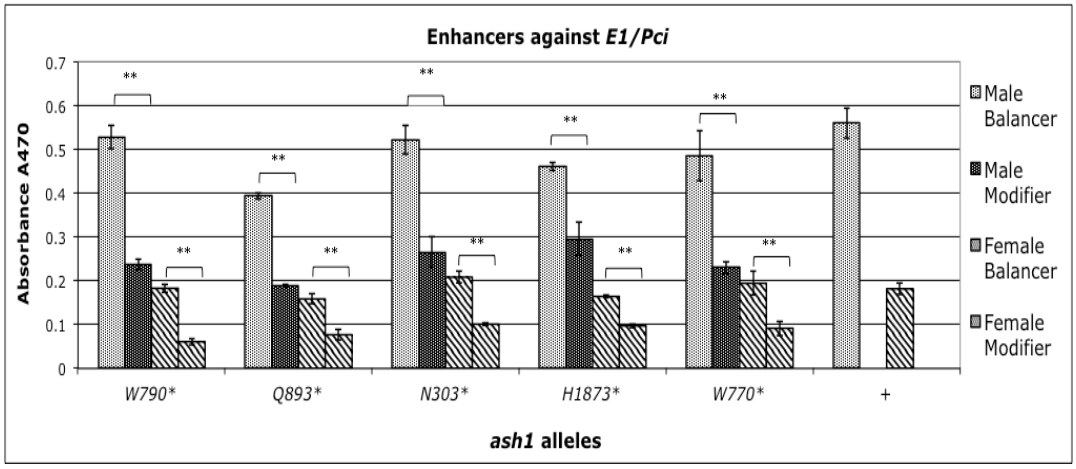


Figure 2.5 Pigment analysis of *ash1* mutants' effect on *E1/Pci.* Pigment assay of heterozygous *ash1*; *Pci* mutants crossed back to the parental *E1* stock used in the mutagenesis. In the progeny (y^- w^-; *ash1* e^-/+;*Pci* ♂ X w^-; dp^-; e^-; *E1*☿), *ebony* versus *ebony*$^+$ flies were compared for each sex. Statistical significance between pairs (T-Test = 1-tailed, independent (unpaired, unequal variance)) is given above each mutant pair tested, * = $p< 0.05$, ** = $p<0.01$, NS= not significant.

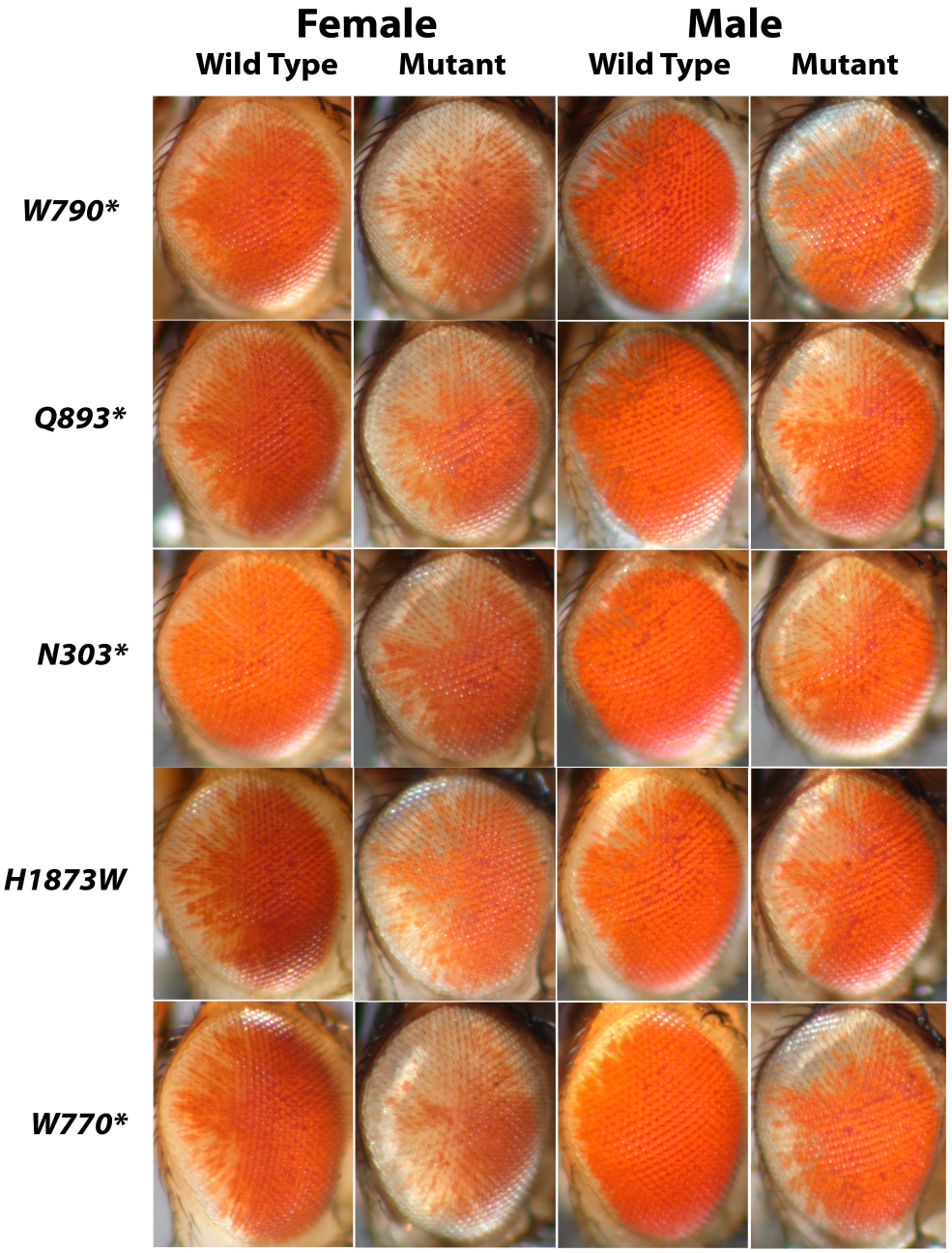


Figure 2.6 Photographs of *ash1* mutants' enhancement of *E1/Pci.* Representative examples from each class of progeny from heterozygous *ash1*; *Pci* mutants crossed back to the parental *E1* stock used in this mutagenesis. Note the pattern of enhancement starting at the posterior edge of the eye weakening as it progresses anteriorly. *ebony* versus *ebony*$^{+}$ flies were compared for each sex. Flies are posed facing right. (*y*$^{-}$ *w*$^{-}$; *ash1 e*$^{-}$/+;*Pci* ♂ X *w*$^{-}$; *dp*$^{-}$; *e*$^{-}$; *E1*♀)

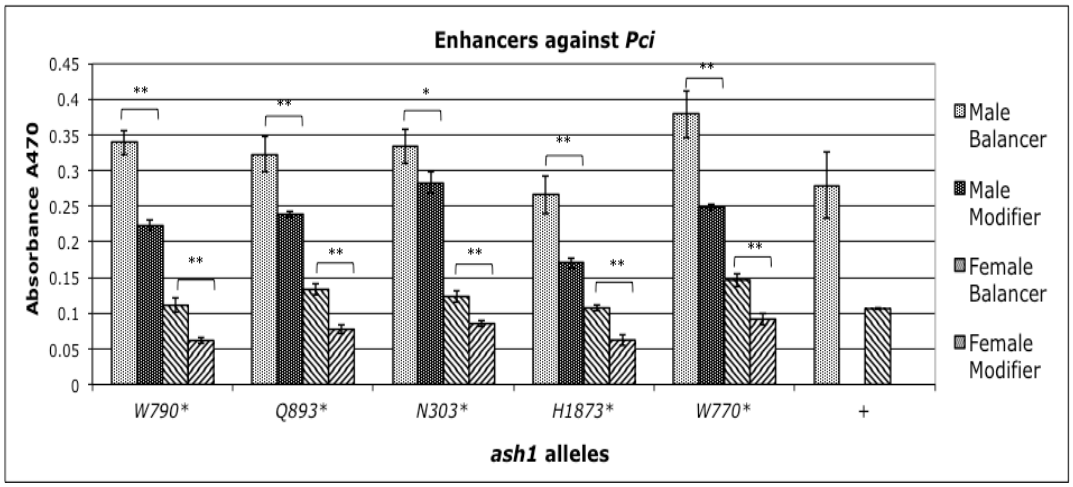


Figure 2.7 Pigment analysis of *ash1* mutants' effect on *Pci*/+.
Heterozygous *ash1* mutants crossed back to the parental *Pci* stock *E1* was derived from. *ebony* versus *ebony*$^{+}$ flies were compared for each sex. (y^{-} w^{-}; *ash1* e^{-}/+ ♂ X *w*; dp^{-}; e^{-}; *Pci*♀) Statistical significance between pairs (T-Test = 1-tailed, independent (unpaired, unequal variance)) is given above each mutant pair tested, * = $p< 0.05$, ** = $p<0.01$, NS= not significant.

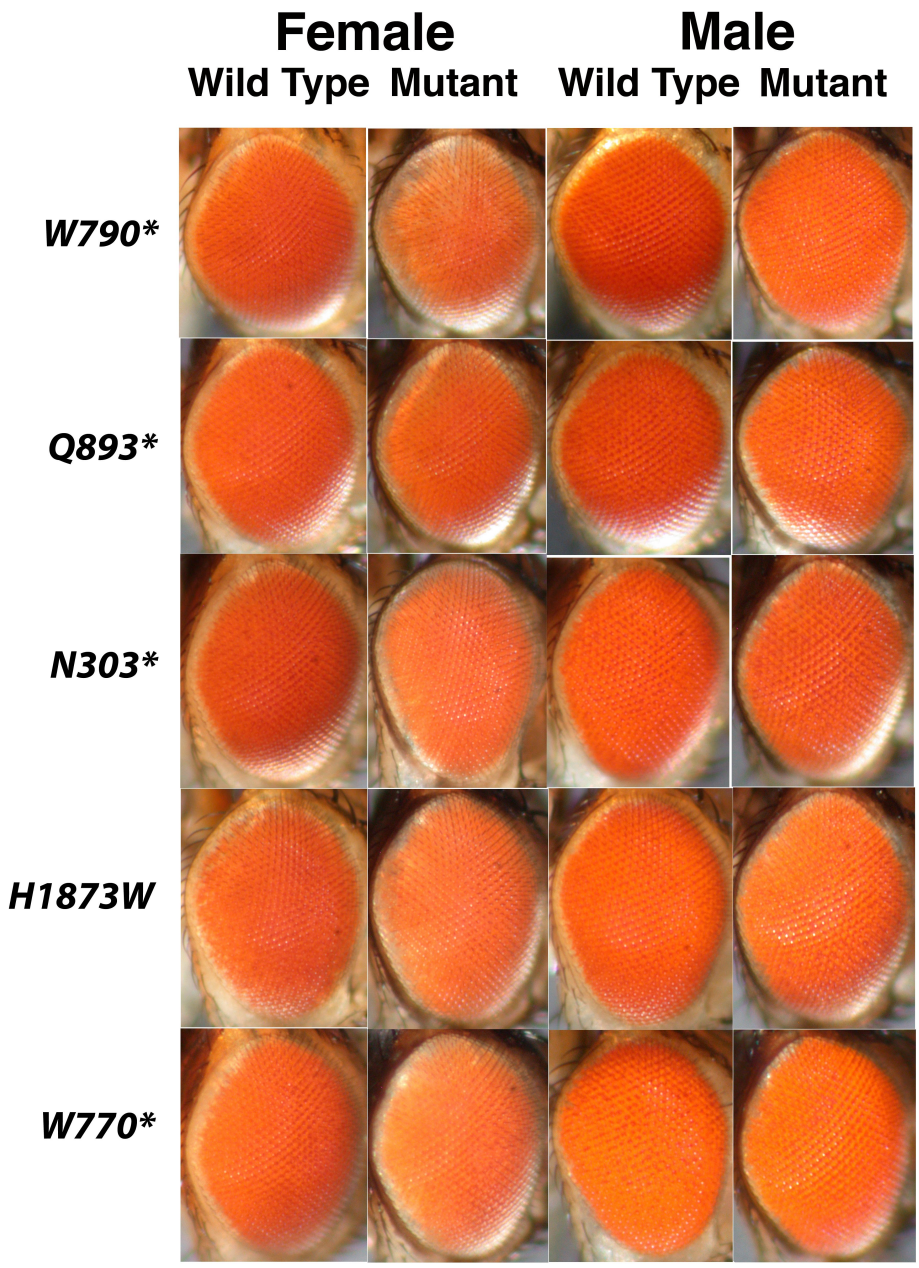


Figure 2.8 *ash1* mutants' enhance *Pci*/+. Photographs of representative examples from each class of progeny from heterozygous *ash1* mutants crossed back to the parental *Pci* stock *E1* was derived from. *ebony* versus *ebony*$^{+}$ flies were compared for each sex. Flies are posed facing right. (y^{-} w^{-}; *ash1* e^{-}/+ ♂ X *w*; dp^{-}; e^{-}; *Pci*♀)

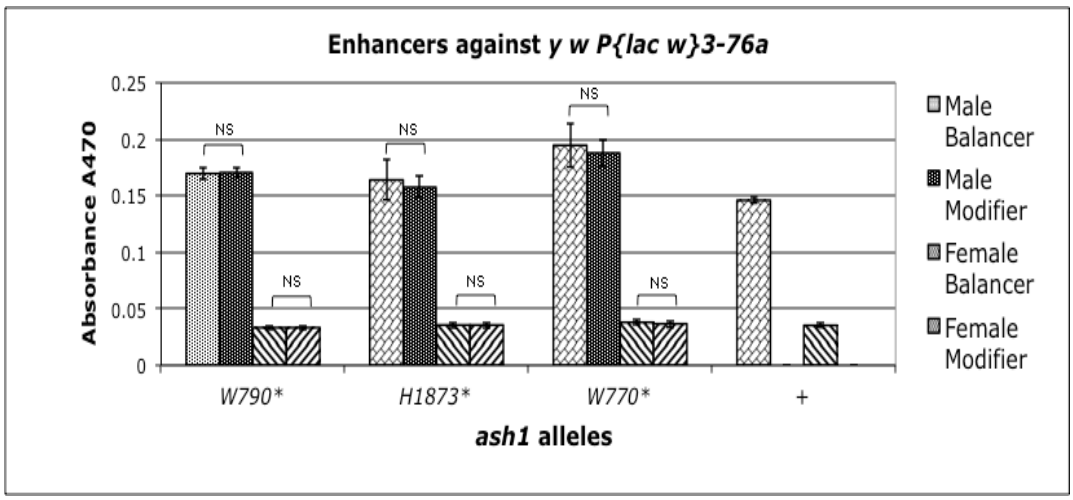


Figure 2.9 Pigment analysis of *ash1* mutants' effect on *P{lacW}3-76a.* Heterozygous *ash1* mutants were crossed to *P{lacW}3-76a*, an insertion of a transgene identical to *P{lacW}cl*^{Dplac} at 18A1 on the X chromosome. *Sb*$^{+}$ *Ser*$^{+}$ flies were compared to *Sb Ser* sibs for each sex. (*y*$^{-}$ *w*$^{-}$; *ash1 e*$^{-}$ /*TM3 Sb Ser e*$^{-}$ ♂ X *y*^{1} *w* P{lacW}3-76a* ☿) Statistical significance between pairs (T-Test = 1-tailed, independent (unpaired, unequal variance)) is given above each mutant pair tested, * = p< 0.05, ** = p<0.01, NS= not significant.)

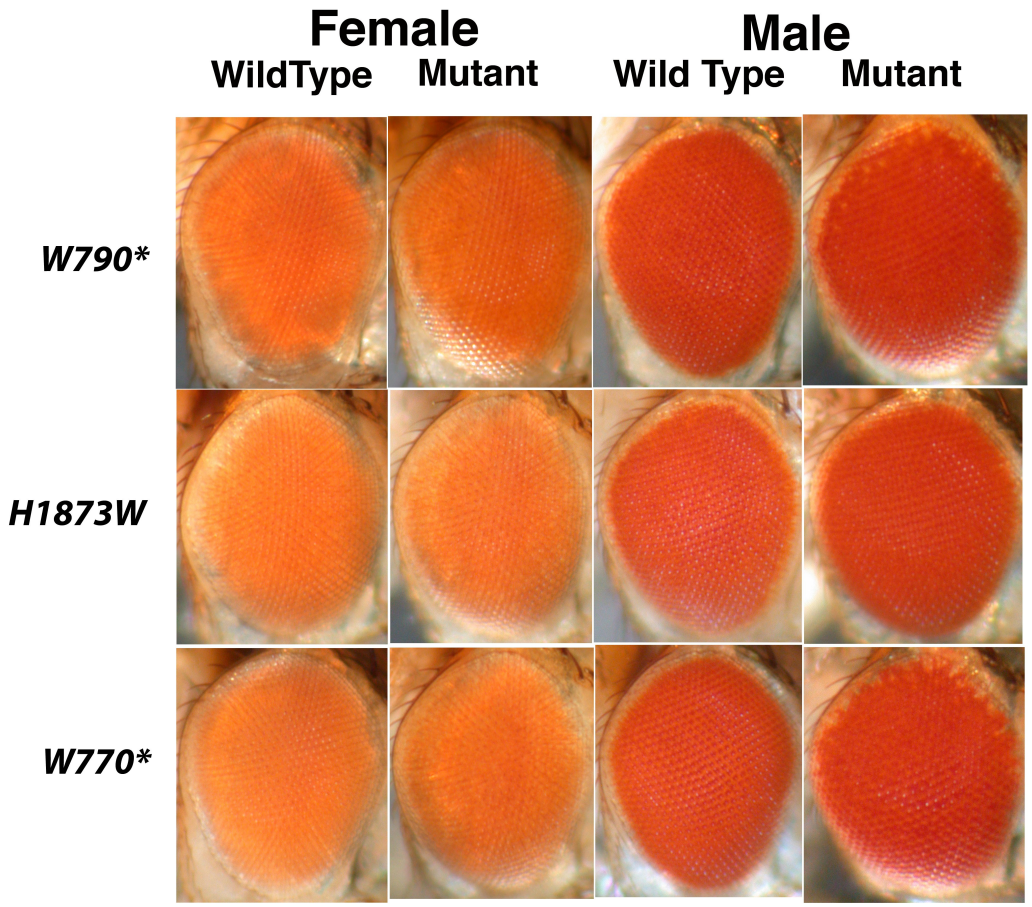


Figure 2.10 *ash1* mutants do not affect *P{lacW}3-76a.* Photographs of representative examples from each class of progeny from heterozygous *ash1* mutants crossed to *P{lacW}3-76a*, an insertion of a transgene identical to *P{lacW}ci*^{Dplac} at 18A1 on the X chromosome. *Sb*$^{+}$ *Ser*$^{+}$ versus *Sb Ser* flies were compared for each sex. (*y*$^{-}$ *w*$^{-}$; *ash1 e*$^{-}$/*TM3 Sb Ser e*$^{-}$ ♂ X *y*^{1} *w* P{lacW}3-76a* ♀)

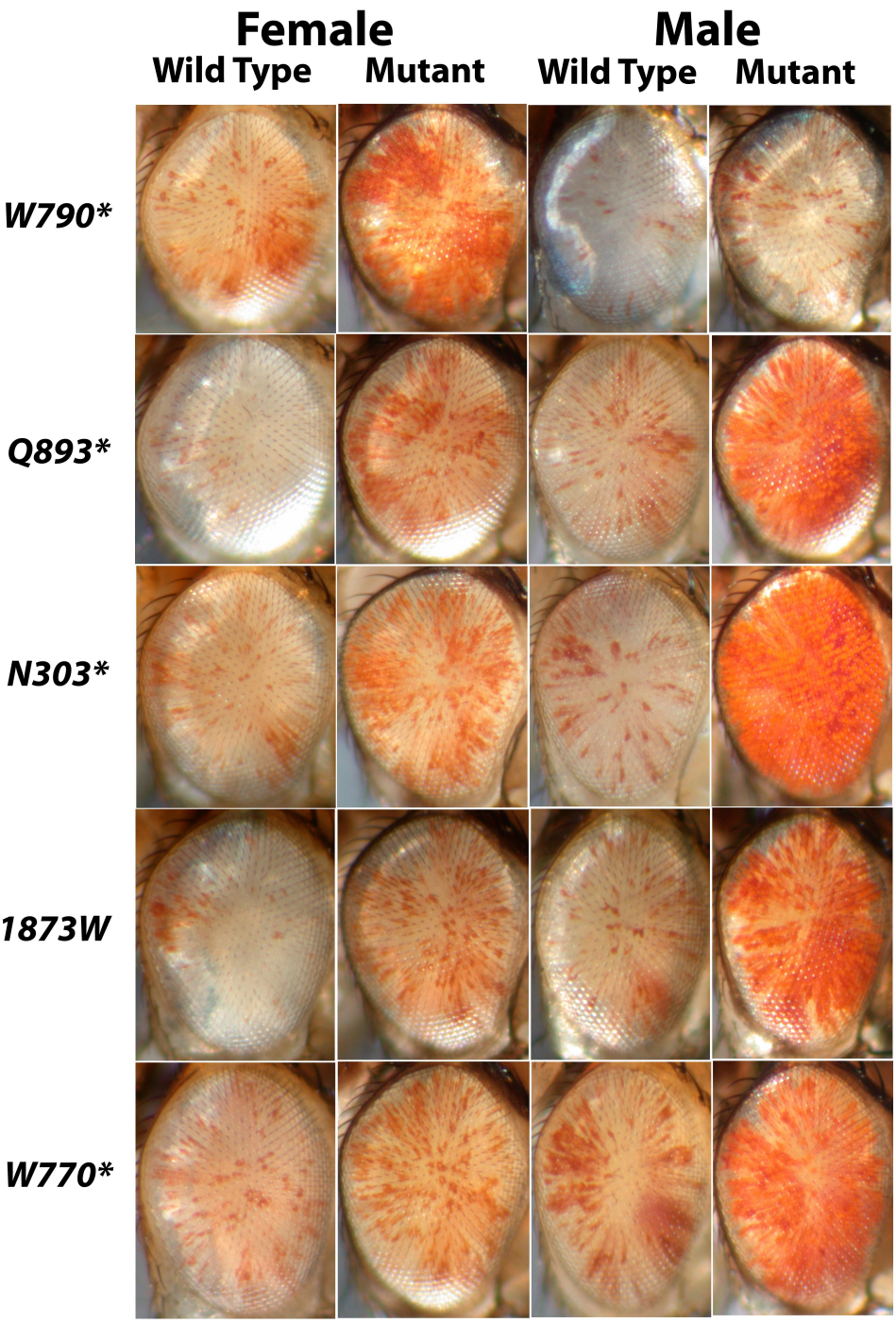


Figure 2.11 Photographs showing ash1 mutants effect on w^{m4}. Representative examples from each class of progeny from heterozygous *ash1* mutants crossed to w^{m4}. *ebony* versus $ebony^{+}$ flies were compared for each sex. Flies are posed facing right. (y^{-} *w*; *ash1* e^{-}/+ ♂ X w^{m4}; e^{-}; ☿)

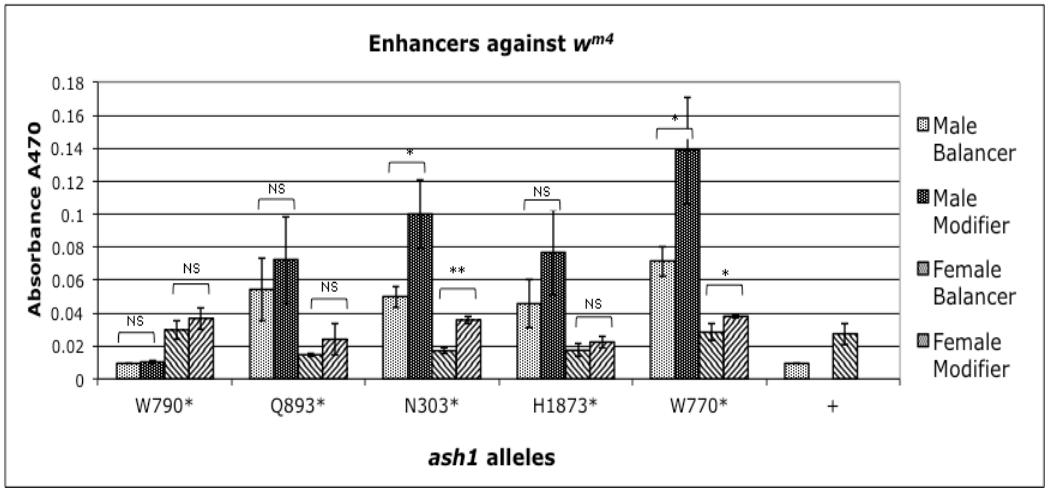


Figure 2.12 Pigment analysis of *ash1* mutants' effect on w^{m4}. Heterozygous *ash1* mutants were crossed to w^{m4}, *ebony* versus $ebony^{+}$ flies were compared for each sex. (y^{-} *w*; *ash1* e^{-}/+ ♂ X w^{m4}; e^{-}; ☿) Statistical significance between pairs (T-Test = 1-tailed, independent (unpaired, unequal variance)) is given above each mutant pair tested, * = $p< 0.05$, ** = $p<0.01$, NS= not significant.)

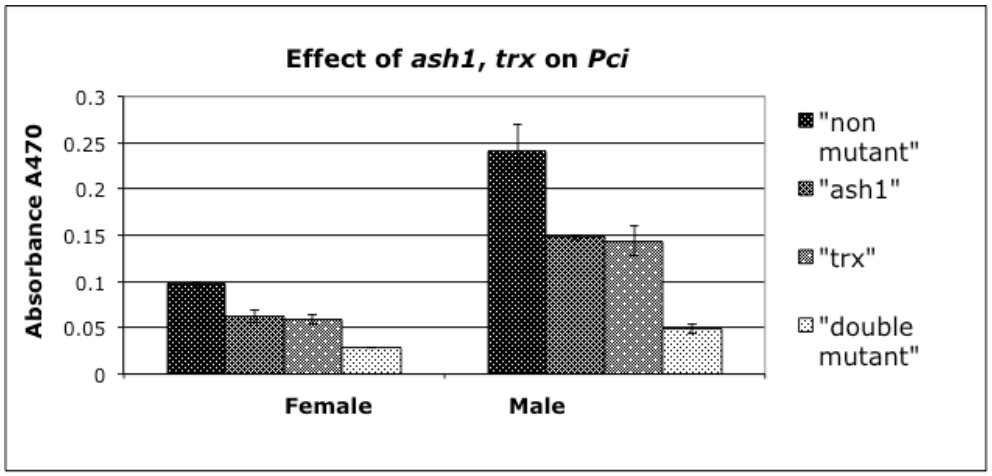


Figure 2.13 The effect of *ash1* and *trx* on *Pci*. Pigment analysis of flies from a cross of y^{-} *w*; $ash1^{H1873W}$ e^{-}/+ ; Pci ♂ X y^{-} w^{-}; trx^{R1583*} e^{-}/*TM3 Sb Ser* e^{-}☿. Statistical significance (T-Test = 1-tailed, independent (unpaired, unequal variance)) between both mutants and balancers $p<0.01$, between double mutants and single mutants $p<0.01$, between $ash1^{N303*}$, trx^{S2582*} not significant, between $ash1^{H1873W}$, trx^{R1583*} $p< 0.05$.

Chapter 3 – Manuscript for *trx* research

This chapter is a version of a manuscript in preparation for submission for publication.

trx mutations enhance *P* element dependent silencing (PDS) in *Drosophila melanogaster*.

Allen McCracken and John Locke*

Department of Biological Sciences,
University of Alberta,
Edmonton, Alberta, Canada, T6G 2E9

*Corresponding Author:
John Locke
Department of Biological Sciences
CW 405, Biological Sciences Bldg
University of Alberta
Edmonton, Alberta
Canada, T6G 2E9
Email: john.locke@ualberta.ca
Phone: 780-492-2193
Fax: 780-492-9234

Abstract

In *Drosophila melanogaster*, transgenes containing the mini-w^+ transgene normally show a uniformly expressed red colour throughout the adult eye; however, in previous work in this lab, we recovered a spontaneous mutation of *P{lacW}ci*^{Dplac} called *P{lacW}ci*$^{Dplac\ E1}$ (*E1*) that has a variegated phenotype thought to be due to an adjacent *gypsy* element insertion. We describe here three mutants from a screen for genetic modifiers of *E1* variegation, which contain lesions in *trx* that enhance this variegated phenotype. In addition to enhancement of variegation at *E1*, these mutant alleles also reduce expression of (silence) *P{lacW}ci*^{Dplac} itself. Loss of *trx* is phenotypically equivalent to loss of *ash1*, and double heterozygote mutants of *ash1* and *trx* dose-dependently enhance variegation in an additive manner. This is consistent with a model where ASH1 recruits TRX to the *ci* region to maintain it in a transcriptionally active state.

Introduction

In *Drosophila melanogaster,* expression of the *white* (*w*) gene is required for the import of colourless precursors into the adult eye where they are converted to the typical red pigments. Lack of white gene expression results in a white eye phenotype. Because the *white* gene is cell autonomous its expression can vary from cell to cell and provides an easily visible marker for changes in gene expression. For example, chromosomal translocation of the normally euchromatic *white* gene next to regions of heterochromatin causes random silencing of *white* gene expression. This silencing is maintained in a semi-stable state through multiple cell divisions (clonal inheritance) during tissue development and a mosaic pattern of gene expression or a variegated (mottled) phenotype. This Position Effect Variegation (PEV - Muller 1930) was the first known example of epigenetic inheritance. A model to explain the variegation is the *cis*-spreading of heterochromatin from a pericentric initiation site past the rearrangement break point to inhibit access by the transcriptional complexes resulting in a transcriptional repression state (Locke *et al*. 1988). Homolog position and pairing also contribute to PEV as shown by the bw^{D} allele of the *brown* (*bw*) eye color gene, which dominantly *trans*-inactivates the wild-type homologue in bw^{+}/bw^{D} heterozygotes depending upon the degree of bw^{D}-*2Rh* association (Dernburg *et al*. 1996; Csink & Henikoff 1996; Belyaeva *et al*. 1997).

Similar variegated phenotypes are observed for some $white^{+}$ transgenes. We are using $P\{lacW\}ci^{Dplac}$ (hereafter *Pci*), a chromosome 4 transgene insert between *cubitus interruptus (ci)* and *Ribosomal protein S3A* (*RpS3A*), which was originally isolated as an enhancer trap matching *ci* gene expression (Eaton & Kornberg 1990). In M strain flies, the *Pci* w^{+} transgene expresses uniformly throughout the adult eye. However, in flies containing *P* elements (*P* strains) or their derivative elements such as *KP*s, variegation occurs. This results in mosaic patches of white

ommatidea on a red eye background. This *P* element dependent silencing (PDS) is a phenomenon phenotypically and mechanistically similar to hPEV. Other w^+ transgenes inserted at this chromosome 4 location near the *Pci* insert, also show PDS variegation. This random silencing of the w^+ minigene during development, indicates a phenomenon similar to hPEV because variegating inserts respond to hPEV modifier mutations in *Su(var)205* and *3-7* in the same manner as w^{m4} and other centromeric PEV lines (Sun *et al*. 2000; Wallrath & Elgin 1995; Bushey & Locke 2004). As in PEV, PDS is also position dependent, as insertions of the same transgene in other locations do not display PDS. Also, PDS is suppressed by the chromosomal translocation of *Pci* away from the centromere (Bushey & Locke 2004). The variegated phenotype and position dependence suggest PDS and heterochromatic PEV share factors in a common mechanism of gene silencing.

A spontaneous mutant of *Pci*, called $P\{lacW\}ci^{DplacE1}$ (hereafter *E1*), was recovered that displayed a variegated eye phenotype in M stain flies (absence of *P* elements) and a complete white eye (silenced) phenotype in combination with *P{ry+ SalI}89D.* As with PDS, variegation of *E1* was weakly suppressed by *Su(var)205*, while *Su(var)3-7* mutant alleles had a much stronger effect. Molecular analysis of *E1* revealed a *gypsy* element insertion approximately 1 kb upstream from $P\{lacW\}ci^{Dplac}$. Testing of *E1* with mutants of *su(Hw)* and *mod(mdg4)* showed that the variable silencing (variegation) was not dependent on the *gypsy* insulator function *per se* (Bushey & Locke 2004). This does not, however, rule out some interaction between the *gypsy* insulator and the recently discovered *wari* element at the 3' end of the w^+ transgene (Chetverina *et al.* 2008). The *E1* allele is able to dominantly *trans*-silence the w^+ expression when paired with *Pci*, however if unpaired, as in translocations of *Pci*, this silencing is absent or weakened (Bushey & Locke 2004). This pairing is similar to that

in bw^{+}/bw^{D} heterozygotes (Dernburg *et al*. 1996; Csink & Henikoff 1996; Belyaeva *et al*. 1997).

The *E1* phenotype of an intermediate variegated phenotype has allowed for an easier screen for enhancers of PDS that show a closer to full white eye phenotype. Using this system, we have screened for and recovered enhancers of *w* expression that include both epigenetic regulators of global heterochromatin and those specific to *ci* regulation as well. We describe here the isolation and characterization of 3 *trx* mutants, members of one of the 5 complementation groups identified as an enhancer in this screen.

Materials and Methods

Drosophila stocks and mutations: *D. melanogaster* mutations are described previously (Lindsley & Zimm 1992) unless otherwise cited,. The *P{lacW}ci*^{Dplac} allele (Eaton & Kornberg 1990) is a *P{lacZ*$^{P\T.W}$*w*$^{+mC}$ *amp*^{R} *ori = lacW}* insertion on chromosome *4* approximately 3 kbp upstream from the *ci* locus while *P{lacW}ci* ^{DplacE1} has an additional gypsy element insertion ~ 1kb further upstream (Bushey & Locke 2004).

Fly stocks (Table A1) were maintained at room temperature on standard yeast/cornmeal medium.

The source of the three mutations in this study was *w; dp; e; P{lacW}ci* ^{DplacE1} males that were mutagenized with 25 mM EMS (Ashburner 1989), mated to *y w* virgin females and screened for an enhanced *E1* eye phenotype (reduced pigment in a variegated eye) in the progeny. Potential mutants were mated to *w*$^{-}$*; dp*$^{-}$*; e*$^{-}$*; P{lacW}ci* ^{Dplac} flies to; 1) confirm transmission of the enhancer; 2) confirm dominant silencing of Pci; and 3) determine chromosomal location for stocking. Mutant *trx* alleles were kept as balanced stocks over *TM6B, Hu e*$^{-}$.

Genetic Mapping: Both the enhancer and recessive lethal phenotypes were genetically mapped by recombination relative to *Gl, Sb, H, and e* markers. During the enhancer mapping recombinants to both sides of the enhancer were collected and tested for retention of the enhancer and lethal phenotypes to confirm linkage. The recessive lethal location for each mutation was refined by complementation testing against deletions in the region. In each test, at least 100 progeny were scored and the lack of mutation/deletion progeny was considered sufficient to deem the combination lethal.

DNA sequencing: Mutants were crossed back to the parental unmutagenized stock, the genomic DNA was isolated and used to amplify overlapping gene segments by PCR (Tables A4, A6), which were then

sequenced. Chromatogram double peaks were used to identify point mutations. All polymorphisms and mutations were confirmed by sequencing both strands.

Eye pigment assays: Relative expression of the w^+ gene was determined using a modification of the method of Ephrussi and Herold (1944) to measure the amount of brown eye pigment. Heads from 6-8 day old adult flies were stored at -20° until extracted. Three replicate samples of 10 heads of each genotype were extracted in 200 μL of acidified ethyl alcohol (1% HCl in ethanol) with shaking for ~48 hours. Absorbance at 470 nm was then measured for 150 μL of each sample in a 96 well Costar flat bottom plate using a Bio-Tek PowerWave XS spectrophotometer. A Zeiss stereo-microscope and Nikon Coolpix 995 digital camera were used to photograph adult fly eyes under mineral. For both assays the parental unmutagenized chromosome was used as a control.

Results

Recovery of *trx* mutations in a screen for enhancers of E1

We screened ~44,000 progeny from a cross designed to produce mutations that phenotypically enhance w^+ silencing in *E1* and recovered 58 mutations that transmitted and segregated with either the second or third chromosome. These mutations were grouped into five straightforward recessive lethal complementation groups and three complex complementation groups, as well as many singles that were not pursued any further. Of a simple group with 3 alleles (labeled *3a26a*, *3b77a*, and *5b19a*), *3b77a* genetically mapped to 3-51.8 (n=704) while *5b19a* mapped to 3-50.2 (n=678) indicating both enhancer phenotypes mapped to the same locus. Recombinants recovered to the left and right of the enhancer, when crossed back to other members of the complementation group, retained the common lethal phenotype indicating linkage between the two phenotypes. The recessive lethality common to *3b77a* and *5b19a* genetically mapped to 3-46.3 (n=474) while that between *3a26a* and *3b77a* mapped to 3-45.5 (n=443) indicating a single common lethality that mapped to the same locus as the enhancer phenotype. Deficiency mapping further refined this position to include 3R:10,103,658… 3R:10,114,795, within Df(3R)Exel6267 but not Df(3R)BSC487, in complementation tests involving both *3a26a* and *5b19a*. This includes only the *trx locus*.

All 3 alleles failed to complement the recessive lethality of the extant alleles trx^1 and trx^{E2} and we therefore conclude these are mutant alleles of *trx*.

We wished to test trx^1 and trx^{E2} for enhancement of *E1*, however PCR tests using internally directed *P* element terminal repeat primers (Tables A4, A8) amplified an approximately 1150 nt band in both trx^1 and trx^{E2} stocks indicating the presence of *KP* elements (O'Hare & Rubin 1983, Black *et al*. 1987). Such elements are known to enhance PDS (Bushey &

Locke 2004; Sameny & Locke 2011). Efforts to recombine out these *KP*s were unsuccessful and thus we couldn't assay trx^{1} and trx^{E2} 's effect on *E1* enhancement.

DNA sequencing of *trx* mutant alleles

DNA sequencing of the entire coding region of our mutants, in heterozygotes with the parental unmutagenized chromosome, confirmed that each had a change within *trx* that altered the predicted amino acid coding sequence by introducing premature stop codons (Table 3.1). Mutation *5b19a* had a thirteen bp deletion resulting in a +1 frameshift and multiple premature stop codons that predicted termination before the most carboxy terminal RING domain/PHD finger and its carboxy-terminal SET domain. Mutations *3a26a* and *3b77a* both had the same serine to amber mutation that predicts termination prior to the carboxy-terminal SET domain. Our three alleles are now designated trx^{R1583*}, $trx^{S2582*a}$, and $trx^{S2582*b}$.

Quantification of Enhancer Phenotype

Variegation of *white* in *E1* heterozygotes was visibly enhanced for all alleles in both sexes, frequently producing flies with near white eyes (Figure 3.1a). Pigment measurements of *E1* enhancement showed that male *trx* mutants had approximately half the pigment of non-mutant internal control flies from the same cross (Figure 3.1b). Female *trx* mutants also showed enhancement compared to non-mutant internal controls from the same cross, with all being significant at the 95% confidence limit (Figure 3.1b). This quantitation confirms the visual appearance that all of our *trx* alleles enhance silencing of *white* in *E1*.

To see if the silencing also affected *Pci* we examined the effect of *trx* mutants in *Pci*/+ flies. All three *trx* mutants showed significantly reduced pigment compared to non-mutant internal control flies from the same cross, with trx^{S1583*} and $trx^{S2582*a}$ being significant at > 99% and $trx^{S2582*b}$ > 95% confidence limits (Figure 3.2b). Flies had a similar pattern of weak

silencing starting at the posterior edge of the eye and decaying rapidly as it progressed anteriorly (Fig. 3.2a).

To address the possibility that *trx* product was interacting directly with the w^+ transgene within the *P{lacW}ci*^{Dplac} insert, we tested y^1 *w* P{lacW}3-76a* /+ flies that lacked *E1* or *Pci* with two of our *trx* mutants. The *P{lacW}3-76a* transgene is the same as in *P{lacW}ci*^{Dplac} but is located at 18A1 (60.7 cM) on the X chromosome and presents a full red eye. Our *trx* mutant alleles were virtually indistinguishable from non-mutant control flies of the same cross indicating that *trx*'s silencing effect is dependent upon the insert location and not on the w^+ transgene *P{lacW}ci*^{Dplac} construct itself (Figure 3.3 a,b). *trx* has been shown not to affect hPEV (Buchner *et al*. 2000).

Given that *trx* mutants enhanced PDS in *P{lacW}ci*^{Dplac}, but did not affect the same construct in *P{lacW}3-76a*, we asked whether or not *trx* had any effect on hPEV at w^{m4} by crossing y^- *w*; *trx** e^-/+ males to w^{m4}; e^- virgin females. Variegation of w^{m4} in flies of both sexes was visibly and quantitatively suppressed by the mutants S2582*[a] and R1583* isolated in this study (Figure 3.4a,b).

<u>Interactions between *ash1* and *trx* - Double Mutant Enhancer Phenotype</u>

One of the other complementation groups that we found in this screen, which influenced the transcriptional status of the *ci* region was *ash1*. This group of 5 *ash1* alleles silenced expression of both *E1/+* and *E1*/*Pci* by approximately 50 %. Silencing of *Pci* was slightly less but still significant at $p<.01$ for all tests but one which was significant at $p<.05$ indicating that silencing by ASH1 did not require the presence of the proximal *gypsy* insert. *P{lacW}3-76a*, a transgene construct identical to that in *P{lacW}ci*^{Dplac} but located at 18A1 on the X chromosome was not affected by any of our *ash1* alleles indicating that silencing was not construct specific. We took these results to indicate that *ash1* normally acted in regulating the *ci* locus (McCracken & Locke 2012). ASH1 is known to

interact with, and presumably recruit, TRX at other loci (Kuzin *et al*. 1994, Rozovskaia *et al*. 1999, Tillib *et al*. 1999); therefore, we asked whether or not they interacted in regulation of *ci.*

To determine if *ash1* and *trx* had synergistic effects at *Pci*, we crossed *y w: ash1*^{N303*} *e/+*; *Pci* males to *y w*; *trx*^{S2582*a} *e/TM3, Sb Ser e* females. The progeny of the appropriate genotypes were collected, photographed, and the eye pigment measured (Figure 3.5a,b). Both *ash1* and *trx* single mutant flies had the same pattern of weak silencing starting at the posterior edge of the eye and decaying rapidly as it progressed anteriorly. However, in the double mutant *y w*: *ash1*^{N303*} *e /trx*^{S1583*a} *e*; *Pci/+* flies variegation was much more noticeable and continued further to the anterior of the eye (Figure 3.5a). Both *ash1* and *trx* single mutants had about half the pigment of non-mutant internal balancer control flies with no significant difference between *ash1* and *trx* flies. The double mutant flies had about one quarter the pigment of non-mutant internal control flies indicating a cumulative, dosage dependant effect on silencing for both *ash1* and *trx* (Figure 3.5b).

Discussion

We have shown that the 3 *En*(*PDS*) in one of the complementation groups described in this study are alleles of *trx*. Both the enhancement of PDS and recessive lethal phenotypes genetically mapped to the *trx* locus and they could not be separated by recombination. All 3 alleles failed to complement all deficiencies tested that uncover *trx* and failed to complement both extant alleles trx^{1} and trx^{E2}. Sequence analysis as described above confirmed that each *trx* mutant had a sequence change resulting in the introduction of a premature stop codon. One, trx^{S1583*}, had a thirteen nucleotide deletion that resulted in a +1 frameshift which introduced multiple premature stop codons that predicted termination before TRX's most carboxy terminal RING domain/PHD finger and its carboxy-terminal SET domain while $trx^{S2582*a}$, and $trx^{S2582*b}$ both had a serine to amber mutation that predicted termination prior to TRX's carboxy-terminal SET domain. All three alleles likely represent at least partial, if not complete, loss of function alleles. Despite trx^{S1583*} occurring ~1000 amino acid residues earlier than $trx^{S2582*a}$, and $trx^{S2582*b}$, it does not represent a phenotypically stronger allele of *trx* (Figures 3.1, 3.2).

trx^{1} is the founding and eponymous member of the *trx*G, and was originally characterized as a gene that caused homeotic transformations when mutated (Ingham & Whittle 1980; Breen & Harte 1991; Ingham 1998). Its phenotypes of transformation of haltere-to-wing, and of posterior abdominal segments towards more anterior segments, resembles loss-of-function mutations in the *Bithorax* complex, one of the two *Drosophila melanogaster* HOX gene clusters, thus implicating *trx* as a positive regulator of HOX gene function. Subsequent genetic and molecular analyses showed that TRX is required to maintain target genes in transcriptionally active states throughout development by counteracting the repressive effects of *Polycomb* group proteins (Ringrose & Paro 2004; Schuettengruber *et al*. 2007). Thus, as for *Pc*G genes, *trx* was postulated

to be involved in the epigenetic inheritance of the expression states of HOX genes (Ingham 1998; Cavalli & Paro 1999). The subsequent identification of other *trx*G proteins such as *ash1* was based on genetic screens for second site mutations with phenotypes indicating a loss of HOX genes function, or for suppressors of *Pc*G-dependent mutant phenotypes (Kennison & Tamkun 1988), which has led to the idea of *trx*G proteins being antagonists of *Pc*G-dependent gene silencing.

trx is a SET domain (***S****u(var)3-9*, ***E****nhancer of zeste*, and ***t****rithorax*) domain containing protein capable of binding to and methylating histone H3K4 specifically in chromatin (Czermin *et al*. 2002; Smith *et al*. 2004; Katsani *et al*. 2001). Methylation of histone H3 on lysine 4 (H3K4) is generally associated with transcriptionally active regions of chromatin (Eissenberg & Shilatifard 2010) implying that TRX maintains target gene activity partly through the methylation of H3K4.

Chinwalla *et al*. (1995) used dual fluorescence confocal microscopy to show co-localization of *trithorax* and *Polycomb* proteins at 30 sites on polytene chromosomes sites indicating that many of their chromosomal binding sites coincide and that interactions between them may be a significant feature of their mode of action. It has since been shown that both *Pc*G and *trx*G regulators are recruited to specific chromosomal elements in the 5' regulatory region of target genes and that the same element could act as either an activating or silencing regulatory element. In the transcriptionally repressed state, the elements facilitate the recruitment of *Pc*G proteins and are therefore designated as *Polycomb* response elements (PREs). While in the activated state, they recruit *trx*G proteins and are termed *trithorax* response elements (TREs) (Orlando 2003; Ringrose & Paro 2007).

The *trx*G and *Pc*G proteins have not been previously shown to regulate *ci*; however, two groups have independently demonstrated PRC1 and

PRC2 binding sites at *ci*, evidence of a PRE/TRE (Schwartz *et al*. 2006, Tolhuis *et al*. 2006). Recent chromatin immunprecipitation studies by Schuettengruber et al. (2009) suggest a complicated relationship between TRX protein binding sites and H3K4me3. TRX contains a cleavage site similar to its human homologue's (MLL) cleavage site 2 (QMD/GVDD vs QLD/GVDD) in an analogous position in the protein. *Drosophila* TASPASE 1 cleaves wild-type TRX, but not TRX with a QMD/GVDD to QMAAVDD mutation in the cleavage site (Capotosti *et al*. 2007). They also found that the SET containing carboxy-terminal fragment of TRX (TRX-C) showed high affinity to PcG binding sites and limited overlap with sites of H3K4me3, whereas the non-SET containing amino-terminal fragment (TRX-N) bound mainly to active promoter regions tri-methylated on H3K4. These distinct distributions of the N- and C-terminal domains of TRX are consistent with its proteolytic cleavage in a manner analogous to that seen with MLL (Hsieh *et al*. 2003). A 271-amino-acid deletion (trx^{E3}) that spans the cleavage site displays defective *antennapedia* complex, but not *bithorax* complex gene expression, suggesting that TRX cleavage plays a selective role in its function (Breen 1999; Mazo *et al*. 1990; Sedkov *et al*. 1994).

Given our findings that: 1) trx acts at the *ci* locus; 2) two of our mutants affect only TRX-C, the TRE binding fragment, while leaving TRX-N intact; and 3) the third mutant, trx^{S1583*}, results in the complete elimination of TRX-C, it is reasonable to infer that the PRE/TRE found at *ci* is active and recruits TRX-C. The elimination of the carboxy terminal RING/PHD domains of TRX-N in trx^{S1583*} indicates that either (1) these domains are not essential for TRX-N binding at *ci*, or (2) that TRX-N and TRX-C functions are not additive at *ci* as trx^{S1583*}displays silencing of $P\{lacW\}ci^{Dplac}$ and of $P\{lacW\}ci^{DplacE1}$ indistinguishable from that of $trx^{S2582*a}$, and $trx^{S2582*b}$.

trx and *ash1* act together at the *ci* locus

We have shown that both *ash1* and *trx* act dose dependently at the *ci* locus, and that that this interaction is cooperative (Figure 3.5). This is consistent with ASH1 acting to recruit TRX to PRE/TREs at *ci*; if TRX acts to "lock" *ci* into a transcriptionally active state, then halving the dose of *trx* should halve the amount of TRX protein available to regulate the *ci* region. Similarly, if ASH1 acts to recruit TRX to *ci*, then halving the dose of *ash1* should halve the amount of ASH1 protein available to recruit TRX, resulting again in half of the amount of TRX protein recruited to regulate the *ci* region. If both *ash1* and *trx* doses are halved, then half the amount of ASH1 protein with half the amount of TRX to recruit should result in one quarter of the normal amount of TRX protein recruited to regulate the *ci* region. This should result in one quarter of the amount of *white* reporter being produced, which is what is seen (Figure 3.5). It has already been established that ASH1 and TRX act together in multimeric protein complexes. Kuzin *et al*. (1994) observed that association of TRX with polytene chromosomes is *ash1* dependent indicating a possible physical interaction between the two proteins. Rozovskaia *et al*. (1999) found that TRX and ASH1 proteins colocalize on salivary gland polytene chromosomes, coimmunoprecipitate from embryonic extracts and bind *in vivo* to *bxd*, which contains several *trx* response elements. In a study similar to this thesis, $ash1^{22}$ heterozygous mutant flies containing a mini-white reporter transgene inserted near *bxd* showed strong *white* reduction relative to wild type, similar to results obtained for trx^{B11} null allele using the same transgenic reporter lines (Tillib *et al*. 1999), suggesting that the activity of this TRE-containing region is both ASH1 and TRX dependent. Finally, using yeast two-hybrid assays they found that TRX and ASH1 SET domains interacted strongly with each other, and that mutations of conserved residues within either SET domain prevented this association. These results suggest that the association previously seen in embryos

between TRX and ASH1 is direct and involves their conserved SET domains. Collectively, these results suggest that TRX and ASH1 interact either within TrxG protein complexes or between complexes in close proximity on *bxd* to maintain Ubx transcription. My results suggest a similar interaction is taking place at *ci*; loss of either of TRX or *ash1*'s SET domain should prevent their association at *ci*, with concomitant reporter gene silencing as seen in Figure 3.4. We leave it to future researchers to fully delineate the binding elements (TREs), and to determine if their transcription is necessary for function as shown for *Ubx* TREs (Sanchez-Elsner 2006).

References

Ashburner, M. 1989. *Drosophila*, A laboratory manual Cold Spring Harbor Press, Cold Spring Harbor, New York.

Belyaeva, E.S., Koryakov, D,E., Pokholkova, G.V., Demakova ,O.V., and Zhimulev, I.F., 1997. Cytological study of the brown dominant position effect. *Chromosoma* **106**:124-132.

Breen, T.R., and Harte, P.J. 1991. Molecular characterization of the Trithorax gene, a positive regulator of homeotic gene expression in *Drosophila. Mech. Dev*. **35**:113–127.

Breen, T.R. 1999. Mutant alleles of the *Drosophila trithorax* gene produce common and unusual homeotic and other developmental phenotypes. *Genetics* **152**: 319-344.

Buchner, K., Roth, P., Schotta, G., Krauss, V., Saumweber, H., Reuter, G., and Dorn, R. 2000. Genetic and molecular complexity of the position effect variegation modifier *mod*(*mdg4*) in *Drosophila*. *Genetics* **155**: 141–157.

Bushey, D. and Locke, J. 2004. Mutations in *Su(var)205* and *Su(var)3-7* Suppress *P*-Element-Dependent Silencing in *Drosophila melanogaster*. *Genetics* **168**: 1395–1411.

Capotosti, F., Hsieh, J.J., and Herr, W. 2007. Species selectivity of mixed-lineage leukemia/trithorax and HCF proteolytic maturation pathways. *Mol. Cell Biol*. **27**:7063–7072.

Cavalli, G., and Paro, R. 1999. Epigenetic inheritance of active chromatin after removal of the main transactivator. *Science* **286:** 955–958.

Chetverina, D., Savitskaya, E., Maksimenko, O., Melnikova, L., Zaytseva, O., Parshikov, A., Galkin, A.V., and Georgiev, P. 2008. Red flag on the white reporter: a versatile insulator abuts the white gene in *Drosophila* and is omnipresent in mini-white constructs *Nucleic Acids Research* **36**: 929–937.

Chinwalla, V., Jane, E.P., and Harte, P.J. 1995. The *Drosophila* Trithorax protein binds to specific chromosomal sites and is colocalized with Polycomb at many sites. *EMBO J.* **14** :2056–65.

Csink, A. K., and Henikoff,S., 1996. Genetic modification of heterochromatic association and nuclear organization in *Drosophila*. *Nature* **381**: 529–531.

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

Dernburg, A.F., Broman, K.W., Fung, J.C., Marshall, W.F., Philips, J., Agard D.A., and Sedat, J.W., 1996 Perturbation of Nuclear Architecture by Long-Distance Chromosome Interactions *Cell* **85**: 745–759.

Eaton, S., and Kornberg, T.B. 1990. Repression of ci-D in posterior compartments of *Drosophila* by *engrailed*. *Genes Dev*. **4**: 1068–1077.

Eissenberg, J. C., and Shilatifard, A**.** 2010. Histone H3 lysine 4 (H3K4) methylation in development and differentiation. *Dev. Biol*. **339:**240–249.

Ephrussi, B., and Herold, J. L. 1944. Studies of eye pigments of *Drosophila*. I. Methods of extraction and quantitative estimation of the pigment components. *Genetics* **29**: 148–175.

Girton, J.R. and Johansen, K.M. 2008. Chromatin Structure and the Regulation of Gene Expression: The Lessons of PEV in *Drosophila. Advances in Genetics*. **61**: 1-43.

Hsieh, J. J., E. H. Cheng, and Korsmeyer, S.J. 2003. Taspase1: a threonine aspartase required for cleavage of MLL and proper HOX gene expression. *Cell* **115:** 293–303.

Ingham, P.W., & Whittle, R. 1980. *Trithorax*: a new homeotic mutation of *Drosophila melanogaster* causing transformations of abdominal and thoracic imaginal segments. *Mol. Gen. Genet.* **179**: 607–614.

Ingham, P.W., 1998. *trithorax* and the regulation of homeotic gene expression in *Drosophila*: a historical perspective. *Int. J. Devel. Biol*. **42**: 423–429.

Katsani, K.R., Arredondo, J.J., Kal, A.J., and Verrijzer, C.P. 2001. A homeotic mutation in the *trithorax* SET domain impedes histone binding. *Genes Dev*. **15**: 2197–2202.

Kennison, J. A. & Tamkun, J. W. 1988. Dosage-dependent modifiers of Polycomb and Antennapedia mutations in *Drosophila*. *Proc. Natl Acad. Sci. USA* **85**: 8136–8140

Kuzin, B., Tillib, S., Sedkov, Y., Mizrokhi, L., and Mazo, A.,1994. The *Drosophila trithorax* gene encodes a chromosomal protein and directly regulates the region specific homeotic gene forkhead. *Genes Devel*. **8**: 2478–2490.

Lindsley, D. L., and Zimm, G. G. 1992. The Genome of *Drosophila melanogaster*. Academic Press, New York.

Locke,J., Kotarski, M.A., and Tartof, K.D. 1988. Dosage dependent modifiers of position effect variegation in *Drosophila* and a mass action model that explains their effect. *Genetics* **120**:181-198.

Mazo, A.M., Huang, D.H., Mozer, B.A., and Dawid, I.B. 1990. The *trithorax* gene, a trans-acting regulator of the *bithorax* complex in *Drosophila*, encodes a protein with zinc-binding domains. *Proc. Natl. Acad. Sci. USA* **87**:2112–2116.

Muller, H. J., 1930. Types of visible variation induced by X-Rays in *Drosophila. J. Genet*. **22**: 299–334.

Orlando, V. 2003. *Polycomb*, epigenomes, and control of cell identity. *Cell* **112**: 599–606.

Ringrose, L. and Paro, R. 2004. Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu. Rev. Genet.* **38:** 413–443.

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

Rozovskaia, T., Tillib, S., Smith, S., Sedkov, Y., Rozenblatt-Rosen, O., Petruk, S., Yano, T., Nakamura, T., Ben-Simchon, L., Gildea, J., Croce, C.M. Shearn, A., Canaani, E., and Mazo, A. 1999. Trithorax and ASH1 Interact Directly and Associate with the Trithorax Group-Responsive *bxd* Region of the *Ultrabithorax* Promoter. *Mol. Cell. Biol*. **19**: 6441–6447.

Sameny, A. and Locke, J. 2011. The *P*-element-induced silencing effect of *KP* transposons is dose dependent in *Drosophila melanogaster*. *Genome* **54:** 752-762.

Sanchez-Elsner, T., Gou, D., Kremmer, E., and Sauer, F. 2006. Noncoding RNAs of Trithorax Response Elements Recruit *Drosophila* Ash1 to Ultrabithorax. *Science* **311:** 1118-1123.

Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc,B., and Cavalli. G., 2007. Genome Regulation by Polycomb and Trithorax Proteins. *Cell* **128**: 735–745.

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

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

Sedkov, Y., Tillib, S., Mizrokhi, L., Mazo, A., 1994. The bithorax complex is regulated by *trithorax* earlier during *Drosophila* embryogenesis than is the Antennapedia complex, correlating with a bithorax-like expression pattern of distinct early *trithorax* transcripts. *Development* **120**:1907–1917.

Smith, S.T., Petruk, S., Sedkov, Y., Cho, E., Tillib, S., Canaani, E., and Mazo, A., 2004. Modulation of heat shock gene expression by the TAC1 chromatin-modifying complex. *Nat. Cell. Biol.* **6**: 162–167.

Sun, F. L., Cuaycong, M. H., Craig, C. A., Wallrath, L. L., Locke, J., and Elgin, S.C.R. 2000. The fourth chromosome of *Drosophila melanogaster*: interspersed euchromatic and heterochromatic domains. *Proc.Natl. Acad. Sci. USA* **97**: 5340–5345.

Tillib, S., Petruk, S., Sedkov, Y., Kuzin, A., Fujioka, M., Goto, T., and Mazo. A. 1999. Trithorax- and Polycomb-group response elements within an *Ultrabithorax* transcription maintenance unit consist of closely situated but separable sequences. *Mol. Cell. Biol.* **19**:5189–5202.

Tolhuis, B., de Wit, E., Muijrers, I., Teunissen, H., Talhout, W., van Steensel, B., and van Lohuizen, M. 2006. Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in *Drosophila melanogaster*. *Nat. Genet*. **38**: 694–699.

Wallrath, L.L., and Elgin, S.C.R., 1995. Position effect variegation in Drosophila is associated with an altered chromatin structure. *Genes & Development* **9**:1263-1277.

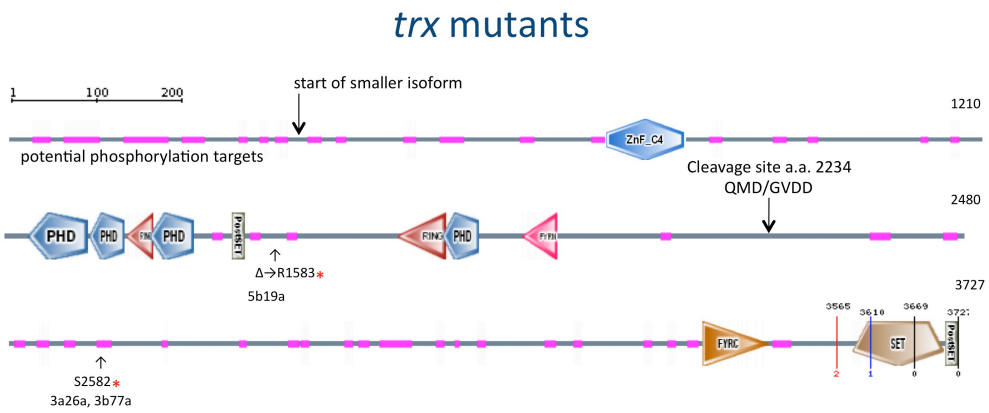
Tables and Figures

Table 3.1 - En (PDS) *trx* mutants

a)

Mutant	Mutagen	Mutation	Amino Acid Change	Designation
3a26a	EMS	C8588T	S2582Amber	$trx^{S2582a*}$
3b77a	EMS	C8588T	S2582Amber	$trx^{S2582b*}$
5b19a	EMS	Δ 5578-90	1578QQQQR to 1578HGMLTAmber	trx^{R1583*}

b)

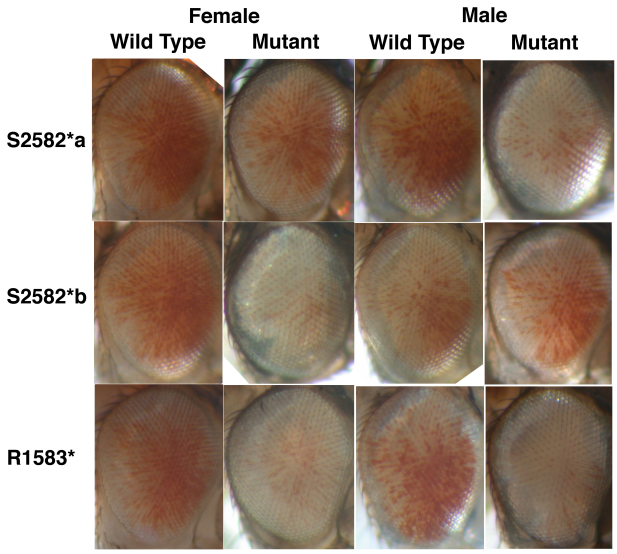


Domain	Start	End	Expected	Domain	Start	End	Expected
ZnF_C4	772	860	1e-12	PHD	1268	1345	2.76e-07
PHD	1346	1391	2.37e-06	RING	1347	1390	2.49e+00
PHD	1423	1480	1.46e-06	PostSET	1528	1544	2.06e+01
RING	1737	1800	1.67e+00	PHD	1796	1842	2.72e-06
FYRN	1897	1942	4.56e-14	FYRC	3388	3476	4.14e-35
SET	3588	3710	2.36e-42	PostSET	3710	3726	7.16e-05

Table 3.1 a) Table showing mutagen used, coding sequence change, amino acid alteration, and new designation of the mutant. (b) Schematic representation of TRX showing important domains and the location of mutants used in this study. Diagram is split into 3 successive sections to fit onto page legibly. Courtesy SMART Heidleberg (http://smart.embl-heidelberg.de/)

Figure 3.1 Effect of *trx* mutants on *E1.*

a)



b)

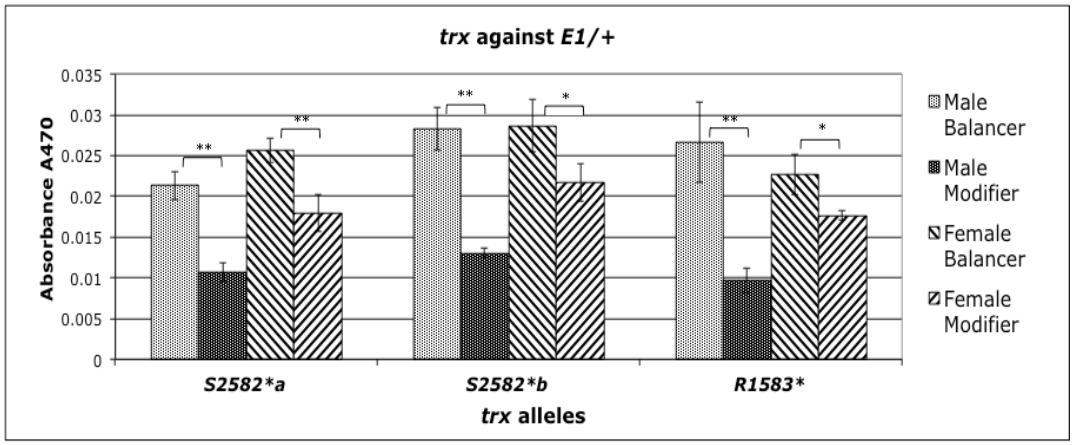
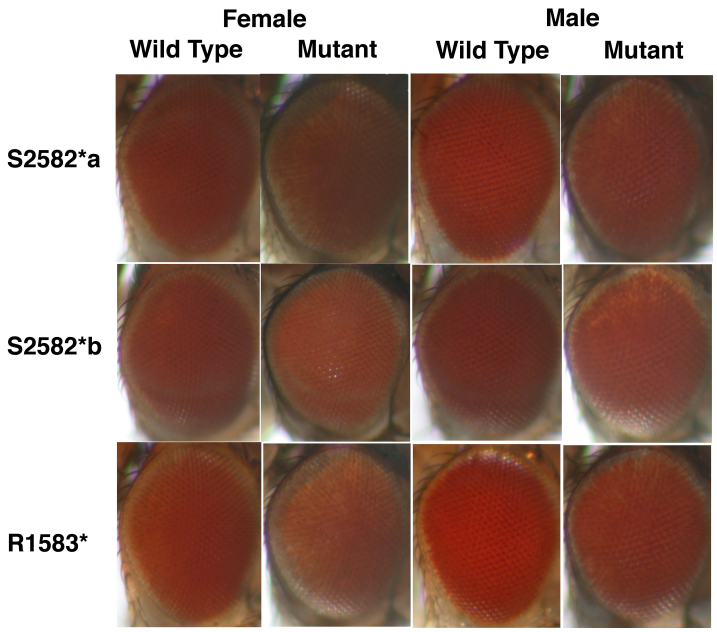


Figure 3.1 a) Photographs of representative examples from each class of progeny from heterozygous *trx* mutants crossed back to the parental *E1* stock used in this mutagenesis. *ebony* versus *ebony*$^{+}$ flies were compared for each sex. Flies are facing right. (y^{-} w^{-}; *trx* e^{-}/+ ♂ X w^{-}; dp^{-}; e^{-}; *E1*♀)

b) Pigment analysis of flies from the same crosses as in a) Statistical significance between pairs (T-Test = 1-tailed, independent (unpaired, unequal variance)) is given above each mutant pair tested, * = $p< 0.05$, ** = $p<0.01$

Figure 3.2 Effect of *trx* mutants on *P{lacW}ci*^{Dplac} (*Pci*).
a)



b)

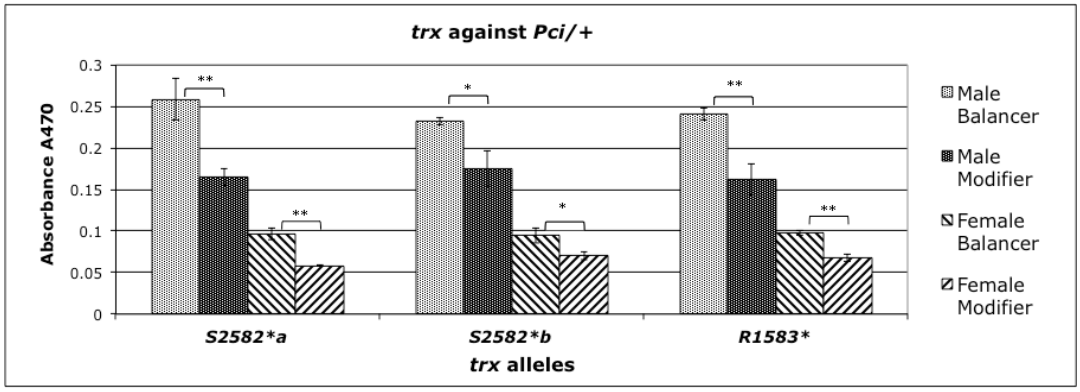
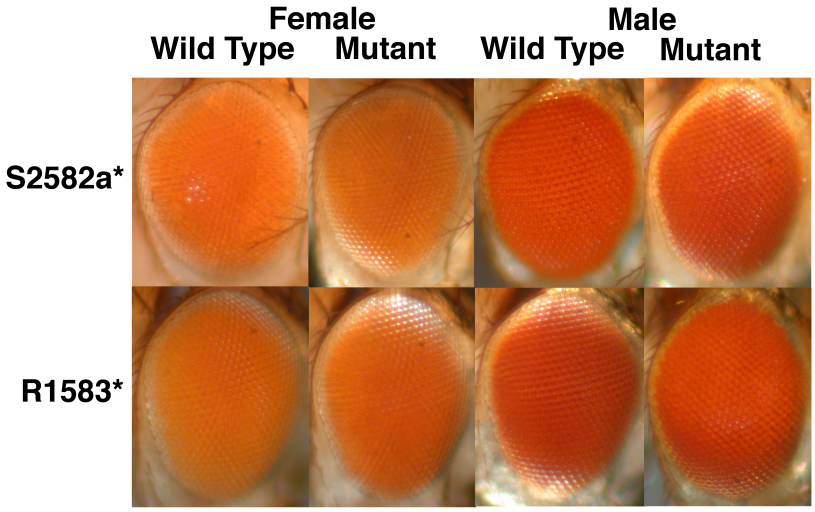
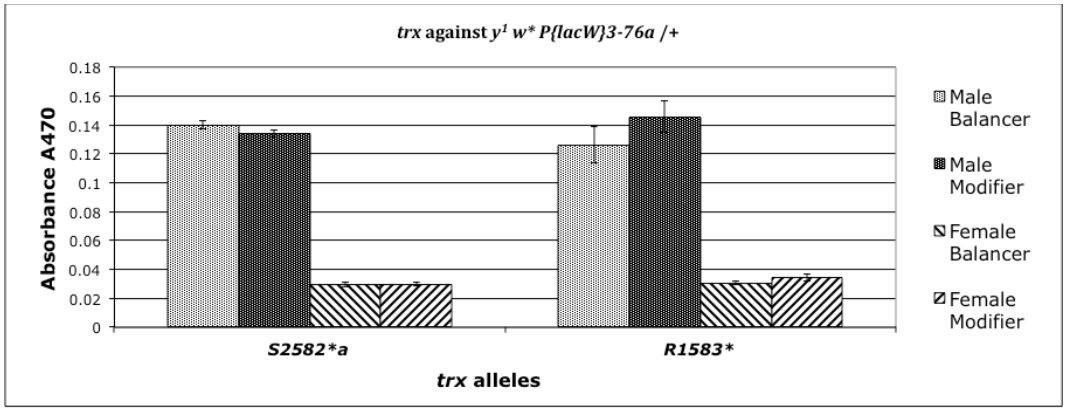


Figure 3.2 a) Photographs of representative examples from each class of progeny from heterozygous *trx* mutants crossed back to the parental *Pci* stock *E1* was derived from. *ebony* versus *ebony*$^{+}$ flies were compared for each sex. Flies are posed facing right. (*y*$^{-}$ *w*$^{-}$; *trx e*$^{-}$/+ ♂ X *w*$^{-}$; *dp*$^{-}$; *e*$^{-}$; *Pci*♀)
b) Pigment analysis of flies from the same crosses as in a) Statistical significance between pairs (T-Test = 1-tailed, independent (unpaired, unequal variance)) is given above each mutant pair tested, * = p< 0.05, ** = p<0.0

Figure 3.3 Effect of *trx* mutants on *P{lacW}3-76a.*
a)



b)



Mutant	Female P-value	Male P-value
*S2582*a*	0.5000	0.0275
*R1583**	0.0463	0.0591

Figure 3.3 a) Photographs showing *trx* mutants do not effect *P{lacW}3-76a.* Photographs of representative examples from each class of progeny from heterozygous *trx* mutants crossed to *P{lacW}3-76a*, an insertion of a transgene identical to *P{lacW}cl*^{Dplac} at 18A1 on the X chromosome. *Sb*$^{+}$ *Ser*$^{+}$ versus *Sb Ser* flies were compared for each sex. (*y*$^{-}$ *w*$^{-}$; *trx*/*TM3 Sb Ser e*$^{-}$ ♂ X *y*^{1} *w* P{lacW}3-76a* ♀)

b) Pigment analysis of flies from the same crosses as in a) Statistical significance between pairs (T-Test = 1-tailed, independent (unpaired, unequal variance)) is given below the chart.

Figure 3.4 Effect of *trx* mutants on w^{m4}.

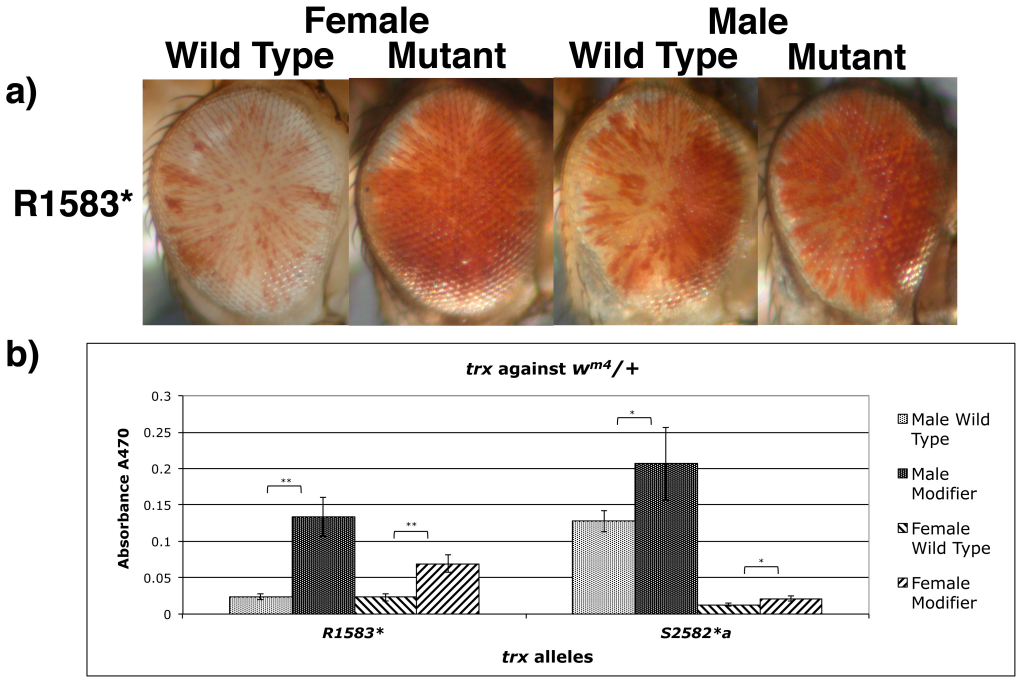
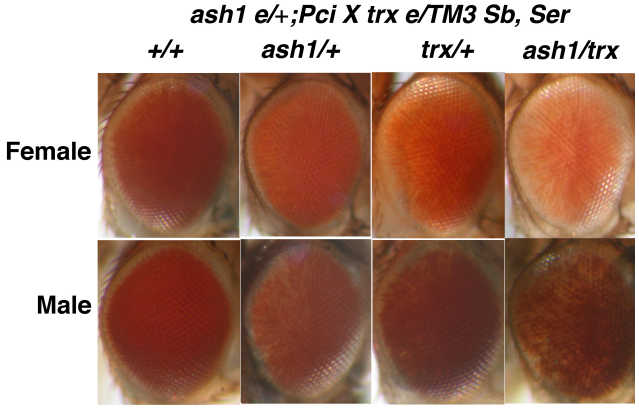


Figure 3.4 a) Photographs showing *trx* mutants' effect on w^{m4}. Representative examples from each class of progeny from heterozygous *trx* mutants crossed to w^{m4}. *ebony* versus *ebony*$^{+}$ flies were compared for each sex. Flies are posed facing right. (y^{-} *w*; *trx* e^{-}/+ ♂ X w^{m4}; e^{-}; ♀)
b) Pigment analysis of *trx* mutants' effect on w^{m4}.
Heterozygous *trx* mutants were crossed to w^{m4}, *ebony* versus *ebony*$^{+}$ flies were compared for each sex. (y^{-} *w*; *trx* e^{-}/+ ♂ X w^{m4}; e^{-}; ♀) Statistical significance between pairs (T-Test = 1-tailed, independent (unpaired, unequal variance)) is given above each mutant pair tested, * = $p< 0.05$, ** = $p<0.01$, NS= not significant.)

Figure 3.5 Synergistic effect of *ash1* and *trx* at *Pci.*

a)



b)

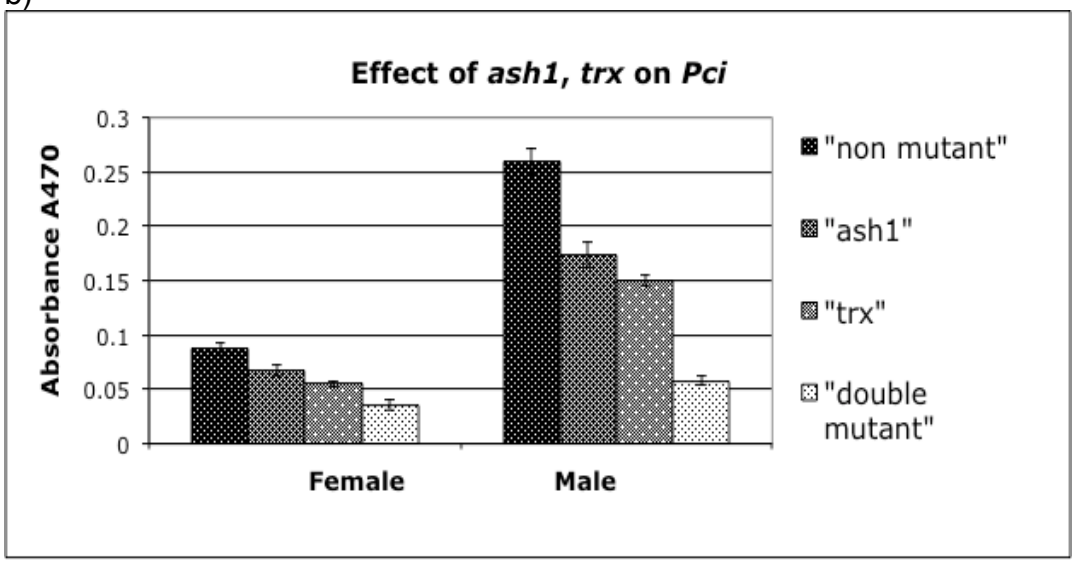


Figure 3.5 a) Photographs of representative examples from each class of progeny from the following cross (y^{-} w^{-}; $ash1^{N303*}$ e^{-}/+ ; Pci ♂ X y^{-} w^{-}; $trx^{S2582a*}$ e^{-}/*TM3 Sb Ser* e^{-}♀). Flies are posed facing right.

b) Pigment analysis of flies from the same cross as in a). Statistical significance (T-Test = 1-tailed, independent (unpaired, unequal variance)) between both mutants and balancers $p<0.01$, between double mutants and single mutants $p<0.01$, between $ash1^{N303*}$, trx^{S2582*} not significant.

Chapter 4 – Manuscript for *CG8878* research

This chapter is a version of a manuscript in preparation for submission for publication.

Mutations in *CG8878* enhance
P element dependent silencing (PDS)
and
Position Effect Variegation (PEV)
in
***Drosophila melanogaster*.**

Allen McCracken and John Locke*

Department of Biological Sciences,
University of Alberta,
Edmonton, Alberta, Canada, T6G 2E9

*Corresponding Author:
John Locke
Department of Biological Sciences
CW 405, Biological Sciences Bldg
University of Alberta
Edmonton, Alberta
Canada, T6G 2E9
Email: john.locke@ualberta.ca
Phone: 780-492-2193
Fax: 780-492-9234

Abstract

In *Drosophila melanogaster*, the mini-*white*$^{+}$ transgene in *P{lacW}ci*^{Dplac} is expressed throughout the adult eye; however, when other *P* or *KP* elements are present, a variegated eye phenotype results indicating random *w*$^{+}$ silencing during development. This *P* element dependent silencing (PDS) can be modified by the haplo-suppressors/triplo-enhancers, *Su(var)205* and *Su(var)3-7*, indicating that these heterochromatic modifiers also act dose dependently in PDS. Here we use a spontaneous derivative mutation of *P{lacW}ci*^{Dplac} called *P{lacW}ci*^{DplacE1} (*E1*) that variegates in the absence of *P* elements, presumably due to an adjacent *gypsy* element insertion, to screen for second-site mutations that enhance variable silencing of *white*$^{+}$ in *E1*. Amongst the modifiers, we isolated 7 mutations in *CG8878*, an essential gene, that enhance the *E1* variegated phenotype. *CG8878* potentially encodes a serine/threonine kinase whose closest Drosophila paralogue, *ballchen* (*nhk-1*), phosphorylates histones. These mutant alleles enhance both PDS at *E1* and Position Effect Variegation (PEV) at *w*^{m4}, indicating a common silencing mechanism between the two.

Introduction

In *Drosophila melanogaster,* expression of the *white*$^{+}$ gene (*w*$^{+}$) is cell autonomous, and necessary for the import of pigment precursors for normal colour in the adult eye. In *white*$^{-}$ mutants, the absence of pigment or white eye phenotype, can be rescued with *w*$^{+}$ containing *P*-element transgenes. However, in some insertion locations expression of *w*$^{+}$ is sensitive to the local chromatin environment, such as adjacent heterochromatin. For example, *P{lacW}ci*^{Dplac} (*Pci*), a transgene inserted proximally on chromosome 4 between *Ribosomal protein S3A* (*RpS3A*) and *cubitus interruptus (ci),* was originally isolated as an enhancer trap of *ci* (Eaton & Kornberg 1990). The *w*$^{+}$ minigene it contains is sensitive to changes in levels of heterochromatin proteins HP1 and SU(VAR)3-7 (Bushey & Locke 2004).

The *P{lacW}ci*^{Dplac} transgene is also sensitive to the presence of P elements in the genome in a phenomenon called *P* element dependent silencing (PDS), which is similar to heterochromatic position effect variegation (hPEV). In flies lacking *P* elements (M strains) the *w*$^{+}$ transgene is expressed in a uniform manner (even red eye phenotype). However, in flies containing *P* elements (P strains) or *KP* s (derivative elements capable of mimicking some of the characteristics of P strains such as modifying P- repressor sensitive alleles, but not enabling *P* element transposition) variegation occurs resulting in a mosaic expression of white ommatidia on a red background in the eye (Sameny & Locke 2011). PDS also occurs when other *w*$^{+}$ transgenes are inserted near this location (Bushey & Locke 2004).

The random silencing of the *w*$^{+}$ minigene in *P{lacW}ci*^{Dplac} during development indicates a phenomenon similar to heterochromatin spreading in hPEV. This is supported by *P{lacW}ci*^{Dplac} responding in a dose sensitive manner to *Su(var) 205* and *3-7,* like *w*^{m4} and other

centromeric hPEV lines (Sun *et al*. 2000, Wallrath & Elgin 1995). As with PEV, position is important as insertions of the same transgene in other locations do not display PDS. Also, chromosomal translocations of *P{lacW}ci*^{Dplac} away from its centromere proximal location suppresses PDS (Bushey & Locke 2004). The position dependence and variegated phenotype suggest silencing may occur from a similar mechanism as heterochromatic PEV.

During investigation of PDS at *P{lacW}ci*^{Dplac} two spontaneous mutants, *P{lacW}ci*^{DplacE1} and *P{lacW}ci*^{DplacE2} (hereafter *E1* and *E2*), were recovered that showed a variegated eye phenotype in the absence of other *P* elements and a complete white eye phenotype when combined with *P* elements, such as *P{ry+ SalI}89D.* hPEV modifier loci, such as *Su(var)205* and *Su(var)3-7,* suppressed variegation at *E1* and *E2*. Molecular analysis of *E1* and *E2* revealed novel *gypsy* element insertions approximately 1 kb distal from the *P{lacW}ci*^{Dplac} insert and in opposite orientations 547 bp apart. Testing of *E1* and *E2* against mutants of *su(Hw)* and *mod(mdg4)* showed that variegation by *E1* and *E2* was not as a result of the *gypsy* insulator function *per se* (Bushey & Locke 2004); however, this does not rule out a possible interaction between the *gypsy* insulator and the *wari* element at the 3' end of the w^+ transgene (Chetverina *et al.* 2008). Both *E1* and *E2 trans*-silence w^+ expression of *P{lacW}ci*^{Dplac} on a paired homolog, but not when present on translocations. Therefore position or pairing contributes to trans-silencing similar to the dominant *trans*-inactivation of the wild-type homologue in bw^+/bw^D heterozygotes (Dernburg *et al*. 1996; Csink & Henikoff 1996; Belyaeva *et al*. 199; Bushey & Locke 2004).

Our previously described modifier of PDS screen (Bushey & Locke, 2004) required controlling two chromosomes, however, the *E1* variegating PDS system is less complicated as all components are on the same chromosome. Furthermore, its intermediate variegated eye phenotype can

visibly reveal not only suppressors, as shown for *Su(var)205* and *Su(var)3-7* mutants, but also enhancers which would have a closer to white eye phenotype, as shown in the presence of other *P* elements. We have used *E1* in a genetic screen to induce and recover second site modifiers of the w^+ variegation. Our screen for enhancers of w^- variegated expression provided several loci for investigation, one of which is a novel locus, *CG8878*. We describe here the isolation and characterization of seven *CG8878* mutants, members of one of the complementation groups identified in this screen that enhances both PDS and hPEV at w^{m4}.

Materials and Methods

Drosophila stocks and mutations: Unless otherwise cited, *D. melanogaster* mutations were described previously (Lindsley & Zimm 1992). The *P{lacW}ci*^{Dplac} allele (Eaton & Kornberg 1990) is a *P{lacZ*$^{P\T.W}$*w*$^{+mC}$ *amp*^{R} *ori = lacW}* construct inserted ~3 kbp upstream (distal) from the *ci* locus on chromosome *4*. *P{lacW}ci* ^{DplacE1} has a gypsy element insertion ~ 1kb further upstream as previously described (Bushey & Locke 2004). *y*^{1} *w* P{lacW}3-76a* is a *lacW* transgene inserted on the X-chromosome and was originally isolated by Y. N. Jan and provided by the Bloomington stock center, while *In(1)w*^{m4}*,w*^{m4} was obtained from K.D. Tartof. Fly stocks (Table A1) were maintained at room temperature on standard yeast/cornmeal medium.

Mutagenesis used *w*$^{-}$*; dp*$^{-}$*; e*$^{-}$*; P{lacW}ci* ^{DplacE1} males treated with 25 mM EMS (Ashburner 1989) mated to *y*$^{-}$ *w*$^{-}$*; +/+* virgin females and screened for a dominant enhanced eye colour phenotype in the progeny. Putative mutants were mated to *w*$^{-}$*; dp*$^{-}$*; e*$^{-}$*; P{lacW}ci* ^{Dplac} flies to confirm transmission and segregation and to determine chromosomal location. Mutant *CG8878* alleles were kept as balanced stocks with *CyO*.

Genetic Mapping: The dominant enhancer of *E1*/*Pci* phenotype trait was used for genetic recombination mapping of mutant *4a7a* relative to *wg*^{Sp} *L Bc* and *Pin* markers as it gave a fuller red eye phenotype which provided more room for enhancement and thus allowed a more reliable visual assessment of enhancement. Recombinants both left and right of the enhancer were collected and tested for retention of the enhancer phenotype by crossing males to *w*$^{-}$*; dp*$^{-}$*; e*$^{-}$*; P{lacW}ci* ^{DplacE1} virgin females, and for retention of the recessive lethal phenotype by crossing to other members of the same complementation group. After establishing absolute linkage between the recessive lethal and dominant enhancer phenotypes, the lethal locus position was refined by complementation analysis against

deficiencies in the region. At least 100 progeny were scored and if the heterozygous mutant/deficiency combination did not occur the combination was considered lethal.

DNA sequencing: A series of overlapping CG8878 gene segments were amplified by PCR (Tables A4, A7) and the product was sequenced. Point mutations were identified as double peaks. All polymorphisms and mutations were confirmed by sequencing both strands.

Eye pigment assays: The amount of w^+ gene activity was assayed by measuring the amount of brown eye pigment using a modification of the method of Ephrussi and Herold (1944). Heads from 5-9 day old adult flies were stored at -20° until extracted. For each genotype, three replicate samples of 10 heads were extracted in 200 μL of acidified ethyl alcohol (1% HCl in 30% ethanol) with shaking for 48 hours. Absorbance at 470 nm was then measured using a 96 well Costar flat bottom plate in a Bio-Tek PowerWave XS spectrophotometer. Photographs of representative adult flies eyes were taken under mineral oil using a Zeiss stereo-microscope and a Nikon Coolpix 995 digital camera. For both the eye pigment assay and the adult eye photographs, the balancer chromosome *CyO* was used as the control.

Results

Screen for second site enhancers of w^+ variegation in *E1*

We screened ~44,000 progeny from EMS treated fathers for dominant enhancement of w^+ silencing in *E1* and recovered 58 confirmed mutations. *Inter se* complementation analysis showed they fell into five simple and three more complex recessive lethal complementation groups, as well as many singles. One of the simple groups, with 7 alleles (*1a27a*, *3a22a*, *3a52a*, *3a66a*, *3a90a*, *3a97a*, *4a7a*), is described here. These seven were examined further and the dominant enhancer of *E1*/*Pci* phenotype for allele *4a7a* genetically mapped to 2-65.4 (n=490) by recombination relative to wg^{Sp} *L Bc* and *Pin* markers as described in Materials and Methods. Linkage between the dominant enhancer of *E1*/*Pci* and lethal phenotypes was demonstrated as described above. Deficiency mapping of the lethal phenotype of this group refined its position to 48E2;48E4, within *Df(2R)BSC199* (7,779,605: 8,059,989) but not *Df(2R)BSC879* (7,779,605: 8,029,867). This includes the *CG8878 locus* as the only candidate capable of influencing gene expression in a heritable as opposed to cell autonomous matter, that is expressed in the correct tissue at the correct time, and is of a size likely to result in 7 independent mutations in a mutagenesis of this size. Note that both *Hen1* and *Prp8* can influence gene expression post translationally: however, that should not lead to clonal inheritence (variegation) as seen here.

DNA sequencing of the mutants

DNA sequencing spanning the entire predicted coding region of our mutants, in heterozygotes with the *CyO* balancer chromosome, showed that five alleles (*1a27a*, *3a22a*, *3a52a*, *3a66a*, *3a97a*) had a base pair change within *CG8878* that altered the predicted amino acid coding sequence (Table 4.1, Figure 4.1). Three of the alleles (*3a27a*, *3a52a* and *3a97a*) had G/C to A/T transitions that resulted in premature stop codons; with *3a52a* being at the amino terminal end of the first predicted STKc

domain and therefore most likely to be a null allele. Allele *3a66a* had a single nucleotide deletion that caused a frame-shift and multiple premature stop codons while *1a27a* had a G/C to A/T transition that caused the loss of an intron donor splice site, a frame-shift and multiple premature stop codons. Two other alleles (*3a90a* and *4a7a*) had the same nineteen base pair deletion in the 5' promoter region that included 4 base pairs of the proximal predicted E box and are thus presumptive transcriptional regulatory mutants.

Phenotypic characterization of the mutants

Visual pigment assessment for the dominant enhancement of white eyed colour variegation in *E1/+* heterozygotes indicated all mutant alleles were enhanced relative to the *CyO* control in both sexes, and frequently produced flies indistinguishable from w^-. Representative photographs of mutant eyes are given in Figure 4.2 a. The extent of enhancement was quantified by pigment assays of our *CG8878* mutants. Both male and female *CG8878* mutants had less than half the pigment of non-mutant internal control flies (*CyO* balancer) from the same cross, with all being significant (95% confidence limit - Figure 4.2 b). All mutant *CG8878* alleles enhance silencing of *E1/+*.

To see if the dominant enhancement (w^+ silencing) was limited to the *E1* allele, we quantitatively assayed the effect of *CG8878* mutants on *Pci*/+ flies, which lack the gypsy element present in *E1*. All *CG8878* mutants reduced the amount of pigment compared to non-mutant internal control flies (*CyO* balancer) from the same cross (Figure 4.3 b). However, only the putative null allele *3a52a* showed significance in both sexes. Visually, variegation was less visible in this cross, with eyes displaying only a pattern of weak silencing starting at the posterior edge, with rapidly decaying anterior progression (Figure 4.3 a).

To address the possibility that loss of the *CG8878* product was directly silencing the w^+ transgene within the *P{lacW}ci*^{Dplac} insert, we assayed pigment in *y*^{1} *w* P{lacW}3-76a* /+ flies with *CG8878* mutants. The transgene in *P{lacW}3-76a* is identical to that in *P{lacW}ci*^{Dplac} but is located at 18A1 (60.7 cM) on the X chromosome and presents a full red eye. Of the three alleles tested, only our *CG8878* putative regulatory mutant males showed a barely significant difference (p=.03) from non-mutant control flies (*CyO* balancer) of the same cross indicating that *CG8878*'s dominant enhancement (silencing) is likely to be dependant upon chromosomal location and not on the w^+ transgene *P{lacW}ci*^{Dplac} construct itself (Figure 4.4). The consistency of our results across all mutants tested indicates that this method of pigment determination is both accurate and precise.

Next, we asked whether *CG8878* mutants had an effect on classical hPEV by by crossing *y- w-*; *dp- CG8878*/CyO, Cy dp-* males to virgin *In(1)w*^{m4}; *dp-; e-* females. Variegation of w^{m4} was visibly enhanced by all three mutants tested (Figure 4.5 a) and quantitatively (95% and 99% confidence limits) enhanced in male and female flies respectively (Figure 4.5 b).

Amino acid sequence comparisons

A comparison of *CG8878*'s predicted amino acid sequence with 12 Drosophila species reveals that homologs are present and highly conserved in all species studied; this supports *CG8878* being an essential, expressed gene (Figure 4.6, Supplemental figure 4.1). This tree parallels one already found for these species (Hahn *et al*. 2007).

CG8878's closest *D. melanogaster* paralogue is *ballchen* (*nhk1*), which encodes a nucleosomal histone 2a kinase that is related to vertebrate Vaccinia Related Kinases (VRK1, VRK2) (Figure 4.7, Supplemental Figure 4.2). Amino acid sequence comparisons suggest both *CG8878* and

ballchen are derived from a common *VRK* like precursor. However, in CG8878, the VRK domain appears to have been split in two by an ~282 amino acid insertion (Figures 4.1, 4.8). CG8878 shows 36% identity and 56% positive correlation to BALLCHEN over both parts of the kinase domain (http://blast.ncbi.nlm.nih.gov/Blast.cgi) indicating a possible functional conservation. Since our mutation *3a22a* (R546Opal) is recessive lethal and enhances variegation at *E1*, *Pci,* and w^{m4} (Figures 4.2 b, 4.3 b, 4.5 b), it appears the second part of CG8878's split VRK-like domain (Fig 4.8) is essential for CG8878's function.

Discussion

We induced, recovered, and characterized seven mutations that dominantly enhance the variable silencing (variegation) of *E1*, whose expression is similar to *P* element dependent silencing (PDS) and showed that they are all alleles of *CG8878.* The dominant enhancement genetically maps at or near the *CG8878* locus and it could not be separated from the lethal phenotype by crossing over. The lethal phenotype deficiency maps to a very fine region that uncovers *CG8878* and no other reasonable candidate gene. Five alleles contain mutations resulting in stop codons; two at the amino terminal end of CG8878's amino proximal predicted STKc domain likely represent null alleles, one between CG8878's two predicted kinase domains, and two in the amino end of *CG8878*'s carboxy proximal predicted kinase domain. Taken together, this shows that loss of the *CG8878* gene function is responsible for the enhanced silencing of w^+ in *E1*.

CG8878 and *Hen1*

The *CG8878* transcription unit is located entirely within the large (5.4 kb) second intron of another gene, *Hen1* (formerly Pimet), in the antisense orientation (Supplemental Figure 4.3). *Hen1* has been shown to mediate 2'-*O*-methylation at the 3' end of *Piwi* interacting RNAs in *Drosophila* (Saito *et al*. 2007; Horwich *et al*. 2007). *Piwi* interacting RNAs are germ-line specific 24-30 nt RNAs that couple with PIWI proteins to silence invading transposable elements (reviewed by Saito *et al*. 2006). Given that *P{lacW}ci*^{Dplac} has P element terminal repeats and, at the 5' end, a P-element transposase *lacZ* fusion, could *Hen1* and not *CG8878* actually be the enhancer isolated in this screen? Several points argue against this: 1) Hen1 is not an essential gene because *PBac(WH)Hen1[f00810]* is a null for *Hen1* (Horwich *et al*. 2007) but is not recessive lethal; 2) *P{lacW}3-76a* appears to be unaffected by our *En*(*var*)s despite being the same construct but at a different location; 3) w^{m4}, which is not *P* element derived,

is significantly affected by our *En*(*var*)s; 4) all seven mutants had sequence lesions in *CG8878*; and 5), all of these sequence changes are entirely inside *Hen1*'s second intron, and therefore should have no effect on *Hen1* expression. The most parsimonious explanation is that *CG8878* is an essential gene and when mutated has a dominant *En*(*var*) phenotype.

Potential molecular function of *CG8878*?

The closest *Drosophila melanogaster* homologue of *CG8878* is *nucleosomal histone kinase-1* (*nhk-1* or *ballchen*) with 41% identity (E= 3e-29), with regions of maximum similarity coinciding with *CG8878*'s putative kinase domains as shown in Figure 4.1 b). NHK-1 has high affinity for chromatin and has been shown to phosphorylate Threonine 119 at the carboxy terminus of nucleosomal, but not free, H2A in *Drosophila* embryos. H2A T119 is phosphorylated during mitosis but not in S phase which coincides with NHK-1's chromatin association as shown by immunostaining and may be a component of the histone code related to cell cycle progression (Aihara *et al*. 2004). Ivanovska *et al*. (2005) described a point mutation, Z3-0437, in the kinase domain of NHK-1 that led to female sterility due to defects in the formation of the karyosome. This led to metaphase I arrest as a result of failure of the synaptonemal complex to disassemble and to load condensin onto chromosomes in the mutant. Mitosis was also shown to be affected, as embryos laid by nhk-$1^{-/-}$ mutant females arrested with aberrant mitotic spindles and polar bodies. They also found a lack of Histone H4K5 and H3K14 acetylation in the karyosomes in *nhk-1* mutant but not control oocytes, implying that Histone H2A threonine 119 phosphorylation is required for meiotic acetylation of these residues. Lancaster *et al*. (2007) found that phosphorylation of *barrier to autointegration factor* protein (BAF) by NHK-1 was necessary for karyosome formation. Loss of NHK-1 or expression of nonphosphorylatable BAF resulted in ectopic chromosome-nuclear

envelope association in oocytes leading the authors to propose that tethering of chromosomes to the nuclear envelope is disrupted by NHK-1 mediated BAF phosphorylation, allowing karyosome formation in oocytes. These findings are intriguing as possible histone phosphorylation by CG8878 would readily explain its action as an *En*(*var*). For example, JIL1 phosphorylation of H3S10 blocks methylation of H3K9 allowing hyperacetylation of Histone 3 and promoting a transcriptionally active chromatin state (Zhang *et al*. 2006). CG8878's expression profile is consistent with it being a genome wide inhibitor of heterochromatin spread as it is expressed in all tissues, at all stages of development, with maxima at times of peak developmental change such as early embryogenesis and prepupariation (Gelbart & Emmert 2011). It remains to be seen what *CG8878*'s target, and mode of action are, but the possibilities are intriguing. The recessive lethal phenotype of *3a66a,* which results in a premature stop codon between CG8878's two predicted kinase domains, and *3a22a*, and *3a97a*, which result in a premature stop codon in the amino end of CG8878's carboxy proximal predicted kinase domain, combined with the enhancer of *E1* and w^{m4} phenotypes of *3a22a* shown in Figures 4.2 b, and 4.5 b, argue that this latter predicted kinase domain is essential for CG8878 function.

CG8878 acts at the *ci* locus

Pci was isolated as an enhancer trap of *ci* and is an allele of *ci.* While *Pci* (and the *E1* gypsy element) are inserted in the *ci* distal regulatory region, both ci^{57g}, a deletion upstream of *Pci*, and ci^{1}, a gypsy insert upstream of *Pci,* exhibit *ci* phenotypes. Expression of the *Pci* enhancer-trap reporter in imaginal discs accurately mimicked that of ci^{D} RNA with both being expressed specifically in anterior compartment cells (Eaton & Kornberg 1990). All alleles isolated of *CG8878* are *En(PDS)* with *E1*, *E1/Pci* (personal observation) and appear to act at *Pci* to lower *w* expression without *E1*. The lack of CG8878 function of the presumptive

null allele, *3a52a*, results in statistically significant increased silencing of the *white* reporter gene which should be mimicking *ci* expression. Since these mutants have little effect on *P{lacW}3-76a* , this effect is not construct dependent and is not due to a direct interaction with the *white* promoter, but with the *ci* regulatory region itself. Since *Pci* reporter expression is approximately halved when *3a52a* is present, and does not depend on the presence of *E1*, we infer that CG8878 normally acts at the *ci* regulatory region to impede the spread of heterochromatin into this region, likely in a dose dependent manner.

Given that our alleles of *CG8878* were isolated in a screen for enhancers of PDS at *ci* that also uncovered *trx*G members *trithorax* and *ash1*, we propose to now call *CG8878 t&a*, for *trithorax* & *ash1* .

References

Aihara, H., Nakagawa,T., Yasui, K., Ohta, T., Hirose, S., Dhomae, N., Takio, K., Kaneko, M., Takeshima, Y., Muramatsu, M., and Ito, T. 2004. Nucleosomal histone kinase-1 phosphorylates H2A Thr 119 during mitosis in the early *Drosophila* embryo. *Genes Dev.* **18**: 877-888

Ashburner, M. 1989. *Drosophila,* A laboratory manual. Cold Spring Harbor Press, Cold Spring Harbor, New York.

Belyaeva,E.S., Koryakov,D,E., Pokholkova,G.V., Demakova,O.V., Zhimulev, I.F., 1997. Cytological study of the *brown* dominant position effect. *Chromosoma* **106**: 124-132.

Bushey, D. and Locke, J. 2004. Mutations in *Su(var)205* and *Su(var)3-7* Suppress P-Element-Dependent Silencing in *Drosophila melanogaster*. Genetics, **168**: 1395–1411.

Chetverina, D., Savitskaya, E., Maksimenko, O., Melnikova, L., Zaytseva, O., Parshikov, A., Galkin, A.V., and Georgiev, P. 2008. Red flag on the white reporter: a versatile insulator abuts the white gene in *Drosophila* and is omnipresent in mini-white constructs *Nucleic Acids Research* **36**:929–937.

Csink, A. K., and Henikoff,S. 1996. Genetic modification of heterochromatic association and nuclear organization in *Drosophila*. *Nature* **381:** 529–531.

Dernburg, A.F., Broman, K.W., Fung, J.C., Marshall, W.F., Philips, J., Agard D.A., and Sedat, J.W. 1996. Perturbation of Nuclear Architecture by Long-Distance Chromosome Interactions *Cell* **85**: 745–759.

Eaton, S., and Kornberg, T.B. 1990. Repression of *ci-D* in posterior compartments of *Drosophila* by *engrailed*. *Genes Dev*. **4:** 1068–1077.

Ephrussi, B., and Herold, J. L. 1944. Studies of eye pigments of *Drosophila*. I. Methods of extraction and quantitative estimation of the pigment components. *Genetics* **29:** 148–175.

Gelbart, W.M., Emmert, D.B. (2011.8.23). Calculation of RPKM to generate quantitative expression data: read-length values for modENCODE developmental timecourse RNA-Seq data. FlyBase analysis.

Hahn M.W., Han M.V., Han S-G. 2007. Gene Family Evolution across 12 *Drosophila* Genomes. PLoS Genet 3(11): e197.

Horwich, M.D., Li, C., Matranga, C., Vagin, V., Farley, G., Wang, P., and Zamore, P.D. 2007. The *Drosophila* RNA Methyltransferase, DmHen1, Modifies Germline piRNAs and Single-Stranded siRNAs in RISC *Current Biology* **17**: 1265–1272.

Ivanovska, I., Khandan, T., Ito, T., and Orr-Weaver, T.L. 2005. A histone code in meiosis: the histone kinase, NHK-1, is required for proper chromosomal architecture in *Drosophila* oocytes. *Genes Dev.* **19**: 2571-2582

Lancaster, O.M., Cullen, C.F., Ohkura, H. 2007. NHK-1 phosphorylates BAF to allow karyosome formation in the *Drosophila* oocyte nucleus. *J. Cell Biol*. **179**: 817--824.

Lindsley, D. L., and Zimm, G.G. 1992. *The Genome of Drosophila melanogaster*. Academic Press, New York.

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

Saito, K., Sakaguchi, Y., Suzuki, T., Siomi, H., and Siomi, M.C. 2007. *Pimet*, the *Drosophila* homolog of HEN1, mediates 2'-*O*-methylation of Piwi- interacting RNAs at their 3' ends. *Genes Dev.* **21**: 1603-1608.

Sameny A, Locke J. 2011. The *P*-element-induced silencing effect of *KP* transposons is dose dependent in *Drosophila melanogaster*. *Genome* *54*:752-62.

Sun, F. L., Cuaycong, M. H., Craig, C. A., Wallrath, L. L., J. Locke and Elgin, S.C.R. 2000. The fourth chromosome of *Drosophila melanogaster*: interspersed euchromatic and heterochromatic domains. *Proc.Natl. Acad. Sci. USA* **97:** 5340–5345.

Wallrath, L.L., and Elgin, S.C.R. 1995. Position effect variegation in *Drosophila* is associated with an altered chromatin structure. *Genes & Development* **9**: 1263-1277

Zhang, W., Deng, H., Bao, X., Lerach, S., Girton, J., Johansen J., and Johansen, K.M. 2006. The JIL-1 histone H3S10 kinase regulates dimethyl H3K9 modifications and heterochromatic spreading in *Drosophila. Development* **133**: 229-235

Tables and Figures

Table 4.1 – List of EMS induced mutations in CG8878 and their predicted changes to the amino acid sequence.

Mutant	Mutation	Type	Effective Change
1a27a	G⟶ A 8037095	Loss of intron donor splice site, frameshift.	R180Opal
3a22a	C⟶ T m1942	Point, transition.	R546Opal
3a52a	G⟶ A m675	Point, transition.	W123Opal
3a66a	1Δ C m1665	Frameshift.	I468Amber
3a90a	19 bp Δ T 8,038,801-19	Frameshift.	4 bp Δ E box
3a97a	C⟶ T m1942	Point, transition.	R546Opal
4a7a	19 bp Δ T 8,038,801-19	Frameshift.	4 bp Δ E box

Table 4.1 Table showing mutagen used, coding sequence change, type of mutation, and resulting effective amino acid alteration of the mutant.

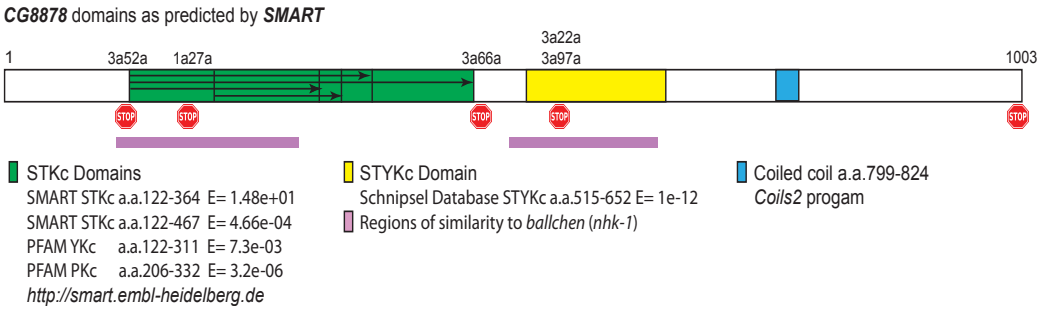


Figure 4.1 Schematic representation of *CG8878* polypeptide sequence. Domains predicted by SMART (University of Heidelberg) and the location of lesions described in this study are shown. Mutant designations are above the polypeptide backbone while the nature of the corresponding mutation is below. Regions of sequence similarity to *ballchen* (*nhk-1*) are shown as mauve bars below the CG8878 sequence.

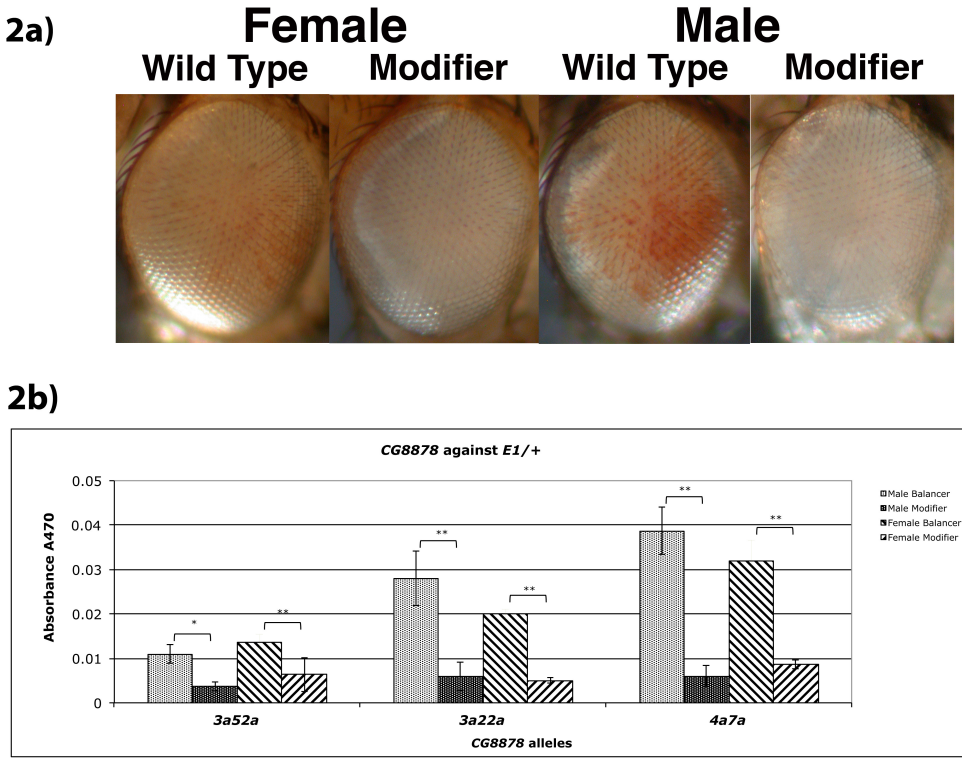


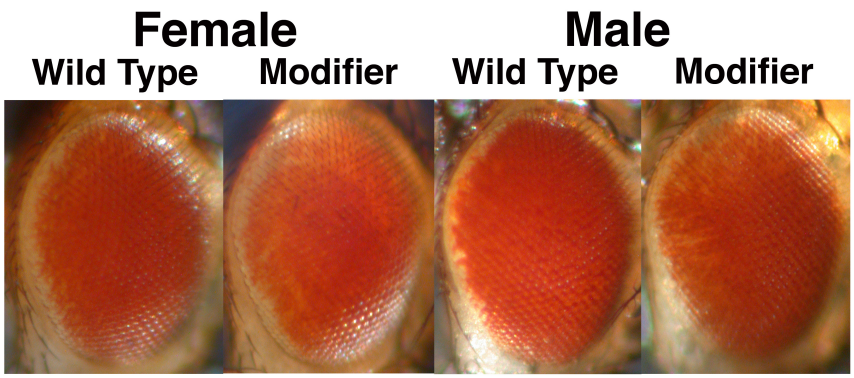
Figure 4.2 – Enhancement of *E1* by CG8878 mutations.

a) Representative photographs of eyes from each class of progeny from heterozygous *CG88781* mutants crossed back to the parental *E1* stock used in this mutagenesis. *Cy* versus Cy^+ flies were compared for each sex. Flies are posed facing right.

($y^- w^-$; dp^- *3a52a* /*CyO, Cy* dp^- ♂ X w^-; dp^-; e^-; *E1*♀)

b) Pigment assays of heterozygous *CG8878* mutants crossed back to the parental *E1* stock used in this mutagenesis. *Cy* versus Cy^+ flies were compared for each sex. ($y^- w^-$; dp^- *CG8878*/*CyO, Cy* dp^- ♂ X w^-; dp^-; e^-; *E1*♀) Statistical significance between pairs (T-Test = 1-tailed, independent (unpaired, unequal variance)) is given above each mutant pair tested, * = $p< 0.05$, ** = $p<0.01$, NS= not significant.

3a)



3b)

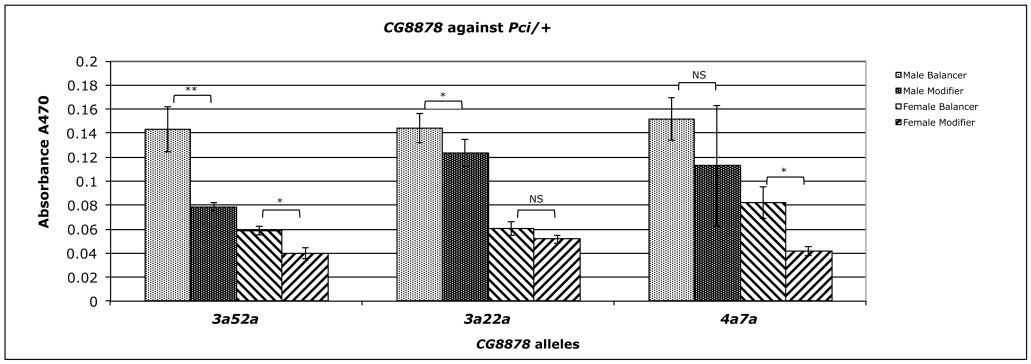


Figure 4.3 – Enhancement of *Pci* by *CG8878* mutations.

a) Photographs of representative examples from each class of progeny from heterozygous *CG8878* mutants crossed back to *Pci*, the parental stock from which *E1* was derived. *Cy* versus Cy^+ flies were compared for each sex. Flies are posed facing right.

(y^- w^-; dp^- *3a52a* /*CyO, Cy* dp^- ♂ X w^-; dp^-; e^-; *Pci*♀)

b) Pigment assays of heterozygous *CG8878* mutants crossed back to the parental *Pci* stock *E1* was derived from. *Cy* versus Cy^+ flies were compared for each sex. (y^- w^-; dp^- *CG8878*/*CyO, Cy* dp^- ♂ X w^-; dp^-; e^-; *Pci*♀) Statistical significance between pairs (T-Test = 1-tailed, independent (unpaired, unequal variance)) is given above each mutant pair tested, * = $p< 0.05$, ** = $p<0.01$, NS= not significant.

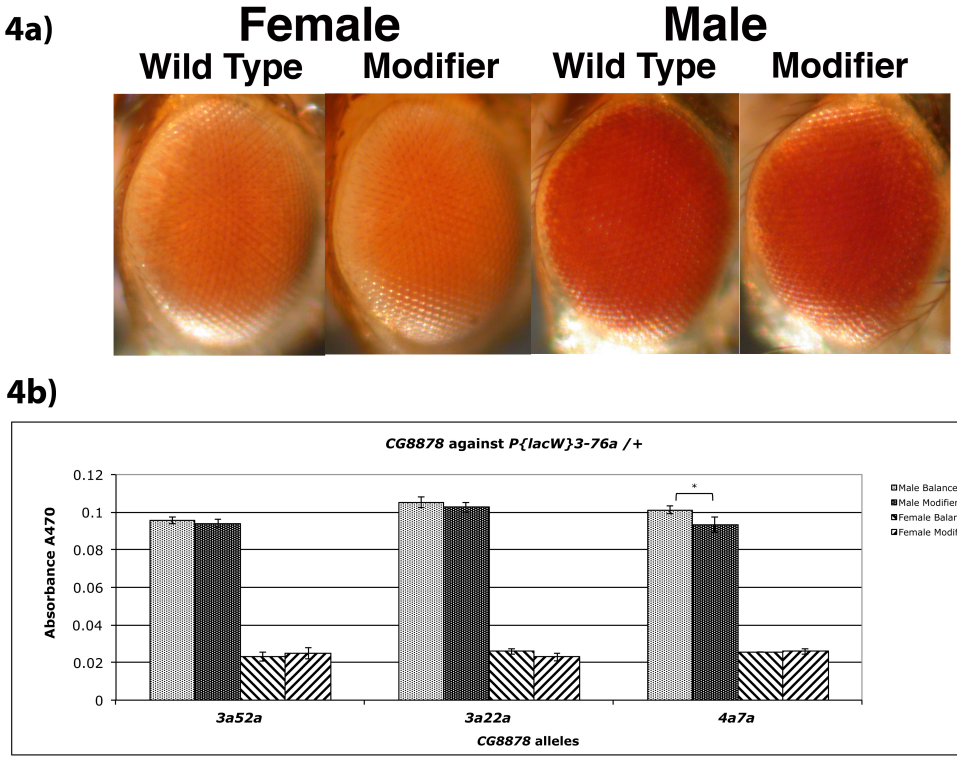
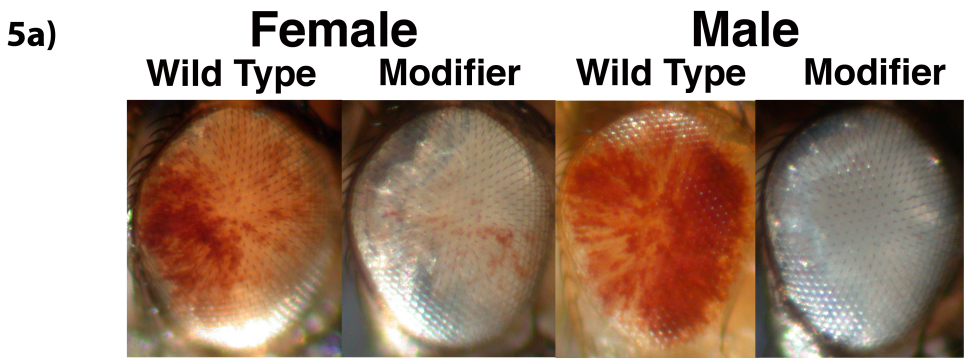


Figure 4.4 – Lack of enhancement of *P{lacW}3-76a* by CG8878 mutations. a) Photographs of representative examples from each class of progeny from heterozygous *CG8878* mutants crossed to *P{lacW}3-76a*, an insertion of the same transgene as *P{lacW}ci*^{Dplac}, but only at 18A1 on the X chromosome. *Cy* versus *Cy*$^{+}$ flies were compared for each sex. Flies are posed facing right. (*y*$^{-}$ *w*$^{-}$; *dp*$^{-}$ *3a52a*/*CyO Cy dp*$^{-}$ ♂ X *y*^{1} *w*$^{-}$ *P{lacW}3-76a* ☿) b) Pigment assays of heterozygous *CG8878* mutants crossed to *P{lacW}3-76a*. *Cy* versus *Cy*$^{+}$ flies were compared for each sex. (*y*$^{-}$ *w*$^{-}$; *dp*$^{-}$ *CG8878*/ *CyO Cy dp*$^{-}$ ♂ X *y*^{1} *w*$^{-}$ *P{lacW}3-76a*☿) Statistical deviation between pairs was insignificant (T-Test = 1-tailed, independent (unpaired, unequal variance)) except for *4a7a* males which was significant at $p < 0.05$ ($* = p < 0.05$).



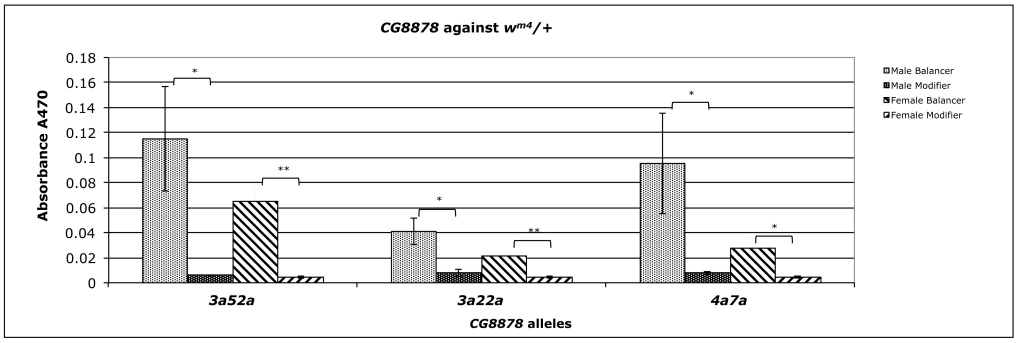


Figure 4.5 – Enhancement of w^{m4} by CG8878 mutations.

a) Photographs of representative examples from each class of progeny from heterozygous *CG8878* mutants crossed to w^{m4}. *Cy* versus Cy^{+} flies were compared for each sex. Flies are posed facing right. (y^{-} w^{-}; dp^{-} *3a52a/ CyO, Cy* dp^{-} ♂ X w^{m4}; dp^{-}; e^{-} ♀)

b) Pigment assays of heterozygous *CG8878* mutants crossed to w^{m4}. *Cy* versus Cy^{+} flies were compared for each sex. (y^{-} w^{-}; dp^{-} *CG8878/ CyO Cy* dp^{-} ♂ X w^{m4}; dp^{-}; e^{-} ♀) Statistical significance between pairs (T-Test = 1-tailed, independent (unpaired, unequal variance)) is given above each mutant pair tested, * = $p< 0.05$, ** = $p<0.01$, NS= not significant.

6 a)

SeqA	Name	Length	SeqB	Name	Length	Score
1	Dmel	1004	2	Dsim	1012	96.0
1	Dmel	1004	3	Dsec	1003	97.0
1	Dmel	1004	4	Dere	1006	92.0
1	Dmel	1004	5	Dyak	1007	93.0
1	Dmel	1004	6	Dana	1030	74.0
1	Dmel	1004	7	Dpse	1035	71.0
1	Dmel	1004	8	Dgri	1069	64.0
1	Dmel	1004	9	Dwil	1025	65.0
1	Dmel	1004	10	Dmoj	1076	64.0
1	Dmel	1004	11	Dvir	1123	63.0
1	Dmel	1004	12	Dper	1035	71.0

6 b)

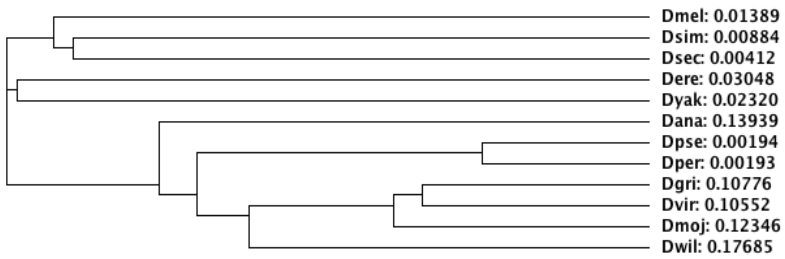


Figure 4.6 - Comparison between *D. melanogaster* CG8878 amino acid sequences and those of 11 other *Drosophila* species. Accession numbers given in Table S 4.1.

a) Score table showing degree of similarity between CG8878 homologues.

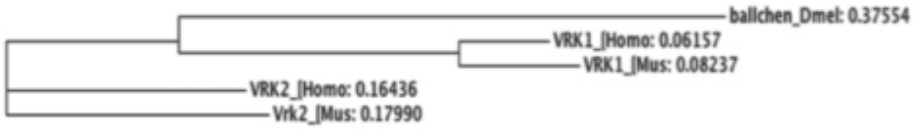
b) Cladogram showing evolutionary distances between CG8878 homologues for various *Drosophila* species (http://www.ebi.ac.uk/Tools/msa/clustalw2/).

Names are abbreviated using the capitalized first letter of the genus followed by the first three letters of the species.

7a)

SeqA	Name	Length	SeqB	Name	Length	Score
1	ballchen_Dmel	599	2	VRK1_[Homo	396	37.0
1	ballchen_Dmel	599	3	VRK1_[Mus	440	34.0
1	ballchen_Dmel	599	4	VRK2_[Homo	427	33.0
1	ballchen_Dmel	599	5	Vrk2_[Mus	454	32.0

7b)



7c)

SeqA	Name	Length	SeqB	Name	Length	Score
1	CG8878_Dmel	1004	2	VRK1_[Homo	396	28.0
1	CG8878_Dmel	1004	3	VRK1_[Mus	440	27.0
1	CG8878_Dmel	1004	4	VRK2_[Homo	427	29.0
1	CG8878_Dmel	1004	5	Vrk2_[Mus	454	25.0

7d)

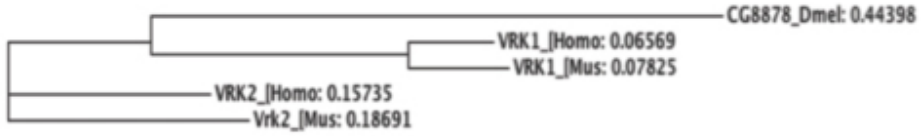


Figure 4.7 Phylogeny of *CG8878* vs *ballchen* and vertebrate VRK genes

a) Score table showing degree of similarity between *ballchen* and mammalian *VRK1,2* .

b) Phylogram showing evolutionary distances between *ballchen* and mammalian *VRK1,2*

c) Score table showing degree of similarity between *CG8878* and mammalian *VRK1,2*

d) Phylogram showing evolutionary distances between *ballchen* and mammalian *VRK1,2* Abbreviations are as follows: *Drosophila melanogaster* (Dmel), *Homo sapiens* (Homo), *Mus musculus* (Mus) (http://www.ebi.ac.uk/Tools/msa/clustalw2/) Accession numbers given in Table S 4.1.

```
CG8878   MGKRLQLERPTTDRSARKRKRSAVKAAEKRQRLSGGSSSANGFEFHENDDEESCSSAGSA 60
ballchen ------------------------------------------------------------

CG8878   AGTEADPPTLLHTPQARSLLLTGASIASDHNNSSVMESPRPVYTLRPSVVNGTILRDVLS 120
ballchen ----------------MPRVAKPKAAAPAKKVVSAKKAKSKLYKMPEKVKEGTVFTDLAK 44
                         . : .  : *. ::  *. ::   :*.:  .* :**:: *: .

CG8878   KAWRLGRPIGKGNFGEIFLASDDTVCPASSETAKYVVKIEPHSNGPLFVEIHCLINTSRN 180
ballchen GQWRIGPSIGVGGFGEIYAA-----CKVGEKNYDAVVKCEPHGNGPLFVEMHFYLRNAKL 99
           **:* .** *.****: *     * ...:. . *** ***.*******:*  :..::

CG8878   NDLSDAAEDAASLPAPQTHVLSRGPPSGIPSFIASGTHYFGDVRYRFLVLPRFDRDLHSL 240
ballchen EDIKQFMQKHG------------LKSLGMPYILANGSVEVNGEKHRFIVMPRYGSDLTKF 147
         :*:.:  :. .              . *:* ::*.*:  ... ::**:*:**:. ** .:

CG8878   IKNS--RVQQKSLLVLAVHIINVLENLHDKGYCHNDIKAQNLMVSKCKYLRRQVVPKGNG 298
ballchen LEQNGKRLPEGTVYRLAIQMLDVYQYMHSNGYVHADLKAANILLG---------LEKGG- 197
         :::.  *: : ::  **::::::* : :*.:** * *:** *::::.         : **.

CG8878   YEDHYEEKQQTTDSGNSSEQETNDDDYFLKSEKFALKKIVDIKQDEDEDDEDFDDGATSN 358
ballchen ------------------------------------------------------------

CG8878   SNNSNSLDVFHTPVNKKRSARNAIQFSGSNPVRACRREKRNSMYEEMVKSHYLRPTKRIS 418
ballchen ------------------------------------------------------------

CG8878   YREEFNEDGYPKETAENSDESPESSDNESDEFIPPSSRRSVIKRGRSAQIATPKKTPVST 478
ballchen ------------------------------------------------------------

CG8878   RASRQEKVKKEPNGDQKLRSRGSKHLDNNPTEYKFLPTEEEHVFLIDFGLASKFQDRGVH 538
ballchen ---------------------------------------AAQAYLVDFGLASHFVTG--- 215
                                                 :.:*:******:*

CG8878   RPFIMDQRRAHDGTLEFTSRDAHLGAHSRRSDLECLGYNLLYWSEGYLPWKDVAQQQQQE 598
ballchen -DFKPDPKKMHNGTIEYTSRDAHLGVPTRRADLEILGYNLIEWLGAELPWVTQKLLAVPP 274
           *  * :: *:**:*:********. :**:*** *****: *  . ***

CG8878   KVHRAKELFMTDVPEMLRQFYGKQVPKYLGEFLLQIGQLAYQERPNYERYRKIFKREYQR 658
ballchen KVQKAKEAFMDNIGESLKTLFPKGVPPPIGDFMKYVSKLTHNQEPDYDKCRSWFSSALKQ 334
         **::*** ** :: * *: :: * **  :*:*:  :.:*::::.*:*:: *. *.   ::

CG8878   LGYDPCQMRLSSEEILRTCVSTKDVVDGSKCDIFELNNKAAVNVMRNSTLSTPFQEHSLT 718
ballchen LKIP--------------------------NNGDLDFKMKPQTSSNNNLSPPGTSKAAT 367
         *                             :  :*: *   :.  *..**.*  .:: *

CG8878   NRVSPKNLRSKSNKKTTKKKFSWAEVLSQDPDQIARERAVKEFEREETICPLESRLPRRY 778
ballchen ARK-AKKIDSPVLNSSLDEKISASEDDEEEEEKSHRKKTAKKVTPSARNAKVSPLKRVAD 426
          *  .*:: *   :.: .:*:* :*  .:: ::  *::::.*:.  .   . :..

CG8878   EGKPTYAILDMEQRRREKGLVVQEHIEEEEEDADEDDEEENQEAMDIDQEEDGEAADSAE 838
ballchen SSPPSQKRVKTEPKSTPRERATPKASPKPRSTPKASPKPQTPTAARLRTPN----AKINF 482
         .. *:   :. * :   :  .. :   : .. .. . : :.  *  :   :    *.
```

```
CG8878   GEDESDRSMEGSDCSDHSQKRARGRPKGTSRKQTTSRQAQPHQNQPPVKVHRGVGRPGKN 898
ballchen SPSISLRGRPGGKTVINDDLTPQPRSKKTYEFNFELDVSMDANVIVNVKRKKKADQD--- 539
         . . * *.  *..   :.:  .: *.* * . :     :   :    ** :: ..:

CG8878   SGVVKLAAGAVSKNRTTPLSAVASNKRGCATRKENSTLASATGEGERKLKSGRTRRALYK 958
ballchen ------KATAVDSRTPSSRSALASSSKEEASPVTRVNLRKVNGHGDSSTPGRSPRTPAVT 593
               * **... .:. **:**..:  *:   . .* ...*.*: .  .  .* .  .

CG8878   TEPKHGEHDAENNSSLLVVQNLYGEYDDENNYGKGRSVHSSRHCRK 1004
ballchen VRKYQG---------------------------------------- 599
         ..  :*
```

Figure 4.8 Pairwise alignment of CG8878, and BALLCHEN, its closest *Drosophila* paralogue.

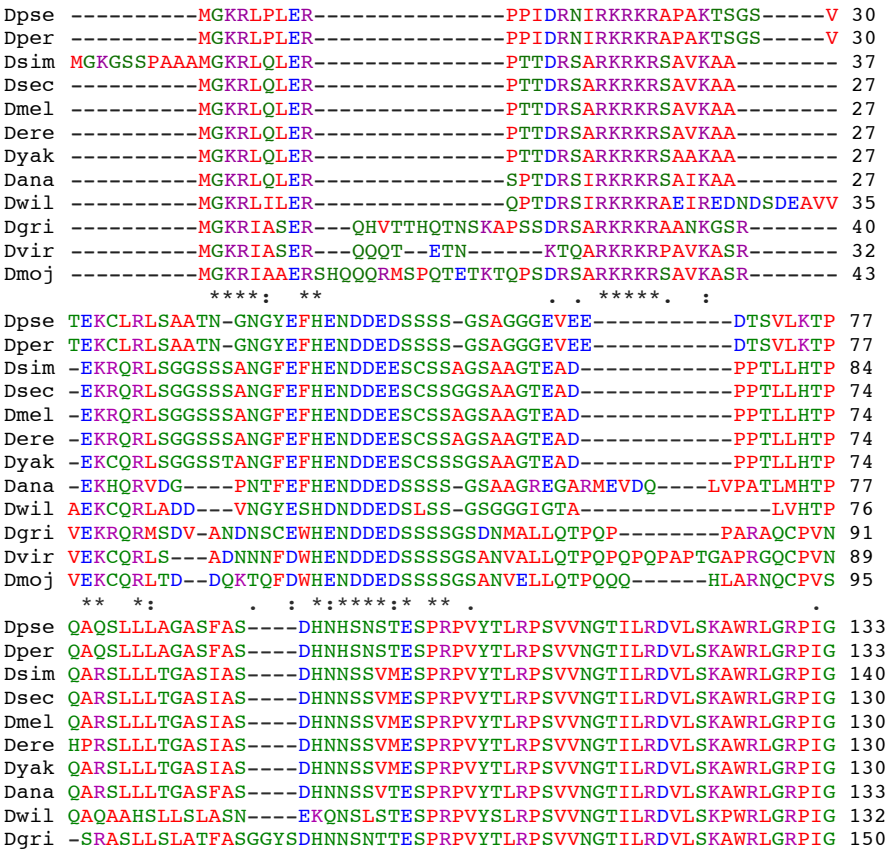
Comparison symbols: * = identity, : = side groups with strongly similar properties, . = side groups with weakly similar properties. Amino acid color code: red = small hydrophobic, blue = acidic, magenta = basic, green = hydroxyl, sulfhydryl, amine, G

Accession numbers given in Table S1.

Supplemental

Species	Protein	Accession #
Drosophila melanogaster	CG8878, isoform A	NP_610733.1
Drosophila simulans	GD15248	XP_002076301.1
Drosophila sechellia	GM20384	XP_002033485.1
Drosophila erecta	GG22606	XP_001975983.1
Drosophila yakuba	GE13474	XP_002091126.1
Drosophila ananassae	GF12642	XP_001958964.1
Drosophila pseudoobscura pseudoobscura	GA21385	XP_001360093.2
Drosophila grimshawi	GH22153	XP_001987873.1
Drosophila willistoni	GK23220	XP_002074728.1
Drosophila mojavensis	GI19382	XP_002004832.1
Drosophila virilis	GJ22443	XP_002050972.1
Drosophila persimilis	GL11027	XP_002015431.1
Drosophila melanogaster	ballchen, isoform A	NP_651508.1
Homo sapiens	VRK1	NP_003375.1
Homo sapiens	VRK2	AAH21663.1
Mus musculus	VRK1, isoform A	NP_035835.1
Mus musculus	VRK2	AAH13520.1

Table 4.S1 Polypeptide accession numbers used.



```
Dvir -SRASLLSLATFASGGFSDHNNSNATESPRPVYTLRPSVVNGTILRDVLSKAWRLGRPIG 148
Dmoj NPRASLLSLATFASGGYSDHNNSNTTESPRPVYTLRPSVVNGTILRDLLSKAWRLGRPIG 155
      .::     :  :.    :::::*   *******:*************:***.********
Dpse KGNFGEIFLASEDTVCPVSSERAKFVVKIEPHSNGPLFVEIHCLINTSQTKKIADDHEDA 193
Dper KGNFGEIFLASEDTVCPVSSERAKFVVKIEPHSNGPLFVEIHCLINTSQTKKIADDHEDA 193
Dsim KGNFGEIFLASDDTVCPASSETAKYVVKIEPHSNGPLFVEIHCLINTSRNNDLSDVAEDA 200
Dsec KGNFGEIFLASDDTVCPASSETAKYVVKIEPHSNGPLFVEIHCLINTSRNNDLSDIAEDA 190
Dmel KGNFGEIFLASDDTVCPASSETAKYVVKIEPHSNGPLFVEIHCLINTSRNNDLSDAAEDA 190
Dere KGNFGEIFLASDDTVCPASSETAKYVVKIEPHSNGPLFVEIHCLINTSQNKDLTDVAEDA 190
Dyak KGNFGEIFLASDDTVCPASSETAKYVVKIEPHSNGPLFVEIHCLINTSQSKDLSDVAEDA 190
Dana KGNFGEIFLASDDTVCPTSSENAKYVVKIEPHSNGPLFVEIHCLINTSQGKDPSGVEDNV 193
Dwil KGNFGEIFLASDETVCPVTSESAKYVVKIEPHSNGPLFVEIHCLIHTSQTKVTNGQENEE 192
Dgri KGNFGEIFLASDDTVCPVSLESAKYVVKIEPHSNGPLFVEIHCLINTNQTPEITS-GNEP 209
Dvir KGNFGEIFLASDDTVCPVSSESAKYVVKIEPHSNGPLFVEIHCLINTNQTQQITN-GSEA 207
Dmoj KGNFGEIFLASDDTVCPVTSESAKYVVKIEPHSNGPLFVEIHCLINTNQSESVTD-GNEA 214
     **********::****.: * **:*********************:*.:     . .:
Dpse ANISAR--RHIISNGPPSGIPSYIASGTHYFGDARYRFLVLPRFDRDLHSLIKNSRVEQK 251
Dper ANISAR--RHIISNGPPSGIPSYIASGTHYFGDARYRFLVLPRFDRDLHSLIKNSRVEQK 251
Dsim ASLPAP-QTHVLSRGPPSGIPSFIASGTHYFGDVRYRFLVLPRFDRDLHSLIKNSRVQQK 259
Dsec ASLPAP-QTHVLSRGPPSGIPSFIASGTHYFGDVRYRFLVLPRFDRDLHSLIKNSRVQQK 249
Dmel ASLPAP-QTHVLSRGPPSGIPSFIASGTHYFGDVRYRFLVLPRFDRDLHSLIKNSRVQQK 249
Dere ASLPAP-QTHALSRGPPSGIPSFIASGTHYFGDVRYRFLVLPRFDRDLHSLIKNSRVQQK 249
Dyak ASLPAL-QTHALSRGPPSGIPSFIASGTHYFGDVRYRFLVLPRFDRDLHSLIKNSRVQQK 249
Dana VNLPAQQQPLALAQGPPSGIPSYIASGTHYFGDARYRFIVLPRFDRDLHSLIRNSRMQQK 253
Dwil AQAKISPVAPHVAQGPPSGIPSYIASGTHYFGDGRYRFLVLPRFDRDLHSLIKNSRVQQK 252
Dgri RSQKIEKLLPRINNGPPSGIPSYIASGTHYFGDGRYRFLVLPRFDRDLHSLIKNTRVAQK 269
Dvir CS--LAQLPPTISQGPPTGIPSYIASGTHYFGDGRYRFLVLPRFDRDLHSLIKNARVAQK 265
Dmoj CA--IEKLPPRISHGPPSGIPSYIASGTHYFGDGRYRFLVLPRFDRDLHSLIKNARVAQK 272
               : .***:****:********** ****:*************:*:*: **
Dpse CLLVLAIHIINVLEHLHDKGYCHNDIKAQNLMISKCKYLKQQAVPAGKVKG--AKSDGYD 309
Dper CLLVLAIHIINVLEHLHDKGYCHNDIKAQNLMISKCKYLKQQAVPAGKVKG--AKSDGYD 309
Dsim SLLVLAVHIINVLENLHDKGYCHNDIKAQNLMVSKCKYLRRQAVP---------KGNGYE 310
Dsec SLLVLAVHIINVLENLHDKGYCHNDIKAQNLMVSKCKYLRRQAVP---------KGNGYE 300
Dmel SLLVLAVHIINVLENLHDKGYCHNDIKAQNLMVSKCKYLRRQVVP---------KGNGYE 300
Dere SLLVLAVHIIDVLENLHDKGYCHNDIKAQNLMVSKCKYLRRQAVP---------KGNGYE 300
Dyak SLLVLAVHIINVLENLHDKGYCHNDIKAQNLMVSKCKYLRRQTVP---------KGNGYE 300
Dana SLLVLAIHIIDVLEHLHDKGYCHNDIKAQNLMVSKCKYLKRQAVRVGAAADVKNKGNNFD 313
Dwil SLLVLAIHIVNVLEYLHDKGYCHNDIKAQNLMISKCKYLQRQTVS-------------RD 299
Dgri SLLLLAIHIINVLEHLHDKGYCHNDIKAQNLMMSKCKYLKRQAMQGGGSSN----RSFAS 325
Dvir SLLVLAINIINVLEHLHDKGYCHNDIKAQNLMISKCKYLKRQPVQQGGGKS----KSYD- 320
Dmoj SLLVLAIHIINVLEHLHDKGYCHNDIKAQNLMISKCKYLKRQAVQPGG-------KANYE 325
     .**:**::*::*** ******************:******::* :
Dpse EHYDEKQQTTDSGNSSEQETTDDVVKFSDGDDDYFMKNK-KLALKQIVDAAQ----EDEE 364
Dper EHYDEKQQTTDSGNSSEQETTDDVVKFSDGDDDYFMKNK-KLALKQIVDAAQ----EDEE 364
Dsim DHYDEKQQTTDSGNSSEQETND---------DDYFLKRE-KFALKKIVDIKQD---EDED 357
Dsec DHYEEKQQTTDSGNSSEQETND---------DDYFLKRE-KFALKKIVDIKQD---EDED 347
Dmel DHYEEKQQTTDSGNSSEQETND---------DDYFLKSE-KFALKKIVDIKQD---EDED 347
Dere DHYEEKQQTTDSGNSSEQETND---------DDYFLKSE-KFALKKIVDIKQD---EDED 347
Dyak DHYEEKQQTTDSGNSSEQETND---------DEYFLKSE-KFALKKIVDIKQD---EDED 347
Dana DHYDEKQQTTDSGNSSEQETND---------DDFFLKRE-RFPLKRLAHIDQE---DDED 360
Dwil KHYEEKQQTTDSGNSSEQETTND--------DEYFLKNE-KFALKRMAHIGKG---KDED 347
Dgri EHYDEKQQTTDSGNSSEQEAND---------EDYFVKSE-KLALKRLVGVKEDNDAEDED 375
Dvir EHYEEKQQTTDSGNSSEQETND---------DDYFAKND-KFALKHIAPIKEDLDAEDED 370
Dmoj EHYEEKQQTTDSGNSSEQETND---------DDYFIKNENKYALKRIAGIKEDVDDVDED 376
     .**:****************:.:          :::* * . : .**::.   :     **:
Dpse D---EDFDDGATSNSNNSNS--LDTYHTPLN-KNKGSVRKTNIEFSGSNPVRSCRREKR- 417
Dper D---EDFDDGATSNSNNSNS--LDTYHTPVN-KNKGSVRKTNIEFSGSNPVRSCRREKR- 417
Dsim D---EDFDDGATSNSNNSNS--LDVFQTPVN-KKRS--VRNAVQFSGSNPVRACRREKR- 408
Dsec D---EDFDDGATSNSNNSNS--LDVFQTPVN-KKRS--VRNAVQFSGSNPVRACRREKR- 398
Dmel D---EDFDDGATSNSNNSNS--LDVFHTPVN-KKRS--ARNAIQFSGSNPVRACRREKR- 398
Dere D---EDFDDGATSNSNNSNS--VDIFQTPVN-KKRS--TRNPVQFSGSNPVRSCRREKR- 398
Dyak D---EDFDDGATSNSNNSNS--LDVFHTPVN-KKRS--VRNPVQFSGSNPVRSCRREKR- 398
Dana ----EDFDDGATSNSNNSNS--MDVYHTPIN-KKRR--GRNNVQFSGSNPVRSCRREKR- 410
Dwil MEEDEDFDDGATTNSNNSNSNSLDCYQTPVN-KKRAR-PRTGVEFSGSNPVRSCRREKR- 404
Dgri ED----FDDGATTNSNNSNS--LDMYQTPVN-KRKGRARTNAIEFSGSNPMRSCRRNDTR 428
Dvir EDEDEDFDDGATTNSNNSNS--LDLYQTPVSNKNKRRARNAPVEFSGSNPMRSCRRNDRR 428
Dmoj ED----FDDGATTNSNNSNS--LDIYQTPVN-KNKRRPRQNAVEFSGSNPMRSCRRNENC 429
           ******:*******  :* ::**:. *.:        ::*******:*:***:.
Dpse --NSMYEEMVKSHYLRPTKRVSYREDFN-EDGYPIKNNED-----KDDQSPVTSDNDSEE 469
Dper --NSMYEEMVKSHYLRPTKRVSYREDFN-EDGYPIKNNED-----KDDQSPVTSDNDSEE 469
Dsim --NSMYEEMVKSHYLRPTKRISYREEFN-EDGYPKETAE------NSDESPESSDNESDE 459
Dsec --NSMYEEMVKSHYLRPTKRISYREEFN-EDGYPKETAE------NSDESPESSDNESDE 449
Dmel --NSMYEEMVKSHYLRPTKRISYREEFN-EDGYPKETAE------NSDESPESSDNESDE 449
```

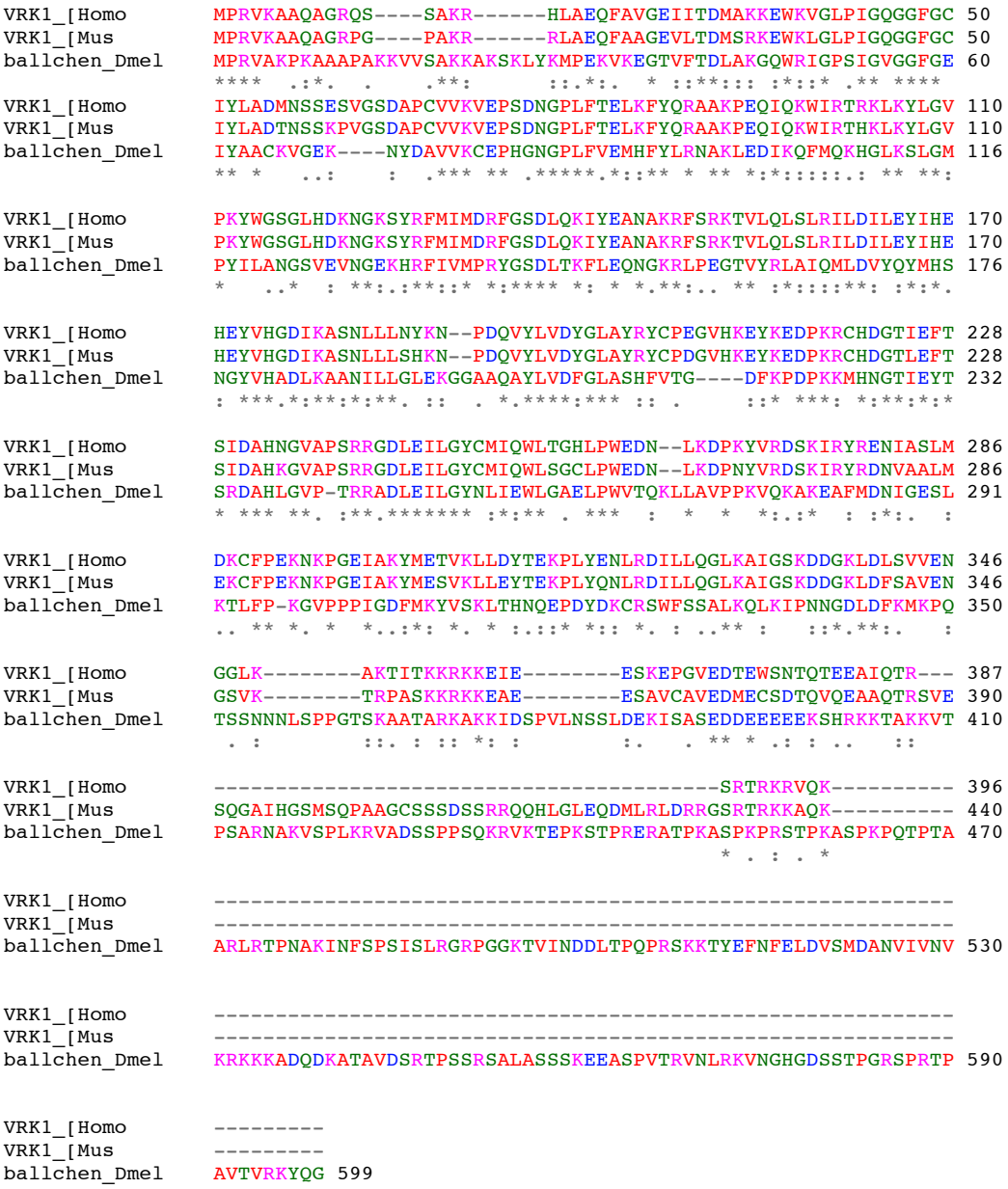
```
Dere --NSMYEEMVKSHYLRPTKRISYREEFN-EDGYPKDTAE------NSDESPESSDNESDE 449
Dyak --NSMYEEMVKSHYLRPTKRISYREEFN-EDGYPKNTAE------NSDESAESSDNESDE 449
Dana --NSMYEEMVKSHYLRPTKRISYREEFN-EAGDPIKASED-----NSDQSAESSDNESEE 462
Dwil --NSMYEEMVKSHYLRPTKRVSYREEFN-EEGYPIKEEQQQAEGANSDQSPVSSDNESEE 461
Dgri NSSSMYEEMVKSHYLRPAKRVSYSELLVNEDGYPVKPDAD------NEQSPVSSDNESEE 482
Dvir NSSSMYEEMVKSHYLRPAKRVSYSELFN-EDGYPVKADAN------SEQSPESSDNESDE 481
Dmoj NSSSMYDEMVKSHYLRPAKRVSYSELFN-EDGYPVKADAK------SEQSAESSDNDSDE 482
       .***:**********:**:** * :  * * * .          .::*. :***:*:*
Dpse FLPPSARRATAAVGKRARHAQAHAHPSTPSKCSITTRATRHQSKLKSEMSEGSK-RSGRR 528
Dper FLPPSARRATAAVGKRARHAQAHAHPSTPSKCSITTRATRHQSKLKSEMSEGSK-RSGRR 528
Dsim FIPPSNRRPAIKRGR-------SAQIATPKKTPVSTRVSR-QEKVKKEPNVEQKLRSRGS 511
Dsec FIPPSNRRPAIKRGR-------SAQIATPKKTPVSTRVSR-QEKVKKEPNGDQK-RSRGS 500
Dmel FIPPSSRRSVIKRGR-------SAQIATPKKTPVSTRASR-QEKVKKEPNGDQKLRSRGS 501
Dere FIPPSSRRSASKRGK-------GTQIATPKKCPVSTRATRHQEKVKKEPNGDQKSRSRGS 502
Dyak FIPPSSRRTASKRGR-------SVQIATPKKCPVSTRATRHQEKVKKEPNGDQKARSRGS 502
Dana FIPPSERRSTIKRGR-------PALASSAKKGPAPTRAKRNQEKVKKEPLGGNRPRGRGN 515
Dwil FLPPSIRRSTKKSSA----------SFSAAKRLGTRRQQKIMKNREMIAMAEESSRSRDG 511
Dgri FLPPSARR-GGGSVA------SKRS-SRSAACKSKTTRAMRRQDQKEELKAELNDPGRNG 534
Dvir FVPPCARRSGGASAA------KQRI-SRSSTRAVTRRQESKQQQQQQHLKSNVSEPVPKA 534
Dmoj FVPAFARRSGGASAA------KRGARARTSNRMVTRRRETKQKQ----LKTEITENPRNG 532
     *:*.  **                    .              ..
Dpse KQDDTQSHAEYQLLPAEEEHVFLIDFGLASKYQDRGVHRPFIMDQRRAHDGTLEFTSRDA 588
Dper KQDDTQSHAEYQLLPAEEEHVFLIDFGLASKYQDRGVHRPFIMDQRRAHDGTLEFTSRDA 588
Dsim KHVDCNP-TEYKFLPTEEEHVFLIDFGLASKFQDRGVHRPFIMDQRRAHDGTLEFTSRDA 570
Dsec KHVDSNP-TEYKFLPTEEEHVFLIDFGLASKFQDRGVHRPFIMDQRRAHDGTLEFTSRDA 559
Dmel KHLDNNP-TEYKFLPTEEEHVFLIDFGLASKFQDRGVHRPFIMDQRRAHDGTLEFTSRDA 560
Dere KHLDNNP-SEYKFLPTEEEHVFLIDFGLASKFQDRGVHRPFIMDQRRAHDGTLEFTSRDA 561
Dyak KHLDNNP-SEYKFLPTEEEHVFLIDFGLASKFQDRGVHRPFIMDQRRAHDGTLEFTSRDA 561
Dana KLVENQPPTEYQFVPVEEEHVFLIDFGLASKYQDRGVHRPFIMDQRRAHDGTLEFTSRDA 575
Dwil KRVTTNP-PQYQLIPVEEEHVFLIDFGLASKYQDRGVHRPFIMDQRRAHDGTLEFTSRDA 570
Dgri GVRNKRTAMQYQLMPVEEEHIFLIDFGLASKYQDRGVHRPFIMDQRRAHDGTLEFTSRDA 594
Dvir MSRSKRAATQYQLVQVEEEHIFLIDFGLASKYQDRGVHRPFIMDQRRAHDGTLEFTSRDA 594
Dmoj GVRSKRTVTQYQHVPVEEEHIFLIDFGLASKYQDRGVHRPFIMDQRRAHDGTLEFTSRDA 592
           ..  :*: : .****:**********:****************************
Dpse HLGAHSRRSDLECLGYNLVYWSEGYLPWKDVAQQQQQEKVHRAKELFMTDVSEMLRQFYG 648
Dper HLGAHSRRSDLECLGYNLVYWSEGYLPWKDVAQQQQQEKVHRAKELFMTDVSEMLRQFYG 648
Dsim HLGAHSRRSDLECLGYNLLYWSEGYLPWKDVAQQQQQEKVHRAKELFMTDVPEMLRQFYG 630
Dsec HLGAHSRRSDLECLGYNLLYWSEGYLPWKDVAQQQQQEKVHRAKELFMTDVPEMLRQFYG 619
Dmel HLGAHSRRSDLECLGYNLLYWSEGYLPWKDVAQQQQQEKVHRAKELFMTDVPEMLRQFYG 620
Dere HLGAHSRRSDLECLGYNLLYWSEGYLPWKDVAQQQQQEKVHRAKELFMTDVPEMLRQFYG 621
Dyak HLGAHSRRSDLECLGYNLLYWSEGYLPWKDVAQQQQQEKVHRAKELFMTDVPEMLRQFYG 621
Dana HLGAHSRRSDLECLGYNLVYWSEGCLPWKDAAQQQQQEKVHRAKELFMTDVPEMLRQFYG 635
Dwil HMGAHSRRSDLECLGYNLLYWSEGFLPWKEVASQQQQEKVHRAKELFMTDVSEMLRQFYG 630
Dgri HLGAHSRRSDLECLGYNLLYWSEGYLPWKDVAHQQQQEKVHRAKELFMTDVCEMLRQFYG 654
Dvir HLGAHSRRSDLECLGYNLLYWSEGYLPWKDVAQQQQQEKVHRAKELFMTDVCEMLRQFYG 654
Dmoj HLGAHSRRSDLECLGYNLLYWSEGYLPWKDVAQQQQQEKVHRAKELFMTDVCEMLRQFYG 652
     *:*****************:***** ****:.* ****************** ********
Dpse KQVPKYLGEFLKQIGQLAYQERPNYERYRNIFKREFHRLGHDPSQMRLNSEEILLTRVA- 707
Dper KQVPKYLGEFLKQIGQLAYQERPNYERYRNIFKREFHRLGHDPSQMRLNSEEILLSRVA- 707
Dsim KQVPKYLGEFLLQIGQLAYQERPNYERYRKIFKREYQRLGYDPNQMRLSSEEILRTCVS- 689
Dsec KQVPKYLGEFLLQIGQLAYQERPNYERYRKIFKREYQRLGYDPNQMRLSSEEILRTCVS- 678
Dmel KQVPKYLGEFLLQIGQLAYQERPNYERYRKIFKREYQRLGYDPCQMRLSSEEILRTCVS- 679
Dere KQVPKYLGEFLLQIGQLAYQERPNYERYRKIFKREYQRLGYDPSQMRLSSDEILRTCVS- 680
Dyak KQVPKYLGEFLLQIGQLAYQERPNYERYRKIFKREYQRLGYDPSQMRLSSDEILRTCVS- 680
Dana KQVPKYLGEFLQQIGKLAYQERPNYERYRNIFRNEYQNLGFDLDKMRLSSEEIQRTCVS- 694
Dwil KQIPKYLGEFLKLIGQLSYQERPNYQRYRNIFKREYRRLGHDPNQMRLSSDEILSTCVN- 689
Dgri KQVPKYLGEFLKEIGQLAYQERPNYERYRNIFQREYRRLGYDCSQMQLSSDEILRTRVS- 713
Dvir KQVPKYLGEFLQQIGQLAYQERPNYERYRSIFKREYRRLGYDCSQMQLSSDEILRTRVSC 714
Dmoj KQVPKYLGEFLQQIGQLAYQERPNYERYRNIFKREYRRLGYDCSQMQLSSHDIQRTRIS- 711
     **:********  **:*:*******:***.**:.*::.**.*  :*:*.*.:*  : :
Dpse VKDEMYG------NKCDIFELNNKISTNVMRNSTLSTPFQEHSITNRVSPKNLRSKSSKK 761
Dper VKDEMYG------NKCDIFELNNKISTNVMRNSTLSTPFQEHSITNRVSPKNLRSKSSKK 761
Dsim TKDVVDG------SKCDIFELNNKAAVNVMRNSTLSTPFQEHSLTNRVSPKNLRSKSNKK 743
Dsec TKDVVDG------SKCDIFELNNKAAVNVMRNSTLSTPFQEHSLTNRVSPKNLRSKSNKK 732
Dmel TKDVVDG------SKCDIFELNNKAAVNVMRNSTLSTPFQEHSLTNRVSPKNLRSKSNKK 733
Dere AKDVVDG------SKCDIFELNNKAAANVMRNSTLSTPFQEHSLTNRVSPKNLRSKSNKK 734
Dyak AKDVVDG------SKCDIFELNNKAAVNVMRNSALSTPFQEHSLTNRVSPKNLRSKSNKK 734
Dana VKDEVDA------SRCDIFELNNKITANVMRNAALSTPFQEHALTNRVSPKNLRSKSSKK 748
Dwil IKDEVDGGATPASNKCDIFDVNNKFFSNAIRNSALHTPFQEHSLTNRVSPKNLRSKSEKK 749
Dgri VKDVVDG-----SAKCDIFELNNKIACNVMRNATLSTPFQEHALTNRVSPKNLRSKSNRK 768
Dvir IKDELDG-----GAKCDIFELNNKIACNVMRSATLSTPFQEHSLTNRVSPKNLRSKSNKK 769
Dmoj IKDEVDG-----IGKCDIFDLNSKSAFNQLRNSTFSTPFQDHSLTNRVSPKNLRSKSNKK 766
      ** : .       :****::*.*   * :*.::: ****:*::**************.:*
```

```
Dpse NVKKK-FSWAEVISQDPDQIARERAAKEFEREETICPLQMRLPKRYEGRPTYAILTVEQS 820
Dper NVKKK-FSWAEVISQDPDQIARERAAKEFEREETICPLQMRLPKRYEGRPTYAILTVEQS 820
Dsim TTKKK-FSWAEVLSQDPDQIARERAVKEFEREETICPLESRLPRRYEGKPTYAILDMEQR 802
Dsec TTKKK-FSWAEVLSQDPDQIARERAVKEFEREETICPLESRLPRRYEGKPTYAILDMEQR 791
Dmel TTKKK-FSWAEVLSQDPDQIARERAVKEFEREETICPLESRLPRRYEGKPTYAILDMEQR 792
Dere TTKKK-FSWAEVLSQDPDQIARERAVKEFEREETICPLESRLPRRYEGKPTYAILDMEQR 793
Dyak TTKKK-FSWAEVLSQDPDQIARERAVKEFEREETICPLESRLPRRYEGKPTYAILDMEQR 793
Dana TTKKK-FSWAEVLSQDPDQIARERAVKEFEREESICPLQSRLPRRYDGKPTYAILAVEQS 807
Dwil NVKKKSFSWAEVLLQDPDQIARERAVKEFEREEEICPLKSRLPRRYEGKATYAILAVEQS 809
Dgri NVKKKIFSWAEVLSQDPDQIARERAVKEFEREEVICPLQSRLPRRYEGYPTYAILAVEQS 828
Dvir NAKKKKFSWAEVLSQDPDQIARERAVKEFEREEVICPLQSRLPRRYEGRPTYAILAVEQS 829
Dmoj NAKKLKFSWAEVLSQDPDQIARERAVKEFEREEVICPLHSRLPRRYEGKPTYAILAVEQS 826
     ..**  ******: **********.******* ****. ***:**:* .***** :**
Dpse RRDKGLVVQEHNEGE----------CHDEVDALTREQDQEEEEDNDAAESTDGEYAEQG 869
Dper RRDKGLVVQEHNEGE----------CHDEVDALTREQDQEEEEDNDAAESTDGEYAEQG 869
Dsim RREKGLVVQEHNEEEEGEAD-------EDDEEENQETIDVEQQE--EAADSEEGEDESDR 853
Dsec RREKGLVVQEHNEEEEGEAD-------EDDEEENQEAMDVEQQEDEEAADSEEGEDESDR 844
Dmel RREKGLVVQEHIEEEEEDAD-------EDDEEENQEAMDIDQEEDGEAADSAEGEDESDR 845
Dere RREKGLVVQEHNEEEEEEEE-----EEEDDEEENQEAVDDEQKDD-EAADSVEGEDESDR 847
Dyak RREKGLVVQEHNEEREEEEED---DEEDDDDEENQEALDEEE-DE-EAADSVEGEDDSDR 848
Dana RRDKGLVVQEHIREELNEDD------EEEDDEQNEEEEEHIDEQEEEGKEEVEEEE---- 857
Dwil RRDKGLTVQEHINED--------------DNQADQEEEDENEDDKDEEEEDEDESNADED 855
Dgri RRDKGLIVQEHKTEEAMDEGNTHKASESDEEYEDDEQEQEEEEEQEEEEEVEEEQEEEG 888
Dvir RRDKGLSVQEHNNEEASETDKYEAAEDNDEDYEEEEEEDEEEDQEEEEEEEEEEEEEEDE 889
Dmoj RRERGLSVQEHNSEEADDQVG-ESAETAEEEEEEEEDEEEAEEEEEDEEEVEEEAEQEA 885
     **::** ****                  :     .  :  : :  :  :. : .
Dpse TSEESD-------------------------YSDHSEQAVRRRVGGGAYRKRTSKQTQTQ 904
Dper TSEESD-------------------------YSDHSEQAVRRRVGGGAYRKRTSKQTQTQ 904
Dsim SMEDSD-------------------------CSDHSQKRARGRPKGTHRKQTTSRQTQYQ 888
Dsec SMEDSD-------------------------CSDHSQKRARGRPKGTHRKQTTSRQTQSQ 879
Dmel SMEGSD-------------------------CSDHSQKRARGRPKGTSRKQTTSRQAQPH 880
Dere AMEESD-------------------------YSDNSQKHARGRTKATTRKRTTNRQTQSQ 882
Dyak AMEESD-------------------------YSDHSQKRARGRPKGTTRKRTTSRQTQSQ 883
Dana -VDGSD-------------------------NSNQIARAGRGRPKGSGRKRTTSKPGQAP 891
Dwil ATEETD-------------------------SEE--GRVSRRGGHGRVRKRTTSKQPTTK 888
Dgri EEREAD----------AASDEE------SDESECTTQSEAVRRGRGRPRNCSKQRQLQQQ 932
Dvir EEEEEDEVDMEKAEPEATSEDEEMDYQQSDESEGTTQSETVRRGRGRPRNNSKLIQVQQA 949
Dmoj MDVDQQ----------------------SVDSESTTESRQTRRARRQPRKNSSSSKASQG 923
         :                          .                 .  .
Dpse ----QLKSNRGVSKINKNIASAKFAGGAVSKSR--------------STPLS-AVASNKR 945
Dper ----QLKSNRGVSRSNKNIASAKFAGGAVSKSR--------------STPLS-AVASNKR 945
Dsim QNQPPVRVHRGVGRPGKNSGVVKLAAGAVSKNR--------------TTPLS-AVASNKR 933
Dsec QNQPPVKVHRGVGRPGKNSGVVKLAAGAVGKNR--------------TTPLS-AVASNKR 924
Dmel QNQPPVKVHRGVGRPGKNSGVVKLAAGAVSKNR--------------TTPLS-AVASNKR 925
Dere QNQPSTKVHRGGSRPSKNLGVVKFAAGAISKNR--------------STPLS-AVASNKR 927
Dyak QNQPSVKNHRGVGRPSKNSGVVKFAAGAVSKNR--------------STPLS-AVASNKR 928
Dana -SQPPVKCS--------SNRVAARRLGAVSKIR--------------STPLS-AVASNKR 927
Dwil --------------LGRVATAGKGKSTNVAVDK--------------STPLSTVVASNKR 920
Dgri QTLAKSTEMRKKTLRP-SGKSRNSNGHSQQ----------------IATPLTAVGGN-KR 974
Dvir QQQQQQLGKGKGATRKRTGKLRNTNGDSQQQQQQQQEEQEQQDHRMIVTPLTTVGGN-KR 1008
Dmoj QQLPQLSKSKGTTRKRSSNKSKNINGHTQQQQQ---------AHRIVGTPLTTVGGNNKR 974
                                                  ***: . .. **
Dpse GCATRKENA------TVSSATDE--VER-------------------KLKSRRSPED--- 975
Dper GCATRKENA------TVSSATDE--VER-------------------KLKSRRSPED--- 975
Dsim GCATRKENS------TLASATGE--GER-------------------KLKSG-------- 958
Dsec GCATRKENS------TLASATGE--GER-------------------KLKSG-------- 949
Dmel GCATRKENS------TLASATGE--GER-------------------KLKSG-------- 950
Dere GCATRKENS------TLASATGE--VER-------------------KLKTT-------- 952
Dyak GCATRKENS------TLASATGE--GEQ-------------------KLKTS-------- 953
Dana GCATRKENTTTTTSTTLASATGEPAGER-------------------KLKTGGRDKSA-- 966
Dwil GCATRKEHA------IAASATGE--TER-------------------KLKASQTRREEDL 953
Dgri GCATRKEN------ATMASATGEVEQHS-------------------KLKSRQ------- 1002
Dvir GCAARKEN------ATMASATGELEQHHHHHHHHHNNNNNNNNNNNNKLKSSR------- 1055
Dmoj GCATRKEQQQT---ATMASATGEVQQQT-----------------VKLKSSRNAAAA-- 1011
     ***:***:         :***.*   .                   ***:
```

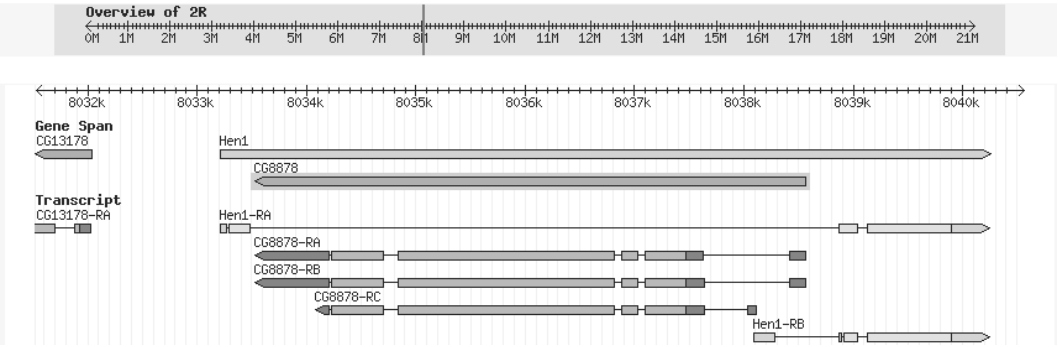
```
Dpse ---------RPKQRRTRRCLYKTESKIG---EHDVENNSSLLE-VQN---LYSEYDDENK 1019
Dper ---------RPKQRRTRRCLYKTESKIG---EHDVENNSSLLE-VQN---LYSEYDDENK 1019
Dsim --------------RTRRALYKTEPKHG---EHDAENNSSLLV-VQN---LYGEYDDENN 997
Dsec --------------RTRRALYKTEPKHG---EHDAENNSSLLV-VQN---LYGEYDDENN 988
Dmel --------------RTRRALYKTEPKHG---EHDAENNSSLLV-VQN---LYGEYDDENN 989
Dere --------------RTRRALYKTEPKHG---EHDVENNSSLLL-VQN---LYGEYDDENN 991
Dyak --------------RTRRALYKTEPKHG---EHDAENNSSLLL-VQN---LYGEYDDENN 992
Dana ----PQTGRQPKQRQTRRCLYKTESQLG---ENDVENNANLLE-VQN---LYGEYDDENN 1015
Dwil HHHQKQQQLLPKQRRPRRCLHKTEAIMGVV-QQDVENNSELLP-MQS---AYGDYDDENN 1008
Dgri ----PVVDQP-KQRRTRRCLYKTESHIRGEKQQDVENNYATN--MQN-VYSVWQYDDENR 1054
Dvir ----SVDQDPTKQRRTRRCLYKTEAHMHGERQQDVENNYATN--MQN-VYSVWQYDDENS 1108
Dmoj ---VAVADDPGKQRRTRRCLYKTEA--------DVENNYATNNNMQNDAYSVWQYDDENS 1060
                   :.**.*:***.         *.***      :*.      :*****

Dpse YI-KGRNVNPS-RHSRKL 1035
Dper YI-KGRNVNPS-RHSRKL 1035
Dsim YG-KGRSVHSS-RHCRK- 1012
Dsec YG-KGRSVHSS-RHCRK- 1003
Dmel YG-KGRSVHSS-RHCRK- 1004
Dere YG-KGRSVHSS-RHCRK- 1006
Dyak YG-KGRSVHSS-RHCRK- 1007
Dana YG-KGRHVHAS-RHCRK- 1030
Dwil YGAKGRNVHASSRHCRK- 1025
Dgri YG-KGRNVNSS-RHCRK- 1069
Dvir YG-KGRNVNSS-RHCRK- 1123
Dmoj YG-KGRNVNSSSRHCRK- 1076
     *  *** *:.* **.**
```

Supplemental Figure 4.1 Pairwise alignment of CG8878 and 12 *Drosophila* homologues. Species names are abbreviated using the capitalized first letter of the genus followed by the first three letters of the species. Comparison symbols: * = identity, : = side groups with strongly similar properties, . = side groups with weakly similar properties. Amino acid color code: red = small hydrophobic, blue = acidic, magenta = basic, green = hydroxyl, sulfhydryl, amine, G (http://www.ebi.ac.uk/Tools/msa/clustalw2/ . Note: for *D. persimilis* a nucleotide was removed (five A's to four A's – a presumed sequencing error) to facilitate amino acid alignment. Accession numbers given in Table S1.



Supplemental Figure 4.2. Pairwise alignment of BALLCHEN and VRK1 from mouse and humans. Comparison symbols: * = identity, : = side groups with strongly similar properties, . = side groups with weakly similar properties. Amino acid color code: red = small hydrophobic, blue = acidic, magenta = basic, green = hydroxyl, sulfhydryl, amine, G (http://www.ebi.ac.uk/Tools/msa/clustalw2/). Accession numbers given in Table S4.1.



Supplemental Figure 4.3. *CG8878* lies within a large intron of *Hen1*.

Top. Location of *CGG8878* on chromosome 2R.

Middle. Expanded view of the orientation and location of *CG8878*, *Hen1* predicted transcripts.

Bottom. *Hen1*, *CG8878* predicted transcripts. Exons are represented by boxes, Introns by lines (light grey = translated, dark grey= nontranslated).

Taken from flybase.

http://flybase.org/cgi-bin/gbrowse/dmel/?Search=1;name=FBgn0027504

Chapter 5 – Discussion

This screen of ~44,000 progeny of EMS mutagenized *y⁻ w⁻; dp⁻; e⁻; E1* males that had been mated to *y⁻ w⁻* virgin females for phenotypic enhancement of w^+ silencing in *E1*/+ yielded 58 mutants that transmitted the enhancer phenotype in transvection with *Pci*. Complementation analysis yielded three simple complementation groups that mapped to the third chromosome and two to the second whose enhancer phenotypes were strong enough to allow recombination mapping of the enhancer phenotype. The weaker and complex recessive lethal complementation groups, as well as mutants that did not fall into a complementation group were not pursued any further. Of the three workable complementation groups on the third chromosome, two were identified as *trxG* members *ash1* and the eponymous *trx* itself, while the third was identified as *Taf4,* a transcription initiation factor TFIID component involved in pre-initiation complex assembly. The two workable complementation groups on the second chromosome were identified as *cg*, a previously known regulator of *ci*, and CG8878, a putative Serine/Threonine/Tyrosine kinase of unknown specificity.

ash1

I have shown that the 5 *En*(*PDS*) in one of the complementation groups described in this study are alleles of *ash1*. The enhancement and lethal phenotypes mapped to the *ash1* locus and could not be separated by recombination. All 5 alleles failed to complement all deficiencies tested that uncover *ash1* and all 5 alleles failed to complement both extant alleles $ash1^{22}$ and $ash1^{B1}$. Four of the alleles contain mutations resulting in stop codons before ASH1's SET domain and therefore likely represent null alleles, the fifth is a mutation in the middle conserved histidine residue in ASH1's PHD finger. Together with 2 conserved cysteine residues they coordinate 2 Zn^{2+} ions; therefore, loss of the histidine residue should abolish this function. The ING2, YNG1 and NURF PHD fingers have been

reported to bind to histone H3 tri-methylated on lysine 4 (H3K4me3), this may therefore be a common property among PHD fingers (Shi *et al*. 2006, Martin *et al*. 2006, Wysocka *et al*. 2006). The modification H3K4me3 is associated with the transcription start site of active genes, and therefore may be part of how ASH1 recognizes a transcriptionally active gene sequence required to be "locked on" by *trxG* proteins.

The gene *absent, small, or homeotic discs* (*ash1*) was first identified by Shearn *et al*. (1971) by lethal mutations that caused a variety of imaginal disc developmental defects in *Drosophila*. It was subsequently identified as a member of the TRX group of genes involved in antagonizing PcG repression and maintaining active transcription of homeotic genes (Shearn 1989). In addition to the aforementioned PHD finger and a Bromo adjacent homology domain, the ASH1 protein contains both pre- and post-SET domains as well as a SET domain with 32% identity to that of SU-(VAR)3–9 (Tripoulis *et al*.,1996). Biesel *et al*. (2002) demonstrated that the isolated ASH1 SET domain with its preSET and postSET domains has HMTase activity *in vitro* with specificity for nucleosomal histone 3 lysines 4 and 9 and, to a lesser extent, lysine 20 of histone H4. However, given that the ASH1 fragment used was lacking both AT hooks and the PHD finger implicated in specificity of chromatin binding, these results should be viewed only as confirming HMTase activity but not specificity. Point mutations resulting in single amino acid substitutions in either the SET domain (*ash1*^{10} (N1458I, previously N1385I)) or preset (associated with SET (AWS)) domain (*ash1*^{21} (E1357K, previously E1284K)) abolish *Drosophila* ASH1 function (Tripoulas *et al.* 1994). *In vitro* HMTase assays using recombinant proteins containing either of these mutations as well as a third mutant (Ash1DeltaN^{1142}) with a SET domain mutation (H1459K) previously shown to abolish SUV39H1 HMTase activity (Rea *et al*. 2000) revealed no significant H3 or H4 methylation. In addition, transgenic flies carrying an Ash1-dependent reporter gene consisting of a 4 kb regulatory

element of the *bxd* region from *Ubx* (an ASH1target) fused to the mini-white gene showed significant reduction in *w* expression in $ash1^{10}$/+ or $ash1^{21}$/+ heterozygous flies compared to wild type (Beisel *et al*. 2002). Taken together, these results indicate that both the pre-SET and SET domains of ASH1 are necessary for HMTase activity and transcriptional activation by ASH1 (Beisel et al., 2002). These authors also investigated the interaction between HP1 and H3 peptides and histone core octamers that had been methylated either by ASH1 or SU(VAR)3-9, an H3 K9 methylase. HP1 bound H3 K9 and H3 K4/K9-methylated peptides and H3 K9-methylated histone core, but did not bind Ash1-methylated core octamers. This suggests that the ASH1 methylation pattern inhibits the interaction of HP1 with chromatin, thus preventing the spread of heterochromatin over transcriptionally active gene regions. Protein-binding assays with Brm and Mor, subunits of a SWI/SNF-like chromatin remodeling complex (Simon & Tamkun, 2002), indicate they interacted with ASH1-methylated peptides but not with peptides methylated at H3 K4 or H3 K9 alone. XChIP indicated that Brm and Mor were present at active but not silent promoters of Ash1 target genes in BCAT5 transfected cells or third leg imaginal discs while repressors HP1 and Pc were only detected at silent promoters. These results imply that transcriptional activation by Ash1 coincides with loss of repressor binding at the promoter of Ash1 target genes and the subsequent recruitment of Brm and Mor. ASH1 methylation both facilitates the interaction of proteins involved with transcriptional activation and prevents the interaction of transcriptional repressor proteins with methylated H3 during epigenetic activation (Beisel et al., 2002). Byrd and Shearn (2003) showed that ASH1 was necessary *in vivo* for most of the detectable H3-K4 methylation, and confirmed the previous groups finding that ASH1's SET domain alone was sufficient for H3-K4 methylation *in vitro*. H3-K4 methylation was virtually abolished in *ash1* mutants lacking preSET, SET, and postSET domains; however,

there was no detectable reduction in H3-K4 methylation caused by an *ash1* allele with intact preSET, SET, and postSET domains. While virtually all H3-K9 methylation was restricted to the heterochromatic chromocenter of wild-type polytene chromosomes as shown by immunofluorescence with antidimethyl-H3K9, in larvae heteroallelic for antimorphic/null alleles of *ash1* this centromeric H3-K9 methylation was slightly reduced suggesting that ASH1 may also be required for some H3-K9 methylation. Immunofluorescence with antidimethyl-H4K20 and antidimethyl-H3K36 gave a distribution of methylated H4-K20 and H3-K36 not detectably different from wild-type. This contradicted the earlier results of Biesel *et al* (2002) and suggested that other enzymes were responsible for nearly all of the H3-K36 and H4-K20 methylation *in vivo*.

After having shown that *ash1* promotes transcriptional activation by presumably trimethylating H3-K4, H3-K9, and H4-K20, and that expression of the homeotic gene Ultrabithorax (Ubx) in third-leg and haltere imaginal discs coincides with this ASH1-mediated histone methylation (Beisel et al., 2002), the same laboratory investigated how ASH1 is recruited to its target genes upon transcriptional activation. It had already been shown that both PcG and trxG regulators are recruited to specific chromosomal elements in the 5' regulatory region of target genes and that the same element could act as either an activating or silencing regulatory element. In the transcriptionally repressed state, the elements facilitate the recruitment of PcG proteins and are therefore designated as Polycomb response elements (PREs) while in the activated state, they recruit trxG proteins and are termed trithorax response elements (TREs) (Orlando 2003; Ringrose & Paro 2004). By default, PRE/TREs recruit PcG complexes and silence associated genes; however, upon transcriptional activation they instead recruit counteracting trxG proteins, which maintain target genes in a transcriptionally active state. This switch from the silenced to the activated state of a PRE/TRE requires transcription of

noncoding RNAs (ncRNAs) from the TRE/PRE elements themselves (Schmitt *et al*. 2005). These authors found that *actin* promoter induced continuous transcription through a PRE/TRE prevented the establishment of PcG-mediated silencing. Maintenance of epigenetic activation at the BX-C requires this PRE/TRE transcription to continue until the end of embryogenesis, indeed, intergenic PRE transcripts can be detected until late larval stages, suggesting that PRE/TRE transcription is continuously required to prevent the recruiting of repressive PcG complexes. Importantly, they found that all other PREs tested (outside the BX-C) were transcribed in the same tissue as the mRNA of the corresponding target gene, suggesting that the antagonizing of PcG repression via PRE/TRE transcription is a fundamental epigenetic mechanism. They did not detect any transcription in the promoter of *rosy* (*ry*), which is not regulated by a PRE (Ringrose *et al*. 2003), leading them to conclude that upstream transcription is not a general property of active promoters but a PRE/TRE-specific phenomenon. Sanchez-Elsner *et al*., (2006) provided the link between the transcription of PRE/TREs and the recruitment of the trxG regulator ASH1 by analyzing the role of three ncRNAs transcribed from different Ubx TRE/PREs within the chromosomal memory element (CME) *bxd* situated 22 kb upstream of the Ubx promoter. They detected transcripts from Ubx and all three TREs in third-leg and haltere imaginal discs but not in wing imaginal discs or Schneider 2 (S2) cells indicating that TRE and Ubx transcription coincide. *In vivo* cross-linked chromatin immunoprecipitation (XChIP) indicated the presence of ASH1 at all three *Ubx* TREs in third-leg and haltere imaginal discs, but not in wing imaginal discs and S2 cells even though they express ash1 (Beisel *et al*., 2002, Tripoulas *et al*., 1996). Comparison of wild-type and homozygous $ash1^{22}$ third-leg discs by XChIP indicated both ASH1 and its characteristic histone methylation pattern at the *Ubx* locus in wild-type discs but not in $ash1^{22}$ mutant discs even though TRE transcripts were detected at comparable

levels in both backgrounds. Thus ASH1 recruitment and histone methylation coincides with activation of Ubx expression in third-leg discs. Since these findings indicate that ASH1 is not a major regulator of TRE transcription in imaginal discs, the question of what maintains PRE/TRE transcription itself remains unanswered.

The finding that ASH1 association with TREs requires the production of TRE transcripts implicates TRE transcripts in the recruitment of ASH1 to Ubx TREs. Sanchez-Elsner (2006) used *in vitro* protein-RNA binding assays to show that the ASH1SET domain consisting of amino acids 1001 to 1619 associates with all three radiolabelled TRE transcripts but not with the antisense RNA, and that this interaction can be outcompeted by excess unlabeled TRE transcripts. Comparison of the ASH1 SET domain and an ASH1 fragment containing the SET domain (ash1ΔN; amino acids 1001 to 2218), with the amino terminal fragment ASH1N (amino acids 1 to 1001) and the carboxy terminal ASH1C (amino acids 1619 to 2218), both of which lack the SET domain led the authors to conclude that the SET domain of ASH1 binds TRE transcripts *in vitro*, not the two terminal fragments. It should be noted however, that according to the authors own data, both terminal fragments did retain all three TRE transcipts, albeit at a much lower level than the SET domain alone; and that the authors did not use intact ASH1 as a control. Thus it is possible that the N-terminal AT hooks and the C-terminal PHD finger contribute to intact ASH1 binding or help to confer specificity, especially given that the point mutation $ash1^{H1873W}$ of the middle conserved histidine residue in ASH1's PHD finger results in a recessive lethal antimorph with an *E(var)* effect indistinguishable from a null allele. PHD fingers have been shown to bind to histone H3 tri-methylated on lysine 4 (H3K4me3), a modification associated with the transcription start site of active genes (Shi et al. 2006, Martin et al. 2006, Wysocka et al. 2006). Thus while *ash1*'s SET domain

may be sufficient for TRE binding *in vitro*, an intact chromatin binding PHD finger is necessary for ASH1 function *in vivo*.

Cross-linked chromatin immunoprecipitation and native chromatin immunoprecipitation (NChIP) comparison of the interaction between ASH1 and the three TREs in mock- and RNase-treated chromatin showed that ASH1 antibodies precipitated all three TREs from mock treated and RNase III but not RNase H or RNase A treated chromatin. Therefore, while association of ASH1 with the *Ubx* TREs is RNA-dependent, dsRNA does not contribute to this interaction. The fact that RNase A and RNase H both disrupt interaction of ASH1 with the *Ubx* TREs indicates that both ssRNA and RNA-DNA hybrids are crucial in ASH1 recruitment to *Ubx* TREs. Given that dsRNA TRE transcripts, dsDNA TREs and DNA-RNA hybrids of TREs and their respective transcripts failed to disrupt the interaction between the *ash1* SET domain and TREs suggesting that ASH1 associates with ssRNA TRE transcripts, the disruption of this association by RNase H suggests that the TRE transcripts are retained at *Ubx* via hybridization with the template DNA. TRE transcripts were detected in cDNA to RNA from 3rd leg imaginal disc chromatin but not the soluble nuclear fraction and ASH1 bound TRE transcripts in the former but not the latter indicating TRE transcript-chromatin association in the cell. RT-PCR of RNA from cross-linked chromatin immunoprecipitation detected chromatin-associated TRE transcripts in both wild-type and $ash1^{22}$ mutant third-leg discs indicating TRE transcript retention at *Ubx* is independent of ASH1. RNase A treatment of chromosome squashes attenuated the association of ASH1 with the majority of the 150 target loci normally detectable (Tripoulas *et al*., 1996) compared to mock-treated chromosomes, indicating that ssRNA is necessary for the recruitment of ASH1 to chromatin targets in general (Sanchez-Elsner 2006). The findings of this lab are somewhat equivocal in that while H3K4 methylation is an activating mark, H4K20 and H3K9 trimethylation in particular are usually repressing marks. It is difficult to

reconcile H3K9 methylation, which normally recruits and is bound by HP1, with the exclusion of HP1 found by these authors. Their argument that ASH1 makes all three marks at once, and that it is a balancing of opposing marks does not agree with the results of Byrd and Shearn (2003) who found that *ash1* was not responsible for either H3K9 or H4K20 methylation *in vivo*. Furthermore, while both groups showed that ASH1's SET domain was sufficient to methylate histones, a certain promiscuity in target methylation is expected by the use of that domain alone without ASH1's DNA and chromatin binding domains. The concept that these DNA and chromatin binding domains are necessary in vivo is further supported by the lethal and E(var) phenotypes of the mutant $ash1^{H1873W}$ which is a point mutation in the middle conserved histidine residue in ASH1's PHD finger. This histidine, together with 2 conserved cysteine residues, coordinates 2 Zn^{2+} ions. The author's proposed targets do not parsimoniously explain the antagonizing of PcG function, and contradict the finding that ASH1L, the mammalian homolog of ASH1 counters PcG function by methylating H3K36, which blocks PcRC2 methylation of H3K37. It is also difficult to reconcile *ash1* H3-K9, and H4-K20 methylation with the finding of Papp and Müller (2006) that trimethylation at H3-K27, H3-K9, and H4-K20 at the *Ubx* promoter and coding region correlates strongly with PRC2 mediated transcriptional repression, and that cross-linked chromatin immunoprecipitation only detected bound ASH1 1 kb downstream of the *bx* transcription start site in transcriptionally active chromatin. ASH1 binding corresponded with a lack of H3-K27, H3-K9, and H4-K20 trimethylated nucleosomes at the promoter and coding region and and the presence of mono and di –methylated nucleosomes instead.

ASH1L, the mammalian homolog of *Drosophila melanogaster* ASH1

ASH1L is a large, multi-domain protein containing four AT hooks, a bromodomain, a bromo- adjacent homology domain (BAH), a PHD finger, a SET domain, and MYND ligand domains, all motifs implicated in

chromatin remodeling (Nakamura *et al*., 2000). Gregory *et al*. (2007) found that ASH1L was associated with the transcribed region of all active genes examined, including Hox genes. They found that the distribution of ASH1L in transcribed chromatin correlated specifically with that of methylated H3K4 but not H3K9 or H4K20, and that prior methylation of H3K9 reduced ASH1L-mediated methylation at H3K4, suggesting cross-regulation of these two marks. ASH1L 's recruitment upon transcription induction correlated with the recruitment of Pol II and H3K4me3. Surprisingly, ASH1L occupancy at the 5' end of an active gene persisted following a transcription elongation block despite diminished H3K4 trimethylation, suggesting that ASH1L alone was insufficient for H3K4 methylation. The SET domain of ASH1L methylated only H3K4 in oligonucleosomes and on synthetic peptides, but had reduced activity on recombinant histone H3 and core histones, *in vitro*. Mutating lysine 4 to arginine in GST fusion proteins containing residues 1 to 46 of histone H3 abolished methyltransferase activity indicating that ASH1L methylates only H3K4. The pattern of ASH1L distribution at housekeeping genes appeared different from at Hox genes; ASH1L distribution was restricted to the 5' portion of *PABPC1* but extended over the entire transcribed region of *HOXA10*. The *trx* homologue MLL1 localized to the transcribed region of *HOXA10* as well, but with a higher correlation to H3K4me3 distribution than that of ASH1L. siRNA knockdown of ASH1L reduced H3K4 me3 by 50% at HoxA10 *in vivo*. Most importantly, both MLL1 and ASH1L associated to *HOXA10* regions containing H3K4 di- and trimethylation suggesting that they are involved in maintenance of these modifications at this locus. The authors further demonstrated that ASH1L's association with chromatin was independent of MLL1 with ASH1L being found at normal levels at promoter and 5' transcribed regions *Mll1*-/- mouse embryonic fibroblasts. Despite the persistence of ASH1L at the *HoxA9* gene in *Mll1*-/- mouse embryonic fibroblasts, there was a dramatic loss of H3K4

trimethylation which the authors attributed to loss of Mll1 HMTase activity. Given that the authors had previously shown ASH1L H3K4 methyltransferase activity, particularly at HOX genes, the inability of ASH1L to compensate for the loss of MLL1 is perplexing.

Klymenko and Muller (2004) demonstrated that removal of PcG complexes in *ash1* and *trx* mutant backgrounds results in transcriptional reactivation suggesting that *trx*G proteins act as *Pc*G antagonists rather than transcriptional activators. Tanaka *et al*. (2007) used recombinant ASH1H SET domain and recombinant core histones and *in vitro* reconstituted nucleosomes to demonstrate that nucleosomes are the preferred targets of mammalian ASH1 and that mutating the conserved histidine (H2113K) of the ASH1 SET domain abolished this activity. They then performed histone methyltransferase assays a series of mutants with lysine to arginine substitutions at one of histone 3-K4, K9, K27, K37, or K79 or histone 4-K20 (targets of SET domain proteins), and mixtures of DNA and core histones. ASH1 did not methylate histone H4, and methylation of histone H3 by mammalian ASH1 was completely abolished for the K36R mutant but not affected by the other H3 mutants. They then constructed an amino terminal deletion, Flag-tagged eukaryotic expression vector for ASH1 containing the bromodomain, bromo-adjacent homology domain and PHD finger as well as the SET domain, motifs they claimed are sufficient for specific target gene trans-activation. *In vitro* histone methyltransferase assay using purified F-ASH1ΔN protein and either wildtype histone H3 or K36R showed that only HeLa cells transfected with F-ASH1ΔN expression vectors contained a histone H3 methyltransferase activity and that substitution of K36R abolished this activity. They therefore concluded that mammalian ASH1 is specific for histone H3 lysine 36. Importantly, they repeated their histone methyltransferase assays using *Drosophila* ASH1 SET domain and the same series of H3-K4R, K9R, K27R, K37R, K79R and H4-K20R mutants. Only substitution of histone H3

K36 with arginine abolished the methyltransferase activity of *Drosophila* ASH1 but not mutations in H3 K4, H3K9, or H4K20 as previously reported (Beisel et al., 2002; Byrd and Shearn, 2003). Tanaka et al. (2007) thus concluded that both mammalian and *Drosophila* ASH1 are histone H3 K36 specific methyltransferases.

Yuan et al. (2011) also demonstrated that ASH1H is a H3K36 dimethylase, and that H3K36me2 inhibited the spread of H3K27 di- and tri-methylation by PRC2. Recombinant human Ash1 SET domain was incubated with radioactive *S*-[*methyl*-^{3}H]adenosylmethionine and wild-type nucleosomes or nucleosomes with alanine substitutions at histone H3 Lys-4, Lys-9, Lys-27, or Lys-36. Only H3 K36A abolished the methyltransferase activity of ASH1, indicating that ASH1 is specific for H3K36. They performed immunoprecipitation experiments with antibodies specific for H3K27me3 or H3K36me3 to demonstrate that H3K27me3 and H3K36me3 rarely co-exist on the same histone and that they may antagonize each other. They then used recombinant human NSD2, a H3K36-specific dimethylase (Li et al., 2009) to pre-install H3K36 methylation onto nucleosomes which were subsequently incubated with reconstituted PRC2 complex. Nucleosomes pretreated with NSD2 without S-adenosylmethionine were still methylated by PRC2 whereas those pretreated with NSD2 plus *S*-adenosylmethionine were not, demonstrating that PRC2 activity is inhibited by pre-existing H3K36 methylation. Substitution of alanine for lysine at H3K36 had the same effect indicating that contact with H3K36 or its neighbours is important for PRC2 activity. Therefore, methylating or mutating H3K36 most likely impairs the enzymatic activity of PRC2 by restricting its binding. Furthermore, the *trx*G protein CBP associates with ASH1 and acetylates H3K27 blocking its methylation by E(z). Taken together these results suggest that one of the modes of action of ASH1 is as a PcG antagonist as well as a recruiter of TRX, which is itself a H3K4 histone methyltransferase.

ASH1 acts at the ci locus

Pci was isolated as an enhancer trap of *ci* and is an allele of *ci*. While *Pci* (and the *E1 gypsy* element) are inserted in the *ci* distal regulatory region, both ci^{57g}, a deletion upstream of *Pci* , and ci^{1}, a *gypsy* insert upstream of *Pci*, exhibit *ci* phenotypes. Expression of the *Pci* enhancer-trap reporter in imaginal discs accurately mimicked that of ci^{D} RNA with both being expressed specifically in anterior compartment cells (Eaton & Kornberg 1990). All alleles isolated of ash1 are *En(PDS)* with *E1*, *E1/Pci* and act at *Pci* to lower *w* expression without *E1*. The loss of ASH1 function results in increased silencing of the *white* reporter gene which should be mimicking *ci* expression. Since these mutants do not affect *P{lacW}3-76a* , this effect is not construct dependent and is not due to a direct interaction with the *white* promoter, but with the *ci* regulatory region itself. Since *Pci* reporter expression is halved when ASH1 dose is halved, and does not depend on the presence of *E1*, I infer that ASH1 normally acts at the *ci* regulatory region, likely in a dose dependant manner. Since *Polycomb* Response Elements (PREs) and *trithorax* Response Elements (TREs) share similar components and distribution (Tillib *et al*.1999; reviewed by Schuettengruber *et al*. 2007; Ringrose & Paro 2007), and *PcG* proteins have been shown to bind at ci (Schwartz *et al*. 2006; Tolhuis *et al*. 2006) it is likely that TRE(s) exist there as well for ASH1 to interact with. The fact that *trx* has been identified as one of the other complementation groups isolated in this screen, further supports the idea of a TRE being present at *ci*. Members of this *trx* complementation group showed a degree of synthetic lethality with those of the *ash1* complementation group (Supplemental Table A3). Genetic evidence has indicated that the activities of ASH1 and TRX are functionally related; for example, flies mutant for both ash1 and trithorax show enhanced penetrance of homeotic phenotypes due to reduced Hox gene expression (Tripoulas *et al*. 1996). Mutants in *ash1* and *trx* may exhibit synthetic lethality; ASH1 co-

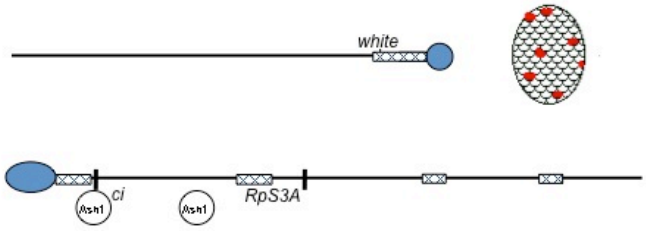
immunoprecipitates with TRX from embryonic nuclear extracts; and ASH1 and Trx co-localize at multiple sites on polytene chromosomes with TRX accumulation being reduced in an *ash1* mutant background. These results suggest a model in which ASH1 binds to H3 and methylates K4 residues via its SET domain; this then recruits TRX which recognizes the methylated H3-K4 residues, and explains the loss of TRX on polytene chromosomes from an *ash1* mutant, and the synthetic lethality.

ASH1 is a histone methyl-transferase (HMTase) with specificity that is currently under debate as discussed above. Methylation of histones in the promoter of ASH1 target genes in Drosophila correlates with their transcriptional activation and subsequent hyperacetylation, and is thought to serve as a binding surface for a chromatin remodelling complex containing the epigenetic activator Brahma (Brm). ChIP analysis of Ultrabithorax transcription in Drosophila indicates that transcriptional activation, trivalent methylation by ASH1 and recruitment of BRM coincide (Beisel *et al*. 2002). We therefore tested a null allele of *brahma*, brm^{2}, against *E1* to see if it enhanced variegation and for complementation against all third chromosome recessive lethal mutants recovered in this screen. Not only did it fail to have any visible effect on *E1* variegation, but it complemented all 29 third chromosome recessive lethal mutants recovered in this screen (data not shown). We take these data as a strong indication that *brahma* is not involved in the epigenetic regulation of *ci*.

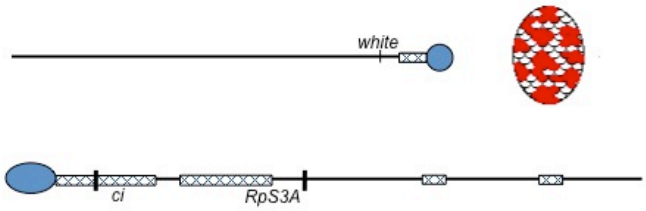
ASH1 protein localizes at over 150 sites on polytene chromosomes of larval salivary glands implying that it functions to maintain the expression pattern of multiple genes (Tripoulas *et al*. 1996). This action of ASH1 at many sites explains our mutants' Su(var) effect on w^{m4}. This opposite effect on w^{m4} is consistent with a model whereby loss of ash1 results in the titration of heterochromatinizing factors away from pericentric heterochromatin. Loss of ASH1 at *ci* would not cause enough ectopic

heterochromatin formation to titre any measureable amount of Su(var) activity away from pericentric heterochromatin, but if this ectopic heterochromatin formation was to occur at many sites the loss of heterochromatin forming factors would add up to a measurable amount as seen at w^{m4} in this study (Figure 5.1).

Figure 5.1 A model for *Ash1*' s effect at w^{m4}.



ASH1 "marks" transcriptionally active domains such as *ci*.
This prevents heterochromatin from spreading into *ci* and other transcriptionally active domains.
This normally allows for the stochastic spread of pericentric heterochromatin across *white*, resulting in a variegating phenotype.



If ASH1 function is lost or altered, heterochromatin spreads into the *ci* region and ~100 other possible sites. This creates a "sink" for heterochromatic proteins and leaves less heterochromatin at w^{m4}.
This results in increased expression of *white*, and a *Su(var)* phenotype.

The fact that we isolated *ash1*^{H1873W} as a recessive lethal member of the *ash1* complementation group indicates that ASH1's PHD finger is necessary for ASH1 function. The dose responsive loss of reporter gene activity in a *ash1*^{H1873W} heterozygote indicates a failure of *ash1*^{H1873W} mutant ASH1 to act at *ci*; most likely due to ASH1 being unable to localize or bind DNA properly. As discussed above, PHD fingers have been shown to bind to histone H3 tri-methylated on lysine 4 (H3K4me3) (Shi et al. 2006, Martin et al. 2006, Wysocka et al. 2006). The modification H3K4me3 is associated with the transcription start site of active genes, and therefore may be how part of how ASH1 recognizes a transcriptionally active gene sequence required to be "locked on" by *trxG* proteins. The mammalian homologue of *Drosophila* t*rithorax* protein, ALL-1/MLL, has been shown to selectively methylate K4 in histone H3 (Milne et al, 2002; Nakamura et al, 2002). Thus TRX and ASH1 may act cooperatively in establishing epigenetic activation with ASH1 blocking PcG repression and initially recruiting TRX whose subsequent methylation of H3K4 allows for enhanced ASH1 chromatin binding.

For all alleles tested, the pattern of loss of *w* expression follows progression of the morphogenetic furrow in *E1/+* and *E1/Pci* flies with ommatidia specified first having the least pigment. This suggests that a lower dose of *ash1* results in a later locking of *Pci* into a transcriptionally active state.

trx

I have shown that the 3 *en*(*PDS*) in one of the complementation groups described in this study are alleles of *trx*. The enhancement and lethal phenotypes mapped to the *trx* locus and could not be separated by recombination. All 3 alleles failed to complement all deficiencies tested that uncover *trx* and failed to complement both extant alleles *trx*^{1} and *trx*^{E2}; therefore I concluded they were alleles of *trx*. Sequence analysis as

described above confirmed that each *trx* mutant had a change resulting in the introduction of a premature stop codon. One, trx^{S1583*}, had a thirteen nucleotide deletion that resulted in a +1 frameshift which introduced multiple premature stop codons that predicted termination before TRX's most amino terminal RING domain/PHD finger and its carboxy-terminal SET domain while $trx^{S2582*a}$, and $trx^{S2582*b}$ both had a serine to amber mutation that predicted termination prior to TRX's carboxy-terminal SET domain.

trx, the founding and eponymous member of the TrxG, was first isolated and characterized as a positive regulator of HOX genes that did not itself initiate HOX gene expression, but maintained it in the appropriate body segments as development progressed (Ingham and Whittle 1980; Ingham 1998). The first definite allele of *trithorax* (*trx*) discovered was a spontaneous mutation that gave a partial transformation of halteres into wings, anterior-ward leg transformations, and rarely, a pair of supernumerary wing-like structures on the thoracic segment anterior to normal. Together with the haltere transformation, this resulted in three pairs of wings, hence the name "*trithorax*" (Ingham and Whittle 1980). As Ingham (1998) later noted, Astauroff (1930) had earlier described the mutation *tetraptera*, which mapped to a similar chromosomal location and had a similar phenotype and and thus was likely to have been the first *trx* allele. Unfortunately, it had since been lost. trx^{1} was originally characterized as a gene that caused homeotic transformations when mutated (Ingham and Whittle 1980; Breen and Harte 1991). Its phenotypes of transformation of haltere-to-wing, and of posterior abdominal segments towards more anterior segments, resembles loss-of-function mutations in the *Bithorax* complex, one of the two *Drosophila melanogaster* HOX gene clusters, thus implicating *trx* as a positive regulator of HOX gene function. Subsequent genetic and molecular analyses showed that TRX is required to maintain target genes in

transcriptionally active states throughout development by counteracting the repressive effects of Polycomb group proteins (Ringrose and Paro 2004; Schuettengruber *et al*. 2007). Thus, as for PcG genes, *trx* was postulated to be involved in the epigenetic inheritance of the expression states of HOX genes (Ingham, 1998; Cavalli & Paro 1999). The subsequent identification of other TrxG proteins such as *ash1* was based on genetic screens for second site mutations with phenotypes indicating a loss of HOX genes function, or for suppressors of PcG-dependent mutant phenotypes (Kennison & Tamkun 1988), which has led to the idea of TrxG proteins being antagonists of PcG-dependent gene silencing.

Cloning and molecular analysis of *trx* identified a common motif shared with other epigenetic regulators, the SET domain, named after the first three founding members; ***S****u(var)3-9*, ***E****nhancer of zeste*, and ***t****rithorax*. SET domain containing proteins are capable of methylating and binding to histones in chromatin, in fact, all known lysine specific HMTases contain a SET domain, and its presence is associated with this activity (Jenuwein and Allis 2001). Czermin *et al*., (2002) demonstrated that embryonic nuclear extract TRX immunoprecipitates contained a H3K4 specific methyltransferase activity attributable to TRX or an associated protein. Smith *et al*. (2004) demonstrated that the *Drosophila* TRX SET domain could methylate histone H3 in the 5' coding region of *hsp70* after induction *in vitro*, and Edman degradation of TRX methylated, tritium-labelled histone H3 showed that Lys 4 (H3-K4) was the only methylated residue detectable. A hypomorphic allele of *trx*, trx^{Z11}, that contains a glycine-to-serine substitution in the SET domain that interferes with histone binding (Katsani *et al*. 2001) also showed little or no histone H3 methylation activity. Methylation of histone H3 on lysine 4 (H3K4) is generally associated with transcriptionally active regions of chromatin (Eissenberg & Shilatifard 2010) Thus, the genetic and biochemical characterization of

TRX indicate it maintains target gene activity partly through the methylation of H3K4.

Antibodies directed against TRX bound to 63 specific sites on salivary gland polytene chromosomes including the sites of its known targets, the HOX gene clusters at the Bithorax and Antennapedia complexes, despite the transcriptionally repressed state of these loci in the salivary gland (Chinwalla *et al*. 1995). Given that 32 of their trx binding site assignments coincided with known binding sites for PcG proteins, the authors used dual fluorescence confocal microscopy to simultaneously localize the *trx* and *Pc* proteins on polytene chromosomes. Co-localization of TRX and PC at 30 sites indicates that many of their chromosomal binding sites coincide and that interactions between them may be a significant feature of their mode of action. However, recent chromatin immunprecipitation studies by Schuettengruber *et al*. (2009) suggest a complicated relationship between TRX protein binding sites and H3K4me3. TRX contains a site similar to human MLL cleavage site 2 (QMD/GVDD vs QLD/GVDD) in an analogous position in the protein to the human homolog's cleavage sites (discussed below). *Drosophila* TASPASE1 cleaves wild-type TRX, but not TRX with a QMD/GVDD to QMAAVDD mutation in the cleavage site (Caposti *et al*., 2007). The importance of this cleavage for TRX's biological function is indicated by the fact that a 271-amino-acid deletion (trx^{E3}) that spans this cleavage site displays defective antennapedia complex gene expression. Interestingly, bithorax complex gene expression is not affected, suggesting TRX cleavage plays a selective role in its function (Breen 1999; Mazo *et al*., 1990; Sedkov *et al*., 1994). Caposti *et al*. (2007) used chromatin immunoprecipitation (ChIP) on chip assays to map the chromosomal distribution of the N- and C-terminal fragments of the Trithorax (TRX) protein as well as histone modifications associated with TRX-mediated activation. They found that the SET containing carboxy-terminal fragment of TRX (TRX-C) showed high affinity to PcG binding sites and limited

overlap with sites of H3K4me3, whereas the non-SET containing amino-terminal fragment (TRX-N) bound mainly to active promoter regions trimethylated on H3K4 and was found at almost every H3K4me3 enriched site. These distinct distributions of the N- and C-terminal domains of TRX are consistent with its proteolytic cleavage in a manner analogous to that seen with MLL (Hsieh *et al*. 2003). Proteolytic cleavage has been demonstrated for TRX and one of the two MLL cleavage sites is conserved in the TRX protein sequence (Capotosti *et al*. 2007).

The mammalian *trx* gene family

The mammalian genome encodes genes for seven *trx*-related proteins, all of which are found in complexes that can methylate H3K4 (reviewed in Eissenberg & Shilatifard 2010). Two of these, Set1A and Set1B, are more similar to *Drosophila melanogaster* Set1 and *Saccharomyces cerevisiae* Set1 in that they are slightly less than half the size of *trx* and are missing several functional domains including PHD fingers, and the FYRN and FYRC domains necessary for self-association, while retaining the RRM, SET and Post SET domains. The actual mammalian homologs of *trx* are the Mixed Lineage Leukemia (MLL) family: MLL, MLL2, MLL3, MLL4 and MLL5. Given that there is only one H3K4 methylase, Set1, in yeast, it is not readily apparent why mammals have six, functionally non-redundant H3K4 HMTases: Set1A and Set1B.

The mammalian homolog most similar to *trx*, Mll (also designated ALL-1 and HRX) was discovered because chromosome rearrangements with their breakpoints within the MLL gene are associated with several acute myeloid and lymphoid leukemias (reviewed in Eissenberg & Shilatifard 2010). MLL is structurally homologous to Drosophila TRX; however it also contains three AT-hook domains that can bind to the minor groove of adenine-thymine (AT) rich DNA (Reeves & Nissen, 1990; Aravind & Landsman, 1998), and a CxxC DNA methyltransferase homology domain missing from TRX. Pradhan *et al*. (2008) demonstrated that the CXXC

region of DNMT1 specifically bound to unmethylated CpG dinucleotides, and that DNA binding was abolished by mutation of the conserved cysteines. A deletion of the CXXC resulted in a significant reduction in catalytic activity, confirming that this domain cooperates with the catalytic domain, and is crucial for DNA methyltransferase activity (Pradhan *et al*., 2008). The CXXC domain of MLL derivatives is essential for oncogenesis; Birke *et al*. (2002) demonstrated that the CXXC domain of MLL also binds unmethylated CpG sequences, and that this binding was hindered by the introduction of 5 methyl cytosine. Given that methylation of CpG islands is rare in Drosophila but a major epigenetic mark in mammals, it is not unreasonable for these domains' absence from TRX.

Yu *et al*. (1995) demonstrated that in addition to roles in hematopoiesis and growth, MLL is necessary for axial segment identity in mice. *Mll+/-* heterozygous knockout mice diplayed bidirectional skeletal transformations, including anterior transformations of C7 to C6 and T3 to T2, and posterior transformations of T13 to L1 and L6 to S1. These mimicked the effects of loss-of-function alleles for multiple HOX genes. Although rare in mice, the authors noted that *Drosophila trx* mutants show bidirectional transformation of both pro- and metathorax towards a mesothorax identity (Ingham & Whittle 1980). The anterior boundaries of the initial axial patterns of representative HOX genes Hoxa-7 and Hoxa-9 were shifted caudally in *Mll+/-* heterozygous embryos and abolished in *Mll-/-* homozygous embryos with concommittent lethality by E10.5. These findings argue for a role for Mll in the maintenance of HOX gene expression analogous to *trx* in *Drosophila* (Yu *et al*. 1998).

It is interesting to note that although complete loss of Mll results in embryonic lethality, mice homozygous for a deletion of Mll's SET domain (Mll1ΔSET) are viable and fertile (Terranova *et al*. 2006), indicating that H3K4 methylation by MLL is not essential. Mll1ΔSET exhibited significant

decreases of Hoxc8, Hoxd4, Hoxa7, and Hoxa5 mRNA level as compared with WT embryos but not Hoxb9 and Hoxd11, and presented skeletal defects resembling HOX gene knockouts. Chromatin immuno-precipitation analysis revealed a significant (~10-fold) decrease in monomethylated H3K4, and a 20 – 40% reduction in dimethylated H3K4, but no change in trimethylated H3K4 at both Hoxd4 and Hoxc8 promoters in the trunks of ΔSET mutant embryos . CpG rich regions of the *Hoxd4* gene were found to be abnormally methylated by bisulfite sequencing, however, a lymphocyte-specific gene, CD3 did not show any change in DNA methylation, nor was there any change in global levels of DNA methylation as determined by immunofluorescence staining using an antibody directed against 5-methyl cytosine. These results suggest that Mll acts at the chromatin level to maintain the expression of select target Hox genes during embryonic development. They also indicate a relationship between histone methylation and DNA methylation that is Mll SET domain dependent (Terranova *et al*. 2006).

Guenther *et al*. (2005) analysed MLL distribution in human monocyte and lymphoblast cell lines by performing chromatin immunoprecipitation using a human proximal promoter and noncoding sequence DNA microarray. They found MLL at ~ 38% of promoters on the array, correlating with 90% of sites occupied by RNA polymerase II and 92% of sites containing trimethylated H3K4. They then used tiling arrays for a 276 gene subset of HOX and non-HOX genes to show that both MLL and trimethylated H3K4 were concentrated at transcription start sites and 5' regions of highly expressed genes, similar to yeast Set1. Interestingly, the distribution of MLL and trimethyl H3K4 extended across a large region of the late HoxA cluster encompassing HoxA1 and the 5' HoxA subcluster including HoxA7, HoxA9, HoxA10, HoxA11, and HoxA13. This is consistent with its high expression in monocytic (U937) cells, and suggests not only that MLL1 is responsible for maintaining a large domain

of active chromatin within the HoxA region, but that MLL acts differently at HOX than at non-HOX loci. In contrast to these findings, Wang *et al*. (2009) demonstrated that MLL is responsible for H3K4 methylation of less than 2% of the genes in mouse embryo fibroblast.

It is worth noting that interpretation of the genetic and chromatin immunoprecipitation data regarding MLL is also complicated by the fact that, like *trx*, MLL is proteolytically cleaved *in vivo* by TASPASE 1 to produce amino- and carboxy-terminal fragments that are capable of non-covalently associating (Hsieh *et al*. 2003). As already noted for the chromatin immunoprecipitation data using antibodies directed against N- and C terminal fragments of *Drosophila* TRX (Schuettengruber *et al*. 2009), the distributions of the MLL N- and C-terminal fragments in vivo may be distinct, as could be the functions of complexes containing either fragment versus those containing both.

Liu *et al*. (2010) discovered a novel function for MLL as an effector in the mammalian S-phase DNA damage checkpoint response; they further postulated that ensuing checkpoint dysfunction contributed to the pathogenesis of MLL leukaemias. They found that MLL accumulated in S phase in all cell types examined due to DNA damage caused by various agents. In the progression of a normal cell cycle, the SCF^{skp2} proteasome recognizes and degrades MLL in S phase: however, following DNA damage, MLL was phosphorylated at serine 516 by ATR, which disrupted its interaction with SCF^{Skp2} E3 ligase, and led to its accumulation. They further showed that stabilized MLL protein accumulated on chromatin, with concomitant methylation of histone H3 lysine 4 at late replication origins. Most importantly, MLL accumulation at late replication origins inhibited CDC45 loading and thus origin firing and DNA replication. Cells lacking MLL had a radioresistant DNA synthesis (RDS) phenotype and chromatid-type genomic abnormalities, which are hallmarks of an S-phase

checkpoint defect. S-phase checkpoint defects in *Mll*$^{-/-}$ mouse embryonic fibroblasts were rescued by knock in of wild-type MLL, but not by alleles with either a serine to alanine substitution of the ATR phosphorylation site (S516A) or minus the SET (ΔSET) domain. Thus both phosphorylation of MLL and its ability to methylate H3K4 are necessary for MLL inhibition of late origin replication initiation after DNA damage. Further phosphorylation of MLL may be necessary for certain aspects of it function; however, www.phosphosite.org gives ~ 200 potential phosphorylation sites for MLL which makes analysis difficult. This is undoubtedly due to mutant/altered MLL's role in oncogenesis, as most of these were discovered in cancerous cell lines.

trx and *ash1* act together at the *ci* locus

I have shown that both *ash1* and *trx* act dose dependently at the *ci* locus and that that this interaction is cooperative (Figures 2.13, 3.3). This is consistent with ASH1 acting to recruit TRX to PRE/TREs at *ci*; if TRX acts to "lock" *ci* into a transcriptionally active state, then halving the dose of *trx* should halve the amount of TRX protein available to regulate the *ci* region. Similarly, if ASH1 acts to recruit TRX to *ci*, then halving the dose of *ash1* should halve the amount of ASH1 protein available to recruit TRX, resulting again in half of the amount of TRX protein recruited to regulate the *ci* region. If both *ash1* and *trx* doses are halved, then half the amount of ASH1 protein with half the amount of TRX to recruit should result in one quarter of the normal amount of TRX protein recruited to regulate the *ci* region. This should result in one quarter of the amount of *white* reporter being produced, which is what we see (Figures 2.13, 3.3). This model relies on there being a PRE/TRE at *ci* to recruit ASH1. Two groups using different techniques have provided evidence that this is indeed the case. Schwartz *et al*. (2006) characterized the genomic distribution of several PRC1 and PRC2 proteins using chromatin immunoprecipitation (ChIP) followed by the analysis of immunoprecipitated DNA using a high-density

genomic tiling microarray. Tolhuis *et al*. (2006) fused several PRC1 and PRC2 proteins to a DNA methyltransferase, then transiently transfected cultured cells. The DAM- Pc/Sce/esc fusion protein thus gave a higher level of DNA methylation at sites it bound to compared to control samples where DAM was expressed alone. Methylated DNA fragments were subsequently immunoprecipitated and quantified via microarray. Both groups found PRC1 and PRC2 binding at *ci*, evidence of a PRE/TRE.

It has already been established that ASH1 and TRX act together in multimeric protein complexes. Kuzin *et al*. (1994) observed that association of TRX with polytene chromosomes is *ash1* dependent indicating a possible physical interaction between the two proteins. Rozovskaia *et al*. (1999) found that TRX and ASH1 proteins colocalize on salivary gland polytene chromosomes, coimmunoprecipitate from embryonic extracts and bind *in vivo* to *bxd*, which contains several TRX response elements. In a study similar to this thesis, $ash1^{22}$ heterozygous mutant flies containing a mini-white reporter transgene inserted near *bxd* showed strong *white* reduction relative to wild type, similar to results obtained for trx^{B11} null allele using the same transgenic reporter lines (Tillib *et al.* 1999), suggesting that the activity of this TRE-containing region is both ASH1 and TRX dependent. Finally, using yeast two-hybrid assays they found that both TRX and ASH1 SET domains can self-associate and that this self-association is prevented by mutations within either SET domain. TRX and ASH1 SET domains interacted strongly with each other in yeast, and GST-linked ASH1 SET bound ~ 20-fold more radiolabeled TRX SET than did GST resin alone. Labeled TRX SET coimmunoprecipitated with unlabeled epitope-tagged T7-ASH1 SET but not with unrelated T7-tagged proteins. In addition, T7-tagged ASH1 SET and HA-tagged TRX SET co-immunoprecipitated from COS cells transiently co-transfected with plasmids encoding the two epitopes. Mutation of conserved residues within either SET domain prevented their

interaction in yeast, but alterations of non-conserved residues within the SET domain or immediately upstream of it did not. These results suggest that the association previously seen in embryos between TRX and ASH1 is direct and involves their conserved SET domains. Collectively, these results suggest that TRX and ASH1 interact either within trxG protein complexes or between complexes in close proximity on *bxd* to maintain Ubx transcription. My results suggest a similar interaction is taking place at *ci*; however, I will leave to future researchers to delineate the binding elements (TREs). The 6XMyc tagged *ash1* transgene I have constructed should facilitate immunoprecipitation of the chromatin region in question. Further to that, replacement of the PHD finger domain in this construct with the same domain from H1873W should clarify whether or not the PHD finger is necessary for ash1 binding to chromatin, or confers some specificity. Given that ASH1 binding to actively transcribed TREs recruits TRX, which then methylates H3K4, the epigenetic mark that PHD fingers have been shown to bind to; it is possible that this circular interaction serves to strengthen the binding of an ASH1/TRX complex to transcriptionally activated regions.

trx and *ash1* act as Su(var)s of w^{m4}.

I have shown that mutations of both *ash1* and *trx* act as dominant Su(var)s of w^{m4} (Figures 2.12, 3.4). This directly contradicts earlier claims by Buchner *et. al*. (2000), who found no effect on w^{m4} variegation by loss of *trx,* and Fanti *et. al.* (2008), who found no dominant effects on w^{m4} variegation by heterozygous mutants of *ash1*, *trx*, and *CBP*. It should be noted however, that neither of these authours performed or presented any quantitative measurements such as pigment analysis to substantiate their observations. Fanti *et. al.* (2008) went on to show that the previously mentioned *trx*-G genes are recessive enhancers of Y chromosome position-effect variegation involved in control of heterochromatic genes *light* and *rolled* expression. They used a Y chromosome rearrangement

carrying an *Hsp70-lacZ*-inducible transgene and mini-white wild-type allele inserted in Y centromeric heterochromatin. This rearrangement has been reported to variegate for both *white* and the inducible *lac-Z* genes in male embryos, salivary glands, and imaginal disks (Lu et al. 1996), thus allowing the analysis of the effect of recessive lethal mutations on variegation. The authours showed that while both *ash1* and *trx* homozygous mutants enhanced *Hsp70-lacZ* variegation, *CBP*, *Trl* and *ISWI* mutants had no effect. Curiously, even though this construct contains the *white* mini gene, which does variegate, they only claim to have tested the effect of a *Trl* mutation on this *white* mini gene's variegation. In contrast to previous work showing *trl* mutations to be strong enhancers of w^{m4} variegation (Farkas et al. 1994), Fanti *et. al.* (2008) found *trl* did not dominantly affect mini-*white* variegation in adults. It is worth noting; however, that after heat stress, TAC1 is recruited to the 5′-coding region of *hsp70* by the elongating Pol II complex, where it is required for high levels of gene expression. TAC1 contains both Trithorax (Trx) and CREB-binding protein (CBP) and is required for methylation and acetylation of nucleosomal histones in the 5′-coding region of *hsp70* after induction (Smith et. al. 2004). This suggests that the enhancement of *Hsp70-lacZ* variegation by mutations in *trx* and *CBP* is a direct effect due to the loss of TAC1 from the *Hsp70-lacZ* promoter itself, rather than any effect on heterochromatin formation or heterochromatic gene regulation.

CG8878

A complementation group of 7 alleles on the second chromosome mapped to a region in which *CG8878*, a putative Serine/Threonine/Tyrosine kinase, was the most logical candidate. The lethal phenotype could not be directly assigned to *CG8878* due to a lack of extant alleles for complementation analysis; therefore, all 7 alleles of this group were sequenced to look for mutations in the *CG8878* gene. Five

were shown to have mutations resulting in premature stop codons prior to the end of *CG8878*'s putative Serine/Threonine/Tyrosine kinase domains; therefore we concluded they were alleles of *CG8878* and were temporarily designated 1a27a', 3a22a, 3a52a, 3a66a, and 3a97a. 3a90a and 4a7a shared a common deletion in the 5' regulatory region that deleted 4 bp of the most proximal E box, and thus are possible regulatory mutants of *CG8878*. Both 1a27a and 3a52a represent null alleles of *CG8878*. 3a52 a had a G⟶ A transition at m675 that changed a Tryptophan to an opal stop codon at aa 123 while1a27a experienced a G⟶ A transition at 2R:8037095 that mutates a donor splice site resulting in a + 61 bp frameshift and the insertion of 32 aa followed by an opal stop codon at aa 180. Both of these should result in a protein truncated prior to both kinase domains; therefore, they represent true null alleles. Both 3a22a and 3a52a had a C⟶ T transition at m1942 that changed an Arginine to an opal stop codon at aa 546. This is in the center of *CG8878*'s carboxy-kinase domain and leads to a truncated protein missing half of one kinase domain. 3a66a had a 1nt deletion (Δ) of C at m1665 that mutated Isoleucine to an amber stop codon at aa 468, prior to *CG8878*'s carboxy-kinase domain which leads to a truncated protein missing one kinase domain. The En(var) phenotype of these three alleles implies that this domain is necessary for normal *CG8878* function at *ci*, while the recessive lethal phenotype argues that this domain is essential. The base changes found and the resulting protein/regulatory changes implied are summarized in Table 5.1.

Even though the entire *CG8878* transcription unit is contained in the antisense direction within the second intron of *Hen1* (formerly Pimet), we interpret the evidence to indicate that *CG8878* is the En(var)/En(PDS). The *prima facie* evidence is: 1) Hen1 is not an essential gene because *PBac(WH)Hen1[f00810]* is a null for *Hen1* (Horwich *et al*. 2007) but is not recessive lethal; 2) all seven mutants had sequence lesions in *CG8878*; and 3), all of these sequence changes are entirely inside *Hen1*'s second

intron, and therefore should have no effect on *Hen1* expression. Given that *Hen1* is involved in the processing of *Piwi* interacting RNAs which are involved in silencing of invading transposable elements in *Drosophila* (reviewed by Saito *et al*. 2006; Saito *et al*. 2007; Horwich *et al*. 2007), and that *P{lacW}ci*^{Dplac} has *P*-element terminal repeats and a *P*-element transposase *lacZ* fusion; if *P{lacW}ci*^{Dplac} enhancement was due to PIWI mediated transposable element silencing we would expect *P{lacW}3-76a* (which is the same construct as *P{lacW}ci*^{Dplac} but at a different location) to be similarly enhanced and *w*^{m4}, which is not *P* element derived, to be unaffected. Instead, *P{lacW}3-76a* expression is unaffected by our *En*(*var*)s while *w*^{m4} is significantly enhanced. The simplest explanation is that *CG8878* is an essential gene and when mutated has a dominant *En*(*var*) phenotype.

Table 5.1 Changes in *CG8878* mutants.

Temporary	Base Change	Change	Designation
1a27a	G→ A 8037095	N180*	*CG8878*^{N180*}
3a22a	C→ T m1942	R→opal aa 546	*CG8878*^{R546a*}
3a52a	G→ A m675	W→opal aa 123	*CG8878*^{W123*}
3a66a	1Δ C m1665	I→amber aa 468	*CG8878*^{I468*}
3a90a	19nt Δ 2R:8038834-8038816	regulatory	*CG8878*^{Rega}
3a97a	C→ T m1942	R→opal aa 546	*CG8878*^{R546b*}
4a7a	19nt Δ 2R:8038834-8038816	regulatory	*CG8878*^{Regb*}

Potential molecular function of *CG8878*?

CG8878 is highly conserved across Drosophila species with two VRK like domains indicating that it is likely to encode an essential kinase of unknown specificity. The closest *Drosophila melanogaster* homologue of *CG8878* is *nucleosomal histone kinase-1* (*nhk-1* or *ballchen*) with 41% identity (E= 3e-29), with regions of maximum similarity coinciding with

CG8878's putative kinase domains. NHK-1 has high affinity for chromatin and phosphorylates Threonine 119 at the carboxy terminus of nucleosomal, but not free, H2A in *Drosophila* embryos during mitosis but not in S phase. This coincides with NHK-1's chromatin association and may be a component of the histone code related to cell cycle progression (Aihara *et al*. 2004). A point mutation, Z3-0437, in the kinase domain of NHK-1 results in female sterility due to defects in the formation of the karyosome as well as a lack of Histone H4K5 and H3K14 acetylation in the karyosomes, implying that Histone H2A threonine 119 phosphorylation is required for meiotic acetylation of these residues Ivanovska *et al*. (2005). Lancaster *et al*. (2007) found that karyosome formation required NHK-1 phosphorylation of *barrier to autointegration factor* protein (BAF). Expression of nonphosphorylatable BAF or loss of NHK-1 resulted in ectopic chromosome-nuclear envelope association in oocytes. This led the authors to propose that NHK-1 mediated BAF phosphorylation disrupts tethering of chromosomes to the nuclear envelope allowing karyosome formation in oocytes. These findings raise the intriguing possibility of CG8878 being a Histone kinase which would readily explain its action as an *En*(*var*). As an example, JIL1 phosphorylation of H3S10 blocks methylation of H3K9 allowing hyperacetylation of Histone 3 and promoting a transcriptionally active chromatin state (Zhang *et al*. 2006). CG8878's expression profile is consistent with it being a genome wide inhibitor of heterochromatin spread as it is expressed in all tissues, at all stages of development, with maxima at times of peak developmental change such as early embryogenesis and prepupariation (Gelbart & Emmert 2011). I leave it to future researchers to determine *CG8878*'s target, and mode of action; the possibilities are certainly interesting.

References – Chapter 5

Aihara, H., Nakagawa,T., Yasui, K., Ohta, T., Hirose, S., Dhomae, N., Takio, K., Kaneko, M., Takeshima, Y., Muramatsu, M., and Ito, T. 2004. Nucleosomal histone kinase-1 phosphorylates H2A Thr 119 during mitosis in the early *Drosophila* embryo. *Genes Dev.* **18**: 877-888

Aravind, L., and Landsman, D. 1998. AT-hook motifs identified in a wide variety of DNA-binding proteins. *Nucleic Acids Res*. **26:**4413–4421.

Astauroff. 1930. Analyse der erblichen Störungsfälle der bilateralen Symmetrieim Zusammenhang mit der selbständigen Variabilität ähnlicher Strukturen. *Z. indukt. Abstamm. Vererb*. **55**: 183-262.

Beisel, C., Imhof, A., Greene, J., Kremmer, E. & Sauer, F. 2002. Histone methylation by the *Drosophila* epigenetic transcriptional regulator Ash1. *Nature* **419**: 857-862

Birke, M., Schreiner, S., Garcia-Cuellar, M.P., Mahr, K., Titgemeyer, F., and Slany, R.K. 2002. The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. *Nucleic Acids Res*. **30**: 958-965.

Breen, T.R., and Harte, P.J. 1991. Molecular characterization of the Trithorax gene, a positive regulator of homeotic gene expression in *Drosophila. Mech. Dev*. **35**: 113–127.

Breen, T.R. 1999. Mutant alleles of the *Drosophila trithorax* gene produce common and unusual homeotic and other developmental phenotypes. *Genetics* **152**: 319-344.

Buchner, K., Roth, P., Schotta, G., Krauss, V., Saumweber, H., Reuter, G., and Dorn, R. 2000. Genetic and molecular complexity of the position effect variegation modifier *mod*(*mdg4*) in *Drosophila*. *Genetics* **155**: 141–157.

Byrd, N.B., and Shearn, A. 2003. ASH1, a *Drosophila* trithorax group protein, is required for methylation of lysine 4 residues on histone H3. *Proc. Natl. Acad. Sci. USA* **10**: 11535–11540

Capotosti, F., Hsieh, J.J., and Herr, W. 2007. Species selectivity of mixed-lineage leukemia/trithorax and HCF proteolytic maturation pathways. *Mol. Cell Biol*. **27**: 7063–7072.

Cavalli, G., and Paro, R. 1999. Epigenetic inheritance of active chromatin after removal of the main transactivator. *Science* **286:** 955–958.

Chinwalla, V., Jane, E.P., and Harte, P.J. 1995. The *Drosophila* Trithorax protein binds to specific chromosomal sites and is colocalized with Polycomb at many sites. *EMBO J.* **14**: 2056–65

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

Eaton, S., and Kornberg, T.B., 1990. Repression of ci-D in posterior compartments of *Drosophila* by *engrailed*. *Genes Dev*. **4**: 1068–1077.

Eissenberg, J. C., and Shilatifard, A**.** 2010. Histone H3 lysine 4 (H3K4) methylation in development and differentiation. *Dev. Biol*. **339:** 240–249.

Fanti, L., Perrini, B., Piacentini, L., Berloco, M., Marchetti, E. Palumbo, G., Pimpinelli, S. 2008. The trithorax group and Pc group proteins are differentially involved in heterochromatin formation in *Drosophila. Chromosoma* **117**:25–39

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

Gelbart, W.M., Emmert, D.B. (2011.8.23). Calculation of RPKM to generate quantitative expression data: read-length values for modENCODE developmental timecourse RNA-Seq data. FlyBase analysis

Gregory, G.D., Vakoc, C.R., Rozovskaia, T., Zheng, X., Patel, S., Nakamura, T., Canaani, E., and Blobel, G.A. 2007. Mammalian

ASH1L Is a Histone Methyltransferase That Occupies the Transcribed Region of Active Genes. *Molecular and Cellular Biology* **27**:8466–8479

Guenther, M.G., Jenner, R.G., Chevalier, B., Nakamura, T., Croce, C.M., Canaani, E., and Young, R. A. 2005. Global and Hox-specific roles for the MLL1 methyltransferase. *Proc. Natl. Acad. Sci. USA* **102:** 8603–8608.

Hsieh, J. J., E. H. Cheng, and Korsmeyer, S.J. 2003. Taspase1: a threonine aspartase required for cleavage of MLL and proper HOX gene expression. *Cell* **115:** 293–303.

Horwich, M.D., Li, C., Matranga, C., Vagin, V., Farley, G., Wang, P., and Zamore, P.D. 2007. The *Drosophila* RNA Methyltransferase, DmHen1, Modifies Germline piRNAs and Single-Stranded siRNAs in RISC *Current Biology* **17**: 1265–1272.

Ingham, P.W., & Whittle, R. 1980 *Trithorax*: a new homeotic mutation of *Drosophila melanogaster* causing transformations of abdominal and thoracic imaginal segments. *Mol. Gen. Genet.* **179**: 607–614.

Ingham, P.W., 1998. *trithorax* and the regulation of homeotic gene expression in *Drosophila*: a historical perspective. *Int. J. Devel. Biol*. **42**: 423–429.

Ivanovska, I., Khandan, T., Ito, T., and Orr-Weaver, T.L. 2005. A histone code in meiosis: the histone kinase, NHK-1, is required for proper chromosomal architecture in *Drosophila* oocytes. *Genes Dev.* **19**: 2571-2582

Jenuwein, T., and Allis, C.D. 2001. Translating the histone code. *Science* **293**:1074-1080.

Katsani, K.R., Arredondo, J.J., Kal, A.J., and Verrijzer, C.P. 2001. A homeotic mutation in the *trithorax* SET domain impedes histone binding. *Genes Dev*. **15**: 2197–2202.

Kennison, J. A. & Tamkun, J. W. 1988. Dosage-dependent modifiers of Polycomb and Antennapedia mutations in *Drosophila*. *Proc. Natl Acad. Sci. USA* **85**: 8136–8140

Klymenko, T. and Muller, J. 2004. The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. *EMBO Rep.* **5**: 373–377.

Kuzin, B., Tillib, S., Sedkov, Y., Mizrokhi, L., and Mazo, A.,1994. The *Drosophila trithorax* gene encodes a chromosomal protein and directly regulates the region specific homeotic gene forkhead. *Genes Devel*. **8**: 2478–2490.

Lancaster, O.M., Cullen, C.F., Ohkura, H. 2007. NHK-1 phosphorylates BAF to allow karyosome formation in the *Drosophila* oocyte nucleus. *J. Cell Biol*. **179**: 817--824.

Li, Y., Trojer, P., Xu, C.F., Cheung, P., Kuo, A., Drury, W.J. 3rd., Qiao, Q., Neubert, T.A., Xu, R.M., Gozani, O., and Reinberg, D. 2009. The target of the NSD family of histone lysine methyl transferases depends on the nature of the substrate. *J Biol Chem*. **284**: 34283-34295

Liu, H., Takeda, S., Kumar, R., Westergard, T.D., Brown, E.J., Pandita, T.K., Cheng, E.H., and Hsieh, J.J. 2010. Phosphorylation of MLL by ATR is required for execution of mammalian S-phase checkpoint. *Nature* **467**: 343–346

Martin, D. G., Baetz, K., Shi, X., Walter, K.L., MacDonald, V.E., Wlodarski, M.J., Gozani, O., Hieter, P., and Howe, L. 2006. The Yng1p plant homeodomain finger is a methyl-histone binding module that recognizes lysine 4-methylated histone H3. *Mol. Cell. Biol*. **26**:7871–7879

Mazo, A.M., Huang, D.H., Mozer, B.A., and Dawid, I.B. 1990. The *trithorax* gene, a trans-acting regulator of the *bithorax* complex in *Drosophila*, encodes a protein with zinc-binding domains. *Proc. Natl. Acad. Sci. USA* **87**: 2112–2116.

Nakamura, T., Blechman, J., Tada, S., Rozovskaia, T., Itoyama, T., Bullrich, F., Mazo, A., Croce, C.M., Geiger, B., and Canaani, E. 2000. huASH1 protein, a putative transcription factor encoded by a human homologue of the *Drosophila ash1* gene, localizes to both nuclei and cell-cell tight junctions. *Proc. Natl. Acad. Sci. USA* **97**: 7284-7289.

Orlando, V. 2003 Polycomb, epigenomes, and control of cell identity. *Cell* **112**: 599-606.

Papp, B., and Müller, J. 2006. Histone tri-methylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins. *Genes Dev.* **20**: 2041-2054.

Pradhan, M., Estève, P.O., Chin, H.G., Samaranayke, M., Kim, G.D., and Pradhan, S. 2008. CXXC domain of human DNMT1 is essential for enzymatic activity. *Biochemistry* **47**: 10000-9.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., and Jenuwein, T. 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**:593–99

Reeves, R., and Nissen, M.S. 1990. The A.T-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure. *J. Biol. Chem*. **265**: 8573–8582.

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

Ringrose, L., and Paro, R. 2004. Epigenetic Regulation Of Cellular Memory By The Polycomb And Trithorax Group Proteins *Annu. Rev. Genet.* **38**:413–43

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

Rozovskaia, T., Tillib, S., Smith, S., Sedkov, Y., Rozenblatt-Rosen, O., Petruk, S., Yano, T., Nakamura, T., Ben-Simchon, L., Gildea, J., Croce, C.M. Shearn, A., Canaani, E., and Mazo, A. 1999. Trithorax and ASH1 Interact Directly and Associate with the Trithorax Group-Responsive *bxd* Region of the *Ultrabithorax* Promoter. *Mol. Cell. Biol*. **19**: 6441–6447

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

Saito, K., Sakaguchi, Y., Suzuki, T., Siomi, H., and Siomi, M.C. 2007. Pimet, the *Drosophila* homolog of HEN1, mediates 2'-*O*-methylation of Piwi- interacting RNAs at their 3' ends. *Genes Dev.* **21**: 1603-1608

Sanchez-Elsner, T., Gou, D., Kremmer, E., and Sauer, F. 2006. Noncoding RNAs of Trithorax Response Elements Recruit *Drosophila* Ash1 to Ultrabithorax. *Science* **311:** 1118-1123

Schuettengruber ,B., Chourrout, D., Vervoort, M., Leblanc,B., and Cavalli. G., 2007. Genome Regulation by Polycomb and Trithorax Proteins. *Cell* **128**: 735–745

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

Schmitt, S., Prestel, M., and Paro, R. 2005. Intergenic transcription through a polycomb group response element counteracts silencing. *Genes Dev.* **19**: 697-708.

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

Sedkov, Y., Tillib, S., Mizrokhi, L., Mazo, A., 1994. The *bithorax* complex is regulated by *trithorax* earlier during *Drosophila* embryogenesis than is the Antennapedia complex, correlating with a *bithorax*-like expression pattern of distinct early *trithorax* transcripts. *Development* **120**: 1907–1917.

Shearn, A. 1989. The *ash-1*, *ash-2* and *trithorax* genes of *Drosophila melanogaster* are functionally related. *Genetics* **121:** 517–525.

Shearn, A., Rice, T., Garen, A., and Gehring, W. 1971. Imaginal Disc Abnormalities in Lethal Mutants of *Drosophila. Proc. Nat. Acad. Sci. USA* **68**: 2594-2598.

Shi, X., Hong, T., Walter, K.L., Ewalt, M., Michishita, E., Hung, T., Carney, D., Peña, P., Lan, F., Kaadige, M.R., Lacoste, N., Cayrou, C., Davrazou, F., Saha, A., Cairns, B.R., Ayer, D.E., Kutateladze, T.G., Shi, Y., Côté, J., Chua, K.F., Gozani, O. 2006. ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature* **442**: 96–99

Simon, J. A., and Tamkun, J.W. 2002 Programming off and on states in chromatin: mechanisms of Polycomb and trithorax group complexes. *Curr. Opin. Genet. Dev*. **12:** 210–218.

Smith, S.T., Petruk, S., Sedkov, Y., Cho, E., Tillib, S., Canaani, E., and Mazo, A., 2004. Modulation of heat shock gene expression by the TAC1 chromatin-modifying complex. *Nat. Cell. Biol.* **6**: 162–167.

Tanaka, Y., Katagiri, Z., Kawahashi, K., Kioussis, D. and Kitajima. S. 2007. Trithorax-group protein ASH1 methylates histone H3 lysine 36. *Gene* **397:**161–168.

Terranova, R., Agherbi, H., Boned, A., Meresse, S., and Djabali, M. 2006. Histone and DNA methylation defects at Hox genes in mice expressing a SET domain-truncated form of Mll. Proc. *Natl. Acad. Sci. USA* **103**: 6629–6634.

Tillib, S., Petruk, S., Sedkov, Y., Kuzin, A., Fujioka, M., Goto, T., and Mazo. A. 1999. Trithorax- and Polycomb-group response elements within an *Ultrabithorax* transcription maintenance unit consist of closely situated but separable sequences. *Mol. Cell. Biol.* **19**:5189–5202.

Tolhuis, B., de Wit, E., Muijrers, I., Teunissen, H., Talhout, W., van Steensel, B., and van Lohuizen, M. 2006 Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in *Drosophila melanogaster*. *Na.t Genet.* **38**: 694–699.

Tripoulas, N.A., Herspergedr, E., Lajeunesse, D., and Shearn, A. 1994. Molecular genetic analysis of the *Drosophila* melanogaster gene absent, *small or* homeotic *discsI (ashl). Genetics* **137:** 1027-1038.

Tripoulas, N., LaJeunesse, D., Gildea, J., and Shearn, A. 1996 The *Drosophila ashl* Gene Product, Which Is Localized at Specific Sites on Polytene Chromosomes, Contains a SET Domain and a PHD Finger. *Genetics* **143**:913-928

Wang, P.F., Lin, C., Simth, E.R., Sanderson, B.W., Wu, M., Gogol, M., Alexander, T., Seidel, C., Wiedemann, L.M., Guo, H., Krumlauf, R., Shilatifard, A., 2009. Global Analysis of H3K4 methylation defines Mll1 (KMT2A) as a gene-specific activator of transcription. *Mol. Cell. Biol*. **29**: 6074–6085.

Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tacket,t A.J., Chait, B.T., Badenhorst, P., Wu, C., Allis, C.D. 2006. A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442**: 86–90

Yu, B.D., Hess, J.L., Horning, S.E., Brown, G.A., and Korsmeyer, S.J. 1995. Altered Hox expression and segmental identity in Mll-mutant mice. *Nature* **378:** 505–508.

Yu, B.D., Hanson, R.D., Hess, J.L., Horning, S.E., and Korsmeyer, S.J. 1998. MLL, a mammalian trithorax-group gene, functions as a

transcriptional maintenance factor in morphogenesis. *Proc. Natl. Acad. Sci. USA* **95:** 10632–10636.

Yuan, W. *Xu, M., Huang, C., Liu, N., Chen, S., and Zhu, B.* 2011. H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J. Biol. Chem.* **286**: 7983–7989 27.

Zhang, W., Deng, H., Bao, X., Lerach, S., Girton, J., Johansen J., and Johansen, K.M. 2006. The JIL-1 histone H3S10 kinase regulates dimethyl H3K9 modifications and heterochromatic spreading in *Drosophila. Development* 133: 229-235

Thesis Appendix

Figure A1 – Mutant Screen Cross Diagram

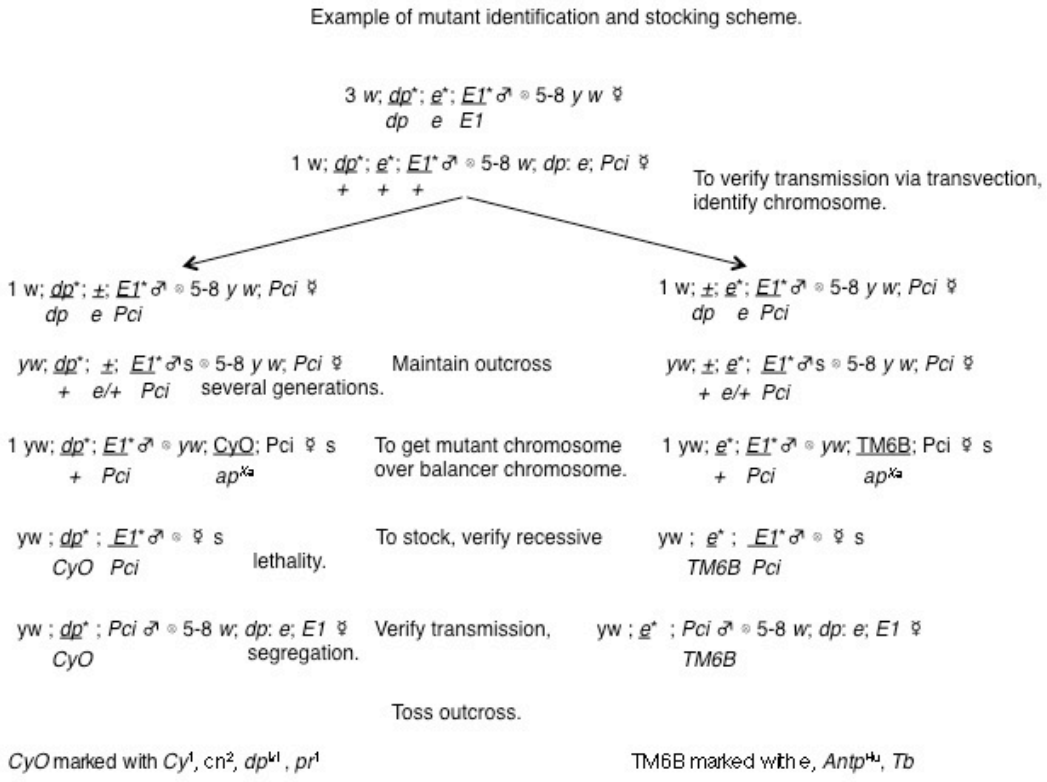


Table A1 – List of Drosophila stocks

Lab designation	Genotype	Purpose	Origin
$167	*y w[67c23(2)]*	To mate with mutagenized males in mutagenesis.	Milian Patel
$454	*w; dp; e; E1*	Mutagenesis, pigment analysis.	D. Bushey
$410	*w; dp; e; Pci*	To verify transmission of enhancement in transvection, pigment analysis.	S. Hanna
$304	*y w; Pci*	Maintain mutants in selectable background prior to stocking.	S. Hanna
$375	*w; CyO/ap^{xa}*	Stocking 2nd chromosome mutants.	S. Hanna
$448	*w; CyO/ap^{xa}; Pci*	Stocking 2nd chromosome mutants.	S. Hanna
$430	*w; ap^{xa}/TM6B*	Stocking 3rd chromosome mutants.	S. Hanna
$449	*w; ap^{xa}/TM6B; Pci*	Stocking 3rd chromosome mutants.	S. Hanna
	y w; Sp L Bc Pin/CyO; Pci	Mapping 2nd chromosome mutants.	Bloomington, $448
	y w; Gl Sb H/TM6B; Pci	Mapping 3rd chromosome mutants.	Bowling Green 2730 £3 X $449
	w; E(var)3-9/TM3, Sb Ser	Provide balancer for pigment analysis.	R.B. Hodgetts
$460	*w^{m4}; dp; e*	Pigment analysis.	Tartoff, $454
p20	*y w, P{lacZ$^{rev:w}$w^{+mC} amp^{R} ori = lacW}*	Pigment analysis.	Bloomington
	w; dp; e	Pigment analysis.	$167, $410
$37	+ Oregon-R	PCR P, KP control	R.B. Hodgetts
π 2	wildtype P strain	Positive KP control	Bowling Green

$ = stock number

Table A2 – Complementation Table for Chromosome 2

Chromosome 2 Enhancer lethal Complementation

Mutant	1a27a	3a22a	3a52a	3a66a	3a90a	3a97a	4a7a	1a67a	2a20a	4a29a	4a30a	4a91c'	2a59g	2a66a	3a95a	1a9a	3b14a	5a94a	1a53b	2a66e	2a69a	2b33a	3a38b	4a42a	4a49a	4a50a	4a53b	4a75a	4a81a	5a44a
1a27a		0/167	0/166	0/219	0/208	0/200	0/153	+	+	+	+	+	+	+	+	+	+	+	+		+	+	+	+	+	+	+	+		+
3a22a	0/167		0/218		0/192	0/235	0/142	+	+	+	7/99	+	+	+	+	+	+	+	+		+	+	+	+	+	+	+	+		+
3a52a	0/166	0/218		0/125	7/192	0/173	0/148	+	+				+	+	+	+			+		+	+				+		+		
3a66a	0/219		0/125			0/178		+	+	+		+	+	+	+	+	+	+	+		+	+	+	+	+	+		+		+
3a90a	0/208	0/192	7/192			0/154	1/141	+m	+	+	+	+	+	+	+	+	+	+	+		+	+		+	+	+	+	+		+
3a97a	0/200	0/235	0/173	0/178	0/154		0/185	+	+	+		+	+	+		+		+	+		(+)	+		+	+	+	+			(+)
4a7a	0/153	0/142	0/148		1/141	0/185		+	(+)	(+)		8/64		+	+	(+)	+	+	+	+	+	+		+	+	+	+			+
1a67a	+	+	+	+	+m	+	+		0/112	+	5/120	(+)1/103	(+)	+	+	+	+	3/30	+		+	+		+	+	+	+	+	+	+
2a20a	+	+	+	+	+	+	(+)	0/112		11/149	(-)	(+)	+	(+)	(+)	(+)	(+)	+	+		+	+	+	+	+	+	(+)	+	5/85	+
4a29a	+	+		+	+	+	(+)	+	11/149		0/140	0/137	+	+	13/104	+	+	4/76	7/64	+	+	+	+	+	+	+	9/71	4/155	0/34	+
4a30a	+	7/99			+			5/120	(-)	0/140		10/149	+	+	+	9/145	+	8/127	+	+	(+)	+	+	+	+	+	2/98	+	0/71	8/138
4a91c'	+	+		+	+	+	8/64	(+)1/103	(+)	0/137	10/149		+	(+)	+	+	+	+	7/119		+	+	+	+	+	+	+	+		+
2a59g	+	+	+	+	+	+		(+)	+	+	+	+		0/365	+	(+)	+	+	+		+	+	+	+	+	+	+	+		+
2a66a	+	+	+	+	+	+	+	+	(+)	+	+	(+)	0/365		(-)0/190	(+)	+	+	+	+	(+)	+	+	22/499	+	19/467	(+)	10/96	(+)	+
3a95a	+	+	+	+	+		+	+	(+)	13/104	+	+	+	(-)0/190		+	+	+	+		+	+	+	+	+	+	+	+		+
1a9a	+	+	+	+	+		(+)	+	(+)	+	9/145	+	(+)	(+)			0/148	0/153	+		+	+		+	+		(+)	+		9/86
3b14a	+	+		+	+		+	+	(+)	+	+	+	+	+	+	0/148		+	+		+	+	+	+	+	+	+	+	(+)	+
5a94a	+	+		+	+	+	+	3/30	+	4/76	8/127	+	+	+	+	0/153	+		+		+	+	+	6/70	+	+	+	+	+	+
1a53b	+	+	+	+	+	+	+	+	+	7/64	+	7/119	+	+	+	+	+	+			+	+	+	+	+	+	+	+	(3/124)	4/42
2a66e							+			+	+			+																
2a69a	+	+	+	+	+	(+)	+	+	+	+	(+)	+	+	(+)	+	+	+	+	+			+	+	+	+	+	+	+		+
2b33a		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		+		+	+	+	+	+	+		+
3a38b	+	+		+					+	+	+	+	+	+	+	+		+	+		+	+		+	+	+	+	2/4		+
4a42a	+	+		+	+	+	+	+	+	+	+	+	+	22/499	+	+	+	6/70	+		+	+	+		+	+	+	+	+		+
4a49a	+	+		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		+	+	+	+		+	+	+		+
4a50a	+	+	+	+	+	+	+	+	+	+	+	+	+	19/467	+	+	+	+	+		+	+	+	+	+		+	+	+	+
4a53b	+	+			+	+	+	+	(+)	9/71	2/98	+	+	(+)		(+) +	+	+	+		+	+	+	+	+	+		+		+
4a75a	+	+	+	+	+			+	+	4/155	+	+	+	10/96	+	+	+	+	+		+	+	2/4	+	+	+	+			+
4a81a								+	5/85	0/34	0/71			(+)			(+)	+	(3/124)											+
5a44a	+	+		+	+	(+)	+	+	+	+	8/138	+	+	+	+	9/86	+	+	4/42		+	+	+	+	+	+	+	+	+		

+ = complements. Groups from left *CG8878*, *cg*, *unknown*, *unknown*.

Table A3 – Complementation Table for Chromosome 3

Chromosome 3 En(PDS): lethal complementation assay

Mutant																													
3a1a'		0/121	0/151	0/53	0/168	+	(+)	+	+	+	14/182	+	+	+	+	+	+	+	+	15/158	(+)	(+)	+	+	+		+	(+)	+
3a31a	0/121			0/65	0/179	+	+	+	+	+	13/139		+	+	+	+	+	+	+	+	(+)	+	+	+	(+)		+	+	
3a42a	0/151					+	+	+	(+)	+	(+)	5/99	+	+		+	+	+	+		13/176		+	+		(+)	(+)	+	(+)
4a5a	0/53	0/65			0/46	dng	+	+		+	13/116		+	+	+	+	+			6/73	3/165			+	+		+	+	
4a89a	0/168	0/179		0/46		+	+		+		(+)		+	+		+	+	+	+	+	+	+	+	+			+	+	+	(+)
2a15a	+	+	+	dng	+		0/176	0/112	0/144	+	+	+	+		+	+	+	+	dng	+	+	+	+	+	+	+	(+)	+	dng
2a46a	(+)	+	+	+	+	0/176		0/44	0/159	+	+	+	+	+	+	+	+	(+)	+	+	+	(+)	+	+	(+)	+	+	(+)	+
2a69a	+	+	+	+		0/112	0/44		0/103	+	+	+	+	+		+	+	+	+	+	+	+	+	+	+	+	+	+	+
4a81a	+	+	(+)		+	0/144	0/159	0/103		+	(+)	+	+	+		+	+	+		+	+		+	+	+	+	(+)	+	(+)
3a26a	+	+	+	+		+	+	+	+		0/149	4/182	+	+	(+)	+	+	+	+	+	+	+	+	+	+	+	+	+	
3b77a	14/182	13/139	(+)	13/116	(+)	+	+	+	(+)	0/149		0/32	+		+	+	+	+	+	+	+	+	(+)	+	+		+	+	+
5b19a	+		5/99			+	+	+	+	4/182	0/32		+	+		+	+	+		+	+		+	+	+		(+)	+	
3a29a	+	+	+	+	+	+	+	+	+	+	+	+		0/123	0/124	+	+	+	+	+	+	+	+		+		+	+	
4a4a	+	+	+	+	+		+	+	+	+		+	0/123		1/49	+	+	(+)	+	+	+	+	+	+	+	+	+	+	
5b63a	+	+		+		+	+			(+)	+		0/124	1/49		+	+	+	+	+	+	+			(+)	+	+	+	
1a27a	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		+	+	+	(+)	+	+	+	+	+	+	+	+	+
1a64a	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		+	+	+	+	+	+	+	+	+	+	+	+
2a13c	+	+		+	+	+	(+)	+	+	+	+	+	+	(+)	+	+	+		+	+	+	+	+	+	+	+	+	+	+
2a49b	+	+	+		+	dng	+	+		+	+		(-)	+	+	+	+	+		+	+	+		+	+	+	+	+	
2a52e	15/158	+		6/73	+	+	+	+	+	+	+	+	+	+	+	(+)	+	+	+		+	+	+	+	+	+	+	+	
2a59c	(+)	(+)	13/176	3/165	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+			+	+	(+)	+	+	+	
2b35a	(+)	+			+	+	(+)	+		+	+		+	+	+	+	+	+	+	+					+	+	+	(+)	+
2b42a	+	+	+		+	+	+	+	+	+	(+)	+	+	+		+	+	+		+	+			+	+	+	+	+	(+)
3a38b	+	+	+	+	+	+	+	+	+	+	+	+	+	+		+	+	+	+	+	+		+		+	+	+	+	+
3a79b	+	(+)		+		+	(+)	+	+	+	+	+	+	+	(+)	+	+	+	+	+	(+)	+	+	+		+	+	+	
3b76a			(+)		+	+	+	+	+	+				+	+	+	+	+	+	+	+	+	+	+	+		+		
4a13a	+	+	(+)	+	+	(+)	+	+	(+)	+	+	(+)	+	+	+	+	+	+	+	+	+	+	+	+	+	+			+	+
4a92a'	(+)	+	+	+	+	+	(+)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	(+)	+	+	+		+		+	
5a63a	+		(+)		(+)	dng	+	+	(+)		+					+	+	+				+	(+)	+			+	+	
Mutant	3a1a'	3a31a	3a42a	4a5a	4a89a	2a15a	2a46a	2a69a	4a81a	3a26a	3b77a	5b19a	3a29a	4a4a	5b63a	1a27a	1a64a	2a13c	2a49b	2a52e	2a59c	2b35a	2b42a	3a38b	3a79b	3b76a	4a13a	4a92a'	5a63a

+ = complements. Groups from left *ash1*, unknown, *trx*, *TAF4*. Box in red indicates *ash1*, *trx* synthetic lethality.

Table A4 - Polymerase chain reaction (PCR).

Two typical reaction mixes were used in this work: (1) low fidelity, and (2) high fidelity. They are described below. Typically, half of one reaction was sufficient for visualizing on an agarose gel, while 4 reactions usually provided enough template for several sequencing reactions.

Component	Low Fidelity Reaction		High Fidelity Reaction	
dNTP	10 mM	0.6 μL	10 mM	0.6 μL
Buffer (supplied)	10 X	3 μL	5 X	1.5 μL
Forward primer	10 μM	1.5 μL	10 μM	1.5 μL
Reverse primer	10 μM	1.5 μL	10 μM	1.5 μL
ddH_2O		22.2 μL		19.1 μL
Taq	NEB	0.2 μL	Phusion®	0.3 μL
Template		1 μL		1 μL

Table A5 - Primers used to characterize *ash1* mutations. (Accession NT_037436.3)

Primer	5' Sequence 3'	T_M °C	Start	End
Ash1 Preprimer	CGAAGCGGAATGGAATGAAGTC	54	19,592,569	19,592,548
Ash1-F0	CGTGATGGGTGCCGTGTTGATA	57	19,593,827	19,593,806
Ash1-R0	CGAAATGTGGAGCCGCAGTACG	58	19,592,512	19,592,533
Ash1-F1	ATTCCGCCGCAAGGTCACACTG	59	19,591,908	19,591,887
Ash1-R1	ACCGCTCTATTTTTGGCCTCCG	57	19,590,976	19,590,997
Ash1-7F	CACCGATGGACTGCGAATGAGA	56	19,591,113	19,591,092
Ash1-7R	GGCGCTTTTTCAACGGCAACTT	58	19,589,015	19,589,036
Ash1-8F	CCAATGGCAGCGGAAGCAGTAA	58	19,589,122	19,589,101
Ash1-8R	CCTCGCCCACGTACTCCAGAAT	56	19,587,008	19,587,029
Ash1-9F	CCCGGAGTGGAGCGCTTTATGA	59	19,587,107	19,587,086
Ash1-9R	TGGCCCGGCTCCTACGTCTAAT	57	19,584,787	19,584,808

Ash1-10F	ACCATCCAGCCTGCACCAAGAG	57	19,584,991	19,584,970
Ash1-10R	GCCATCAGACCGGGAGGAGTTA	56	19,582,953	19,582,974
Ash1F1 MF	GCCCAGAAAACAATCAAGCGTA	53	19,591,504	19,591,483
Ash1F1 MR	TTCGATTTGGTGGCACTCTTAA	51	19,591,394	19,591,415
Ash1F2 MF	ATGAGTGTGGGTGCGGCTAGTG	55	19,590,107	19,590,086
Ash1F2 MR	GCAGCTGGCATCGGACTTGTAA	56	19,589,952	19,589,973
Ash1F3 MF	GCGTCCTCAAACACCAGCTAGA	53	19,588,104	19,588,083
Ash1F3 MR	CGTTTTTTGGCTGGCCTTATCT	53	19,587,982	19,588,003
Ash1F4 MF	CCAGGTGGAGCAGGGACATTAC	54	19,586,001	19,585,980
Ash1F4 MR	CTGGCGATCTTCTGTTGCTCAT	52	19,585,846	19,585,867
Ash1F5 MF	CCCCCTCTCGTCTTTGCAATTG	56	19,584,067	19,584,046
Ash1F5 MR	ACGCGACAGTTAGGCAATCCAA	55	19,583,807	19,583,828
Ash1F2 M2F	AAGCGAGTGGAGAGCGATACGG	56	19,590,704	19,590,683
Ash1F2	TGGGTCTACCTCTGCGCTTTCG'	57	19,590,473	19,590,494

M2R				
Ash1F0 MF	GGGCCCAGTGATTGTCCAGCTC	58	19,593,291	19,593,270
Ash1F0 MR	TGGTTCGTGCGCTTCTGGATAA	56	19,593,098	19,593,119
8F-F3MR-MF	TAGCTATCCTCCCCCCGGTGTG	58	19,588,620	19,588,599
8F-F3MR-MR	GCTGGCCAACCTCCTTTTTCGT	57	19,588,460	19,588,481
9F-F4MR-MF	ACGGCAGGCGGGAAAAGATATT	57	19,586,571	19,586,550

Table A6 - Primers used to characterize *trx* mutations. (Accession NT_033777.2)

Primer	5' Sequence 3'	TM °C	Start	End
trx 1 F	TGCGCCAAAAGGAGTGAGTGAG	56	10103539	10103518
trx 1 R	CCGGAATTGCCATCTGGAGAGG	58	10102539	10102560
trx 2 F	CGGTGACGGTGATAAAACGAGA	53	10102706	10102685
trx 2 R	TCGGACATGGTGGGTTACTGAG	53	10101536	10101557
trx 2b F	ATCGCACATCGTCACCCACCAC	58	10105433	10105412
trx 2b R	GGATTTTCTCGTTCGGCTTTTG	54	10104727	10104748
trx 3 F	GCGCTTGCAACTCTTCCACATG	56	10101683	10101662
trx 3 R	CGCCACCCTCATCATCTTCATC	55	10100544	10100565
trx 4 F	TTTGGCTTATCCCGCGCTTTTG	59	10100598	10100577
trx 4 R	TGGCTCTTGGCTGAATGAATGC	55	10099539	10099560
trx 5 F	TCCTCCGGCAGCCAACAGTATC	57	10099679	10099658
trx 5 R	GCTGCACTTGCTGTCCCTTCTT	54	10098510	10098531
trx 6 F	CGTCTCCTGTTGGTGATGAATC	50	10098615	10098594
trx 6 R	AACGCTGTAGAGACTGGCTTTG	50	10097519	10097540
trx 7 F	TGTGCCCGCAGGAATGAGAGTT	57	10097608	10097587
trx 7 R	GGCTTTCAAGGCATTTCCATTC	53	10096493	10096514
trx 8 F	GAGCCACTGTCGGATGCAATGT	55	10096632	10096611
trx 8 R	TGTGGCCATCTCCTTTTTGTTT	52	10095512	10095533
trx 9 F	GCTGTCGAAAAACTTGGCTGAA	53	10095587	10095566

trx 9 R	GATGGGCTGTTGCACCTGAGTC	55	10094486	10094507
trx 10 F	AGCAGGCGGCCGTCAAACAGAG	61	10094610	10094589
trx 10 R	ATTCCCGCTGCCTTGGCTTGAG	60	10093545	10093566
trx 11 F	TGCCCACATCTTCAGGAGCCTC	57	10093665	10093644
trx 11 R	CCGGCTGTTCCTCCAAGATTGC	58	10092580	10092601
trx 12 F	CCAAGTGCTATGCCCAAAAGTC	52	10092705	10092684
trx 12 R	ACCATTGCGCTTGTGATACTTG	51	10091618	10091639
trx 13 F	AGGCGGGCACATGGACTTACTC	56	10091782	10091761
trx 13 R	ACACCTTTTCGAACCGCATGAG	54	10090674	10090695
trx 14 F	GGGCACAAGCACATCATCATCT	53	10090796	10090775
trx 14 R	TGTTGTTGTGGTTGCCTGTTCA	53	10089504	10089525
trx 15 F	AGAGAACGAAACCAAGCATCAC	49	10089566	10089545
trx 15 R	CGACAAGACGGCACTATTAGAG	48	10088809	10088830

Table A7 - Primers used to characterize *CG8878* mutations. (Accession NT_033778.3)

Primer	5' Sequence 3'	T_M °C	Start	End
CG8878 0 F	CGCGCGACTCCGTATATAATC	51	8038948	8038928
CG8878 0 R	GGCCGGCATTATTAAACAAAGAC	53	8038155	8038177
CG8878 0 -1 R	TTTCGCTTCCATTTGGCTTTTC	54	8038475	8038496
CG88781 F	TTCCCTCGCATATTTATCAGTTTTAG	52	8037748	8037723
CG8878 1 R	GGATACAAATGTGACAATGCCGAGTT	56	8036838	8036863
CG8878 2 F	TTGCCCTGCCAGCTCGGAGACA	62	8036987	8036966
CG8878 2 R	GCGACGTGACGAGGGGGGAATA	60	8035989	8036010
CG8878 3 F	CGCGAGGAGTTTAACGAAGACG	55	8036102	8036081
CG8878 3 R	CGCGGTTCGTAAGTGAGTGCTC	55	8035199	8035220
CG8878 4 F	GGTGGATGGCAGCAAGTGTGAC	55	8035311	8035290
CG8878 4 R	CTCGCCGTGTTTGGGTTCAGTC	57	8034327	8034348
CG8878 5 F	CACCGGCGAGGGAGAGCGAAAG	62	8034405	8034384
CG8878 5 R	GGTCGTGGAGTTCGGTTGTGCC	59	8033408	8033429

Table A8 - Primers used to identify *P*, *KP* contamination.

Primer	5' Sequence 3'	Application
PRPT	TAACATAAGGTGGTCCCGTCG	*P* or *KP* element detection
THE3	GTACTCCCATGGTATAGCC	*P* or *KP* element detection
RP49-F1	AGCATACAGGCCCAAGATCG	Control
RP49-R1	AGTAAACGCGGGTTCTGCAT	Control

END