INTER-KINGDOM EPIGENETICS: CHARACTERIZATION OF MAIZE *B1* TANDEM REPEAT-MEDIATED SILENCING IN *DROSOPHILA MELANOGASTER*

by

Lori A. McEachern

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
August 2010

DALHOUSIE UNIVERSITY

DEPARTMENT OF BIOLOGY

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "INTER-KINGDOM EPIGENETICS: CHARACTERIZATION OF MAIZE *B1* TANDEM REPEAT-MEDIATED SILENCING IN *DROSOPHILA MELANOGASTER*" by Lori A. McEachern in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Dated: August 19, 2010

External Examiner: _____

Research Supervisor: _____

Examining Committee: _____

_____

Departmental Representative: _____

DALHOUSIE UNIVERSITY

DATE: August 19, 2010

AUTHOR: Lori A. McEachern

TITLE: INTER-KINGDOM EPIGENETICS: CHARACTERIZATION OF MAIZE *B1* TANDEM REPEAT-MEDIATED SILENCING IN *DROSOPHILA MELANOGASTER*

DEPARTMENT OR SCHOOL: Department of Biology

DEGREE: Ph.D. CONVOCATION: October YEAR: 2010

DEDICATION

For

my mom, whose strength is inspiring

and

my dad, who I miss every day.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Transgenic organisms are a valuable tool for studying epigenetics, as they provide significant insight into the evolutionary conservation of epigenetic control sequences, the interacting proteins, and the underlying molecular mechanisms. Paramutation is an epigenetic phenomenon in which the epigenetic status and expression level of one allele is heritably altered after pairing with another. At the *b1* locus in maize, a control region consisting of seven 853 bp tandem repeats is required for paramutation. To study the conservation of the epigenetic mechanisms underlying maize *b1* paramutation, I created transgenic *Drosophila* carrying the maize *b1* control region flanked by FRT sites and adjacent to the *Drosophila white* reporter gene. The maize *b1* tandem repeats caused epigenetic silencing in *Drosophila*, as *white* expression consistently increased following repeat removal. A single copy of the tandem repeat sequence was sufficient to cause silencing, and silencing strength increased as the number of repeats increased. *Trans* interactions, such as pairing-sensitive silencing, were also observed and appear to require a threshold number of *b1* tandem repeats, similar to paramutation in maize. Analysis of transcription from the repeats showed that the *b1* tandem repeats are transcribed from both strands in *Drosophila*, as they are in maize. Bidirectional transcription was found to extend to the regions flanking the repeats, and persisted in "repeats-out" transgenes following repeat removal. However, aberrant transcription was lost when a zero-repeat transgene was moved to a new genomic position, suggesting that it may be due to an epigenetic mark that is retained from the previous silenced state. A search for modifiers of *b1* repeat-mediated silencing demonstrated that Polycomb group proteins are involved. Together, these results indicate considerable conservation of an epigenetic silencing process between the plant and animal kingdoms. Genomic imprinting is a related epigenetic process in which parent-specific epigenetic states are inherited and maintained in progeny. The conservation of epigenetic mechanisms was further explored via an in-depth review of the molecular mechanisms underlying genomic imprinting in plants, mammals and insects, and identification of potentially imprinted genes in *Drosophila* by microarray analysis.

LIST OF ABBREVIATIONS USED

ABC	ATP-binding cassette
AGO	Argonaute
ANOVA	Analysis of variance
AP-3	Adaptor-protein 3
ASH	Absent, small, or homeotic discs
AUB	Aubergine
BLOC	Biogenesis of lysosome-related organelle complexes
bp	Basepair
BRM	Brahma
CBBP	CXC domain *b1*-repeat binding protein
cDNA	Complementary DNA
CpG	Cytosine guanine dinucleotide
CTCF	CCCTC-binding factor
Df	Deficiency
DL	Delta
DME	Demeter
DMR	Differentially methylated region
DRM	Domains rearranged methyltransferase
dsRNA	Double stranded RNA
eGFP	Enhanced green fluorescent protein
E(PC)	Enhancer of polycomb
ESC	Extra sex combs
E(VAR)	Enhancer of variegation
FIE	Fertilization-independent endosperm
FIS	Fertilization independent seed
FLP	Flipase recombinase enzyme
FRT	Flipase recombinase target
GAF	GAGA Factor
GRH	Grainyhead
His2Av	Histone H2A variant
HOPS	Homotypic vacuolar protein sorting
HP1	Heterochromatin protein 1
HPS	Hermansky-Pudlak syndrome
ICR	Imprint control region
IGF2	Insulin-like growth factor 2
kb	Kilobases
LAM	Lamin
LAMP	Lysosomal-associated membrane protein
MEA	Medea
miRNA	Micro RNA
MOP	Mediator of paramutation
MOR	Moira
mRNA	Messenger RNA
MSI1	Multicopy suppressor of ira 1

MSL	Male-specific lethal
NCD	Non-claret disjunctional
ncRNA	Non-coding RNA
PAL	Paternal loss inducer
PC	Polycomb
PcG	Polycomb group
PEV	Position-effect variegation
PHO	Pleiohomeotic
PHO-RC	PHO repressive complex
piRNA	Piwi-interacting RNA
PRC	Polycomb repressive complex
PRE	Polycomb response element
PSC	Posterior sex combs
PSQ	Pipsqueak
rasiRNA	Repeat associated short interfering RNA
RMR	Required to maintain repression
RNAi	RNA interference
roX	RNA on the X
RPTS	Repeats
RT-PCR	Reverse transcriptase PCR
siRNA	Small interfering RNA
SNR1	SNF5-related 1
spn-E	Spindle-E
SU(HW)	Suppressor of hairy wing
SU(VAR)	Suppressor of variegation
SU(Z)	Suppressor of zeste
TAS	Telomere associated sequence
TNA	Tonalli
TPE	Telomere position effect
TRE	Trithorax response element
TRX	Trithorax
trxG	Trithorax group
UTR	Untranslated region
VAS	Vasa
VPS	Vacuolar protein sorting
Xist	X inactive specific transcript

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Vett Lloyd, for her enthusiasm, expertise, support and advice. Thank you to my committee members, Dr. Bill Pohajdak and Dr. Ian Meinerzhagen, for providing me with their valuable time, suggestions, and direction. I would also like to thank Dr.Tom MacRae for his support within the department, and Will MacDonald, Tara Edwards, and Dr. Andy Haigh for their assistance within the lab. I thank my parents, Les and Debbie, and my brother, Jeremy, for their love, guidance, and support. Finally, I would like to thank my best friend, soulmate, and 24-hour science advisor, Adam MacNeil, for always believing in me, and for providing me with strength and motivation when I needed it most.

CHAPTER 1

INTRODUCTION

1.1 EPIGENETICS AND EPIGENETIC MECHANISMS

The origin of the term "epigenetics" is credited to Conrad H. Waddington, who in 1942 used this term to describe the "causal mechanisms" that act during development and cause a genotype to give rise to a particular phenotype (Waddington, 1942). Waddington's definition remains most applicable in the field of developmental biology, in the examination of the pathways and processes that occur during development and link a genotype, or a gene's expression, to phenotypic characteristics. In molecular biology and genetics, the term epigenetics is further defined as changes in gene expression that occur without alteration of the underlying DNA sequence, and are heritable through mitosis and/or meiosis. It is this definition of epigenetics that serves as the basis for the studies presented in this thesis, which focus on analyzing an epigenetic control region from maize using a transgenic *Drosophila* system.

The model organism *Drosophila melanogaster* has been widely used to study a variety of epigenetic processes and mechanisms. As many of the proteins that participate in epigenetic processes are evolutionarily conserved, these studies have broad applicability. The molecular basis of epigenetics includes a variety of mechanisms that lead to gene expression or repression, such as histone modifications, changes in higher-order chromatin structure, DNA methylation, RNA interference (RNAi), and non-coding RNAs.

Histone modifications are at the very core of epigenetic gene regulation, and many other epigenetic processes ultimately contribute to the epigenetic status of a locus by directing or targeting modifications of histone proteins. DNA is packaged within the nucleus by its association with nucleosomes, protein structures that consists of two copies

of four different histone proteins (H2A, H2B, H3 and H4). This complex of DNA and protein is termed chromatin; densely-packed "inactive" chromatin is termed heterochromatin, while loosely packed chromatin is termed euchromatin. Chemical modifications of amino acids in the histone proteins, such as methylation, acetylation, phosphorylation, sumoylation, ubiquitination, and ribosylation, can lead to the formation of heterochromatin or euchromatin, depending on the nature and position of the modification. The inclusion of variant histones can also contribute to changes in chromatin structure. Further changes in higher-order chromatin structure may be facilitated by DNA-binding proteins that mediate the formation of chromatin loops or other complex chromatin structures, and thereby modify the access of regulatory proteins, chromatin remodelling proteins, and histone modification enzymes, to their target sequences or sites. These DNA-binding proteins and higher-order chromatin structures may also contribute to epigenetic gene expression by localizing the target sequences to a particular region within the nucleus.

Maintenance of silent or active chromatin states often also involves the well characterized Polycomb group (PcG) and trithorax group (trxG) proteins, which exhibit extensive evolutionary conservation in eukaryotes, with homologues identified in fungi, plants, and animals (Schuettengruber *et al.*, 2007). These proteins form large multimeric complexes that maintain transcriptional repression and activation, primarily by directing histone modifications and chromatin remodelling (Schuettengruber *et al.*, 2007). Epigenetic silencing and *trans*-acting silencing (silencing enhanced and/or mediated by sequences at a distant genomic site) by PcG and other chromatin proteins has also been observed to involve non-coding RNAs, small RNAs, and the RNAi pathway, demonstrating the inter-connectedness of these epigenetic mechanisms (Grimaud *et al.*,

2006a; Kavi *et al.*, 2006; Schmitt and Paro, 2006). While PcG proteins are primarily transcriptional repressors and trxG proteins transcriptional activators, accumulating evidence suggests that some proteins that participate in these epigenetic processes exhibit dual functions (Fujioka *et al.*, 2008; Grimaud *et al.*, 2006b).

DNA methylation is the process through which a methyl group is added to nucleotides in the DNA sequence. The most frequent target of DNA methylation in animals is cytosine bases present in CpG dinucleotides (Bird, 2002), although non-CpG methylation also occurs (Haines *et al.*, 2001; Ramsahoye *et al.*, 2000), and is quite common in plants and some insect species (Chan *et al.*, 2005; Field *et al.*, 2004; Gruenbaum *et al.*, 1981; Lyko *et al.*, 2000). In most organisms that exhibit DNA methylation, *de novo* methyltransferases establish DNA methylation, while maintenance methyltransferases replicate pre-existing methylation patterns as the DNA is replicated. DNA methylation at promoter sequences is frequently associated with repression of gene expression; however, methylation-requiring enhancers, repressors, and protein-binding sequences are also important in epigenetic gene regulation. Evidence suggests that DNA methylation and histone modifications frequently exhibit epigenetic "cross-talk", with DNA methylation guiding histone modifications, and histone modifications similarly influencing DNA methylation (Fuks, 2005; Vaissiere *et al.*, 2008). These two epigenetic processes therefore often function in a mutually reinforcing epigenetic loop that ensures maintenance of a repressive chromatin state.

RNAi pathways involve the processing of large coding or non-coding RNAs into small RNAs. These small RNAs can modify gene expression post-transcriptionally, by degrading an mRNA transcript or inhibiting its translation, or transcriptionally, by mediating chromatin modifications that promote the formation of heterochromatin and

thereby inhibit transcription (Matzke and Birchler, 2005). The molecular mechanisms underlying RNAi-directed heterochromatin formation have been most thoroughly studied in yeast, where transcripts from heterochromatic regions of the genome were found to be processed into small interfering RNAs (siRNAs), which then recruited histone methylation that contributed to heterochromatin formation (Kloc *et al.*, 2008). Both transcriptional and post-transcriptional RNAi-mediated silencing have been observed in a wide range of eukaryotic organisms, and key components of the RNAi machinery are conserved in plants, yeast, and animals (Cerutti and Casas-Mollano, 2006). The diverse range of RNAi-mediated pathways and processes that have been reported throughout the eukaryotic kingdom are therefore likely based on an evolutionarily conserved silencing process that was present in ancient eukaryotes.

Non-coding RNA transcripts may also orchestrate changes in chromatin structure directly, rather than through an RNAi pathway, by mediating protein recruitment, histone modifications and DNA methylation at a target site (Bernstein and Allis, 2005; Matzke and Birchler, 2005; Zaratiegui *et al.*, 2007). A well studied example of a non-coding RNA that mediates changes in chromatin structure is the 17 kb non-coding mammalian *Xist* transcript, which is essential for X chromosome inactivation. *Xist* is expressed from the X chromosome that will be inactivated and subsequently coats that chromosome, which triggers a variety of chromatin remodelling events, including histone modifications and the incorporation of a specialized histone variant. These modifications ensure epigenetic silencing of the inactive X (Bernstein and Allis, 2005). Non-coding RNAs are also essential for dosage compensation of the X chromosome in *Drosophila melanogaster*, which is accomplished by hypertranscription of the X chromosome in males. In *Drosophila* males, the non-coding *roX* RNAs are highly expressed from the X

chromosome, and are incorporated into the MSL chromatin remodelling complex, which binds to hundreds of sites along the X chromosome and catalyzes acetylation of histone H4 at lysine 16, a modification that promotes a transcriptionally active chromatin state (Scott and Li, 2008). Similarly, intergenic non-coding RNA transcripts from a ribosomal RNA gene cluster in mice have been found to interact with a chromatin remodelling complex to establish and maintain a specific heterochromatin structure that impairs transcription at that locus (Mayer *et al.*, 2006).

The molecular mechanism by which small or non-coding RNAs direct DNA modifications, histone modifications, and changes in chromatin structure is not yet fully understood, but it is hypothesized that these RNAs recruit DNA-binding and chromatin-modifying proteins to a target site via their interaction with both the protein complex and either nascent RNA transcripts or genomic DNA (Buhler *et al.*, 2006; Grewal and Moazed, 2003; Irvine *et al.*, 2006; Mayer *et al.*, 2006). An RNA-RNA interaction model is supported by the observation that DNA methylation of a group of genes in *Arabidopsis* is directed by a small RNA that targets an exon-exon junction (Bao *et al.*, 2004). Alternatively, the incorporation of small or non-coding RNAs into a chromatin remodelling protein complex may induce a conformational change in the complex that alters its binding specificity and target sites (Scott and Li, 2008).

1.2 STUDYING EPIGENETIC PHENOMENA IN *DROSOPHILA MELANOGASTER*

Epigenetic effects on gene expression have been extensively studied in *Drosophila melanogaster* (Schulze and Wallrath, 2007). These include position effect variegation (PEV), which occurs when a euchromatic gene is relocated to a genomic

position within or adjacent to a region of heterochromatin. The normally euchromatic gene adopts a heterochromatic structure and is silenced in a subset of cells, resulting in mosaic or variegated expression (Wallrath and Elgin, 1995). Analysis of PEV in *Drosophila* provided the first identification of many genes that encode chromatin modifying proteins, originally termed *Suppressors* or *Enhancers of variegation* (*Su(var)*s or *E(var)*s).

The characteristics of heterochromatin-induced gene silencing may vary depending on the position and nature of the heterochromatin. While silencing due to the insertion or juxtaposition of euchromatic marker genes into centric heterochromatin has been most extensively studied, insertion of marker genes into the tandem repeats that form the *Drosophila* subtelomeric heterochromatin domains produces a distinct type of epigenetic silencing termed telomeric position effect (TPE) (Biessmann *et al.*, 2005). Transgenes that are inserted into subtelomeric heterochromatin can also induce silencing of euchromatic transgenes inserted elsewhere in the genome (Roche and Rio, 1998; Ronsseray *et al.*, 2003; Ronsseray *et al.*, 1998). This process, termed telomeric *trans*-silencing effect, requires both proteins that participate in heterochromatin formation and those that participate in a small-RNA silencing pathway (Josse *et al.*, 2007), further exemplifying the inter-connectedness of these core epigenetic mechanisms in a variety of epigenetic phenomena. When relocated to a euchromatic region, subtelomeric tandem repeats can also cause pairing-sensitive silencing of adjacent marker genes (Boivin *et al.*, 2003), a phenomenon in which silencing is enhanced and overall expression is decreased in flies homozygous for the marker gene and silencing element. This effect is also frequently observed for transgenes containing sequences of known PcG or trxG response elements (Kassis, 2002).

Additional epigenetic phenomena with *trans*-effects have also been observed in *Drosophila*. For example, a heterochromatic insertion into one allele can cause silencing of a paired wild-type allele in *trans,* a process termed *trans*-inactivation (Dreesen *et al.*, 1991). Evidence suggests that this silencing correlates with nuclear localization of the wild-type allele to a region of the nucleus containing centric heterochromatin (Csink *et al.*, 2002). Similarly, an epigenetically silenced transgene array can silence a paired wild-type gene in *trans* (Dorer and Henikoff, 1997). Enhancers or silencers have also been observed to influence gene expression in *trans*, through pairing of homologous chromosomes, in a process called transvection (Duncan, 2002).

Given the evolutionary conservation of many epigenetic proteins and core epigenetic silencing mechanisms, *Drosophila* have proven an invaluable tool for analysing the molecular basis and conservation of epigenetic phenomena observed in other organisms. For example, transgenic *Drosophila* strains have successfully been used to study several examples of mammalian imprinting. Genomic imprinting is an epigenetic process in which an allele is marked based on the sex of the parent transmitting it. This epigenetic mark can lead to transcriptional repression of fully functional alleles, based strictly on whether they were inherited through the male or female germline. Imprinting has been observed in a wide range of eukaryotic organisms, including plants, insects, and mammals (reviewed in chapter 6).

One of the best characterized examples of mammalian imprinting is that of the *H19/Igf2* locus. Using transgenic *Drosophila*, a 1.2 kb silencing element was identified within the *H19/Igf2* imprint control region (ICR) (Lyko *et al.*, 1997). Subsequent experiments showed that this 1.2 kb element also functions as a silencer at the endogenous mouse locus (Drewell *et al.*, 2000). At both the endogenous mouse locus and

in the transgenic *Drosophila* system, the ICR is biallelically transcribed and produces sense and anti-sense RNA (Schoenfelder *et al.*, 2007). Analysis of reporter gene expression in transgenic *Drosophila* indicated that these ICR transcripts induce gene silencing in an RNAi-independent manner (Schoenfelder *et al.*, 2007), providing significant insight into the mechanism of mouse *H19/Igf2* imprinting. Transgenic *Drosophila* experiments also identified a 1.5 kb silencer element at the 3' end of the human *H19* ICR (Arney *et al.*, 2006). Similarly, a 740 bp sequence within the human Prader-Willi imprint centre functions as a silencer in *Drosophila* (Lyko *et al.*, 1998).

Overall these results indicate that several core epigenetic mechanisms underlying mammalian imprinting are highly conserved. The fact that the analyzed ICRs function as silencers but do not confer imprinting of marker genes in *Drosophila*, may indicate that a silenced epigenetic state is the default at these imprinted loci. That is, imprinting may function via gamete-specific exploitation of conserved epigenetic silencing processes. In the studies presented in this thesis, I set out to examine the evolutionary conservation of the molecular mechanisms underlying paramutation, an epigenetic process that similarly exhibits trans-generational epigenetic silencing.

1.3 PARAMUTATION

Paramutation is an epigenetic phenomenon that results in a meiotically stable change in expression of one allele after it has been paired with another. Several examples of paramutation exist in plants (Chandler and Stam, 2004), and it has also been described at the *Kit* locus in mice (Rassoulzadegan *et al.*, 2007). The maize *b1* locus provides one of the best characterized examples of paramutation. The *b1* gene encodes a transcription

factor that regulates expression of genes required for the synthesis of purple anthocyanin pigments (Chandler *et al.*, 1989). Changes in *b1* expression are easily detected; high levels of *b1* lead to darkly pigmented purple plants, while low levels result in lightly pigmented plants. Two alleles at the *b1* locus participate in paramutation: the highly transcribed *B-I* allele, and the weakly transcribed *B'* allele. The DNA sequence of the two alleles is identical (Patterson *et al.*, 1993), indicating that this difference in expression is accomplished via epigenetic mechanisms. The epigenetic status of the silenced *B'* allele is extremely stable, while that of the highly expressed *B-I* allele is not; *B-I* has been observed to spontaneously switch its epigenetic state, and convert to the silenced *B'* allele at a frequency of 1-10% (Coe, 1966; Patterson *et al.*, 1993).

Paramutation at the *b1* locus occurs in heterozygous plants, when a *B-I* allele is combined with a *B'* allele. The presence of the silenced *B'* allele in *trans* results in epigenetic silencing of the normally highly expressed *B-I* allele, and an absence of purple pigment in the heterozygous plants (Coe, 1966; Patterson *et al.*, 1993). One of the most intriguing features of *b1* paramutation is the stability and heritability of this epigenetic change. The change in the epigenetic status of *B-I*, from highly expressed to silenced, is meiotically stable, and is transmitted to the next generation of plants. Thus following heterozygosity with *B'*, the *B-I* allele is paramutated, or converted, to the silenced *B'* allele. The new *B'* allele (also termed *B'**) is functionally equivalent to *B'*, and is equally capable of paramutating another *B-I* allele in subsequent generations (Coe, 1966; Patterson *et al.*, 1993).

A control region located 100 kb upstream of the *b1* transcription start site is required for both paramutation and high expression of the *B-I* allele. At this genomic position, neutral *b1* alleles that do not participate in paramutation contain a single copy of

an 853 bp sequence that is otherwise unique in the maize genome. However, the two alleles that participate in paramutation, *B-I* and *B'*, each contain an identical 6 kb sequence consisting of seven tandem repeats of the 853 bp sequence (Stam *et al.*, 2002). Despite containing identical DNA sequences, the *B'* tandem repeats exhibit a closed chromatin structure, repressive histone modifications, and higher levels of DNA methylation, whereas the *B-I* tandem repeats have an open chromatin structure and histone H3 acetylation, an activating histone modification (Haring *et al.*, 2010; Stam *et al.*, 2002). A recombinant allele with five tandem repeats appears to function similarly to the endogenous alleles that have seven tandem repeats. However, a recombinant allele with only one repeat is not paramutagenic, and a recombinant allele with three tandem repeats exhibits impaired paramutation and a less stable methylation pattern at the control region, suggesting that paramutation strength increases as the number of repeats increases, with stable paramutation requiring a threshold number of repeats (Stam *et al.*, 2002). The requirement of multiple copies of the tandem repeat sequence for stable paramutation may be related to the concentration of differential epigenetic modifications at the repeat junctions (Haring *et al.*, 2010).

Long-distance interactions between the *b1* tandem repeats and the transcription start site have recently been detected, with notable differences between the two epialleles (Louwers *et al.*, 2009). The high-expressing *B-I* allele exhibits a higher frequency of chromatin interactions than *B'*, involving the transcription start site, the tandem repeats, and several additional upstream regulatory regions, suggesting the formation of a complex multi-loop structure that facilitates *b1* expression. In contrast, the weakly expressed *B'* allele exhibits less frequent interactions involving only the transcription start site and the tandem repeats, suggesting the formation of a less stable single-loop

structure. In the presence of a single tandem repeat, no interactions between the repeat sequence and transcription start site are detected, indicating that multiple tandem repeats are likely required for stable interaction between the *b1* repeats and transcriptional start site (Louwers *et al.*, 2009).

Paramutation at the *b1* locus also requires several proteins, including *mediator of paramutation 1*, or *mop1* (Dorweiler *et al.*, 2000), which encodes an RNA dependent RNA polymerase (Alleman *et al.*, 2006). MOP1 is required for maintenance of the silenced *B'* state, and *B'* plants homozygous for a *mop1* mutation have increased *b1* expression, resulting in purple pigmentation. Furthermore, in *mop1* mutant plants, *B-I* fails to paramutate to *B'* indicating that MOP1 is also required for the establishment of paramutation (Dorweiler *et al.*, 2000). Mutations in *mop1* also cause a substantial reduction in the overall level of 24 nt siRNAs (Nobuta *et al.*, 2008), and a loss of transcriptional silencing of Mutator transposons and transgenes (Lisch *et al.*, 2002; McGinnis *et al.*, 2006), suggesting that these epigenetic processes and paramutation share underlying silencing mechanisms.

Recent analysis has found that the *b1* tandem repeats are bound by a protein called CXC domain *b1*-repeat binding protein, or CBBP (Brzeska *et al.*, 2010). At least two CBBP binding sites are present per repeat, and CBBP was shown to form multimers, which could provide a mechanism of tandem repeat "counting". CBBP protein is involved in establishing silencing, but is not required to maintain the silenced state (Brzeska *et al.*, 2010). Additional genes required for *b1* paramutation and transcriptional repression of *B'* include *required to maintain repression 6* (*rmr6*) (Hollick *et al.*, 2005), which encodes the largest subunit of RNA polymerase IV (Erhard *et al.*, 2009), and *mediator of paramutation 2* (*mop2*, also known as *rmr7*), which encodes the second

largest subunit of both RNA polymerases IV and V (Sidorenko *et al.*, 2009). In *Arabidopsis*, these RNA polymerases participate in the production of siRNAs and non-coding RNAs, transcriptional gene silencing, silencing of transposons and repetitive DNA, RNA-directed DNA methylation, and heterochromatin formation (Herr *et al.*, 2005; Kanno *et al.*, 2005; Onodera *et al.*, 2005; Pikaard *et al.*, 2008; Pontier *et al.*, 2005; Wierzbicki *et al.*, 2008).

Examination of transcription from the maize *b1* tandem repeats revealed that the repeats are transcribed from both strands. Interestingly, the level of transcription is similar in *B-I* and *B'* plants, which both have seven tandem repeats, as well as in plants with a neutral allele that contains only a single copy of the repeat (Alleman *et al.*, 2006). 24 nt siRNAs from the tandem repeats have been detected in plants with *B-I*, *B'*, and the single-repeat allele, but are reduced in the presence of a *mop1* or *mop2* mutation (Chandler, 2007; Sidorenko *et al.*, 2009). It is thus likely that the bidirectional transcription of the repeats produces double stranded RNA (dsRNA), and that MOP1 and MOP2 are required to produce significant levels of siRNAs from the dsRNA molecules. Consistent with this, transcription of the *b1* tandem repeats was found to be unaffected in *mop2* mutants, despite the reduction in siRNAs (Sidorenko *et al.*, 2009). Interestingly, a unique *mop2* mutation that prevents paramutation and reduces siRNA production, but does not affect *B'* silencing, was isolated, suggesting that the primary role of the siRNAs is to mediate *trans* communication and establish the epigenetic states of the alleles, rather than maintain the epigenetic silencing of the *B'* allele (Sidorenko *et al.*, 2009). However, given that mutations in *mop1*, *rmr6,* and other *mop2* mutations cause a loss of *B'* silencing in addition to a loss of paramutation, the roles these proteins play in establishment and/or maintenance is not currently clear.

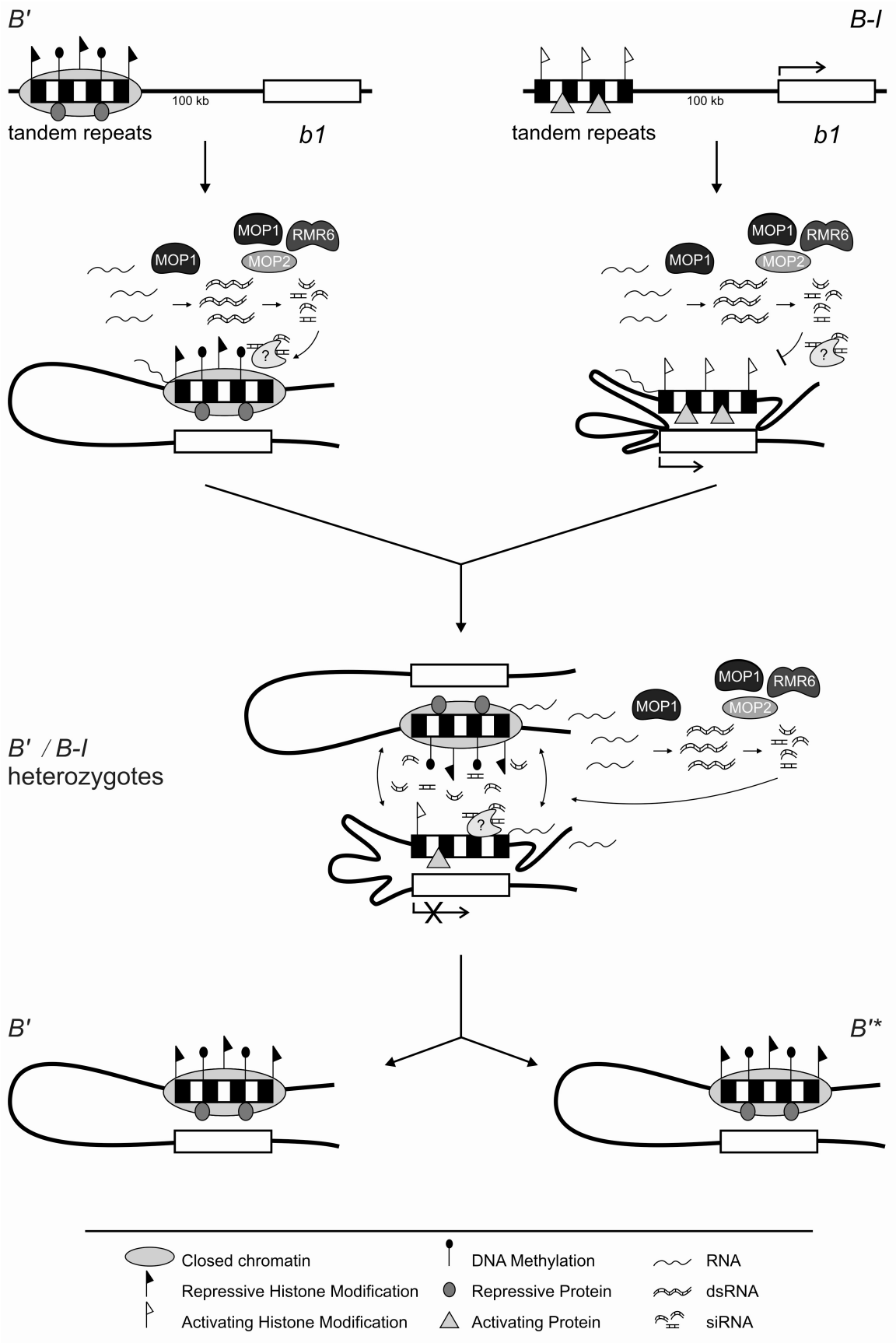
The current model for paramutation at the *b1* locus is that RNA-mediated communication between the *B-I* and *B'* alleles establishes the chromatin states of the control regions, which thereby determines the level of *b1* transcription (Figure 1.1; Arteaga-Vazquez and Chandler, 2010; Chandler, 2007). The open chromatin structure of the *B-I* tandem repeats, and the multi-chromatin loops that are formed at this allele, may promote *b1* transcription, whereas the closed chromatin structure of the *B'* tandem repeats and single chromatin loop may inhibit or prevent *b1* transcription. Importantly, however, the presence of siRNAs in the non-paramutating single-repeat allele suggests that the siRNAs alone are not sufficient to induce paramutation. The number of tandem repeats is also important, and may mediate or stabilize pairing-interactions between alleles, potentially via increased accumulation of proteins or chromatin modifications. In addition to RNA-mediated communication, interactions between the DNA sequences, proteins bound to the DNA, or the formation of higher-order protein complexes (such as CBBP multimers) may also play a role in paramutation by bringing the two alleles together physically, or localizing them to a particular nuclear compartment where silencing and a heritable chromatin state can be established by the siRNAs (Arteaga-Vazquez and Chandler, 2010).

Given that the highly expressed *B-I* allele contains seven tandem repeats that are transcribed and produce siRNAs, there is necessarily an additional mechanism that normally prevents silencing at this allele. The active chromatin structure of the repeats, or specific proteins that bind to the active chromatin structure, may inhibit the formation of the silenced epigenetic state, or the allele may be localized to a different nuclear environment that inhibits silencing (Arteaga-Vazquez and Chandler, 2010; Chandler, 2007). Alternatively, pre-existing repressive modifications at the *B'* allele may make it

susceptible to further siRNA-directed modifications (Haring *et al.*, 2010). This may be similar to the mechanism at the *Arabidopsis FWA* locus, where siRNAs direct DNA methylation at the tandem repeats of silenced alleles with pre-existing methylation, but not at active epialleles (Chan *et al.*, 2006). Spontaneous conversion of the *B-I* allele to the silenced *B'* state indicates that whatever the mechanism used to ensure high levels of transcription, this mechanism occasionally fails, causing this allele to adopt the default silenced epigenetic state (Chandler, 2007).

In order to analyze the conservation of epigenetic mechanisms underlying maize *b1* paramutation, I generated transgenic *Drosophila* carrying the seven maize *b1* tandem repeats adjacent to the *Drosophila white* and *eGFP* reporter genes. In this transgenic system, the *b1* repeats are located between two Flipase Recombinase Target (FRT) sequences, which mediate site-specific removal of the intervening sequences. This experimental design allows for the removal of the repeats by crossing to a source of Flipase recombinase (FLP), enabling analysis of reporter gene expression from transgenes with and without the *b1* repeats, at identical chromosomal positions. The results presented in the following chapters demonstrate that the maize *b1* repeats function as an epigenetic silencer, and a pairing-sensitive silencer, in transgenic *Drosophila*. Furthermore, the maize tandem repeats produce bidirectional transcripts in the transgenic *Drosophila* system. Modifiers of this silencing are examined and include PcG/trxG proteins. I also present a review on the conservation of a related epigenetic process, genomic imprinting, and a preliminary analysis of imprinted genes in *Drosophila*. Transgenic analysis of a conserved intracellular trafficking gene is also presented. The results presented herein highlight the extraordinary evolutionary conservation of eukaryotic epigenetic silencing.

Figure 1.1 Paramutation at the maize *b1* locus. The two alleles that participate in paramutation at the *b1* locus are identical in sequence; however, the *B-I* allele is highly transcribed while the *B'* allele is not. Paramutation requires a control region consisting of seven tandem repeats (white and black boxes), located 100 kb upstream of *b1*. While both *B-I* and *B'* contain identical tandem repeat sequences, the repeats exhibit epigenetic differences in chromatin structure, histone modifications, and DNA methylation, and may be associated with distinct proteins that maintain these epigenetic states. In *B'* plants, a single chromatin loop is formed between the repeats and *b1* transcription start site, while in *B-I* plants a more complicated multi-loop chromatin structure is formed, involving the *b1* tandem repeats, the transcription start site, and several upstream regulatory regions. The tandem repeats are bidirectionally transcribed in both *B-I* and *B'* plants, producing repeat RNA that then forms dsRNA and is processed into siRNAs. MOP1, an RNA-dependent RNA polymerase, may be involved in the production and amplification of the dsRNA, as well as the production of siRNAs, which likely also involves RMR6 and MOP2, components of RNA polymerases IV and V. The siRNAs are hypothesized to direct chromatin modifications at the tandem repeats via mechanisms and proteins that are currently unknown, but this process is blocked at the *B-I* allele, potentially by the active chromatin state, bound proteins, or nuclear environment. In heterozygous plants, the highly transcribed *B-I* allele is "paramutated", or converted, to the silenced *B'* state. siRNAs produced from the tandem repeats are hypothesized to mediate *trans*-interactions or communication between the alleles, as well as direct the establishment of a closed chromatin structure at the *B-I* tandem repeats. The conversion of *B-I* to a silenced epigenetic state is meiotically stable, and in the next generation all progeny will inherit a silenced *B'* allele. The newly paramutated allele is termed *B'**.


B'
B-I
100 kb
tandem repeats
b1
100 kb
tandem repeats
b1
MOP1
RMR6
MOP2
MOP1
MOP1
RMR6
MOP2
MOP1
?
?
B' / B-I
heterozygotes
MOP1
RMR6
MOP2
MOP1
?
B'
B'*
Closed chromatin
Repressive Histone Modification
Activating Histone Modification
DNA Methylation
Repressive Protein
Activating Protein
RNA
dsRNA
siRNA

1.4 TRANSITION TO CHAPTER 2

The following chapter details the generation of transgenic *Drosophila* carrying the maize *b1* tandem repeats, and characterization of the epigenetic silencing induced by these repeats. The maize tandem repeats were observed to cause epigenetic silencing of *white* in *cis* in all transgenic lines containing one to seven tandem repeats, with the strength of silencing increasing as the number of repeats increased. Evidence of *trans* interactions, dependent on repeat number, were also observed. These findings are significant, as they are the first to demonstrate that an epigenetic control sequence from plants is recognized and targeted for silencing in *Drosophila*, illustrating conservation of an epigenetic silencing process between the plant and animal kingdoms.

CHAPTER 2

THE MAIZE *B1* PARAMUTATION CONTROL REGION CAUSES EPIGENETIC SILENCING IN *DROSOPHILA*

Lori A. McEachern and Vett K. Lloyd

Manuscript prepared for submission to *Epigenetics*.

2.1 ABSTRACT

Paramutation is in epigenetic phenomenon in which a combination of alleles in a heterozygous organism results in a meiotically stable change in expression of one of the alleles. At the well-studied maize *b1* locus, paramutation requires a control region consisting of seven 853 bp tandem repeats. To study the conservation of the epigenetic mechanisms underlying maize *b1* paramutation, we created transgenic *Drosophila melanogaster* carrying the maize *b1* control region flanked by FRT sites and adjacent to the *Drosophila white* reporter gene. Here we show that the *b1* tandem repeats cause silencing of *white* in *Drosophila*. A single copy of the tandem repeat sequence is sufficient to cause silencing, and silencing strength increases as the number of tandem repeats increases. Additionally, transgenic lines with the full seven tandem repeats can demonstrate pairing-sensitive silencing and silencing in *trans*, while other lines demonstrate evidence of epigenetic activation in *trans*. These *trans* interactions are reduced or lost when the number of tandem repeats decreases. These results indicate that in *Drosophila*, the maize *b1* tandem repeats function similarly to the silenced *B'* allele in maize, and suggest that the epigenetic mechanisms underlying silencing and *trans*-interactions at the maize *b1* locus are evolutionary conserved.

2.2 INTRODUCTION

Epigenetic gene regulation is essential for the normal development of all eukaryotic organisms, and also results in a variety of unique epigenetic phenomena. Paramutation is an intriguing epigenetic phenomenon in which the epigenetic status and

expression level of one allele changes when it is paired with another. This process is distinguished from many other epigenetic phenomena by the fact that the newly established "paramutated" state is heritable through meiosis and is maintained in subsequent generations. Paramutation has been most extensively studied in plants, where it has been observed for several genes and transgenes in different plant species (Chandler and Stam, 2004), but it has also been described at the *Kit* locus in mice (Rassoulzadegan *et al.*, 2007), suggesting that paramutation and its underlying epigenetic mechanisms may be widespread.

One of the best characterized examples of paramutation is provided by the maize *b1* gene, which encodes a transcription factor that activates the anthocyanin biosynthetic pathway (Chandler *et al.*, 1989). Two alleles at the *b1* locus participate in paramutation: the highly transcribed *B-I* allele, and the weakly transcribed *B'* allele. The *B-I* and *B'* alleles are identical in DNA sequence (Patterson *et al.*, 1993; Stam *et al.*, 2002), yet exhibit very distinct expression patterns, indicating that they are epialleles that differ in their epigenetic status. While the silenced *B'* allele is extremely stable, the highly transcribed *B-I* allele is not, and spontaneously converts to *B'* at a frequency of 1 – 10% (Coe, 1966; Patterson *et al.*, 1993). Paramutation at the *b1* locus occurs when *B-I* and *B'* are paired in heterozygous plants. In the presence of the silenced *B'* allele, the normally highly expressed *B-I* is also silenced (Coe, 1966; Patterson *et al.*, 1993). This change in the epigenetic status of *B-I* is extremely stable, and the newly silenced allele is not only transmitted to the next generation of plants, but also has the ability to silence, or "paramutate", other active *B-I* alleles.

A control region required for both paramutation and expression of the *B-I* allele is located 100 kb upstream of the *b1* transcription start site. At this position, the two alleles

that participate in paramutation contain an identical 6 kb sequence consisting of seven 853 bp tandem repeats, whereas neutral *b1* alleles that do not participate in paramutation contain only a single copy of the 853 bp sequence (Stam *et al.*, 2002). Despite being identical in DNA sequence, the *B'* tandem repeats are marked by repressive histone modifications and exhibit a more closed chromatin structure and a higher level of DNA methylation than *B-I*, which exhibits an open chromatin structure with activating histone acetylation marks (Haring *et al.*, 2010; Stam *et al.*, 2002). Differences in physical interactions between the tandem repeats, the *b1* transcription start site, and additional upstream regulatory regions have also been detected for the *B'* and *B-I* alleles (Louwers *et al.*, 2009).

Transgenic organisms are a valuable tool for studying the molecular basis and evolutionary conservation of the epigenetic processes that act upon specific epigenetic control regions. The observation that epigenetic control regions essential for genomic imprinting in mammals function as silencers in transgenic *Drosophila* (Arney *et al.*, 2006; Lyko *et al.*, 1997; Lyko *et al.*, 1998) suggests that unique or species-specific epigenetic phenomena function by exploiting conserved core epigenetic mechanisms. Here we examine the conservation of epigenetic mechanisms between the plant and animal kingdoms by creating transgenic *Drosophila* carrying the maize *b1* paramutation control region adjacent to reporter genes. We show that the maize *b1* paramutation control region functions as a silencer in transgenic *Drosophila*. In addition, the control region also mediates *trans*-interactions between homozygous transgenes, paralleling its endogenous role in mediating allelic interactions during paramutation. The conservation of the *cis* and *trans*-silencing functions of the *b1* tandem repeats in *Drosophila* shows remarkable fidelity of an epigenetic control region between the plant and animal

kingdoms, and suggests that the epigenetic mechanisms underlying paramutation are evolutionarily conserved.

2.3 MATERIALS AND METHODS

2.3.1 *DROSOPHILA* CULTURE

All fly stocks were maintained at 21 ±3°C on a standard cornmeal, yeast, and sugar medium supplemented with 0.15% methylbenzoate (Sigma) as a mould inhibitor. All crosses were conducted in vials with 5 – 10 females and 3 – 5 males. Crosses were transferred to fresh food after 3 – 5 days, up to three times, before the parents were discarded. All standard stocks used are described in FlyBase (Tweedie *et al.*, 2009) and were provided by the Bloomington *Drosophila* Stock Center.

2.3.2 RECOMBINANT PLASMIDS

The *P{FRT-RPTS-FRT, hsp70-white, hsp70-eGFP}* P-element vector was constructed by subcloning a 6856 bp fragment containing the maize repeats from the *p{MS14}* vector (a generous gift from V. Chandler, University of Arizona) into the *KpnI* site of *pP{WhiteOut2}* (provided by J. Sekelsky, University of North Carolina at Chapel Hill), between two Flipase recombinase target (FRT) sites. The cytoplasmic *eGFP* gene was amplified from *pP{Green-H-Pelican}* (*Drosophila* Genomics Resource Centre) using the primers 5'-ATCGTTCGAAGAGCGCCGGAGTATAAATAG-3' and 5'-CCGCTTCGAATTTACGCCTTAAGATACATTG-3', and was subcloned into the *BstBI* site of the *pP{whiteOut2}* vector, following the *white* cDNA sequence. A 591 bp *EcoRI* fragment located between the last maize repeat and the FRT site was subsequently

removed to decrease the distance between the maize repeats and the *white* gene. The final *P{FRT-RPTS-FRT, hsp70-white, hsp70-eGFP}*, plasmid contains the *b1* maize repeats between two FRT sites, with the *Drosophila white* and cytoplasmic *eGFP* genes located downstream, outside of the FRT sites (Figure 2.1). For brevity, this vector will be referred to as *P{RPT}*.

2.3.3 TRANSGENIC FLIES

Transgenic *Drosophila* were generated by microinjection into a w^{1118}; {Δ2-3}99B host stock (Robertson *et al.*, 1988), using standard methods (Spradling and Rubin, 1982). Additional transgenic lines were generated by crossing to a stable source of Δ2-3 transposase and isolating flies with changed eye pigmentation. The chromosomal insertion site of the transgenic lines was determined by crossing to a *w*; *dp*; *e* marker stock. X chromosome inserts were maintained as homozygous stocks, while 2nd and 3rd chromosome inserts were balanced over *CyO* and *TM3, Sb^1* respectively. Lines were named according to the number of repeats present and the insertion site.

Lines with fewer than seven repeats were obtained via tandem repeat loss, either in the bacterial vector host prior to microinjection (*P{3RPT}10D6* and *P{2RPT}89B9*), during DNA replication following integration (*P{3RPT}17C7*), or during P-element mobilization of an integrated transgene (*P{1RPT}9D3*). Line *P{3RPT}17C7* has been isolated from line *P{7RPT}17C7* since the initial recovery of transgenic flies following microinjection, but the fact that the two transgenes are at the identical integration site suggests that *P{3RPT}17C7* was initially a copy of *P{7RPT}17C7*, and subsequently lost four tandem repeats during stock propagation. Line *P{2RPT}49E4* was obtained by P-element mobilization of *P{2RPT}89B9*.

The maize repeats were removed from the transgenic inserts by crossing adult females from the *P{RPT}* lines to males from a w^{1118}*; P{70FLP}7* stock, which expresses Flipase recombinase (FLP) under control of a *hsp70* promoter. Progeny from the cross were heat-shocked at 37°C for 40 minutes at day 3 – 6 of development to stimulate FLP expression and removal of the FRT-flanked maize repeats. Eye pigmentation of F1 adult flies that inherited both the *P{RPT}* and *P{70FLP}* transgenes was compared with that of sibling flies that inherited the *P{RPT}* transgene only. Flies with the *P{RPT}* and *P{70FLP}* transgenes were backcrossed to a w^{1118} stock to remove the *P{70FLP}* transgene by segregation, and obtain stable transgenic lines containing *hsp70-white* and *hsp70-eGFP* at the same insert location, but without the maize repeats. These lines are termed *P{0RPT}*.

2.3.4 Analysis of transgene insertion sites and the number of tandem repeats

The integration sites of the transgenes were determined by inverse PCR (Huang *et al.*, 2000). Genomic DNA was digested with *Msp*I (Fermentas), ligated with T4 DNA Ligase (Fermentas), and PCR amplified using primers specific to either the 5' or 3' end of the vector (supplementary Table 2.S1). PCR products were visualized on a 1% agarose gel, and strong bands were purified by either PEG precipitation or with a gel-band extraction kit (Biobasic) and were sequenced directly, while weak bands were cloned into pDrive cloning vector (Qiagen) and sequenced. DNA sequencing was conducted at The Centre for Applied Genomics (The Hospital for Sick Children, Toronto, Canada).

The presence or absence of the maize repeats, as well as the number of repeats present, was confirmed by PCR analysis. Long PCR Enzyme Mix (Fermentas) was used

according to the manufacturer's protocol with the primers 5'-GCAAGAGACATCCACTTAACG-3' and 5'-GTGAGAGAGCAATAGTACAGAGAGG-3', which bind outside of the maize repeats and FRT sites. The number of repeats was determined by the size of the resulting band: lines with the full seven repeats exhibited a 6767 bp band, and in lines with fewer repeats, this band size decreased by 853 bp for each repeat missing. The removal of the maize repeats by FLP was confirmed by the presence of a 366 bp band with the above primer combination and the absence of a PCR product when tested with primers specific to the repeat sequence.

2.3.5 Assessment of eye and *eGFP* phenotypes

Eyes of adult males or females were photographed at 3 – 6 days post-eclosion, using a Zeiss Stemi 2000-C microscope and Sony DSC-S70 Camera. Eye pigmentation was quantified using an assay modified from Khesin and Leibovitch (1978). For each genotype analyzed, 3- to 6-day-old adult males and females were isolated separately. Flies exhibiting eye pigmentation drastically different from the majority, likely due to aberrant stimulation of the *hsp70* promoter, were excluded from analysis. Heads were removed by brief freezing at -70°C followed by agitation. Four replicate tubes containing 10 heads each were prepared, and the heads were manually homogenized in 150 µL of 10 mM HCl in 100% ethanol. The homogenate was kept at 4°C overnight, heated to 50°C for 5 minutes, and centrifuged for 10 minutes. The supernatant was transferred to a new tube, and the absorbance at 480 nm was measured in an Amersham Pharmacia Biotech Ultraspec 2000. Mean values and standard deviations were calculated, and significance

was determined by a one-way ANOVA and subsequent two-tailed unpaired Student's t-tests.

To measure pairing-sensitive silencing in the *P{7RPT}86B2* line, homozygous flies with light eye colour were selected from the balanced stock. To analyze eye pigmentation of hemizygous flies in lines containing the transgene on the X chromosome, homozygous transgenic females were crossed to w^{1118}/Y males and hemizygous F1 flies were selected. To analyze eye pigmentation of heterozygous flies (repeats-in / repeats-out), balanced hemizygous (2^{nd} and 3^{rd} chromosome inserts) or homozygous (X chromosome inserts) females containing the maize repeats were crossed to hemizygous males with the same transgene, but with the repeats removed. All images and pigment assays presented herein are of female flies, with the exception of the hemizygous males presented in Figure 2.3. These males were isolated from their respective homozygous stocks, but are hemizygous due to the location of this transgenic insert on the X chromosome.

The *eGFP* reporter gene was observed to be very poorly expressed in most of the transgenic lines, with the level of fluorescence in most larvae similar to background levels, and so was excluded from further analysis.

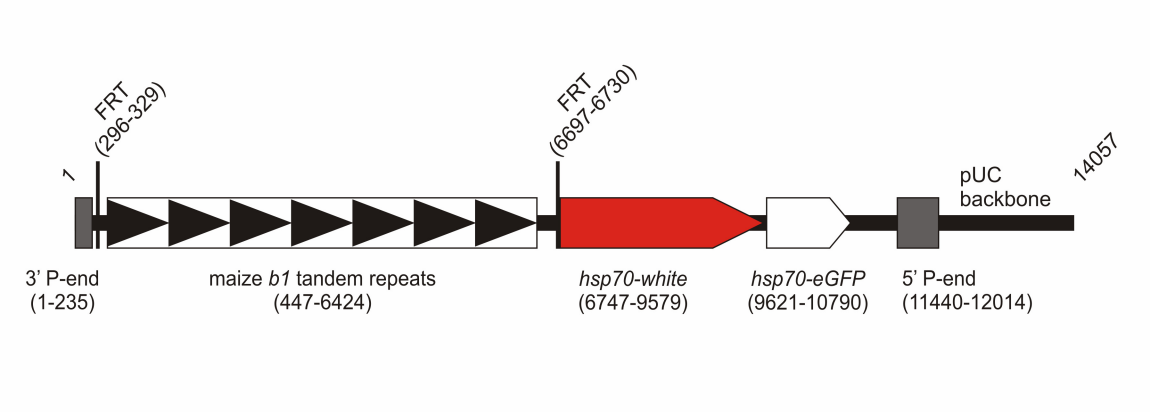


Figure 2.1 The P-element construct used to generate transgenic *Drosophila* with the maize *b1* tandem repeats. The maize *b1* repeats are inserted between two FRT sequences. The *hsp70-white* and *hsp70-eGFP* reporter genes are located adjacent to the repeats, but outside of the FRT sites. Crossing to a source of FLP enzyme mediates removal of the maize *b1* repeats, while leaving the reporter genes intact at the transgene insertion site.

2.4 RESULTS

2.4.1 MAIZE *B1* TANDEM REPEATS CAUSE SILENCING IN *DROSOPHILA*

In order to examine the evolutionary conservation of epigenetic mechanisms between the plant and animal kingdoms, we created transgenic *Drosophila* carrying the maize *b1* tandem repeat control region between two FRT sites, and adjacent to the *Drosophila white* eye colour reporter gene (Figure 2.1). Through initial microinjection and subsequent mobilization of transgene inserts, 11 transgenic lines were obtained, including six lines with seven tandem repeats and five lines with one, two, or three tandem repeats (Table 2.1). Removal of the maize repeats allowed for a comparison of reporter gene expression at identical chromosomal positions, with and without the *b1* repeats present. If the epigenetic mechanisms underlying paramutation-based silencing at the maize *b1* locus are evolutionary conserved, then the *b1* epigenetic control region should silence adjacent reporter genes in transgenic *Drosophila*, and removal of the *b1* tandem repeats should result in an increase in reporter gene expression.

Transgenic *Drosophila* lines containing the full seven maize *b1* tandem repeats exhibited silencing of the *white* reporter gene, as demonstrated by crossing the transgenic flies to a source of FLP enzyme, which excises FRT-flanked DNA sequences by catalyzing recombination between the FRT sites (Golic and Lindquist, 1989). Removal of the maize repeats to produce "repeats-out" strains consistently resulted in an increase in *white* expression and a darker eye phenotype, compared with sibling flies with the same transgene, but with the repeats present ("repeats-in"; Figure 2.2A). Heat-shock driven expression of FLP early in larval development, or targeted removal of the repeats in the eyes without heat-shocking by crossing to an FLP source driven by an eye-specific

enhancer (*P{ey-FLP.B}*), resulted in uniform dark eyes in all flies containing both the *P{7RPT}* transgene and FLP enzyme. Similarly, FLP expression and repeat-removal later in larval development frequently caused a variegated eye phenotype, with the dark patches of eye pigmentation presumably representing cell lineages from which the repeats have been removed (Figure 2.2A).

Transgenic lines with fewer tandem repeats also exhibited an increase in eye pigmentation following removal of the maize *b1* tandem repeats, again indicative of repeat-mediated silencing. Lines with one or three tandem repeats exhibited increased *white* expression, observed as a darker eye colour, following removal of the repeat sequences by FLP-mediated recombination (Figure 2.2B). The same effect of increased *white* expression following removal of the maize *b1* tandem repeats was observed for all transgenic lines containing one to seven tandem repeats (Table 2.1). Thus, the maize *b1* tandem repeats consistently mediated epigenetic silencing in transgenic *Drosophila*.

2.4.2 REPEAT NUMBER DETERMINES *CIS* SILENCING STRENGTH

Analysis of hemizygous males with transgenes at the same genomic position on the X chromosome, but with seven, three, or zero *b1* tandem repeats present, revealed that the strength of silencing increases as the number of repeats increases (Figure 2.3). Line *P{3RPT}17C7* is a stable derivative line with three tandem repeats, at the same genomic position as *P{7RPT}17C7*, which has seven tandem repeats. Epigenetic silencing of *white* was observed with both three and seven *b1* tandem repeats. However, males with seven *b1* tandem repeats have lighter eyes (Figure 2.3A), and less than half the amount of pigment (Figure 2.3B), as males with three *b1* tandem repeats. *white* expression is greatest, and eye colour is darkest, when zero tandem repeats are present upstream of

white. Thus, the greater the number of maize repeats, the more severe the epigenetic silencing of *white*.

2.4.3 *TRANS* SILENCING EFFECTS

In maize *b1* paramutation, the active *B-I* allele is heritably silenced when it is heterozygous with the silenced *B'* allele, via a process that is hypothesized to include *trans* communication between the tandem repeat control regions of the two alleles (Arteaga-Vazquez and Chandler, 2010; Stam, 2009). Pairing-sensitive silencing is a type of *trans* interaction in which silencing is enhanced in homozygotes due to pairing of epigenetic silencing sequences. Pairing-sensitive silencing of *white* in *Drosophila* is observed as a lighter eye colour in homozygous flies with two paired copies of a *white*-containing transgene, compared with hemizygous flies with one unpaired copy of the same transgene. This is in contrast to normal additive *white* expression, in which homozygous flies are expected to have darker eyes than hemizygous flies, due to the increased dosage of the *white* reporter gene.

Pairing-sensitive silencing was observed for four of the six transgenic lines with seven *b1* tandem repeats. In these lines, *white* expression is reduced and eye colour is lighter when the transgene insert is homozygous (Figure 2.4A, column 2) compared with when it is hemizygous (Figure 2.4A, column 1). Pairing-sensitive silencing is highly consistent (i.e. observed in all homozygotes) in all transgenic lines that exhibit this effect, except for line *P{7RPT}86B2*, where it is observed in only 30-50% of homozygous flies. In the tandem repeat lines that exhibited pairing-sensitive silencing, *white* expression decreased by approximately 1.4 fold in flies homozygous for the maize repeats and *white* reporter, compared with hemizygous flies with a single copy of the maize repeats and

white reporter gene (Figure 2.4B). Following removal of the repeats, pairing-sensitive silencing was generally lost (Figure 2.4A and B). However, line *P{7RPT}44D4* has continued to exhibit strong pairing-sensitive silencing of the *white* gene in the absence of the maize repeats for more than three years.

Additional evidence of repeat-mediated silencing in *trans* was obtained by combining a repeats-out transgene with its progenitor seven-repeat transgene at the same genomic position, to generate repeats-in / repeats-out heterozygous flies. This type of *trans* silencing differs from pairing-sensitive silencing, in that, while the *white* reporter genes are present in paired transgenes on homologous chromosomes, the maize *b1* tandem repeats, which mediate epigenetic silencing, are present on one homologue only. Thus, any decrease in *white* marker gene expression, compared with repeats-out hemizygous flies, would indicate silencing of the repeats-out *white* gene in *trans* by the repeats-in allele.

Eye pigmentation analysis indicates that the maize repeats can mediate silencing of *white* in trans. Heterozygous flies frequently exhibited eye pigmentation that was approximately equivalent to, or lighter than, hemizygous flies with a single copy of the repeats-out transgene, despite having two copies of *white* and only a single copy of the seven *b1* tandem repeats (Figure 2.4B and C). Repeat-mediated silencing of the *white* reporter gene in *trans* was especially evident for the *P{7RPT}86B2* transgene. In this line, combining the repeats-in transgene (pigment level = 0.053) with the repeats-out transgene (pigment level = 0.232), resulted in a dramatic reduction in *white* expression in heterozygous flies (pigment level = 0.133). Thus the maize tandem repeats can silence genes that are present in *trans* on the paired chromosomal homologue.

Combining two transgenes with seven tandem repeats, but inserted at different chromosomal positions, did not result in enhanced silencing of the *white* reporter gene (Figure 2.4D). Instead, these flies exhibited a darker eye colour, indicative of increased *white* expression. The observed *trans* silencing effects are thus dependent on homologous pairing of the transgenes, and not the cumulative number of the transgenes.

2.4.4 REPEAT NUMBER DETERMINES *TRANS* SILENCING STRENGTH

Given the observation that *cis* silencing strength is affected by the number of tandem repeats (Figure 2.3), we analyzed the effect of repeat number on silencing in *trans*. Pairing-sensitive silencing of homozygous transgenes, and silencing of *white* in *trans* in heterozygous repeats-in / repeats-out flies, were examined in females with the transgene at the same insertion site (17C7), but with seven, three, or zero tandem repeats present. Strong pairing-sensitive silencing was observed for this transgene insertion site with seven tandem repeats, and pairing-sensitive silencing was lost when the maize repeats were removed (Figure 2.5A and B). Thus, the tandem repeats are required for pairing-sensitive silencing. Eye pigmentation in the line with three tandem repeats is variable. However, *P{3RPT}17C7* homozygous females are generally darker than, or equivalent to, hemizygous females, suggesting that pairing-sensitive silencing is impaired when there are only three tandem repeats present. Pigment quantification agrees with this assessment. Unlike the pairing-sensitive silencing that is observed when seven tandem repeats are present, homozygous females with three tandem repeats exhibit no decrease in eye pigmentation compared with hemizygous females (Figure 2.5B). This result therefore suggests that the strength of pairing-sensitive silencing increases as the number

of tandem repeats increases. Further, pairing-sensitive silencing may require a threshold number of repeats.

Analysis of eye pigmentation in repeats-in / repeats-out heterozygous females provides additional evidence that silencing in *trans* occurs with seven tandem repeats, but is impaired with only three tandem repeats. When the seven-repeat transgene was heterozygous with the repeats-out transgene, females exhibited eye pigmentation that was approximately equivalent to repeats-out hemizygous females (Figure 2.5B and C). The fact that *white* expression did not increase despite the presence of an additional *white* reporter gene suggests that the seven tandem repeats exert a silencing effect on the *white* reporter genes on both homologues. In contrast, when there are only three repeats present, combining the repeats-in transgene with the repeats-out transgene resulted in a significant increase in eye pigmentation compared with repeats-out hemizygous females (Figure 2.5B and C). The transgene containing seven tandem repeats therefore demonstrates a greater capacity for silencing in *trans* than the transgene containing three tandem repeats.

Consistent with the observation that pairing-sensitive silencing is reduced or lost when the number of tandem repeats is reduced from seven to three, three additional transgenic lines with 1 – 3 tandem repeats do not exhibit pairing-sensitive silencing. In these lines, homozygous flies have darker eye pigmentation than hemizygous flies, both with the repeats present and following repeat removal (Figure 2.6A). *white* expression in lines with 1 – 3 maize *b1* tandem repeats was 1.4 – 2.2 fold higher in homozygous flies than in hemizygous flies, both with the repeats present and with the repeats removed (Figure 2.6B). Additionally, in the lines with one and two tandem repeats, heterozygous repeats-in / repeats-out flies exhibited eye pigmentation intermediate to that of the

repeats-out hemizygous and homozygous flies, as would be expected for normal additive *white* expression (Figure 2.6B and C). For example, in line *P{1RPT}9D3*, combining the repeats-in transgene, which has a pigment level of 0.012, with the repeats-out transgene, which has a pigment level of 0.029, resulted in heterozygous flies with an additive pigment level of 0.042.

2.4.5 EPIGENETIC ACTIVATION IN *TRANS*

Heterozygous repeats-in / repeats-out flies with unusually dark eyes that are not consistent with normal additive *white* expression patterns, may be indicative of epigenetic activation in *trans.* This type of *trans* activation effect was observed in line *P{3RPT}10D6*, which contains three tandem repeats, and produces heterozygotes with darker eye pigmentation than repeats-out homozygous flies (Figure 2.6). Further evidence that the maize *b1* tandem repeats can mediate epigenetic activation in *trans* is obtained via examination of two additional transgenic lines with seven tandem repeats (Figure 2.7). In these seven-repeat lines, *cis*-silencing of *white* by the maize *b1* tandem repeats was observed, as eye colour and *white* expression increased following repeat removal. However, pairing-sensitive silencing was not observed, as homozygous flies have darker eye pigmentation (Figure 2.7A) and increased *white* expression (Figure 2.7B) compared with hemizygous flies, both with seven *b1* tandem repeats and following repeat removal. Surprisingly, similar to *P{3RPT}10D6*, at these two genomic positions heterozygous repeats-in / repeats-out flies had darker eye pigmentation and a higher level of *white* expression than even the repeats-out homozygotes, which have two copies of *white* and no *b1* tandem repeats (Figure 2.7B and C). For example, in line *P{7RPT}8B6*, combining the repeats-in transgene (pigment level = 0.016) with the repeats-out transgene

(pigment level = 0.059) resulted in heterozygous flies with a drastic increase in *white* expression that is not consistent with an additive *white* expression pattern (pigment level = 0.142). This result implies that in some genomic contexts, the maize repeats have the capacity for epigenetic activation in *trans*.

In summary, our results show that the maize *b1* tandem repeats function analogously in *Drosophila* and maize. In transgenic *Drosophila,* the repeats silence adjacent genes both on the same chromosome, and in *trans* at the same position on the homologous chromosome. The number of tandem repeats is important in determining the strength of silencing, both in *cis* and in *trans*. In addition, evidence suggests the maize repeats may be targeted by both activating and repressive epigenetic forces in *Drosophila*.

Table 2.1 Insertion sites of the *b1* tandem repeat transgenes, and the observed effect on *white* expression.

Line [a]	Chromosome Arm	Chromosome coordinate [b]	Number of Repeats	Silencing [c]	Pairing-sensitive silencing [c]
P{7RPT}8B6	X	8788229 (−)	7	Y	N
P{7RPT}12C1	X	13656788 (−)	7	Y	Y
P{7RPT}17C7	X	18542227 (−)	7	Y	Y
P{7RPT}37A2	2L	18743079 (+)	7	Y	N
P{7RPT}44D4	2R	4487944 (−)	7	Y	Y
P{7RPT}86B2	3R	6209635 (+)	7	Y	Y
P{3RPT}10D6	X	11608910 (+)	3	Y	N
P{3RPT}17C7	X	18542227 (−)	3	Y	N / =
P{2RPT}49E4	2R	8844629 (+)	2	Y	N
P{2RPT}89B9	3R (*TM6C*)	12075778 (−)	2	Y	nv
P{1RPT}9D3	X	10441735 (+)	1	Y	N

[a] Lines were named according to the number of repeats present and the insertion site, e.g. *P{7RPT}17C7* has 7 repeats and is integrated into cytoband 17C7.

[b] First base off 5' P-end, with + or − indicating the orientation according to the 5' and 3' P-element ends.

[c] Y, Yes; N, No; N / =, hemizygotes lighter than or equivalent to homozygotes; nv, homozygotes not viable.

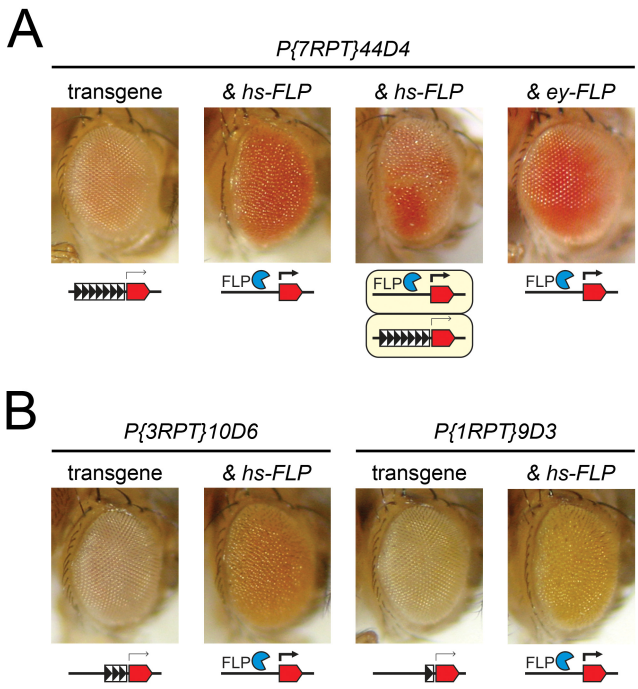


Figure 2.2 The maize *b1* tandem repeats cause silencing of the *white* reporter gene in *Drosophila*. **[A]** A representative transgenic line with the full seven tandem repeats exhibits silencing of *white*, as demonstrated by an increase in *white* expression, producing a darker eye colour, when the repeats are removed by FLP enzyme. Left to right: Transgene only; transgene and *hs-FLP* (heat-shock early in development); transgene and *hs-FLP* (heat-shock later in development); transgene and *ey-FLP*. **[B]** Lines with one or three tandem repeats also exhibit silencing of *white*, illustrated by a darker eye colour when the transgenic lines are crossed to a source of *hs-FLP* enzyme and heat-shocked early in larval development to remove the repeat sequences. Left to right: Transgene only (three repeats); transgene and *hs-FLP*; transgene only (one repeat); transgene and *hs-FLP*.

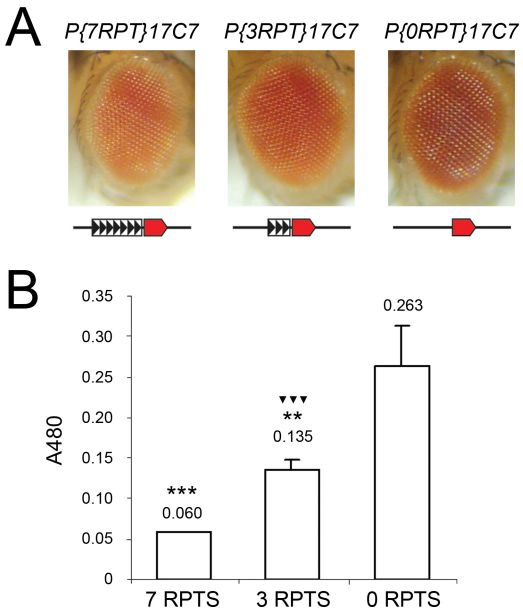
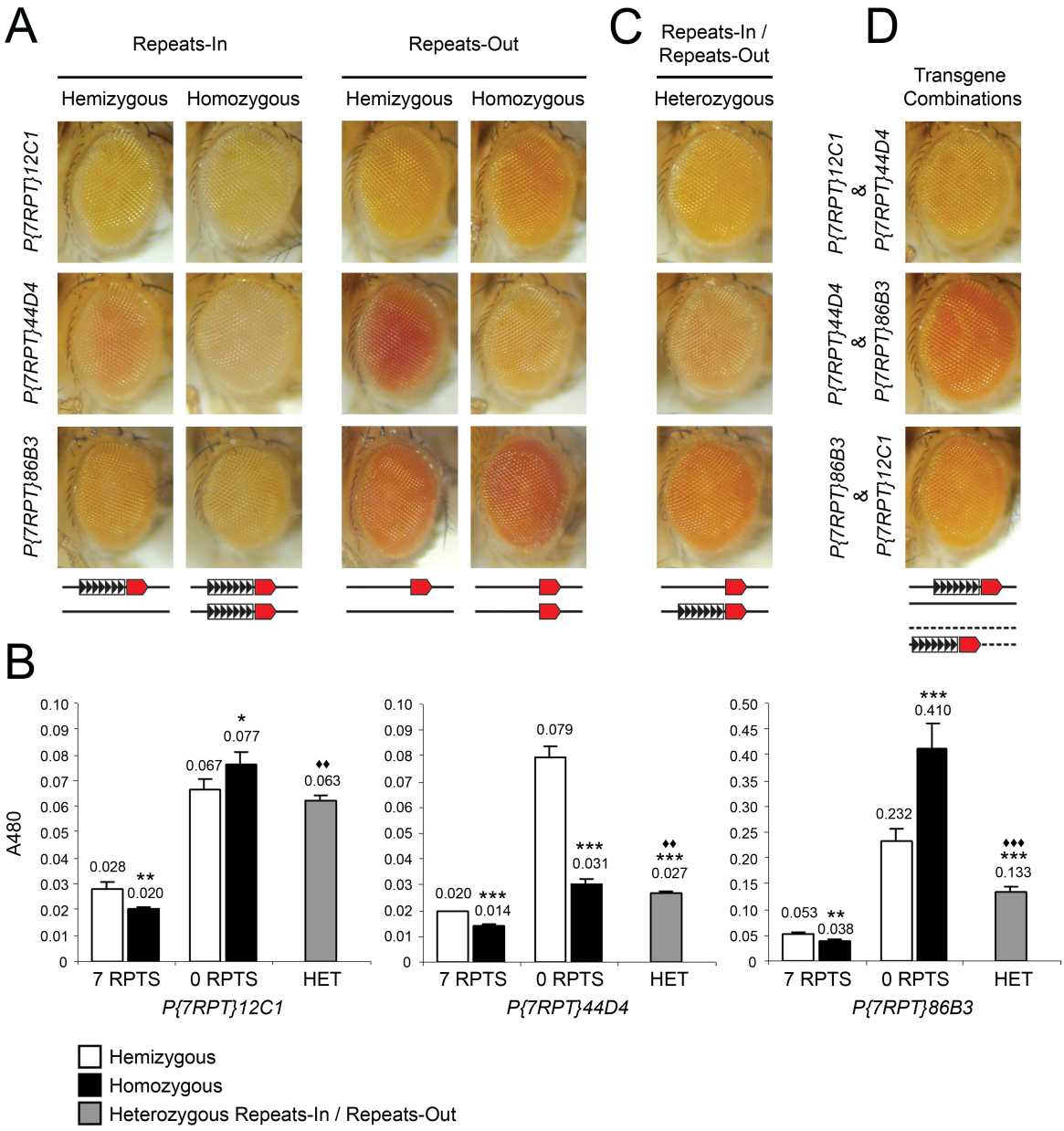


Figure 2.3 Silencing strength increases as the number of repeats increases. **[A]** For transgenes at the same genomic position, eye pigmentation of hemizygous males is lightest with seven tandem repeats adjacent to the *white* reporter gene, darker with three tandem repeats present, and darkest with zero tandem repeats. **[B]** Pigment assay quantification of hemizygous males demonstrates that *white* expression increases as the number of repeats decreases. Statistical significance is indicated for "repeats-in" values significantly different from "repeats-out" (*), and 3 RPTS values significantly different from 7 RPTS (▾). Significances are indicated at the level of $p<0.05$ (*), $p<0.01$ (**) or $p<0.001$ (***), for each symbol.

Figure 2.4 Lines with seven tandem repeats can exhibit pairing-sensitive silencing, in addition to *cis*-silencing, of the *white* reporter gene. **[A]** Pairing-sensitive silencing is observed in several transgenic lines with seven tandem repeats, illustrated by lighter eye pigmentation (reduced *white* expression) in repeats-in homozygous flies compared with hemizygous flies. Pairing-sensitive silencing is generally lost when the maize repeats are removed, but line *P{7RPT}44D4* continues to exhibit pairing-sensitive silencing in the absence of the maize *b1* tandem repeats. **[B]** Pigment assay quantification of *white* expression in hemizygous, homozygous, and heterozygous ("HET") flies, with the repeats present ("7 RPTS") and following repeat removal ("0 RPTS"), for the three transgenic lines illustrated in [A]. All pigment assay measurements of flies with the seven tandem repeats present are significantly reduced compared with those with the repeats removed ($p<0.001$). Additional statistical significance is indicated for homozygous values significantly different than hemizygous (*), and heterozygous values significantly different from repeats out hemizygous (*) or homozygous (♦). Significances are indicated at the level of $p<0.05$ (*), $p<0.01$ (**) or $p<0.001$ (***), for each symbol. **[C]** Heterozygous repeats-in / repeats-out flies exhibit eye pigmentation equal to or lighter than hemizygous repeats-out flies, indicating that the maize repeats can also silence repeats-out *white* in *trans*. **[D]** Combining transgenes with seven tandem repeats but at different chromosomal positions does not result in enhanced silencing. Rather, eye pigmentation is increased compared with the repeats-in hemizygotes, indicative of increased *white* expression.


A
Repeats-In
Repeats-Out
Hemizygous
Homozygous
Hemizygous
Homozygous
C
Repeats-In /
Repeats-Out
Heterozygous
D
Transgene
Combinations
P{7RPT}12C1
P{7RPT}44D4
P{7RPT}86B3
P{7RPT}12C1
&
P{7RPT}44D4
P{7RPT}44D4
&
P{7RPT}86B3
P{7RPT}86B3
&
P{7RPT}12C1
B
A480
0.10
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0.028
**
0.020
0.067
*
0.077
♦♦
0.063
7 RPTS
0 RPTS
HET
P{7RPT}12C1
0.020
***
0.014
0.079
***
0.031
♦♦
***
0.027
P{7RPT}44D4
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0
0.053
**
0.038
0.232
***
0.410
♦♦♦
***
0.133
P{7RPT}86B3
Hemizygous
Homozygous
Heterozygous Repeats-In / Repeats-Out

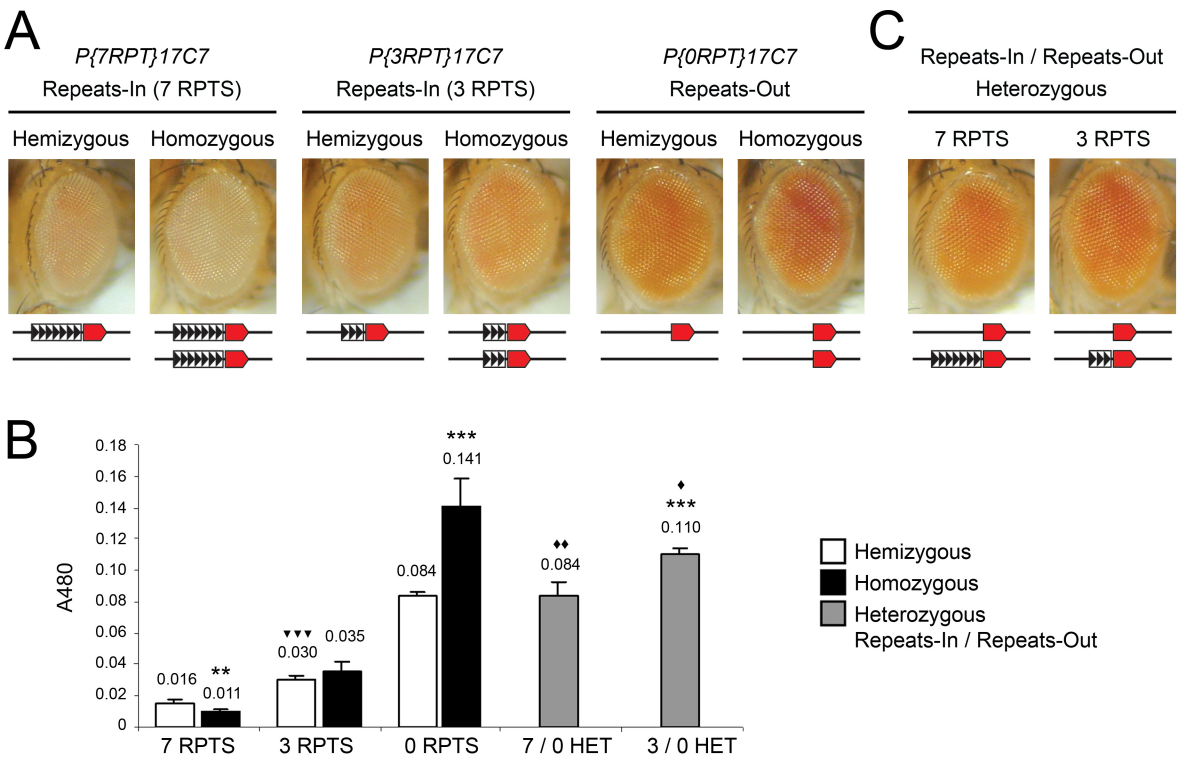
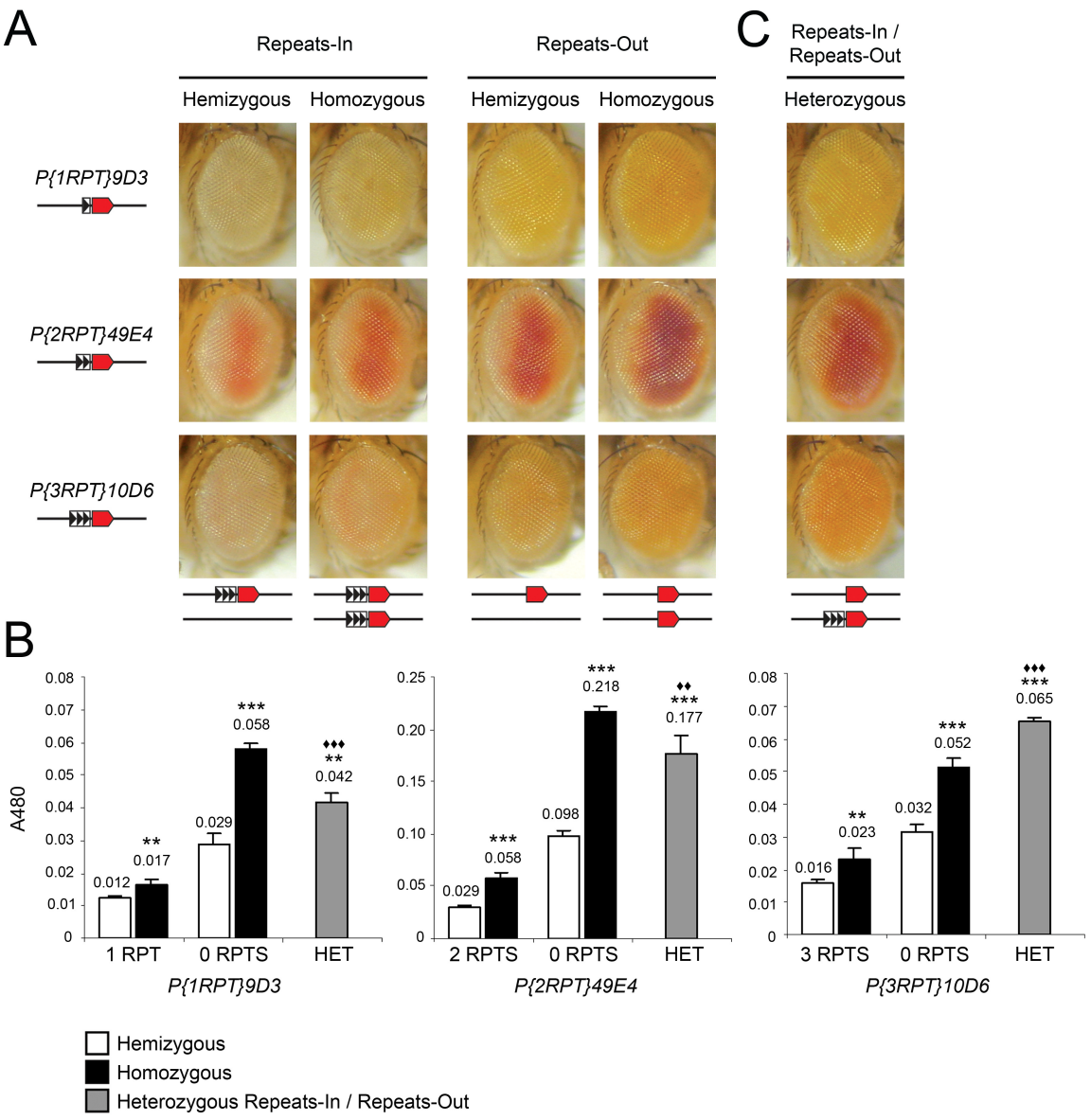


Figure 2.5 Silencing strength in *trans* increases as the number of repeats increases. **[A]** For transgenes at the same genomic position, pairing-sensitive silencing is observed when seven tandem repeats are present, and is lost when the repeats are removed. When three tandem repeats are present, hemizygous females are lighter than, or equivalent to, homozygous females. **[B]** Pigment assay quantification of hemizygous, homozygous, and heterozygous females with seven, three, and zero tandem repeats confirms that pairing-sensitive silencing and heterozygous *trans*-silencing are strongest with the seven tandem repeat transgene, and are reduced or lost when the number of repeats decreases to three and zero. All differences between repeats-in and repeats-out flies are statistically significant, with $p<0.001$. Additional statistical significance is indicated for homozygous values significantly different than hemizygous (*), hemizygous 3 RPTS values significantly different than hemizygous 7 RPTS (▼), and heterozygous values significantly different from zero repeats hemizygous (*) or homozygous (♦). Significances are indicated at the level of $p<0.05$ (*), $p<0.01$ (**) or $p<0.001$ (***), for each symbol. **[C]** Eye pigmentation in heterozygous repeats-in / repeats-out flies is approximately equal to repeats-out hemizygous flies with the seven-repeat transgene in *trans*, but is increased with the three-repeat transgene in *trans*, indicating that the seven-repeat transgene has a greater capacity for silencing in *trans*.

Figure 2.6 Lines with 1 – 3 tandem repeats exhibit silencing, but not pairing-sensitive silencing, of the *white* reporter gene. **[A]** Flies with 1 – 3 tandem repeats exhibit darker eye pigmentation following repeat removal, indicating *cis*-silencing of the white reporter gene. Homozygotes have darker eyes than hemizygotes, indicating a lack of pairing-sensitive silencing. Schematics beneath the eyes are illustrated using the *P{3RPT}* transgene as an example. **[B]** Pigment assay quantification of *white* expression in hemizygous, homozygous, and heterozygous ("HET") flies, with 1 – 3 tandem repeats present and following repeat removal ("0 RPTS"), for the three transgenic lines illustrated in [A]. All pigment assay measurements of flies containing the maize *b1* tandem repeats are significantly reduced compared with those with the repeats removed ($p<0.001$). Additional statistical significance is indicated for homozygous values significantly different than hemizygous (*), and heterozygous values significantly different from zero repeats hemizygous (*) or homozygous (♦). Significances are indicated at the level of $p<0.05$ (*), $p<0.01$ (**) or $p<0.001$ (***), for each symbol. **[C]** Heterozygous eye colour is intermediate to repeats-out hemizygotes and homozygotes, as expected, for transgenic lines with one or two tandem repeats, but is darker than repeats-out homozygotes for the three tandem repeat line, indicating potential activation in *trans*.


A
Repeats-In
Repeats-Out
C
Repeats-In / Repeats-Out
Hemizygous
Homozygous
Hemizygous
Homozygous
Heterozygous
P{1RPT}9D3
P{2RPT}49E4
P{3RPT}10D6
B
A480
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0.012
0.017
0.029
0.058
0.042
1 RPT
0 RPTS
HET
P{1RPT}9D3
0.25
0.20
0.15
0.10
0.05
0
0.029
0.058
0.098
0.218
0.177
2 RPTS
0 RPTS
HET
P{2RPT}49E4
0.016
0.023
0.032
0.052
0.065
3 RPTS
0 RPTS
HET
P{3RPT}10D6
Hemizygous
Homozygous
Heterozygous Repeats-In / Repeats-Out

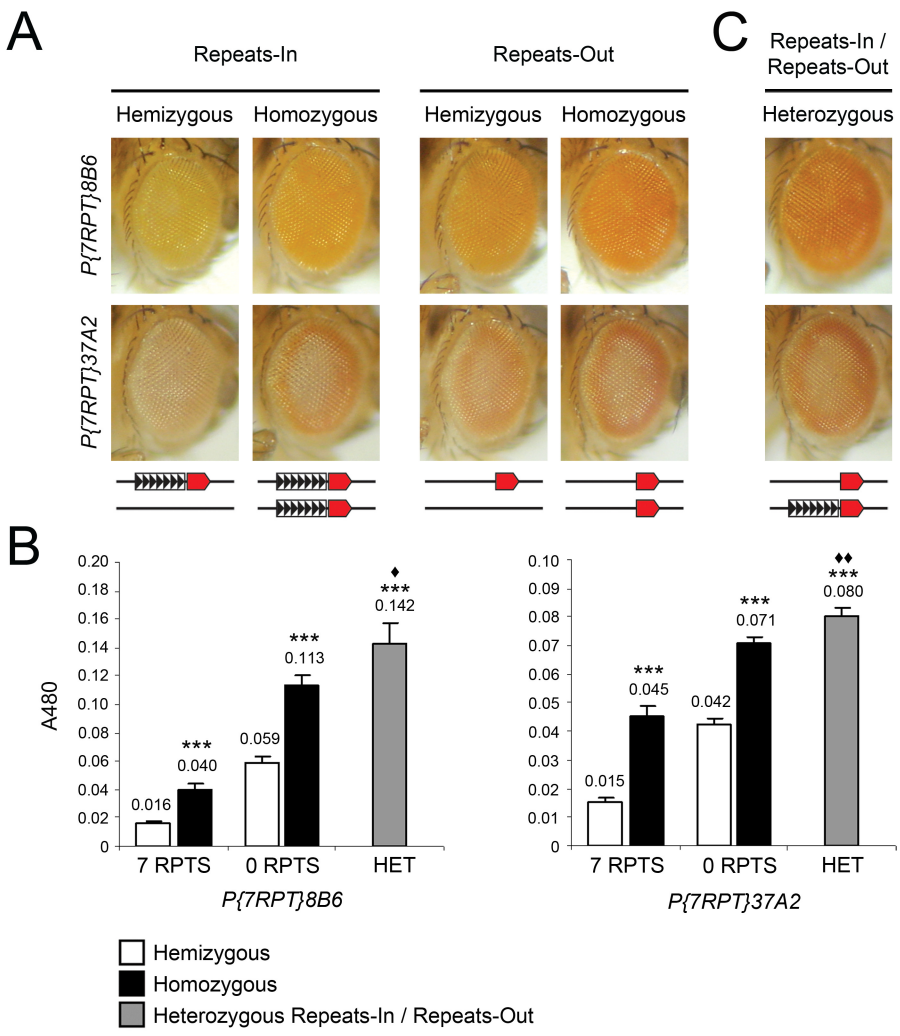


Figure 2.7 Two transgenic lines with seven tandem repeats and *cis*-silencing of *white* demonstrate epigenetic activation of *white* expression in *trans*. **[A]** Visual inspection of eye pigmentation shows that homozygotes are darker than hemizygotes (no pairing-sensitive silencing), and that *white* expression increases following repeat removal (*cis*-silencing). **[B]** Pigment assay quantification of *white* expression in hemizygous, homozygous, and heterozygous ("HET") flies, with the repeats present ("7 RPTS") and following repeat removal ("0 RPTS"). All differences between repeats-in and repeats-out flies are statistically significant, with $p<0.001$. Additional statistical significance is indicated for homozygous values significantly different than hemizygous (*), and heterozygous values significantly different from repeats out hemizygous (*) or homozygous (♦). Significances are indicated at the level of $p<0.05$ (*), $p<0.01$ (**) or $p<0.001$ (***), for each symbol. **[C]** Repeats-in / repeats-out heterozygotes exhibit the darkest eye pigmentation, indicating that the presence of the maize repeats in *trans* may cause epigenetic activation of *white*.

2.5 DISCUSSION

The results presented here demonstrate that the maize *b1* paramutation control region functions as an epigenetic silencer in *Drosophila*. Removal of the *b1* repeats in all transgenic lines containing one, two, three or seven repeats resulted in a loss of silencing and an increase in marker gene expression. While silencing occurred with even one repeat, the strength of silencing was dependent on the number of *b1* repeat sequences. For transgenes inserted at the same genomic position, *white* expression increased 4.4 and 5.4 fold following removal of seven tandem repeats, but only 2.0 and 2.8 fold following removal of three tandem repeats, for males and females respectively (Figures 2.3B and 2.5B). For transgene inserts at different genomic positions, a similar pattern was observed. Upon removal of the repeats, *white* expression increased approximately 2.4 – 5.4 fold in transgenic lines with seven tandem repeats (Figures 2.4B, 2.5B and 2.7B), while in lines with 1 – 3 tandem repeats, *white* expression increased only 2.0 – 3.5 fold (Figures 2.5B and 2.6B).

In maize, the silenced *B'* allele is extremely stable, but the highly expressed *B-I* allele spontaneously converts to *B'*, so it is likely that the silenced epigenetic conformation is the default (Arteaga-Vazquez and Chandler, 2010; Coe, 1966; Patterson *et al.*, 1993). Our results are consistent with this hypothesis, as the maize *b1* repeats consistently caused gene silencing in transgenic *Drosophila*. The observation that a single copy of the tandem repeat sequence can cause observable gene silencing in *Drosophila* suggests that evolutionarily conserved silencing sequences are contained within each 853 bp repeat, and that the specific repeat sequence is sufficient for silencing, while the number of tandem repeats enhances this silencing. This is consistent with

epigenetic silencing of the imprinted *FWA* gene in *Arabidopsis thaliana,* which is mediated by two direct repeats at the gene promoter that contain sequence similarity to a SINE retroelement. Evidence indicates that the SINE-related sequence is required for DNA methylation, epigenetic silencing, and imprinting of Arabidopsis *FWA,* whereas the tandem repeat structure is dispensable (Fujimoto *et al.*, 2008). Similarly, transgenic experiments examining silencing induced by *Drosophila* subtelomeric heterochromatin (also known as telomere-associated sequence, or TAS), showed that silencing was stronger with a greater number of tandem repeats, and that the tandem repeat sequence is important, as a tandem repeat array in itself is not necessarily sufficient to cause silencing (Kurenova *et al.*, 1998). Interestingly, *Drosophila* TAS shares a number of additional features with the *b1* tandem repeat silencing observed here, including reporter gene silencing in *cis* (Boivin *et al.*, 2003; Cryderman *et al.*, 1999), pairing-sensitive silencing (Boivin *et al.*, 2003), *trans*-silencing (Josse *et al.*, 2007; Ronsseray *et al.*, 2003), and *trans* communication (Frydrychova *et al.*, 2007; Mason *et al.*, 2003).

In maize, the *trans*-interactions that occur when the highly expressed *B-I* allele and the weakly expressed *B'* allele are paired in heterozygotes result in meiotically stable epigenetic silencing of the *B-I* allele. The maize *b1* tandem repeats can also mediate *trans* interactions that result in increased silencing in *Drosophila.* Evidence for this includes pairing-sensitive silencing of *white* in repeats-in homozygotes, and *trans*-silencing of *white* in repeats-in / repeats-out heterozygotes (Figure 2.4). At the *b1* locus in maize, paramutation successfully occurs with recombinant alleles that have only five tandem repeats, but is impaired with recombinant alleles that have three tandem repeats, and does not occur when only a single copy of the repeat sequence is present (Stam *et al.*, 2002). This suggests that a minimum number of tandem repeats are required to mediate

or stabilize the *trans*-interactions and establishment of silencing that occurs during paramutation. Consistent with this hypothesis, we observed pairing-sensitive silencing, and silencing of *white* in *trans* in repeats-in / repeats-out heterozygotes, in *Drosophila* lines with the full seven tandem repeats, but not in lines containing one, two, or three tandem repeats. Furthermore, analysis of transgenes at the same genomic position, but with three or seven tandem repeats present, demonstrated that the strength of *trans*-silencing increases as the number of repeats increases (Figure 2.5). Thus, consistent with the endogenous maize system, a threshold number of repeats may be required for these *trans* silencing mechanisms to be established and/or stabilized in *Drosophila*. The fact that two transgenic lines with the full seven *b1* tandem repeats do not exhibit pairing-sensitive silencing, suggests that while the presence of seven *b1* tandem repeats may be required to induce pairing-sensitive silencing, it is not always sufficient to do so. It is likely that other factors at the genomic insertion site, such as the local chromatin environment and/or the proximity to endogenous promoters or enhancers, may influence or inhibit the development or stability of pairing-sensitive silencing.

Such context-dependent regulation is shown by the three transgenic lines, two with seven tandem repeats and one with three tandem repeats, that exhibit evidence of epigenetic activation in *trans* (Figures 2.6 and 2.7). This result demonstrates that the *b1* tandem repeats may be targeted by both activating and repressive epigenetic modifications in transgenic *Drosophila*, similar to the endogenous maize locus where both active and silenced epialleles are formed. Recent epigenetic analysis has shown that the silenced *B'* allele exhibits both activating and repressive histone modifications at the *b1* coding region (Haring *et al.*, 2010; Stam *et al.*, 2002). Similarly, a dynamic epigenetic state consisting of both activating and repressive epigenetic forces also occurs at

Drosophila subtelomeric tandem repeats (Yin and Lin, 2007). In *Drosophila*, the *b1* tandem repeats may form a similar dynamic epigenetic domain. While the repressive epigenetic modifications at the *b1* tandem repeats consistently result in silencing of the *white* reporter gene in *cis*, the balance of the activating and repressive epigenetic forces, as well as external factors at the genomic insertion site, may determine whether silencing or activation is observed when the tandem repeats are present in *trans*.

Intriguingly, pairing-sensitive silencing can persist following repeat-removal, as was seen in line *P{7RPT}44D4* (Figure 2.4). Evidence of persistent epigenetic marks has also been found for the *b1* locus in maize. A silenced *B'* allele becomes highly expressed in *mop1* mutants, but appears to carry a heritable mark that "remembers" its previous silenced epigenetic state, as it reverts back to the silenced *B'* epiallele after reintroduction to wild-type MOP1 protein (Chandler and Alleman, 2008; Dorweiler *et al.*, 2000). Thus, the persistence of pairing-sensitive silencing in *P{0RPT}44D4* might similarly be the result of an epigenetic mark, in the form of a unique chromatin structure or histone modification, or bound protein or RNA, that is retained from the previous heterochromatic state.

While the molecular mechanisms underlying *b1* paramutation are not fully understood, a variety of evidence indicates that RNA-based mechanisms are involved. In maize, the tandem repeats are transcribed from both strands, with a similar level of transcription detected in *B'* and *B-I* plants, as well as in plants with neutral alleles containing a single tandem repeat (Alleman *et al.*, 2006). Paramutation and silencing of *B'* require *mop1,* an RNA-dependent RNA-polymerase (Alleman *et al.*, 2006), and *rmr6* and *mop2,* which encode subunits of RNA polymerases IV and V (Erhard *et al.*, 2009; Hollick *et al.*, 2005; Sidorenko *et al.*, 2009). Small interfering RNAs (siRNAs) from the

tandem repeats are reduced in the presence of mutations that disrupt paramutation (Chandler, 2007; Sidorenko *et al.*, 2009), suggesting that they may mediate the *trans*-communication and establishment of silencing that occurs during paramutation. Interestingly, Polycomb group (PcG) response elements (PREs) frequently cause pairing-sensitive silencing in *Drosophila* (Kassis, 2002), and Polycomb-mediated pairing and epigenetic silencing is reminiscent of paramutation. *Drosophila* PREs produce sense and anti-sense transcripts (Grimaud *et al.*, 2006), and mediate chromosomal interactions via a mechanism that involves the RNAi machinery (Bantignies *et al.*, 2003; Grimaud *et al.*, 2006; Vazquez *et al.*, 2006). The clustering of PREs, PcG proteins, and the RNAi machinery is hypothesized to maintain and stabilize *trans*-interactions between PREs, resulting in enhanced silencing (Grimaud *et al.*, 2006; Kavi *et al.*, 2006).

The results presented here indicate considerable conservation of an epigenetic silencing process between the plant and animal kingdoms. The maize *b1* paramutation control region functions as an epigenetic silencer in *Drosophila*, causing both *cis* and *trans* silencing. The extensive evolutionary distance between maize, an angiosperm plant, and *Drosophila*, a dipteran insect, provides support for the hypothesis that seemingly unique epigenetic processes function by utilizing core mechanisms that are widespread throughout the eukaryotic kingdom.

2.6 ACKNOWLEDGEMENTS

The authors would like to thank Vicki Chandler for the maize *b1* tandem repeat vector, J. Sekelsky for the *pP{WhiteOut2}* vector, the Bloomington *Drosophila* Stock Center, the Lett Fund, and NSERC Canada for a Discovery grant to V. K. Lloyd and a postgraduate scholarship to L. A. McEachern.

2.7 REFERENCES

Alleman M, Sidorenko L, McGinnis K, Seshadri V, Dorweiler JE, White J, Sikkink K, Chandler VL (2006) An RNA-dependent RNA polymerase is required for paramutation in maize. *Nature* **442**: 295-298.

Arney KL, Bae E, Olsen C, Drewell RA (2006) The human and mouse H19 imprinting control regions harbor an evolutionarily conserved silencer element that functions on transgenes in Drosophila. *Dev Genes Evol* **216**: 811-819.

Arteaga-Vazquez MA, Chandler VL (2010) Paramutation in maize: RNA mediated trans-generational gene silencing. *Curr Opin Genet Dev* **20**: 156-163.

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

Boivin A, Gally C, Netter S, Anxolabehere D, Ronsseray S (2003) Telomeric associated sequences of Drosophila recruit polycomb-group proteins in vivo and can induce pairing-sensitive repression. *Genetics* **164**: 195-208.

Chandler V, Alleman M (2008) Paramutation: epigenetic instructions passed across generations. *Genetics* **178**: 1839-1844.

Chandler VL (2007) Paramutation: from maize to mice. *Cell* **128**: 641-645.

Chandler VL, Radicella JP, Robbins TP, Chen J, Turks D (1989) Two regulatory genes of the maize anthocyanin pathway are homologous: isolation of B utilizing R genomic sequences. *Plant Cell* **1**: 1175-1183.

Chandler VL, Stam M (2004) Chromatin conversations: mechanisms and implications of paramutation. *Nat Rev Genet* **5**: 532-544.

Coe EH (1966) The Properties, Origin, and Mechanism of Conversion-Type Inheritance at the B Locus in Maize. *Genetics* **53**: 1035-1063.

Cryderman DE, Morris EJ, Biessmann H, Elgin SC, Wallrath LL (1999) Silencing at Drosophila telomeres: nuclear organization and chromatin structure play critical roles. *Embo J* **18**: 3724-3735.

Dorweiler JE, Carey CC, Kubo KM, Hollick JB, Kermicle JL, Chandler VL (2000) mediator of paramutation1 is required for establishment and maintenance of paramutation at multiple maize loci. *Plant Cell* **12**: 2101-2118.

Erhard KF, Jr., Stonaker JL, Parkinson SE, Lim JP, Hale CJ, Hollick JB (2009) RNA polymerase IV functions in paramutation in Zea mays. *Science* **323**: 1201-1205.

Frydrychova RC, Biessmann H, Konev AY, Golubovsky MD, Johnson J, Archer TK, Mason JM (2007) Transcriptional activity of the telomeric retrotransposon HeT-A in Drosophila melanogaster is stimulated as a consequence of subterminal deficiencies at homologous and nonhomologous telomeres. *Mol Cell Biol* **27**: 4991-5001.

Fujimoto R, Kinoshita Y, Kawabe A, Kinoshita T, Takashima K, Nordborg M, Nasrallah ME, Shimizu KK, Kudoh H, Kakutani T (2008) Evolution and control of imprinted FWA genes in the genus Arabidopsis. *PLoS Genet* **4**: e1000048.

Golic KG, Lindquist S (1989) The FLP recombinase of yeast catalyzes site-specific recombination in the Drosophila genome. *Cell* **59**: 499-509.

Grimaud C, Bantignies F, Pal-Bhadra M, Ghana P, Bhadra U, Cavalli G (2006) RNAi components are required for nuclear clustering of Polycomb group response elements. *Cell* **124**: 957-971.

Haring M, Bader R, Louwers M, Schwabe A, van Driel R, Stam M (2010) The role of DNA methylation, nucleosome occupancy and histone modifications in paramutation. *Plant J*

Hollick JB, Kermicle JL, Parkinson SE (2005) Rmr6 maintains meiotic inheritance of paramutant states in Zea mays. *Genetics* **171**: 725-740.

Huang AM, Rehm EJ, Rubin GM (2000) Recovery of DNA Sequences Flanking P-element Insertions: Inverse PCR and Plasmid Rescue. In *Drosophila Protocols*, Sullivan W, Ashburner M, Hawley RS (eds) pp 429-439. Cold Spring Harbor: CSHL Press.

Josse T, Teysset L, Todeschini AL, Sidor CM, Anxolabehere D, Ronsseray S (2007) Telomeric trans-silencing: an epigenetic repression combining RNA silencing and heterochromatin formation. *PLoS Genet* **3**: 1633-1643.

Kassis JA (2002) Pairing-sensitive silencing, polycomb group response elements, and transposon homing in Drosophila. *Adv Genet* **46**: 421-438.

Kavi HH, Fernandez HR, Xie W, Birchler JA (2006) Polycomb, pairing and PIWI--RNA silencing and nuclear interactions. *Trends Biochem Sci* **31**: 485-487.

Khesin RB, Leibovitch BA (1978) Influence of deficiency of the histone gene-containing 38B-40 region on X-chromosome template activity and the white gene position effect variegation in Drosophila melanogaster. *Mol Gen Genet* **162**: 323-328.

Kurenova E, Champion L, Biessmann H, Mason JM (1998) Directional gene silencing induced by a complex subtelomeric satellite from Drosophila. *Chromosoma* **107**: 311-320.

Louwers M, Bader R, Haring M, van Driel R, de Laat W, Stam M (2009) Tissue- and expression level-specific chromatin looping at maize *b1* epialleles. *Plant Cell* **21**: 832-842.

Lyko F, Brenton JD, Surani MA, Paro R (1997) An imprinting element from the mouse H19 locus functions as a silencer in Drosophila. *Nature Genetics* **16**: 171-173.

Lyko F, Buiting K, Horsthemke B, Paro R (1998) Identification of a silencing element in the human 15q11-q13 imprinting center by using transgenic Drosophila. *Proc Natl Acad Sci U S A* **95**: 1698-1702.

Mason JM, Konev AY, Golubovsky MD, Biessmann H (2003) Cis- and trans-acting influences on telomeric position effect in Drosophila melanogaster detected with a subterminal transgene. *Genetics* **163**: 917-930.

Patterson GI, Thorpe CJ, Chandler VL (1993) Paramutation, an allelic interaction, is associated with a stable and heritable reduction of transcription of the maize b regulatory gene. *Genetics* **135**: 881-894.

Rassoulzadegan M, Grandjean V, Gounon P, Cuzin F (2007) Inheritance of an epigenetic change in the mouse: a new role for RNA. *Biochem Soc Trans* **35**: 623-625.

Robertson HM, Preston CR, Phillis RW, Johnson-Schlitz DM, Benz WK, Engels WR (1988) A stable genomic source of P element transposase in Drosophila melanogaster. *Genetics* **118**: 461-470.

Ronsseray S, Josse T, Boivin A, Anxolabehere D (2003) Telomeric transgenes and trans-silencing in Drosophila. *Genetica* **117**: 327-335.

Sidorenko L, Dorweiler JE, Cigan AM, Arteaga-Vazquez M, Vyas M, Kermicle J, Jurcin D, Brzeski J, Cai Y, Chandler VL (2009) A dominant mutation in mediator of paramutation2, one of three second-largest subunits of a plant-specific RNA polymerase, disrupts multiple siRNA silencing processes. *PLoS Genet* **5**: e1000725.

Spradling AC, Rubin GM (1982) Transposition of cloned P elements into Drosophila germ line chromosomes. *Science* **218**: 341-347.

Stam M (2009) Paramutation: a heritable change in gene expression by allelic interactions in trans. *Mol Plant* **2**: 578-588.

Stam M, Belele C, Dorweiler JE, Chandler VL (2002) Differential chromatin structure within a tandem array 100 kb upstream of the maize *b1* locus is associated with paramutation. *Genes Dev* **16**: 1906-1918.

Tweedie S, Ashburner M, Falls K, Leyland P, McQuilton P, Marygold S, Millburn G, Osumi-Sutherland D, Schroeder A, Seal R, Zhang H (2009) FlyBase: enhancing Drosophila Gene Ontology annotations. *Nucleic Acids Res* **37**: D555-559.

Vazquez J, Muller M, Pirrotta V, Sedat JW (2006) The mcp element mediates stable long-range chromosome-chromosome interactions in Drosophila. *Mol Biol Cell* **17**: 2158-2165.

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

2.8 SUPPLEMENTARY MATERIAL

Table 2.S1 Primers used in inverse PCR and sequencing reactions to determine transgene insertion site.

P-end [a]	Orientation	Name [b]	Sequence (5' → 3')	Primer Type
3'	F	P1-Map-F	CACACCACAAATATACTGTTGCCGAGC	PCR
		P1-FSEQ	AACTTCTAGGGATAACAGGGTAATGG	Sequencing
	R	P1-Map-R	TCAAACCCCACGGACATGCTAAGG	PCR
		P1-RSEQ	CGTTAAGTGGATGTCTCTTGCCG	PCR & Sequencing
5'	F	Plac1	CACCCAAGGCTCTGCTCCCACAAT	PCR
		P2-Set1-F	CTCTTAATAGCACACTTCGGCACG	PCR
		P2-SEQ-F	ATTCGTCCGCACACAACCTTTCC	PCR
		Sp1	ACACAACCTTTCCTCTCAACAA	Sequencing
	R	Pwht1	GTAACGCTAATCACTCCGAACAGGTCACA	PCR
		P2-RSEQ	GGAACCATTTGAGCGAACCGAA	PCR
		P2-Set1-R	AGGTGAATGTGTTGCGGAGAGC	PCR

[a] Some primers bind outside the P-end sequence, but within that end of the vector, before the first *Msp*I restriction site.

[b] Primers Plac1, Sp1, and Pwht1 are described in Huang et al. (2000).

2.9 TRANSITION TO CHAPTER 3

The following chapter details bidirectional transcription at the maize *b1* tandem repeats in transgenic *Drosophila*. I show that the tandem repeats are transcribed from both strands, and that aberrant transcription persists following repeat removal, but is lost upon moving the transgene to a new genomic position. These findings are significant because they demonstrate conservation of the RNA-based mechanisms underlying *b1* paramutation. In addition, the persistent transcription in "repeats-out" transgenes is a unique and exciting result, as it appears to be an epigenetic mark that is retained from the previous epigenetic state, and is meiotically transmitted for many generations.

CHAPTER 3

MAIZE *B1* PARAMUTATION-ASSOCIATED REPEATS ARE BIDIRECTIONALLY TRANSCRIBED IN *DROSOPHILA*.

Lori A. McEachern and Vett K. Lloyd

Manuscript prepared for submission to *Genetics*.

3.1 ABSTRACT

Paramutation is a phenomenon in which the epigenetic state of one allele heritably changes when it is paired with another. In maize, paramutation at the *b1* locus appears to involve RNA-based mechanisms, and requires a control region consisting of seven tandem repeats that are bidirectionally transcribed. Previously we have shown that tandem repeats required for *b1* paramutation in maize cause silencing in *cis* and in *trans* in transgenic *Drosophila.* Here we examine transgenic *Drosophila* lines for transcription of the maize *b1* tandem repeats. Using RT-PCR, bidirectional transcription of the *b1* tandem repeats was detected in transgenic lines with the full seven tandem repeats, as well as transgenic lines with two or three tandem repeats. Transcription from both strands was also detected in the regions flanking the tandem repeats. Following repeat removal, aberrant bidirectional transcription in the region previously flanking the repeats persisted in most transgenic lines. However, a transgenic line with zero tandem repeats moved to a novel genomic position appeared to be "reset" and demonstrated no detectable transcription originating from this region. These results suggest that the RNA-based mechanisms underlying *b1* paramutation are conserved in *Drosophila,* and can direct epigenetic changes that persist for many generations, even following repeat removal.

3.2 INTRODUCTION

Paramutation is an epigenetic process that results in a heritable change in expression of one allele after it has been paired with another. Paramutation has been described most extensively in plants (Chandler and Stam, 2004), but it has also been

observed in mice (Rassoulzadegan et al., 2007), suggesting that it may be based on an ancient and evolutionarily conserved process. At the maize *b1* locus, paramutation occurs when the highly expressed *B-I* allele is present in *trans* to the weakly expressed *B'* allele. The *B-I* and *B'* alleles are identical in sequence, and contain an identical paramutation control region located 100 kb upstream of the *b1* transcriptional start site, indicating that they differ only in their epigenetic status (Patterson et al., 1993; Stam et al., 2002). When the two alleles are combined in a heterozygote, the normally highly expressed *B-I* allele is epigenetically silenced, and the newly silenced *B-I* allele (now equivalent to *B'*, and also termed *B'**) has the ability to silence or "paramutate" additional *B-I* alleles when transmitted to the next generation (Coe, 1966; Patterson et al., 1993).

The *b1* paramutation control region consists of seven 853 bp tandem repeats. Neutral alleles that do not participate in paramutation contain only a single copy of the repeat sequence. In *B'* plants, the tandem repeats exhibit increased methylation relative to *B-I*, repressive histone modifications, and a closed chromatin structure, while in *B-I* plants the tandem repeats are characterized by histone acetylation and a more accessible chromatin structure (Haring et al., 2010; Stam et al., 2002). Transcription from both strands of the tandem repeats has been detected for both the *B-I* and *B'* alleles, as well as a single repeat, neutral allele (Alleman et al., 2006). The tandem repeat transcripts are hypothesized to mediate the *trans* communication and establishment of silencing that occurs during paramutation, likely by directing siRNA-based chromatin modifications (Arteaga-Vazquez and Chandler, 2010; Chandler, 2007). siRNAs from the tandem repeats have been detected from *B'*, *B-I,* and a neutral allele, but are reduced in the presence of mutations that disrupt paramutation (Arteaga-Vazquez and Chandler, 2010; Sidorenko et al., 2009), underscoring their importance in paramutation. The importance

of RNA-based mechanisms in paramutation is further emphasized by the requirement of MOP1, an RNA-dependent RNA polymerase (Alleman et al., 2006), and RMR6 and MOP2, subunits of RNA polymerases IV and V, for paramutation and epigenetic silencing of *B'*.

The detection of siRNAs from the neutral allele that does not participate in paramutation indicates that while repeat siRNAs may be required for paramutation, they are not always sufficient, with the number of tandem repeats being a critical factor for paramutation. The tandem repeats may mediate or stabilize the pairing-interactions between alleles, potentially by enhancing the accumulation of proteins or important chromatin modifications at critical regions of the repeats. Consistent with this, a protein called CXC domain *b1*-repeat binding protein, or CBBP, was recently found to bind to the tandem repeats in maize, with at least two binding sites per repeat. This protein can form multimers, which could provide a mechanism of repeat "counting" (Brzeska et al., 2010). In addition, epigenetic differences between the *B-I* and *B'* tandem repeats were found to be concentrated at repeat junctions, indicating that these regions may be uniquely important for paramutation, potentially by harbouring a high concentration of binding sites for histone modifying- or DNA methylating-enzymes (Haring et al., 2010). The highly expressed *B-I* allele, which both produces siRNAs and contains seven tandem repeats, may utilize an additional mechanism to prevent silencing. The active chromatin structure of the repeats, or specific proteins that bind to the active chromatin structure, may inhibit the formation of the silenced epigenetic state, or the allele may be localized to a different nuclear environment that inhibits silencing (Arteaga-Vazquez and Chandler, 2010; Chandler, 2007).

We have recently shown that the *b1* paramutation control region causes epigenetic silencing in *Drosophila* (chapter 2). Silencing of an adjacent marker gene in *cis* was observed in all transgenic lines with 1 – 7 tandem repeats, with the strength of silencing increasing as the number of repeats increased. In transgenic lines with the full seven tandem repeats, evidence of *trans* interactions, such as pairing-sensitive silencing, were also frequently observed. Silencing in *trans* was reduced or lost when the number of tandem repeats decreased. Here we examine transgenic *Drosophila* for transcription of the maize *b1* tandem repeats, and show that the repeats are transcribed from both strands, as they are at the endogenous maize locus. Bidirectional transcription was also detected at the regions flanking the repeats. Following removal of the repeats, aberrant transcription at the region previously encompassing the repeats was detected in all transgenic lines. However, a line with zero tandem repeats, but moved to a new genomic position, did not exhibit detectable transcription at this region, suggesting that this transcription may be a mark, or the result of a mark, persisting from the previous epigenetic state.

3.3 MATERIALS AND METHODS

3.3.1 *DROSOPHILA* STOCKS

All *Drosophila* stocks were maintained at 21±3°C on a standard cornmeal medium. Standard stocks used are described in FlyBase (Crosby et al. 2007) and were provided by the Bloomington *Drosophila* Stock Center. All transgenic *Drosophila* lines used in this study are described in chapter 2, except Line *P{0RPT.n}83C5*, which was obtained by P-element mobilization of *P{0RPT}44D4*, a stable transgenic line from which

the maize tandem repeats had previously been removed by Flipase recombinase (FLP). *P{0RPT.n}83C5* is inserted on the 3rd chromosome at cytoband 83C5, at chromosome coordinate 1,682,023 with the P-ends in "+" orientation. All zero-repeat transgenic lines were propagated for a minimum of eight generations after repeat removal before RNA isolation.

3.3.2 RNA ISOLATION

For each genotype examined, approximately 30 adult males and females, aged 0- to 24-hours, were collected. Hemizygous flies were selected from balanced stocks with autosomal inserts, while homozygous flies were selected from stocks containing an X chromosome insert. RNA was isolated using Trizol reagent (Invitrogen) according to manufacturer's instructions. 10 μg of RNA was treated with DNAseI (Fermentas) for 30 minutes at 37°C immediately prior to reverse transcription.

3.3.3 REVERSE TRANSCRIPTION AND PCR

Reverse transcription was conducted using Superscript II RT (Invitrogen). Control reactions containing the identical reaction components, but lacking the Reverse Transcriptase (RT) enzyme, were carried out in tandem. The RT reactions (but not the controls lacking RT enzyme) were found to produce PCR bands in the absence of primers in the RT reaction, likely due to nonspecific-, self-, or small RNA-priming of the RT enzyme. To eliminate this effect and confirm strand-specific transcription of the maize repeats, an adapter sequence of 5'-GCCTGCCCCAACCTCC-3' (Shpiz et al., 2009) was added to the 5' end of a primer specific to the maize repeats or surrounding DNA sequence. A single adapter-specific primer was used in each RT reaction, except for the

actin 88f RT reaction, which used an *actin 88f* specific RT primer. The primer sequences and combinations used to detect each band are listed in supplementary Tables 3.S1 and 3.S2. Each RT reaction contained 2.5 μg of RNA and 2 pmole of primer in a 20 μL volume. Forty PCR cycles were carried out using 2 μL of the RT reaction in a 10 μL PCR reaction containing 0.25 μM primers. For all RT-PCRs except *actin*, the adapter primer and one specific primer were used in each PCR reaction. 2 μL of load dye was added to each PCR, and 6 μL of this was run on a 1.7 – 2% agarose gel in Sodium Borate (SB) buffer. Gels were stained with Sybr Green I (Lonza) for visualization.

3.4 RESULTS

3.4.1 THE MAIZE *B1* TANDEM REPEATS ARE TRANSCRIBED FROM BOTH STRANDS IN TRANSGENIC *DROSOPHILA*.

Analysis of *b1* tandem repeat transcription in maize previously demonstrated that the repeats are transcribed from both strands, with the highest levels of transcription observed at the start and end of each repeat sequence (Alleman et al., 2006; Sidorenko et al., 2009). We therefore analyzed transgenic *Drosophila* for transcription of the *b1* tandem repeats using an RT-PCR adapter-primer approach, with primers designed to amplify sequences present at the start (Region 1) or end (Region 2) of each tandem repeat (Figure 3.1A). Transcription analysis of three transgenic *Drosophila* lines, each containing seven tandem repeats, demonstrated that transcription occurs from both strands within the *b1* tandem repeats, with forward and reverse transcription detected for all lines at both ends of the repeat units (Region 1 and Region 2; Figure 3.1B). Similarly, transgenic lines with either two or three *b1* tandem repeats also exhibited bidirectional transcription at both regions (Figure 3.1C).

3.4.2 Bidirectional transcription is also detected at the regions flanking the repeats.

To determine whether bidirectional transcription extended to or from the regions flanking the *b1* tandem repeats, forward and reverse transcription at the region upstream of the first tandem repeat (5' Flanking Region) and downstream of the last tandem repeat (3' Flanking Region) were assessed (Figure 3.2A). Forward and reverse transcription at the 5' Flanking Region was detected for the seven tandem repeat transgenes (Figure 3.2B). At the 3' Flanking Region, forward transcription was relatively weak but detectable. However, significant reverse transcription was only detected for two of the three transgenes assessed (Figure 3.2B). In line *P{7RPT}44D4*, reverse transcription at this region was either very weak or absent, or the transcript was inaccessible or degraded, as only a faint smudge was visible, even after 40 PCR cycles. Forward and reverse transcription at both flanking regions was detected for both transgenes containing either two or three *b1* tandem repeats (Figure 3.2C).

3.4.3 Aberrant transcription persists following repeat removal, but is lost at a new genomic position.

Previously we have shown that removal of the *b1* tandem repeats by FLP recombinase relieves silencing of the adjacent *white* reporter gene. To examine the "repeats-out" transgenes for aberrant forward or reverse transcription persisting following repeat removal, we used adapter-primer RT-PCR to detect transcription in the region encompassing the FRT site, which previously contained the *b1* tandem repeats (Figure 3.3A). All repeats-out transgenes exhibited aberrant reverse transcription in this region (Figure 3.3B). The observation that reverse transcription is detectable in this region in *P{0RPT}44D4* (Figure 3.3B), suggests that there is likely a reverse transcript being

synthesized in the 3' flanking region of the *P{7RPT}44D4* transgene (Figure 3.2B), but is eluding detection due to the transcript's scarcity, degradation, processing or inaccessibility, rather than its absence.

Forward transcription in this region of the repeats-out transgenes was highly variable. Weak band-like smudges were observed for two of the three transgenic lines that previously contained seven tandem repeats, while the third contained no detectable forward transcription. The two transgenic lines that previously contained two or three tandem repeats exhibited more significant forward transcription in this region (Figure 3.3B).

To assess whether these transcripts are inherent to the transgenic construct, or are persisting in the repeats-out lines due to their previous epigenetic state, we analyzed transcription from a transgenic line that contains zero-tandem repeats, but has been moved to a new genomic position. This transgenic line, termed *P{0RPT.n}83C5*, was obtained by P-element mobilization of *P{0RPT}44D4* (Figure 3.3B, second line). It is identical in sequence to the repeats-out transgenes, and is distinguished only by the fact that it is inserted at a site that did not previously contain the *b1* tandem repeats. Interestingly, no aberrant forward or reverse transcription was detected from the *P{0RPT.n}83C5* transgene in the region encompassing the FRT site (Figure 3.3C). RT-PCR of *actin* confirms the integrity of the RNA samples (Figure 3.3), and PCR of genomic DNA confirms that this region is intact and readily amplifiable from all zero-repeat transgenes, including *P{0RPT.n}83C5* (supplementary Figure 3.S1). Transcription of the FRT flanking region is thus not an inherent feature of the zero-repeats transgenes, suggesting that the aberrant transcription detected from the repeats-out transgenes (Figure

3.3B) may be an epigenetic mark left-over from the previous epigenetic state when the *b1* tandem repeats were present.

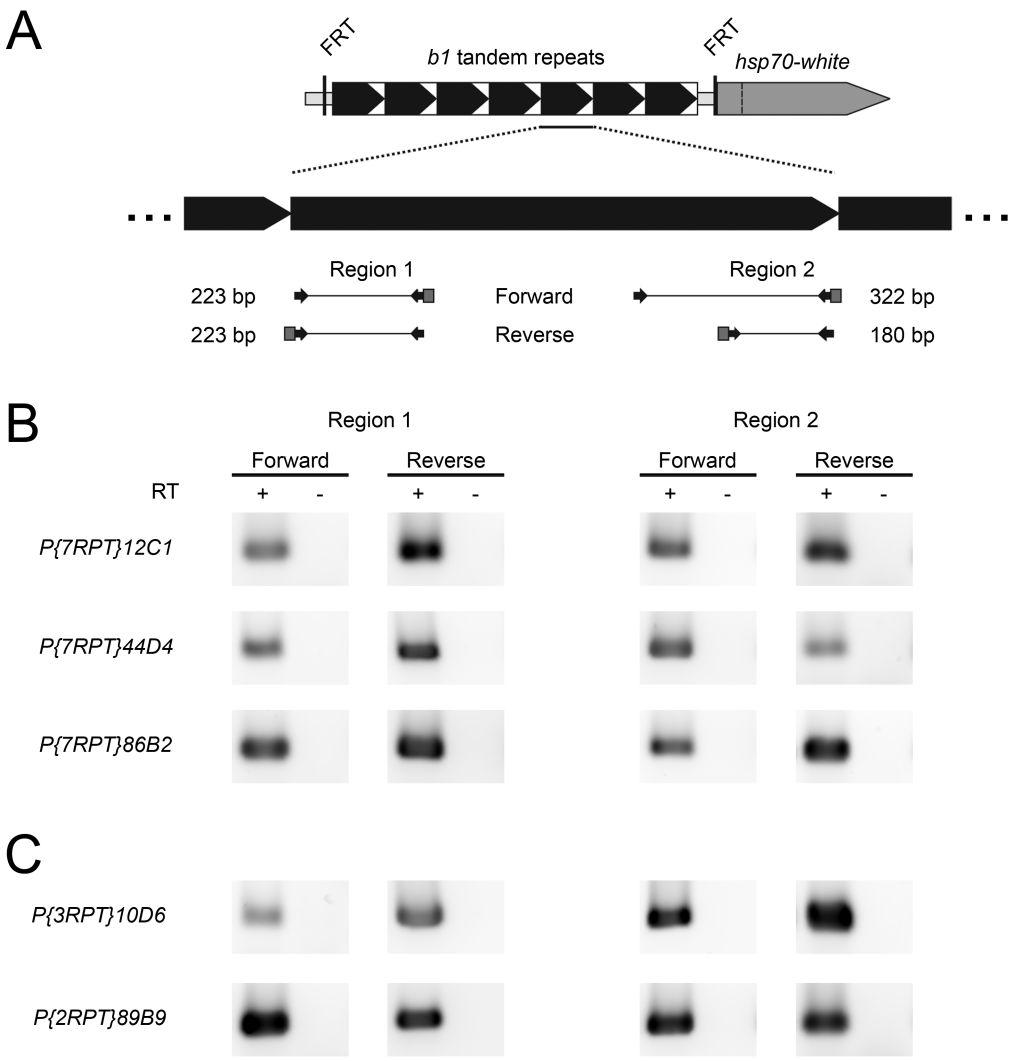


Figure 3.1 The maize *b1* tandem repeats are transcribed from both strands in transgenic *Drosophila*. **[A]** The transgenic sequence contains the maize *b1* tandem repeats between two FRT sites and adjacent to *Drosophila hsp70-white*. Two regions within the sequence of each tandem repeat were analyzed for forward and reverse transcription. The size and direction of the transcripts are indicated, with forward transcripts corresponding to transcription in the left to right direction as drawn, and reverse transcripts corresponding to transcription from right to left. For each transcript, RT- and PCR-primer positions are indicated with black arrows, and the adapter sequence, which is added on during the RT reaction and then used as a PCR primer, is indicated with a grey block. **[B]** Three transgenic *Drosophila* lines, each containing seven tandem repeats, demonstrated bidirectional transcription of the repeat sequence for both regions analyzed. The presence (+) or absence (-) of RT enzyme in the RT reaction is indicated above each lane. **[C]** Two additional transgenic lines, with three or two tandem repeats, also exhibited transcription from both strands, for both repeat regions.

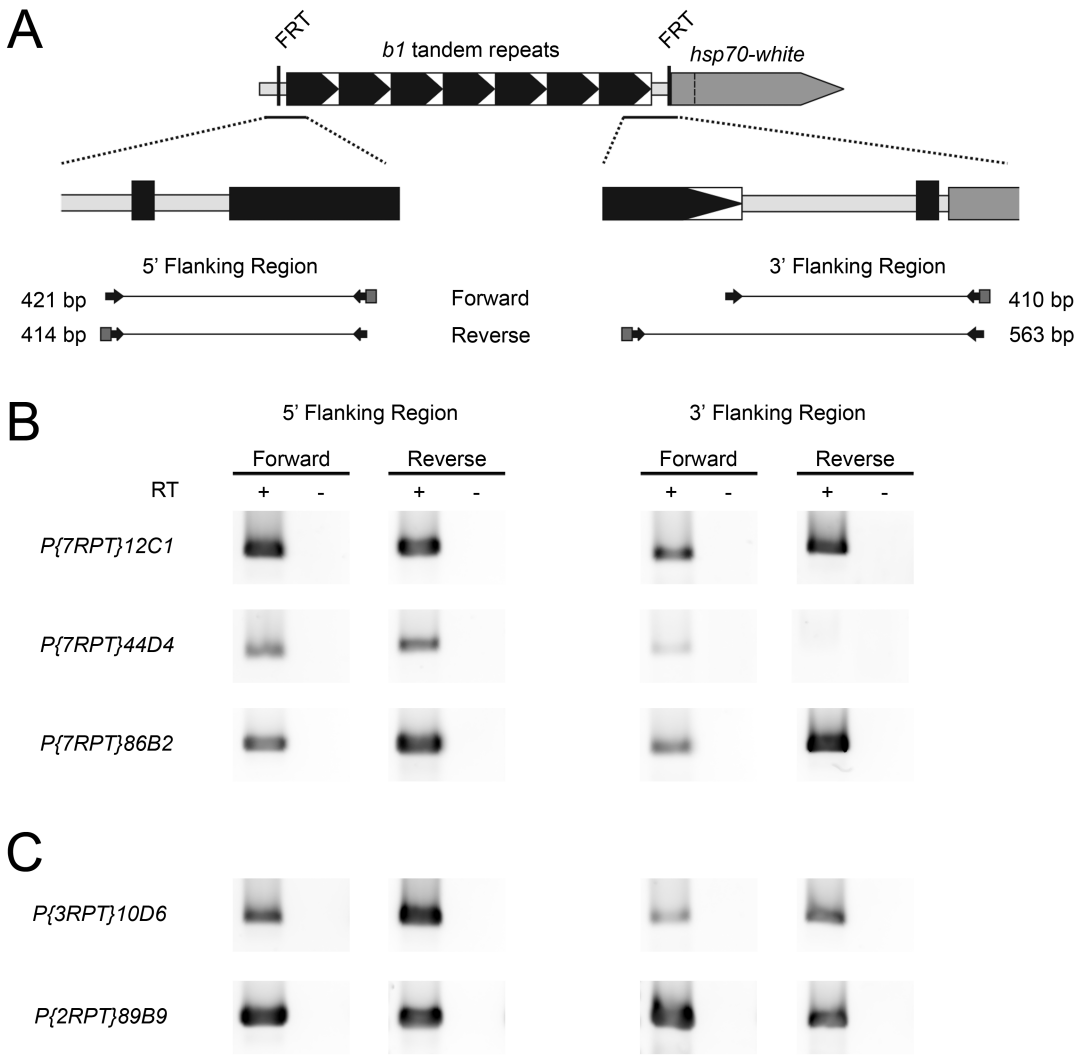


Figure 3.2 Transcripts from both strands are detected at the 5' and 3' regions flanking the repeats. **[A]** The 5' flanking region analyzed includes a portion of the vector backbone, the FRT site, and the start of the first tandem repeat. The 3' flanking region analyzed includes the end of the last repeat, a portion of the vector backbone, the second FRT site, and the start of the *hsp70* promoter (*white* cDNA begins at the dashed line in the top schematic). The positions of the primer and adapter sequences are indicated, as in Figure 3.1. **[B]** Two transgenic *Drosophila* lines with seven *b1* tandem repeats exhibited bidirectional transcription of both flanking regions. In line *P{7RPT}44D4* transcription from both strands was detected for the 5' flanking region; however at the 3' flanking region, transcription in the reverse direction was either very weak or absent. The presence (+) or absence (-) of RT enzyme in the RT reaction is indicated above each lane. **[C]** Two additional transgenic lines, with three or two tandem repeats, exhibited transcription from both strands, for both the 5' and 3' regions flanking the *b1* repeats.

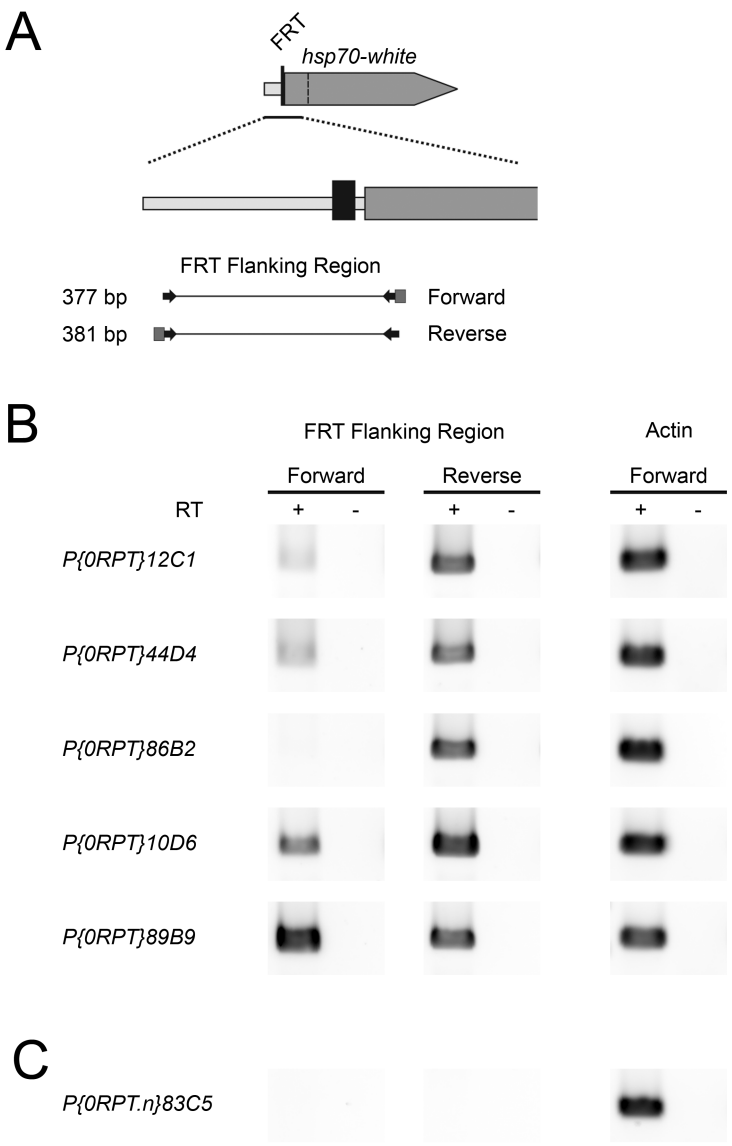


Figure 3.3 Aberrant transcription persists in some lines after repeat removal, but is not detected in a line where the zero repeat transgene was moved to a new genomic position. **[A]** Following repeat removal, a single FRT site is present adjacent to the *hsp70-white* gene. The region analyzed for transcription includes a portion of the P-end and vector sequence, the FRT site, and the start of the *hsp70* promoter (*white* cDNA begins at the dashed line in the top schematic). Primer and adapter positions are indicated as in Figure 3.1. **[B]** In transgenic lines with zero tandem repeats, reverse transcription was detected in all lines. Significant forward transcription was only detected in lines *P{0RPT}89B9* and *P{0RPT}10D6*, although lines *P{0RPT}44D4* and *P{0RPT}12C1* exhibited weak band-like smudges. RT-PCR of *actin* confirms the integrity of the RNA. The presence (+) or absence (-) of RT enzyme in the RT reaction is indicated above each lane. **[C]** No forward or reverse transcription at the region flanking the FRT site could be detected for a zero-repeat transgene that had been moved to a new genomic position, indicating that this transcription is not inherent to the transgenic construct.

3.5 DISCUSSION

We have previously shown that the epigenetic silencing and *trans* communication functions of the maize *b1* tandem repeats are conserved in transgenic *Drosophila* (chapter 2). In maize, bidirectional transcripts from the repeats are believed to be essential for the *trans* communication and establishment of epigenetic states that occurs during paramutation. Here we show that the *b1* tandem repeats are transcribed from both strands in *Drosophila,* as they are at the endogenous locus in maize. Our finding is consistent with the observation that bidirectional transcription and epigenetic silencing from the mouse *H19* imprint control region is also conserved in transgenic *Drosophila* (Drewell et al., 2000; Lyko et al., 1997; Schoenfelder et al., 2007). Further, it has recently been proposed that paramutation is simply an extreme manifestation of RNA-directed DNA methylation or RNAi-directed chromatin modifications (Teixeira and Colot, 2010). RNA-directed chromatin modifications and heterochromatin formation are conserved epigenetic processes that are widespread throughout the eukaryotic kingdom (Djupedal and Ekwall, 2009; Matzke and Birchler, 2005; Zaratiegui et al., 2007). Thus, bidirectional transcription at the maize *b1* tandem repeats in *Drosophila* may trigger epigenetic processes similar to those occurring at the endogenous maize locus, resulting in heterochromatinization, and *cis* and *trans* silencing.

Our analysis of repeat transcription in transgenic *Drosophila* demonstrates that the maize *b1* tandem repeats are transcribed from both strands, with forward and reverse transcription detected for two distinct regions flanking the repeat junctions (Figure 3.1). In maize, these regions exhibit the highest levels of transcription (Alleman et al., 2006; Sidorenko et al., 2009), and epigenetic differences between the *B-I* and *B'* tandem repeats

are concentrated at these regions (Haring et al., 2010). Bidirectional transcription at both regions was detected for transgenic *Drosophila* lines with seven tandem repeats, as well as lines with only two or three tandem repeats (Figure 3.1). This is consistent with the endogenous maize locus, where bidirectional repeat transcription is detected in plants with the highly expressed *B-I* allele and the weakly expressed *B'* allele, both of which both contain seven tandem repeats, as well as in plants with a neutral allele that contains a single copy of the repeat sequence (Alleman et al., 2006).

In transgenic *Drosophila*, we have previously observed that a single tandem repeat can silence adjacent sequences in *cis*, with the strength of silencing increasing as the number of repeats increases (chapter 2). Double stranded RNA produced by bidirectional transcription is frequently processed by the RNAi machinery to produce small RNAs, which can then trigger epigenetic silencing via transcriptional or post-transcriptional mechanisms (Kavi et al., 2008; Matzke and Birchler, 2005). In *Drosophila*, the tandem repeat transcripts may therefore direct heterochromatin formation and epigenetic silencing at the *b1* repeat sequences, with all transgenes containing even one tandem repeat acquiring chromatin modifications that result in visible epigenetic silencing. A greater number of repeats would result in a greater accumulation of silencing factors, and therefore a greater silencing strength.

In *Drosophila*, there is also substantial evidence that the repeats can mediate *trans* interactions between paired sequences, including pairing-sensitive silencing. This process appears to require a threshold number of tandem repeats (chapter 2). Similarly, in maize, paramutation is hypothesized to occur by RNA-mediated *trans* communication between alleles that establishes the epigenetic states of the tandem repeat control sequences, and converts the highly expressed *B-I* allele to a silent state (Arteaga-Vazquez

and Chandler, 2010). The number of tandem repeats is important for the strength of paramutation (Stam *et al.*, 2002). In our transgenic system, the repeat transcripts may play a similar role in establishing silencing and mediating interactions between paired alleles. Potentially, repeat-mediated interactions between homozygous transgenes results in pairing-sensitive silencing in transgenic lines containing seven tandem repeats, with the number of repeats stabilizing the chromosomal interactions via an increased accumulation of heterochromatic proteins or epigenetic modifications at the repeats or repeat junctions. Consistent with this, pairing-sensitive silencing in *Drosophila* is frequently observed from transgenes containing Polycomb Response Elements (PREs), which produce sense and anti-sense transcripts (Grimaud et al., 2006), and mediate chromosomal interactions via a mechanism that involves the RNAi machinery (Bantignies et al., 2003; Grimaud et al., 2006; Vazquez et al., 2006). The clustering of PREs, Polycomb Group proteins, and the RNAi machinery is hypothesized to create a localized concentration of RNAs that triggers RNAi pathways after reaching a critical threshold level, with the RNAs, RNAi machinery, and PcG proteins functioning together to maintain and stabilize the *trans*-interactions between PREs (Grimaud et al., 2006; Kavi et al., 2006).

The detection of bidirectional transcription at the regions flanking the tandem repeats (Figure 3.2) may indicate that the repeats recruit transcriptional factors that increase aberrant bidirectional transcription over a large chromosomal region encompassing the repeats themselves, and the adjacent sequences. The extension of bidirectional transcription into the adjacent chromosomal regions may contribute to silencing of adjacent genes by triggering degradation of the endogenous gene transcripts, or by causing chromatin modifications and heterochromatinization to encompass the adjacent gene regions.

Transcripts from the repeat flanking regions may also contribute to the *trans* interactions that we have previously observed between repeats-in / repeats-out heterozygotes (chapter 2). As the repeats-out transgenes no longer contain the *b1* tandem repeat sequences, homology between the non-coding RNAs produced from repeats-in and repeats-out transgenes would be at the regions flanking the tandem repeats. These transcripts may therefore be essential in mediating communication between repeats-in and repeats-out transgenes, and triggering silencing of *white* in *trans*. Similarly, transgenes inserted into the tandem repeats of subtelomeric heterochromatin can trigger *trans*-silencing of euchromatic transgenes through mechanisms that require sequence similarity and utilize proteins involved in heterochromatin formation and RNAi (Josse et al., 2007). RNAs from the "silencer" transgene contained in the telomeric tandem repeats are hypothesized to interact with the euchromatic "target" transgene, or facilitate pairing-interactions between the two transgenes, and then trigger heterochromatinization that spreads throughout the chromosomal region encompassing the target (Josse et al., 2007). Sequence similarity between the repeats-in and repeats-out transgenes may similarly trigger pairing-interactions and the transfer of a heterochromatic state from the repeats-in "silencer" transgene to the repeats-out "target" transgene, with aberrant RNAs from homologous regions shared between the transgenes mediating the pairing-interactions or the targeting of chromatin modifications, which then spread throughout the repeats-out transgene.

Importantly, following repeat removal, all transgenic lines examined continued to produce reverse transcripts at the region flanking the FRT site, which previously contained the *b1* tandem repeats (Figure 3.3). The repeats had been removed from these transgenic lines for a minimum of eight generations, and in some cases several years,

prior to RNA isolation. The absence of transcripts at this region following relocation of the transgene to a new genomic position (Figure 3.3C) indicates that this transcription is not inherent to the transgenic sequence. This suggests that the transcription might therefore be an epigenetic mark, or the result of an epigenetic mark, that persists and is propagated at this region, even after repeat removal. The aberrant transcription may be established when the *b1* repeats are present and then continues to be produced following repeat removal, potentially by the propagation of distinct histone modifications, a unique chromatin structure, or the assembly of transcription factors at the aberrant transcription start site. Moving the transgene to a new genomic position presumably disrupts this epigenetic mark.

The persistence of an aberrant epigenetic mark through many meiotic generations is similar to paramutation, in which an epigenetic change, once established, is stable through meiosis. Transmission of the RNA or unique chromatin marks into the gametes could result in the propagation of this epigenetic memory for many generations. Indeed, paramutation mediated by RNAs that are transferred to progeny via the gametes has been documented at the *Kit* locus in mice (Rassoulzadegan et al., 2007).

Aberrant forward transcription appears to be more readily lost following repeat removal. Forward transcripts originating in the flanking DNA were absent or very weak in lines *P{0RPT}86B2*, *P{0RPT}12C1* and *P{0RPT}44D4* (Figure 3.3B). This confirms that the relatively robust forward transcription detected in the *b1* repeats and flanking regions of the repeats-in versions of these lines (Figures 3.1B and 3.2B) is specifically recruited to the *b1* tandem repeat sequences.

In summary, the maize *b1* tandem repeats are transcribed from both strands in *Drosophila*, as at the endogenous maize locus. In *Drosophila,* the transcription extends to

adjacent chromosomal regions and potentially directs the establishment of the previously observed epigenetic silencing and *trans* interactions. Aberrant reverse transcription persists following repeat removal, but is lost upon relocation to a new genomic position, suggesting that this may be an epigenetic mark that confers a meiotically-stable memory of the previous heterochromatic state. Overall these results indicate conservation of the RNA-based epigenetic mechanisms underlying maize *b1* paramutation, and suggest that seemingly unique epigenetic phenomena, such as paramutation, function by exploiting evolutionarily conserved core epigenetic mechanisms that are found in a wide range of eukaryotic species.

3.6 REFERENCES

Alleman M, Sidorenko L, McGinnis K, Seshadri V, Dorweiler JE, White J, Sikkink K, Chandler VL (2006) An RNA-dependent RNA polymerase is required for paramutation in maize. *Nature* **442**: 295-298.

Arteaga-Vazquez MA, Chandler VL (2010) Paramutation in maize: RNA mediated trans-generational gene silencing. *Curr Opin Genet Dev* **20**: 156-163.

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

Brzeska K, Brzeski J, Smith J, Chandler VL (2010) Transgenic expression of CBBP, a CXC domain protein, establishes paramutation in maize. *Proc Natl Acad Sci U S A* **107**: 5516-5521.

Chandler VL (2007) Paramutation: from maize to mice. *Cell* **128**: 641-645.

Chandler VL, Stam M (2004) Chromatin conversations: mechanisms and implications of paramutation. *Nat Rev Genet* **5**: 532-544.

Coe EH (1966) The Properties, Origin, and Mechanism of Conversion-Type Inheritance at the B Locus in Maize. *Genetics* **53**: 1035-1063.

Djupedal I, Ekwall K (2009) Epigenetics: heterochromatin meets RNAi. *Cell Res* **19**: 282-295.

Drewell RA, Brenton JD, Ainscough JF, Barton SC, Hilton KJ, Arney KL, Dandolo L, Surani MA (2000) Deletion of a silencer element disrupts H19 imprinting independently of a DNA methylation epigenetic switch. *Development* **127**: 3419-3428.

Grimaud C, Bantignies F, Pal-Bhadra M, Ghana P, Bhadra U, Cavalli G (2006) RNAi components are required for nuclear clustering of Polycomb group response elements. *Cell* **124**: 957-971.

Haring M, Bader R, Louwers M, Schwabe A, van Driel R, Stam M (2010) The role of DNA methylation, nucleosome occupancy and histone modifications in paramutation. *Plant J*

Josse T, Teysset L, Todeschini AL, Sidor CM, Anxolabehere D, Ronsseray S (2007) Telomeric trans-silencing: an epigenetic repression combining RNA silencing and heterochromatin formation. *PLoS Genet* **3**: 1633-1643.

Kavi HH, Fernandez H, Xie W, Birchler JA (2008) Genetics and biochemistry of RNAi in Drosophila. *Curr Top Microbiol Immunol* **320**: 37-75.

Kavi HH, Fernandez HR, Xie W, Birchler JA (2006) Polycomb, pairing and PIWI--RNA silencing and nuclear interactions. *Trends Biochem Sci* **31**: 485-487.

Lyko F, Brenton JD, Surani MA, Paro R (1997) An imprinting element from the mouse H19 locus functions as a silencer in Drosophila. *Nature Genetics* **16**: 171-173.

Matzke MA, Birchler JA (2005) RNAi-mediated pathways in the nucleus. *Nat Rev Genet* **6**: 24-35.

Patterson GI, Thorpe CJ, Chandler VL (1993) Paramutation, an allelic interaction, is associated with a stable and heritable reduction of transcription of the maize b regulatory gene. *Genetics* **135**: 881-894.

Rassoulzadegan M, Grandjean V, Gounon P, Cuzin F (2007) Inheritance of an epigenetic change in the mouse: a new role for RNA. *Biochem Soc Trans* **35**: 623-625.

Schoenfelder S, Smits G, Fraser P, Reik W, Paro R (2007) Non-coding transcripts in the H19 imprinting control region mediate gene silencing in transgenic Drosophila. *EMBO Rep*

Shpiz S, Kwon D, Rozovsky Y, Kalmykova A (2009) rasiRNA pathway controls antisense expression of Drosophila telomeric retrotransposons in the nucleus. *Nucleic Acids Res* **37**: 268-278.

Sidorenko L, Dorweiler JE, Cigan AM, Arteaga-Vazquez M, Vyas M, Kermicle J, Jurcin D, Brzeski J, Cai Y, Chandler VL (2009) A dominant mutation in mediator of paramutation2, one of three second-largest subunits of a plant-specific RNA polymerase, disrupts multiple siRNA silencing processes. *PLoS Genet* **5**: e1000725.

Stam M, Belele C, Dorweiler JE, Chandler VL (2002) Differential chromatin structure within a tandem array 100 kb upstream of the maize *b1* locus is associated with paramutation. *Genes Dev* **16**: 1906-1918.

Teixeira FK, Colot V (2010) Repeat elements and the Arabidopsis DNA methylation landscape. *Heredity*

Vazquez J, Muller M, Pirrotta V, Sedat JW (2006) The mcp element mediates stable long-range chromosome-chromosome interactions in Drosophila. *Mol Biol Cell* **17**: 2158-2165.

Zaratiegui M, Irvine DV, Martienssen RA (2007) Noncoding RNAs and gene silencing. *Cell* **128**: 763-776.

3.7 SUPPLEMENTARY MATERIAL

Table 3.S1 Primers used in RT-PCR and PCR reactions.

Primer name	Primer sequence (5' → 3')
RT Primers:	
5F1-Ad	GCCTGCCCCAACCTCCAAGAGACATCCACTTAACG
5F2-Ad	GCCTGCCCCAACCTCCAAATATACTGTTGCCGAGC
Reg1-F1-Ad	GCCTGCCCCAACCTCCGTTTGCTGCATCCTTGACC
Reg2-F2-Ad	GCCTGCCCCAACCTCCATTAGAGGGCTCCAAGAGG
Reg1-R1-Ad	GCCTGCCCCAACCTCCGTGAGGGTGAGGTGAATGC
Reg2-R1-Ad	GCCTGCCCCAACCTCCGTATAAAAGTTGTGTACTGC
3R1-Ad	GCCTGCCCCAACCTCCGAGCAATAGTACAGAGAGG
actin-RT	AAATGGCCATGAAGGATGAG
PCR Primers:	
Adapter	GCCTGCCCCAACCTCC
5F1	GCAAGAGACATCCACTTAACG
5F2	CACACCACAAATATACTGTTGCCGAGC
Reg1-F1	GGTTTGCTGCATCCTTGACC
Reg2-F1	AGGGTGTTAAATCCTGAGCG
Reg2-F2	AATTAGAGGGCTCCAAGAGG
Reg2-F3	CACAACTTTTATACCGAATACTTGG
Reg1-R1	TGTGAGGGTGAGGTGAATGC
Reg2-R1	TCGGTATAAAAGTTGTGTACTGC
3R1	GTGAGAGAGCAATAGTACAGAGAGG
actin-F	CACCATGTACCCTGGTATTG
actin-R	TTAAAAGCATTTGCGGTGAAC

Table 3.S2 Primer combinations used in RT and PCR reactions.

Region Analyzed	Transcript orientation	RT Primer	PCR primer used in combination with Adapter primer	Expected band size (bp)
Region 1	Forward	Reg1-R1-Ad	Reg1-F1	223
	Reverse	Reg1-F1-Ad	Reg1-R1	223
Region 2	Forward	Reg2-R1-Ad	Reg2-F1	322
	Reverse	Reg2-F2-Ad	Reg2-R1	180
5' Flanking Region	Forward	Reg1-R1-Ad	5F2	421
	Reverse	5F2-Ad	Reg1-R1	414
3' Flanking Region	Forward	3R1-Ad	Reg2-F3	410
	Reverse	Reg2-F2-Ad	3R1	563
Repeats-Out FRT Region	Forward	3R1-Ad	5F1	377
	Reverse	5F1-Ad	3R1	381

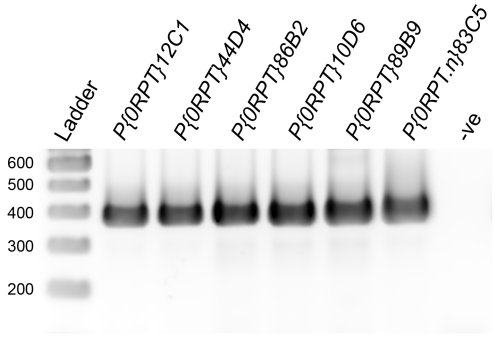


Figure 3.S1 PCR of genomic DNA from transgenic lines with zero tandem repeats, using primers 5F1 and 3R1, confirms that all lines contain the FRT flanking region that was amplified via RT-PCR in Figure 3.3. Expected band size is 367 bp.

3.8 TRANSITION TO CHAPTER 4

In this chapter I present the results of testing various genetic modifiers on the *b1* tandem repeat-mediated silencing in transgenic *Drosophila*. These results indicate that Polycomb group proteins play a role in determining the epigenetic state of the tandem repeats. The role of additional proteins may vary depending on the transgene's insertion site. Additional work is needed to confirm these results in more transgenic lines, to determine the effect of additional RNAi mutations, and to assess the effect of these mutations on the "repeats-out" transgenes.

CHAPTER 4

MODIFIERS OF *B1* REPEAT-MEDIATED SILENCING IN *DROSOPHILA*

4.1 INTRODUCTION

The model organism *Drosophila melanogaster* has been widely used to study a variety of epigenetic mechanisms and processes. Given the conservation of both the core epigenetic mechanisms and the proteins that participate in epigenetic regulation of gene expression, transgenic *Drosophila* have proven a useful tool for examining epigenetic control regions from other species. For example, mammalian imprint control regions function as silencers in transgenic *Drosophila* (Arney *et al.*, 2006; Lyko *et al.*, 1997; Lyko *et al.*, 1998), and bidirectional transcription of the mouse *H19/Igf2* ICR is conserved in *Drosophila* (Schoenfelder *et al.*, 2007). In addition, molecular and genetic analysis of the mouse *H19/Igf2* ICR in *Drosophila* provided significant insight into the mechanism of silencing and the function of the bidirectional transcripts (Schoenfelder *et al.*, 2007).

I have previously shown that the maize *b1* paramutation control region functions as an epigenetic silencer, capable of both *cis* and *trans* silencing, in transgenic *Drosophila* (chapter 2). Further, it is bidirectionally transcribed in *Drosophila*, as it is in maize (chapter 3). Analysis of the proteins involved in repeat-mediated silencing in *Drosophila* may therefore provide insight into the mechanism of silencing at the endogenous maize locus, in addition to the transgenic system. Here I report preliminary results of genetic crosses examining the effects of mutations in genes encoding *Drosophila* RNAi, Polycomb Group (PcG), trithorax Group (trxG), chromatin-modifying, and heterochromatin proteins, on *b1* repeat-mediated epigenetic silencing.

4.2 MATERIALS AND METHODS

All fly stocks were maintained at 21±3°C on a standard cornmeal, yeast, and sugar medium supplemented with 0.15% methylbenzoate (Sigma) as a mould inhibitor. The *P{RPT}* transgenic lines used in modifier testing have previously been described (chapter 2). All standard mutations are described in FlyBase (Tweedie et al., 2009) and were provided by the Bloomington Indiana *Drosophila* Stock Center. Mutant alleles used in modifier testing, with corresponding FlyBase ID number and Bloomington stock number, are indicated in Table 4.1. Mutations generated by P-element insertion into a gene were not tested if the P-element contained a copy of *white,* as this would be expected to increase eye pigmentation due to the increased *white* dosage.

All crosses were conducted in vials with 5 – 10 w^{1118}; *P{2RPT}49E4 / CyO* or w^{1118}; *P{2RPT}89B9, TM3 /* + transgenic females and 3 – 5 males carrying the mutant allele being tested. Crosses were subcultured onto fresh food up to three times after 3-5 days of egg-laying. F1 progeny from these crosses were assessed by comparing eye pigmentation of males with the *P{2RPT}* transgene and mutation, to control sibling males with the *P{2RPT}* transgene and balancer chromosome. Given the position of *P{2RPT}89B9* on the *TM3* balancer chromosome, several 3rd chromosome mutations were tested by generating balanced w^{1118}; *P{2RPT}89B9*, *TM3* / mutation stocks. w^{1118}; *P{2RPT}89B9, TM3 /* + females were crossed to a males carrying a 3rd chromosome mutation, and F1 progeny with the genotype w^{1118}; *P{2RPT}89B9, TM3* / mutation were selected and crossed together. These balanced stocks were propagated for a minimum of 2 – 3 generations before assessing eye phenotype. The effect of modifiers on *white*

expression from *P{0RPT}89B9* was assessed by crossing w^{1118}; *P{2RPT}89B9, TM3 /* mutation females to w^{1118}; *P{0RPT}89B9, TM3* / + males.

Eyes of F1 mutation males, F1 control males, or w^{1118}; *P{2RPT}89B9, TM3 /* mutation balanced males, were photographed at 3 to 6 days post-eclosion, using a Zeiss Stemi 2000-C microscope and Sony DSC-S70 Camera. Wings of males and females were similarly photographed, but without strict aging.

Table 4.1 Mutant alleles used in modifier testing, with corresponding FlyBase Identifier and Bloomington stock number.

Mutant Allele	FlyBase Allele ID	Stock Number
AGO1^{04845}	FBti0005279	11388
ash1^{B1}	FBal0039152	5045
ash2^{1}	FBal0000757	4584
aub^{HN}	FBal0000798	8517
brm^{2}	FBal0001296	3619
brm^{2}, *trx*^{E2}	FBal0001296, FBal0017174	3622
Dl^{7}	FBal0002467	485
Df(2R)vg-B	FBab0002233	752
Df(2R)vg-D	FBab0002235	434
esc^{5}, *E(Pc)*^{1}	FBal0003825, FBal0003311	3142
His2Av^{810}	FBal0005491	9264
JIL-1^{3}	FBal0016510	6347
Lam^{sz18}	FBal0008356	6392
mor^{1}	FBal0012411	3615
osa^{2}	FBal0013299	3616
Pc^{1}	FBal0013551	1728
pho^{b}	FBal0013788	1140
piwi^{06843}	FBti0004305	12225
Psc^{1}	FBal0013980	4200
Rpd3^{04556}	FBti0005513	11633
r2d2^{1}	FBal0151615	8518
Sin3A^{08269}	FBti0005335	12350
Snr1^{01319}	FBti0005425	11529
spn-E^{1}	FBal0016041	3327
su(Hw)^{3}	FBal0016320	672
Su(var)205^{5}	FBal0016507	6234
Su(var)2-1^{1}	FBal0016489	6232
Su(var)3-9^{1}	FBal0016557	6209
Su(var)3-9^{2}	FBal0016558	6210
Su(z)2$^{1.A1}$	FBal0045153	5549
Su(z)3^{1}	FBal0030292	5550
Su(z)7^{1}	FBal0030295	6239
tna^{rl075}	FBti0004238	12080
trx^{E2}	FBal0017174	3621
vas^{1}	FBal0017845	284

4.3 RESULTS

Initial testing of *white* expression in transgenic lines with seven or two *b1* tandem repeats demonstrated that *white* expression in lines with two tandem repeats was more responsive to the presence of a genetic modifier (data not shown). In lines with seven tandem repeats, *white* expression level appeared to be quite stable, with expression changes occasionally observed in some individuals from a modifier cross, but not consistently observed for most individuals from any particular cross. Consistent with this, we have previously shown that epigenetic silencing of *white* is stronger at the same genomic position with seven *b1* tandem repeats than with three *b1* tandem repeats (chapter 2), and that aberrant transcription within the transgenes persists even following repeat removal (chapter 3), suggesting that *b1* mediated epigenetic silencing is robust in *Drosophila*. It is also important to note that many of the tested modifiers are hypomorphic alleles, rather than true amorphs. In addition, most mutations are homozygous lethal, and thus modifier effect on transgenic *white* expression was necessarily examined in flies heterozygous for the mutation (containing one mutant allele and one wild-type allele). A stronger effect on silencing, or disruption of the stable epigenetic silencing observed in the seven-repeat lines, may have been obtained with homozygous null modifiers, but as these are generally inviable, their effect on *white* expression could not be examined. The majority of extensive modifier testing was therefore carried out with two transgenic lines, *P{2RPT}49E4* and *P{2RPT}89B9*, which each contain two tandem repeats.

4.3.1 MODIFIERS OF *P{2RPT}49E4* EXPRESSION

white expression in *P{2RPT}49E4* is normally variable, ranging from light to mid-pink, with the strongest pigmentation observed in a vertical line at the mid-region of the eye (Figure 4.1A). Many mutations in histone-modifying or chromatin genes had no visible effect on *white* expression from the *P{2RPT}49E4* transgene in F1 progeny, including mutations in *Suppressor of variegation 3-9* (*Su(var)3-9*), *Suppressor of variegation 2-1* (*Su(var)2-1*), *Suppressor of zeste 3* (*Su(z)3*), *Rpd3*, and *Sin3A* (Figure 4.1B). However, a mutation in the Notch ligand *Delta* (*Dl*), a mutation in *Suppressor of variegation 205* (*Su(var)205*), which encodes Heterochromatin Protein 1 (HP1), and a mutation in *Jil-1*, caused increased *white* expression from the *P{2RPT}49E4* transgene, indicating that they may normally participate in *b1* repeat mediated silencing (Figure 4.1B).

Mutations in several RNAi genes were also tested and had no readily observable effect, including *Argonaute 1 (AGO1)*, *r2d2*, *spindle E (spn-E)*, *aubergine* (*aub*) and *piwi* (Figure 4.1C). A mutation in *vasa* (*vas*), which encodes an RNA helicase, appeared to cause darker eyes than most control sibling flies (Figure 4.1C). However, several dark-eyed control flies were also observed from this cross. Because the *vasa* mutation was introduced from the male parent, this cannot be due to a maternal effect, so this interaction should be examined further.

Of the trithorax group (trxG) genes tested, mutations in *trithorax* (*trx*), *tonalli* (*tna*), *SNF5-Related 1* (*Snr1*) and *osa* had no effect on eye pigmentation in F1 progeny (Figure 4.1D). Flies with a mutation in *absent, small, or homeotic discs 1* (*ash1*) were lighter than normal for *P{2RPT}49E4*, however, control sibling flies from this cross were lighter than normal as well, with no consistent pigmentation difference observed between

the control and mutant flies. A mutation in the trxG gene *brahma* (*brm*) caused a reduction in *white* expression resulting in lighter eyes, indicating that BRM may normally promote transcription of *white* at the transgene (Figure 4.1D). However, a mutation in *moira* (*mor*), a trxG gene that interacts with *brm,* caused darker eyes, suggesting a role in repeat-mediated silencing.

Several mutations in PcG genes were also tested. A mutation in *Polycomb* (*Pc*), and a double mutation in *extra sex combs* (*esc*) and *Enhancer of Polycomb* (*E(Pc)*) had no visible effect on *P{2RPT}49E4 white* expression (Figure 4.1E). However, a mutation in the PcG gene *pleiohomeotic* (*pho*) caused an increase in *white* expression from the *P{2RPT}49E4* transgene. The increase in *white* expression observed with the pho^{b} allele was more apparent when compared with sibling control flies from the cross, which consistently had lighter eyes than the *pho* mutant flies. Similarly, a mutation in the PcG gene *Posterior sex combs* (*Psc*) also caused an increase in expression. The most drastic increase in expression was observed with deficiency *Df(2R)vgB*, which deletes both *Psc* and the PcG gene *Suppressor of zeste 2* (*Su(z)2*) (Figure 4.1E). This suggests that PcG genes participate in *b1* repeat-mediated silencing from the *P{2RPT}49E4* transgene.

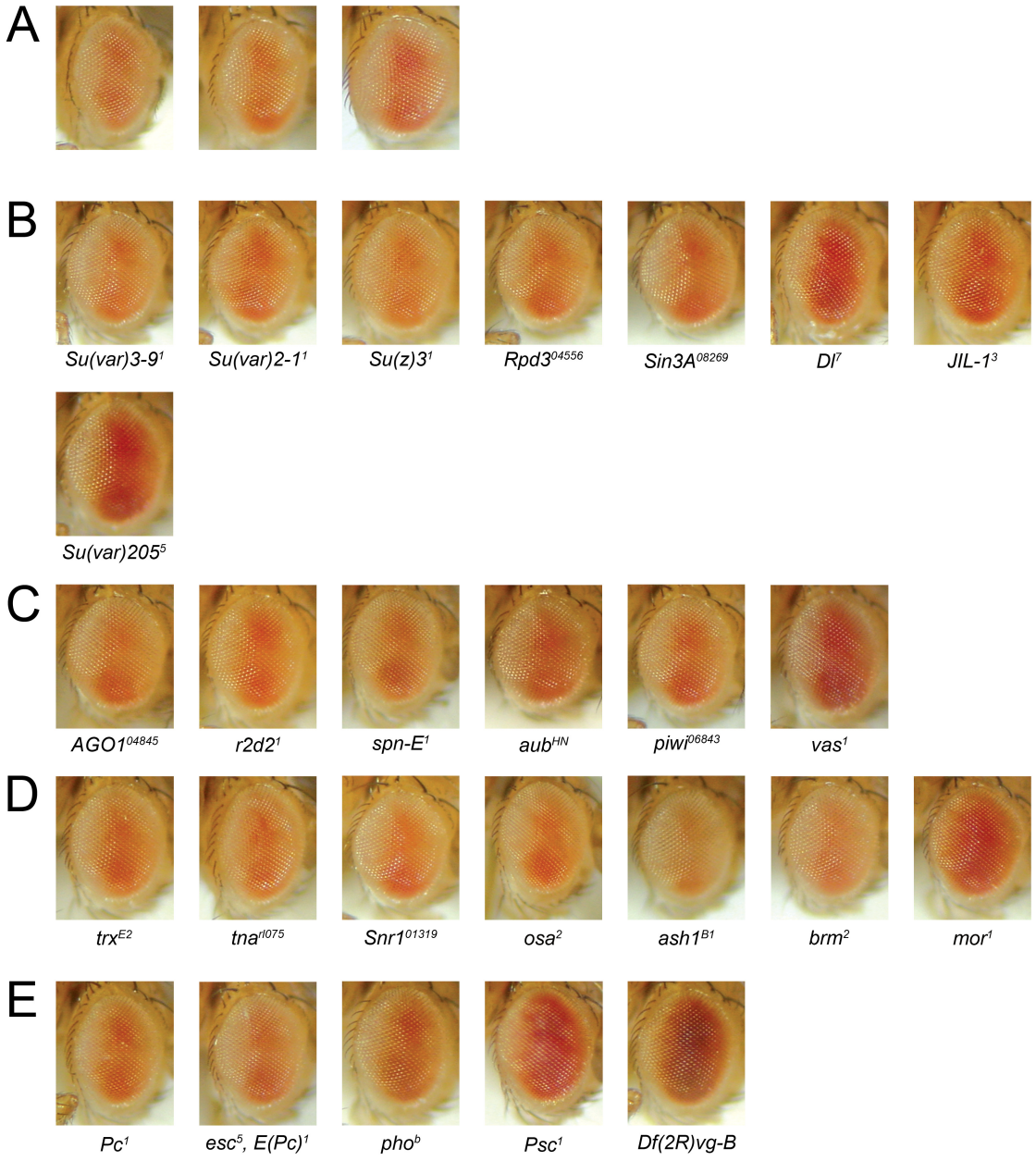


Figure 4.1 Modifiers of *white* expression from the *P{2RPT}49E4* transgene. **[A]** Normal range or *white* expression observed in *P{2RPT}49E4* flies without additional modifiers. **[B-E]** *P{2RPT}49E4 white* expression in the presence of mutations in **[B]** histone- or chromatin-modifying genes, **[C]** RNAi genes, **[D]** trxG genes, and **[E]** PcG genes.

4.3.2 MODIFIERS OF *P{2RPT}89B9* EXPRESSION

The insertion of *P{2RPT}89B9* on the *TM3* balancer chromosome allowed for the generation of balanced stocks containing a 3rd chromosome mutation and the *P{2RPT}89B9* transgene. The normal range of *P{2RPT}89B9 white* expression in the absence of a third chromosome modifier is illustrated in Figure 4.2A. A reduction in *white* expression due to increased silencing is observed as white or light beige eyes with reduced yellow tone, while an increase in *white* expression due to reduced silencing produces darker eye pigmentation with pink, orange or red tones.

Mutations in several 3rd chromosome histone-modifying or chromatin genes were assessed from *P{2RPT}*89B9 / mutation balanced stocks. Two different alleles of the histone methyltransferase *Su(var)3-9* had no visible effect on *white* expression (Figure 4.2B). In *P{2RPT}89B9/Rpd3* stocks, flies with both normal pigmentation and decreased pigmentation (indicative of increased silencing and reduced *white* expression) were observed (Figure 4.2B; $Rpd3^{04556}$ and $Rpd3^{04556}$ HS). Crossing together males and females with decreased pigmentation yielded only light-eyed flies, indicating that the hyper-silenced state is stable once established. Similarly, a mutation in *histone H2A variant* (*His2AV*), also caused increased silencing and reduced *white* expression (Figure 4.2B). In contrast, the $Jil\text{-}1^{3}$ allele, as well as a mutation in *suppressor of Hairy wing* (*su(Hw)*) caused reduced silencing and increased *white* expression, indicating that these wild-type proteins may normally be involved in establishing or maintaining the silenced epigenetic state (Figure 4.2B).

Mutations in 3rd chromosome trxG genes were also assessed from balanced *P{2RPT}*89B9 / mutation stocks. Mutations in *ash1*, *ash2*, *Snr1*, *tna*, and *osa* had no effect on *white* expression from the *P{2RPT}89B9* transgene (Figure 4.2C). A mutation

in *brm*, a mutation and *trx*, as well as a double mutation in both *brm* and *trx*, caused reduced *white* expression, indicating that these proteins may normally promote active transcription at the *P{2RPT}89B9* transgene (Figure 4.2C). Similar to the observed effect on the *P{2RPT}49E4* transgene, a mutation in *mor* caused increased *white* expression from *P{2RPT}89B9*, suggesting it may participate in repeat-mediated silencing (Figure 4.2C).

Additional mutations were examined in F1 progeny inheriting both the transgene and mutation from a cross between *P{2RPT}89B9* females and mutation males. Mutations in chromatin-organizing genes *Su(var)205*, *Suppressor of zeste 7* (*Su(z)7*), and *Lamin* (*Lam*) had no readily apparent effect on *white* expression (Figure 4.2D). A mutation in the signalling ligand *Dl* also had no effect on *P{2RPT}89B9 white* expression (Figure 4.2D).

A mutation in the RNAi gene *AGO1* had no observable effect on *white* expression. However, a mutation in the RNA helicase *spn-E* caused a reduction in *white* expression (Figure 4.2E).

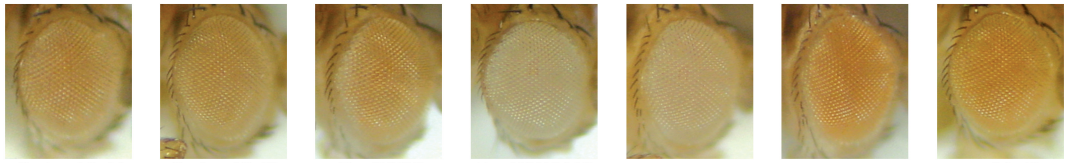
Mutations in PcG genes were also assessed for effect on *P{2RPT}89B9 white* expression (Figure 4.2F). A mutation in *Psc* caused no readily observable effect. However, both normal and high-expressing progeny were observed from a cross between *P{2RPT}89B9* and a mutation stock carrying deficiency *Df(2R)vg-D,* which deletes *Psc* but causes a gain of function in *Su(z)2* (Figure 4.2F). A loss of function mutation in *Su(z)2*, and a chromosome carrying mutations in both *esc* and *E(Pc)*, also caused a moderate but consistent increase in *white* expression in F1 progeny inheriting both the *P{2RPT}89B9* transgene and mutation(s), compared with sibling progeny inheriting the *P{2RPT}89B9* transgene only (Figure 4.2F).

Figure 4.2 Modifiers of *white* expression from the *P{2RPT}89B9* transgene. **[A]** Normal range or *white* expression observed in *P{2RPT}89B9* flies without additional modifiers. **[B-C]** *P{2RPT}89B9 white* expression assessed from balanced stocks carrying the *P{2RPT}89B9* transgene on the *TM3* balancer chromosome, and mutations in **[B]** histone- or chromatin-modifying genes, and **[C]** trxG genes. **[D – F]** *P{2RPT}89B9 white* expression assessed in F1 progeny in the presence of mutations in **[D]** chromatin-organizing or signalling genes, **[E]** RNAi genes, and **[F]** PcG genes.

A

From balanced stocks

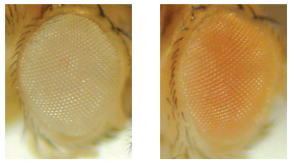
B



Su(var)3-9^{1} *Su(var)3-9*^{2} *Rpd3*^{04556} *Rpd3*^{04556} HS *His2Av*^{810} *JIL-1*^{3} *su(Hw)*^{3}

C

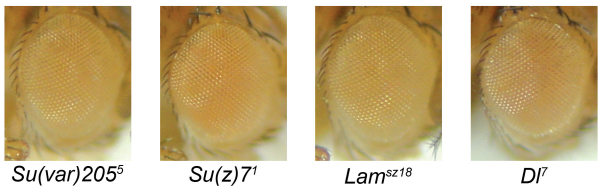
ash1^{B1} *ash2*^{1} *Snr1*^{01319} *tna*^{rl075} *osa*^{2} *brm*^{2} *trx*^{E2}



brm^{2}, *trx*^{E2} *mor*^{1}

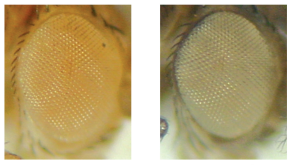
F1 generation from crosses

D



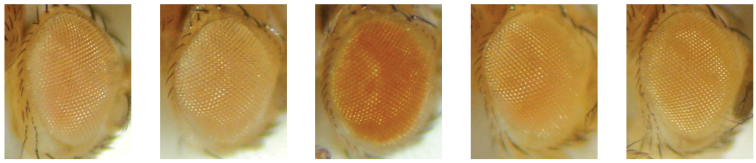
Su(var)205^{5} *Su(z)7*^{1} *Lam*^{sz18} *Dl*^{7}

E



AGO1^{04845} *spn-E*^{1}

F



Psc^{1} *DF(2R)vg-D* *DF(2R)vg-D* *Su(z)2*$^{1.a1}$ *esc*^{5}, *E(Pc)*^{1}

4.3.3 MODIFIERS OF *P{0RPT}89B9* EXPRESSION

To determine whether modifiers were affecting repeat-mediated silencing of *white*, or *white* expression in general, a subset of the mutations tested with the repeats-in strains were crossed to the "repeats-out" version of *P{2RPT}89B9, P{0RPT}89B9,* which normally exhibits orangey-red eyes (Figure 4.3A). A chromosome carrying mutations in both *brm* and *trx*, and a mutation in *His2Av,* caused a reduction in *white* expression from the *P{0RPT}89B9* transgene (Figure 4.3B). Similarly, a mutation in only *brm,* as well as the *spn-E* mutation, also caused a reduction in *white* expression (data not shown). No effect on repeats-out *white* expression was readily visible when *Jil-1* and *su(Hw)* mutations were introduced (Figure 4.3C).

The reduction in *white* expression observed upon the introduction of *His2Av*, *brm*, *spn-E*, and potentially *trx,* mutations may indicate that these proteins affect *white* expression from the *P{2RPT}89B9* transgene independently of the tandem repeats. However, additional testing should be undertaken to confirm these results, as these crosses were conducted within a few generations of repeat removal and isolation of the *P{0RPT}89B9* stock, and so *white* expression or the local chromatin environment may have been affected by a persistent epigenetic mark propagated from the previous heterochromatic state (as shown in chapter 3, Figure 3.3). Similarly, the chromatin structure may have been affected by the recent DNA recombination at the FRT site. The *P{0RPT}89B9* stock was also very weak at this time, and so most crosses only yielded a few *P{0RPT}89B9/*mutation flies to assess. In addition, these crosses were conducted by crossing *P{2RPT}89B9*/mutation balanced females to *P{0RPT}89B9* males and analyzing *P{0RPT}89B9*/mutation progeny. If the protein affected by the mutation binds to the repeats, this may have also affected the results, as the mutation and repeats-in version of

the transgene were present together in the maternal environment, potentially titrating the protein available for deposition into the egg. This could be resolved by crossing *P{0RPT}89B9* males to mutation stocks that have not previously been exposed to the maize tandem repeats. The effect of modifiers could also be assessed with the zero-repeat transgene moved to a new genomic position, *P{0RPT.n}83C5,* to determine whether a perduring epigenetic mark is affecting the results.

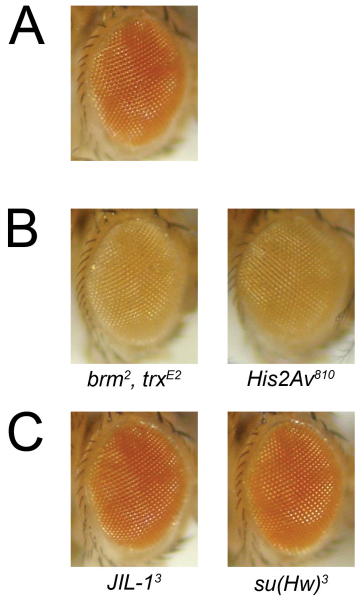


Figure 4.3 Modifiers of *white* expression from the *P{0RPT}89B9* transgene. **[A]** The normal level of *white* expression observed in *P{0RPT}89B9* flies without additional modifiers. **[B]** A double mutation in *brm* and *trx*, and a *His2Av* mutation, caused reduced *white* expression from the *P{0RPT}89B9* transgene. The brm^2 allele and the $spn\text{-}E^1$ allele caused a similar reduction in expression (not shown). **[C]** Mutations in *Jil-1* and *su(Hw)* had no visible effect on *white* expression.

4.3.4 WING PHENOTYPES OBSERVED IN COMBINATION WITH THE *Pc*^{1} ALLELE

Upon crossing several transgenic stocks to the *Pc*^{1} mutation, a high frequency of mutant wing phenotypes were observed in progeny inheriting both the maize repeats transgene and the *Pc*^{1} mutation, despite no readily observable effect on *white* expression. Mutant wing phenotypes were not observed in F1 progeny inheriting only the *Pc*^{1} mutation or the maize repeats transgene. Crossing together flies with these wing phenotypes exacerbated the effect. Wing phenotypes included wings that curve up, under, or are held out, short L4 wing veins, ectopic wing vein material, and wing margin defects (Figure 4.4). Mutant wing phenotypes were observed when three different seven-tandem repeat transgenes were crossed to *Pc*^{1}, including *P{7RPT}17C7*, *P{7RPT}44D4*, and *P{7RPT}89B9*. No mutant wing phenotypes were noted upon crossing to *P{3RPT}10D6* and *P{2RPT}49E4*, which may indicate that a stronger phenotype is obtained with a higher number of *b1* repeats. Overall, this result seems to indicate that the combination of the maize repeats transgene with the *Pc*^{1} mutation is enhancing mutant phenotypes of genes that require PC for normal expression patterns, potentially by accumulating PC protein at the maize repeats and depleting its binding elsewhere. As PC is associated with maintaining transcriptional repression, these phenotypes may be the result of ectopic or aberrant expression of genes that are normally repressed by PcG proteins during development.

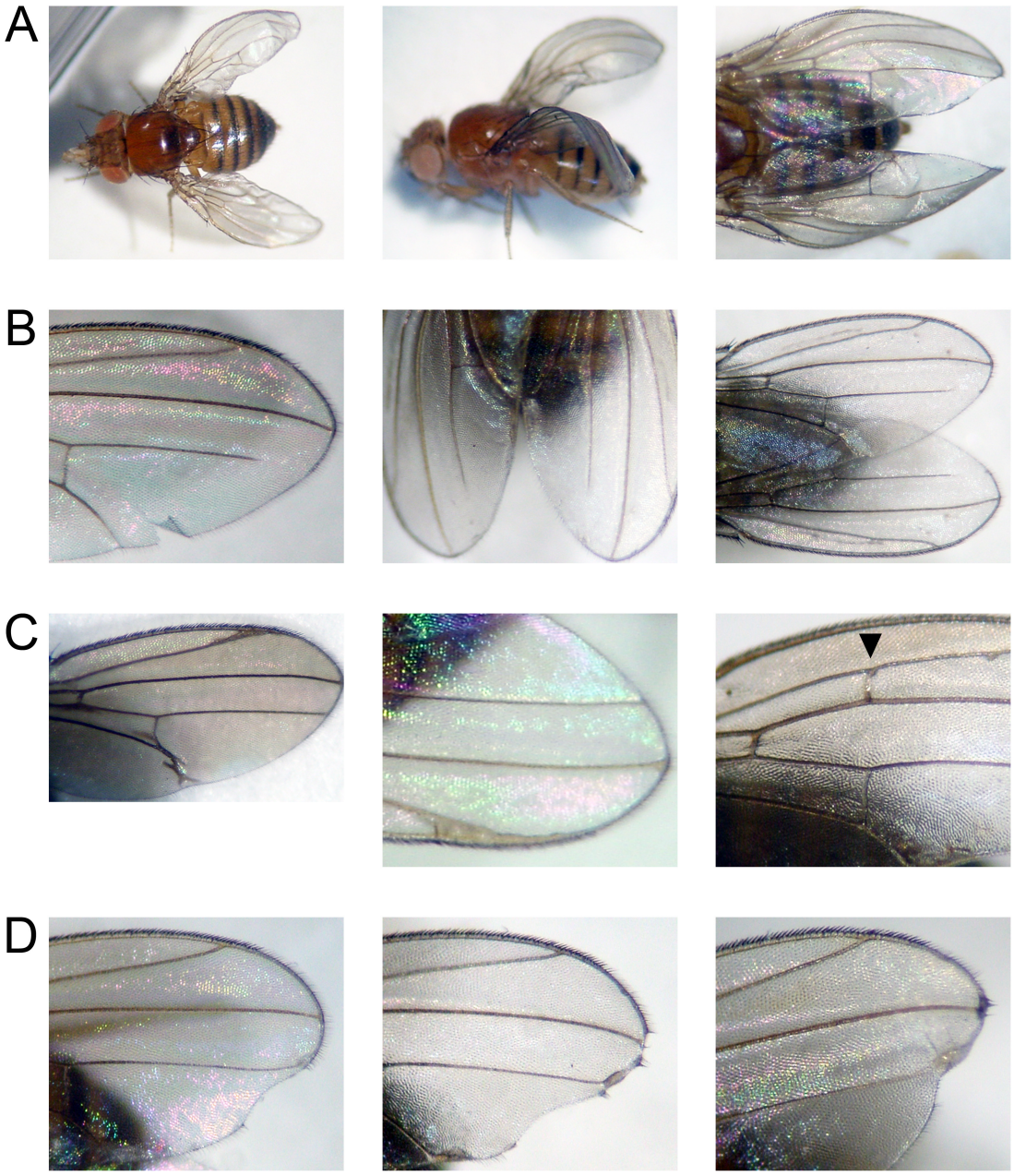


Figure 4.4 Mutant wing phenotypes observed upon crossing *P{7RPT}* transgenes to Pc^{1}. **[A]** Held-out, curved up and curved under wing phenotypes, observed in transgenic lines *P{7RPT}17C7*, *P{7RPT}44D4*, and *P{7RPT}89B9* (from left to right). **[B]** Truncated L4 wing veins, **[C]** ectopic wing veins, and **[D]** wing margin defects observed in *P{7RPT}89B9*.

4.3.5 PRE CONSENSUS MOTIFS WITHIN THE *B1* REPEAT SEQUENCE

PcG proteins associate with regulatory sequences termed PcG response elements (PREs). At these regulatory elements, sequences that bind PcG proteins and trxG proteins are often closely intermingled (Tillib *et al.*, 1999), and the regulatory sequences are thus often termed PRE/TREs. Several DNA motifs have been shown to be important for the function of PRE/TREs (Ringrose and Paro, 2007). These motifs are bound by DNA-binding proteins that then recruit PcG or trxG proteins to maintain transcriptional repression or activation. DNA-binding proteins known to associate with PRE/TREs include PHO, Dorsal Switch Protein 1 (DSP1), GAGA Factor (GAF, also known as Trithorax-like), Pipsqueak (PSQ, which binds to the same target sequence as GAF), Zeste, Grainyhead (GRH), and the SP1/KLF family of DNA binding proteins. Current evidence suggests that GAF, PSQ, DSP1 and Zeste can elicit either activation or repression, while PHO, GRH and SP1/KLF proteins recruit PcG proteins and mediate repression at their target sites (reviewed in Ringrose and Paro, 2004; Ringrose and Paro, 2007). I analyzed the *b1* repeat sequence for putative binding sites of these PRE/TRE-associated DNA-binding proteins (Ringrose and Paro, 2007). Within a single *b1* repeat sequence, there are two potential PHO binding sites, two potential DSP1 binding sites, one potential GAF/ PSQ site, one potential SP1/KLF binding site, one potential Zeste binding site, and three potential GRH sites (Figure 4.5).

CGCCATGGGTTTGCTGCATCCTTGACCCGTAGCCTCACTCACAGCTATGCA
ACTGCTGCAGCCTGTGCAGGCTTAGCCTCAGCCTATCGTGGCCCGACAACT
AAACAGGTCGTGCAGTGCTCTCCCCAAGTCCCGACCACTAATAGTCGTGAT
CCCTGTTTGGAGACGATGACTCGTGGACAAATAGTGCATTCACCTCACCCT
CACACATATTTTTTTTTTGAAATCAAGATCCATTGAACATCTTGTCCAGTT
AAATCACTGGACACCGTGACAGCCACATTGGTTAGTTCAGTTCGTGGTGGA
CCGATGGTTCGCAGTCGCAGCATCACCCTCACACATGGTCCGCATGGCTAC
GCGTATCTATGTTCGTGCGAAGGGTGGTTCTATATATTTAGACTACCTTCA
GTGGATAGGGTAAAATTAATAAAGGACACATAATTATAGATGAGACTATTT
AATATTTTTTTTTACATAAATGAGATTTGAAATGATGACTATGACATGGGA
TTAGATAGGGGATCTGGATTATTGCAGGGTGTTAAATCCTGAGCGATTTTC
GCGGGTGACCAGTGTTTTTTTTACCGTTTTCGATACTAGACGAAATCAAAT
TTCAAAAATCTCACTTAAAATTTTTTATTTTAAAAAAAACCATAAAAATGG
GTAAAAATTTTTATTAAAAATTAGAGGGCTCCAAGAGGTCTATAAAAATTT
GGTGTTTAAAAATTCATGTTTTTGTGCCAAAAAACAAAAGGTTCTAGACC
TGTTAACCCAAATGGATATTGTTGCATCTCCCCCAAATTCTTTATCTACCT
AACACTGCAGTACACAACTTTTATACCGAATACTTGG

PHO	GCCAT
DSP1	GAAAA
GRH	TGTTTTT
ZESTE	YGAGYG
GAF/PSQ	GAGAG
Sp1/KLF	RRGGYGY

Figure 4.5 Putative binding sites for proteins that bind to PRE/TREs in the maize *b1* repeat sequence. The sequence of one *b1* repeat is shown, with the consensus binding sequences indicated below.

4.4 DISCUSSION

In maize, relatively little is known about the mechanisms underlying paramutation and the proteins required to establish and maintain the epigenetic states of the *b1* tandem repeats. Three proteins essential for *b1* paramutation have been identified, and all participate in RNA and siRNA production and amplification (Alleman *et al.*, 2006; Erhard *et al.*, 2009; Hollick *et al.*, 2005; Sidorenko *et al.*, 2009). However, additional unidentified proteins are hypothesized to participate in the establishment and maintenance of the distinct chromatin states (Arteaga-Vazquez and Chandler, 2010). Here I show that the epigenetic state of the *b1*-maize repeats in *Drosophila* appears to require both PcG and trxG genes. PcG and trxG proteins bind to DNA sequences termed Polycomb or Trithorax group response elements (PREs and TREs) to maintain transcriptional repression or activation, respectively (Ringrose and Paro, 2007; Schuettengruber *et al.*, 2007). PREs frequently cause pairing-sensitive silencing (Kassis, 2002), a *trans* silencing effect that I have previously described for several transgenes containing seven *b1* tandem repeats (chapter 2). The observation that *white* expression increased in the presence of a number of *PcG* mutations, and that mutant wing phenotypes were observed upon combining seven-repeat transgenes with the Pc^{1} allele, is consistent with the tandem repeats acting as PREs that are targeted by PcG proteins for transcriptional repression.

The fact that some modifiers of *b1* repeat-mediated silencing in *Drosophila* differ between the two transgenic lines assessed is not surprising. Heterochromatic transgenes in *Drosophila* have previously been shown to be influenced by different modifiers at different genomic positions (Haynes *et al.*, 2007). Spatial positioning within the genome, proximity to other heterochromatin domains, and local determinants at the insertion site,

are all likely important in determining a transgene's ability to recruit, or not recruit, certain proteins (Haynes *et al.*, 2006; Haynes *et al.*, 2007; Sabl and Henikoff, 1996). Nevertheless, some common themes emerge in the modifier analysis of *b1* repeat-mediated silencing in *Drosophila*, including the involvement of PcG proteins.

Mutations in the PcG genes *Su(z)2* and *Psc*, and *pho*, as well as a double mutation in *esc* and *E(Pc)* caused an increase in *white* marker gene expression in one or both of the transgenic lines tested. PSC is a component of Polycomb Repressive Complex 1 (PRC1), a repressive complex that triggers chromatin compaction and transcriptional repression and prevents chromatin remodelling (Francis *et al.*, 2004; Francis *et al.*, 2001). SU(Z)2 is a functional homologue of PSC and exhibits similar effects on chromatin compaction and inhibition of chromatin remodelling. SU(Z)2 can also interact with the members of PRC1, indicating that it may form a PRC1-like complex that could affect different target sites, or function at different developmental times (Lo *et al.*, 2009). Consistent with this, homozygous *Psc* mutations are embryonic lethal, while homozygous *Su(z)2* mutations are lethal later in development, during the larval and pupal stages (Wu and Howe, 1995). Similarly, *Psc* and *Su(z)2* double mutants cause more severe misexpression of Hox genes than single mutants, indicating that the two proteins may substitute for each other (Beuchle *et al.*, 2001). Functional substitution of PSC and SU(Z)2 and/or the activity of the two proteins at different developmental times, likely explains the observation that the most drastic increase in expression was observed in Line *P{2RPT}49E4* when combined with a deficiency that deletes both *Su(z)2* and *Psc* (Figure 6.1F).

A chromosome with *esc* and *e(Pc)* mutations also affected *white* expression from *P{2RPT}89B9*. E(PC) is a suppressor of position-effect variegation (PEV) that also enhances homeotic phenotypes associated with PcG mutations (Sato *et al.*, 1983; Sinclair

et al., 1998). It is a component of the *Drosophila* Tip60 complex, which contains both histone acetylation and chromatin remodelling activities, and is involved in both gene activation and repression (Kusch *et al.*, 2004; Qi *et al.*, 2006; Schirling *et al.*, 2010). ESC is a member of the PRC2 complex, which represses gene transcription via methylation of Lysine 27 on histone H3 (Czermin *et al.*, 2002; Muller *et al.*, 2002). The components of the PRC2 complex and its role in the maintenance of silencing by catalyzing H3K27 methylation is conserved in multicellular eukaryotes, including plants, *C. elegans*, flies, and humans (Bender *et al.*, 2004; Cao *et al.*, 2002; Czermin *et al.*, 2002; Kuzmichev *et al.*, 2002; Muller *et al.*, 2002; Pien and Grossniklaus, 2007; Schubert *et al.*, 2006). In addition, the PRC2 complex has been shown to play an important role in mediating other epigenetic phenomena in plants, including genomic imprinting (Jullien and Berger, 2009). Given the extensive conservation of the PRC2 complex, this complex is a likely candidate for modifying the tandem repeat chromatin structure in both maize and *Drosophila*. Further molecular analysis could be undertaken to confirm whether the PRC2 complex is recruited by the *b1* tandem repeats in *Drosophila.*

A mutation in *pho* caused increased *white* expression from line *P{2RPT}49E4*, but has not been tested on *P{2RPT}89B9.* PHO is a DNA binding PcG protein that is a member of a third PcG protein complex termed PHO-RC (Brown *et al.*, 1998; Klymenko *et al.*, 2006). PHO can interact with PRC2 members ESC and E(Z), as well as PRC1 members PC and PSC, and plays a key role in binding to PcG target sites and recruiting PcG-mediated transcriptional repression (Mohd-Sarip *et al.*, 2002; Wang *et al.*, 2004). Importantly, PHO has been shown to be essential for pairing-sensitive silencing from a PRE at the *Drosophila even skipped* locus (Fujioka *et al.*, 2008), and may thus similarly contribute to pairing-sensitive silencing at the *b1* tandem repeats. Interestingly, PHO

binding to the *even skipped* PRE is also required for transcriptional activation, and a PRE from the *engrailed* gene similarly causes pairing-sensitive silencing but is also required for transcriptional activation (DeVido *et al.*, 2008; Fujioka *et al.*, 2008). An additional PRE from the *abdominal-B* gene causes pairing-sensitive silencing, but mutation or deletion of a DSP1 binding site causes it to switch to a pairing-sensitive activator (Dejardin *et al.*, 2005). The DSP1 protein regulates a wide range of genes, including homeotic genes, and can trigger both activation and repression (Decoville *et al.*, 2001; Dejardin *et al.*, 2005; Lehming *et al.*, 1994). We have previously described both pairing-sensitive silencing and pairing-sensitive activation in our experiments with the *b1* tandem repeats (chapter 2). Potentially, the two seven-repeat transgenes that exhibit activation in *trans* fail to recruit a protein that is essential for establishing pairing-sensitive silencing, and thus the epigenetic state "switches" to activation, similar to the effect of disrupting DSP1 binding at the *abdominal-B* PRE. Overall, several features of the *b1* tandem repeat transgenes are consistent with *Drosophila* PREs, and support the hypothesis that the *b1* tandem repeats function as a PRE and recruit PcG proteins in *Drosophila*. Consistent with this, sequence analysis demonstrates that each *b1* tandem repeat has two copies of the PHO consensus binding sequence, GCCAT, and two copies of DSP1 consensus binding sequence GAAAA (Figure 4.5). Several additional putative binding sites for other DNA-binding proteins known to associate with PREs are also present within the *b1* sequence, including three binding sites for the protein GRH, which can interact with PHO to enhance the binding of both proteins to their target sites (Blastyak *et al.*, 2006).

The wing phenotypes observed upon crossing transgenic stocks to the Pc^{1} allele further supports our hypothesis that the tandem repeats recruit PcG proteins. Wing phenotypes including elevated or crinkled wings, and gaps in the L4 wing vein, have

previously been observed for some *Pc* mutations (Duncan and Lewis, 1982). However, these phenotypes are not normally observed in either the Pc^1 or maize repeats transgenic stocks, and were not observed in control flies inheriting only the transgene or the Pc^1 allele from the crosses. The fact that the wing phenotypes arise when combining the repeats and the Pc^1 mutation, suggests that the repeats are enhancing the phenotypic effect of the Pc^1 allele, potentially by accumulating PC proteins and thereby reducing their binding elsewhere. This would be consistent with the observation that *white* expression was not visibly changed in the presence of the Pc^1 mutation, as the PC protein would still accumulate normally at the tandem repeats, but its binding elsewhere would be reduced.

Mutations in the trxG genes *brm* and *trx* caused a decrease in *white* expression from both *P{2RPT}89B9* and *P{2RPT}49E4* . TRX is a member of trithorax acetylation complex (TAC1), which promotes active gene expression via histone H3K4 methyltransferase activity, and histone acetyltransferase activity (Petruk *et al.*, 2001; Smith *et al.*, 2004). BRM is a member of the ATP-dependent BRM chromatin remodelling complex, a complex that is widely associated with regions of transcriptionally active chromatin in *Drosophila* (Armstrong *et al.*, 2002; Papoulas *et al.*, 1998). The observed effects of these two mutations on *white* expression are consistent with the role of trxG proteins in the maintenance of endogenous active epigenetic states. Sequences that bind PcG and trxG proteins are frequently intermingled (Tillib *et al.*, 1999), and thus the maize *b1* repeats may contain response elements that recruit both groups of proteins. Several of the PRE/TRE-associated proteins with putative binding sites in the *b1* repeats (Figure 4.5) are able to mediate gene repression or activation, including GAF, PSQ, DSP1 and Zeste (reviewed in Ringrose and Paro, 2007), and may

thus recruit trxG proteins and promote the formation of active chromatin at the repeats. However, given the preliminary testing results with *P{0RPT}89B9,* additional testing should be undertaken to confirm whether these trxG proteins are affecting *white* expression by targeting sequences in the maize tandem repeats, or by more generally increasing *white* expression from the transgenes. Indeed, these proteins may be recruited to the *hsp70* promoter of the *white* gene rather than the *b1* tandem repeats, as TRX has been shown to be recruited to *hsp70* promoters following a heat-shock (Smith *et al.*, 2004). While our transgenes were not heat-shocked, the *hsp70* promoter may similarly recruit trxG complexes under basal conditions. However, heat shock loci are not associated with the BRM complex, and so the *hsp70* promoter would not be expected to recruit BRM and the BRM complex (Armstrong *et al.*, 2002).

An intriguing result is that the trithorax group gene *moira*, which is an essential component of the BRM chromatin remodelling complex, appears to have the opposite effect on the maize *b1* repeats. A mutation in *moira* resulted in an increase in *white* expression from both *P{2RPT}49E4* and *P{2RPT}89B9*, suggesting that it participates in repeat-mediated silencing rather than mediating an active chromatin structure. The BRM complex is a dynamic chromatin remodelling complex that participates in both gene activation and repression (Armstrong *et al.*, 2002; Marenda *et al.*, 2004; Moller *et al.*, 2005), and thus it may serve a dual role at the maize *b1* repeats. Alternatively, mutations in *mor* or *brm* may be causing an increase in *white* expression indirectly, via the misexpression of other genes that are targeted by the BRM complex, with the two mutations potentially having different effects on other gene expression targets. MOR may also play a role in epigenetic gene repression that has not yet been described. It is also possible that one of these modifier results is an artifact and the other is the true effect

of the BRM complex at the maize *b1* tandem repeats. Given the observation that the *brm* mutation caused reduced expression of the *P{0RPT}89B9* transgene, the effect of the *mor* mutation is perhaps the best reflection of the role of the BRM complex in *b1* silencing; that is, that the BRM complex contributes to the formation of a repressive chromatin structure.

Mutations in the histone deacetylase *Rpd3* also result in increased silencing from the *white* transgene, suggesting that RPD3 promotes active chromatin at the *b1* tandem repeats. Histones are generally acetylated throughout active chromatin regions, and hypoacetylated in heterochromatin; therefore histone deacetylases are frequently implicated in mediating heterochromatin formation and gene repression. While RPD3 is often involved in *Drosophila* gene repression (Pile and Wassarman, 2000), it has also implicated in the deacetylation of Lysine 12 on histone H4 (Rundlett *et al.*, 1996), which is an abundant histone modification in heterochromatin, and it appears to promote active chromatin and gene expression at other loci (Cho *et al.*, 2005). In agreement with our results, mutations in *Rpd3* have previously been shown to increase silencing associated with PEV in *Drosophila*, and telomeric position effect in yeast (De Rubertis *et al.*, 1996).

The histone methyltransferase SU(VAR)3-9 does not appear to affect silencing from the *b1* repeats. This is consistent with the observation that modifiers of PEV and PcG proteins rarely overlap (Sinclair *et al.*, 1998). Similarly, PcG proteins appear to participate in tandem repeat silencing at *Drosophila* subtelomeric heterochromatin, but typical modifiers of PEV such as SU(VAR)3-9 do not (Boivin *et al.*, 2003; Cryderman *et al.*, 1999; Doheny *et al.*, 2008; Wallrath and Elgin, 1995). However, the *Jil-1*^{3} allele, which has previously been reported to be a strong suppressor of PEV (Ebert *et al.*, 2004), caused an increase in *white* expression for both transgenic lines. JIL-1 encodes a histone

H3 kinase; however, the *Jil-1*^{3} mutation generates a truncated protein that retains kinase function (Ebert *et al.*, 2004), but is mislocalized to ectopic sites (Zhang *et al.*, 2006). Recent research has demonstrated that ectopic histone phosphorylation by JIL-1 can induce a dramatic change in chromatin structure from condensed heterochromatin to open euchromatin (Deng *et al.*, 2008). Thus it is likely that aberrant histone phosphorylation in *Jil-1*^{3} mutants results in decreased heterochromatinization at the *b1* transgenes, causing a loss of silencing.

A mutation in *Su(var)205*, which encodes HP1, caused increased *white* expression from *P{2RPT}49E4* but not *P{2RPT}89B9*, which could indicate that protein accumulation at the *b1* tandem repeats may be affected by the genomic position of the transgene insert, as has been described for other heterochromatic *Drosophila* transgenes (Haynes et al., 2006; Haynes et al., 2007; Sabl and Henikoff, 1996). Similarly, a mutation in the DNA-binding insulator protein SU(HW) caused increased *white* expression from *P{2RPT}89B9* but has not yet been assessed with line *P{2RPT}49E4*. SU(HW) has previously been shown to regulate silencing from the mouse *H19* imprint control region in *Drosophila* (Schoenfelder and Paro, 2004), and may thus play a similar role in the establishment or maintenance of *b1* repeat-mediated silencing.

The effect of a mutation in *His2Av*, a histone H2A variant that localizes to centromeric heterochromatin (Swaminathan *et al.*, 2005), was assessed with line *P{2RPT}89B9*, in which it caused a reduction in transgenic *white* expression. *His2Av* mutations have been shown to have varying effects on PEV, either increasing, decreasing, or having no effect on *white* expression, depending on the position that is assessed (Haynes *et al.*, 2007; Swaminathan *et al.*, 2005). H2AV may thus play a versatile and dynamic role in the formation of distinct chromatin domains. However, the reduction in

white expression from *P{0RPT}89B9* may indicate that the *His2Av*^{810} mutation affects *white* expression more generally, potentially by modulating the expression of other target genes that influence *white* expression, rather than influencing the epigenetic status of the maize repeats directly.

A mutation in the RNAi gene *spn-E*, an RNA helicase required for repeat-associated small interfering RNA (rasiRNA) mediated silencing (Klenov *et al.*, 2007), caused a reduction in *white* expression from *P{2RPT}89B9.* rasiRNAs are a subset of piwi-interacting RNAs (piRNAs) (Faehnle and Joshua-Tor, 2007), and recent research at the tandem repeats of the 3R subtelomeric region (3R TAS) has demonstrated a similar result, with piRNAs produced from the 3R TAS counteracting heterochromatinization (Yin and Lin, 2007). However, no effect was observed on *white* expression from *P{2RPT}49E4*, and a reduction in *white* expression was observed from *P{0RPT}89B9*, the "repeats-out" version of *P{2RPT}89B9,* which may indicate that this result is an artifact of the 89B9 insertion site.

Similar to *spn-E, vasa* encodes an RNA helicase that has been implicated in retrotransposon silencing in ovaries and the developing oocytes (Vagin *et al.*, 2004). Given the potential effect of the *vas*^{1} mutation on *white* expression from *P{2RPT}49E4,* as well as the observed bidirectional transcription of the *b1* tandem repeats (chapter 3), additional RNAi mutants should be tested further to determine whether RNAi pathways participate in either establishment or maintenance of *b1* mediated repeat silencing. As members of the PIWI subfamily of RNAi proteins are expressed at high levels in female germ cells as well as somatic cells within the ovary (Saito *et al.*, 2006; Williams and Rubin, 2002), assessing *eGFP* expression from a *P{RPT}* transgene that expresses *eGFP* at detectable levels, such as *P{7RPT}12C1*, in the ovaries of females in combination with

various RNAi mutations may prove a valuable strategy. Reciprocal crosses, where the mutation is introduced maternally and the *P{RPT}* transgene is introduced paternally, should also be useful in examining maternal or early effects of the RNAi mutations.

Overall, these results indicate that *b1* tandem repeat silencing in *Drosophila* is likely mediated by PcG proteins. This hypothesis is consistent with that observation that several transgenic lines with the *b1* tandem repeats exhibit pairing-sensitive silencing, an epigenetic effect that is frequently observed with transgenes carrying PREs. Accumulation of additional silencing proteins may vary depending on the transgene insertion site. Additional modifier testing with other transgenic lines and "repeats-out" lines should confirm the role of PcG proteins in *b1* tandem repeat silencing.

4.5 TRANSITION TO CHAPTER 5

Here I describe unique *white* expression patterns from two variant transgenic lines. A transgene inserted at the same genomic position as *P{7RPT}44D4*, but in the opposite orientation, exhibits pairing-sensitive silencing with itself, but impaired pairing-sensitive silencing when combined with *P{7RPT}44D4.* This result demonstrates that pairing-sensitive silencing of *white* in the *b1* transgenes is dependent on the orientation of the maize *b1* repeats and/or *white* genes that are present on the paired homologous chromosomes. I also describe a repeats-out variant line that contains a rearrangement within the transgene and exhibits a higher level of *white* expression, but is still susceptible to silencing in *trans* by the "repeats-in" progenitor line. This result indicates that heterozygous silencing of repeats-out *white* in *trans* by the *b1* repeats is relatively unaffected by the expression level of the repeats-out *white* gene.

CHAPTER 5

CHARACTERIZATION OF *WHITE* EXPRESSION IN VARIANT LINES

5.1 INTRODUCTION

Two variant transgenic lines, exhibiting distinct expression patterns that differ from those of their respective progenitor transgenic lines, have been isolated and characterized. Line *P{7RPT}44D4+* is a variant line of Line *P{7RPT}44D4*, inserted at the same genomic position and containing the same number of tandem repeats, but with the transgenic sequence inserted in the opposite orientation relative to the surrounding genomic sequence. Line *P{0RPT.V}12C1* is a variant line of *P{0RPT}12C1*, the "repeats-out" version of *P{7RPT }12C1*. It is inserted at same genomic position as *P{7RPT}12C1* and *P{0RPT}12C1*, and contains zero tandem repeats, but contains a rearrangement within the transgenic sequence.

5.2 MATERIALS AND METHODS

Fly stocks were maintained and fly crosses were conducted as previously described in section 4.1.1. Line *P{7RPT}44D4+* was isolated by crossing line *P{7RPT}44D4* to w^{1118}; ry^{506}, Dr^{1}, *P{Δ2-3}99B / TM6C,* Sb^{1} (FBst0005908), which expresses Delta 2-3 transposase, and selecting for changed eye pigmentation. Line *P{0RPT.V}12C1* arose spontaneously during stock propagation of *P{0RPT}12C1*. The transgene insertion sites of *P{7RPT}44D4+* and *P{0RPT.V}12C1* were confirmed using inverse PCR, as previously described in chapter 2. For inverse PCR of *P{0RPT.V}12C1*, *Pvu*II was used for restriction digestion instead of *Msp*I, as *Msp*I failed to produce a significant band for the 3' P-element. Inverse PCR sequencing data of *P{0RPT.V}12C1* provided the DNA sequence of the first 746 bp of the *P{0RPT.V}12C1* transgene. The

orientation of the *eGFP* gene was determined by PCR using the primers 5F2 (5'-CACACCACAAATATACTGTTGCCGAGC-3') and GFP-F (5'-ATCGTTCGAAGAGCGCCGGAGTATAAATAG-3'), which yields a 2565 bp band from the rearranged *P{0RPT.V}12C1* transgene, but no band from the non-rearranged *P{0RPT}12C1* transgene.

Fly eyes were photographed as described previously in section 4.1.1. Eye images presented in Figure 5.1 are of adult females, while both males and female eye images are presented in Figure 5.2. Repeats-in / repeats-out heterozygotes were created by crossing repeats-in females to repeat out males. *P{7RPT}44D4+* / *P{7RPT}44D4* heterozygotes were generated by crossing *P{7RPT}44D4+* / *CyO* females to *P{7RPT}44D4 / CyO* males. *P{0RPT}44D4+* / *P{0RPT}44D4* heterozygotes were generated in the same manner, using flies from the *P{0RPT}* stocks rather than *P{7RPT}*. Pigment assays were conducted as previously described previously (chapter 2). Statistical significances were determined using ANOVA and two-tailed unpaired student's t-tests.

PCR and RT-PCR of *P{0RPT.V}12C1* and *P{0RPT}12C1* was conducted as described in chapter 3, but with 35 PCR cycles instead of 40.

5.3 RESULTS AND DISCUSSION

5.3.1 *WHITE* EXPRESSION IN LINE *P{7RPT}44D4+*

While generating additional transgenic lines by exposing line *P{7RPT}44D4* to a source of Delta 2-3 transposase, line P*{7RPT}44D4+* was isolated. In line *P{7RPT}44D4+*, the transgene is inserted at the same genomic position as *P{7RPT}44D4*, but with the full transgenic sequence (3' P-end, repeats, *white*, *eGFP*, 5'

P-end) in the opposite orientation relative to the surrounding genomic sequence (Figure 5.1A). This line was named *P{7RPT}44D4+* to denote this orientation change.

As is the case with line *P{7RPT}44D4*, line *P{7RPT}44D4+* exhibits pairing-sensitive silencing in repeats-in homozygotes, and *trans* silencing of *white* in repeats-in / repeats-out heterozygotes. Following repeat removal, homozygotes were not viable, and so pairing-sensitive silencing in repeats-out homozygotes could not be assessed (Figure 5.1B).

Significantly, when *P{7RPT}44D4* and *P{7RPT}44D4+* were combined in heterozygotes, consistent pairing-sensitive silencing of *white* was not observed. Heterozygous *P{7RPT}44D4 / P{7RPT}44D4+* flies were variable in eye pigmentation, but always darker than *P{7RPT}44D4* hemizygotes, with most flies also darker than or approximately equivalent to *P{7RPT}44D4+* hemizygotes (Figure 5.1C). However, approximately 10-20% of heterozygous flies exhibited eye pigmentation that was lighter than *P{7RPT}44D4+* hemizygotes, suggesting that weak pairing-sensitive silencing may occasionally be established when the two transgenes are combined. This result indicates that the orientation of the tandem repeats and/or the *white* marker gene is important in establishing pairing-sensitive silencing between the *b1* repeat and *white* sequences on paired homologues. Similarly, the orientation of tandem repeats from the 2L TAS was previously found to be important for the establishment of silencing (Kurenova *et al.*, 1998), and the orientation of tandem repeats from the X chromosome TAS was important for the establishment of pairing-sensitive silencing (Boivin *et al.*, 2003). In the case of the 44D4 insertions, the orientation of the tandem repeats relative to the *white* gene would be correct to establish silencing and pairing-sensitive silencing of each transgene when

paired with itself, but may be inhibitory in establishing pairing-sensitive with a paired transgene in which the repeats are in the opposite orientation.

No appreciable silencing in *trans* was observed when the two repeats-out versions of the transgenes were combined, despite the pairing-sensitive silencing that is observed in the *P{0RPT}44D4* line (Figure 5.1C). Whatever silencing factors continue to accumulate at *P{0RPT}44D4* and result in pairing-sensitive silencing, must fail to accumulate at *P{0RPT}44D4+*, and so no substantial reduction in expression is observed upon combining the two repeats-out transgenes.

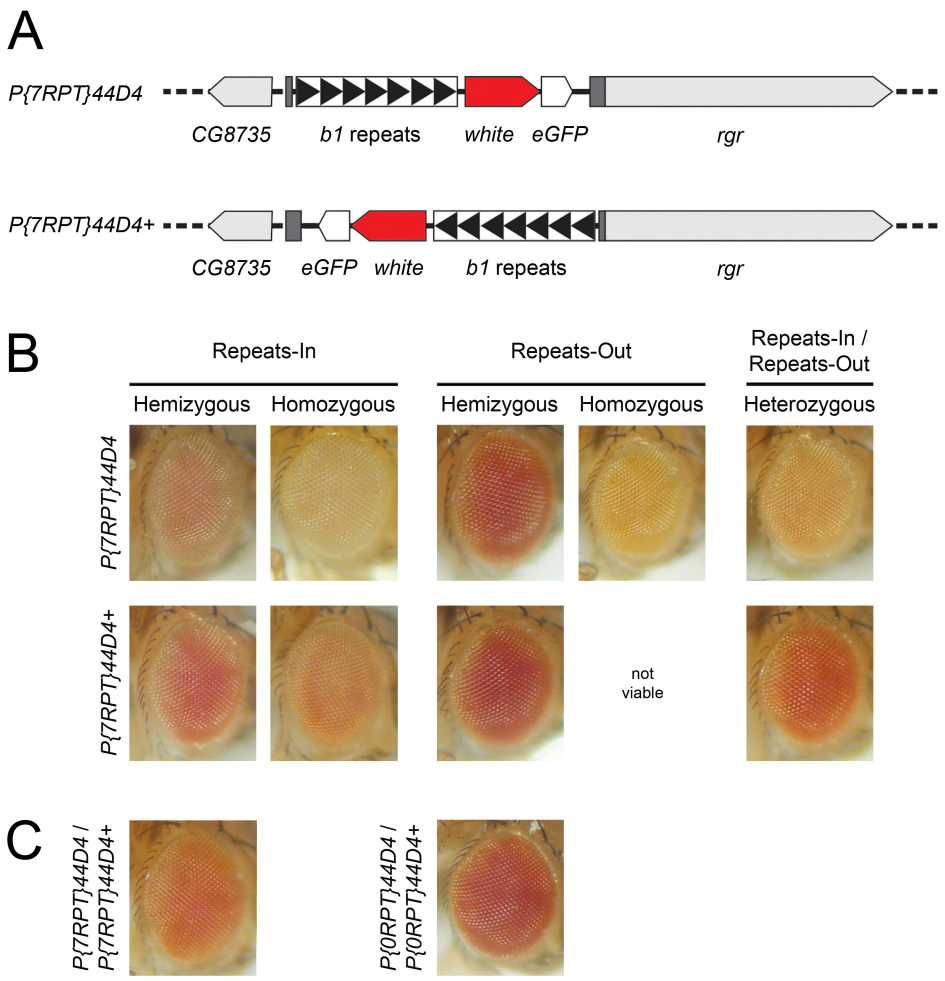


Figure 5.1 Characterization of the *P{7RPT}44D4+* transgene. **[A]** *P{7RPT}44D4+* is inserted at the same genomic position as *P{7RPT}44D4,* but with the full transgene in the opposite orientation relative to the surrounding genomic sequence. **[B]** Eye pigmentation of *P{7RPT}44D4* and *P{7RPT}44D4+* repeats-in and repeats-out hemizygous and homozygous flies, as well as repeats-in / repeats-out heterozygous flies. **[C]** Average observed eye pigmentation of adult females carrying paired copies of the two transgene versions, with either the repeats-in or repeats-out.

5.3.2 *WHITE* EXPRESSION IN LINE *P{0RPT.V}12C1*

Line *P{0RPT.V}12C1* is a repeats-out variant of line *P{7RPT}12C1*. It has been isolated as an independent transgenic stock without the maize *b1* tandem repeats for approximately five years, and recent sequence analysis indicated that a rearrangement has occurred within this transgene insertion. As this rearrangement is not detected in the repeats-in version of this line, it must have occurred within this specific isolate of the *P{0RPT}12C1* stock following repeat removal. The current *P{0RPT}12C1* stock used for eye pigment analysis has had the repeats removed for approximately three years. Full sequence data of the rearranged transgene is not currently known, but initial DNA sequencing and PCR analysis suggests that the first 422 bp of the transgene is intact, followed by an rearrangement that places the 5' P-end inverted and adjacent to the break site at 422 bp (Figure 5.2A). Accordingly, this line was named *P{0RPT.V}12C1* to denote that it is a variant of *P{0RPT}12C1*.

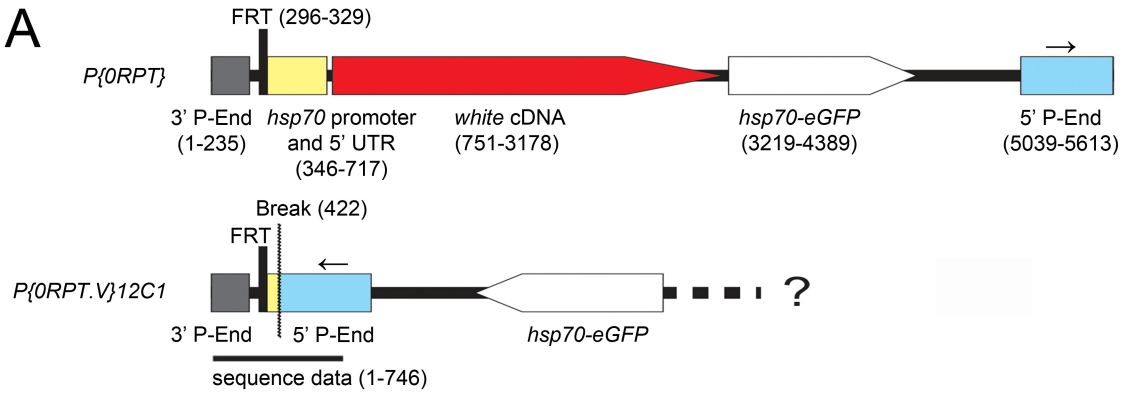
white expression in line *P{0RPT.V}12C1* is increased compared with the non-rearranged repeats-out transgene, in both males and females (Figure 5.2B – E). In addition, line *P{0RPT.V}12C1* exhibits a greater difference in *white* expression between hemizygous and homozygous females than the non-rearranged repeats-out transgene, indicating a greater loss of pairing-sensitive silencing (Figure 5.2B and D). Intriguingly, however, when *P{0RPT.V}12C1* is combined with *P{7RPT}12C1* in heterozygous females, a significant decrease in *white* expression is observed, compared with *P{0RPT.V}12C1* hemizygous females (Figure 5.2B and D). The level of *white* expression in *P{0RPT.V}/P{7RPT}* heterozygotes is similar to that of *P{0RPT}/P{7RPT}* heterozygotes; however, because the normal level of *white* expression in *P{0RPT.V}12C1* hemizygotes is higher, the fold decrease and observed effect on eye pigmentation when

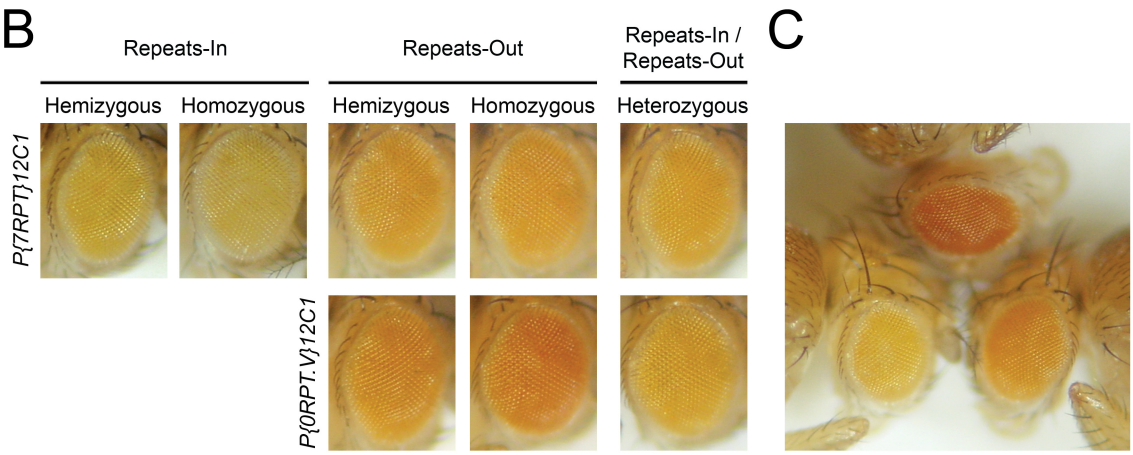
combined with the repeats-in transgene is more substantial and noticeable. *white* expression in *P{0RPT.V}/P{7RPT}* heterozygotes exhibits a 1.35 fold decrease compared with *P{0RPT.V}* hemizygotes, whereas *white* expression in *P{0RPT}/P{7RPT}* heterozygotes is approximately equivalent to *P{0RPT}* hemizygotes (1.06 fold decrease; Figure 5.2D). This result appears to indicate that when the repeats-in and repeats-out transgenes are combined in *trans*, the accumulation of silencing factors reduces *white* expression to a lower "set expression level", rather than reducing it by a certain degree or fold decrease, depending on how much it is normally transcribed. The observed level of expression in heterozygotes may be reflective of the basal level of transcription that occurs from the paired transgenes due to the accumulated epigenetic modifications, and may be relatively unaffected by the original transcription level of the repeats-out transgene. Thus repeats-out lines with a higher overall level of *white* expression (observed as darker eye pigment) may exhibit a more drastic decrease in expression when combined with the repeats-in version. Consistent with this, the reduction in *white* expression in *trans* in repeats-in / repeats-out heterozygotes is most readily apparent in line *P{7RPT}86B2*, which exhibits the highest level of *white* expression in *P{0RPT}86B2* hemizygotes (chapter 2). Additional sequencing should confirm the exact position and orientation of the *white* gene in the recombined *P{0RPT.V}12C1* transgene.

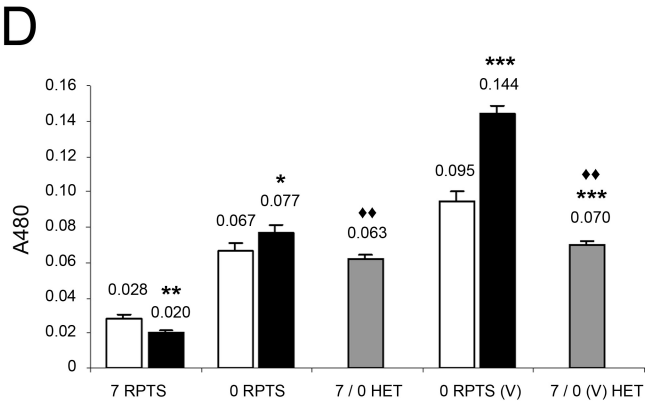
I have previously shown that bidirectional transcription within the non-rearranged *P{0RPT}12C1* transgene can be detected for the 368 bp region encompassing the FRT site, which previously contained the maize *b1* tandem repeats. This region remains intact in the rearranged *P{0RPT.V}12C1* transgene (Figure 5.2F), and can be PCR amplified from *P{0RPT.V}12C1* DNA (Figure 5.2G). However, only trace amounts of either forward or reverse transcription could be detected for this region of the rearranged

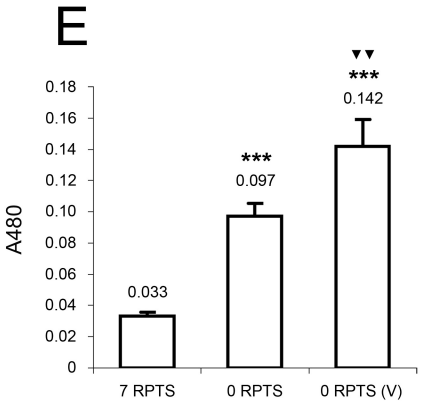
transgene. The loss of this transcription may be due to the increased time the stock has been isolated from the repeats-in version, potentially due to the loss of an epigenetic mark or memory of the previous epigenetic state. Alternatively, the origin of the relatively robust reverse transcript that is detected in *P{0RPT}12C1* and other repeats-out lines may be within the rearranged portion of the *hsp70* promoter, UTR, or *white* gene. Additional characterization of the *P{0RPT.V}12C1* transgene could determine the full sequence of the rearrangement, and RT-PCR analysis of the moved *hsp70-white* region could confirm whether a reverse transcript persists at this region following the genomic rearrangement.

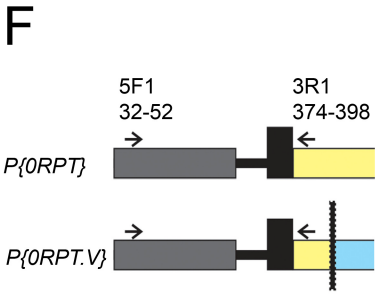
Figure 5.2 Characterization of the *P{0RPT.V}12C1* transgene. **[A]** The sequence of all *P{0RPT}* transgenes is illustrated to scale. *P{0RPT.V}12C1* contains a break within the *hsp70* promoter that places the 5' P-end adjacent to the break site, and in the opposite orientation (illustrated with an arrow). Sequence analysis confirms the first 746 bp of *P{0RPT.V}12C1*. The orientation of the *eGFP* gene is inferred from PCR analysis. The sequence of the remainder of the transgene is currently unknown. **[B]** Eye pigmentation of *P{7RPT}12C1* repeats-in and repeats-out hemizygous and homozygous females (top row), *P{0RPT.V}12C1* hemizygous and homozygous females (bottom row), and both versions of repeats-in / repeats-out heterozygous females. *P{0RPT.V}12C1* females exhibit a greater increase in *white* expression and a greater difference between hemizygotes and homozygotes than *P{0RPT}12C1*, but an observable reduction in pigmentation when combined with the repeats-in transgene. **[C]** Eye pigmentation of *P{7RPT}12C1* (left), *P{0RPT}12C1* (right) and *P{0RPT.V}12C1* (top) males. Eye pigmentation is greatest in the repeats-out variant line (*P{0RPT.V}12C1*). **[D]** Pigment assay quantification of *white* expression in hemizygous (white bars), homozygous (black bars), and heterozygous (grey bars, "HET") females, with seven tandem repeats, zero tandem repeats, or the rearranged zero repeat variant transgene (V). Pigmentation in both the zero repeat and zero repeat variant lines is significantly increased from the transgenic line with seven tandem repeats ($p<0.001$). Pigmentation in the zero repeat variant line is also significantly increased from the zero repeat line ($p<0.001$). Additional statistical significance is indicated for homozygous values significantly different than hemizygous (*), and heterozygous values significantly different from the corresponding zero repeats hemizygous (*) or homozygous (♦) values. Significances are indicated at the level of $p<0.05$ (*), $p<0.01$ (**) or $p<0.001$ (***), for each symbol. **[E]** Pigment assay quantification of hemizygous males. *white* expression is significantly increased in zero-repeat males compared with seven repeat males (*), and in zero repeat variant males (V) compared with zero-repeat males (▼). Significances are indicated at the level of $p<0.05$ (*), $p<0.01$ (**) or $p<0.001$ (***), for each symbol. **[F]** A schematic illustrating the DNA region and primers used to measure transcription in the zero repeat and zero repeat variant lines. **[G]** PCR analysis confirms the region is intact and amplifiable in both *P{0RPT}12C1* and *P{0RPT.V}12C1*. **[H]** RT-PCR analysis demonstrates that aberrant transcription persists at this region in the *P{0RPT}12C1* line, but is mostly undetectable in *P{0RPT.V}12C1*.

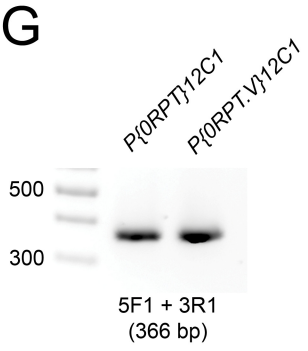

A
FRT (296-329)
P{0RPT}
3' P-End (1-235)
hsp70 promoter and 5' UTR (346-717)
white cDNA (751-3178)
hsp70-eGFP (3219-4389)
5' P-End (5039-5613)
Break (422)
FRT
P{0RPT.V}12C1
3' P-End
5' P-End
hsp70-eGFP
?
sequence data (1-746)

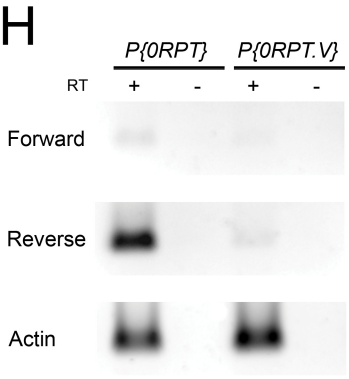

B
Repeats-In
Repeats-Out
Repeats-In / Repeats-Out
Hemizygous
Homozygous
Hemizygous
Homozygous
Heterozygous
P{7RPT}12C1
P{0RPT.V}12C1
C


D
A480
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0
0.028
**
0.020
0.067
*
0.077
♦♦
0.063
0.095
***
0.144
♦♦
***
0.070
7 RPTS
0 RPTS
7 / 0 HET
0 RPTS (V)
7 / 0 (V) HET


E
A480
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0
0.033
***
0.097
♦♦
***
0.142
7 RPTS
0 RPTS
0 RPTS (V)


F
5F1
32-52
3R1
374-398
P{0RPT}
P{0RPT.V}


G
P{0RPT}12C1
P{0RPT.V}12C1
500
300
5F1 + 3R1
(366 bp)


H
P{0RPT}
P{0RPT.V}
RT
+
-
+
-
Forward
Reverse
Actin

5.4 TRANSITION TO CHAPTER 6

The following chapter provides an extensive review of the molecular mechanisms underlying genomic imprinting in plants, insects, and mammals. Genomic imprinting is an epigenetic process similar to paramutation, in which parent-specific epigenetic marks are stably transmitted to progeny. This review highlights similarities and evidence that demonstrates that the molecular mechanisms of imprinting are based on core epigenetic processes that are evolutionary conserved, and can be exploited to produce seemingly unique patterns of gene expression. This chapter will be included in the text book *Epigenetics: Linking Genotype and Phenotype in Development and Evolution* (University of California Press), currently in production and scheduled for 2010 publication.

CHAPTER 6

THE EPIGENETICS OF GENOMIC IMPRINTING:

CORE EPIGENETIC PROCESSES ARE CONSERVED IN MAMMALS, INSECTS AND PLANTS

Lori A. McEachern and Vett Lloyd

In: Epigenetics: Linking Genotype and Phenotype in Development and Evolution. Edited by B. Hallgrímsson and B.K. Hall. University of California Press, Berkeley. In Press.

6.1 INTRODUCTION: WHAT IS GENOMIC IMPRINTING?

Genomic imprinting is an epigenetic process in which an allele is marked according to the sex of the parent transmitting it. These sex-specific marks may affect single genes, gene clusters, or entire chromosomes, and result in maternal and paternal alleles or chromosomes that are epigenetically distinct from one another. This difference in epigenetic status can lead to differential transcriptional activity, chromosome loss, or chromosome inactivation. In an organism, allelic differences that result due to genomic imprinting can be observed as the exclusive or preferential expression of a gene when it is inherited from one parent, but not the other. Thus, in contrast to classic mechanisms of gene expression and regulation, in genomic imprinting it is the allele's parent-of-origin, and not the underlying DNA sequence, that determines its activity. In this chapter, we therefore use the term epigenetic in the limited but specific sense employed in molecular biology to mean processes that affect gene expression, without changing DNA sequence.

Genomic imprinting has been most extensively studied in mammals (Morison *et al.*, 2005; Wood and Oakey, 2006), but has been observed in a wide range of organisms, including plants (Alleman and Doctor, 2000; Scott and Spielman, 2006), insects (Khosla *et al.*, 2006; Lloyd, 2000), *C. elegans* (Bean *et al.*, 2004), and zebrafish (Martin and McGowan, 1995). Many similarities exist in the epigenetic mechanisms that underlie genomic imprinting in these species. These mechanisms include, but are not limited to, DNA methylation, histone modifications, changes in higher order chromatin structure, non-coding RNA, and RNA interference (RNAi). Accumulating evidence suggests that these epigenetic mechanisms are frequently interrelated and mutually reinforcing.

DNA methylation is an epigenetic process in which methyl groups are added to nucleotides, often cytosines present in CpG dinucleotides, without affecting the underlying DNA sequence. When DNA methylation encompasses the promoter of a gene, it frequently results in transcriptional repression. However, methylation-sensitive enhancers, repressors, and protein binding sequences, are also common and important in mediating epigenetic gene expression.

In the cell, DNA is wrapped around nucleosomes, a protein structure that consists of two copies of four different histone proteins (H2A, H2B, H3 and H4). Chemical modifications including methylation, acetylation, and phosphorylation, of amino acids in the histone sequences can contribute to the formation of inactive or active chromatin structures (Figure 6.1A). Evidence suggests that DNA methylation and histone modifications are intimately linked and exhibit extensive epigenetic "cross talk", with information flowing from DNA to histones and from histones to DNA. Given that DNA methylation can guide histone modifications, and histone modifications can influence DNA methylation, it is likely that these processes function in a mutually reinforcing epigenetic loop that ensures maintenance of a repressive chromatin state (Fuks, 2005; Vaissiere *et al.*, 2008).

Epigenetic gene regulation by non-coding RNAs can involve RNAi-mediated pathways, in which the non-coding RNAs are processed into small RNAs, but can also be RNAi-independent. RNAi-mediated epigenetic gene regulation may occur at the post-transcriptional level, with the small RNAs guiding degradation of an mRNA transcript or inhibiting its translation, or it may occur at the transcriptional level, with the small RNAs mediating chromatin modifications that inhibit transcription. The exact function of many non-coding RNAs remains elusive, but a variety of evidence connects non-coding RNAs

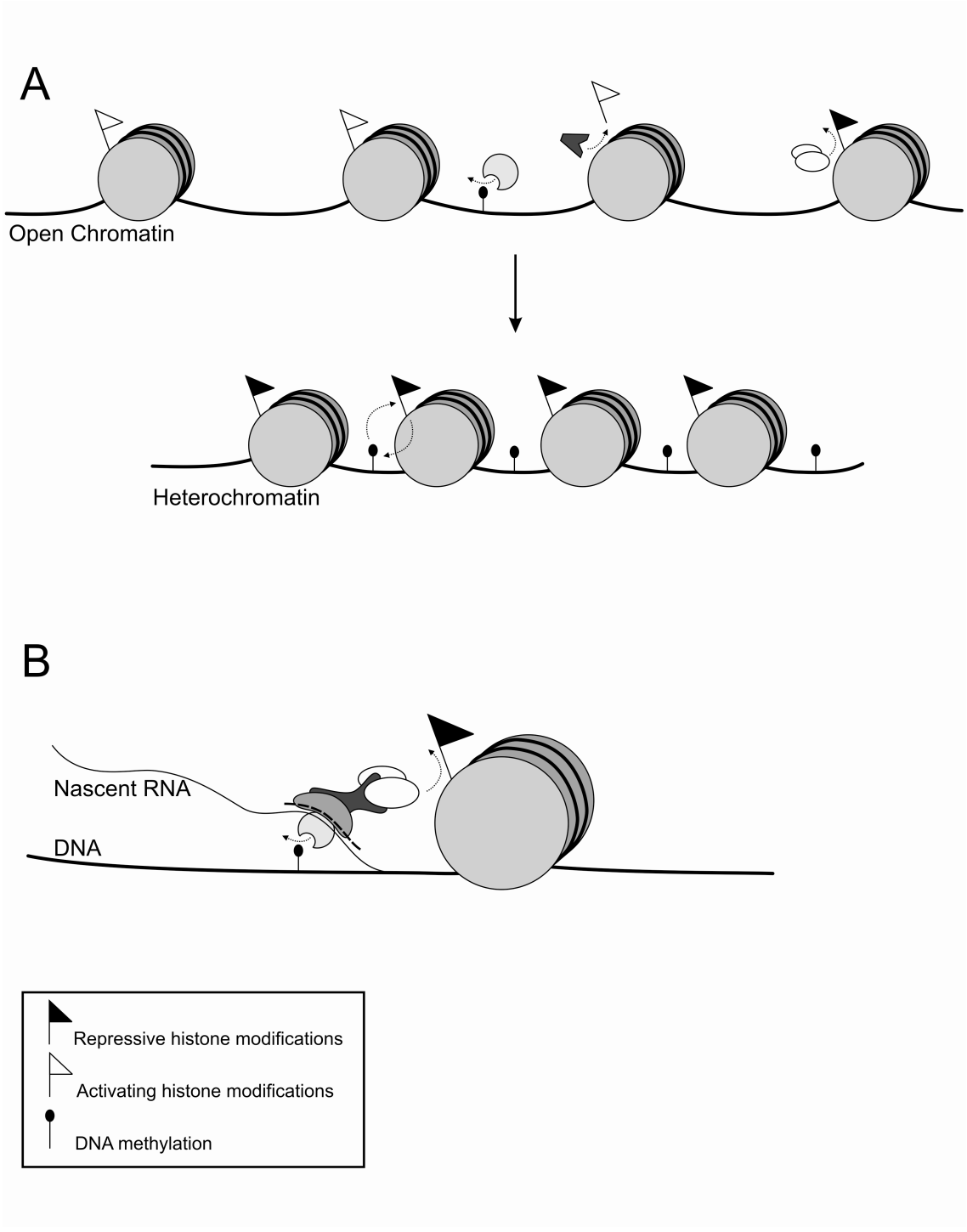
with other epigenetic mechanisms, including histone modifications, DNA methylation, and heterochromatin formation (Bernstein and Allis, 2005; Matzke and Birchler, 2005; Zaratiegui *et al.*, 2007). Thus, it is likely that many non-coding RNA transcripts are important in mediating higher-order chromatin structure. For example, in mammalian X chromosome inactivation, a 17 kb non-coding RNA called *Xist* is expressed from the X chromosome that will be inactivated, and subsequently coats that chromosome. This initiates a variety of chromatin remodelling events, including histone modifications and the incorporation of a specialized histone variant, which ensure silencing of the inactive X (Bernstein and Allis, 2005).

Recent work in yeast has revealed further details of the mechanism by which RNAi can direct heterochromatin formation. This analysis demonstrated that transcripts from heterochromatic regions of the genome accumulate during S-phase of the cell cycle, and are processed into short interfering RNAs (siRNAs), which then recruit histone methylation that contributes to heterochromatin formation (Kloc *et al.*, 2008). RNAi-mediated heterochromatin formation has also been reported in plants and animals, and *de novo* DNA methylation in plants is RNA-directed (Lippman and Martienssen, 2004; Matzke and Birchler, 2005; Zaratiegui *et al.*, 2007). How small or non-coding RNAs direct DNA modifications is not yet fully understood, but recent evidence in yeast favours a model in which small RNA interaction with nascent RNA transcripts recruits a complex of chromatin and DNA binding and modifying proteins (Figure 6.1B; Buhler *et al.*, 2006; Irvine *et al.*, 2006). The discovery that DNA methylation of a group of genes in *Arabidopsis* is directed by a small RNA that targets an exon-exon junction, also supports an RNA-RNA interaction model (Bao *et al.*, 2004). Alternatively, small or non-coding

RNAs may direct these modifications via base-pairing interactions with genomic DNA (Grewal and Moazed, 2003; Mayer *et al.*, 2006).

The observation of imprinting in such an extensive range of animal and plant species, combined with the utilization of many of the same mechanisms to establish and maintain imprinted expression, suggests that genomic imprinting is a widespread occurrence based on phylogenetically conserved core epigenetic processes that can be adapted to serve different functions in different species. The conservation of these epigenetic processes is further emphasized by transgenic experiments in which an epigenetic control region from one organism successfully functions in another. In this chapter we will examine examples of genomic imprinting from mammals, insects, and plants, with a focus on imprinting mechanisms, and the conservation of core epigenetic processes.

Figure 6.1 Genomic imprinting utilizes several interrelated and conserved epigenetic mechanisms. **[A]** Epigenetic gene regulation is often mediated by histone and DNA modifications that contribute to higher order chromatin structure. In the cell, DNA is wrapped around nucleosomes (grey cylinders), which each contain eight histone proteins. Histones can acquire activating or repressive modifications. Activating histone modifications lead to an open chromatin structure that promotes gene transcription. Removal of activating histone modifications, and the addition of repressive histone modifications and DNA methylation, leads to a condensed heterochromatic structure that hinders transcription. DNA methylation and histone modifications are often mutually reinforcing, with each modification influencing and contributing to the other. **[B]** Non-coding and small RNAs can regulate gene expression and contribute to the formation of a compacted heterochromatin structure by directing histone modifications and DNA methylation. Illustrated is a possible model for small RNA mediated chromatin modifications. Small RNAs (dashed line) interact with a nascent transcript at the target gene locus according to sequence homology, and tether chromatin modifying enzymes to the locus. These may include DNA methyltransferases (light grey pacman) or histone modifying enzymes (white ellipses). These RNA-directed DNA and histone modifications may then alter the chromatin structure of the locus.


A
Open Chromatin
Heterochromatin
B
Nascent RNA
DNA
Repressive histone modifications
Activating histone modifications
DNA methylation

6.2 THE EVOLUTION OF GENOMIC IMPRINTING

Genomic imprinting has been most extensively studied in mammals, and thus the majority of hypotheses of the selective forces leading to imprinted gene expression are based on mammalian imprinted genes. The **parental conflict hypothesis** is the most thoroughly debated, and is based on a reproductive mode involving multiple paternity within a litter of mammals (Moore and Haig, 1991). In this reproductive scenario, it is beneficial to the mother to distribute nutrients evenly to her offspring, as all share her genes, while it is in the father's genetic interests for only his offspring to receive maximal resources. This hypothesis predicts that paternal imprinting should enhance the expression of fetal growth promoters, while maternal imprinting should have the opposite effect. While the early identification of several imprinted genes involved in fetal growth resulted in great enthusiasm for this hypothesis, the discovery of imprinted genes with a variety of functions makes it increasingly unlikely that the parental conflict hypothesis can account for the imprinting of all genes.

The **ovarian time bomb hypothesis**, based on the supposition that genomic imprinting was selected to prevent ovarian trophoblastic disease (Varmuza and Mann, 1994), and the role of imprinting in preventing parthenogenesis, have also been debated. In mice, imprinting is a major barrier to parthenogenesis (Kono, 2006), and uniparental inheritance of chromosomal regions containing imprinted genes can lead to embryonic lethality, and postnatal growth and developmental defects (Cattanach and Kirk, 1985). However, at least six imprinted genes have been identified in chromosomal regions that have no obvious phenotype when inherited uniparentally (Peters and Beechey, 2004), and additional imprinted genes have been shown to have only behavioural or cognitive effects

(Davies *et al.*, 2005; Plagge *et al.*, 2005). Thus, it is likely that imprinting affects a wide range of genes in the mouse and other mammals, not only those involved in growth and development. Prevention of parthenogenesis or ovarian trophoblastic disease is therefore unlikely to be the selective force behind the imprinting of all mammalian genes.

Additional hypotheses relate to the benefit of imprinting in establishing appropriate gene dosages or functional haploidy (Holliday, 1990; Ohlsson *et al.*, 2001; Okamura and Ito, 2006), enhancing the "evolvability" of a population (Beaudet and Jiang, 2002), and modifying expression of genes that have different optimal expression levels and selective pressures in males and females (Day and Bonduriansky, 2004; Iwasa and Pomiankowski, 1999). Overall, as each new imprinted gene is discovered, and as imprinting is discovered in animals and plants with reproductive strategies differing from those in mammals, it seems increasingly unlikely that a single evolutionary hypothesis will explain the occurrence of all imprinted genes, even within a single species. It is more likely that a variety of selective forces have contributed to the evolution of imprinted gene expression.

6.3 GENOMIC IMPRINTING IN MAMMALS

Genomic imprinting has been observed in the eutherian (Khatib *et al.*, 2007) and marsupial (Suzuki *et al.*, 2005) lineages of mammals, and is most frequently studied in mice and humans. Imprinting in humans has been of particular interest to the medical and research communities due to the association of imprinted genes and aberrant imprinting with a variety of diseases and cancers. In both mice and humans, many of the identified imprinted genes occur in clusters that contain shared regulatory regions and/or transcripts

that control the imprinted expression of multiple genes in the cluster. Imprinted genes that reside within introns, or have originated from retrotransposition events, are also common (Morison *et al.*, 2005). Many of the identified mammalian imprinted genes are not imprinted in all tissues at all times, indicating that tissue- and temporal-specific imprinting may be frequent, and may complicate the identification of imprinted genes with unique patterns of expression.

At the molecular level, the protein-coding imprinted genes identified in mice and humans participate in a wide range of cellular processes, with no obvious function or theme in common (Morison *et al.*, 2005; Peters and Beechey, 2004). In addition, approximately 30% of imprinted genes correspond to non-coding RNA transcripts, many of which are involved in regulating the imprinted expression of other genes. The imprinted genes influenced by the transcript may overlap, be located nearby, or at a distance from the transcript, and may be imprinted in the same, or opposite direction. Other imprinted RNAs encode small nucleolar RNAs (snoRNAs) and microRNAs, and still others have no known function.

While there is much overlap between the genes that are imprinted in mice and humans, there are also significant differences. Several genes are reported to be imprinted in mice but not in humans, or *vice versa*, and at least two genes are reported to be oppositely imprinted in the two species (*COPG2* and *ZIM2*). Additional genes are imprinted in one species, but lack an orthologue in the other (Morison *et al.*, 2005). These differences may suggest that for many genes, the loss or gain of imprinting during a species' evolution may occur somewhat easily and without drastic effects. Thus, while the epigenetic processes underlying imprinting may be conserved, there likely exists many species-specific differences in the genes that are affected.

6.3.1 IMPRINTING AT THE *H19/IGF2* LOCUS

H19 and *Insulin-like growth factor 2* (*Igf2*) are perhaps the most extensively studied and best characterized imprinted genes. Located approximately 90 kb apart, these two genes are reciprocally imprinted, with the non-coding *H19* transcript expressed only from the maternal allele, and *Igf2* expressed only from the paternal allele (Bartolomei *et al.*, 1991; DeChiara *et al.*, 1991). Imprinted expression of *H19* and *Igf2* is controlled by a shared imprint control region (ICR) located approximately 2 kb upstream of the *H19* transcription start site. Deletion of this ICR results in a loss of imprinting of both genes (Thorvaldsen *et al.*, 1998). In addition to the ICR, expression of both *Igf2* and *H19* requires several tissue specific enhancers spread over at least three regions 10-120 kb downstream of the *H19* gene (Ainscough *et al.*, 2000a; Davies *et al.*, 2002; Kaffer *et al.*, 2000; Leighton *et al.*, 1995).

6.3.1.1 DNA METHYLATION

The *H19*/*Igf2* ICR contains multiple binding sites for the enhancer-blocking, insulator protein CTCF, which can only bind when these sites are unmethylated (Bell and Felsenfeld, 2000; Hark *et al.*, 2000). Methylation of the ICR is present in sperm but not ova (Tremblay *et al.*, 1995), enabling CTCF to bind to the maternally inherited ICR, but not the paternally inherited ICR. CTCF binding to the unmethylated maternal ICR prevents the downstream enhancers from activating *Igf2*, and instead the enhancers stimulate transcription of *H19*. Conversely, CTCF cannot bind to the methylated ICR on the paternal allele, and the downstream enhancers activate expression of *Igf2* on the paternally inherited chromosome (Bell and Felsenfeld, 2000; Hark *et al.*, 2000).

Once imprinted expression is established, the ICR is required to maintain *Igf2* silencing on the maternal chromosome, but not *H19* silencing on the paternally inherited chromosome (Srivastava *et al.*, 2000). The *H19* promoter acquires methylation on the paternal allele during embryogenesis (Bartolomei *et al.*, 1993; Tremblay *et al.*, 1997; Tremblay *et al.*, 1995), which is likely sufficient to maintain its silenced state. Mutation or deletion of the CTCF binding sites does not affect differential methylation of the ICR in sperm and ova, and thus the methylated paternal allele imprints appropriately in the absence of CTCF binding sites. However, in the absence of CTCF binding sites on the maternal allele, the *H19* promoter and gene region acquire methylation post implantation, *H19* expression is reduced, and *Igf2* is expressed biallelically (transcribed from both parental alleles) (Engel *et al.*, 2006; Szabo *et al.*, 2004).

6.3.1.2 CONSERVED EPIGENETIC MECHANISMS

The epigenetic processes that cause *H19/Igf2* imprinting appear to be conserved between mammals and *Drosophila*, as demonstrated by transgenic experiments. A silencer element within the mouse *H19* ICR was originally discovered using transgenic *Drosophila* containing the mouse *H19* upstream region adjacent to *lacZ* and mini-*white* reporter genes (Lyko *et al.*, 1997). Deletion constructs delineated the silencing element to a 1.2 kb region (Lyko *et al.*, 1997), and subsequent experiments showed that this 1.2 kb element also functions specifically in *H19* silencing, and not imprinting, at endogenous mouse locus (Drewell *et al.*, 2000). Targeted deletion of the silencer resulted in a loss of *H19* silencing following paternal transmission, while paternal *Igf2* expression, differential methylation, and expression of both genes following maternal transmission were unaffected. A similar 1.5 kb silencer element appears to exist at the 3' end of the human

H19 ICR (Arney *et al.*, 2006). This region silenced a mini-*white* reporter gene in transgenic *Drosophila*, and functioned as a silencer in transient transfection assays using a human embryonic kidney cell line (Arney *et al.*, 2006), while additional regions from the human ICR did not.

Additional insight into the complexity of the *H19* ICR and imprinting mechanism, stems from recent evidence demonstrating that the mouse ICR is biallelically transcribed and produces both sense and anti-sense RNA (Schoenfelder *et al.*, 2007). Biallelic transcription was also detected in the transgenic *Drosophila* system, where further analysis indicated that the *H19* ICR transcripts induce gene silencing in an RNAi-independent manner (Schoenfelder *et al.*, 2007). In transgenic *Drosophila*, mutations in RNAi genes failed to relieve reporter gene silencing, and no siRNAs were detected from the *H19* ICR. Furthermore, artificially producing *H19* ICR siRNAs resulted in a significant reduction of *H19* ICR transcripts, which was accompanied by a more than 5-fold increase in mini-*white* expression. In their endogenous context, these ICR transcripts may be involved in forming a repressive chromatin structure and mediating *H19* repression on the paternal allele. This would be similar to the model for imprinting at the mammalian *Cdkn1c-Kcnq1* imprinted domain, where a non-coding RNA transcript is believed to mediate a repressive chromatin structure on the paternal allele (Umlauf *et al.*, 2004). Non-coding RNA transcripts are similarly required to establish and maintain a heterochromatic structure at a ribosomal RNA gene cluster in mice (Mayer *et al.*, 2006), and evidence suggests that the higher-order chromatin structure of mouse pericentric heterochromatin involves an RNA component (Maison *et al.*, 2002). On the maternal allele, CTCF binding may prevent the repressive effect of the *H19* ICR transcripts, or alternatively, the ICR transcripts could serve a different functional role that is undetected

in the transgenic *Drosophila* system, which does not imprint and acts most similarly to the silenced paternal allele.

The central A6-A4 region — also termed the centrally conserved domain — is located between the two genes, and is unmethylated, DNAseI hypersensitive, and GC rich (Koide *et al.*, 1994). Within this region, two sub-regions show a high level of homology between humans and mice. Region 1 is necessary for maintaining repression of *Igf2* from the maternal allele in skeletal muscle (Ainscough *et al.*, 2000b), while Region 2 appears to be an enhancer for *Igf2* expression in the choroid plexus, where it is normally expressed biallelically (Jones *et al.*, 2001). Analysis of transgenic *Drosophila* containing the central A6-A4 region (Erhardt *et al.*, 2003) may provide additional insight into the mechanism of repression within this region. Transgenic flies containing this region adjacent to mini-*white* and *lacZ* show overall silencing of both reporter genes, as well as eye pigment variegation in some lines, indicating the formation of compact chromatin domains (Erhardt *et al.*, 2003). Silencing increased in *Enhancer of Zeste (E(z))* mutant flies, and decreased in *Posterior Sex Combs (Psc)* mutant flies, which were also observed to bind to the transgene integration site (Erhardt *et al.*, 2003). Both E(z) and Psc are highly conserved proteins involved in chromatin remodelling and maintaining silenced and/or active gene states (LaJeunesse and Shearn, 1996), suggesting that similar genes or the mammalian gene homologues may be involved in modulating the chromatin structure at the A6-A4 region in mice.

6.3.1.3 ADDITIONAL REGULATORY REGIONS

In addition to the ICR and downstream enhancers that are essential for normal imprinted expression of both genes, several additional sequences are required for

appropriate tissue-specific expression and repression of *H19* and *Igf2* (Figure 6.2A). The central A6-A4 region has been discussed above. Two conserved sequences upstream of *H19* and the *ICR*, termed HUC1 and HUC2, are biallelically transcribed in both mice and humans, and appear to be mesoderm-specific enhancers (Drewell *et al.*, 2002). Additional differentially methylated regions (DMRs) surround the *Igf2* gene and affect its expression. In mice, this region contains three DMRs: DMR0, DMR1 and DMR2. DMR0 encompasses the promoter region of a placental-specific transcript (P0) and is maternally hypermethylated in the placenta, but biallelically methylated in the fetus (Moore *et al.*, 1997). Both DMR1 and DMR2 are methylated on the active paternal allele, but function oppositely. DMR1 is a mesodermal repressor located upstream of the *Igf2* gene. Deletion of DMR1 results in biallelic expression of *Igf2* in several mesoderm derived tissues (Constancia *et al.*, 2000). Conversely, DMR2 is an *Igf2* enhancer located in the sixth exon, and while deletion does not affect imprinting, it results in reduced *Igf2* expression from the paternal allele (Murrell *et al.*, 2001). *In vitro* experiments confirm that the methylation status of these two DMRs is important for their function, and is conducive to paternal *Igf2* expression. Methylation of DMR1 causes a loss of reporter gene silencing (Eden *et al.*, 2001), while *in vitro* methylation of DMR2 increases reporter gene expression (Murrell *et al.*, 2001). The *Igf2* region of humans only contains DMR0 and DMR2; however, recent analysis indicates that DMR0 is methylated on the active paternal allele in all tissues and may function similarly to mouse DMR1 (Murrell *et al.*, 2008).

6.3.1.4 CHROMATIN LOOPING

Further analysis in mice demonstrated that physical long-range interactions between the ICR and *Igf2* DMRs likely establish parent-of-origin specific chromatin loops. On the maternal chromosome, the ICR physically interacts with both DMR1 and a matrix attachment region located 3' to the *Igf2* gene termed MAR3, while on the paternal chromosome the ICR interacts with DMR2 (Kurukuti *et al.*, 2006; Murrell *et al.*, 2004). CTCF binding to the ICR is necessary to mediate the higher order chromatin structure on the maternal allele, and mutation of the CTCF binding sites abolishes these physical interactions and causes the region to adopt the paternal chromatin structure (Kurukuti *et al.*, 2006). Elimination of CTCF binding to the maternal ICR also causes a loss of CTCF binding within DMR1, as well as *de novo* methylation of the DMR1 and DMR2 regions. It thus appears that CTCF is recruited to DMR1 through the physical interaction with the ICR, and that this recruitment protects the region from the acquisition of methylation on the maternal allele (Kurukuti *et al.*, 2006).

Differential interactions between the gene promoters and the shared downstream enhancers have also been detected (Yoon *et al.*, 2007). On the paternally inherited chromosome from which *Igf2* is normally expressed, physical interactions between the *Igf2* promoters and the downstream enhancers are observed. Conversely, on the maternally silenced chromosome, the *Igf2* promoters physically interact with the ICR, and ICR-enhancer interactions are also detected (Yoon *et al.*, 2007). The maternal chromosome also exhibits physical interactions between the *H19* promoter and the downstream enhancers, consistent with its expression of this gene (Yoon *et al.*, 2007).

6.3.1.5 HISTONE MODIFICATIONS

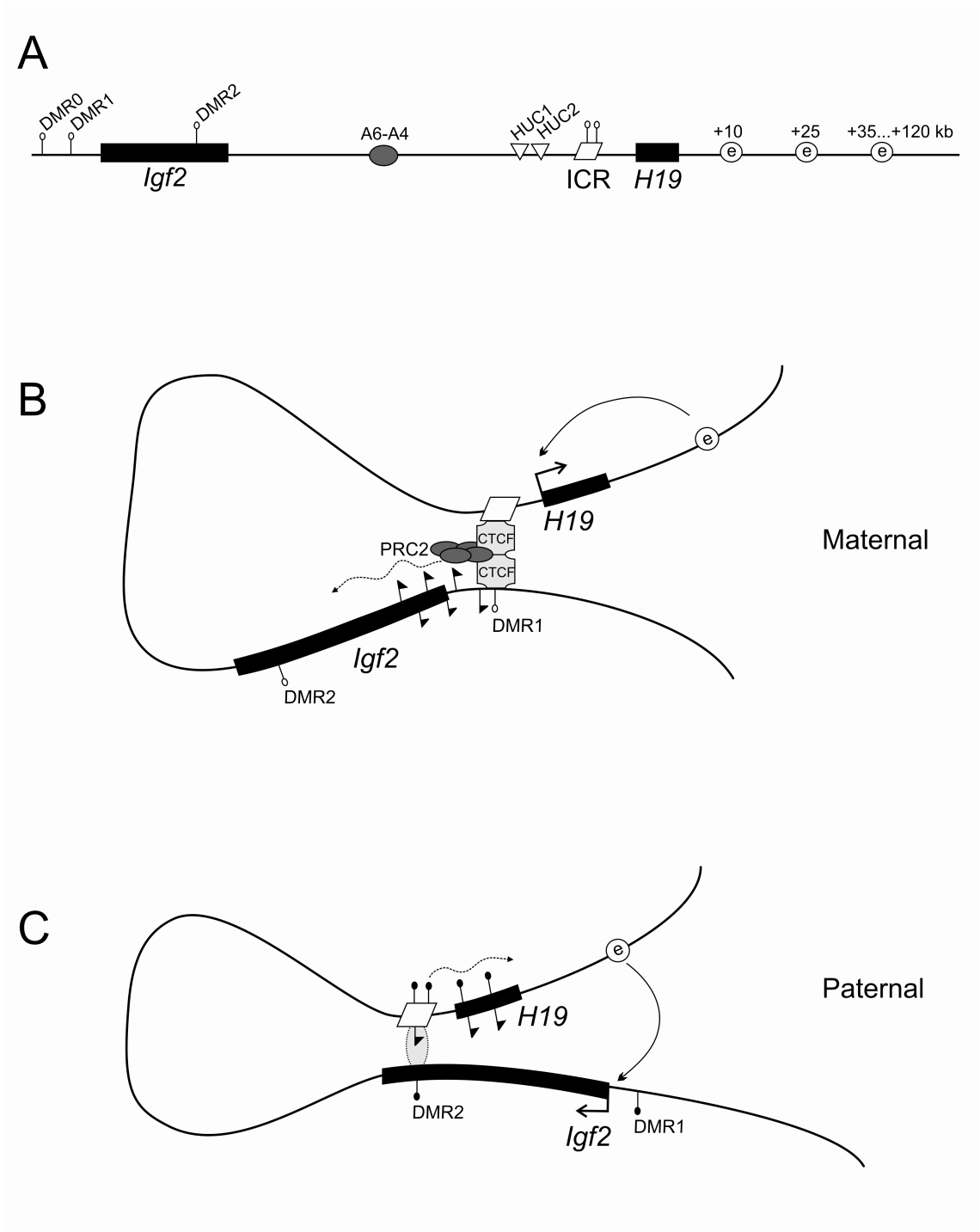
Using chromatin immunoprecipitation, CTCF was found to bind to the *Igf2* promoter region on the maternal chromosome, beginning at DMR1 and exhibiting the strongest binding at the two major promoters, P2 and P3. This region was also found to be hypermethylated at Lysine 27 of Histone H3 (H3K27), a modification mediated by the Polycomb repressive complex 2 (PRC2) in mammals. Consistent with this, CTCF was found to directly interact with Suz12, an essential component of PRC2. RNAi knockdown of Suz12 resulted in hypomethylation of H3K27 at the *Igf2* promoters of the maternal allele, and biallelic *Igf2* expression (Li *et al.*, 2008). CTCF binding to the unmethylated maternal ICR therefore mediates long range interactions with the *Igf2* promoter, and then subsequently recruits the PRC2 complex, which results in histone methylation and repression of *Igf2* on the maternal allele only.

Differential histone modifications have also recently been detected throughout the rest of the *Igf2/H19* imprinted domain (Han *et al.*, 2008; Verona *et al.*, 2008). In addition to H3K27 methylation, the silenced maternal *Igf2* region is also enriched for repressive methylation at Lysine 9 of Histone H3 (H3K9), and the heterochromatic histone variant macroH2A1 (Han *et al.*, 2008). Activating histone marks, including histone acetylation and Histone H3 Lysine 4 (H3K4) methylation, are predominant on the maternal chromosome at the ICR and *H19* promoter/gene region, and on the paternal chromosome at the *Igf2* promoter/gene region (Han *et al.*, 2008). Both ICRs contain H3K27 methylation, and the paternal ICR and *H19* gene are also strongly enriched for H3K9 methylation and macroH2A1 (Han *et al.*, 2008; Verona *et al.*, 2008). H3K27 methylation at the *H19* promoter and gene is unclear, and is either enriched on the paternal allele (Han *et al.*, 2008), or equivalent on the maternal and paternal alleles (Verona *et al.*, 2008).

Abolishing CTCF binding to the ICR caused the maternal chromosome to adopt the normal paternal histone composition throughout both *Igf2* and *H19*, suggesting that CTCF is essential for organizing the maternal chromatin structure (Han *et al.*, 2008). Transcription of *H19* from the maternal allele also appears to be required for establishing H3K4 methylation and histone acetylation throughout this region, as the maternal allele will lose active chromatin modifications in cells with an ICR deletion where *H19* is not expressed, but not in cells with the same ICR deletion where *H19* is expressed (Verona *et al.*, 2008).

Although much of *H19/Igf2* imprinting research has focused on the role of differential DNA methylation in establishing and maintaining imprinted expression of these two genes, it is now clear that *H19/Igf2* imprinting is much more complex. Given that DNA methylation, histone modifications, higher-order chromatin structure, and RNA-directed modifications are often mutually reinforcing epigenetic processes, it is not surprising that recent results have indicated that all of these processes are essential in *H19/Igf2* imprinting. A model for *Igf2* and *H19* imprinting is illustrated in Figure 6.2.

Figure 6.2 A model for imprinting of the mammalian *H19* and *Igf2* genes. **[A]** Relative positions of the imprinted *Igf2* and *H19* genes, the imprint control region (ICR), downstream enhancers, and additional regulatory regions. The shared ICR is located upstream of *H19*. The positions of the downstream enhancers required for imprinted expression of both genes are given relative to the *H19* transcriptional start site. Sites of differential methylation are indicated with lollipops. **[B]** On the maternally inherited allele, CTCF binds to the unmethylated ICR and orchestrates chromatin looping and CTCF binding to DMR1 of *Igf2*, potentially through CTCF homodimerization. CTCF recruits PRC2 to the *Igf2* promoters via its interaction with Suz12. The PRC2 complex mediates H3K27 methylation throughout DMR1 and the *Igf2* promoters and gene region (indicated with filled flags and dotted arrow). This repressive histone modification leads to maternal *Igf2* silencing. The downstream enhancers interact with the *H19* promoters, leading to maternal-specific *H19* expression. **[C]** On the paternal chromosome, the ICR is methylated in sperm (filled lollipops), which prevents binding of the insulator protein CTCF. During embryogenesis, this methylation spreads to encompass the *H19* promoter and gene region (dotted arrow). The ICR, *H19* promoter, and *H19* gene region also contain repressive histone modifications that likely contribute to repression (filled flags). DMR1 and DMR2 are methylated (filled lollipops), and the ICR can interact with DMR2 (dotted ellipse), which may contribute to the formation of a chromatin loop that facilitates enhancer access to the *Igf2* promoters, resulting in paternal-specific *Igf2* expression.


A
DMR0
DMR1
DMR2
A6-A4
HUC1
HUC2
+10
+25
+35...+120 kb
Igf2
ICR
H19
B
H19
CTCF
CTCF
PRC2
DMR1
Igf2
DMR2
Maternal
C
H19
DMR2
Igf2
DMR1
Paternal

6.4 GENOMIC IMPRINTING IN INSECTS

Although much of the recent research in the field of genomic imprinting has focused on mammals, imprinting was first described in two insect systems: *Sciara* (black fungus gnats), and coccids (scale insects). The process of imprinting in these insects, as well as in the model organism *Drosophila melanogaster*, is the same as that of mammals. The imprint is differentially established depending on the sex of the germ line, maintained throughout embryonic development, and then erased in gametogenesis so that the adult organism properly transmits the appropriate imprint to his/her progeny. However, while most documented examples of imprinting in mammals affect individual genes or several genes grouped in a large cluster, imprinting in *Sciara* and coccids results in a parent-of-origin effect on whole chromosomes. This can lead to elimination or heterochromatinization of chromosomes based strictly on whether they were inherited through the male or female germ line.

Imprinting in *Drosophila* has been observed for marker genes on rearranged chromosomes, as well as for transgenes inserted at heterochromatic positions (Lloyd, 2000; Maggert and Golic, 2002). Paternal-specific chromosomal loss has also been observed in *Drosophila* in the presence of certain mutations, indicating that the chromosomes must carry a parent-of-origin specific imprint, despite this imprint not normally causing an obvious consequence on chromosomal behaviour. Together this evidence indicates that *Drosophila* is fully capable of imprinting, and can generate both smaller imprinted domains, similar to mammalian imprinting centres, that may result in imprinted expression of non-developmentally essential genes, and imprints that can cause the loss of whole chromosomes, similar to those observed in other insects.

6.4.1 *SCIARA*

The term "imprint" was first used in describing the complex process of chromosome elimination in *Sciara* (Crouse, 1960). Following the observation that the developing *Sciara* embryo specifically eliminated chromosomes of paternal origin, it was hypothesized that the transmission of chromosomes through the male and female germ lines resulted in an "imprint" that marked the chromosomes based on their inheritance. This imprint was concluded to be unrelated to the genetic content of the chromosome, and therefore solely determined by sex of the parent who had transmitted it.

The complex process of chromosome elimination in *Sciara* occurs in three distinct elimination events (reviewed in Goday and Esteban, 2001). *Sciara* embryos inherit three X chromosomes: one maternally, and two paternally. During early embryonic development, either one or both of the paternal X chromosomes are eliminated from somatic cells, depending on whether the sex of the embryo is female or male, respectively. The germ line retains all three X chromosomes until later in embryonic development, when a single paternal X chromosome is eliminated from the germ nuclei of both males and females by expulsion through the nuclear membrane and into the cytoplasm where it is degraded. Female meiosis then proceeds normally. However, during male meiosis a third elimination event occurs. This elimination discards all remaining paternal chromosomes, including the autosomes, into a cytoplasmic bud that is extruded from the developing sperm nuclei. Non-disjunction of the maternal X chromosome during meiosis II results in the inclusion of both maternal X chromatids into a single mature sperm cell. Male *Sciara* therefore produce sperm containing only the chromosomes that they inherited maternally (Figure 6.3). It is important to note that both males and females must re-imprint their chromosomes at some point during germ cell

development so that their chromosomes are properly recognized as either “maternal” or “paternal” in the next generation.

Early cytological analysis provided evidence that maternal and paternal chromosomes exhibit distinct characteristics in the early germ line. Following elimination of the paternal X from the early germ nuclei, the paternal chromosomes appear unravelled and light-stained, while the maternal chromosomes appear condensed and dark stained (Berry, 1941; Rieffel and Crouse, 1966). This difference is evident until just prior to gonial mitosis, at which point the two chromosome sets appear equally condensed. Maternal and paternal chromosomes have also been observed to occupy distinct nuclear compartments in developing germ cells until meiosis, resulting in physical separation of the two chromosome sets within the nucleus (Kubai, 1987). The unique behaviours of the X chromosome, including both elimination and non-disjunction, requires a distinct controlling element that maps to a block of X chromosome heterochromatin (Crouse, 1960) that also contains ribosomal RNA genes (Crouse, 1979; Crouse *et al.*, 1977).

6.4.1.1 HISTONE MODIFICATIONS

Molecular analysis of maternal and paternal chromatin modifications during germ cell development discovered several differences in histone acetylation (Goday and Ruiz, 2002) and histone methylation (Greciano and Goday, 2006) during the chromosome elimination process. In early germ cells, prior to elimination, the paternal chromosomes are highly acetylated on histones H3 and H4, with the exception of the paternal X that is eliminated. The paternal X that will be eliminated, and the entire maternal chromosome complement, are hypoacetylated (Goday and Ruiz, 2002). In addition, the maternal

chromosomes exhibit H3K4 methylation, while the H3 histones of the paternal chromosomes are unmethylated (Greciano and Goday, 2006). The paternal X chromosome that is eliminated is therefore the only chromosome that is both unmethylated and hypoacetylated, which may be the distinguishing factor that identifies it for elimination. Hypoacetylation in particular, is hypothesized to be required for its interaction with inner nuclear membrane proteins, and subsequent elimination from the nucleus (Goday and Ruiz, 2002).

Histone acetylation differences are maintained during the X chromosome elimination, and subsequent decondensation of the remaining paternal chromosomes. Post-decondensation, the maternal chromosomes acquire histone acetylation, rendering both chromosome sets equally acetylated (Goday and Ruiz, 2002). At this stage, the H3 histones of the two chromosome sets are also equivalently methylated (Greciano and Goday, 2006). However, during male meiosis, the paternal chromosome set becomes hypoacetylated and hypermethylated. Interestingly, both the acetylation and methylation differences are reversed from those observed in early germ nuclei, where the maternal chromosomes were hypoacetylated and hypermethylated. In addition to methylation at H3K4, the paternal chromosomes also acquire methylation at Histone H4 Lysine 20 (H4K20) during male meiosis (Greciano and Goday, 2006).

Localization of the maternal and paternal chromosome sets to distinct nuclear compartments is likely an essential component of the chromosome modification and elimination processes. These nuclear compartments may be associated with the activity of specific histone acetyltransferases, deacetylases, methyltransferases, or demethylases, resulting in histone modifications for one parental chromosome set, but not the other. These modifications may then, in turn, participate in localizing or identifying the

chromosomes for elimination or retention. The roles of other core epigenetic processes, such as RNA-mediated modifications and DNA methylation, have not yet been investigated. The current model of *Sciara* chromosome imprinting, which incorporates histone modifications, chromosome distribution, and elimination, is diagrammed in Figure 6.3.

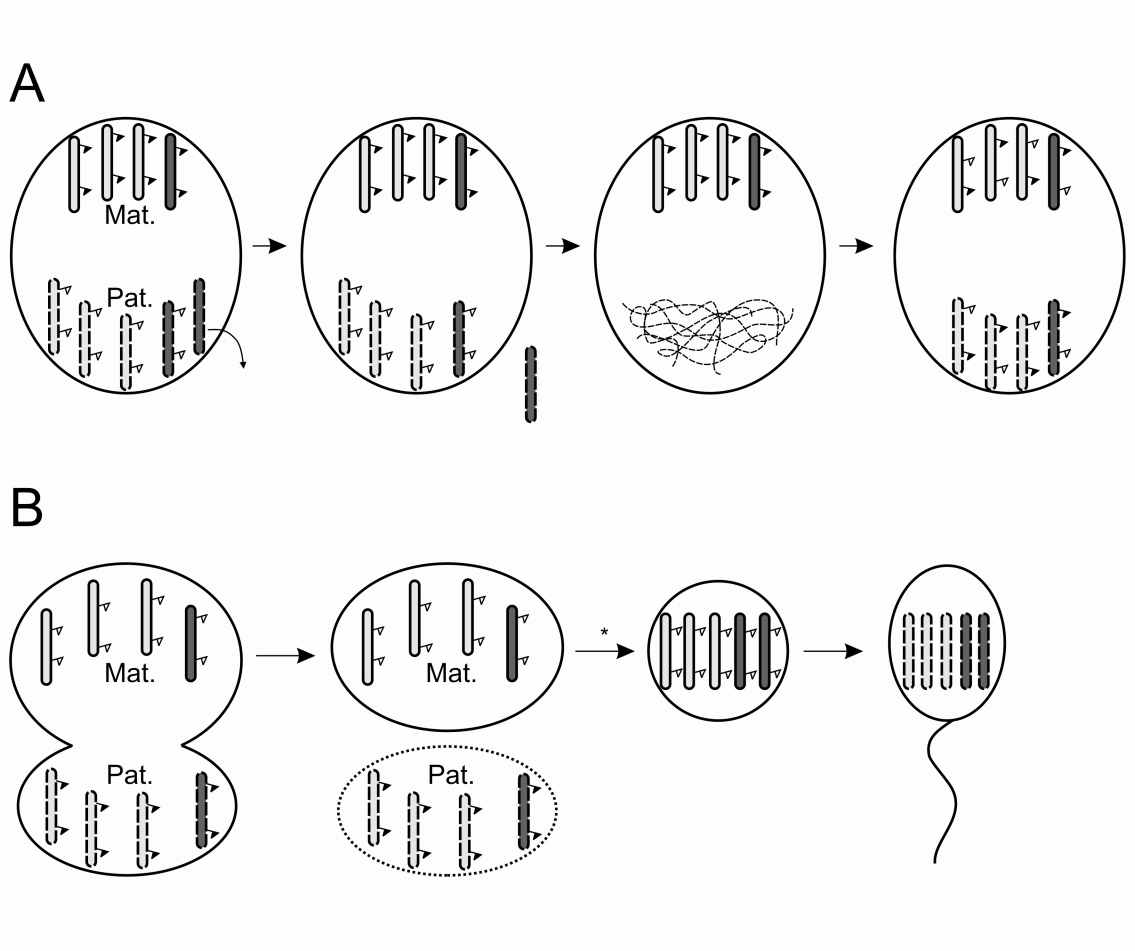


Figure 6.3 Genomic imprinting and chromosome elimination in the *Sciara* germ line (adapted from Greciano and Goday, 2006). Maternally inherited chromosomes are outlined with a solid line, and paternally inherited chromosomes are outlined with a dashed line. X-chromosomes are filled with dark grey, and autosomal chromosomes with light grey. **[A]** In the early germ nucleus, paternal and maternal chromosomes are separated into distinct nuclear locations. Maternally inherited chromosomes exhibit histone hypermethylation (filled flags), and paternally inherited chromosomes exhibit histone hyperacetylation (white flags), with the exception of one paternally inherited X chromosome, which is neither hyperacetylated nor hypermethylated. This X chromosome is eliminated from the early germ nucleus of both sexes. Following this elimination, the paternal chromosomes undergo decondensation. This decondensation is maintained until just prior to gonial mitosis, at which point all chromosomes are equally condensed, methylated, and acetylated. Female meiosis then proceeds normally. **[B]** During male meiosis, the maternally inherited chromosomes are hyperacetylated, and the paternally inherited chromosomes are hypermethylated. The paternally inherited chromosomes are eliminated during meiosis I. During meiosis II (*), non-disjunction of the maternal X chromosome occurs. The mature sperm nucleus contains the maternally inherited autosomes, and two copies of the maternal X chromosome, which are re-marked as paternal so that they are recognized as being inherited from a male in the next generation.

6.4.2 COCCIDS

Imprinting of chromosomes has also been studied in coccids (superfamily *Coccoidea*), a group of insects that includes the *Pseudococcidae* family of mealybugs, and the *Diaspididae* family of armored scale insects. Three complex genetic systems involving imprinting and chromosome elimination or inactivation, have been studied in the coccid insects (reviewed in Khosla *et al.*, 2006). Of these, the lecanoid chromosome system exhibited by a diverse group of *Coccoidea* families, including the mealybugs, has been the most thoroughly investigated.

Sex chromosomes are absent in lecanoid coccids, and thus the chromosomal complement of all zygotes is initially identical. However, in male-determined embryos, a full haploid chromosome set is inactivated via heterochromatinization during embryogenesis. This chromosomal inactivation is non-random; it is consistently the paternally inherited chromosome set that becomes heterochromatic (Brown and Nelson-Rees, 1961). These heterochromatic chromosomes are transcriptionally inactive (Brown and Nelson-Rees, 1961), and thus males are functionally haploid, with the exception of a few tissues that exhibit reversal of heterochromatinization (Nur, 1967). In spermatogenesis of lecanoid coccids, the paternally-inherited heterochromatic chromosome set disintegrates, and thus males transmit only the chromosomes that were inherited maternally. Imprinting in Diaspidoid coccids is similar, but in this case the entire paternally inherited haploid genome is eliminated early in male development, rather than inactivated in heterochromatin (reviewed in Khosla *et al.*, 2006).

6.4.2.1 DNA METHYLATION

Analysis of the role of DNA methylation in imprinting in mealybugs has produced conflicting results. In analyzing methylation at CCGG sequences in the mealybug *Planococcus citri*, one study determined that the paternally inherited chromosomes are hypomethylated in both males and females (Bongiorni *et al.*, 1999). Thus, methylation could serve as a mark that distinguishes the parental origin of the chromosomes, but would probably not contribute directly to the silencing (Bongiorni and Prantera, 2003). However, a second study found no significant difference in the methylation of paternally and maternally inherited chromosomes (Buglia *et al.*, 1999). Sequence-specific analysis of CpG methylation in the mealybug *Planococcus lilacinus* found that male-specific methylation occurs more frequently than female-specific methylation (Mohan and Chandra, 2005). In addition, these sex-specific methylated sequences were associated with transcriptionally silent chromatin, but only in the sex exhibiting methylation, which may suggest a direct link between sex-specific DNA methylation and transcriptional silencing in mealybugs (Mohan and Chandra, 2005). A higher frequency of 5-methylcytosine in males compared to females was also found of two additional species of mealybugs, although only one was deemed statistically significant (Scarbrough *et al.*, 1984). Interestingly, mealybugs also exhibit a significant amount of 5-methylcytosine in other dinucleotide combinations, and some species have been shown to also contain a high frequency of the normally rare 6-methyladenosine and 7-methylguanosine modified bases (Achwal *et al.*, 1983; Deobagkar *et al.*, 1982). The characterization of an active CpA methylase that methylates both CpG and CpA dinucleotides (Devajyothi and Brahmachari, 1992) confirms that mealybugs have the capacity for DNA methylation.

However, the exact role of DNA methylation in imprinting in coccids remains to be elucidated.

6.4.2.2 CHROMATIN STRUCTURE AND HISTONE MODIFICATIONS

Chromatin analysis of the genomes of male and female mealybugs with Micrococcal nuclease, an enzyme impeded by condensed chromatin, demonstrated that approximately 5-10% of the genome is organized into nuclease-resistant chromatin in males, but not females (Khosla *et al.*, 1999; Khosla *et al.*, 1996). The nuclease resistant chromatin sequences were found to be associated with the nuclear matrix (Khosla *et al.*, 1996), and include unique sequences, as well as middle-repetitive sequences distributed throughout the genome (Khosla *et al.*, 1999). Detailed analysis of two middle-repetitive sequences found that these genome-wide sequences are enriched within the nuclease resistant chromatin of male mealybugs, and exhibit different chromatin organization between males and females, and also within male nuclei (Khosla *et al.*, 1999). The specialized organization of these sequences into nuclease resistant chromatin is therefore likely a characteristic of the condensed paternal chromosomes, consistent with their cytologically visible heterochromatic structure. As only 10% of the male genome is organized into this nuclease resistant chromatin, and not 50% as would be expected if it were a property of the entire heterochromatinized chromosome set, it has been hypothesized that these sequences may function as initiation centres for heterochromatinization (Khosla *et al.*, 1999; Khosla *et al.*, 1996). Similarly, they may mediate differential organization of homologous chromosomes within the nucleus, which could serve as an epigenetic mark that distinguishes the genomes and triggers the heterochromatinization of the paternal genome (Khosla *et al.*, 1999).

A protein with similarity to *Drosophila* Heterochromatin Protein 1 (HP1) is encoded by the *pchet2* gene in mealybugs (Bongiorni *et al.*, 2007; Epstein *et al.*, 1992). In males, this protein accumulates at the distinct chromocentre that contains the heterochromatic paternal chromosomes, with little binding elsewhere (Bongiorni *et al.*, 2001). The heterochromatic chromocentre is also strongly enriched for H3K9 (Cowell *et al.*, 2002), and H4K20 methylation (Kourmouli *et al.*, 2004). In females, PCHET2 protein, H3K9 methylation, and H4K20 methylation exhibit a scattered distribution throughout all of the chromosomes. The heterochromatic paternal genome in male mealybugs was also found to be hypoacetylated at histone H4 compared to the euchromatic maternal genome, with an increase in acetylation accompanied by a decrease in condensation (Ferraro *et al.*, 2001).

During male embryogenesis, a dense PCHET2 signal was evident prior to the formation of the chromocenter, indicating that PCHET2 accumulation precedes and likely contributes to, heterochromatinization (Bongiorni *et al.*, 2001). Consistent with this, a knock-down of *pchet2* was accompanied by a decondensation of the paternal chromosomes, a loss of H4K20 methylation, and overall genome instability (Bongiorni *et al.*, 2007). In mammals, constitutive heterochromatin formation is thought to require HP1 binding to H3K9 methylation, which subsequently recruits H4K20 methyltransferases (Schotta *et al.*, 2004). The facultative heterochromatinization of paternal chromosomes observed in male mealybugs appears to be consistent with this model. Analysis of cells undergoing reversal of heterochromatinization found that H3K9 methylation remains associated with the decondensing paternal chromosomes, while H4K20 methylation is lost, and PCHET2 become dispersed (Bongiorni *et al.*, 2007). H3K9 methylation may therefore be the primary epigenetic modification of the paternal chromosomes, and may

be the "imprint" carried by the paternal chromosomes, leading to their heterochromatinization (Bongiorni *et al.*, 2007). The role of non-coding RNAs in coccid imprinting has not yet been studied.

6.4.3 *DROSOPHILA MELANOGASTER*

Drosophila melanogaster is a model organism widely used to study gene expression. While endogenous imprinted genes remain yet to be identified, imprinted expression of transgenes, or marker genes on rearranged chromosomes, indicates that the capacity for differential gene expression based on parental inheritance is certainly present and mechanistically possible. Of the imprinting examples studied in *Drosophila*, all are associated with gene-poor regions of constitutive heterochromatin (reviewed in Lloyd, 2000). For example, imprinting has been observed on rearranged chromosomes when a chromosomal breakage results in the juxtaposition of a euchromatic marker gene with a region of broken heterochromatin. While this type of a disruption frequently results in variegation of the marker gene due to its new position adjacent to heterochromatin, imprinting of the marker gene is only observed in a small number of cases, suggesting it is a unique characteristic of only certain heterochromatin regions or segments. Imprinting is identified when transmission of the rearranged chromosome through one parent causes a significantly different level of marker gene expression than when transmitted through the other parent. In these cases, the imprint's origin appears to involve discrete regions at which the key epigenetic processes act, similar to mammalian imprint centres.

Imprinting has also been observed at a high frequency for transgenes inserted into the heterochromatic Y chromosome (Golic *et al.*, 1998; Maggert and Golic, 2002). While most imprinted transgene insertions on the Y chromosome exhibit increased silencing

when transmitted paternally, the reverse has also been observed, with increased silencing following maternal transmission. Furthermore, as with the imprinted domains in centric heterochromatin, some transgenes exhibit opposite imprinting of the two marker genes, or imprinting of only one of the two marker genes, despite their insertion at the same genomic position. Imprinting in *Drosophila*, may therefore include reciprocal imprinting of closely linked genes, or differential gene response to an imprinting centre, similar to many imprinting clusters in mammals (Lloyd, 2000; Maggert and Golic, 2002). The sequestering of imprinted domains to heterochromatic regions of the genome with low gene density appears to be both mechanistic, with heterochromatic repeat sequences nucleating the imprint, and a result of selection against the inclusion of too many genes in the imprinted domain (Anaka *et al.*, 2009).

The observation of chromosomal loss in the presence of mutations in the *paternal loss inducer* (*pal*) gene provides additional evidence of imprinting in *Drosophila.* In these cases, chromosomal loss is not random, but instead specifically affects chromosomes that were paternally inherited (Baker, 1975; Fitch *et al.*, 1998). The *pal* gene acts exclusively in males and is hypothesized to encode a sperm specific protein that could distinguish the paternal from the maternal chromosomes in the zygote after fertilization. *Drosophila* chromosomes therefore likely carry an imprint that distinguishes them based on their parent-of-origin, despite not normally causing chromosomal loss or inactivation, as in coccids or *Sciara*.

6.4.3.1 HISTONE MODIFICATIONS

The best studied example of imprinting in *Drosophila* is the *Dp(1;f)LJ9* mini-X chromosome (Lloyd *et al.*, 1999). Among the genes imprinted on this mini-X

chromosome is the easily observable eye colour gene, *garnet*. Imprinting of such an easily monitored gene allowed for identification of genes involved in *Drosophila* imprinting (Joanis and Lloyd, 2002). Mutations in several *Suppressor of variegation* (*Su(var)*) genes resulted in a loss of the maintenance of the paternal imprint. These included the well characterized *Su(var)3-9* and *Su(var)2-5* (Joanis and Lloyd, 2002). *Su(var)3-9* encodes a histone methyltransferase that contributes to heterochromatin formation by catalyzing H3K9 methylation (Schotta *et al.*, 2002), a mark that is recognized and bound by the *Su(var)2-5* gene product, Heterochromatin Protein 1 (HP1) (Lachner *et al.*, 2001). Mutations in the gene *Su(var)3-3*, which encodes a histone demethylase that associates with prospective heterochromatic regions and removes the H3K4 methylation mark that is normally associated with active chromatin (Rudolph *et al.*, 2007), also resulted in a loss of paternal silencing (Joanis and Lloyd, 2002). Mutations in two *trithorax* group genes, *trithorax* and *brahma*, exhibited the opposite effect on the paternal imprint (Joanis and Lloyd, 2002), consistent with the role of these two proteins in complexes that participate in the formation and maintenance of active chromatin via activating histone modifications and chromatin remodelling (Simon and Tamkun, 2002). Overall, these results indicate that imprinting in *Drosophila* is likely accomplished by histone modifications that mediate the formation of a repressive heterochromatin structure upon passage through one germ line but not the other. The role of DNA methylation and antisense/non-coding RNA is under active investigation; preliminary results indicate that these epigenetic processes are also involved (MacDonald and Lloyd, 2004; Maggert and Golic, 2004).

6.5 GENOMIC IMPRINTING IN PLANTS

Several endogenous plant genes in both *Arabidopsis thaliana* (thale cress) and *Zea mays* (maize) are imprinted in the endosperm, a triploid tissue formed by the fusion of a haploid sperm from a pollen grain with the diploid central cell in the ovule. While the endosperm does not contribute its genome to the next generation, it nourishes the developing embryo and is an essential component of an angiosperm seed. Four of the five known imprinted genes in *Arabidopsis*, and five of the six known imprinted genes in maize, are expressed maternally and silenced paternally in the endosperm. The thorough examination of several endosperm-specific imprinted genes does not preclude the existence of non-developmentally essential imprinted genes in the embryo or adult plant. Imprinting of such genes could be specific to certain tissues or developmental stages, or could involve partial, rather than complete, silencing of one parental allele. In support of this hypothesis, several paternally-inherited genes and transgenes have been shown to be down-regulated or silenced in the early embryo in *Arabidopsis*, providing evidence that the maternal and paternal genomes are non-equivalent during early embryogenesis (Baroux *et al.*, 2001; Vielle-Calzada *et al.*, 2000).

6.5.1 *MEDEA* IMPRINTING IN *ARABIDOPSIS*

Arabidopsis MEDEA (*MEA*) is an imprinted gene that encodes a SET domain-containing Polycomb group protein homologous to *Drosophila Enhancer of Zeste (E(z))* (Grossniklaus et al., 1998). Polycomb group proteins function in multimeric protein complexes that maintain transcriptional repression by modifying chromatin structure (Orlando, 2003). *E(z)* and *E(z)* homologues are members of the Polycomb Repressive

Complex 2 (PRC2), which exhibits histone methyltransferase activity through the SET-domain of *E(z)* (Muller *et al.*, 2002). In *Arabidopsis*, many of the core components of PRC2 are represented by small gene families rather than single copy genes, and it is hypothesized that diversification of the ancestral PRC2 complex has led to multiple distinct PRC2 complexes that target different genes for repression (Chanvivattana *et al.*, 2004). *MEA* is one of three *E(z)* homologues that has been identified, each of which has at least partially diverged in gene expression pattern and protein function (Chanvivattana *et al.*, 2004).

The *Arabidopsis* Polycomb Group complex that contains MEA also includes the proteins Fertilization-Independent Endosperm (FIE) and Multicopy Suppressor of Ira 1 (MSI1), which are homologues of *Drosophila Extra Sex Combs (Esc)* and *p55,* respectively. This complex is hypothesized to also include Fertilization Independent Seed2 (FIS2), a homologue of *Drosophila Suppressor of Zeste 12* (Chanvivattana *et al.*, 2004; Kohler *et al.*, 2003a). These four genes are members of the *Arabidopsis* Fertilization Independent Seed (FIS) group, a class of genes that are characterized by a mutant phenotype that includes seed development in the absence of fertilization. This Polycomb group complex is therefore also termed the FIS complex. Interestingly, both *FIS2* and *PHERES1*, a known target of the FIS complex, are also imprinted in *Arabidopsis*. Maternal inheritance of a *mea* mutation results in seed abortion, aberrant proliferation of the central cell in the absence of fertilization, and overproliferation of the endosperm following fertilization (Kiyosue *et al.*, 1999). The normal development of seeds inheriting a *mea* mutation paternally provided early evidence of a parent-of-origin effect (Grossniklaus *et al.*, 1998; Kiyosue *et al.*, 1999) that was later shown to be a result of *MEA* imprinting (Kinoshita *et al.*, 1999). Paternal inheritance of a non-functional *mea*

allele has no effect because the imprinted *MEA* gene is not normally expressed from the paternal allele. Similar asymmetrical consequences resulting from the inheritance of mutant alleles have led to the discovery of several mammalian imprinted genes.

6.5.1.1 DNA METHYLATION

Like other imprinted genes identified in plants, imprinting of *MEA* occurs only in the endosperm, where it is expressed from the maternal allele only. Conversely, biallelic expression of *MEA* occurs in the embryo and other tissues of the adult plant (Kinoshita *et al.*, 1999). Early genetic analysis revealed that imprinting of *MEA* requires a maternal copy of the *DEMETER* (*DME*) gene, a DNA glycosylase that is primarily expressed in the central cell prior to fertilization. Maternal *dme* mutations resulted in a lack of maternal *MEA* expression in the central cell and endosperm, while ectopic *DME* expression in the endosperm resulted in *MEA* expression from the paternal allele (Choi *et al.*, 2002). The role of DNA glycosylases in removing mismatched or altered bases from DNA, and the nicks discovered at the *MEA* promoter upon ectopic expression of *DME*, led to the hypothesis that DME may contribute to *MEA* imprinting by excising DNA methylation at the maternal allele.

This hypothesis was supported by the discovery that a maternal mutation in the DNA maintenance methyltransferase *met1* could suppress the *dme* mutant phenotype if a wild-type maternal *MEA* allele was also present (Xiao *et al.*, 2003). The combination of maternal *met1* and *dme* mutations restored *MEA* expression to normal levels from the maternal allele in the endosperm, indicating that MET1 and DME act antagonistically in controlling *MEA* imprinting (Xiao *et al.*, 2003). Three regions of methylation were detected in the *MEA* promoter, with a decrease in methylation detected in the presence of

a *met1* mutation (Xiao *et al.*, 2003). In the endosperm, the maternal alleles were found to be hypomethylated compared to the paternal allele, with maternal methylation increasing in *dme* mutant seeds (Gehring *et al.*, 2006). Thus, expression of *DME* in the central cell prior to fertilization appears to remove *MEA* methylation on the maternal allele and establish a hypomethylated state that is required for its expression. Consistent with this, DME was found to excise 5-methylcytosines *in vitro* (Gehring *et al.*, 2006).

Intriguingly, DME-mediated hypomethylation appears to be a unique requirement for *MEA* expression in the central cell and early endosperm. Hypomethylation of *MEA* is not required in the embryo, where *MEA* is biallelically expressed but exhibits methylation comparable to the silenced paternal allele in the endosperm. Furthermore, in *dme* mutants, the maternal allele is expressed late in endosperm development despite being hypermethylated (Gehring *et al.*, 2006). Why then is DME required to establish hypomethylation of *MEA* in the central cell? Removal of maternal methylation may be a prerequisite for additional modifications required for expression in that environment such as the removal of histone methylation or changes in chromatin structure. Alternatively, the DME enzyme may directly mediate removal of both DNA and histone methylation at the *MEA* promoter, or it may activate *MEA* indirectly by removing methylation on an additional gene that, in turn, further modifies and activates the *MEA* locus (Jullien *et al.*, 2006a).

6.5.1.2 HISTONE MODIFICATIONS

Hypomethylation of the paternal allele does not result in its expression in the endosperm, indicating that DNA methylation does not directly maintain paternal *MEA* silencing. Instead, the FIS Polycomb group complex containing MEA itself was found to

be essential for this repression. Maternal mutations in *mea*, *fie*, *fis2*, or *msl1* resulted in expression from the paternal allele (Gehring *et al.*, 2006; Jullien *et al.*, 2006a), and chromatin immunoprecipitation analysis confirmed that MEA can physically interact with its own promoter (Baroux *et al.*, 2006). Furthermore, the paternal allele was found to be enriched in H3K27 methylation (Gehring *et al.*, 2006; Jullien *et al.*, 2006a), a repressive histone mark that is well characterized in *Drosophila* and mammals where it is catalyzed by the *MEA* homologues *E(z)* and *E(z)H2* (Cao *et al.*, 2002; Muller *et al.*, 2002). The function of *MEA* and H3K27 methylation appears to be conserved in *Arabidopsis*, and a maternal *mea* mutation resulted in a decrease in H3K27 methylation at the paternal *MEA* allele and a loss of silencing (Gehring *et al.*, 2006). *MEA* is therefore a gene that controls its own imprinting, with the maternally expressed protein contributing to the silencing of the paternal allele. The observation that a paternal mutation in *fie*, a single copy gene essential for all known *Arabidopsis* Polycomb complexes, resulted in *MEA* expression from the paternal allele in the endosperm, suggests that *MEA* silencing must also be maintained by a Polycomb group complex during male gametogenesis for successful imprinting in the endosperm (Jullien *et al.*, 2006a). The current model of *MEA* imprinting in *Arabidopsis* is summarized in Figure 6.4.

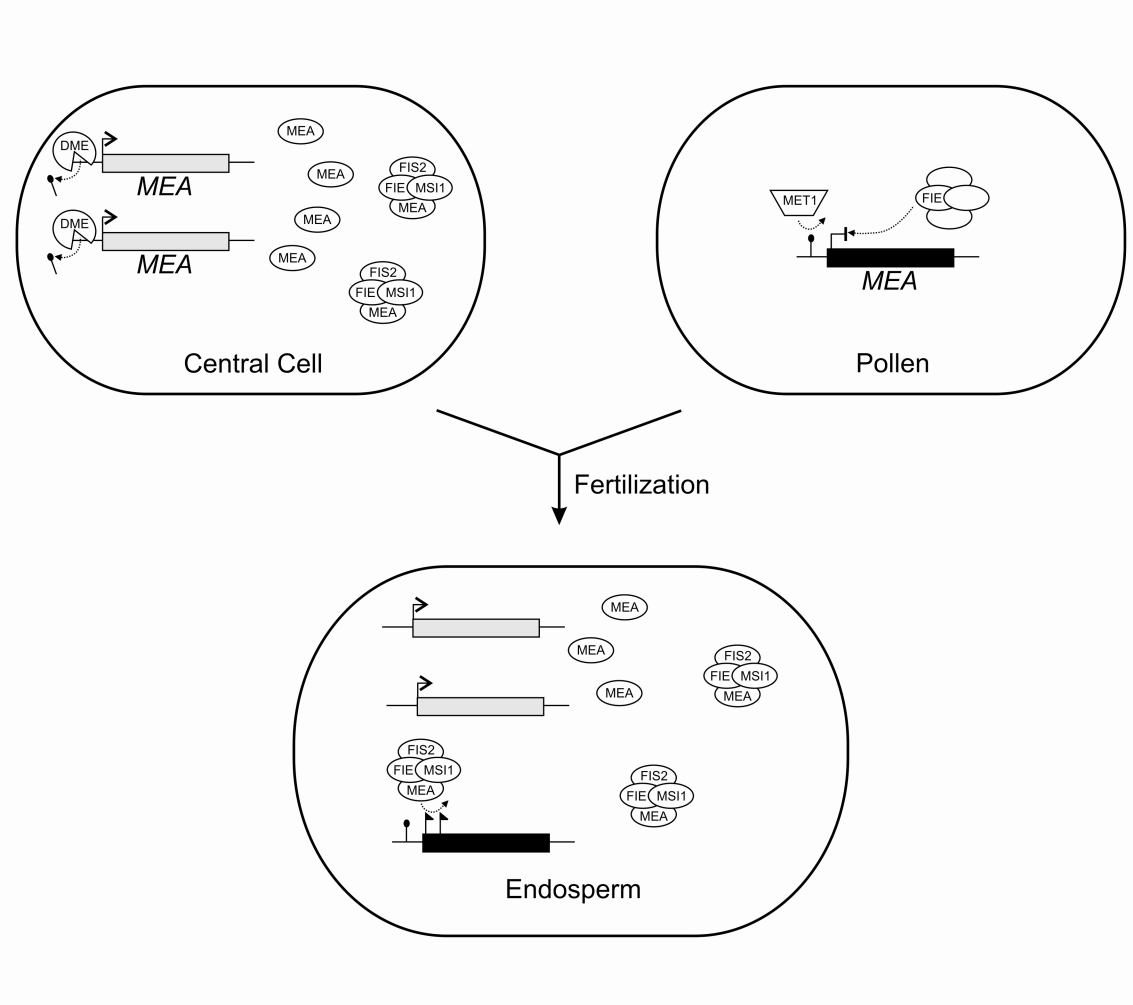


Figure 6.4 Imprinting of the *MEA* gene in Arabidopsis. In the central cell prior to fertilization, DEMETER (DME) removes DNA methylation (filled lollipops) from the maternal *MEA* alleles (light grey). The *MEA* gene is expressed and produces MEA protein, which assembles into the FIS Polycomb group complex. In the male pollen, DNA methylation is maintained by MET1, and a polycomb group complex containing FIE is necessary to maintain *MEA* silencing. Following fertilization, silencing of the paternal *MEA* allele in the endosperm is maintained by the FIS Polycomb group complex that contains the maternally expressed MEA protein. This complex catalyzes H3K27 methylation (filled flags) at the paternal *MEA* locus, a modification that inhibits paternal *MEA* expression via the formation of a repressive chromatin structure.

6.5.2 IMPRINTING OF OTHER PLANT GENES

Imprinting of the other known imprinted genes in *Arabidopsis* is accomplished using similar mechanisms to those acting at the *MEA* locus. The maternally expressed imprinted genes *FWA*, *FIS2*, and *MPC* require the maintenance methyltransferase MET1 to maintain methylation and paternal allele repression, and require maternal DME to excise methylation and activate expression of the maternal allele (Jullien *et al.*, 2006b; Kinoshita *et al.*, 2004; Tiwari *et al.*, 2008). The *PHERES1* gene is currently the only known imprinted gene in *Arabidopsis* that is paternally expressed in the endosperm. *PHERES1* expression requires DNA methylation at a region located 3' to the gene. Conversely, repression of *PHERES1* requires both hypomethylation of the 3' region and the Arabidopsis FIS Polycomb complex, which catalyzes H3K27 methylation at the *PHERES1* promoter (Kohler *et al.*, 2003b; Makarevich *et al.*, 2006; Makarevich *et al.*, 2008). Repression of the maternal allele may require hypomethylation of the 3' sequence in order to facilitate the binding of a methylation-sensitive chromatin-binding protein. Such a protein could mediate the formation of a repressive chromatin loop and/or the recruitment of the FIS Polycomb complex to the *PHERES1* promoter. This would be similar to the mammalian *Igf2/H19* imprint centre, where CTCF binds to the unmethylated ICR and orchestrates a chromatin loop that recruits PRC2 to *Igf2*, leading to its repression.

6.5.2.1 *DE NOVO* DNA METHYLATION AND SMALL RNAS

Thus far, *de novo* methylation has not been observed in imprinting in *Arabidopsis*. Instead, methylation must be maintained, and then selectively removed in the endosperm. As the endosperm is terminally differentiated and does not contribute its genome to the

next generation, re-methylation is not required. However, there is evidence that *de novo* methylation may play a role in imprinting in maize. In maize, two orthologues of the *Arabidopsis FIE* gene, *FIE1* and *FIE2,* are imprinted and expressed from the maternal alleles during endosperm development (Danilevskaya *et al.*, 2003; Gutierrez-Marcos *et al.*, 2003). Similar to imprinted genes in *Arabidopsis*, *FIE1* is hypermethylated on the silenced paternal allele in the endosperm, with methylation also detected in the sperm, egg, and embryo, but not in the central cell that will contribute to the endosperm (Gutierrez-Marcos *et al.*, 2006). It is thus likely that *FIE1* methylation is actively removed from the maternal alleles in the central cell prior to fertilization. Examination of *FIE1* histone modifications have demonstrated that the silenced paternal allele is enriched for H3K27 methylation, while the expressed maternal allele is enriched for H3 and H4 acetylation, and H3K4 methylation, consistent with the histone patterns of many other imprinted genes (Haun and Springer, 2008). While the status of histone modifications at *FIE2* is unknown, its methylation pattern is currently unique among the known imprinted genes in plants. Methylation is absent from *FIE2* in the sperm, egg, embryo, and central cell, but the silenced paternal *FIE2* allele exhibits hypermethylation in the endosperm (Gutierrez-Marcos *et al.*, 2006). The imprinted paternal *FIE2* allele must therefore acquire *de novo* methylation in the endosperm. This observation also further implies that the paternal *FIE2* allele carries a non-DNA methylation based imprint that identifies it for hypermethylation and repression in the endosperm (Scott and Spielman, 2006).

There is considerable evidence that *de novo* methylation in plants, mediated by the *de novo* methyltransferases Domains Rearranged Methyltransferase 1 (DRM1) and Domains Rearranged Methyltransferase 2 (DRM2), is directed by small RNAs. While this pathway has been extensively studied for genes and transgenes that are methylated

and silenced in the adult plant, it also has applicability to any imprinted genes that may acquire *de novo* methylation, such as the *FIE2* gene in maize. The *FWA* gene that is imprinted in the *Arabidopsis* endosperm is biallelically silenced in the embryo and adult plant, and provides an excellent system for methylation studies, as it exhibits CpG and non-CpG methylation at direct repeats located 5' to the gene (Soppe *et al.*, 2000). *FWA* transgenes have been observed to acquire *de novo* methylation and silencing in wild type plants, but not in *drm1/drm2* mutants (Cao and Jacobsen, 2002b). Interestingly, *de novo* methylation of *FWA* transgenes also requires a functional RNAi pathway. Mutations in several genes in the siRNA-generating pathway cause a similar loss of *de novo* methylation and silencing of the *FWA* transgenes (Chan *et al.*, 2004). In the *drm1/drm2* and RNAi mutants, non-CpG methylation at the endogenous *FWA* locus was also lost (Cao and Jacobsen, 2002a; Chan *et al.*, 2004), suggesting that the *de novo* methyltransferase and RNAi pathways may be required to maintain these modifications, in addition to establishing *de novo* methylation. The observation that the paternally inherited *FIE2* allele is extensively methylated at CpG and non-CpG sites in the endosperm, whereas *FIE1* methylation is almost entirely restricted to CpG sites (Gutierrez-Marcos *et al.*, 2006), likely indicates that a similar RNA-mediated methylation pathway directs *de novo* methylation of the imprinted paternal *FIE2* allele in maize.

There is also evidence that *trans*-communication between alleles is important in the methylation and silencing processes, and may similarly be important in plant imprinting. Introduction of an *FWA* transgene into *fwa-1* mutant plants, in which the endogenous *FWA* gene is hypomethylated and expressed, occasionally results in methylation and silencing of the endogenous *fwa-1* mutant allele, and rescue of the mutant phenotype (Chan *et al.*, 2006; Soppe *et al.*, 2000). Furthermore, while an *FWA*

transgene very consistently acquires methylation and silencing when introduced into wild-type plants, in *fwa-1* plants, methylation and silencing of the introduced *FWA* transgene is inefficient (Chan *et al.*, 2006). The epigenetic status of the endogenous *FWA* locus can therefore influence that of the *FWA* transgene, and vice versa. Given that siRNAs accumulate equally in wild-type and *fwa-1* mutant plants, additional chromatin or DNA modifications are likely required for efficient RNA-directed DNA methylation and epigenetic silencing (Chan *et al.*, 2006). A similar mechanism involving RNA-mediated allelic communication and chromatin modifications has been proposed to function in maize *b1* paramutation, an epigenetic process related to imprinting that produces meiotically stable changes in chromatin structure and gene expression (Chandler, 2007). Overall, this evidence indicates that all three core epigenetic processes – DNA methylation, histone modifications, and small RNA-mediated modifications – underlie genomic imprinting in plants.

6.6 EVOLUTIONARY CONSERVATION OF GENOMIC IMPRINTING

Examination of genomic imprinting in organisms as diverse as mammals, insects, and plants suggests that imprinting is accomplished using phylogenetically conserved epigenetic mechanisms. Histone modifications are a common theme in genomic imprinting, and are utilized in all species examined in order to establish higher-order chromatin structures that contribute to the imprinting of whole chromosomes, single genes, or gene clusters. At the mammalian *H19/Igf2* imprint centre, the maternal chromosome is enriched for repressive histone modifications at *Igf2*, and activating histone modifications at *H19*. Conversely, and consistent with the gene expression

patterns, the paternal chromosome is enriched for activating histone modifications at *Igf2*, and repressive histone modifications at *H19*. A similar role for histone modifications is observed in insect imprinting. In *Sciara*, unique patterns of histone acetylation and methylation contribute to chromosome elimination. The inactive paternal chromosomes in coccids are associated with repressive histone methylation, and histone H4 hypoacetylation. On the *Drosophila Dp(1;f)LJ9* imprinted chromosome, the silenced paternal imprint requires a protein that catalyzes repressive histone methylation, and a protein that removes activating histone methylation. The importance of histone modifications is also demonstrated for imprinted genes in plants, in which the silenced alleles are associated with repressive histone modifications and the expressed alleles are associated with activating histone modifications. In addition, both plants and mammals utilize a homologous Polycomb complex, the PRC2 complex in mammals and the FIS complex in plants, to catalyze repressive H3K27 methylation at silenced imprinted alleles.

DNA methylation is another common theme in plant and mammalian imprinting, and may also play a role in insect imprinting systems. Non-coding and small RNAs exhibit conserved functional roles in catalyzing heterochromatin formation in mammals, insects and plants; it is therefore likely that these will be demonstrated to be important in many imprinting systems in different species. In mammals, non-coding RNAs are very common in imprinting clusters, and are both imprinted, and regulate imprinted expression of other genes. Within the *H19/Igf2* imprint centre, for example, the *H19* gene is an imprinted non-coding RNA, the ICR is biallelically transcribed and produces sense and antisense transcripts, and the HUC1 and HUC2 sequences are biallelically transcribed. In plants, *de novo* DNA methylation can be guided by small RNAs, and thus imprinted plant

genes that exhibit *de novo* methylation, such as *FIE2* in maize, likely use this RNA-mediated mechanism. Given that DNA methylation, histone modifications, chromatin structure and non-coding RNAs are frequently interconnected and mutually reinforcing, it is not surprising that all of these conserved mechanisms have been implicated in genomic imprinting.

In addition to the similarities apparent in the mechanisms underlying imprinting, the broad consequences of imprinting are also similar across this diverse group of organisms. In both plants and mammals, inactivation of one parental allele via imprinting is common. There is accumulating evidence that this type of imprinting also occurs in *Drosophila*. In addition, the imprinted inactivation of whole chromosomes that are paternal in origin is not unique to insects. In marsupials, the paternal X chromosome is inactivated in females as a method of dosage compensation (VandeBerg *et al.*, 1987). In eutherian mammals, X chromosome inactivation in the somatic tissues of females is random; however, in the extraembryonic tissue of mice, the paternal X chromosome is imprinted, and is preferentially inactivated (Takagi and Sasaki, 1975). In marsupials and in the extraembryonic tissues of mice, the inactivated X is hypoacetylated at histone H4 (Wakefield *et al.*, 1997). This mark is similarly associated with the inactive paternal chromosomes in coccids, and the eliminated paternal chromosomes in *Sciara*.

The evolutionary conservation of the epigenetic mechanisms that underlie imprinting is further exemplified by the results of transgenic *Drosophila* experiments. Silencer elements from both the mouse and human *H19/Igf2* ICRs, and the mouse A6-A4 region, silence reporter genes in transgenic *Drosophila*. The mouse ICR is even biallelically transcribed and produces non-coding RNAs in *Drosophila*, as it does at its endogenous locus. Similarly, the human Prader-Willi imprint centre also functions as a

silencer in *Drosophila* (Lyko *et al.*, 1998). Transgenic organisms are therefore valuable tools in studying genomic imprinting, as they can be used to analyze the function of genetic sequences and epigenetic processes, as well as the conservation of epigenetic mechanisms. The fact that mammalian ICRs frequently function as epigenetic silencers but do not imprint in *Drosophila* may indicate that a silenced epigenetic state is often the default. Silencing may use core epigenetic mechanisms that are highly conserved from one species to another, while imprinting is a more divergent, gamete-specific, modification of these conserved silencing processes.

Discordances in imprinting within the mammalian lineage may indicate that imprinting of genes can evolve rapidly, and thus genes that are subject to imprinting may differ in imprinted status, expression pattern, and regulatory sequences, from one species to another. Nevertheless, the evidence presented here indicates that at the core, genomic imprinting occurs via exploiting conserved epigenetic silencing mechanisms in order to establish distinct patterns of epigenetic gene regulation.

6.7 REFERENCES

Achwal CW, Iyer CA, Chandra HS (1983) Immunochemical evidence for the presence of 5mC, 6mA and 7mG in human, Drosophila and mealybug DNA. *FEBS Lett* **158**: 353-358.

Ainscough JF, Dandolo L, Surani MA (2000a) Appropriate expression of the mouse H19 gene utilises three or more distinct enhancer regions spread over more than 130 kb. *Mech Dev* **91**: 365-368.

Ainscough JF, John RM, Barton SC, Surani MA (2000b) A skeletal muscle-specific mouse Igf2 repressor lies 40 kb downstream of the gene. *Development* **127**: 3923-3930.

Alleman M, Doctor J (2000) Genomic imprinting in plants: observations and evolutionary implications. *Plant Mol Biol* **43**: 147-161.

Anaka M, Lynn A, McGinn P, Lloyd VK (2009) Genomic Imprinting in Drosophila has properties of both mammalian and insect imprinting. *Dev Genes Evol* **219**: 59-66.

Arney KL, Bae E, Olsen C, Drewell RA (2006) The human and mouse H19 imprinting control regions harbor an evolutionarily conserved silencer element that functions on transgenes in Drosophila. *Dev Genes Evol* **216**: 811-819.

Baker BS (1975) Paternal loss (pal): a meiotic mutant in Drosophila melanogaster causing loss of paternal chromosomes. *Genetics* **80**: 267-296.

Bao N, Lye KW, Barton MK (2004) MicroRNA binding sites in Arabidopsis class III HD-ZIP mRNAs are required for methylation of the template chromosome. *Dev Cell* **7**: 653-662.

Baroux C, Blanvillain R, Gallois P (2001) Paternally inherited transgenes are down-regulated but retain low activity during early embryogenesis in Arabidopsis. *FEBS Lett* **509**: 11-16.

Baroux C, Gagliardini V, Page DR, Grossniklaus U (2006) Dynamic regulatory interactions of Polycomb group genes: MEDEA autoregulation is required for imprinted gene expression in Arabidopsis. *Genes Dev* **20**: 1081-1086.

Bartolomei MS, Webber AL, Brunkow ME, Tilghman SM (1993) Epigenetic mechanisms underlying the imprinting of the mouse H19 gene. *Genes Dev* **7**: 1663-1673.

Bartolomei MS, Zemel S, Tilghman SM (1991) Parental imprinting of the mouse H19 gene. *Nature* **351**: 153-155.

Bean CJ, Schaner CE, Kelly WG (2004) Meiotic pairing and imprinted X chromatin assembly in Caenorhabditis elegans. *Nat Genet* **36**: 100-105.

Beaudet AL, Jiang YH (2002) A rheostat model for a rapid and reversible form of imprinting-dependent evolution. *Am J Hum Genet* **70**: 1389-1397.

Bell AC, Felsenfeld G (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* **405**: 482-485.

Bernstein E, Allis CD (2005) RNA meets chromatin. *Genes Dev* **19**: 1635-1655.

Berry RO (1941) Chromosome behavior in the germ cells and development of the gonads in Sciara ocellaris. *Journal of Morphology* **68**: 547-583.

Bongiorni S, Cintio O, Prantera G (1999) The relationship between DNA methylation and chromosome imprinting in the coccid Planococcus citri. *Genetics* **151**: 1471-1478.

Bongiorni S, Mazzuoli M, Masci S, Prantera G (2001) Facultative heterochromatization in parahaploid male mealybugs: involvement of a heterochromatin-associated protein. *Development* **128**: 3809-3817.

Bongiorni S, Pasqualini B, Taranta M, Singh PB, Prantera G (2007) Epigenetic regulation of facultative heterochromatinisation in Planococcus citri via the Me(3)K9H3-HP1-Me(3)K20H4 pathway. *J Cell Sci* **120**: 1072-1080.

Bongiorni S, Prantera G (2003) Imprinted facultative heterochromatization in mealybugs. *Genetica* **117**: 271-279.

Brown SW, Nelson-Rees WA (1961) Radiation Analysis of a Lecanoid Genetic System. *Genetics* **46**: 983-1007.

Buglia G, Predazzi V, Ferraro M (1999) Cytosine methylation is not involved in the heterochromatization of the paternal genome of mealybug Planococcus citri. *Chromosome Res* **7**: 71-73.

Buhler M, Verdel A, Moazed D (2006) Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing. *Cell* **125**: 873-886.

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

Cao X, Jacobsen SE (2002a) Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. *Proc Natl Acad Sci U S A* **99 Suppl 4**: 16491-16498.

Cao X, Jacobsen SE (2002b) Role of the arabidopsis DRM methyltransferases in de novo DNA methylation and gene silencing. *Curr Biol* **12**: 1138-1144.

Cattanach BM, Kirk M (1985) Differential activity of maternally and paternally derived chromosome regions in mice. *Nature* **315**: 496-498.

Chan SW, Zhang X, Bernatavichute YV, Jacobsen SE (2006) Two-step recruitment of RNA-directed DNA methylation to tandem repeats. *PLoS Biol* **4**: e363.

Chan SW, Zilberman D, Xie Z, Johansen LK, Carrington JC, Jacobsen SE (2004) RNA silencing genes control de novo DNA methylation. *Science* **303**: 1336.

Chandler VL (2007) Paramutation: from maize to mice. *Cell* **128**: 641-645.

Chanvivattana Y, Bishopp A, Schubert D, Stock C, Moon YH, Sung ZR, Goodrich J (2004) Interaction of Polycomb-group proteins controlling flowering in Arabidopsis. *Development* **131**: 5263-5276.

Choi Y, Gehring M, Johnson L, Hannon M, Harada JJ, Goldberg RB, Jacobsen SE, Fischer RL (2002) DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in arabidopsis. *Cell* **110**: 33-42.

Constancia M, Dean W, Lopes S, Moore T, Kelsey G, Reik W (2000) Deletion of a silencer element in Igf2 results in loss of imprinting independent of H19. *Nature Genetics* **26**: 203-206.

Cowell IG, Aucott R, Mahadevaiah SK, Burgoyne PS, Huskisson N, Bongiorni S, Prantera G, Fanti L, Pimpinelli S, Wu R, Gilbert DM, Shi W, Fundele R, Morrison H, Jeppesen P, Singh PB (2002) Heterochromatin, HP1 and methylation at lysine 9 of histone H3 in animals. *Chromosoma* **111**: 22-36.

Crouse HV (1960) The Controlling Element in Sex Chromosome Behavior in Sciara. *Genetics* **45**: 1429-1443.

Crouse HV (1979) X heterochromatin subdivision and cytogenetic analysis in Sciara coprophila (diptera, sciaridae). *Chromosoma* **74**: 219-239.

Crouse HV, Gerbi SA, Liang CM, Magnus L, Mercer IM (1977) Localization of ribosomal DNA within the proximal X heterochromatin of Sciara coprophila (Diptera, Sciaridae). *Chromosoma* **64**: 305-318.

Danilevskaya ON, Hermon P, Hantke S, Muszynski MG, Kollipara K, Ananiev EV (2003) Duplicated fie genes in maize: expression pattern and imprinting suggest distinct functions. *Plant Cell* **15**: 425-438.

Davies K, Bowden L, Smith P, Dean W, Hill D, Furuumi H, Sasaki H, Cattanach B, Reik W (2002) Disruption of mesodermal enhancers for Igf2 in the minute mutant. *Development* **129**: 1657-1668.

Davies W, Isles A, Smith R, Karunadasa D, Burrmann D, Humby T, Ojarikre O, Biggin C, Skuse D, Burgoyne P, Wilkinson L (2005) Xlr3b is a new imprinted candidate for X-linked parent-of-origin effects on cognitive function in mice. *Nat Genet* **37**: 625-629.

Day T, Bonduriansky R (2004) Intralocus sexual conflict can drive the evolution of genomic imprinting. *Genetics* **167**: 1537-1546.

DeChiara TM, Robertson EJ, Efstratiadis A (1991) Parental imprinting of the mouse insulin-like growth factor II gene. *Cell* **64**: 849-859.

Deobagkar D, Muralidharan K, Devare S, Kalghatgi K, Chandra H (1982) The mealybug chromosome system I: Unusual methylated bases and dinucleotides in DNA of a Planococcus species. *Journal of Biosciences* **4**: 513-526.

Devajyothi C, Brahmachari V (1992) Detection of a CpA methylase in an insect system: characterization and substrate specificity. *Mol Cell Biochem* **110**: 103-111.

Drewell RA, Arney KL, Arima T, Barton SC, Brenton JD, Surani MA (2002) Novel conserved elements upstream of the H19 gene are transcribed and act as mesodermal enhancers. *Development* **129**: 1205-1213.

Drewell RA, Brenton JD, Ainscough JF, Barton SC, Hilton KJ, Arney KL, Dandolo L, Surani MA (2000) Deletion of a silencer element disrupts H19 imprinting independently of a DNA methylation epigenetic switch. *Development* **127**: 3419-3428.

Eden S, Constancia M, Hashimshony T, Dean W, Goldstein B, Johnson AC, Keshet I, Reik W, Cedar H (2001) An upstream repressor element plays a role in Igf2 imprinting. *Embo J* **20**: 3518-3525.

Engel N, Thorvaldsen JL, Bartolomei MS (2006) CTCF binding sites promote transcription initiation and prevent DNA methylation on the maternal allele at the imprinted H19/Igf2 locus. *Hum Mol Genet* **15**: 2945-2954.

Epstein H, James TC, Singh PB (1992) Cloning and expression of Drosophila HP1 homologs from a mealybug, Planococcus citri. *J Cell Sci* **101 (Pt 2)**: 463-474.

Erhardt S, Lyko F, Ainscough JF, Surani MA, Paro R (2003) Polycomb-group proteins are involved in silencing processes caused by a transgenic element from the murine imprinted H19/Igf2 region in Drosophila. *Dev Genes Evol* **213**: 336-344.

Ferraro M, Buglia GL, Romano F (2001) Involvement of histone H4 acetylation in the epigenetic inheritance of different activity states of maternally and paternally derived genomes in the mealybug Planococcus citri. *Chromosoma* **110**: 93-101.

Fitch KR, Yasuda GK, Owens KN, Wakimoto BT (1998) Paternal effects in Drosophila: implications for mechanisms of early development. *Curr Top Dev Biol* **38**: 1-34.

Fuks F (2005) DNA methylation and histone modifications: teaming up to silence genes. *Curr Opin Genet Dev* **15**: 490-495.

Gehring M, Huh JH, Hsieh TF, Penterman J, Choi Y, Harada JJ, Goldberg RB, Fischer RL (2006) DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation. *Cell* **124**: 495-506.

Goday C, Esteban MR (2001) Chromosome elimination in sciarid flies. *Bioessays* **23**: 242-250.

Goday C, Ruiz MF (2002) Differential acetylation of histones H3 and H4 in paternal and maternal germline chromosomes during development of sciarid flies. *J Cell Sci* **115**: 4765-4775.

Golic KG, Golic MM, Pimpinelli S (1998) Imprinted control of gene activity in Drosophila. *Curr Biol* **8**: 1273-1276.

Greciano PG, Goday C (2006) Methylation of histone H3 at Lys4 differs between paternal and maternal chromosomes in Sciara ocellaris germline development. *J Cell Sci* **119**: 4667-4677.

Grewal SI, Moazed D (2003) Heterochromatin and epigenetic control of gene expression. *Science* **301**: 798-802.

Grossniklaus U, Vielle-Calzada JP, Hoeppner MA, Gagliano WB (1998) Maternal control of embryogenesis by MEDEA, a polycomb group gene in Arabidopsis. *Science* **280**: 446-450.

Gutierrez-Marcos JF, Costa LM, Pra MD, Scholten S, Kranz E, Perez P, Dickinson HG (2006) Epigenetic asymmetry of imprinted genes in plant gametes. *Nature Genetics* **38**: 876-878.

Gutierrez-Marcos JF, Pennington PD, Costa LM, Dickinson HG (2003) Imprinting in the endosperm: a possible role in preventing wide hybridization. *Philos Trans R Soc Lond B Biol Sci* **358**: 1105-1111.

Han L, Lee DH, Szabo PE (2008) CTCF is the master organizer of domain-wide allele-specific chromatin at the H19/Igf2 imprinted region. *Mol Cell Biol* **28**: 1124-1135.

Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM (2000) CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* **405**: 486-489.

Haun WJ, Springer NM (2008) Maternal and paternal alleles exhibit differential histone methylation and acetylation at maize imprinted genes. *Plant J* **56**: 903-912.

Holliday R (1990) Genomic imprinting and allelic exclusion. *Dev Suppl*: 125-129.

Irvine DV, Zaratiegui M, Tolia NH, Goto DB, Chitwood DH, Vaughn MW, Joshua-Tor L, Martienssen RA (2006) Argonaute slicing is required for heterochromatic silencing and spreading. *Science* **313**: 1134-1137.

Iwasa Y, Pomiankowski A (1999) Sex specific X chromosome expression caused by genomic imprinting. *J Theor Biol* **197**: 487-495.

Joanis V, Lloyd VK (2002) Genomic imprinting in Drosophila is maintained by the products of Suppressor of variegation and trithorax group, but not Polycomb group, genes. *Mol Genet Genomics* **268**: 103-112.

Jones BK, Levorse J, Tilghman SM (2001) Deletion of a nuclease-sensitive region between the Igf2 and H19 genes leads to Igf2 misregulation and increased adiposity. *Hum Mol Genet* **10**: 807-814.

Jullien PE, Katz A, Oliva M, Ohad N, Berger F (2006a) Polycomb group complexes self-regulate imprinting of the Polycomb group gene MEDEA in Arabidopsis. *Curr Biol* **16**: 486-492.

Jullien PE, Kinoshita T, Ohad N, Berger F (2006b) Maintenance of DNA methylation during the Arabidopsis life cycle is essential for parental imprinting. *Plant Cell* **18**: 1360-1372.

Kaffer CR, Srivastava M, Park KY, Ives E, Hsieh S, Batlle J, Grinberg A, Huang SP, Pfeifer K (2000) A transcriptional insulator at the imprinted H19/Igf2 locus. *Genes Dev* **14**: 1908-1919.

Khatib H, Zaitoun I, Kim ES (2007) Comparative analysis of sequence characteristics of imprinted genes in human, mouse, and cattle. *Mamm Genome* **18**: 538-547.

Khosla S, Augustus M, Brahmachari V (1999) Sex-specific organisation of middle repetitive DNA sequences in the mealybug Planococcus lilacinus. *Nucleic Acids Res* **27**: 3745-3751.

Khosla S, Kantheti P, Brahmachari V, Chandra HS (1996) A male-specific nuclease-resistant chromatin fraction in the mealybug Planococcus lilacinus. *Chromosoma* **104**: 386-392.

Khosla S, Mendiratta G, Brahmachari V (2006) Genomic imprinting in the mealybugs. *Cytogenetic & Genome Research* **113**: 41-52.

Kinoshita T, Miura A, Choi Y, Kinoshita Y, Cao X, Jacobsen SE, Fischer RL, Kakutani T (2004) One-way control of FWA imprinting in Arabidopsis endosperm by DNA methylation. *Science* **303**: 521-523.

Kinoshita T, Yadegari R, Harada JJ, Goldberg RB, Fischer RL (1999) Imprinting of the MEDEA polycomb gene in the Arabidopsis endosperm. *Plant Cell* **11**: 1945-1952.

Kiyosue T, Ohad N, Yadegari R, Hannon M, Dinneny J, Wells D, Katz A, Margossian L, Harada JJ, Goldberg RB, Fischer RL (1999) Control of fertilization-independent endosperm development by the MEDEA polycomb gene in Arabidopsis. *Proc Natl Acad Sci U S A* **96**: 4186-4191.

Kloc A, Zaratiegui M, Nora E, Martienssen R (2008) RNA interference guides histone modification during the S phase of chromosomal replication. *Curr Biol* **18**: 490-495.

Kohler C, Hennig L, Bouveret R, Gheyselinck J, Grossniklaus U, Gruissem W (2003a) Arabidopsis MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed development. *Embo J* **22**: 4804-4814.

Kohler C, Hennig L, Spillane C, Pien S, Gruissem W, Grossniklaus U (2003b) The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene PHERES1. *Genes Dev* **17**: 1540-1553.

Koide T, Ainscough J, Wijgerde M, Surani MA (1994) Comparative analysis of Igf-2/H19 imprinted domain: identification of a highly conserved intergenic DNase I hypersensitive region. *Genomics* **24**: 1-8.

Kono T (2006) Genomic imprinting is a barrier to parthenogenesis in mammals. *Cytogenetic & Genome Research* **113**: 31-35.

Kourmouli N, Jeppesen P, Mahadevhaiah S, Burgoyne P, Wu R, Gilbert DM, Bongiorni S, Prantera G, Fanti L, Pimpinelli S, Shi W, Fundele R, Singh PB (2004) Heterochromatin and tri-methylated lysine 20 of histone H4 in animals. *J Cell Sci* **117**: 2491-2501.

Kubai DF (1987) Nonrandom chromosome arrangements in germ line nuclei of Sciara coprophila males: the basis for nonrandom chromosome segregation on the meiosis I spindle. *J Cell Biol* **105**: 2433-2446.

Kurukuti S, Tiwari VK, Tavoosidana G, Pugacheva E, Murrell A, Zhao Z, Lobanenkov V, Reik W, Ohlsson R (2006) CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. *Proc Natl Acad Sci U S A* **103**: 10684-10689.

Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**: 116-120.

LaJeunesse D, Shearn A (1996) E(z): a polycomb group gene or a trithorax group gene? *Development* **122**: 2189-2197.

Leighton PA, Saam JR, Ingram RS, Stewart CL, Tilghman SM (1995) An enhancer deletion affects both H19 and Igf2 expression. *Genes Dev* **9**: 2079-2089.

Li T, Hu JF, Qiu X, Ling J, Chen H, Wang S, Hou A, Vu TH, Hoffman AR (2008) CTCF regulates allelic expression of Igf2 by orchestrating a promoter-polycomb repressive complex-2 intrachromosomal loop. *Mol Cell Biol* **28**: 6473-6482.

Lippman Z, Martienssen R (2004) The role of RNA interference in heterochromatic silencing. *Nature* **431**: 364-370.

Lloyd V (2000) Parental imprinting in Drosophila. *Genetica* **109**: 35-44.

Lloyd VK, Sinclair DA, Grigliatti TA (1999) Genomic imprinting and position-effect variegation in Drosophila melanogaster. *Genetics* **151**: 1503-1516.

Lyko F, Brenton JD, Surani MA, Paro R (1997) An imprinting element from the mouse H19 locus functions as a silencer in Drosophila. *Nature Genetics* **16**: 171-173.

Lyko F, Buiting K, Horsthemke B, Paro R (1998) Identification of a silencing element in the human 15q11-q13 imprinting center by using transgenic Drosophila. *Proc Natl Acad Sci U S A* **95**: 1698-1702.

MacDonald WA, Lloyd VK (2004) Genomic imprinting in Drosophila and the role of non-coding RNAs. *45th Annual Drosophila Research Conference,* **Abstract 323B**

Maggert KA, Golic KG (2002) The Y chromosome of Drosophila melanogaster exhibits chromosome-wide imprinting. *Genetics* **162**: 1245-1258.

Maggert KA, Golic KG (2004) DNA methylation affects genomic imprinting. *45th Annual Drosophila Research Conference,* **Abstract 133**

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

Makarevich G, Leroy O, Akinci U, Schubert D, Clarenz O, Goodrich J, Grossniklaus U, Kohler C (2006) Different Polycomb group complexes regulate common target genes in Arabidopsis. *EMBO Rep* **7**: 947-952.

Makarevich G, Villar CB, Erilova A, Kohler C (2008) Mechanism of PHERES1 imprinting in Arabidopsis. *J Cell Sci* **121**: 906-912.

Martin CC, McGowan R (1995) Genotype-specific modifiers of transgene methylation and expression in the zebrafish, Danio rerio. *Genet Res* **65**: 21-28.

Matzke MA, Birchler JA (2005) RNAi-mediated pathways in the nucleus. *Nat Rev Genet* **6**: 24-35.

Mayer C, Schmitz KM, Li J, Grummt I, Santoro R (2006) Intergenic transcripts regulate the epigenetic state of rRNA genes. *Mol Cell* **22**: 351-361.

Mohan KN, Chandra HS (2005) Isolation and analysis of sequences showing sex-specific cytosine methylation in the mealybug Planococcus lilacinus. *Mol Genet Genomics* **274**: 557-568.

Moore T, Constancia M, Zubair M, Bailleul B, Feil R, Sasaki H, Reik W (1997) Multiple imprinted sense and antisense transcripts, differential methylation and tandem repeats in a putative imprinting control region upstream of mouse Igf2. *Proc Natl Acad Sci U S A* **94**: 12509-12514.

Moore T, Haig D (1991) Genomic imprinting in mammalian development: a parental tug-of-war. *Trends Genet* **7**: 45-49.

Morison IM, Ramsay JP, Spencer HG (2005) A census of mammalian imprinting. *Trends Genet* **21**: 457-465.

Muller J, Hart CM, Francis NJ, Vargas ML, Sengupta A, Wild B, Miller EL, O'Connor MB, Kingston RE, Simon JA (2002) Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. *Cell* **111**: 197-208.

Murrell A, Heeson S, Bowden L, Constancia M, Dean W, Kelsey G, Reik W (2001) An intragenic methylated region in the imprinted Igf2 gene augments transcription. *EMBO Rep* **2**: 1101-1106.

Murrell A, Heeson S, Reik W (2004) Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. *Nat Genet* **36**: 889-893.

Murrell A, Ito Y, Verde G, Huddleston J, Woodfine K, Silengo MC, Spreafico F, Perotti D, De Crescenzo A, Sparago A, Cerrato F, Riccio A (2008) Distinct methylation changes at the IGF2-H19 locus in congenital growth disorders and cancer. *PLoS ONE* **3**: e1849.

Nur U (1967) Reversal of heterochromatization and the activity of the paternal chromosome set in the male mealy bug. *Genetics* **56**: 375-389.

Ohlsson R, Paldi A, Graves JA (2001) Did genomic imprinting and X chromosome inactivation arise from stochastic expression? *Trends Genet* **17**: 136-141.

Okamura K, Ito T (2006) Lessons from comparative analysis of species-specific imprinted genes. *Cytogenet Genome Res* **113**: 159-164.

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

Peters J, Beechey C (2004) Identification and characterisation of imprinted genes in the mouse. *Brief Funct Genomic Proteomic* **2**: 320-333.

Plagge A, Isles AR, Gordon E, Humby T, Dean W, Gritsch S, Fischer-Colbrie R, Wilkinson LS, Kelsey G (2005) Imprinted Nesp55 influences behavioral reactivity to novel environments. *Mol Cell Biol* **25**: 3019-3026.

Rieffel SM, Crouse HV (1966) The elimination and differentiation of chromosomes in the germ line of sciara. *Chromosoma* **19**: 231-276.

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.

Scarbrough K, Hattman S, Nur U (1984) Relationship of DNA methylation level to the presence of heterochromatin in mealybugs. *Mol Cell Biol* **4**: 599-603.

Schoenfelder S, Smits G, Fraser P, Reik W, Paro R (2007) Non-coding transcripts in the H19 imprinting control region mediate gene silencing in transgenic Drosophila. *EMBO Rep*

Schotta G, Ebert A, Krauss V, Fischer A, Hoffmann J, Rea S, Jenuwein T, Dorn R, 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, Lachner M, Sarma K, Ebert A, Sengupta R, Reuter G, Reinberg D, Jenuwein T (2004) A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev* **18**: 1251-1262.

Scott RJ, Spielman M (2006) Deeper into the maize: new insights into genomic imprinting in plants. *Bioessays* **28**: 1167-1171.

Simon JA, Tamkun JW (2002) Programming off and on states in chromatin: mechanisms of Polycomb and trithorax group complexes. *Current Opinion in Genetics & Development* **12**: 210-218.

Soppe WJ, Jacobsen SE, Alonso-Blanco C, Jackson JP, Kakutani T, Koornneef M, Peeters AJ (2000) The late flowering phenotype of fwa mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. *Mol Cell* **6**: 791-802.

Srivastava M, Hsieh S, Grinberg A, Williams-Simons L, Huang SP, Pfeifer K (2000) H19 and Igf2 monoallelic expression is regulated in two distinct ways by a shared cis acting regulatory region upstream of H19. *Genes Dev* **14**: 1186-1195.

Suzuki S, Renfree MB, Pask AJ, Shaw G, Kobayashi S, Kohda T, Kaneko-Ishino T, Ishino F (2005) Genomic imprinting of IGF2, p57(KIP2) and PEG1/MEST in a marsupial, the tammar wallaby. *Mech Dev* **122**: 213-222.

Szabo PE, Tang SH, Silva FJ, Tsark WM, Mann JR (2004) Role of CTCF binding sites in the Igf2/H19 imprinting control region. *Mol Cell Biol* **24**: 4791-4800.

Takagi N, Sasaki M (1975) Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature* **256**: 640-642.

Thorvaldsen JL, Duran KL, Bartolomei MS (1998) Deletion of the H19 differentially methylated domain results in loss of imprinted expression of H19 and Igf2. *Genes Dev* **12**: 3693-3702.

Tiwari S, Schulz R, Ikeda Y, Dytham L, Bravo J, Mathers L, Spielman M, Guzman P, Oakey RJ, Kinoshita T, Scott RJ (2008) MATERNALLY EXPRESSED PAB C-TERMINAL, a Novel Imprinted Gene in Arabidopsis, Encodes the Conserved C-Terminal Domain of Polyadenylate Binding Proteins. *Plant Cell* **20**: 2387-2398.

Tremblay KD, Duran KL, Bartolomei MS (1997) A 5' 2-kilobase-pair region of the imprinted mouse H19 gene exhibits exclusive paternal methylation throughout development. *Mol Cell Biol* **17**: 4322-4329.

Tremblay KD, Saam JR, Ingram RS, Tilghman SM, Bartolomei MS (1995) A paternal-specific methylation imprint marks the alleles of the mouse H19 gene. *Nature Genetics* **9**: 407-413.

Umlauf D, Goto Y, Cao R, Cerqueira F, Wagschal A, Zhang Y, Feil R (2004) Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nature Genetics* **36**: 1296-1300.

Vaissiere T, Sawan C, Herceg Z (2008) Epigenetic interplay between histone modifications and DNA methylation in gene silencing. *Mutat Res* **659**: 40-48.

VandeBerg JL, Robinson ES, Samollow PB, Johnston PG (1987) X-linked gene expression and X-chromosome inactivation: marsupials, mouse, and man compared. *Isozymes Curr Top Biol Med Res* **15**: 225-253.

Varmuza S, Mann M (1994) Genomic imprinting--defusing the ovarian time bomb. *Trends Genet* **10**: 118-123.

Verona RI, Thorvaldsen JL, Reese KJ, Bartolomei MS (2008) The transcriptional status but not the imprinting control region determines allele-specific histone modifications at the imprinted H19 locus. *Mol Cell Biol* **28**: 71-82.

Vielle-Calzada JP, Baskar R, Grossniklaus U (2000) Delayed activation of the paternal genome during seed development. *Nature* **404**: 91-94.

Wakefield MJ, Keohane AM, Turner BM, Graves JA (1997) Histone underacetylation is an ancient component of mammalian X chromosome inactivation. *Proc Natl Acad Sci U S A* **94**: 9665-9668.

Wood AJ, Oakey RJ (2006) Genomic imprinting in mammals: emerging themes and established theories. *PLoS Genet* **2**: e147.

Xiao W, Gehring M, Choi Y, Margossian L, Pu H, Harada JJ, Goldberg RB, Pennell RI, Fischer RL (2003) Imprinting of the MEA Polycomb gene is controlled by antagonism between MET1 methyltransferase and DME glycosylase. *Dev Cell* **5**: 891-901.

Yoon YS, Jeong S, Rong Q, Park KY, Chung JH, Pfeifer K (2007) Analysis of the H19ICR insulator. *Mol Cell Biol* **27**: 3499-3510.

Zaratiegui M, Irvine DV, Martienssen RA (2007) Noncoding RNAs and gene silencing. *Cell* **128**: 763-776.

6.8 TRANSITION TO CHAPTER 7

The work presented in this chapter includes microarray gene expression analysis that was undertaken to search for candidate imprinted genes in *Drosophila*. A complete data set highlighting gene expression differences between gynogenetic and sexually produced *Drosophila* is included. This analysis has identified several interesting potentially imprinted genes. Preliminary microarray analysis of an androgenetic fly, and flies with uniparental inheritance of compound second or third chromosomes, is also included. This data provides a significant step towards identifying endogenously imprinted genes in *Drosophila*.

CHAPTER 7

IDENTIFICATION OF CANDIDATE IMPRINTED GENES IN *DROSOPHILA*

7.1 INTRODUCTION

Genomic imprinting is an epigenetic process in which an allele, chromosome, or chromosomal region, is marked according to the sex of the parent through which it is being transmitted. Imprinting can lead to epigenetic silencing of a gene or gene cluster, or chromosomal loss or heterochromatinization, based strictly on the gene or chromosome's pattern of inheritance. The process of imprinting has been observed in a wide range of species, including plants (Alleman and Doctor, 2000; Scott and Spielman, 2006), nematodes (Bean *et al.*, 2004), insects (Khosla *et al.*, 2006; Lloyd, 2000), fish (Martin and McGowan, 1995) and mammals (Morison *et al.*, 2005; Wood and Oakey, 2006).

The existence of imprinted genes in mammals first became evident in 1984, when two research groups demonstrated that gynogenetic and androgenetic mice are not viable, and thus the maternal and paternal genomes are non-equivalent (McGrath and Solter, 1984; Surani *et al.*, 1984). Gynogenetic and androgenetic animals are genetically diploid, but the genetic material in gynogenetic or parthenogenetic individuals is entirely maternal in origin, while in androgenetic individuals it is paternal in origin. Genomic imprinting is believed to be the primary barrier to parthenogenesis in mammals, however, the development of viable and fertile parthenogenetic mice was accomplished by using a modified oocyte donor from which the maternally expressed *H19* gene was deleted, and the normally maternally silenced *Igf2* gene was expressed (Kono *et al.*, 2004). The development of parthenogenetic mice by the appropriate expression of only two imprinted genes therefore suggests that most mammalian imprinted genes are not developmentally essential. Consistent with this, as an increasing number of imprinted

genes are described, it has become clear that mammalian imprinted genes have a wide range of molecular functions (Morison *et al.*, 2005).

In *Drosophila*, imprinting has been observed for transgenes inserted into Y-chromosome heterochromatin, and marker genes and transgenes positioned adjacent to various heterochromatic domains on rearranged chromosomes (Lloyd, 2000; Maggert and Golic, 2002). For example, on the *Dp(1;f)LJ9* mini-X chromosome, the *garnet* eye colour gene is juxtaposed to the centric heterochromatin of the X chromosome, and transmission through the paternal germ line results in epigenetic silencing of *garnet*, while transmission through the maternal germ line results in normal *garnet* expression (Lloyd *et al.*, 1999). In addition, paternal-specific chromosome loss has been observed in *Drosophila* in the presence of certain mutations, indicating that the maternal and paternal chromosome sets carry parent-specific marks that can distinguish them in the embryo (Baker, 1975; Fitch *et al.*, 1998). Despite this evidence that *Drosophila* has the capacity to imprint, no endogenous imprinted genes have yet been identified. Both gynogenetic and androgenetic *Drosophila* are viable, indicating that developmentally essential genes are not imprinted. Imprinted genes may thus include late-acting genes, redundant genes, genes with a subtle phenotype or, given the association of *Drosophila* imprinting with heterochromatin, non-coding transcripts.

Here we set out to identify candidate imprinted genes by conducting microarray analysis of gene expression in gynogenetic and androgenetic flies. Gynogenetic offspring contain only maternal DNA, as is the case for parthenogenetic offspring, but they are produced following mating with a male who does not contribute his genome. In *Drosophila*, gynogenetic flies can be obtained by mating *gyn-2*; *gyn-3* mutant females with a male *ms(3)K81* mutant defective in syngamy (Fuyama, 1984). The *ms(3)K81* gene

is hypothesized to encode a sperm-specific protein that is essential for proper mitosis and migration of the male pronucleus (Yasuda *et al.*, 1995), and the *gyn-2*; *gyn-3* mutations induce nondisjunction in the embryo to allow diploidization of haploid maternal nuclei. Androgenetic offspring have been observed in *Drosophila* by mating $\alpha Tub67C^{3}$, ncd^{D} / ncd^{D} mutant females with w^{1118} males (Komma and Endow, 1995). *Non-claret disjunctional* (*ncd*) encodes an oocyte and early embryo kinesin microtubule motor protein (McDonald and Goldstein, 1990; McDonald *et al.*, 1990), while *αTub67C* encodes an early embryo-specific α-tubulin (Kalfayan and Wensink, 1982; Matthews *et al.*, 1989). Following this mating, at a low frequency (reported to be ~1.5%) the female pronucleus is lost and the two male nuclei from the first division fuse to form diploid androgenetic progeny (Komma and Endow, 1995).

The rationale for using this strategy to identify imprinted genes in *Drosophila* is based on the premise that imprinted genes should be oppositely expressed in gynogenetic and androgenetic flies. For example, a paternally imprinted gene, in which the paternally inherited allele is silenced and the maternally inherited allele is expressed, should be upregulated in gynogenetic flies, which have two maternal alleles, compared with genetically identical control flies which have the normal complement of one maternal allele and one paternal allele. Similarly, this gene would be downregulated in androgenetic flies, which have two paternally inherited alleles, as compared with genetically identical control flies. The opposite would hold true for maternally imprinted genes, which are maternally silenced and paternally expressed.

7.2 MATERIALS AND METHODS

7.2.1 FLY CARE

All fly stocks were maintained at 21 ±3°C on a standard cornmeal, yeast, and sugar medium supplemented with 0.15% methylbenzoate (Sigma) as a mould inhibitor. All crosses were conducted in vials with 5 – 10 females and 3 – 5 males. Crosses were transferred to fresh food up to three times after 3-5 days of egg-laying. All standard stocks are described in FlyBase (Tweedie et al., 2009) and were provided by the Bloomington *Drosophila* Stock Center.

7.2.2 GENERATION OF GYNOGENETIC AND ANDROGENETIC FLIES

Gynogenetic females were obtained by mating w^1/w^1; $gyn\text{-}2^1$; $gyn\text{-}3^1$ females (FBst0005353) to +/Y; $ms(3)K81^1$ / $TM3, Sb^1, Ser^1$ males (FBst0005352). Control, sexually-produced, w^1/w^1; $gyn\text{-}2^1$; $gyn\text{-}3^1$ females were isolated from w^1/w^1; $gyn\text{-}2^1$; $gyn\text{-}3^1$ females crossed with w^1/Y; $gyn\text{-}2^1$; $gyn\text{-}3^1$ males.

To obtain androgenetic females, st^1, ncd^D females (FBst0002243) were mated to *Dp(1;Y)BS*; h^1, $\alpha Tub67C^3$, st^1, ncd^D / *TM3*, males (FBst0002245). h^1, $\alpha Tub67C^3$, st^1, ncd^D / st^1, ncd^D females were selected from this cross, crossed to w^{1118} males, and w^{1118} androgenetic females were selected. However, using this strategy, only a single androgenetic fly was obtained from several months of crosses (a minimum of 20,000 screened progeny).

7.2.3 MICROARRAY HYBRIDIZATION

RNA was extracted from 50 – 100 gynogenetic and control females, isolated as virgins and aged approximately three days, using Trizol reagent (Invitrogen) according to manufacturer's instructions. The RNA samples were shipped to the Canadian *Drosophila* Microarray Centre, where they were fluorescently labelled with different fluorophores during reverse-transcription, combined, and then hybridized to cDNA-based microarrays (as described by Neal *et al.*, 2003). A total of six microarrays were used in gynogenetic analysis, including one 7Kv3 array (estimated to represent approximately 5,500 genes) and five 12Kv1 arrays (estimated to represent approximately 10,500 genes). For all arrays except one, gynogenetic RNA was labelled with Cy5 and control RNA was labelled with Cy3. One 12Kv1 array was conducted with a "dye-flip", where gynogenetic RNA was labelled with Cy3 and control RNA was labelled with Cy5.

RNA was similarly extracted from the single androgenetic fly and a single w^{1118} control fly, and shipped to the Canadian *Drosophila* Microarray Centre where it was amplified using a MessageAmp II amplification kit (Ambion) and hybridized to a 12K array. Androgenetic RNA was labelled with Cy5 and control RNA was labelled with Cy3. Given the lack of biological replicates, the androgenetic data was excluded from additional analysis. However, preliminary gene lists of up- and down-regulated genes identified on the first array are presented in supplementary Table 7.S1.

7.2.4 DATA ANALYSIS

All data analysis was conducted using GeneTraffic software. Data points with a spot to background intensity ratio of less than 2, a spot intensity less than 2 times the

average background, or a spot intensity less than 128 were flagged and excluded from analysis. Arrays were normalized using the Lowess sub-grid method.

Differentially expressed genes in gynogenetic flies were selected using the following filter parameters: number of valid spots greater than or equal to six, coefficient of variance less than or equal to one, and mean Log2 ratio greater than or equal to 0.5 for upregulated genes, or less than or equal to -0.5 for downregulated genes (approximately 1.4 fold increase or decrease). A control gene list consisting of genes that were unchanged in gynogenetic flies was generated using the same search criteria, but Log2 ratios between -0.15 and 0.15 (less than 1.1 fold change).

Upregulated and downregulated gene lists were uploaded into the PANTHER Classification System batch ID search tool (Thomas *et al.*, 2003, http://panther6.ai.sri.com/genes/batchIdSearch.jsp) to identify trends in the biological processes and molecular functions of the identified genes. Gene functions or processes listed as "unknown" on PANTHER were manually checked on FlyBase (Tweedie et al., 2009), and classified into a process or function category, if this information was available.

7.3 RESULTS

Microarray analysis of gene expression in gynogenetic females identified a total of 37 genes or cDNA clones upregulated in gynogenetic flies (Table 7.1) and 42 downregulated genes or cDNA clones (Table 7.2), relative to genotypically identical sexually produced *gyn-2*; *gyn-3* females. A control gene list of 43 genes or cDNA clones that were unchanged in gynogenetic flies is presented in Table 7.3 for comparison. Many mean Log2 ratios had high standard deviations and coefficient of variances, which may

indicate that the data are of relatively weak strength (not shown). Both upregulated and downregulated genes were distributed across all chromosome arms, with the exception of the small heterochromatic chromosome 4, which contained no identified downregulated genes (Table 7.4). While the upregulated genes were distributed relatively evenly across the chromosomes, more downregulated genes were found on chromosome arms 2R and 3R than the other chromosomal regions (Table 7.4). The unchanged gene list contained genes with a relatively even distribution across all chromosome arms (Table 7.4).

A wide range of molecular functions were identified for the genes from the upregulated and downregulated gene lists, as well as from the unchanged control gene list (Table 7.5). The molecular functions of many genes are still unknown, which is reflected in the relatively high number of this class of gene in all categories. However, genes with unknown molecular functions were especially predominant on the downregulated gene list (17/38 = 44.7%), compared with both the upregulated gene list (5/48 = 10.4%) and the unchanged gene list (3/49 = 6.1%). On the upregulated gene list, proteins with nucleic acid binding (12.5%), hydrolase (8.3%), or transcription factor activity (8.3%) were predominant. Similarly, nucleic acid binding (7.9%) and transferase activity (7.9%) were the most frequent molecular functions identified on the downregulated gene list (Table 7.5). On the unchanged gene list, proteins with select regulatory (12.2%), protein binding (10.2%), protease (8.2%), or transcription factor activity (8.2%) were most frequently observed (Table 7.5).

In addition to classifying genes according to molecular function, genes may also be categorized according to their associated biological processes, which include broad biological roles such as pathways, cellular activities and physiological functions. Trends in the biological processes identified from the upregulated gene list include nucleic acid

metabolism (15.5%), signal transduction (12.1%), developmental processes (12.1%) and protein metabolism and modification (8.6%; Table 7.6). The downregulated genes are mostly from unknown biological processes (31.8%), but also include proteins that participate in immunity and defense (11.4%), signal transduction (11.4%), protein metabolism and modification (9.1%), and transport (9.1%; Table 7.6). Frequent biological processes on the unchanged gene list include protein metabolism and modification (14.5%), developmental processes (10.1%), intracellular protein traffic (8.7%), signal transduction (8.7%), and transport (8.7%; Table 7.6).

Table 7.1 Genes or clones identified as upregulated in gynogenetic *Drosophila* using microarray analysis.

Gene or Clone ID	Gene Name or Clone Notes	Chromo-some	Cytoband	# Valid Spots	Mean Log2 Ratio	Molecular functions	Biological Processes
CG10143	Adenosine deaminase-related growth factor E	2L	51B7	6	**1.52**	Adenosine deaminase activity; growth factor activity; deaminase activity.	Purine ribonucleoside monophosphate biosynthetic process.
RH01124	*Contains transposon sequences*	NA	NA	6	**1.48**	NA	NA
RH35946	*Contains transposon sequences*	NA	NA	6	**1.16**	NA	NA
CG11129	Yolk protein 3	X	12C1	10	**1.15**	Structural molecule activity; ATP binding; catalytic activity.	Sex differentiation; vitellogenesis; embryonic development; lipid metabolic process.
RH08140	Computed gene = CG11129, Yolk protein 3	X	12C1	10	**1.07**	(Listed above)	(Listed above)
RH36841	Clone contains similarity to CG6469, Cuticular protein 66D and	3L	66D9	10	**1.04**	Structural constituent of chitin-based larval cuticle; structural constituent of chitin-based cuticle.	Unknown.
	CG2238, Elongation factor 2b	2L	39E7			Translation elongation factor activity; GTP binding; GTPase activity.	Translational elongation; translation; mitotic spindle elongation; mitotic spindle organization.
CG32016	NA	4	102F3	12	**0.83**	Unknown.	Unknown.
CG6806	Larval serum protein 2	3L	68F5	8	**0.83**	Unknown.	Unknown.
GH19165	*Contains transposon sequences*	NA	NA	6	**0.81**	NA	NA
CG11064	Retinoid- and fatty-acid binding glycoprotein	4	102F8	22	**0.80**	Fatty acid binding; retinoid binding; microtubule binding; heme binding; retinol binding; structural molecule activity; lipid transporter activity; receptor binding.	Transport; Wnt receptor signaling pathway; smoothened signaling pathway; lipid transport.
CG2985	Yolk protein 1	X	9A5	10	**0.80**	Structural molecule activity; catalytic activity.	Sex differentiation; vitellogenesis; lipid metabolic process
CG12410	crossveinless	X	5A13	10	**0.77**	Unknown.	Unknown.
CG5711	Arrestin 1	2L	36E3	10	**0.71**	Metarhodopsin binding.	Deactivation of rhodopsin mediated signaling; metarhodopsin inactivation; phototransduction; endocytosis; photoreceptor cell maintenance.
CG17580	NA	2R	49D3	7	**0.67**	Unknown.	Unknown.

Table 7.1 (continued, 2 of 3)

Gene or Clone ID	Gene Name or Clone Notes	Chromo-some	Cytoband	# Valid Spots	Mean Log2 Ratio	Molecular functions	Biological Processes
CG9539	Sec61alpha	2L	26D7	12	**0.65**	Protein transporter activity; P-P-bond-hydrolysis-driven protein transmembrane transporter activity.	Cell death; SRP-dependent cotranslational protein targeting to membrane, translocation; cuticle development; dorsal closure; head involution.
CG10483	NA	3L	64F5	6	**0.64**	Unknown.	Unknown.
CG1895	Cyp28c1	X	10F1	7	**0.63**	Electron carrier activity; heme binding; cysteine-type endopeptidase activity; monooxygenase activity.	Proteolysis.
CG3680	NA	3L	77E8	8	**0.63**	Catalytic activity.	Metabolic process.
RH53853	Computed gene = CG1793, Mediator complex subunit 26	4	102E1	10	**0.62**	RNA polymerase II transcription mediator activity; transcription factor activity; DNA binding; protein binding; transcription regulator activity.	Transcription initiation from RNA polymerase II promoter; phagocytosis, engulfment; transcription.
CG6349	DNA polymerase Alpha 180kD	2L	93F2	6	**0.60**	DNA-directed DNA polymerase activity; 3'-5'-exodeoxyribonuclease activity; DNA binding; nucleoside binding; nucleotide binding.	DNA strand elongation during DNA replication; DNA replication.
CG9285	Dipeptidase B	3R	87F13	12	**0.57**	Dipeptidyl-peptidase activity; tripeptidyl-peptidase activity; aminopeptidase activity.	Proteolysis.
LD04986	Computed gene = CG17704, Nipped-B	2R	41B3	12	**0.56**	Transcription activator activity; binding.	Mitotic sister chromatid cohesion; regulation of transcription, DNA-dependent.
CG4264	Heat shock protein cognate 4	3R	88E4	12	**0.55**	ATPase activity; unfolded protein binding; chaperone binding; ATP binding.	Embryonic development via the syncytial blastoderm; axon guidance; axonal fasciculation; nervous system development; neurotransmitter secretion; vesicle-mediated transport; protein folding; RNA interference.
CG5083	Retinoblastoma-family protein 2	3R	89A8	8	**0.55**	DNA binding; promoter binding.	Negative regulation of transcription from RNA polymerase II promoter; regulation of cell cycle.
CG5354	pineapple eye	2L	31E1	10	**0.54**	Electron carrier activity; protein binding; zinc ion binding.	Compound eye development.
CG7415	Dipeptidyl aminopeptidase III	3R	84F13	12	**0.54**	Dipeptidyl-peptidase activity.	Proteolysis.

Table 7.1 (continued, 3 of 3)

Gene or Clone ID	Gene Name or Clone Notes	Chromo-some	Cytoband	# Valid Spots	Mean Log2 Ratio	Molecular functions	Biological Processes
CG8390	vulcan	2R	41F9	12	**0.54**	Unknown.	Imaginal disc-derived leg morphogenesis; cell-cell signaling.
CG8759	Nascent polypeptide associated complex protein alpha subunit	2R	49C2	12	**0.54**	Protein binding.	Regulation of pole plasm oskar mRNA localization; oogenesis.
CG10160	Ecdysone-inducible gene L3	3L	65A11	6	**0.53**	L-lactate dehydrogenase activity; binding.	Glycolysis; anaerobic glycolysis.
CG1389	torso	2R	43E11	8	**0.53**	Transmembrane receptor protein tyrosine kinase activity; protein tyrosine kinase activity; ATP binding.	Terminal region determination; protein amino acid phosphorylation; gastrulation; pole cell migration; chorion-containing eggshell pattern formation; cell fate determination; determination of anterior/posterior axis, embryo; torso signaling pathway; Malpighian tubule morphogenesis.
CG32251	NA	3L	64B1	10	**0.52**	Unknown.	Unknown.
CG14641	NA	3R	82A1	12	**0.51**	mRNA binding; zinc ion binding; nucleic acid binding; nucleotide binding.	Regulation of alternative nuclear mRNA splicing, via spliceosome.
CG17369	Vacuolar H+-ATPase 55kD B subunit	3R	87C2	12	**0.51**	Hydrogen-exporting ATPase activity, phosphorylative mechanism; proton-transporting ATPase activity, rotational mechanism; hydrogen ion transporting ATP synthase activity, rotational mechanism.	Proton transport; ATP synthesis coupled proton transport.
CG2939	sloppy paired 2	2L	24C7	9	**0.51**	RNA polymerase II transcription factor activity; transcription factor activity; sequence-specific DNA binding.	Regulation of transcription; periodic partitioning; regulation of transcription, DNA-dependent.
CG6187	RluA-2	2L	31F1	10	**0.51**	DRAP deaminase activity; pseudouridine synthase activity; RNA binding.	Pseudouridine synthesis.
CG4600	yippee interacting protein 2	2L	30E4	12	**0.50**	Acetyl-CoA C-acyltransferase activity.	Fatty acid beta-oxidation.

Table 7.2 Genes or clones identified as downregulated in gynogenetic *Drosophila* using microarray analysis.

Gene or Clone ID	Gene Name or Clone Notes	Chromo-some	Cytoband	# Valid Spots	Mean Log2 Ratio	Molecular functions	Biological Processes
CG11671	NA	3R	84F8	8	**-1.18**	Unknown.	Unknown.
CG15043	NA	X	17B5	9	**-0.79**	Unknown.	Unknown.
CG8175	Metchnikowin	2R	52A1	7	**-0.79**	Unknown.	Defense response to Gram-positive bacterium; defense response to fungus; defense response; antifungal humoral response; antibacterial humoral response; defense response to Gram-negative bacterium.
CG13335	NA	2R	50B2	20	**-0.75**	Unknown.	Unknown.
CG9568	NA	2L	29F7	10	**-0.75**	Unknown.	Unknown.
CG5413	Cellular Repressor of E1A-stimulated Genes	3R	90A5	10	**-0.7**	Protein binding; transcription repressor activity.	Negative regulation of transcription.
LP12092	Clone contains similarity to CG2043, larval cuticle protein 3 and	2R	44C6	10	**-0.69**	Structural constituent of chitin-based larval cuticle; structural constituent of chitin-based cuticle.	Unknown.
	CG16836	2R	55C4			Unknown.	Unknown.
CG15066	Immune induced molecule 23	2R	55C4	6	**-0.68**	Unknown.	Toll signaling pathway; antibacterial humoral response; defense response.
CG5732	Gld2	3R	94A3	6	**-0.68**	Polynucleotide adenylyltransferase activity.	Long-term memory.
CG18609	NA	2R	55E11	8	**-0.64**	Unknown.	Unknown.
CG31483	PFTAIRE-interacting factor 2	3R	83F1	6	**-0.63**	Protein binding.	Unknown.
CG10816	Drosocin	2R	51C1	6	**-0.62**	Unknown.	Defense response to Gram-negative bacterium; defense response; defense response to Gram-positive bacterium; defense response to bacterium; antibacterial humoral response.
CG1315	NA	3R	84A1	10	**-0.62**	Argininosuccinate synthase activity; ATP binding.	Argininosuccinate metabolic process; arginine biosynthetic process; urea cycle.
CG7929	ocnus	3R	99D3	12	**-0.62**	Unknown.	Unknown.
CG16704	NA	2L	24B1	10	**-0.61**	Serine-type endopeptidase inhibitor activity; transporter activity.	Transport.
CG17190	NA	3R	92C1	6	**-0.61**	Unknown.	Unknown.

Table 7.2 (continued, 2 of 3)

Gene or Clone ID	Gene Name or Clone Notes	Chromo-some	Cytoband	# Valid Spots	Mean Log2 Ratio	Molecular functions	Biological Processes
RE31004	Clone may correspond to CR41111 (Non-protein coding gene)	NA	Unknown	10	**-0.61**	Unknown.	Unknown.
CG17321	NA	2L	37B1	6	**-0.59**	Unknown.	Unknown.
CG31386	NA	3R	86D9	6	**-0.59**	Unknown.	Unknown.
CG13061	Neuropeptide-like precursor 3	3L	72E1	10	**-0.58**	Neuropeptide hormone activity.	Neuropeptide signaling pathway.
CG14930	NA	2L	32F2	6	**-0.58**	DNA binding.	Unknown.
CG4000	NA	3R	92F6	10	**-0.57**	Unknown.	Unknown.
CG7593	NA	3R	99B9	10	**-0.57**	N-acetyltransferase activity.	Phagocytosis, engulfment; metabolic process.
RH23262	Clone may correspond to CG41623, Ubiquinol-cytochrome C reductase complex 11 kDa protein	3L	Het	10	**-0.57**	Ubiquinol-cytochrome-c reductase activity.	Mitochondrial electron transport, ubiquinol to cytochrome c.
CG13499	NA	2R	58B1	6	**-0.56**	Unknown.	Unknown.
CG31758	Phosphodiesterase 1c	2L	33A8	10	**-0.56**	Calcium- and calmodulin-regulated 3',5'-cyclic-GMP phosphodiesterase activity; 3',5'-cyclic-AMP phosphodiesterase activity; 3',5'-cyclic-GMP phosphodiesterase activity; calmodulin-dependent cyclic-nucleotide phosphodiesterase activity; 3',5'-cyclic-nucleotide phosphodiesterase activity; serine-type endopeptidase inhibitor activity.	Cyclic nucleotide metabolic process; regulation of proteolysis; signal transduction.
CG6640	NA	3L	67D8	9	**-0.56**	Transporter activity.	Transport.
RH39775	Computed gene = CG5596, Myosin alkali light chain 1	3R	98A14	10	**-0.56**	Microfilament motor activity; ATPase activity, coupled; calcium ion binding; cysteine-type endopeptidase activity.	Muscle contraction; mesoderm development; proteolysis.
CG5896	Gram-positive Specific Serine protease	3R	97E5	6	**-0.55**	Serine-type endopeptidase activity.	Proteolysis; defense response to Gram-positive bacterium; innate immune response; positive regulation of Toll signaling pathway
CG9120	Lysozyme X	3L	61F3	10	**-0.55**	Lysozyme activity.	Defense response; antimicrobial humoral response; cell wall catabolic process.

Table 7.2 (continued, 3 of 3)

Gene or Clone ID	Gene Name or Clone Notes	Chromo-some	Cytoband	# Valid Spots	Mean Log2 Ratio	Molecular functions	Biological Processes
LP07070	Clone contains similarity to CG17146, adenylate kinase-1 and	3L	69A2	10	**-0.55**	Adenylate kinase activity; uridine kinase activity; ATP binding.	ATP metabolic process.
	CG30425, Ribosomal protein L41	2R	60E5			Structural constituent of ribosome	Translation.
CG14872	NA	3R	88F7	12	**-0.54**	Transporter activity; binding.	Transport.
CG17420	Ribosomal protein L15	3L	80F (Het.)	10	**-0.53**	Structural constituent of ribosome.	Translation.
CG32576	NA	X	14B7	10	**-0.53**	Unknown.	Vesicle-mediated transport.
CG3454	Histidine decarboxylase	2R	46F7	10	**-0.53**	Histidine decarboxylase activity; pyridoxal phosphate binding.	Compound eye photoreceptor development; cellular amino acid and derivative metabolic process; carboxylic acid metabolic process.
CG4759	Ribosomal protein L27	3R	96E9	10	**-0.53**	Structural constituent of ribosome.	Translation; mitotic spindle elongation; mitotic spindle organization.
CG7875	transient receptor potential	3R	99C6	6	**-0.53**	Calcium channel activity; calmodulin binding; intracellular ligand-gated calcium channel activity; store-operated calcium channel activity; light-activated voltage-gated calcium channel activity; protein binding.	Calcium ion transport; response to light stimulus; light-induced release of internally sequestered calcium ion; calcium-mediated signaling; phototransduction; inhibition of phospholipase C activity; detection of light stimulus involved in sensory perception; protein localization; olfactory learning; detection of light stimulus involved in visual perception.
RH07289	Computed gene = CG18279, Immune induced molecule 10	2R	50A5	10	**-0.52**	Unknown.	Toll signaling pathway; antibacterial humoral response; defense response.
RH09134	Computed gene = CG9691	X	9A3	10	**-0.52**	Unknown.	Unknown.
CG31676	NA	2L	38F2	6	**-0.51**	Unknown.	Unknown.

Table 7.3 Genes or clones identified as having no change in gynogenetic *Drosophila* using microarray analysis.

Gene or Clone ID	Gene Name or Clone Notes	Chromo-some	Cytoband	# Valid Spots	Mean Log2 Ratio	Molecular functions	Biological Processes
CG13503	Verprolin 1	2R	58B8	10	**-0.14**	Actin filament binding.p^;`-	Regulation of cell shape; myoblast fusion; actin filament organization; positive regulation of actin filament polymerization
CG14437	COQ7	X	6C13	6	**-0.14**	Transition metal ion binding; oxidoreductase activity.	Oxidation reduction; ubiquinone biosynthetic process; protein metabolic process
CG1900	Rab40	X	11A11	12	**-0.14**	GTPase activity.	Regulation of cell shape.
CG7811	black	2L	34D1	9	**-0.14**	Aspartate 1-decarboxylase activity.	Visual behaviour.
CG10145	M-spondin	2R	51B9	6	**-0.13**	Unknown.	Unknown.
CG17246	Succinate dehydrogenase A	2R	56D3	12	**-0.13**	Succinate dehydrogenase (ubiquinone) activity.	Tricarboxylic acid cycle.
CG2258	NA	X	7D6	8	**-0.13**	Protein binding.	Unknown.
CG4698	Wnt oncogene analog 4	2L	27E7	6	**-0.13**	Signal transducer activity; receptor binding.	Female gonad development; cellular component movement; establishment of imaginal disc-derived wing hair orientation; cell migration; establishment of ommatidial polarity; salivary gland morphogenesis.
GH19628	Computed gene = CG10998, ryder cup	X	19E7	6	**-0.13**	Binding.	Male meiosis.
RH69856	Computed gene = CG14880	3R	86B13	6	**-0.13**	Chitin binding.	Chitin metabolic process.
CG15381	dpr3	2L	22C1	9	**-0.12**	Unknown.	Regulation of calcium ion transport via store-operated calcium channel activity.
CG1698	NA	2R	46B3	8	**-0.12**	Neurotransmitter transporter activity; potassium:amino acid symporter activity.	Neurotransmitter transport.
CG31102	NA	3R	96D1	7	**-0.12**	Unknown.	Unknown.
CG5237	NA	3R	91F12	7	**-0.12**	Protein binding; zinc ion binding.	Unknown.
LD06565	Computed gene = CG5481, leak	2L	22A1	8	**-0.12**	Axon guidance receptor activity; receptor activity.	Salivary gland boundary specification; axon guidance; cardioblast cell fate specification; positive regulation of cell-cell adhesion; mushroom body development.
RE26306	Computed gene = CG12399, Mothers against dpp	2L	23D3	10	**-0.11**	Protein binding; sequence-specific DNA binding; specific transcriptional repressor activity; RNA polymerase II transcription factor activity.	Anatomical structure development; regulation of developmental process; organ morphogenesis; stem cell differentiation; organ development; positive regulation of cellular component biogenesis; regulation of S phase of mitotic cell cycle.

Table 7.3 (continued, 2 of 3)

Gene or Clone ID	Gene Name or Clone Notes	Chromo-some	Cytoband	# Valid Spots	Mean Log2 Ratio	Molecular functions	Biological Processes
CG1827	NA	2R	45F5	8	**-0.1**	N4-(beta-N-acetylglucosaminyl)-L-asparaginase activity.	Unknown.
CG3320	Rab-protein 1	3R	93D2	10	**-0.1**	GTPase activity.	Regulation of cell shape; negative regulation of dendrite morphogenesis; cell adhesion; protein secretion; actin filament organization.
CG6025	Arflike at 72A	3L	72C1	12	**-0.1**	GTP binding.	dsRNA transport.
CG1609	Gcn2	3R	100C3	8	**-0.09**	Elongation factor-2 kinase activity.	Protein amino acid phosphorylation.
RE74025	Computed gene = CG4994, Mitochondrial phosphate carrier protein	3L	70E1	10	**-0.08**	Phosphate transmembrane transporter activity; transmembrane transporter activity.	Wing disc dorsal/ventral pattern formation.
CG1950	NA	X	11A2	6	**0.08**	Ubiquitin thiolesterase activity.	Ubiquitin-dependent protein catabolic process.
CG31722	NA	2L	32B1	6	**0.08**	Voltage-gated anion channel activity.	Anion transport.
CG33084	NA	2L	34D1	8	**0.08**	Unknown.	Regulation of translational fidelity.
CG4207	bonsai	2R	58F3	12	**0.08**	Structural constituent of ribosome.	Regulation of growth.
RH38767	Clone contains similarity to CG34104and	X	9B1	7	**0.09**	GTPase activity.	Small GTP mediated signal transduction.
	CG11556, Rabphilin	X	9E2			Protein transporter activity.	Synaptic vesicle endocytosis and exocytosis.
RE45331	Computed gene = CG17469, MitF	4	102F8	10	**0.1**	Transcription regulator activity.	Regulation of transcription; compound eye morphogenesis.
CG11071	NA	X	12C7	10	**0.11**	Transcription factor activity.	Unknown.
CG4265	Ubiquitin carboxy-terminal hydrolase	2L	22D4	12	**0.11**	Ubiquitin thiolesterase activity.	Protein deubiquitination
CG4670	NA	2R	49F11	7	**0.11**	Flavin-linked sulfhydryl oxidase activity	Oxidation reduction; cell redox homeostasis.
CG15013	dusky-like	3L	64B1	8	**0.12**	structural constituent of chitin-based cuticle.	Unknown.
CG15438	NA	2L	24F2	12	**0.12**	Sodium-dependent phosphate transporter activity.	Transmembrane transport.
CG17216	NA	3R	86E2	8	**0.12**	Protein serine/threonine kinase activity.	Protein amino acid phosphorylation.
CG9115	myotubularin	2L	26B3	8	**0.12**	Protein tyrosine/serine/threonine phosphatase activity.	Mitotic cell cycle; chromosome segregation.

Table 7.3 (continued, 3 of 3)

Gene or Clone ID	Gene Name or Clone Notes	Chromo-some	Cytoband	# Valid Spots	Mean Log2 Ratio	Molecular functions	Biological Processes
CG4654	DP transcription factor	2R	49F10	14	**0.13**	Protein binding; DNA binding.	Positive regulation of gene expression; positive regulation of nurse cell apoptosis; oogenesis; dorsal/ventral axis specification, ovarian follicular epithelium; regulation of cell cycle; cell cycle
CG5021	NA	3L	66E5	6	**0.13**	Unknown.	Unknown.
CG5848	cactus	2L	35F1	12	**0.13**	Protein binding.	Antifungal humoral response; nervous system development; dorsal/ventral axis specification; oogenesis; antimicrobial humoral response; positive regulation of antifungal peptide biosynthetic process; response to DNA damage stimulus; dorsal appendage formation; phagocytosis, engulfment.
CG1404	ran	X	10A8	24	**0.14**	GTP binding.	Regulation of cell shape; cell adhesion; cell cycle; actin filament organization.
CG4585	NA	2R	60A5	12	**0.14**	Phosphotransferase activity, for other substituted phosphate groups.	Phospholipid biosynthetic process.
CG4847	NA	2R	54C1	8	**0.14**	Cysteine-type endopeptidase activity; cysteine-type peptidase activity.	Proteolysis.
CG6276	NA	3R	88F1	12	**0.14**	DNA binding.	Unknown.
CG9393	NA	3R	85D24	12	**0.14**	Unknown.	Protein targetting to mitochondrion.

Table 7.4 Chromosomal distribution of the upregulated and downregulated genes identified in gynogenetic *Drosophila*, as well as the genes from the control unchanged gene list.

Chromosome Arm	Downregulated in gynogenetic flies	Upregulated in gynogenetic flies	Number on unchanged gene list
X	3	5	9
2L	6	9	11
2R	11	5	10
3L	6	6	4
3R	15	6	8
4	0	3	1
Unassigned	1	3	0
Total	**42**	**37**	**43**

Table 7.5 Molecular functions observed for the upregulated, downregulated, and unchanged genes.

Molecular Functions	Number of hits on upregulated gene list	Number of hits on downregulated gene list	Number of hits on unchanged gene list
Calcium binding protein	0	1	0
Chaperone	1	0	0
Cytoskeletal protein	0	1	1
Defense/immunity protein	0	1	1
Extracellular matrix	1	0	1
Hormone activity	0	1	0
Hydrolase	4	1	1
Ion channel	1	1	1
Isomerase	1	0	0
Kinase	1	0	2
Ligase	1	1	1
Lyase	2	1	1
Membrane traffic protein	1	0	0
Nucleic acid binding	6	3	3
Oxidoreductase	3	0	3
Phosphatase	0	0	1
Protease	2	1	4
Protein binding	0	2	5
Receptor	2	0	1
Select regulatory molecule	2	2	6
Signaling molecule	1	0	1
Storage protein	3	0	0
Structural molecule	0	0	1
Synthase and synthetase	3	0	1
Transcription factor	4	0	4
Transfer/carrier protein	0	0	2
Transferase	2	3	1
Transporter	2	2	4
Unknown	5	17	3
Total	**48**	**38**	**49**

Table 7.6 Biological processes observed for the upregulated, downregulated and unchanged genes.

Biological Process	Number of hits on upregulated gene list	Number of hits on downregulated gene list	Number of hits on unchanged gene list
Amino acid metabolism	0	2	3
Apoptosis	0	0	1
Carbohydrate metabolism	3	1	1
Cell adhesion	0	0	1
Cell cycle	3	0	2
Cell proliferation and differentiation	3	0	2
Cell structure and motility	1	0	3
Coenzyme and prosthetic group metabolism	1	0	1
Developmental processes	7	0	7
Electron transport	1	0	2
Immunity and defense	1	5	1
Intracellular protein traffic	2	0	6
Lipid, fatty acid and steroid metabolism	4	1	2
Memory	0	1	0
Muscle contraction	0	1	0
Neuronal activities	0	0	1
Nitrogen metabolism	0	1	0
Nucleoside, nucleotide and nucleic acid metabolism	9	2	4
Oncogenesis	1	0	1
Other metabolism	0	2	0
Phosphate metabolism	0	0	2
Protein metabolism and modification	5	4	10
Sensory perception	2	1	0
Signal transduction	7	5	6
Transport	4	4	6
Unknown	4	14	7
Total	58	44	69

7.4 DISCUSSION

Microarray analysis identified 79 genes that are at least 1.4 fold up- or down-regulated in gynogenetic flies (Tables 7.1 and 7.2). A control gene list containing 43 genes that do not exhibit changed expression in gynogenetic flies was also extracted from the microarray data to use as a reference (Table 7.3). In the absence of androgenetic data, these gene lists cannot be used to reliably identify imprinted genes, as they would also be expected to include genes regulated by imprinted genes, and genes involved in the processes leading to chromosomal non-disjunction that produces the gynogenetic females. However, any imprinted genes represented on the cDNA microarrays should theoretically be included in these lists. The differentially expressed genes identified in Tables 7.1 and 7.2 are therefore a starting point that could be used to select candidate imprinted genes for further investigation, or combined with additional microarray or molecular analysis to identify potentially imprinted genes.

A significant drawback in using microarrays for this analysis is that non-coding transcripts and heterochromatic genes may be underrepresented. An alternative method such as cDNA subtraction may prove valuable in identifying these types of imprinted transcripts. If the generation of androgenetic flies continues to prove difficult, a valid strategy that has undergone preliminary testing in our lab is to conduct reciprocal crosses between stocks carrying compound second or third chromosomes. Compound chromosomes consist of two homologous autosomes joined together (Novitski *et al.*, 1981). By crossing females with compound second (or third) chromosomes, to males carrying differently marked compound second (or third) chromosomes, progeny that have inherited the entire second (or third) complement from either the maternal or paternal

parent can be selected for expression analysis. Using this method, I have completed one microarray replicate for each the second and third compound chromosomes (supplementary Tables 7.S3 – 7.S6). In addition, these crosses yielded evidence that maternally and paternally inherited second chromosomes are not equivalent, possibly due to genomic imprinting. Of approximately 1450 adult females scored from two separate rounds of reciprocal crosses, 71% had inherited the compound second chromosomes from the maternal parent, while only 29% had inherited them from the paternal parent (supplementary Table 7.S7). Additional microarray replicates of these flies to generate a full data set, in combination with cDNA subtraction and the gynogenetic data, should be a significant step towards identifying imprinted genes in *Drosophila.* Confirmation of imprinting can be obtained by RT-PCR of progeny from reciprocal crosses in which the wild-type copy of the gene has been inherited from one parent, and the corresponding chromosomal deficiency has been inherited from the other. Reciprocal crosses between mutant alleles could also be examined for subtle phenotypic effects.

The high level of variability observed in the gynogenetic data may be due to the difficulty in standardizing the experimental conditions over the extended time it took to acquire sufficient flies for microarray analysis. While the experimental and control RNA used in a single hybridization was isolated in tandem, the RNA for the microarray replicates was isolated from crosses conducted at different times over a 20 month period, which may have introduced variation due to the accumulation of modifiers and changes in gene expression levels within the stocks over time. The environmental conditions of the gynogenetic and control crosses were also difficult to control, due to differences in egg-laying and embryonic viability. There may have also been differences in the microarray chips, labelling, and hybridization procedures, over the time-period that the microarray

data was produced (October 2003 - June 2005). Relatively high variance may also be expected for dual-labelled microarrays, due to competition between the two labelled probes during hybridization for the same target sequences spotted on the arrays.

In general, the identified differentially expressed genes are distributed throughout the *Drosophila* genome, without significant clustering at any particular gene region or cytoband (Tables 7.1 – 7.4). Similarly, heterochromatic regions conferring imprinting have been described for all *Drosophila* chromosomes except the fourth (Golic *et al.*, 1998; Haller and Woodruff, 2000; Lloyd, 2000; Maggert and Golic, 2002). The lack of previously described imprinting on the fourth chromosome may be due to the fact that this chromosome is small in size and mostly heterochromatic, and therefore more difficult to study than other chromosomes. The fact that the fourth chromosome has closely associated heterochromatic and euchromatic regions interspersed throughout its length (Sun *et al.*, 2000) makes it a strong candidate for harbouring imprinted genes, as all previously described imprinted domains in *Drosophila* are contained within heterochromatin (Lloyd, 2000; Maggert and Golic, 2002). The observation that three genes on chromosome four were upregulated in gynogenetic flies, but none was downregulated, may indicate that one or more heterochromatic regions on chromosome four confers a paternal imprint (silencing when passed through the paternal germ line).

Two genes identified in Table 7.2 are from known regions of heterochromatin and should be investigated further. These include CG41623, which encodes Ubiquinol-cytochrome C reductase complex 11 kDa protein, is located in 3L heterochromatin and was downregulated, and CG17420 which encodes Ribosomal protein L15, is located in 3L heterochromatin (80F) and was also downregulated in gynogenetic flies (Table 7.2).

Additionally, two clones containing transposon sequences were quite highly upregulated in gynogenetic flies, and a third was more moderately upregulated (Table 7.1). Transposable elements are preferentially clustered in *Drosophila* heterochromatin (Pimpinelli *et al.*, 1995). This result may therefore be reflective of differentially expressed transposable element sequences positioned near a paternal imprint centre located within *Drosophila* heterochromatin.

Two genes located in close proximity to one another were both identified as downregulated in gynogenetic flies. *Immune induced molecule 23* and *CG16836* are located on chromosome 2R at cytoband 55C4, approximately 6 kb apart and in opposite orientations. Both genes exhibited an approximately 1.6 fold reduction in expression compared with sexually produced controls. The molecular function of both genes is currently unknown. As imprinted genes are often found in clusters, these two genes may be additional candidates for further examination.

The identified genes varied widely in their molecular function and biological processes (Tables 7.5 and 7.6). While the molecular function and biological processes of many genes in the *Drosophila* genome have not yet been identified, genes with unknown functions and processes are especially prevalent on the downregulated gene list (Tables 7.5 and 7.6). This is perhaps not surprising, as the lack of prior discovery of any imprinted genes despite the extensive characterization of many mutant alleles and phenotypes, suggests that imprinted genes in *Drosophila* are likely to be rare transcripts or uncharacterized genes without readily observable phenotypic effects. The relatively high number of upregulated genes involved in nucleic acid binding (12.5% vs. 6.1% on the control gene list, Table 7.5) may indicate a trend towards imprinted DNA- or RNA-interacting proteins. In mammals, imprinted genes with a wide range of functions have

been observed. These include genes involved in growth and development, genes with behavioural or cognitive effects, and non-coding RNA transcripts (Davies *et al.*, 2005; Peters and Beechey, 2004; Plagge *et al.*, 2005). Identified imprinted genes in plants include Polycomb group genes, transcription factors, and a gene that encodes the C-terminal domain of poly(A) binding proteins (Jullien and Berger, 2009). Imprinted genes in *Drosophila* are likely to be similarly diverse in function.

7.5 SUPPLEMENTARY MATERIAL

7.5.1 SUPPLEMENTARY MATERIALS AND METHODS

Flies with uniparental inheritance of compound second chromosomes were generated via reciprocal crosses between the stocks *C(2)EN*, b^1, pr^1 (FBst0001112) and *C(2)EN*, bw^1, sp^1 (FBst0001020), with *C(2)EN*, b^1, pr^1 female progeny used in microarray analysis. Flies with uniparental inheritance of compound third chromosomes were generated via reciprocal crosses between stocks *C(3)EN*, th^1, st^1 (FBst0001114) and *C(3)EN*, st^1, cu^1, e^s (FBst0001117), with *C(3)EN*, st^1, cu^1, e^s female progeny used in microarray analysis. RNA isolation and microarray analysis was conducted as previously described in section 7.2. For both compound chromosome microarrays, maternal RNA was labelled with Cy5 and paternal RNA was labelled with Cy3.

7.5.2 SUPPLEMENTARY RESULTS

Table 7.S1 Genes or clones identified as upregulated in androgenetic *Drosophila* using microarray analysis.

Gene or Clone ID	Gene Name or Clone Notes	Chromo-some	Cytoband	# Valid Spots	Mean Log2 Ratio	Molecular functions
LP03565	Contains similarity to: CG31775 and	2L	35B5	2	2.99	unknown
	CG42586	2L	35B5			unknown
CG2979	Yolk protein 2	X	9B1	2	2.96	structural molecule activity
CG32654	Sec16	X	11B2	2	2.82	unknown
CG9538	Antigen 5-related	X	12F4	2	2.71	unknown
CG8197	NA	2R	45A1	2	2.42	protein binding
RH08140	Computed gene = CG11129,Yolk Protein 3	X	12C2	2	2.4	structural molecule activity
LD01252	Computed gene = CG4916, maternal expression at 31B	2L	31B1	2	2.39	RNA binding
CG11129	Yolk protein 3	X	12C2	2	2.38	structural molecule activity
CG9468	NA	2L	29F1	2	2.36	alpha-mannosidase activity
CG18240	NA	2R	47D4	2	2.3	electron carrier activity
CG4466	Heat shock protein 27	3L	67A8	2	2.25	heat shock protein activity
CG11796	NA	3L	77B6	2	2.14	4-hydroxyphenylpyruvate dioxygenase activity
LP05416	Computed gene = CG7916	2L	34D7	2	2.13	unknown
CG11819	NA	3R	96B19	2	2.1	unknown
CG8505	NA	2R	49A2	2	2.09	structural constituent of cuticle
CG31366	Heat-shock-protein-70Aa	3R	87A2	2	2.01	heat shock protein activity
CG6287	NA	2L	32D2	2	1.98	phosphoglycerate dehydrogenase activity
CG3333	Nucleolar protein at 60B	2R	60C1	2	1.97	pseudouridylate synthase activity
CG11064	Retinoid- and fatty-acid binding protein	4	102F2	2	1.93	microtubule binding, structural molecule activity, retinoid and fatty acid binding
CG17012	NA	2L	22D1	2	1.86	serine-type peptidase activity, serine-type endopeptidase activity
CG8994	exuperantia	2R	57A9	2	1.84	unknown
CG7008	NA	3L	61B3	2	1.81	transcription co-activator activity
CG6226	FK506-binding protein 1	3R	88F1	2	1.78	FK506 binding
CG7917	Nucleoplasmin	3R	99C6	2	1.78	nucleoplasmin ATPase activity, histone binding
CG8579	Jonah 44E	2R	44E2	2	1.77	serine-type endopeptidase activity
CG17161	grapes	2L	36A10	2	1.75	protein kinase activity; protein serine/threonine kinase activity.
CG1070	Alhambra	3R	84C1	2	1.74	transcription factor activity, DNA binding
CG33111	NA	3R	95B7	2	1.73	transition metal ion binding, oxidoreductase activity

Table 7.S1 (continued, 2 of 2)

Gene or Clone ID	Gene Name or Clone Notes	Chromo-some	Cytoband	# Valid Spots	Mean Log2 Ratio	Molecular functions
RH08424	Computed gene = CG11413	2R	60D4	2	1.73	unknown
CG11844	vig2	3R	96C1	2	1.7	unknown
CG18180	NA	3L	67C3	2	1.68	serine-type endopeptidase activity
CG7214	NA	2L	28C2	4	1.67	unknown
CG10297	NA	3L	65A6	2	1.66	structural constituent of adult cuticle
CG2050	modulo	3R	100C6	2	1.66	DNA binding, RNA binding
CG13849	Nop56	3R	94A12	2	1.63	unknown
CG7118	NA	3L	66C8	2	1.62	serine-type endopeptidase activity, elastase activity
CG15365	NA	X	8C4	2	1.58	unknown
CG30008	NA	2R	46C2	2	1.58	unknown
CG8129	NA	3R	85D11	2	1.56	threonine dehydratase activity
CG4886	cyclophilin-33	2R	54C3	2	1.53	RNA binding, peptidyl-prolyl cis-trans isomerase activity, cyclophilin-type peptidy-prolyl cis-trans isomerase activity
CG32793	NA	X	3C3	2	1.51	unknown
LD08227	Computed gene = CG3793	2L	35D4	2	1.51	methyltransferase activity
CG14222	NA	X	18D8	2	1.46	acyl-CoA N-acyltransferase activity
CG3902	NA	3L	75E1	2	1.46	short branched-chain acyl-CoA dehydrogenase activity
RE50740	Computed gene = CG6375, pitchoune	3R	94A1	2	1.46	TP-dependent RNA helicase activity, RNA helicase activity
CG4817	Structure specific recognition protein	2R	60A2	2	1.45	DNA secondary structure binding, single-stranded DNA binding, single-stranded RNA binding
CG13841	NA	3R	94C4	2	1.44	unknown
CG13095	NA	2L	29C3	2	1.42	NOT aspartic-type endopeptidase activity
CG4178	Larval serum protein 1 Beta	2L	21D4	2	1.42	nutrient reservoir activity
CG18543	matrimony	3L	66C11	2	1.4	unknown
CG6543	NA	2R	50C6	2	1.4	enoyl-CoA hydratase activity
CG7269	Helicase at 25E	2L	25E6	2	1.39	RNA helicase activity, ATP dependent RNA helicase activity

Table 7.S2 Genes or clones identified as downregulated in androgenetic *Drosophila* using microarray analysis.

Gene or Clone ID	Gene Name or Clone Notes	Chromo-some	Cytoband	# Valid Spots	Mean Log2 Ratio	Molecular functions
CG9686	NA	X	9A3	2	-4.54	unknown
CG3331	ebony	3R	93C2	2	-3.84	beta-alanyl-dopamine synthase activity
CG9307	NA	3R	87F11	2	-3.54	chitinase activity
CG8743	NA	3L	76B11	2	-3.46	calcium channel activity
CG9701	NA	3L	73C2	2	-3.41	cation binding, hydrolase activity, hydrolyzing O-glycosyl compounds
CG31200	NA	3R	92F10	2	-3.35	serine-type endopeptidase activity
CG7778	NA	2L	29B1	2	-3.02	unknown
CG31199	NA	3R	92F10	2	-3	serine-type endopeptidase activity
CG1600	NA	2R	43D2	2	-2.92	oxioreductase activity, zinc ion binding
CG12045	Cuticular protein 100A	3R	100A6	2	-2.86	structural constituent of chitin-based cuticle
CG8091	Nedd2-like caspase	3L	67D2	2	-2.8	endopeptidase activity, cysteine-type endopeptidase activity, protein homodimerization activity
CG17052	NA	X	19B3	2	-2.78	structural constituent of peritrophic membrane
CG4377	NA	2R	58A2	2	-2.77	unknown
CG9441	Punch	2R	57C7	2	-2.76	GTP cyclohydrolase activity, GTP cyclohydrolase I activity
CG7811	black	2L	34D1	2	-2.69	aspartate 1-decarboxylase activity, glutamate decarboxylase activity
CG12818	NA	3R	86A8	2	-2.69	unknown
CG14945	NA	2L	33B11	2	-2.59	glycosylphosphatidylinositol diacylglycerol-lyase activity, phospholipase C activity
RH05209	Computed gene = CG8736	2R	44D4	2	-2.59	unknown
RE72980	Computed gene = CG8696, Larval visceral protein H	2R	44D1	2	-2.56	alpha-glucosidase activity
CG2471	Sclp	X	10F7	2	-2.54	protein binding
SD04061	Computed gene = CG32850	4	102B3	2	-2.39	protein binding, zinc ion binding
CG10550	NA	3R	96D1	2	-2.38	unknown
RH03309	Computed gene = CG40485	X	20F3	2	-2.27	oxioreductase activity
RH64511	Computed gene = CG40485	X	20F3	2	-2.27	oxioreductase activity
CG11765	Peroxiredoxin 2540	2R	47A7	3	-2.24	antioxidant activity, non-selenium glutathione peroxidase activity, peroxidase activity
CG15884	Cuticular protein 97Eb	3R	97E2	2	-2.24	structural constituent of chitin-based cuticle
CG6781	NA	3L	66D1	2	-2.23	glutathione transferase activity
CG10570	NA	2L	37A4	2	-2.18	unknown

Table 7.S2 (continued, 2 of 2)

Gene or Clone ID	Gene Name or Clone Notes	Chromo-some	Cytoband	# Valid Spots	Mean Log2 Ratio	Molecular functions
CG8756	vermiform	3L	76C2	2	-2.14	chitin binding, chitin deacetylase activity
CG4795	Calphotin	3R	87A9	2	-2.1	calcium ion binding
CG8483	NA	3R	87E2	2	-2.09	unknown
CG1561	NA	X	10C7	2	-2.08	unknown
CG33158	NA	3L	72F1	2	-2.03	translation elongation activity
CG3957	wing morphogenesis defect	2R	59E3	2	-1.96	receptor binding
CG32209	serpentine	3L	76C2	2	-1.96	chitin binding, chitin deacetylase activity
CG5618	NA	3L	77B1	2	-1.88	sulfinoalanine decarboxylase activity
CG31764	virus-induced RNA 1	2L	33E3	4	-1.84	unknown
CG32499	Chitin deacetylase-like 4	X	20C3	2	-1.84	chitin binding
CG10912	NA	2R	55B2	2	-1.82	unknown
CG4551	smell impaired 35A	2L	34F4	2	-1.8	protein serine/threonine kinase activity, protein kinase activity, protein amino acid phosphorylation
CG8785	NA	2R	49B10	2	-1.8	amino acid transmembrane transporter activity
CG6891	NA	X	17D4	2	-1.78	actin binding
CG10911	NA	2R	55B2	2	-1.77	unknown
CG9399	NA	3R	85D25	4	-1.76	unknown
CG10562	NA	3R	96D1	2	-1.74	unknown
CG7365	NA	3L	76F2	2	-1.73	phospholipase activity
CG8588	pastrel	3L	65F6	2	-1.69	unknown
CG5646	NA	3R	98A2	2	-1.68	carrier activity
RE41137	Computed gene = CG32850	4	102B3	2	-1.67	protein binding, zinc ion binding
CG10680	NA	2L	38A8	2	-1.66	unknown
CG1803	NA	X	11A2	2	-1.63	calcium ion binding
CG1885	NA	X	8C1	2	-1.63	uroporphyrinogen-III synthase activity

Table 7.S3 Genes or clones identified as upregulated when compound second chromosomes are inherited maternally.

Gene or Clone ID	Gene Name or Clone Notes	Chromo-some	Cytoband	# Valid Spots	Mean Log2 Ratio	Molecular functions
CG15848	Sarcoplasmic calcium-binding protein 1	2R Het	80A	2	1.96	calcium ion binding
RE43488	Computed gene = CG3186, eIF-5A	2R	60B7	2	1.91	translation regulator activity
CG7929	ocnus	3R	99D3	2	1.81	unknown
GM07660	NA	NA	NA	2	1.81	unknown
LP12092	Computed gene = CG2043, Larval cuticle protein 3	2R	44C6	2	1.76	structural constituent of chitin-based cuticle
	and					
	CG16836	2R	55C4			unknown
CG6547	mitochondrial ribosomal protein L2a	3R	86C6	2	1.74	structural constituent of ribosome
CR30425	Ribosomal protein L41	2R	60E5	2	1.73	structural constituent of ribosome
RH61355	Computed gene = CG1821, Ribosomal protein L31	2R	45F5	2	1.68	structural constituent of ribosome
CG8369	NA	3R	85B2	2	1.67	unknown
CG40228	NA	3L	3L Het	2	1.64	unknown
CG13585	NA	2R	60D4	2	1.6	unknown
CG40410	Apoptosis-linked gene-2	3R	3R Het	2	1.6	calcium ion binding
CG5557	squeeze	3R	91F8-91F9	2	1.58	transcription factor activity
CG9568	NA	2L	29F7	2	1.57	unknown
CG4184	NA	2L	21D1	2	1.53	RNA polymerase II transcription mediator activity
CG31921	NA	2L	21C1	2	1.52	unknown
LD29015	Computed gene = CG14464	2R	41C3	2	1.52	unknown
CG4494	NA	2L	27C7	2	1.51	protein binding
CG7170	Jonah 66Cii	3L	66C10	2	1.51	chymotrypsin activity
GH16768	Computed gene = CG9946, eukaryotic translation Initiation Factor 2α	X	14C6	2	1.51	translation initiation factor activity
RH26620	Computed gene = CG30425, Ribosomal protein L41	2R	60E5	2	1.51	structural constituent of ribosome
CG2219	NA	4	102A6	2	1.5	GTPase activity
CG9894	NA	2L	23A3	2	1.49	unknown
LP07070	Computed gene = CG17146, Adenylate kinase-1	3L	69A2	2	1.49	uridine kinase activity
CG1829	Cyp6v1	X	19E1	2	1.47	electron carrier activity

Table 7.S3 (continued, 2 of 2)

Gene or Clone ID	Gene Name or Clone Notes	Chromo-some	Cytoband	# Valid Spots	Mean Log2 Ratio	Molecular functions
CG17420	Ribosomal protein L15	3L	3L Het, 80F	2	1.47	structural constituent of ribosome
CG17996	NA	2L	36A12	2	1.46	unknown
CG12372	spt4	2R	49B10	2	1.46	chromatin binding
CG7085	lethal (2) s5379	2L	22D4	2	1.45	unknown
CG9537	Daxx-like protein	2L	26D8	2	1.44	receptor signaling protein activity
RH56668	Computed gene = CG5163, Transcription-factor-IIA-S	3R	95C8	2	1.44	general RNA polymerase II transcription factor activity
CG1789	NA	3R	8A3	2	1.43	unknown
CG4759	Ribosomal protein L27	3R	96E9	2	1.43	nucleic acid binding, structural constituent of ribosome
CG14482	NA	2R	54C9	2	1.42	ubiquinol-cytochrome-c reductase activity
CG18315	Adenine phosphoribosyltransferase	3L	62B9	2	1.41	adenine phosphoribosyltransferase activity
CG4409	NA	2R	53C4	2	1.4	unknown
CG5323	NA	2R	55E2	2	1.39	unknown
CG32267	NA	3L	63E2	2	1.38	unknown
CG15323	NA	X	19B1	2	1.38	unknown
CG7712	NA	2R	47C6	2	1.36	NADH dehydrogenase activity, NADH dehydrogenase (ubiquinone) activity
CG8044	Heterochromatin protein 4	3L	66A21	2	1.35	unknown
CG13993	NA	2L	26B4	2	1.34	unfolded protein binding
CG1315	NA	3R	84A1	2	1.34	argininosuccinate synthase activity
CG16765	pasilla	3R	85D16	2	1.34	mRNA binding
CG2981	Troponin C at 41C	2R	41D1	2	1.33	calmodulin binding, calcium ion binding
CG18778	Cuticular protein 65Au	3L	65A6	2	1.31	structural constituent of chitin-based cuticle
CG40127	NA	2R	NA	2	1.31	unknown
CG32230	NA	3L	80F2	2	1.3	NADH dehydrogenase activity
CG5219	mitochondrial ribosomal protein L15	3L	77C3	2	1.29	structural constituent of ribosome
CG8808	Pyruvate dehydrogenase kinase	2R	45D5	2	1.29	pyruvate dehydrogenase (acetyl-transferring) kinase activity
CG11912	NA	2L	21C1	2	1.28	serine-type endopeptidase activity

Table 7.S4 Genes or clones identified as downregulated when compound second chromosomes are inherited maternally.

Gene or Clone ID	Gene Name or Clone Notes	Chromo-some	Cytoband	# Valid Spots	Mean Log2 Ratio	Molecular functions
LD33274	Computed gene = CG17484, Adherens junction protein p120	2R	41B1	2	-3.25	binding
CG2207	Decondensation factor 31	2L	39E3	2	-2.59	histone binding
CG2238	Elongation factor 2b	2L	39E7	2	-2.27	translation elongation factor activity
CG11512	Glutathione S transferase D4	3R	87B8	2	-2.25	glutathione transferase activity
CG32016	NA	4	102F3	2	-2.19	unknown
LD04986	Computed gene = CG17704, Nipped-B	2R	41B3	2	-2.19	transcription activator activity
CG10067	Actin 57B	2R	57B5	2	-2.12	structural constituent of cytoskeleton
CG11064	Retinoid- and fatty-acid binding protein	4	102D3	4	-2.06	microtubule binding, heme binding, structural molecule activity, retinoid binding, fatty acid binding
CG18240	NA	2R	47D4	2	-1.99	electron carrier activity
RE72980	Computed gene = CG8696, Larval visceral protein H	2R	44D1	2	-1.99	alpha-glucosidase activity
CG7439	Argonaute 2	3L	71C3	2	-1.98	protein binding, endoribonuclease activity, siRNA binding
CG4600	yippee interacting protein 2	2L	30E4	2	-1.95	Acetyl-CoA C-acyltransferase activity
LD12894	Computed gene = CG4993, PRL-1	2L	35F1	2	-1.93	prenylated protein tyrosine phosphatase activity
RE51330	Computed gene = CG31400, Heat shock RNA ω	3R	93D4	2	-1.92	unknown
CG11154	ATP synthase-β	4	102F6	2	-1.91	hydrogen-exporting ATPase activity, phosphorylative mechanism
CG9998	U2 small nuclear riboprotein auxiliary factor 50	X	14B17	2	-1.89	poly-pyrimidine tract binding, pre-mRNA splicing factor activity, RNA binding, protein heterodimerization activity
CG2682	NA	2R	41F2	2	-1.89	transcription factor activity
CG3572	visceral mesodermal armadillo-repeats	2R	42E1	2	-1.87	Ral GTPase binding
RH36841	Clone contains similarity to CG32029, Cuticular protein 66D	3L	66D9	2	-1.87	structural constituent of chitin-based cuticle, structural constituent of chitin-based larval cuticle
	and					
	CG2238, Elongation factor 2b	2L	39E7			translation elongation factor activity
CG6058	Aldolase	3R	97A6	4	-1.84	fructose-bisphosphate aldolase activity
RE33126	Clone contains similarity to CG10067, Actin 57B	2R	57B5	2	-1.8	structural constituent of cytoskeleton
	and					
	CG17489, Ribosomal protein L5	2L	40F7			protein binding

Table 7.S4 (continued, 2 of 2)

Gene or Clone ID	Gene Name or Clone Notes	Chromo-some	Cytoband	# Valid Spots	Mean Log2 Ratio	Molecular functions
RE38067	Computed gene = CG10067, Actin 57B	2R	57B5	2	-1.79	structural constituent of cytoskeleton
CG8635	NA	2R	44D5	4	-1.77	zinc ion binding, nucleic acid binding
CG10243	NA	2R	51D1	3	-1.76	electron transporter activity
CG8759	Nascent polypeptide associated complex protein alpha subunit	2R	49C2	2	-1.76	protein binding
CG4485	Cytochrome P450-9b1	2R	42F3	2	-1.74	oxidoreductase activity, electron transporter activity
CG3927	NA	2R	58F1	2	-1.73	nucleic acid binding, RNA binding
CG11129	Yolk protein 3	X	12C1	2	-1.71	structural molecule activity
LD18112	Computed gene = CG17486	2R	41B1	2	-1.71	asparagine synthase (glutamine-hydrolyzing) activity
CG8345	Cyp6w1	2R	42A12	2	-1.7	electron carrier activity
RH28847	Computed gene = CG6949, mitochondrial ribosomal protein L45	3R	94B6	2	-1.7	structural constituent of ribosome
CG5277	Intronic Protein 259	2L	31E1	2	-1.67	unknown
CG10911	NA	2R	55B2	2	-1.66	unknown
CG3699	NA	X	1D3	2	-1.65	oxidoreductase activity, acting on CH-OH group of donors
CG9252	deadlock	2L	39A4	2	-1.64	unknown
CG2875	NA	X	3D6	2	-1.64	unknown
CG7668	NA	3L	76E1	2	-1.61	receptor binding
CG3989	NA	X	11B16	2	-1.6	phosphoribosylaminoimidazole carboxylase activity
CG15081	lethal (2) 03709	2R	55F5	2	-1.59	unknown
CG7008	NA	3L	61B3	2	-1.59	transcription coactivator activity
CG31916	Muscle-specific protein 300	2L	25C6	3	-1.59	actin binding
CG7490	Ribosomal protein LP0	3L	79B2	2	-1.58	RNA binding, structural constituent of ribosome
CG1873	Elongation factor 1α100E	3R	100D2	2	-1.56	translation elongation factor activity
CG7878	NA	3R	84E13	2	-1.54	helicase activity, ATP-dependent RNA helicase activity
CG11081	plexin A	4	102D1	2	-1.54	transmembrane receptor protein tyrosine kinase activity, semaphorin receptor activity, axon guidance receptor activity
GH13704	Computed gene = CG34166	2L	35B5	2	-1.52	unknown
CG6339	rad50	2R	58E1	2	-1.51	nuclease activity, zinc ion binding, ATP binding
CG32793	NA	X	3C3	2	-1.5	unknown
CG7415	Dipeptidyl aminopeptidase III	3R	84F13	2	-1.48	dipeptidyl-peptidase III activity
LD03740	Computed gene = CG17484, Adherens junction protein p120	2R	41B1	2	-1.48	binding
CG2229	Jonah 99Fii	3R	99F6	2	-1.47	serine-type endopeptidase activity

Table 7.S5 Genes or clones identified as upregulated when compound third chromosomes are inherited maternally.

Gene or Clone ID	Gene Name or Clone Notes	Chromo-some	Cytoband	# Valid Spots	Mean Log2 Ratio	Molecular functions
CG1718	NA	X	19F3	2	2.32	transporter activity, ATPase activity, coupled to transmembrane movement of substances
CG18234	NA	3L	75D3	2	2.1	oxidoreductase activity
CG7637	NA	2R	47C1	2	2.03	unknown
GM27176	Computed gene = lethal (2) 35Bd	2L	35B8	2	1.54	mRNA (guanine-N7-)-methyltransferase activity; RNA binding
CG10809	NA	3L	67C11	2	1.21	unknown
CG9434	Frost	3R	85E2	2	1.15	unknown
CG11671	NA	3R	84F8	2	0.97	unknown
CG4004	NA	X	11B14	2	0.84	unknown
CG5189	NA	2R	55C9	2	0.73	unknown
CG18764	NA	3R	86E13	2	0.72	zinc ion binding, nucleic acid binding
CG6164	Niemann-Pick type C-2f i	3R	95E1	2	0.67	unknown
CG4377	NA	2R	58A2	2	0.66	unknown
CG9588	NA	3R	87F7	2	0.66	protein binding
CG10910	NA	2R	55B1	2	0.63	unknown
RH23262	Computed gene = CG41623, Ubiquinol-cytochrome C reductase complex 11 kDa protein	3L	3L Het	2	0.63	ubiquinol-cytochrome-c reductase activity
CG8264	Bx42	X	8C9	2	0.62	protein binding
CG13969	brain washing	2L	38B2	2	0.61	ceramidase activity
CG31509	Turandot A	3R	93A2	2	0.59	unknown
CG4646	NA	2R	49F10	2	0.57	unknown
CG14375	NA	3R	87E8	2	0.57	unknown
CG15220	NA	X	10D1	2	0.56	DNA binding
CG15347	NA	X	7E11	2	0.56	unknown
CG8221	Amyrel	2R	53C10	2	0.55	alpha-amylase activity, oligo-1,6-glucosidase activity
CG4109	Syntaxin 8	3L	73A10	2	0.55	SNAP receptor activity
CG11986	NA	3R	85C2	2	0.55	unknown
CG3699	NA	X	1D2	2	0.55	oxidoreductase activity, acting on CH-OH group of donors
CG12919	eiger	2R	46E1	2	0.54	protein binding
CG10075	NA	3L	65D4	2	0.54	unknown
CG5161	NA	3L	72D9	2	0.54	unknown
CG17820	female-specific independent of transformer	3R	93F14	2	0.54	unknown
CG14543	NA	3R	96F10	2	0.54	unknown

Table 7.S5 (continued, 2 of 2)

Gene or Clone ID	Gene Name or Clone Notes	Chromo-some	Cytoband	# Valid Spots	Mean Log2 Ratio	Molecular functions
CG30339	NA	2R	45F6	2	0.53	transporter activity
CG1340	NA	3R	100A3	2	0.53	mRNA binding, translation initiation factor activity
CG1609	Gcn2	3R	100C3	2	0.53	elongation factor-2 kinase activity
CG14989	NA	3L	64A5	2	0.52	unknown
CG4208	XRCC1	X	4F5	2	0.52	damaged DNA binding
CG7562	TBP-related factor	2L	28E1	2	0.51	RNA polymerase III transcription factor activity, transcription factor binding,RNA polymerase II transcription factor activity, DNA binding
CG5676	NA	2L	31B1	2	0.51	unknown
CG3358	NA	2R	42E1	2	0.51	endodeoxyribonuclease activity, producing 5'-phosphomonoesters
CG4140	lethal (2) 35Be	2L	35B8	2	0.5	unknown
CG10531	Cht9	2R	57B16	2	0.5	chitinase activity
CG14637	abstrakt	3R	82A1	2	0.5	ATP-dependent helicase activity; helicase activity
CG5934	NA	3R	97F1	2	0.5	unknown

Table 7.S6 Genes or clones identified as downregulated when compound third chromosomes are inherited maternally.

Gene or Clone ID	Gene Name or Clone Notes	Chromo-some	Cytoband	# Valid Spots	Mean Log2 Ratio	Molecular functions
CG6988	Protein disulfide isomerase	3L	71B5	2	-0.93	protein disulfide isomerase activity
CG17697	frizzled	3L	70D4	2	-0.93	transmembrane receptor activity, Wnt-protein binding
CG17342	Lk6	3R	86E18	2	-0.87	protein kinase activity, kinase activity
CG6050	Elongation factor Tu mitochondrial	2R	50A14	2	-0.8	translation elongation factor activity
CG8219	NA	3L	65A8	2	-0.77	protein transmembrane transporter activity
CG11771	NA	3R	86B10	2	-0.76	metalloendopeptidase activity
CG9244	Aconitase	2L	39A7	2	-0.74	aconitate hydratase activity
RE14563	*contains transposon sequences*			2	-0.74	NA
CG1765	Ecdysone receptor	2R	42A9	2	-0.73	ecdysteroid hormone receptor activity, protein binding, protein heterodimerization activity
CG5436	Heat shock protein 68	3R	95D11	2	-0.71	unfolded protein binding
CG12101	Heat shock protein 60	X	10A4	2	-0.7	unfolded protein binding
CG2081	Vago	X	10A4	2	-0.7	unknown
CG8745	NA	3L	70C5	2	-0.68	alanine-glyoxylate transaminase activity, ornithine-oxo-acid transaminase activity
CG5711	Arrestin 1	2L	36E3	2	-0.68	metarhodopsin binding
CG1372	yolkless	X	12E3	2	-0.67	vitellogenin receptor activity
CG33123	NA	2L	23F6	2	-0.67	leucine-tRNA ligase activity
CG8180	NA	2R	52A4	2	-0.63	unknown
CG7470	NA	3L	79A5	2	-0.63	delta1-pyrroline-5-carboxylate synthetase activity, glutamate 5-kinase activity, glutamate-5-semialdehyde dehydrogenase activity
LP04654	Computed gene = CG42600	2R	45F5	2	-0.62	unknown
CG33060	NA	3L	72E1	2	-0.61	unknown
CG3605	NA	2L	23C4	2	-0.61	unknown
CG8487	gartenzwerg	2R	48F8	2	-0.61	guanyl-nucleotide exchange factor activity
CG9485	NA	2R	57D11	2	-0.61	amylo-alpha-1,6-glucosidase activity, 4-alpha-glucanotransferase activity.
CG11129	Yolk protein 3	X	12C1	2	-0.6	structural molecule activity
CG13124	NA	2L	30E1	2	-0.6	protein binding
CG10363	Thiolester containing protein IV	2L	37F1	2	-0.6	peptidase inhibitor activity
CG6223	β-coatomer protein	X	17A9	2	-0.6	protein binding, structural molecule activity
RH07164	Computed gene = Heat shock protein cognate 4	3R	88E4	2	-0.6	chaperone binding
CG32477	NA	3L	61B2	2	-0.58	unknown
CG32562	xmas-2	X	15E7	2	-0.56	mRNA binding
CG2718	Glutamine synthetase 1	2L	21B2	2	-0.56	glutamate-ammonia ligase activity

Table 7.S6 (continued, 2 of 2)

Gene or Clone ID	Gene Name or Clone Notes	Chromo-some	Cytoband	# Valid Spots	Mean Log2 Ratio	Molecular functions
CG9127	adenosine 2	2L	26B3	2	-0.56	phosphoribosylformylglycinamidine synthase activity
CG2139	aralar 1	3R	99F4	4	-0.56	calcium ion binding, transmembrane transporter activity
CG4264	Heat shock protein cognate 4	3R	88E4	2	-0.56	chaperone binding
CG5004	NA	X	15E3	2	-0.55	unknown
CG7238	septin interacting protein 1	2L	25F3	2	-0.55	nucleic acid binding
CG7897	gp210	2R	41F9	2	-0.55	unknown
CG9805	eIF3-S10	3R	82B1	2	-0.55	translation initiation factor activity
CG7719	NA	3R	91C1	2	-0.55	unknown
CG18255	Stretchin-Mlck	2R	52D4	4	-0.55	myosin light chain kinase activity, calmodulin-dependent protein kinase activity
CG12050	NA	2L	39A1	2	-0.54	unknown
CG1782	Ubiquitin activating enzyme 1	2R	46A1	2	-0.54	ubiquitin activating enzyme activity
CG31705	NA	2L	32F2	2	-0.53	unknown
CG31998	NA	4	102A7	2	-0.53	unknown
CG11376	NA	2L	21B1	2	-0.52	GTP binding, GTPase binding, guanyl-nucleotide exchange factor activity
CG8098	Picot	2R	53C15	2	-0.52	high affinity inorganic phosphate:sodium symporter activity
CG17646	NA	2L	22B1	2	-0.52	ATPase activity, coupled to transmembrane movement of substances; transporter activity.
CG32145	omega	3L	70E7	2	-0.52	dipeptidyl-peptidase activity
CG18446	NA	2R	46C1	2	-0.51	zinc ion binding
CG10321	NA	2R	57F5	2	-0.51	zinc ion binding
CG1106	Gelsolin	3R	82A1	2	-0.51	actin binding
CG1884	Not1	2R	45F4	4	-0.5	basal transcription repressor activity
CG32048	NA	3L	67C1	4	-0.5	unknown
CG4900	Iron regulatory protein 1A	3R	94C3	2	-0.5	iron-responsive element binding, aconitate hydratase activity
RH01124	*contains transposon sequences*			2	-0.5	
CG3658	NA	X	1E1	2	-0.5	3'-5' DNA helicase activity; chromatin binding

Table 7.S7 Inheritance of compound second chromosomes is affected by the parent of origin.

Genotype of maternal parent	Cross	Number of adult females with maternally inherited chromosomes	Number of adult females with paternally inherited chromosomes
C(2)EN, b^1, pr^1	1	426	118
	2	327	59
C(2)EN, bw^1, sp^1	1	126	66
	2	150	176
Total		**1029**	**419**
Percent		71.1%	28.9%

7.6 TRANSITION TO CHAPTER 8

The manuscript presented here includes work that was undertaken as a side-project during my Ph.D., and was subsequently published in *Genome*. I am listed as an equal-contribution first author on the publication. Hermansky-Pudlak syndrome is an autosomal recessive disorder that results from defects in cellular trafficking. This work shows that the *Drosophila pink* gene is the orthologue of the human *Hermansky-Pudlak syndrome 5* (*Hps5*) gene. For this project I generated *P{pink-eGFP}* and *P{UAST-pink-EGFP}* transformation vectors, and created transgenic *P{UAS-pink-EGFP}* flies. I used immunofluorescence to characterize *pink-EGFP* distribution in COS-1 cells and *Drosophila* malpighian tubules, and conducted sequence comparison analysis. Identification of the *Drosophila* orthologue of the *Hps5* gene provides the basis for future studies to analyze and characterize the *Drosophila* HPS5 (pink) protein, and the effects of different mutations on this protein's function.

CHAPTER 8

THE *PINK* GENE ENCODES THE *DROSOPHILA* ORTHOLOGUE OF THE HUMAN HERMANSKY-PUDLAK SYNDROME 5 (*HPS5*) GENE.

Monika Syrzycka[1], Lori A. McEachern[1], Jennifer Kinneard, Kristel Prabhu, Kathleen Fitzpatrick, Sandra Schulze, John M. Rawls, Vett K. Lloyd, Donald A. R. Sinclair and Barry M. Honda.

[1] These authors contributed equally to this work.

Manuscript submitted January 9, 2007. Available online June 27, 2007
Published in *Genome* 50(6): 548-556.

8.1 ABSTRACT

Hermansky-Pudlak syndrome (HPS) consists of a set of human autosomal recessive disorders, with symptoms resulting from defects in genes required for protein trafficking in lysosome-related organelles, such as melanosomes and platelet-dense granules. A number of human HPS genes and rodent orthologues have been identified whose protein products are key components of one of four different protein complexes (AP-3 or BLOC-1, -2 and -3) that are key participants in the process. *Drosophila melanogaster* has been a key model organism in demonstrating the *in vivo* significance of many genes involved in protein trafficking pathways; for example, mutations in the "granule group" genes lead to changes in eye colour arising from improper protein trafficking to pigment granules in the developing eye. An examination of the chromosomal positioning of *Drosophila* HPS orthologues suggested that *CG9770*, the *Drosophila HPS5* orthologue, might correspond to the *pink* locus. Here we confirm this gene assignment, making *pink* as the first eye colour gene in flies to be identified as a BLOC complex gene.

8.2 INTRODUCTION

Hermansky-**P**udlak **s**yndrome (HPS) is a set of clinically related, genetically heterogeneous autosomal recessive disorders. All are characterized by defects of both melanosomes, leading to (variably severe) albinism, and platelet dense granules, leading to prolonged bleeding (2004; Hermansky and Pudlak, 1959; Huizing and Gahl, 2002). Both phenotypes (hypopigmentation and bleeding effects) are due to defects in trafficking

to, and maturation of, the specialized lysosome-related organelles found in melanocytes and blood cells: pigment granules in melanocytes and platelet dense granules in blood cells (Huizing *et al.*, 2004). Lysosomal and lysosome-related organelle defects are also seen in a variety of other cell types (Li *et al.*, 2004). These cellular defects can result in pulmonary, neurological, immune or hematopoetic clinical complications; for example, patients with HPS-1 or HPS-4 often develop progressive pulmonary fibrosis which can be attributed to defects in lysosome-related organelles, the lamellar bodies of type II lung epithelial cells (Nakatani *et al.*, 2000). In addition, HPS-5 patients have increased serum cholesterol levels, a symptom not documented in other forms of the disease (Huizing *et al.*, 2004; Zhang *et al.*, 2003).

Defects in eight human genes (*HPS1-8*) account for approximately half of the HPS cases known (Wei, 2006). Rodent coat colour mutants have been pivotal in identifying additional HPS genes (see Dell'Angelica, 2004; Di Pietro and Dell'Angelica, 2005; Li *et al.*, 2004; Odorizzi *et al.*, 1998; Wei, 2006). The role of vesicle trafficking in normal lysosomal function was illustrated by the finding that *HPS2* encodes the beta subunit of the **a**daptor-**p**rotein 3 (AP-3) complex that directs trafficking from the trans-Golgi network to endosomes (Dell'Angelica *et al.*, 1999; Odorizzi *et al.*, 1998; Simpson *et al.*, 1997). In addition to the AP-3 complex, a number of other complexes are involved in trafficking to, and maturation or stability of, lysosome-related organelles (Table 8.1). Two other complexes, containing orthologues of either VPS (vacuolar protein sorting) or Rab genes of yeast, when mutant, give rise to defects in mouse coat colour or *Drosophila* eye colour (Huizing *et al.*, 2002a). Members of these complexes specifically implicated in HPS include a component of the homotypic vacuolar protein sorting (HOPS) complex (Vsp33a; Suzuki *et al.*, 2003), and mammalian Rab38 (Loftus *et al.*, 2002).

Studies in the model organism *Drosophila* have been invaluable in providing demonstrations *in vivo* of the physiological relevance of some of these genes and complexes, due to the readily observed eye colour phenotypes associated with mutations affecting protein trafficking in pigment granules, e.g. *garnet* (Lloyd *et al.*, 1998; Lloyd *et al.*, 1999; Ooi *et al.*, 1997; Simpson *et al.*, 1997), *deep orange* and *carnation* (Sevrioukov *et al.*, 1999; Shestopal *et al.*, 1997), *light* (Warner *et al.*, 1998), *carmine* (Mullins *et al.*, 1999), and *ruby* (Kretzschmar *et al.*, 2000). Mutations in the *Drosophila ruby* gene (the β subunit of the AP-3 complex) also show behavioural defects (Kretzschmar *et al.*, 2000); this β subunit corresponds to the *HPS2* gene product (Huizing *et al.*, 2002b). Several additional fly genes have been predicted to encode orthologues of human HPS genes (Dell'Angelica, 2004; Huizing *et al.*, 2002a).

Analyses of other HPS genes and their mouse orthologues have identified three additional, novel complexes involved in the function of specialized lysosome-like organelles in metazoans: the **b**iogenesis of **l**ysosome-related **o**rganelle **c**omplexes (BLOC) 1-3. The mouse *HPS7* (*sandy*) and *HPS8* (*reduced pigmentation*) genes encode subunits of BLOC-1, which interacts with the cytoskeleton. Mutations in *HPS7* (which encodes the protein dysbindin) have also been implicated in muscular dystrophy and schizophrenia, consistent with a possible role in neuronal function (see Nazarian *et al.*, 2006 and references therein). Mouse *HPS3* (*cocoa*), *HPS5* (*ruby-eye 2*) and *HPS6* (*ruby-eye*) encode subunits of the BLOC-2 complex, the molecular function of which is unknown, and *HPS4* (*light ear*) and *HPS1* (*pale ear*) encode subunits of BLOC-3, a complex that may regulate positioning of lysosomes and related organelles (reviewed by Li *et al.*, 2004; Wei, 2006). Investigation of the remaining mouse and *Drosophila*

pigmentation mutants is likely to lead not only to the identification of additional human HPS genes but also to a better understanding of the cellular events leading to the formation and correct functioning of the wide variety of specialized lysosome-like organelles in metazoan animals.

Here, we report that the *Drosophila pink* (*p*) gene is the orthologue of the human *HPS5* and mouse *ruby-eye 2* genes. This was established by transgenic rescue and sequencing of three existing *pink* alleles. Consistent with the endosomal location of the Pink protein, mutations in the *pink* gene result in decreased pigmentation and behavioural phenotypes, as observed with other members of the granule group of genes.

Table 8.1 Different protein complexes associated with HPS genes.

Complex	Human gene name	Associated HPS syndrome	Mouse strain	*Drosophila* orthologue
AP-3	*HPS2/AP3B1*[a]	HPS-2	Pearl	*ruby*
	AP3D1[b]	?	Mocha	*garnet*
Rab geranylgeranyl transferase	*RABGGTA*[c]	?	Gunmetal	*?*
HOPS	*VPS33a*[c,d]	?	Buff	*carnation*
BLOC-1	*BLOC1S1*[e]	?	?	*CG30077*
	BLOC1S2[e]	?	?	*CG14145*
	BLOC1S3[e]	HPS-8	Reduced pigmentation	*?*
	CNO[f]		Cappuccino	*CG14149*
	DTNBP1[g]	HPS-7	Sandy	*CG6856*
	MUTED[h]	?	Muted	*CG34131*
	PLDN[h]	?	Pallid	*CG14133*
	Snapap[h]	?	?	*CG9958*
BLOC-2	*HPS3*[i]	HPS-3	Cocoa	*CG14562*
	HPS5[h]	HPS-5	Ruby-eye 2	*pink, CG9770*
	HPS6[h]	HPS-6	Ruby-eye	*?*
BLOC-3	*HPS1*[h]	HPS-1	Pale ear	*CG12855*
	HPS4[h]	HPS-4	Light ear	*CG4966*

[a] Huizing *et al.*, 2002b
[b] Kantheti *et al.*, 1998
[c] Detter *et al.*, 2000
[d] Suzuki *et al.*, 2003
[e] Starcevic and Dell'Angelica, 2004
[f] Ciciotte *et al.*, 2003
[g] Li *et al.*, 2004
[h] Nazarian *et al.*, 2006
[i] Di Pietro *et al.*, 2004

8.3 MATERIALS AND METHODS

8.3.1 *DROSOPHILA* GENETICS

All the mutations used in this study (p^1, p^p, p^{snb}, *Df(3R)p* ^{25}, *In(3R)p*^{419}, *T(Y;3)p*^{X14} and *P{EPgy2}CG9770*^{EY00893}) are described in FlyBase (Grumbling and Strelets, 2006) and were obtained from the Bloomington *Drosophila* Stock Center (Indiana University). The construction of EGFP (enhanced green fluorescent protein)-tagged *pink*$^+$ transgenic strain, hereafter referred to as *pink-EGFP*, is described below. All crosses were performed at 22º C unless otherwise stated. Culture medium was standard cornmeal-molasses medium supplemented with propionic acid as a mould inhibitor.

8.3.2 STANDARD RECOMBINANT DNA METHODS

Restriction digests, ligations, DNA isolation and cloning, and polymerase chain reactions were performed with enzymes and buffers supplied by MBI Fermentas following standard procedures (Sambrook *et al.*, 1989), or via commercial kits (GE Healthcare, Invitrogen, Qiagen) following the respective manufacturer's instructions.

8.3.3 SEQUENCE ANALYSIS AND COMPARISON

The *pink/CG9770* genomic region (GenBank accession number NM_141553) was PCR amplified from p^p, p^1 and p^{snb} genomic DNA (see Figure 8.2 legend for primers) and the resulting PCR fragments were cloned into the TA cloning kit vector pCR2.1 (Invitrogen). The fragments were sequenced at the University of Calgary University Core DNA services facility (http://www.ucalgary.ca/~dnalab) on an ABI 3730 capillary DNA Sequencer, using the same primers. Sequences were analyzed using the FlyBase BLAST

server with the following default settings: Genome assembly (NT) database, program blastn, expect = 10 (http://flybase.net/blast/). WD40 repeats were identified using the EMBL-EBI (European Molecular Biology Organization – European Bioinformatics Institute) Interpro website and database; for further details see: http://www.ebi.ac.uk/interpro/ISpy?mode=single&ac=Q9VHN9.

8.3.4 Pigment Assays

Pteridine assays were performed as described by Ashburner (1989). Ten heads were homogenized per sample, the absorbance at 485 nm was measured in triplicate, and the values were averaged.

8.3.5 Construction of, and Assays with, Transgene Constructs

8.3.5.1 *pink-EGFP* plasmid vector for COS-1 cell transformation

The *pink-EGFP* plasmid, p+-EGFP, was constructed using the insert from cDNA clone RE38137 (5850 bp in the vector pFLC-I; GenBank accession number AY089625), obtained from the Berkeley *Drosophila* Genome Project via the Canadian *Drosophila* Microarray Centre (University of Toronto at Mississauga). Clone RE38137 differs from the sequence for *CG9770* given in the BDGP database at six different single nucleotide positions: one in the 5'UT region, four in the protein-coding region, and one in the 3'UT region. This cDNA also contains a poly(A) tail sequence and is lacking two, presumably linker, nucleotides at the 5' end. Of the four nucleotide changes in the protein coding region, three result in predicted amino acid changes: substitutions at positions 700, 1196 and 1856 change the predicted amino acid from T to A, L to Q, and K to R, respectively,

in the CG9770 to RE38137 sequences. As the clone was used for localization studies, rather than rescue, these changes are likely not relevant.

The *pink* cDNA was isolated as a 2036 bp *Eco*RI-*Bam*HI fragment, and was then ligated into the pEGFP-N1 vector (Clontech) multiple cloning site between the *Eco*RI and *Bam*HI restriction sites. Successful ligation was confirmed by restriction digestion with *Eco*RI, yielding a fragment of 6738 bp. Construction of the Pink-EGFP fusion protein involved removal of the predicted 203 C-terminal amino acids from the 826 residue Pink protein and addition of four linker amino acids, PVAT, derived from the multiple cloning site of the pEGFP-N1 vector, followed by the EGFP protein.

8.3.5.2 TRANSFORMATION OF COS-1 CELLS

COS-1 African green monkey kidney cells were maintained following standard procedures in a water-jacketed incubator with 5% carbon dioxide in 25 mL vented culture flasks (Falcon) to allow gas exchange. Cells were raised on Dulbecco's Modified Eagles Medium (DMEM) (Gibco) supplemented with 10% fetal bovine serum (FBS) (Gibco) and 2% penicillin-streptomycin antibiotic mixture (Gibco). Cells were subcultured approximately every 4 to 5 days. For transfection, 1.0 µg plasmid DNA and 4µL Superfect^{TM} transfection reagent (Qiagen) were added to 60µL DMEM (without FBS or antibiotics) and vigorously mixed. The solution was incubated for 10-15 minutes at room temperature before being added to a COS-1 subculture in a sterile six-well culture plate (Falcon), each well containing a sterile cover slip. Cells were then cultured for 24 to 48 hours before fixation, DAPI (4'6-diamidino-2-phenylindole) and antibody staining, and inspection using a Zeiss LSM 410 laser scanning confocal microscope with a 63X oil objective lens. The LAMP - 3 (lysosomal-associated membrane protein 3) antibody

H5C6 was obtained from the Developmental Studies Hybridoma Bank at the University of Iowa (http://www.uiowa.edu/~dshbwww/h5c6.html). The LAMP-3 antibody was detected with a Cy3-labelled goat anti-mouse secondary antibody (Jackson Immunoresearch).

8.3.5.3 GENOMIC RESCUE CONSTRUCTS

All genomic transgene constructs except *P{Wd}* were cloned into the vector pCaSpeR4 (Thummel *et al.*, 1988); *P{Wd}* was cloned into pW8 (Klemenz *et al.*, 1987). *P{Mkk4}* is a 4.9 kb *SacI/BglII* fragment and *P{Wd}* is a 5 kb *Xba*I/*Eco*RI fragment, subcloned from genomic clone λ50-8, described in Jones and Rawls (1988). P*{PX1}* is an 8.2 kb *Pst*I/*Xho*I fragment, *P{EK1}* is a 5.9 kb *Eco*RI/*Kpn*I fragment and *P{XE1}* is a 6.5 kb *Xho*I/*Eco*RI fragment, all subcloned from the genomic clone λ3-10, also described in Jones and Rawls (1988).

8.3.5.4 PP(UAST)-*PINK*-*EGFP* VECTOR FOR GERMLINE TRANSFORMATION

The *pink-EGFP* fragment was isolated from the p^+-EGFP plasmid described above as a 2777 bp *Eco*RI-*Not*I fragment. This fragment was ligated into the pUAST vector multiple cloning site between the *Eco*RI and *Not*I restriction sites. Successful ligation was confirmed by restriction digestion.

8.3.5.5 GENERATION OF TRANSGENIC *DROSOPHILA* LINES

For both genomic and cDNA transgenes, transgenic *Drosophila* were generated essentially as described by Spradling (1986). *Drosophila* w^{1118}*;Ki Δ2-3* embryos (Sved *et*

al., 1990) were injected with highly purified DNA, and embryos surviving injection were selected as first instar larvae, cultured on standard *Drosophila* medium, and crossed to w^{1118} individuals of the appropriate sex to detect successful transformants. The chromosome into which the transgene had inserted was determined through standard crosses with the *w, dp, e, ci* strain, which is marked on all four chromosome sets.

8.3.5.6 INTRACELLULAR LOCALIZATION OF EGFP-TAGGED PINK

Malpighian tubules were isolated from P(UAST)-p^{+}-*EGFP* individuals carrying a 3rd chromosome with a P(UAST)-p^{+}-*EGFP* transgene insert, balanced against the P{w^{+mC}=*tubP-GAL4*}LL7 strain, which expresses GAL4 ubiquitously from the tubulin promoter. Fluorescence and DIC (differential interference contrast) microscopy were performed using a Zeiss LSM 510 laser scanning confocal microscope with a 25X oil objective lens.

8.4 RESULTS

8.4.1 IDENTIFICATION OF THE *PINK/HPS5* GENE

A preliminary examination of the chromosomal positioning of *Drosophila* HPS gene orthologues (D.A.R. Sinclair, unpublished results) suggested that *CG9770*, the *Drosophila HPS5* orthologue, might correspond to the *pink* locus. In addition, a P-element insertion in exon 1 of *CG9770* (*P{EPgy2}CG9770*^{EY00893}, Bloomington stock number 15046; hereafter referred to as *EY00893*) produced a moderate eye colour phenotype in combination with *Df(3R)p*^{25}, which deletes the *pink* gene (data not shown, but see pigment assays below), indicating that *CG9770* was likely to correspond to *pink*. Finally, Southern blot analysis of radiation or P-element induced *pink* mutations showed

band shifts (data not shown), correlating loss of *pink* gene function with DNA rearrangement in the region of *CG9770* (e.g. *In(3R)p*^{419} and *T(Y;3)p*^{X14} indicated in Figure 8.1; Jones and Rawls, 1988)

We therefore undertook to rescue *pink* mutations with genomic clones to confirm the molecular identity of the *pink* locus. Figure 8.1 shows the genomic organization of the region near *CG9770*, including the genomic fragments cloned into P-vector transgenes and used to make transgenic fly lines. Transgene rescue (complementation) of the *pink* eye colour phenotype was tested by appropriate crosses to generate homozygous *pink* siblings, with and without the transgene. Transgenes were tested with the *pink* mutation within the *TM3* balancer chromosome as well as with a *pink-peach (p*^{p}*)* stock. For example, crosses of *P{EK1}/CyO; p*^{p}*/p*^{p} males to *+/+; p*^{p}*/p*^{p} females produced progeny displaying complementation of *pink* by the transgene: half of the animals were normal-wing red-eye (*P{EK1}/+; p*^{p}*/p*^{p}) and the remainder were Curly-wing pink-eye (*+/CyO; p*^{p}*/p*^{p}). Similar crosses using lines with transgenes other than *P{EK1}* (see Figure 8.1) produced normal-wing pink-eye and Curly-wing pink-eye animals (i.e. those transgenes failed to complement *pink* mutation).

Only *P{EK1}* restored the eyes to wild type; this is the only genomic clone which contains the full length *CG9770* sequence, and *CG9770* is the only full length predicted gene contained within the clone. Thus, we conclude that *CG9770* corresponds to the *pink* locus, and we will hereafter refer to the *CG9770* gene as *pink*.

The predicted Pink protein is 826 amino acids, and encoded by a gene with three exons and two introns of 82 bp and 78 bp (FlyBase: http://flybase.bio.indiana.edu/reports/FBgn0037605.html; see Figure 8.2 for a schematic diagram). There appear to be three regions near the N-terminus with homology to WD40

domains (amino acids 14-52, 54-93, and 101-140); see http://www.ebi.ac.uk/interpro/ISpy?mode=single&ac=Q9VHN9 for further details. As has been observed in metazoan orthologues (Zhang *et al.*, 2003), the C-terminal portion of the protein is conserved across evolution, although no clear functional domain has yet been identified (see supplementary Figure 8.S1 for sequence alignments).

We also sequenced three alleles of *pink* (p^{1}, p^{snb} and p^{p}), each of which contains a single base pair deletion within codon Q648, causing a frameshift mutation. This results in a truncated protein (648 wild-type amino acids, followed by 39 additional amino acids and a stop codon; for schematic diagram see Figure 8.2). The correlation between loss of protein function and altered C-terminus is consistent with the high level of C-terminal sequence conservation across evolution, as noted above. Each allele also has a silent base change (in the triplet encoding proline residue 62), and another base change in an upstream region, where the second nucleotide of a POU domain (Finney *et al.*, 1988) has been changed from A to C.

It seemed very unusual that all three, reportedly independent mutations in *pink* would be identical. To rule out contamination, stock mislabelling, or other artifacts, sequencing was repeated using mutant DNA from re-acquired stocks of each the three alleles, with the same results. Since it is highly unlikely that the same two sequence changes arose three times independently, we believe it more likely that contamination/loss of some *pink* mutant stocks may have occurred at some point in their history (see Discussion).

8.4.2 PHENOTYPIC ANALYSIS OF *PINK* MUTANTS: PIGMENT ASSAYS OF DIFFERENT *PINK* ALLELES

We measured eye pigment levels in adult flies bearing the three *pink* alleles as well as the P insertion mutation (Figure 8.3). The similar reduction in red eye colour pigment levels in p^1, p^p and p^{snb} strains is consistent with their having a common molecular defect. All three show levels around 17% of the $pink^+/Df(3R)p^{25}$ controls. The p^p allele shows similar levels of pigment when homozygous (p^p/p^p) or hemizygous for a deletion of the *pink* region ($p^p/Df(3R)p^{25}$), suggesting that the three *pink* alleles are nulls, at least with respect to eye pigmentation. $P\{EPgy2\}CG9770^{EY00893}/Df(3R)p^{25}$ flies also have reduced levels of pteridines (approximately 60% of the wild type pigment levels), which is consistent with the aforementioned observation that these flies display a modest eye colour phenotype.

8.4.3 INTRACELLULAR LOCALIZATION OF THE PINK PROTEIN

Malpighian tubules from transgenic flies containing a ubiquitously-expressed (see Materials and Methods) EGFP-tagged *pink* gene were examined (Figure 8.4A), because Malpighian tubules are active in endo- and exo-cytosis (Sullivan and Sullivan, 1975). Figure 8.4A shows that the Pink-EGFP fluorescence is punctate, suggesting a cytoplasmic endosomal or lysosomal localization, and Figure 8.4B clearly shows that this staining is not nuclear. Further analysis of Pink-EGFP in COS cells is also consistent with a cytoplasmic/endosomal location, with no Pink-EGFP fusion protein in lysosomes: LAMP-3, a lysosomal marker, shows no appreciable co-localization (Figure 8.4C).

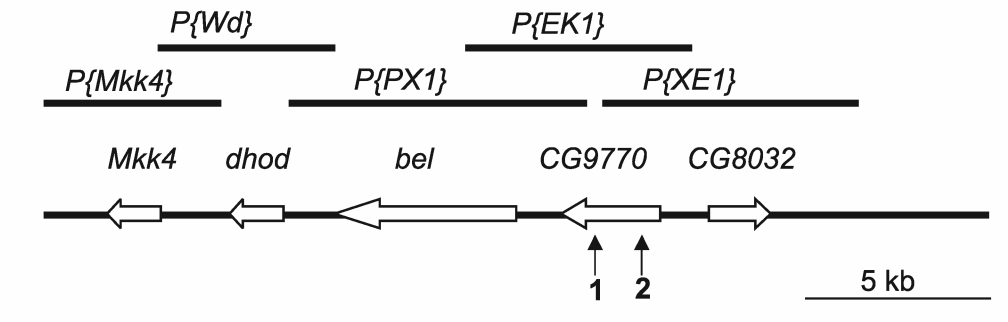


Figure 8.1 85A6 region map. The lowest line with open arrows shows the genomic organization of the region, including positions and directions of transcription for five genes (open horizontal arrows). The lines above the genomic region show the positions and extent of inserts used in different rescue constructs. Only the *P{EK1}* transgene rescues the eye colour phenotype of *pink* mutants. The numbered vertical arrows indicate approximate breakpoints of two chromosome rearrangements that produce *pink* mutant phenotypes: 1, $In(3R)p^{419}$; 2, $T(Y;3)p^{X14}$.

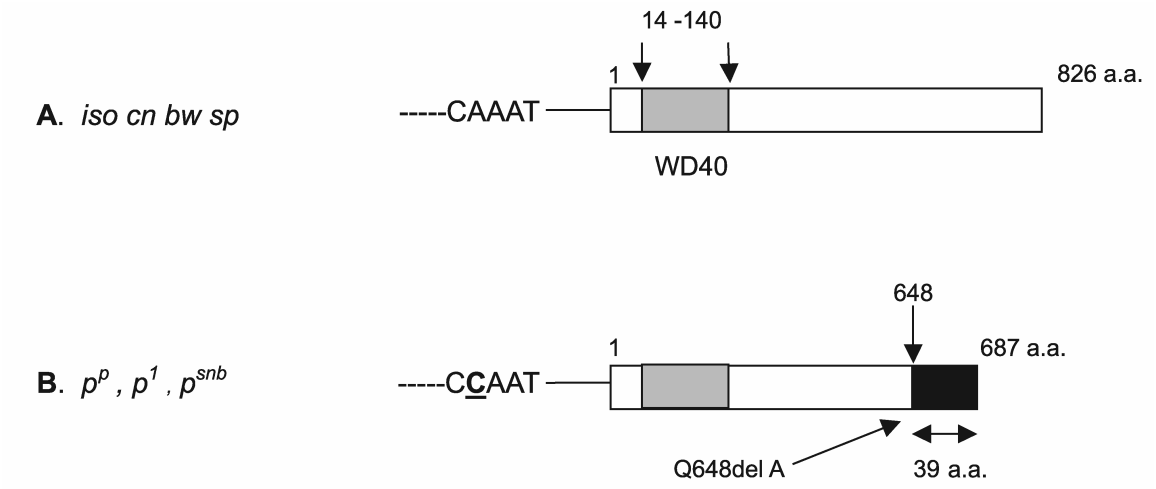


Figure 8.2 Schematic diagram showing nucleotide variations and predicted structure of the protein product of the *Drosophila pink* gene, *CG9770*. Schematic diagrams are based on the sequence from **[A]** the nominal "wild type" (BDGP), or **[B]** the three *pink* alleles. The putative POU domain nucleotide sequence, CAAAT, is shown upstream of the schematic box representing the protein. The region of WD40 repeats is shaded in grey. Pink primers used are as follows: Pink 5'552, 5'-TTGACAGTAGGTTTATGGGTT-GGG-3'; Pink 5', 5'-GGACGAGAAGCAGTTCTACTATGG-3'; Pink 5'down, 5'-AGCTCGGGAGGCAGTTCAGCCACC-3'; Pink 3'2284, 5'-ATTCTCTCGA-ACTGATTGGACCAC-3'; Pink 3'3106, 5'-GCTAGTTTCTCGAGCAGATTCACC-3'; and Pink 3'4240, 5'-TGGAAGAGAAATGAGGCTGGGCAG-3'.

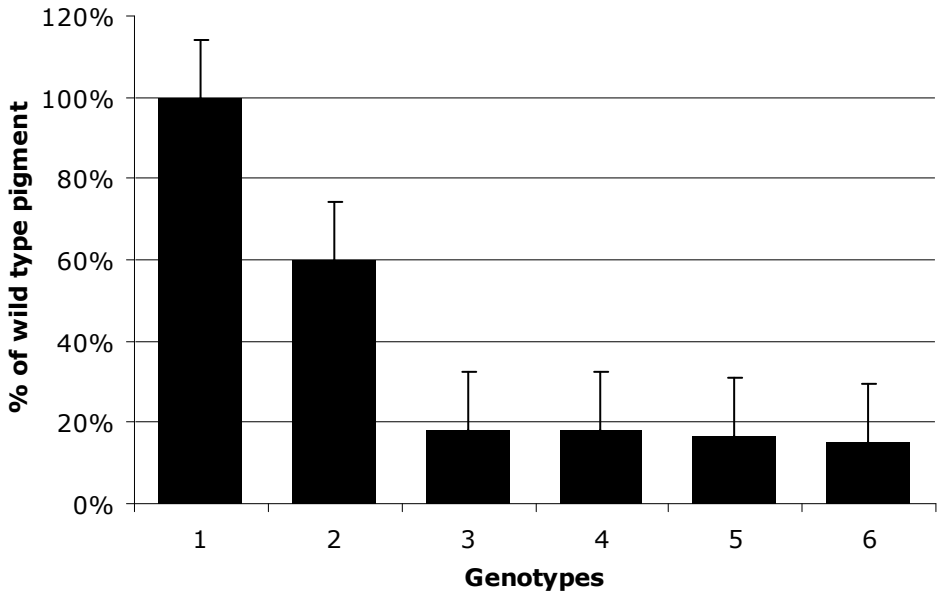


Figure 8.3 Pteridine pigment assays of *CG9770/pink* mutants. Assays were performed as described in Materials and methods, with standard error (SE) bars shown above each bar. Genotypes and % pigment levels were as follows: 1, *OR/Df(3R)p*^{25} (100%); 2, *EY00893/Df(3R)p*^{25} (60%); 3, *p*^{snb}*/Df(3R)p*^{25} (18%); 4, *p*^{p}*/Df(3R)p*^{25} (18%); 5, *p*^{1}*/Df(3R)p*^{25} (17%); 6, *p*^{p}*/ p*^{p} (15%).

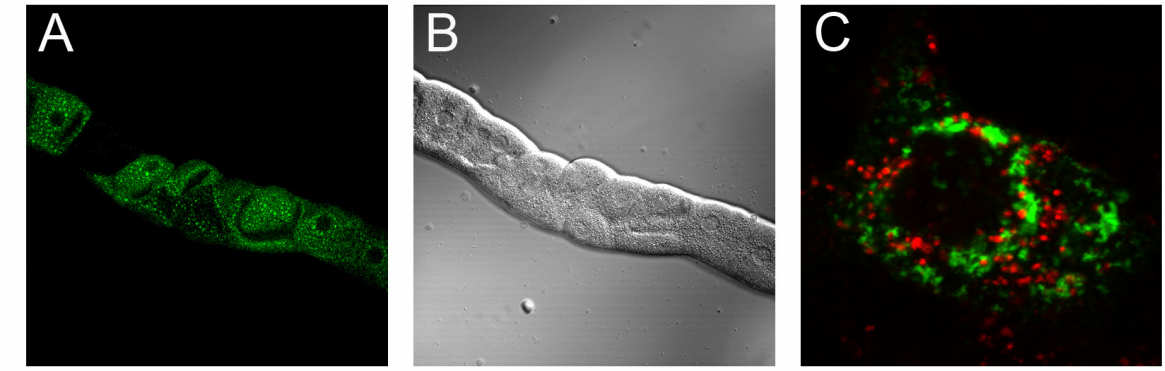


Figure 8.4 Localization of EGFP-tagged Pink protein. **[A]** Malpighian tubules from a *Drosophila* line constitutively expressing Pink-EGFP show fluorescence in a dispersed, punctate pattern consistent with a cytoplasmic location. **[B]** Corresponding DIC (differential interference contrast) microscopy image, indicating the positions of nuclei of some of the cells in panel A. **[C]** A COS-1 African green monkey kidney cell transformed with Pink-EGFP shows punctate perinuclear fluorescence (green) that does not co-localize with fluorescence from the lysosomal marker LAMP-3 (red).

8.5 DISCUSSION

The *Drosophila HPS5* orthologue *CG9770* corresponds to the *pink* locus, thus confirming an interesting link between eye and coat colour in mice and eye colour in flies, as well as to disorders involving lysosome biogenesis or maturation in a number of metazoans. In an independent study, Falcon-Perez *et al.* (2007) have also come to the same conclusion.

Pigment assays demonstrate that pteridine pigments are reduced, as would be expected for this type of mutation. Such a decrease is characteristic of the granule group of eye colour mutants and is likely due to decreased transport of the White ABC transporter protein to pigment granules (Lloyd *et al.*, 2002). The White protein is a transmembrane ABC-type transporter, responsible for loading metabolites into lysosome-related organelles, and affects both pigment pathways (Mackenzie *et al.*, 1999). The White protein complexes, which would normally be transported to pigment granules, are mislocalized in the developing eye cell. In neurons, mislocalization of White protein complexes induces male-male courtship behaviour (Campbell and Nash, 2001; Lloyd *et al.*, 2002; Zhang and Odenwald, 1995). p^1 and p^p males also show levels of male-male courtship significantly higher than wild-type (V.K. Lloyd, unpublished results).

HPS5 is part of the BLOC-2 complex, along with HPS3 and 6; however the function of the complex is unknown. In mammals, BLOC-2 protein complexes localize to early melanosomes; the observed cytoplasmic distribution of the *Drosophila* Pink-EGFP (Figure 8.4) is consistent with this observation. The Pink-EGFP fusion product does not appear to be treated by the cell as an abnormal protein: if so, it would be

expected to co-localize with LAMP-3 to the lysosome, and it clearly does not (Figure 8.4B).

HPS symptoms associated with mutations in BLOC-2 complex members tend to be less severe than some of the others, but HPS-5 patients have lipid metabolism defects – elevated levels of serum cholesterol (Huizing *et al.*, 2004; Zhang *et al.*, 2003). In humans, the White orthologue ABC8 is involved in cholesterol metabolism (Klucken *et al.*, 2000; Schmitz *et al.*, 2001). As noted above, mutations in *pink* may affect White transport (Lloyd *et al.*, 2002); a similar effect in mammals would result in a failure to properly transport cholesterol, and thus serum levels would increase in patients who have *HPS5* mutations. Consistent with this connection between *pink* and defects in lipid metabolism, the *C. elegans* orthologue of *pink* (W09G3.6) has been identified in a genome-wide RNAi screen for genes that regulate fat metabolism (Ashrafi *et al.*, 2003).

The most striking common molecular defect in the p^{1}, p^{snb} and p^{p} alleles is a single base pair deletion causing a frameshift that results in significant changes in the C terminus, in which some residues are conserved across metazoa (Figure 8.S1). An additional mutation was found in an upstream region, where a nucleotide change from A to C is observed in a POU domain (Finney *et al.*, 1988), but its function is currently unknown.

It was wholly unexpected to find these same sequence changes in all three alleles of *pink* that we tested. Sequence analysis of newly acquired stocks of the alleles ensured that no mistakes had been made in stock labelling, PCR, cloning, or sequencing. While it is formally possible that the same pair of sequence changes arose spontaneously on three separate occasions, it seems more likely that stock contamination, mislabelling, or loss occurred sometime during past handling of the stocks.

Many of the human and mouse mutations that have been characterized are similar to the *Drosophila* mutations described here, in that they also affect the C-terminal region (Figure 8.2), suggesting that this region performs some important function. However, in the 10 mammalian (7 human and 3 mouse; Huizing *et al.*, 2004; Zhang *et al.*, 2003) HPS genes described to date, no mutations have yet been identified in or 5' to the WD40 repeats, which are putative protein-protein interaction domains. This suggests that such mutations might be lethal in mammals. It would therefore be of considerable interest to study the effects of such mutations on the *Drosophila HPS5* gene. Future work with this and other HPS orthologues in *Drosophila* should provide powerful molecular and genetic tools to further study this set of disorders and associated physiological processes.

8.6 ACKNOWLEDGEMENTS

We wish to thank the Bloomington *Drosophila* Stock Center, The Canadian *Drosophila* microarray facility for cDNA clones, Anthony Howells for unpublished data on *pink*, discussion and encouragement, and NSERC Canada for Discovery grants to B.H. and V.K.L., and postgraduate studentships to M.S. and L.A.M.

8.7 REFERENCES

Ashburner M. (1989) *Drosophila: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Ashrafi K, Chang FY, Watts JL, Fraser AG, Kamath RS, Ahringer J, Ruvkun G (2003) Genome-wide RNAi analysis of Caenorhabditis elegans fat regulatory genes. *Nature* **421**: 268-272.

Campbell JL, Nash HA (2001) Volatile general anesthetics reveal a neurobiological role for the white and brown genes of Drosophila melanogaster. *J Neurobiol* **49**: 339-349.

Ciciotte SL, Gwynn B, Moriyama K, Huizing M, Gahl WA, Bonifacino JS, Peters LL (2003) Cappuccino, a mouse model of Hermansky-Pudlak syndrome, encodes a novel protein that is part of the pallidin-muted complex (BLOC-1). *Blood* **101**: 4402-4407.

Dell'Angelica EC (2004) The building BLOC(k)s of lysosomes and related organelles. *Curr Opin Cell Biol* **16**: 458-464.

Dell'Angelica EC, Shotelersuk V, Aguilar RC, Gahl WA, Bonifacino JS (1999) Altered trafficking of lysosomal proteins in Hermansky-Pudlak syndrome due to mutations in the beta 3A subunit of the AP-3 adaptor. *Mol Cell* **3**: 11-21.

Detter JC, Zhang Q, Mules EH, Novak EK, Mishra VS, Li W, McMurtrie EB, Tchernev VT, Wallace MR, Seabra MC, Swank RT, Kingsmore SF (2000) Rab geranylgeranyl transferase alpha mutation in the gunmetal mouse reduces Rab prenylation and platelet synthesis. *Proc Natl Acad Sci U S A* **97**: 4144-4149.

Di Pietro SM, Dell'Angelica EC (2005) The cell biology of Hermansky-Pudlak syndrome: recent advances. *Traffic* **6**: 525-533.

Di Pietro SM, Falcon-Perez JM, Dell'Angelica EC (2004) Characterization of BLOC-2, a complex containing the Hermansky-Pudlak syndrome proteins HPS3, HPS5 and HPS6. *Traffic* **5**: 276-283.

Falcon-Perez JM, Romero-Calderon R, Brooks ES, Krantz DE, Dell'Angelica EC (2007) The Drosophila pigmentation gene pink (p) encodes a homologue of human Hermansky-Pudlak syndrome 5 (HPS5). *Traffic* **8**: 154-168.

Finney M, Ruvkun G, Horvitz HR (1988) The C. elegans cell lineage and differentiation gene unc-86 encodes a protein with a homeodomain and extended similarity to transcription factors. *Cell* **55**: 757-769.

Grumbling G, Strelets V (2006) FlyBase: anatomical data, images and queries. *Nucleic Acids Res* **34**: D484-488.

Hermansky F, Pudlak P (1959) Albinism associated with hemorrhagic diathesis and unusual pigmented reticular cells in the bone marrow: report of two cases with histochemical studies. *Blood* **14**: 162-169.

Huizing M, Boissy RE, Gahl WA (2002a) Hermansky-Pudlak syndrome: vesicle formation from yeast to man. *Pigment Cell Res* **15**: 405-419.

Huizing M, Gahl WA (2002) Disorders of vesicles of lysosomal lineage: the Hermansky-Pudlak syndromes. *Curr Mol Med* **2**: 451-467.

Huizing M, Hess R, Dorward H, Claassen DA, Helip-Wooley A, Kleta R, Kaiser-Kupfer MI, White JG, Gahl WA (2004) Cellular, molecular and clinical characterization of patients with Hermansky-Pudlak syndrome type 5. *Traffic* **5**: 711-722.

Huizing M, Scher CD, Strovel E, Fitzpatrick DL, Hartnell LM, Anikster Y, Gahl WA (2002b) Nonsense mutations in ADTB3A cause complete deficiency of the beta3A subunit of adaptor complex-3 and severe Hermansky-Pudlak syndrome type 2. *Pediatr Res* **51**: 150-158.

Jones WK, Rawls JM, Jr. (1988) Genetic and molecular mapping of chromosome region 85A in Drosophila melanogaster. *Genetics* **120**: 733-742.

Kantheti P, Qiao X, Diaz ME, Peden AA, Meyer GE, Carskadon SL, Kapfhamer D, Sufalko D, Robinson MS, Noebels JL, Burmeister M (1998) Mutation in AP-3 delta in the mocha mouse links endosomal transport to storage deficiency in platelets, melanosomes, and synaptic vesicles. *Neuron* **21**: 111-122.

Klemenz R, Weber U, Gehring WJ (1987) The white gene as a marker in a new P-element vector for gene transfer in Drosophila. *Nucleic Acids Res* **15**: 3947-3959.

Klucken J, Buchler C, Orso E, Kaminski WE, Porsch-Ozcurumez M, Liebisch G, Kapinsky M, Diederich W, Drobnik W, Dean M, Allikmets R, Schmitz G (2000) ABCG1 (ABC8), the human homolog of the Drosophila white gene, is a regulator of macrophage cholesterol and phospholipid transport. *Proc Natl Acad Sci U S A* **97**: 817-822.

Kretzschmar D, Poeck B, Roth H, Ernst R, Keller A, Porsch M, Strauss R, Pflugfelder GO (2000) Defective pigment granule biogenesis and aberrant behavior caused by mutations in the Drosophila AP-3beta adaptin gene ruby. *Genetics* **155**: 213-223.

Li W, Rusiniak ME, Chintala S, Gautam R, Novak EK, Swank RT (2004) Murine Hermansky-Pudlak syndrome genes: regulators of lysosome-related organelles. *Bioessays* **26**: 616-628.

Lloyd V, Ramaswami M, Kramer H (1998) Not just pretty eyes: Drosophila eye-colour mutations and lysosomal delivery. *Trends Cell Biol* **8**: 257-259.

Lloyd VK, Sinclair DA, Alperyn M, Grigliatti TA (2002) Enhancer of garnet/deltaAP-3 is a cryptic allele of the white gene and identifies the intracellular transport system for the white protein. *Genome* **45**: 296-312.

Lloyd VK, Sinclair DA, Wennberg R, Warner TS, Honda BM, Grigliatti TA (1999) A genetic and molecular characterization of the garnet gene of Drosophila melanogaster. *Genome* **42**: 1183-1193.

Loftus SK, Larson DM, Baxter LL, Antonellis A, Chen Y, Wu X, Jiang Y, Bittner M, Hammer JA, 3rd, Pavan WJ (2002) Mutation of melanosome protein RAB38 in chocolate mice. *Proc Natl Acad Sci U S A* **99**: 4471-4476.

Mackenzie SM, Brooker MR, Gill TR, Cox GB, Howells AJ, Ewart GD (1999) Mutations in the white gene of Drosophila melanogaster affecting ABC transporters that determine eye colouration. *Biochim Biophys Acta* **1419**: 173-185.

Mullins C, Hartnell LM, Wassarman DA, Bonifacino JS (1999) Defective expression of the mu3 subunit of the AP-3 adaptor complex in the Drosophila pigmentation mutant carmine. *Mol Gen Genet* **262**: 401-412.

Nakatani Y, Nakamura N, Sano J, Inayama Y, Kawano N, Yamanaka S, Miyagi Y, Nagashima Y, Ohbayashi C, Mizushima M, Manabe T, Kuroda M, Yokoi T, Matsubara O (2000) Interstitial pneumonia in Hermansky-Pudlak syndrome: significance of florid foamy swelling/degeneration (giant lamellar body degeneration) of type-2 pneumocytes. *Virchows Arch* **437**: 304-313.

Nazarian R, Starcevic M, Spencer MJ, Dell'Angelica EC (2006) Reinvestigation of the dysbindin subunit of BLOC-1 (biogenesis of lysosome-related organelles complex-1) as a dystrobrevin-binding protein. *Biochem J* **395**: 587-598.

Odorizzi G, Cowles CR, Emr SD (1998) The AP-3 complex: a coat of many colours. *Trends Cell Biol* **8**: 282-288.

Ooi CE, Moreira JE, Dell'Angelica EC, Poy G, Wassarman DA, Bonifacino JS (1997) Altered expression of a novel adaptin leads to defective pigment granule biogenesis in the Drosophila eye color mutant garnet. *Embo J* **16**: 4508-4518.

Sambrook J, Fritsch EF, Maniatis T. (1989) *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Schmitz G, Langmann T, Heimerl S (2001) Role of ABCG1 and other ABCG family members in lipid metabolism. *J Lipid Res* **42**: 1513-1520.

Sevrioukov EA, He JP, Moghrabi N, Sunio A, Kramer H (1999) A role for the deep orange and carnation eye color genes in lysosomal delivery in Drosophila. *Mol Cell* **4**: 479-486.

Shestopal SA, Makunin IV, Belyaeva ES, Ashburner M, Zhimulev IF (1997) Molecular characterization of the deep orange (dor) gene of Drosophila melanogaster. *Mol Gen Genet* **253**: 642-648.

Simpson F, Peden AA, Christopoulou L, Robinson MS (1997) Characterization of the adaptor-related protein complex, AP-3. *J Cell Biol* **137**: 835-845.

Spradling AC (1986) P element-mediated transformation. In *Drosophila: A Practical Approach*, Roberts DB (ed), pp 175-197. Oxford: IRL Press.

Starcevic M, Dell'Angelica EC (2004) Identification of snapin and three novel proteins (BLOS1, BLOS2, and BLOS3/reduced pigmentation) as subunits of biogenesis of lysosome-related organelles complex-1 (BLOC-1). *J Biol Chem* **279**: 28393-28401.

Sullivan DT, Sullivan MC (1975) Transport defects as the physiological basis for eye color mutants of Drosophila melanogaster. *Biochem Genet* **13**: 603-613.

Suzuki T, Oiso N, Gautam R, Novak EK, Panthier JJ, Suprabha PG, Vida T, Swank RT, Spritz RA (2003) The mouse organellar biogenesis mutant buff results from a mutation in Vps33a, a homologue of yeast vps33 and Drosophila carnation. *Proc Natl Acad Sci U S A* **100**: 1146-1150.

Sved JA, Eggleston WB, Engels WR (1990) Germ-line and somatic recombination induced by in vitro modified P elements in Drosophila melanogaster. *Genetics* **124**: 331-337.

Thummel CS, Boulet AM, Lipshitz HD (1988) Vectors for Drosophila P-element-mediated transformation and tissue culture transfection. *Gene* **74**: 445-456.

Warner TS, Sinclair DA, Fitzpatrick KA, Singh M, Devlin RH, Honda BM (1998) The light gene of Drosophila melanogaster encodes a homologue of VPS41, a yeast gene involved in cellular-protein trafficking. *Genome* **41**: 236-243.

Wei ML (2006) Hermansky-Pudlak syndrome: a disease of protein trafficking and organelle function. *Pigment Cell Res* **19**: 19-42.

Zhang Q, Zhao B, Li W, Oiso N, Novak EK, Rusiniak ME, Gautam R, Chintala S, O'Brien EP, Zhang Y, Roe BA, Elliott RW, Eicher EM, Liang P, Kratz C, Legius E, Spritz RA, O'Sullivan TN, Copeland NG, Jenkins NA, Swank RT (2003) Ru2 and Ru encode mouse orthologs of the genes mutated in human Hermansky-Pudlak syndrome types 5 and 6. *Nat Genet* **33**: 145-153.

Zhang SD, Odenwald WF (1995) Misexpression of the white (w) gene triggers male-male courtship in Drosophila. *Proc Natl Acad Sci U S A* **92**: 5525-5529.

8.8 SUPPLEMENTARY MATERIAL

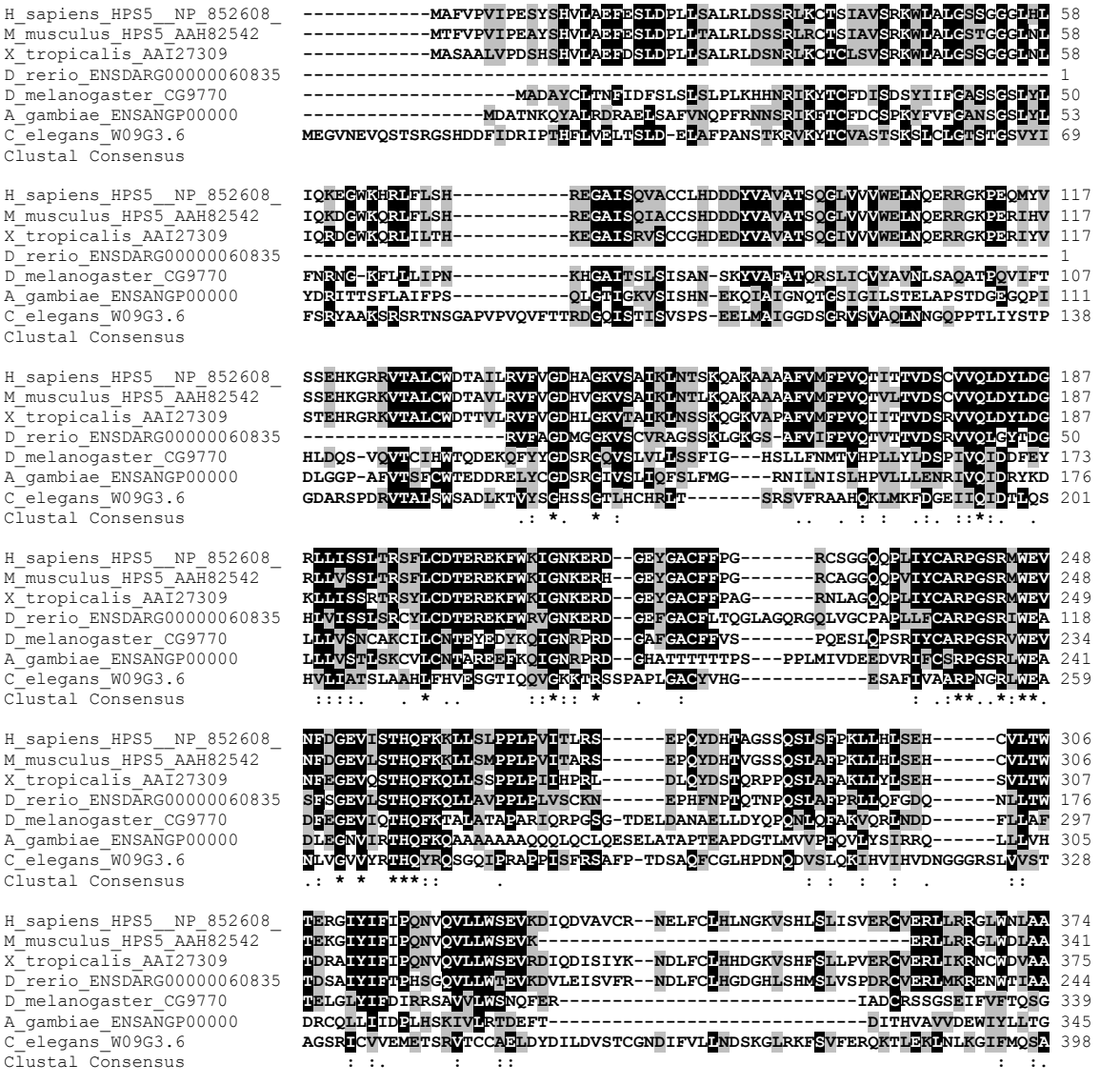


Figure 8.S1 Sequence alignments of HPS5 orthologues across species. The *Drosophila* CG9770 amino acid sequence was used to find orthologues via the NCBI protein-protein BLAST server (http://www.ncbi.nlm.nih.gov/BLAST/). Multiple sequence alignments were obtained using the ClustalW and BOXSHADE programs available through the San Diego Supercomputer Centre (SDSC) Biology Workbench (http://workbench.sdsc.edu/). Homo_sapiens_HP = Homo sapiens HPS5 (NP 852608), Mus_musculus_HP = mouse HPS5 (AAH82542), Xenopus_tropica = Xenopus tropicalis AAI27309, Danio_rerio_ENS = Danio rerio ENSDARG00000060835, D_melanogaster_ = Drosophila melanogaster CG9770, Anopheles_gambi = Anopheles gambiae ENSANGP00000027610, C_elegans_W09G3 = Caenorhabditis elegans W09G3.6.

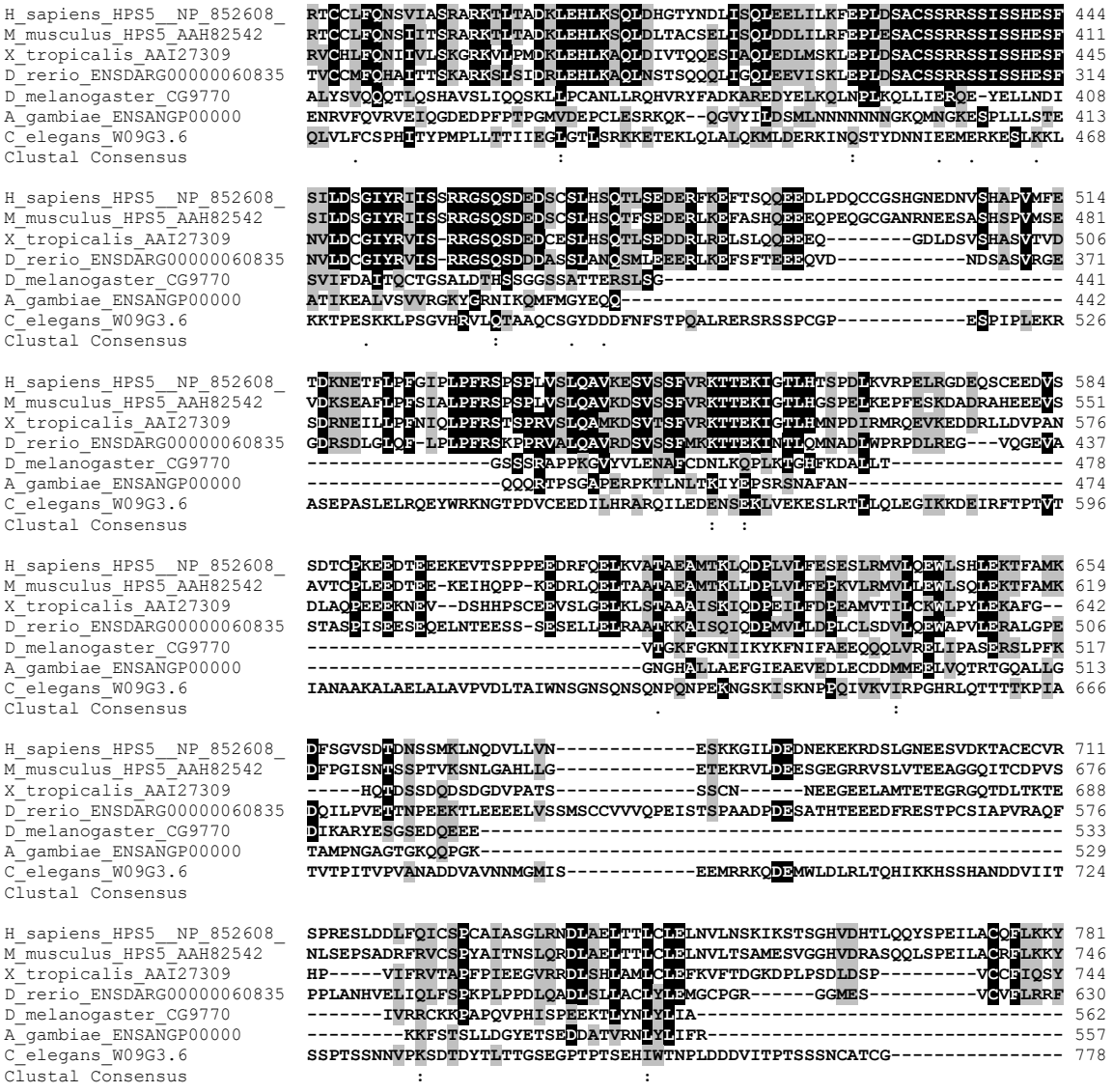


Figure 8.S1 (continued, 2 of 3) Sequence alignments of HPS5 orthologues across species.

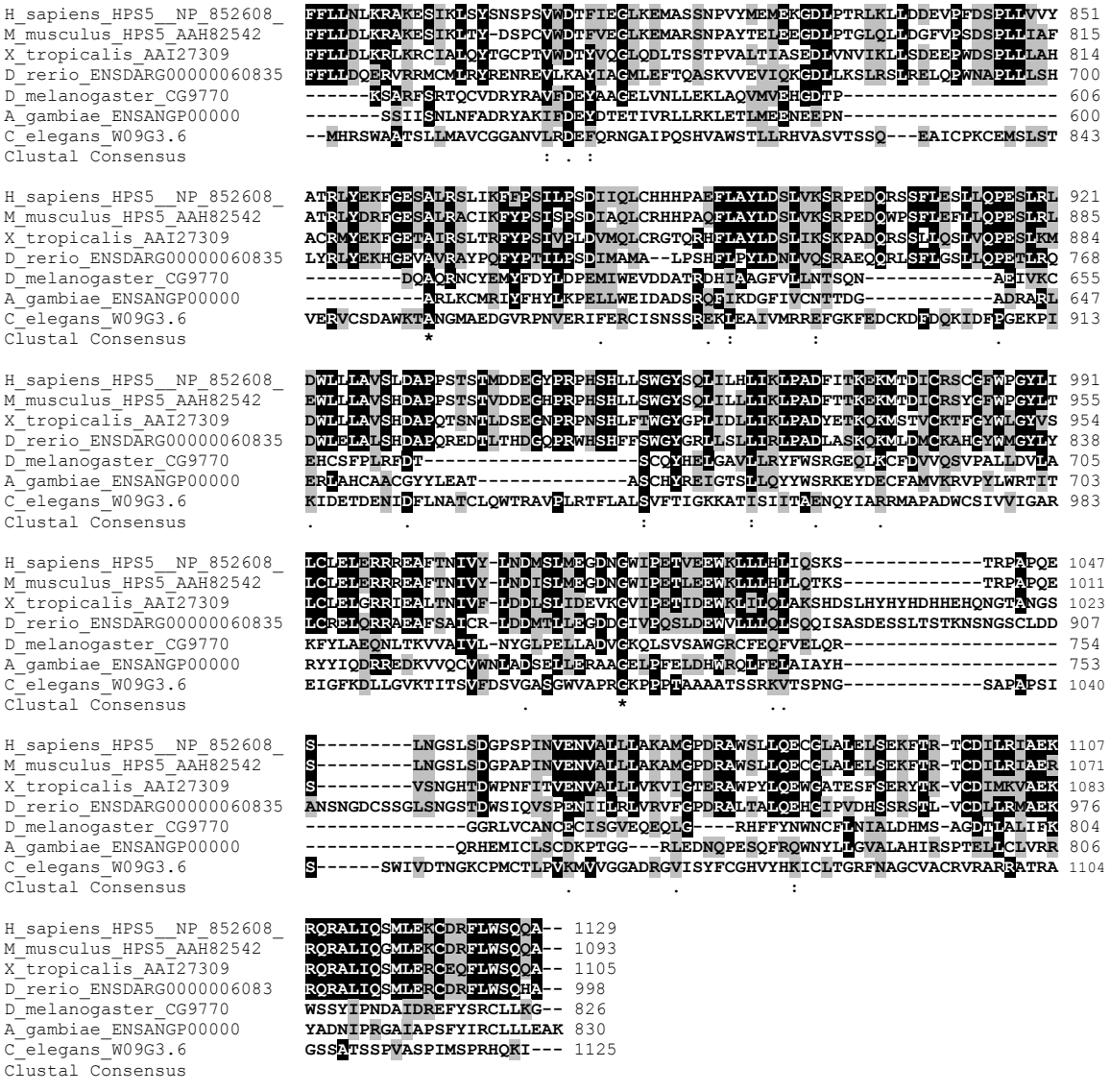


Figure 8.S1 (continued, 3 of 3) Sequence alignments of HPS5 orthologues across species.

8.9 TRANSITION TO CHAPTER 9

The following chapter is a summary of the research presented in this thesis, highlighting the significance of my findings. I compile and elaborate on evidence that has been presented in this thesis that supports the hypothesis that seemingly unique epigenetic phenomena, such as paramutation and genomic imprinting, function by utilizing core epigenetic mechanisms that are highly conserved.

CHAPTER 9

CONCLUSION

The work presented and discussed in this thesis demonstrates significant conservation of epigenetic mechanisms between the plant and animal kingdoms. While previous studies have demonstrated conservation of epigenetic mechanisms within the animal kingdom, by showing that mammalian imprint centres function as epigenetic silencers in transgenic *Drosophila* (Arney *et al.*, 2006; Lyko *et al.*, 1997; Lyko *et al.*, 1998), the work presented here is the first to test for evolutionary conservation of epigenetic silencing between plants and animals using a transgenic system. The extensive evolutionary distance between maize, an angiosperm plant, and *Drosophila,* a dipteran insect, provides strong support for the hypothesis that core epigenetic mechanisms are conserved throughout the eukaryotic kingdom, and that seemingly unique epigenetic phenomena function by exploiting these core mechanisms.

In the work presented here, I have shown that the maize *b1* tandem repeat control region functions as an epigenetic silencer in *cis* in *Drosophila*, and also demonstrates significant evidence of retaining other epigenetic features of the endogenous maize locus. For example, the *b1* tandem repeats mediate interactions between paired sequences (chapter 2) and are bidirectionally transcribed (chapter 3). These results indicate considerable conservation of epigenetic processes in the transgenic *Drosophila* system.

In maize, pairing of a silenced *B'* allele with a highly expressed *B-I* allele causes heritable epigenetic silencing of *B-I* (Coe, 1966; Patterson *et al.*, 1993). The tandem repeat control region is required for this process, and the strength of paramutation decreases as the number of tandem repeats decreases (Stam *et al.*, 2002). Similarly, I have shown that transgenes with the full seven tandem repeats frequently demonstrate pairing-sensitive silencing, a *trans* silencing effect in which paired transgenes exhibit

reduced expression compared with a single hemizygous copy. Pairing-sensitive silencing is reduced or lost when the number of tandem repeats decreases from seven to three or zero, indicating the requirement of the tandem repeat array in stabilizing this interaction, as is the case for the *b1* locus in maize.

Trans-communication between alleles is essential for paramutation, and requires a threshold number of repeats (Stam *et al.*, 2002). Similarly, I have shown that the presence of the *b1* tandem repeats on one chromosomal homologue can mediate epigenetic silencing or activation of a "repeats-out" transgene present in *trans* on the paired chromosomal homologue. This repeats-in / repeats-out *trans* effect was observed for all seven tandem repeat transgenes, but only one of two transgenes with three tandem repeats, and no transgenes with one or two tandem repeats, further emphasizing the importance of the number of tandem repeats in mediating *trans*-interactions in *Drosophila*.

In *Drosophila*, pairing-sensitive silencing is frequently observed for transgenes containing Polycomb Response Elements (PREs), which recruit and bind Polycomb group (PcG) proteins to mediate epigenetic silencing (Kassis, 2002). *Drosophila* subtelomeric tandem repeats (also known as telomere-associated sequences, or TAS) have been shown to act as PREs and share several similarities with *b1* tandem repeat-mediated silencing. TAS tandem repeats cause reporter gene silencing (Boivin *et al.*, 2003; Cryderman *et al.*, 1999) that increases in strength as the number of tandem repeats increases (Kurenova *et al.*, 1998), and can mediate pairing-sensitive silencing (Boivin *et al.*, 2003), *trans*-silencing (Josse *et al.*, 2007; Ronsseray *et al.*, 2003) and *trans*-communication (Frydrychova *et al.*, 2007; Mason *et al.*, 2003). TAS tandem repeats also exhibit both heterochromatic and euchromatic histone modifications (Yin and Lin, 2007),

and a similar dual-epigenetic state has recently been described for the *b1* locus in maize (Haring *et al.*, 2010). Given the similarities between *Drosophila* TAS-mediated epigenetic silencing and *b1* repeat-mediated epigenetic silencing, I hypothesized that the *b1* tandem repeats act as a PRE and recruit PcG proteins in *Drosophila*. Consistent with this, preliminary mutational analysis suggests that PcG proteins participate in epigenetic regulation of the *b1* tandem repeats in *Drosophila* (chapter 4). PcG proteins are highly conserved throughout eukaryotes, and have been shown to play important roles in other epigenetic phenomena in plants, including genomic imprinting (Jullien and Berger, 2009). The involvement of PcG proteins in *b1* silencing in *Drosophila* may indicate that these proteins are important at the endogenous locus as well. The transgenic *Drosophila* system also provides a unique opportunity to study the molecular mechanisms of *b1* paramutation in more detail, as a tremendous number of *Drosophila* mutant strains are readily available, and can be tested with relative speed and ease.

Currently, relatively little is known about the molecular mechanisms and proteins involved in *b1* paramutation in maize. The observations that the tandem repeats are bidirectionally transcribed, and that paramutation requires an RNA-dependent RNA polymerase (Alleman *et al.*, 2006), as well as components of RNA polymerases IV and V (Erhard *et al.*, 2009; Hollick *et al.*, 2005; Sidorenko *et al.*, 2009), were significant breakthroughs in determining that RNA based mechanisms are involved. In the work presented here, I have characterized bidirectional transcription at the *b1* tandem repeats in *Drosophila* (chapter 3). Interestingly, aberrant transcription persists at transgenes following repeat removal. This suggests that the transgenes acquire an epigenetic mark that serves as a "memory" of its previous epigenetic state, and is stable through meiosis for many generations. This epigenetic mark is, however, disrupted by changing the local

chromatin environment. Stability of an epigenetic mark through meiosis is a characteristic of paramutation, further underscoring the conservation of the epigenetic mechanisms underlying paramutation, and also providing a platform to study a unique case of epigenetic memory not previously described in *Drosophila*.

A detailed analysis of the molecular mechanisms underlying genomic imprinting, an epigenetic phenomenon that is observed in a wide range of plant and animal species, further underscores that unique epigenetic phenomena function by exploiting core epigenetic mechanisms (chapter 6). Plants, mammals and insects utilize many of the same mechanisms to establish and maintain imprinted expression. These mechanisms include DNA methylation, histone modifications, changes in higher order chromatin structure, non-coding RNA and RNA interference, all of which are frequently interrelated and mutually reinforcing. Histone modifications have been observed to play an essential role in plant, insect, and mammalian imprinting, and can result in parent of origin specific higher order chromatin structures or modifications, that contribute to the imprinting of genes, gene clusters, or chromosomes (Goday and Ruiz, 2002; Greciano and Goday, 2006; Haun and Springer, 2008; Joanis and Lloyd, 2002; Li *et al.*, 2008). A homologous Polycomb complex participates in both plant and mammalian imprinting, further emphasizing their relatedness (Jullien *et al.*, 2006; Kohler *et al.*, 2005; Li *et al.*, 2008).

The conservation of genomic imprinting was first revealed by transgenic experiments in which imprint control regions from mice and humans successfully functioned in *Drosophila* (Lyko *et al.*, 1997; Lyko *et al.*, 1998). Interestingly, while *Drosophila* exhibit extensive evidence of genomic imprinting, endogenously imprinted genes have not yet been identified. In this thesis, I have presented gene expression analysis of gynogenetic *Drosophila*, and preliminary analysis of an androgenetic fly, as

well as flies inheriting compound second or third chromosomes maternally or paternally (chapter 7). This analysis has identified several interesting candidate imprinted genes that can now be examined further to determine their imprinted status.

Drosophila have long been a valuable resource for a wide range of genetic and epigenetic studies. Initially, transgenic experiments examined genetic conservation by assessing the function of mammalian genes and proteins in transgenic *Drosophila* (Jowett et al., 1991). More recently, transgenic *Drosophila* have provided valuable insight into the evolutionary conservation of molecular mechanisms underlying epigenetic processes. The conservation of core epigenetic mechanisms enhances the utility of *Drosophila* in epigenetic studies, as the transgenic system can then be used to advance the understanding of the molecular mechanisms and proteins that function at the endogenous locus. Here I have provided significant insight into the conservation of the mechanisms underlying *b1* paramutation in maize. I have shown that the silencing and *trans* interaction functions, and the transcriptional status, of the *b1* repeats are conserved in *Drosophila*. This is the greatest evolutionary distance tested for a complex epigenetic process, and provides substantial evidence that epigenetic mechanisms are conserved between the plant and animal kingdoms.

REFERENCES

Achwal CW, Iyer CA, Chandra HS (1983) Immunochemical evidence for the presence of 5mC, 6mA and 7mG in human, Drosophila and mealybug DNA. *FEBS Lett* **158**: 353-358.

Ainscough JF, Dandolo L, Surani MA (2000a) Appropriate expression of the mouse H19 gene utilises three or more distinct enhancer regions spread over more than 130 kb. *Mech Dev* **91**: 365-368.

Ainscough JF, John RM, Barton SC, Surani MA (2000b) A skeletal muscle-specific mouse Igf2 repressor lies 40 kb downstream of the gene. *Development* **127**: 3923-3930.

Alleman M, Doctor J (2000) Genomic imprinting in plants: observations and evolutionary implications. *Plant Mol Biol* **43**: 147-161.

Alleman M, Sidorenko L, McGinnis K, Seshadri V, Dorweiler JE, White J, Sikkink K, Chandler VL (2006) An RNA-dependent RNA polymerase is required for paramutation in maize. *Nature* **442**: 295-298.

Anaka M, Lynn A, McGinn P, Lloyd VK (2009) Genomic Imprinting in Drosophila has properties of both mammalian and insect imprinting. *Dev Genes Evol* **219**: 59-66.

Armstrong JA, Papoulas O, Daubresse G, Sperling AS, Lis JT, Scott MP, Tamkun JW (2002) The Drosophila BRM complex facilitates global transcription by RNA polymerase II. *Embo J* **21**: 5245-5254.

Arney KL, Bae E, Olsen C, Drewell RA (2006) The human and mouse H19 imprinting control regions harbor an evolutionarily conserved silencer element that functions on transgenes in Drosophila. *Dev Genes Evol* **216**: 811-819.

Arteaga-Vazquez MA, Chandler VL (2010) Paramutation in maize: RNA mediated trans-generational gene silencing. *Curr Opin Genet Dev* **20**: 156-163.

Ashburner M. (1989) *Drosophila: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Ashrafi K, Chang FY, Watts JL, Fraser AG, Kamath RS, Ahringer J, Ruvkun G (2003) Genome-wide RNAi analysis of Caenorhabditis elegans fat regulatory genes. *Nature* **421**: 268-272.

Baker BS (1975) Paternal loss (pal): a meiotic mutant in Drosophila melanogaster causing loss of paternal chromosomes. *Genetics* **80**: 267-296.

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

Bao N, Lye KW, Barton MK (2004) MicroRNA binding sites in Arabidopsis class III HD-ZIP mRNAs are required for methylation of the template chromosome. *Dev Cell* **7**: 653-662.

Baroux C, Blanvillain R, Gallois P (2001) Paternally inherited transgenes are down-regulated but retain low activity during early embryogenesis in Arabidopsis. *FEBS Lett* **509**: 11-16.

Baroux C, Gagliardini V, Page DR, Grossniklaus U (2006) Dynamic regulatory interactions of Polycomb group genes: MEDEA autoregulation is required for imprinted gene expression in Arabidopsis. *Genes Dev* **20**: 1081-1086.

Bartolomei MS, Webber AL, Brunkow ME, Tilghman SM (1993) Epigenetic mechanisms underlying the imprinting of the mouse H19 gene. *Genes Dev* **7**: 1663-1673.

Bartolomei MS, Zemel S, Tilghman SM (1991) Parental imprinting of the mouse H19 gene. *Nature* **351**: 153-155.

Bean CJ, Schaner CE, Kelly WG (2004) Meiotic pairing and imprinted X chromatin assembly in Caenorhabditis elegans. *Nat Genet* **36**: 100-105.

Beaudet AL, Jiang YH (2002) A rheostat model for a rapid and reversible form of imprinting-dependent evolution. *Am J Hum Genet* **70**: 1389-1397.

Bell AC, Felsenfeld G (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* **405**: 482-485.

Bender LB, Cao R, Zhang Y, Strome S (2004) The MES-2/MES-3/MES-6 complex and regulation of histone H3 methylation in C. elegans. *Curr Biol* **14**: 1639-1643.

Bernstein E, Allis CD (2005) RNA meets chromatin. *Genes Dev* **19**: 1635-1655.

Berry RO (1941) Chromosome behavior in the germ cells and development of the gonads in Sciara ocellaris. *Journal of Morphology* **68**: 547-583.

Beuchle D, Struhl G, Muller J (2001) Polycomb group proteins and heritable silencing of Drosophila Hox genes. *Development* **128**: 993-1004.

Biessmann H, Prasad S, Semeshin VF, Andreyeva EN, Nguyen Q, Walter MF, Mason JM (2005) Two distinct domains in Drosophila melanogaster telomeres. *Genetics* **171**: 1767-1777.

Bird A (2002) DNA methylation patterns and epigenetic memory. *Genes Dev* **16**: 6-21.

Blastyak A, Mishra RK, Karch F, Gyurkovics H (2006) Efficient and specific targeting of Polycomb group proteins requires cooperative interaction between Grainyhead and Pleiohomeotic. *Mol Cell Biol* **26**: 1434-1444.

Boivin A, Gally C, Netter S, Anxolabehere D, Ronsseray S (2003) Telomeric associated sequences of Drosophila recruit polycomb-group proteins in vivo and can induce pairing-sensitive repression. *Genetics* **164**: 195-208.

Bongiorni S, Cintio O, Prantera G (1999) The relationship between DNA methylation and chromosome imprinting in the coccid Planococcus citri. *Genetics* **151**: 1471-1478.

Bongiorni S, Mazzuoli M, Masci S, Prantera G (2001) Facultative heterochromatization in parahaploid male mealybugs: involvement of a heterochromatin-associated protein. *Development* **128**: 3809-3817.

Bongiorni S, Pasqualini B, Taranta M, Singh PB, Prantera G (2007) Epigenetic regulation of facultative heterochromatinisation in Planococcus citri via the Me(3)K9H3-HP1-Me(3)K20H4 pathway. *J Cell Sci* **120**: 1072-1080.

Bongiorni S, Prantera G (2003) Imprinted facultative heterochromatization in mealybugs. *Genetica* **117**: 271-279.

Brown JL, Mucci D, Whiteley M, Dirksen ML, Kassis JA (1998) The Drosophila Polycomb group gene pleiohomeotic encodes a DNA binding protein with homology to the transcription factor YY1. *Mol Cell* **1**: 1057-1064.

Brown SW, Nelson-Rees WA (1961) Radiation Analysis of a Lecanoid Genetic System. *Genetics* **46**: 983-1007.

Brzeska K, Brzeski J, Smith J, Chandler VL (2010) Transgenic expression of CBBP, a CXC domain protein, establishes paramutation in maize. *Proc Natl Acad Sci U S A* **107**: 5516-5521.

Buglia G, Predazzi V, Ferraro M (1999) Cytosine methylation is not involved in the heterochromatization of the paternal genome of mealybug Planococcus citri. *Chromosome Res* **7**: 71-73.

Buhler M, Verdel A, Moazed D (2006) Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing. *Cell* **125**: 873-886.

Campbell JL, Nash HA (2001) Volatile general anesthetics reveal a neurobiological role for the white and brown genes of Drosophila melanogaster. *J Neurobiol* **49**: 339-349.

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

Cao X, Jacobsen SE (2002a) Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. *Proc Natl Acad Sci U S A* **99 Suppl 4**: 16491-16498.

Cao X, Jacobsen SE (2002b) Role of the arabidopsis DRM methyltransferases in de novo DNA methylation and gene silencing. *Curr Biol* **12**: 1138-1144.

Cattanach BM, Kirk M (1985) Differential activity of maternally and paternally derived chromosome regions in mice. *Nature* **315**: 496-498.

Cerutti H, Casas-Mollano JA (2006) On the origin and functions of RNA-mediated silencing: from protists to man. *Curr Genet* **50**: 81-99.

Chan SW, Henderson IR, Jacobsen SE (2005) Gardening the genome: DNA methylation in Arabidopsis thaliana. *Nat Rev Genet* **6**: 351-360.

Chan SW, Zhang X, Bernatavichute YV, Jacobsen SE (2006) Two-step recruitment of RNA-directed DNA methylation to tandem repeats. *PLoS Biol* **4**: e363.

Chan SW, Zilberman D, Xie Z, Johansen LK, Carrington JC, Jacobsen SE (2004) RNA silencing genes control de novo DNA methylation. *Science* **303**: 1336.

Chandler V, Alleman M (2008) Paramutation: epigenetic instructions passed across generations. *Genetics* **178**: 1839-1844.

Chandler VL (2007) Paramutation: from maize to mice. *Cell* **128**: 641-645.

Chandler VL, Radicella JP, Robbins TP, Chen J, Turks D (1989) Two regulatory genes of the maize anthocyanin pathway are homologous: isolation of B utilizing R genomic sequences. *Plant Cell* **1**: 1175-1183.

Chandler VL, Stam M (2004) Chromatin conversations: mechanisms and implications of paramutation. *Nat Rev Genet* **5**: 532-544.

Chanvivattana Y, Bishopp A, Schubert D, Stock C, Moon YH, Sung ZR, Goodrich J (2004) Interaction of Polycomb-group proteins controlling flowering in Arabidopsis. *Development* **131**: 5263-5276.

Cho Y, Griswold A, Campbell C, Min KT (2005) Individual histone deacetylases in Drosophila modulate transcription of distinct genes. *Genomics* **86**: 606-617.

Choi Y, Gehring M, Johnson L, Hannon M, Harada JJ, Goldberg RB, Jacobsen SE, Fischer RL (2002) DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in arabidopsis. *Cell* **110**: 33-42.

Ciciotte SL, Gwynn B, Moriyama K, Huizing M, Gahl WA, Bonifacino JS, Peters LL (2003) Cappuccino, a mouse model of Hermansky-Pudlak syndrome, encodes a novel protein that is part of the pallidin-muted complex (BLOC-1). *Blood* **101**: 4402-4407.

Coe EH (1966) The Properties, Origin, and Mechanism of Conversion-Type Inheritance at the B Locus in Maize. *Genetics* **53**: 1035-1063.

Constancia M, Dean W, Lopes S, Moore T, Kelsey G, Reik W (2000) Deletion of a silencer element in Igf2 results in loss of imprinting independent of H19. *Nature Genetics* **26**: 203-206.

Cowell IG, Aucott R, Mahadevaiah SK, Burgoyne PS, Huskisson N, Bongiorni S, Prantera G, Fanti L, Pimpinelli S, Wu R, Gilbert DM, Shi W, Fundele R, Morrison H, Jeppesen P, Singh PB (2002) Heterochromatin, HP1 and methylation at lysine 9 of histone H3 in animals. *Chromosoma* **111**: 22-36.

Crouse HV (1960) The Controlling Element in Sex Chromosome Behavior in Sciara. *Genetics* **45**: 1429-1443.

Crouse HV (1979) X heterochromatin subdivision and cytogenetic analysis in Sciara coprophila (diptera, sciaridae). *Chromosoma* **74**: 219-239.

Crouse HV, Gerbi SA, Liang CM, Magnus L, Mercer IM (1977) Localization of ribosomal DNA within the proximal X heterochromatin of Sciara coprophila (Diptera, Sciaridae). *Chromosoma* **64**: 305-318.

Cryderman DE, Morris EJ, Biessmann H, Elgin SC, Wallrath LL (1999) Silencing at Drosophila telomeres: nuclear organization and chromatin structure play critical roles. *Embo J* **18**: 3724-3735.

Csink AK, Bounoutas A, Griffith ML, Sabl JF, Sage BT (2002) Differential gene silencing by trans-heterochromatin in Drosophila melanogaster. *Genetics* **160**: 257-269.

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

Danilevskaya ON, Hermon P, Hantke S, Muszynski MG, Kollipara K, Ananiev EV (2003) Duplicated fie genes in maize: expression pattern and imprinting suggest distinct functions. *Plant Cell* **15**: 425-438.

Davies K, Bowden L, Smith P, Dean W, Hill D, Furuumi H, Sasaki H, Cattanach B, Reik W (2002) Disruption of mesodermal enhancers for Igf2 in the minute mutant. *Development* **129**: 1657-1668.

Davies W, Isles A, Smith R, Karunadasa D, Burrmann D, Humby T, Ojarikre O, Biggin C, Skuse D, Burgoyne P, Wilkinson L (2005) Xlr3b is a new imprinted candidate for X-linked parent-of-origin effects on cognitive function in mice. *Nat Genet* **37**: 625-629.

Day T, Bonduriansky R (2004) Intralocus sexual conflict can drive the evolution of genomic imprinting. *Genetics* **167**: 1537-1546.

De Rubertis F, Kadosh D, Henchoz S, Pauli D, Reuter G, Struhl K, Spierer P (1996) The histone deacetylase RPD3 counteracts genomic silencing in Drosophila and yeast. *Nature* **384**: 589-591.

DeChiara TM, Robertson EJ, Efstratiadis A (1991) Parental imprinting of the mouse insulin-like growth factor II gene. *Cell* **64**: 849-859.

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

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

Dell'Angelica EC (2004) The building BLOC(k)s of lysosomes and related organelles. *Curr Opin Cell Biol* **16**: 458-464.

Dell'Angelica EC, Shotelersuk V, Aguilar RC, Gahl WA, Bonifacino JS (1999) Altered trafficking of lysosomal proteins in Hermansky-Pudlak syndrome due to mutations in the beta 3A subunit of the AP-3 adaptor. *Mol Cell* **3**: 11-21.

Deng H, Bao X, Cai W, Blacketer MJ, Belmont AS, Girton J, Johansen J, Johansen KM (2008) Ectopic histone H3S10 phosphorylation causes chromatin structure remodeling in Drosophila. *Development* **135**: 699-705.

Deobagkar D, Muralidharan K, Devare S, Kalghatgi K, Chandra H (1982) The mealybug chromosome system I: Unusual methylated bases and dinucleotides in DNA of a Planococcus species. *Journal of Biosciences* **4**: 513-526.

Detter JC, Zhang Q, Mules EH, Novak EK, Mishra VS, Li W, McMurtrie EB, Tchernev VT, Wallace MR, Seabra MC, Swank RT, Kingsmore SF (2000) Rab geranylgeranyl transferase alpha mutation in the gunmetal mouse reduces Rab prenylation and platelet synthesis. *Proc Natl Acad Sci U S A* **97**: 4144-4149.

Devajyothi C, Brahmachari V (1992) Detection of a CpA methylase in an insect system: characterization and substrate specificity. *Mol Cell Biochem* **110**: 103-111.

DeVido SK, Kwon D, Brown JL, Kassis JA (2008) The role of Polycomb-group response elements in regulation of engrailed transcription in Drosophila. *Development* **135**: 669-676.

Di Pietro SM, Dell'Angelica EC (2005) The cell biology of Hermansky-Pudlak syndrome: recent advances. *Traffic* **6**: 525-533.

Di Pietro SM, Falcon-Perez JM, Dell'Angelica EC (2004) Characterization of BLOC-2, a complex containing the Hermansky-Pudlak syndrome proteins HPS3, HPS5 and HPS6. *Traffic* **5**: 276-283.

Djupedal I, Ekwall K (2009) Epigenetics: heterochromatin meets RNAi. *Cell Res* **19**: 282-295.

Doheny JG, Mottus R, Grigliatti TA (2008) Telomeric position effect--a third silencing mechanism in eukaryotes. *PLoS One* **3**: e3864.

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

Dorweiler JE, Carey CC, Kubo KM, Hollick JB, Kermicle JL, Chandler VL (2000) mediator of paramutation1 is required for establishment and maintenance of paramutation at multiple maize loci. *Plant Cell* **12**: 2101-2118.

Dreesen TD, Henikoff S, Loughney K (1991) A pairing-sensitive element that mediates trans-inactivation is associated with the Drosophila brown gene. *Genes Dev* **5**: 331-340.

Drewell RA, Arney KL, Arima T, Barton SC, Brenton JD, Surani MA (2002) Novel conserved elements upstream of the H19 gene are transcribed and act as mesodermal enhancers. *Development* **129**: 1205-1213.

Drewell RA, Brenton JD, Ainscough JF, Barton SC, Hilton KJ, Arney KL, Dandolo L, Surani MA (2000) Deletion of a silencer element disrupts H19 imprinting independently of a DNA methylation epigenetic switch. *Development* **127**: 3419-3428.

Duncan IW (2002) Transvection effects in Drosophila. *Annu Rev Genet* **36**: 521-556.

Duncan L, Lewis EB (1982) Genetic control of body segment differentiation in Drosophila. In *Developmental Order: Its Origin and Regulation*, Subtelny S (ed), pp 533-554. New York: Liss.

Ebert A, Schotta G, Lein S, Kubicek S, Krauss V, Jenuwein T, Reuter G (2004) Su(var) genes regulate the balance between euchromatin and heterochromatin in Drosophila. *Genes Dev* **18**: 2973-2983.

Eden S, Constancia M, Hashimshony T, Dean W, Goldstein B, Johnson AC, Keshet I, Reik W, Cedar H (2001) An upstream repressor element plays a role in Igf2 imprinting. *Embo J* **20**: 3518-3525.

Engel N, Thorvaldsen JL, Bartolomei MS (2006) CTCF binding sites promote transcription initiation and prevent DNA methylation on the maternal allele at the imprinted H19/Igf2 locus. *Hum Mol Genet* **15**: 2945-2954.

Epstein H, James TC, Singh PB (1992) Cloning and expression of Drosophila HP1 homologs from a mealybug, Planococcus citri. *J Cell Sci* **101 (Pt 2)**: 463-474.

Erhard KF, Jr., Stonaker JL, Parkinson SE, Lim JP, Hale CJ, Hollick JB (2009) RNA polymerase IV functions in paramutation in Zea mays. *Science* **323**: 1201-1205.

Erhardt S, Lyko F, Ainscough JF, Surani MA, Paro R (2003) Polycomb-group proteins are involved in silencing processes caused by a transgenic element from the murine imprinted H19/Igf2 region in Drosophila. *Dev Genes Evol* **213**: 336-344.

Faehnle CR, Joshua-Tor L (2007) Argonautes confront new small RNAs. *Curr Opin Chem Biol* **11**: 569-577.

Falcon-Perez JM, Romero-Calderon R, Brooks ES, Krantz DE, Dell'Angelica EC (2007) The Drosophila pigmentation gene pink (p) encodes a homologue of human Hermansky-Pudlak syndrome 5 (HPS5). *Traffic* **8**: 154-168.

Ferraro M, Buglia GL, Romano F (2001) Involvement of histone H4 acetylation in the epigenetic inheritance of different activity states of maternally and paternally derived genomes in the mealybug Planococcus citri. *Chromosoma* **110**: 93-101.

Field LM, Lyko F, Mandrioli M, Prantera G (2004) DNA methylation in insects. *Insect Mol Biol* **13**: 109-115.

Finney M, Ruvkun G, Horvitz HR (1988) The C. elegans cell lineage and differentiation gene unc-86 encodes a protein with a homeodomain and extended similarity to transcription factors. *Cell* **55**: 757-769.

Fitch KR, Yasuda GK, Owens KN, Wakimoto BT (1998) Paternal effects in Drosophila: implications for mechanisms of early development. *Curr Top Dev Biol* **38**: 1-34.

Francis NJ, Kingston RE, Woodcock CL (2004) Chromatin compaction by a polycomb group protein complex. *Science* **306**: 1574-1577.

Francis NJ, Saurin AJ, Shao Z, Kingston RE (2001) Reconstitution of a functional core polycomb repressive complex. *Mol Cell* **8**: 545-556.

Frydrychova RC, Biessmann H, Konev AY, Golubovsky MD, Johnson J, Archer TK, Mason JM (2007) Transcriptional activity of the telomeric retrotransposon HeT-A in Drosophila melanogaster is stimulated as a consequence of subterminal deficiencies at homologous and nonhomologous telomeres. *Mol Cell Biol* **27**: 4991-5001.

Fujimoto R, Kinoshita Y, Kawabe A, Kinoshita T, Takashima K, Nordborg M, Nasrallah ME, Shimizu KK, Kudoh H, Kakutani T (2008) Evolution and control of imprinted FWA genes in the genus Arabidopsis. *PLoS Genet* **4**: e1000048.

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

Fuks F (2005) DNA methylation and histone modifications: teaming up to silence genes. *Curr Opin Genet Dev* **15**: 490-495.

Fuyama Y (1984) Gynogenesis in Drosophila Melanogaster. *Jpn. J. Genet.* **59**: 91-96.

Fuyama Y (1986) Genetics of Parthenogenesis in DROSOPHILA MELANOGASTER. II. Characterization of a Gynogenetically Reproducing Strain. *Genetics* **114**: 495-509.

Gehring M, Huh JH, Hsieh TF, Penterman J, Choi Y, Harada JJ, Goldberg RB, Fischer RL (2006) DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation. *Cell* **124**: 495-506.

Goday C, Esteban MR (2001) Chromosome elimination in sciarid flies. *Bioessays* **23**: 242-250.

Goday C, Ruiz MF (2002) Differential acetylation of histones H3 and H4 in paternal and maternal germline chromosomes during development of sciarid flies. *J Cell Sci* **115**: 4765-4775.

Golic KG, Golic MM, Pimpinelli S (1998) Imprinted control of gene activity in Drosophila. *Curr Biol* **8**: 1273-1276.

Golic KG, Lindquist S (1989) The FLP recombinase of yeast catalyzes site-specific recombination in the Drosophila genome. *Cell* **59**: 499-509.

Greciano PG, Goday C (2006) Methylation of histone H3 at Lys4 differs between paternal and maternal chromosomes in Sciara ocellaris germline development. *J Cell Sci* **119**: 4667-4677.

Grewal SI, Moazed D (2003) Heterochromatin and epigenetic control of gene expression. *Science* **301**: 798-802.

Grimaud C, Bantignies F, Pal-Bhadra M, Ghana P, Bhadra U, Cavalli G (2006a) RNAi components are required for nuclear clustering of Polycomb group response elements. *Cell* **124**: 957-971.

Grimaud C, Negre N, Cavalli G (2006b) From genetics to epigenetics: the tale of Polycomb group and trithorax group genes. *Chromosome Res* **14**: 363-375.

Grossniklaus U, Vielle-Calzada JP, Hoeppner MA, Gagliano WB (1998) Maternal control of embryogenesis by MEDEA, a polycomb group gene in Arabidopsis. *Science* **280**: 446-450.

Gruenbaum Y, Naveh-Many T, Cedar H, Razin A (1981) Sequence specificity of methylation in higher plant DNA. *Nature* **292**: 860-862.

Grumbling G, Strelets V (2006) FlyBase: anatomical data, images and queries. *Nucleic Acids Res* **34**: D484-488.

Gutierrez-Marcos JF, Costa LM, Pra MD, Scholten S, Kranz E, Perez P, Dickinson HG (2006) Epigenetic asymmetry of imprinted genes in plant gametes. *Nature Genetics* **38**: 876-878.

Gutierrez-Marcos JF, Pennington PD, Costa LM, Dickinson HG (2003) Imprinting in the endosperm: a possible role in preventing wide hybridization. *Philos Trans R Soc Lond B Biol Sci* **358**: 1105-1111.

Haines TR, Rodenhiser DI, Ainsworth PJ (2001) Allele-specific non-CpG methylation of the Nf1 gene during early mouse development. *Dev Biol* **240**: 585-598.

Haller BS, Woodruff RC (2000) Varied expression of a Y-linked P[w+] insert due to imprinting in Drosophila melanogaster. *Genome* **43**: 285-292.

Han L, Lee DH, Szabo PE (2008) CTCF is the master organizer of domain-wide allele-specific chromatin at the H19/Igf2 imprinted region. *Mol Cell Biol* **28**: 1124-1135.

Haring M, Bader R, Louwers M, Schwabe A, van Driel R, Stam M (2010) The role of DNA methylation, nucleosome occupancy and histone modifications in paramutation. *Plant J* **Apr 29**: [Epub ahead of print].

Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM (2000) CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* **405**: 486-489.

Haun WJ, Springer NM (2008) Maternal and paternal alleles exhibit differential histone methylation and acetylation at maize imprinted genes. *Plant J* **56**: 903-912.

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

Haynes KA, Gracheva E, Elgin SC (2007) A Distinct type of heterochromatin within Drosophila melanogaster chromosome 4. *Genetics* **175**: 1539-1542.

Hermansky F, Pudlak P (1959) Albinism associated with hemorrhagic diathesis and unusual pigmented reticular cells in the bone marrow: report of two cases with histochemical studies. *Blood* **14**: 162-169.

Herr AJ, Jensen MB, Dalmay T, Baulcombe DC (2005) RNA polymerase IV directs silencing of endogenous DNA. *Science* **308**: 118-120.

Hollick JB, Kermicle JL, Parkinson SE (2005) Rmr6 maintains meiotic inheritance of paramutant states in Zea mays. *Genetics* **171**: 725-740.

Holliday R (1990) Genomic imprinting and allelic exclusion. *Dev Suppl*: 125-129.

Huang AM, Rehm EJ, Rubin GM (2000) Recovery of DNA Sequences Flanking P-element Insertions: Inverse PCR and Plasmid Rescue. In *Drosophila Protocols*, Sullivan W, Ashburner M, Hawley RS (eds) pp 429-439. Cold Spring Harbor: CSHL Press.

Huizing M, Boissy RE, Gahl WA (2002a) Hermansky-Pudlak syndrome: vesicle formation from yeast to man. *Pigment Cell Res* **15**: 405-419.

Huizing M, Gahl WA (2002) Disorders of vesicles of lysosomal lineage: the Hermansky-Pudlak syndromes. *Curr Mol Med* **2**: 451-467.

Huizing M, Hess R, Dorward H, Claassen DA, Helip-Wooley A, Kleta R, Kaiser-Kupfer MI, White JG, Gahl WA (2004) Cellular, molecular and clinical characterization of patients with Hermansky-Pudlak syndrome type 5. *Traffic* **5**: 711-722.

Huizing M, Scher CD, Strovel E, Fitzpatrick DL, Hartnell LM, Anikster Y, Gahl WA (2002b) Nonsense mutations in ADTB3A cause complete deficiency of the beta3A subunit of adaptor complex-3 and severe Hermansky-Pudlak syndrome type 2. *Pediatr Res* **51**: 150-158.

Irvine DV, Zaratiegui M, Tolia NH, Goto DB, Chitwood DH, Vaughn MW, Joshua-Tor L, Martienssen RA (2006) Argonaute slicing is required for heterochromatic silencing and spreading. *Science* **313**: 1134-1137.

Iwasa Y, Pomiankowski A (1999) Sex specific X chromosome expression caused by genomic imprinting. *J Theor Biol* **197**: 487-495.

Joanis V, Lloyd VK (2002) Genomic imprinting in Drosophila is maintained by the products of Suppressor of variegation and trithorax group, but not Polycomb group, genes. *Mol Genet Genomics* **268**: 103-112.

Jones BK, Levorse J, Tilghman SM (2001) Deletion of a nuclease-sensitive region between the Igf2 and H19 genes leads to Igf2 misregulation and increased adiposity. *Hum Mol Genet* **10**: 807-814.

Jones WK, Rawls JM, Jr. (1988) Genetic and molecular mapping of chromosome region 85A in Drosophila melanogaster. *Genetics* **120**: 733-742.

Josse T, Teysset L, Todeschini AL, Sidor CM, Anxolabehere D, Ronsseray S (2007) Telomeric trans-silencing: an epigenetic repression combining RNA silencing and heterochromatin formation. *PLoS Genet* **3**: 1633-1643.

Jowett T, Wajidi MF, Oxtoby E, Wolf CR (1991) Mammalian genes expressed in Drosophila: a transgenic model for the study of mechanisms of chemical mutagenesis and metabolism. *Embo J* **10**: 1075-1081.

Jullien PE, Berger F (2009) Gamete-specific epigenetic mechanisms shape genomic imprinting. *Curr Opin Plant Biol* **12**: 637-642.

Jullien PE, Katz A, Oliva M, Ohad N, Berger F (2006a) Polycomb group complexes self-regulate imprinting of the Polycomb group gene MEDEA in Arabidopsis. *Curr Biol* **16**: 486-492.

Jullien PE, Kinoshita T, Ohad N, Berger F (2006b) Maintenance of DNA methylation during the Arabidopsis life cycle is essential for parental imprinting. *Plant Cell* **18**: 1360-1372.

Kaffer CR, Srivastava M, Park KY, Ives E, Hsieh S, Batlle J, Grinberg A, Huang SP, Pfeifer K (2000) A transcriptional insulator at the imprinted H19/Igf2 locus. *Genes Dev* **14**: 1908-1919.

Kalfayan L, Wensink PC (1982) Developmental regulation of Drosophila alpha-tubulin genes. *Cell* **29**: 91-98.

Kanno T, Huettel B, Mette MF, Aufsatz W, Jaligot E, Daxinger L, Kreil DP, Matzke M, Matzke AJ (2005) Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nat Genet* **37**: 761-765.

Kantheti P, Qiao X, Diaz ME, Peden AA, Meyer GE, Carskadon SL, Kapfhamer D, Sufalko D, Robinson MS, Noebels JL, Burmeister M (1998) Mutation in AP-3 delta in the mocha mouse links endosomal transport to storage deficiency in platelets, melanosomes, and synaptic vesicles. *Neuron* **21**: 111-122.

Kassis JA (2002) Pairing-sensitive silencing, polycomb group response elements, and transposon homing in Drosophila. *Adv Genet* **46**: 421-438.

Kavi HH, Fernandez H, Xie W, Birchler JA (2008) Genetics and biochemistry of RNAi in Drosophila. *Curr Top Microbiol Immunol* **320**: 37-75.

Kavi HH, Fernandez HR, Xie W, Birchler JA (2006) Polycomb, pairing and PIWI--RNA silencing and nuclear interactions. *Trends Biochem Sci* **31**: 485-487.

Khatib H, Zaitoun I, Kim ES (2007) Comparative analysis of sequence characteristics of imprinted genes in human, mouse, and cattle. *Mamm Genome* **18**: 538-547.

Khesin RB, Leibovitch BA (1978) Influence of deficiency of the histone gene-containing 38B-40 region on X-chromosome template activity and the white gene position effect variegation in Drosophila melanogaster. *Mol Gen Genet* **162**: 323-328.

Khosla S, Augustus M, Brahmachari V (1999) Sex-specific organisation of middle repetitive DNA sequences in the mealybug Planococcus lilacinus. *Nucleic Acids Res* **27**: 3745-3751.

Khosla S, Kantheti P, Brahmachari V, Chandra HS (1996) A male-specific nuclease-resistant chromatin fraction in the mealybug Planococcus lilacinus. *Chromosoma* **104**: 386-392.

Khosla S, Mendiratta G, Brahmachari V (2006) Genomic imprinting in the mealybugs. *Cytogenetic & Genome Research* **113**: 41-52.

Kinoshita T, Miura A, Choi Y, Kinoshita Y, Cao X, Jacobsen SE, Fischer RL, Kakutani T (2004) One-way control of FWA imprinting in Arabidopsis endosperm by DNA methylation. *Science* **303**: 521-523.

Kinoshita T, Yadegari R, Harada JJ, Goldberg RB, Fischer RL (1999) Imprinting of the MEDEA polycomb gene in the Arabidopsis endosperm. *Plant Cell* **11**: 1945-1952.

Kiyosue T, Ohad N, Yadegari R, Hannon M, Dinneny J, Wells D, Katz A, Margossian L, Harada JJ, Goldberg RB, Fischer RL (1999) Control of fertilization-independent endosperm development by the MEDEA polycomb gene in Arabidopsis. *Proc Natl Acad Sci U S A* **96**: 4186-4191.

Klemenz R, Weber U, Gehring WJ (1987) The white gene as a marker in a new P-element vector for gene transfer in Drosophila. *Nucleic Acids Res* **15**: 3947-3959.

Klenov MS, Lavrov SA, Stolyarenko AD, Ryazansky SS, Aravin AA, Tuschl T, Gvozdev VA (2007) Repeat-associated siRNAs cause chromatin silencing of retrotransposons in the Drosophila melanogaster germline. *Nucleic Acids Res* **35**: 5430-5438.

Kloc A, Zaratiegui M, Nora E, Martienssen R (2008) RNA interference guides histone modification during the S phase of chromosomal replication. *Curr Biol* **18**: 490-495.

Klucken J, Buchler C, Orso E, Kaminski WE, Porsch-Ozcurumez M, Liebisch G, Kapinsky M, Diederich W, Drobnik W, Dean M, Allikmets R, Schmitz G (2000) ABCG1 (ABC8), the human homolog of the Drosophila white gene, is a regulator of macrophage cholesterol and phospholipid transport. *Proc Natl Acad Sci U S A* **97**: 817-822.

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

Kohler C, Hennig L, Bouveret R, Gheyselinck J, Grossniklaus U, Gruissem W (2003a) Arabidopsis MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed development. *Embo J* **22**: 4804-4814.

Kohler C, Hennig L, Spillane C, Pien S, Gruissem W, Grossniklaus U (2003b) The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene PHERES1. *Genes Dev* **17**: 1540-1553.

Kohler C, Page DR, Gagliardini V, Grossniklaus U (2005) The Arabidopsis thaliana MEDEA Polycomb group protein controls expression of PHERES1 by parental imprinting. *Nat Genet* **37**: 28-30.

Koide T, Ainscough J, Wijgerde M, Surani MA (1994) Comparative analysis of Igf-2/H19 imprinted domain: identification of a highly conserved intergenic DNase I hypersensitive region. *Genomics* **24**: 1-8.

Komma DJ, Endow SA (1995) Haploidy and androgenesis in Drosophila. *Proc Natl Acad Sci U S A* **92**: 11884-11888.

Kono T (2006) Genomic imprinting is a barrier to parthenogenesis in mammals. *Cytogenetic & Genome Research* **113**: 31-35.

Kono T, Obata Y, Wu Q, Niwa K, Ono Y, Yamamoto Y, Park ES, Seo JS, Ogawa H (2004) Birth of parthenogenetic mice that can develop to adulthood. *Nature* **428**: 860-864.

Kourmouli N, Jeppesen P, Mahadevhaiah S, Burgoyne P, Wu R, Gilbert DM, Bongiorni S, Prantera G, Fanti L, Pimpinelli S, Shi W, Fundele R, Singh PB (2004) Heterochromatin and tri-methylated lysine 20 of histone H4 in animals. *J Cell Sci* **117**: 2491-2501.

Kretzschmar D, Poeck B, Roth H, Ernst R, Keller A, Porsch M, Strauss R, Pflugfelder GO (2000) Defective pigment granule biogenesis and aberrant behavior caused by mutations in the Drosophila AP-3beta adaptin gene ruby. *Genetics* **155**: 213-223.

Kubai DF (1987) Nonrandom chromosome arrangements in germ line nuclei of Sciara coprophila males: the basis for nonrandom chromosome segregation on the meiosis I spindle. *J Cell Biol* **105**: 2433-2446.

Kurenova E, Champion L, Biessmann H, Mason JM (1998) Directional gene silencing induced by a complex subtelomeric satellite from Drosophila. *Chromosoma* **107**: 311-320.

Kurukuti S, Tiwari VK, Tavoosidana G, Pugacheva E, Murrell A, Zhao Z, Lobanenkov V, Reik W, Ohlsson R (2006) CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. *Proc Natl Acad Sci U S A* **103**: 10684-10689.

Kusch T, Florens L, Macdonald WH, Swanson SK, Glaser RL, Yates JR, 3rd, Abmayr SM, Washburn MP, Workman JL (2004) Acetylation by Tip60 is required for selective histone variant exchange at DNA lesions. *Science* **306**: 2084-2087.

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

Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**: 116-120.

LaJeunesse D, Shearn A (1996) E(z): a polycomb group gene or a trithorax group gene? *Development* **122**: 2189-2197.

Lehming N, Thanos D, Brickman JM, Ma J, Maniatis T, Ptashne M (1994) An HMG-like protein that can switch a transcriptional activator to a repressor. *Nature* **371**: 175-179.

Leighton PA, Saam JR, Ingram RS, Stewart CL, Tilghman SM (1995) An enhancer deletion affects both H19 and Igf2 expression. *Genes Dev* **9**: 2079-2089.

Li T, Hu JF, Qiu X, Ling J, Chen H, Wang S, Hou A, Vu TH, Hoffman AR (2008) CTCF regulates allelic expression of Igf2 by orchestrating a promoter-polycomb repressive complex-2 intrachromosomal loop. *Mol Cell Biol* **28**: 6473-6482.

Li W, Rusiniak ME, Chintala S, Gautam R, Novak EK, Swank RT (2004) Murine Hermansky-Pudlak syndrome genes: regulators of lysosome-related organelles. *Bioessays* **26**: 616-628.

Lippman Z, Martienssen R (2004) The role of RNA interference in heterochromatic silencing. *Nature* **431**: 364-370.

Lisch D, Carey CC, Dorweiler JE, Chandler VL (2002) A mutation that prevents paramutation in maize also reverses Mutator transposon methylation and silencing. *Proc Natl Acad Sci U S A* **99**: 6130-6135.

Lloyd V (2000) Parental imprinting in Drosophila. *Genetica* **109**: 35-44.

Lloyd V, Ramaswami M, Kramer H (1998) Not just pretty eyes: Drosophila eye-colour mutations and lysosomal delivery. *Trends Cell Biol* **8**: 257-259.

Lloyd VK, Sinclair DA, Alperyn M, Grigliatti TA (2002) Enhancer of garnet/deltaAP-3 is a cryptic allele of the white gene and identifies the intracellular transport system for the white protein. *Genome* **45**: 296-312.

Lloyd VK, Sinclair DA, Grigliatti TA (1999a) Genomic imprinting and position-effect variegation in Drosophila melanogaster. *Genetics* **151**: 1503-1516.

Lloyd VK, Sinclair DA, Wennberg R, Warner TS, Honda BM, Grigliatti TA (1999b) A genetic and molecular characterization of the garnet gene of Drosophila melanogaster. *Genome* **42**: 1183-1193.

Lo SM, Ahuja NK, Francis NJ (2009) Polycomb group protein Suppressor 2 of zeste is a functional homolog of Posterior Sex Combs. *Mol Cell Biol* **29**: 515-525.

Loftus SK, Larson DM, Baxter LL, Antonellis A, Chen Y, Wu X, Jiang Y, Bittner M, Hammer JA, 3rd, Pavan WJ (2002) Mutation of melanosome protein RAB38 in chocolate mice. *Proc Natl Acad Sci U S A* **99**: 4471-4476.

Louwers M, Bader R, Haring M, van Driel R, de Laat W, Stam M (2009) Tissue- and expression level-specific chromatin looping at maize *b1* epialleles. *Plant Cell* **21**: 832-842.

Lyko F, Brenton JD, Surani MA, Paro R (1997) An imprinting element from the mouse H19 locus functions as a silencer in Drosophila. *Nature Genetics* **16**: 171-173.

Lyko F, Buiting K, Horsthemke B, Paro R (1998) Identification of a silencing element in the human 15q11-q13 imprinting center by using transgenic Drosophila. *Proc Natl Acad Sci U S A* **95**: 1698-1702.

Lyko F, Ramsahoye BH, Jaenisch R (2000) DNA methylation in Drosophila melanogaster. *Nature* **408**: 538-540.

MacDonald WA, Lloyd VK (2004) Genomic imprinting in Drosophila and the role of non-coding RNAs. *45th Annual Drosophila Research Conference,* **Abstract 323B**

Mackenzie SM, Brooker MR, Gill TR, Cox GB, Howells AJ, Ewart GD (1999) Mutations in the white gene of Drosophila melanogaster affecting ABC transporters that determine eye colouration. *Biochim Biophys Acta* **1419**: 173-185.

Maggert KA, Golic KG (2002) The Y chromosome of Drosophila melanogaster exhibits chromosome-wide imprinting. *Genetics* **162**: 1245-1258.

Maggert KA, Golic KG (2004) DNA methylation affects genomic imprinting. *45th Annual Drosophila Research Conference,* **Abstract 133**

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

Makarevich G, Leroy O, Akinci U, Schubert D, Clarenz O, Goodrich J, Grossniklaus U, Kohler C (2006) Different Polycomb group complexes regulate common target genes in Arabidopsis. *EMBO Rep* **7**: 947-952.

Makarevich G, Villar CB, Erilova A, Kohler C (2008) Mechanism of PHERES1 imprinting in Arabidopsis. *J Cell Sci* **121**: 906-912.

Marenda DR, Zraly CB, Dingwall AK (2004) The Drosophila Brahma (SWI/SNF) chromatin remodeling complex exhibits cell-type specific activation and repression functions. *Dev Biol* **267**: 279-293.

Martin CC, McGowan R (1995) Genotype-specific modifiers of transgene methylation and expression in the zebrafish, Danio rerio. *Genet Res* **65**: 21-28.

Mason JM, Konev AY, Golubovsky MD, Biessmann H (2003) Cis- and trans-acting influences on telomeric position effect in Drosophila melanogaster detected with a subterminal transgene. *Genetics* **163**: 917-930.

Matthews KA, Miller DF, Kaufman TC (1989) Developmental distribution of RNA and protein products of the Drosophila alpha-tubulin gene family. *Dev Biol* **132**: 45-61.

Matzke MA, Birchler JA (2005) RNAi-mediated pathways in the nucleus. *Nat Rev Genet* **6**: 24-35.

Mayer C, Schmitz KM, Li J, Grummt I, Santoro R (2006) Intergenic transcripts regulate the epigenetic state of rRNA genes. *Mol Cell* **22**: 351-361.

McDonald HB, Goldstein LS (1990) Identification and characterization of a gene encoding a kinesin-like protein in Drosophila. *Cell* **61**: 991-1000.

McDonald HB, Stewart RJ, Goldstein LS (1990) The kinesin-like ncd protein of Drosophila is a minus end-directed microtubule motor. *Cell* **63**: 1159-1165.

McGinnis KM, Springer C, Lin Y, Carey CC, Chandler VL (2006) Transcriptionally Silenced Transgenes in Maize are Activated by Three Mutations Defective in Paramutation. *Genetics* **173**: 1637-1647.

McGrath J, Solter D (1984) Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell* **37**: 179-183.

Mohan KN, Chandra HS (2005) Isolation and analysis of sequences showing sex-specific cytosine methylation in the mealybug Planococcus lilacinus. *Mol Genet Genomics* **274**: 557-568.

Mohd-Sarip A, Venturini F, Chalkley GE, Verrijzer CP (2002) Pleiohomeotic can link polycomb to DNA and mediate transcriptional repression. *Mol Cell Biol* **22**: 7473-7483.

Moller A, Avila FW, Erickson JW, Jackle H (2005) Drosophila BAP60 is an essential component of the Brahma complex, required for gene activation and repression. *J Mol Biol* **352**: 329-337.

Moore T, Constancia M, Zubair M, Bailleul B, Feil R, Sasaki H, Reik W (1997) Multiple imprinted sense and antisense transcripts, differential methylation and tandem repeats in a putative imprinting control region upstream of mouse Igf2. *Proc Natl Acad Sci U S A* **94**: 12509-12514.

Moore T, Haig D (1991) Genomic imprinting in mammalian development: a parental tug-of-war. *Trends Genet* **7**: 45-49.

Morison IM, Ramsay JP, Spencer HG (2005) A census of mammalian imprinting. *Trends Genet* **21**: 457-465.

Muller J, Hart CM, Francis NJ, Vargas ML, Sengupta A, Wild B, Miller EL, O'Connor MB, Kingston RE, Simon JA (2002) Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. *Cell* **111**: 197-208.

Mullins C, Hartnell LM, Wassarman DA, Bonifacino JS (1999) Defective expression of the mu3 subunit of the AP-3 adaptor complex in the Drosophila pigmentation mutant carmine. *Mol Gen Genet* **262**: 401-412.

Murrell A, Heeson S, Bowden L, Constancia M, Dean W, Kelsey G, Reik W (2001) An intragenic methylated region in the imprinted Igf2 gene augments transcription. *EMBO Rep* **2**: 1101-1106.

Murrell A, Heeson S, Reik W (2004) Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. *Nat Genet* **36**: 889-893.

Murrell A, Ito Y, Verde G, Huddleston J, Woodfine K, Silengo MC, Spreafico F, Perotti D, De Crescenzo A, Sparago A, Cerrato F, Riccio A (2008) Distinct methylation changes at the IGF2-H19 locus in congenital growth disorders and cancer. *PLoS ONE* **3**: e1849.

Nakatani Y, Nakamura N, Sano J, Inayama Y, Kawano N, Yamanaka S, Miyagi Y, Nagashima Y, Ohbayashi C, Mizushima M, Manabe T, Kuroda M, Yokoi T, Matsubara O (2000) Interstitial pneumonia in Hermansky-Pudlak syndrome: significance of florid foamy swelling/degeneration (giant lamellar body degeneration) of type-2 pneumocytes. *Virchows Arch* **437**: 304-313.

Nazarian R, Starcevic M, Spencer MJ, Dell'Angelica EC (2006) Reinvestigation of the dysbindin subunit of BLOC-1 (biogenesis of lysosome-related organelles complex-1) as a dystrobrevin-binding protein. *Biochem J* **395**: 587-598.

Neal SJ, Gibson ML, So AK, Westwood JT (2003) Construction of a cDNA-based microarray for Drosophila melanogaster: a comparison of gene transcription profiles from SL2 and Kc167 cells. *Genome* **46**: 879-892.

Nobuta K, Lu C, Shrivastava R, Pillay M, De Paoli E, Accerbi M, Arteaga-Vazquez M, Sidorenko L, Jeong DH, Yen Y, Green PJ, Chandler VL, Meyers BC (2008) Distinct size distribution of endogeneous siRNAs in maize: Evidence from deep sequencing in the mop1-1 mutant. *Proc Natl Acad Sci U S A* **105**: 14958-14963.

Novitski E, Grace D, Strommen C (1981) The entire compound autosomes of Drosophila melanogaster. *Genetics* **98**: 257-273.

Nur U (1967) Reversal of heterochromatization and the activity of the paternal chromosome set in the male mealy bug. *Genetics* **56**: 375-389.

Odorizzi G, Cowles CR, Emr SD (1998) The AP-3 complex: a coat of many colours. *Trends Cell Biol* **8**: 282-288.

Ohlsson R, Paldi A, Graves JA (2001) Did genomic imprinting and X chromosome inactivation arise from stochastic expression? *Trends Genet* **17**: 136-141.

Okamura K, Ito T (2006) Lessons from comparative analysis of species-specific imprinted genes. *Cytogenet Genome Res* **113**: 159-164.

Onodera Y, Haag JR, Ream T, Nunes PC, Pontes O, Pikaard CS (2005) Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* **120**: 613-622.

Ooi CE, Moreira JE, Dell'Angelica EC, Poy G, Wassarman DA, Bonifacino JS (1997) Altered expression of a novel adaptin leads to defective pigment granule biogenesis in the Drosophila eye color mutant garnet. *Embo J* **16**: 4508-4518.

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

Papoulas O, Beek SJ, Moseley SL, McCallum CM, Sarte M, Shearn A, Tamkun JW (1998) The Drosophila trithorax group proteins BRM, ASH1 and ASH2 are subunits of distinct protein complexes. *Development* **125**: 3955-3966.

Patterson GI, Thorpe CJ, Chandler VL (1993) Paramutation, an allelic interaction, is associated with a stable and heritable reduction of transcription of the maize b regulatory gene. *Genetics* **135**: 881-894.

Peters J, Beechey C (2004) Identification and characterisation of imprinted genes in the mouse. *Brief Funct Genomic Proteomic* **2**: 320-333.

Petruk S, Sedkov Y, Smith S, Tillib S, Kraevski V, Nakamura T, Canaani E, Croce CM, Mazo A (2001) Trithorax and dCBP acting in a complex to maintain expression of a homeotic gene. *Science* **294**: 1331-1334.

Pien S, Grossniklaus U (2007) Polycomb group and trithorax group proteins in Arabidopsis. *Biochim Biophys Acta* **1769**: 375-382.

Pikaard CS, Haag JR, Ream T, Wierzbicki AT (2008) Roles of RNA polymerase IV in gene silencing. *Trends Plant Sci* **13**: 390-397.

Pile LA, Wassarman DA (2000) Chromosomal localization links the SIN3-RPD3 complex to the regulation of chromatin condensation, histone acetylation and gene expression. *Embo J* **19**: 6131-6140.

Pimpinelli S, Berloco M, Fanti L, Dimitri P, Bonaccorsi S, Marchetti E, Caizzi R, Caggese C, Gatti M (1995) Transposable elements are stable structural components of Drosophila melanogaster heterochromatin. *Proc Natl Acad Sci U S A* **92**: 3804-3808.

Plagge A, Isles AR, Gordon E, Humby T, Dean W, Gritsch S, Fischer-Colbrie R, Wilkinson LS, Kelsey G (2005) Imprinted Nesp55 influences behavioral reactivity to novel environments. *Mol Cell Biol* **25**: 3019-3026.

Pontier D, Yahubyan G, Vega D, Bulski A, Saez-Vasquez J, Hakimi MA, Lerbs-Mache S, Colot V, Lagrange T (2005) Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in Arabidopsis. *Genes Dev* **19**: 2030-2040.

Qi D, Jin H, Lilja T, Mannervik M (2006) Drosophila Reptin and other TIP60 complex components promote generation of silent chromatin. *Genetics* **174**: 241-251.

Ramsahoye BH, Biniszkiewicz D, Lyko F, Clark V, Bird AP, Jaenisch R (2000) Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc Natl Acad Sci U S A* **97**: 5237-5242.

Rassoulzadegan M, Grandjean V, Gounon P, Cuzin F (2007) Inheritance of an epigenetic change in the mouse: a new role for RNA. *Biochem Soc Trans* **35**: 623-625.

Rieffel SM, Crouse HV (1966) The elimination and differentiation of chromosomes in the germ line of sciara. *Chromosoma* **19**: 231-276.

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

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

Robertson HM, Preston CR, Phillis RW, Johnson-Schlitz DM, Benz WK, Engels WR (1988) A stable genomic source of P element transposase in Drosophila melanogaster. *Genetics* **118**: 461-470.

Roche SE, Rio DC (1998) Trans-silencing by P elements inserted in subtelomeric heterochromatin involves the Drosophila Polycomb group gene, Enhancer of zeste. *Genetics* **149**: 1839-1855.

Ronsseray S, Josse T, Boivin A, Anxolabehere D (2003) Telomeric transgenes and trans-silencing in Drosophila. *Genetica* **117**: 327-335.

Ronsseray S, Marin L, Lehmann M, Anxolabehere D (1998) Repression of hybrid dysgenesis in Drosophila melanogaster by combinations of telomeric P-element reporters and naturally occurring P elements. *Genetics* **149**: 1857-1866.

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.

Rundlett SE, Carmen AA, Kobayashi R, Bavykin S, Turner BM, Grunstein M (1996) HDA1 and RPD3 are members of distinct yeast histone deacetylase complexes that regulate silencing and transcription. *Proc Natl Acad Sci U S A* **93**: 14503-14508.

Sabl JF, Henikoff S (1996) Copy number and orientation determine the susceptibility of a gene to silencing by nearby heterochromatin in Drosophila. *Genetics* **142**: 447-458.

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

Sambrook J, Fritsch EF, Maniatis T. (1989) *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Sato T, Russell MA, Denell RE (1983) Homoeosis in Drosophila: A New Enhancer of Polycomb and Related Homoeotic Mutations. *Genetics* **105**: 357-370.

Scarbrough K, Hattman S, Nur U (1984) Relationship of DNA methylation level to the presence of heterochromatin in mealybugs. *Mol Cell Biol* **4**: 599-603.

Schirling C, Heseding C, Heise F, Kesper D, Klebes A, Klein-Hitpass L, Vortkamp A, Hoffmann D, Saumweber H, Ehrenhofer-Murray AE (2010) Widespread regulation of gene expression in the Drosophila genome by the histone acetyltransferase dTip60. *Chromosoma* **119**: 99-113.

Schmitt S, Paro R (2006) RNA at the steering wheel. *Genome Biol* **7**: 218.

Schmitz G, Langmann T, Heimerl S (2001) Role of ABCG1 and other ABCG family members in lipid metabolism. *J Lipid Res* **42**: 1513-1520.

Schoenfelder S, Paro R (2004) Drosophila Su(Hw) regulates an evolutionarily conserved silencer from the mouse H19 imprinting control region. *Cold Spring Harb Symp Quant Biol* **69**: 47-54.

Schoenfelder S, Smits G, Fraser P, Reik W, Paro R (2007a) Non-coding transcripts in the H19 imprinting control region mediate gene silencing in transgenic Drosophila. *EMBO Rep*

Schoenfelder S, Smits G, Fraser P, Reik W, Paro R (2007b) Non-coding transcripts in the H19 imprinting control region mediate gene silencing in transgenic Drosophila. *EMBO Rep* **8**: 1068-1073.

Schotta G, Ebert A, Krauss V, Fischer A, Hoffmann J, Rea S, Jenuwein T, Dorn R, 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, Lachner M, Sarma K, Ebert A, Sengupta R, Reuter G, Reinberg D, Jenuwein T (2004) A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev* **18**: 1251-1262.

Schubert D, Primavesi L, Bishopp A, Roberts G, Doonan J, Jenuwein T, Goodrich J (2006) Silencing by plant Polycomb-group genes requires dispersed trimethylation of histone H3 at lysine 27. *Embo J* **25**: 4638-4649.

Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G (2007) Genome regulation by polycomb and trithorax proteins. *Cell* **128**: 735-745.

Schulze SR, Wallrath LL (2007) Gene regulation by chromatin structure: paradigms established in Drosophila melanogaster. *Annu Rev Entomol* **52**: 171-192.

Scott MJ, Li F (2008) How do ncRNAs guide chromatin-modifying complexes to specific locations within the nucleus? *RNA Biol* **5**: 13-16.

Scott RJ, Spielman M (2006) Deeper into the maize: new insights into genomic imprinting in plants. *Bioessays* **28**: 1167-1171.

Sevrioukov EA, He JP, Moghrabi N, Sunio A, Kramer H (1999) A role for the deep orange and carnation eye color genes in lysosomal delivery in Drosophila. *Mol Cell* **4**: 479-486.

Shestopal SA, Makunin IV, Belyaeva ES, Ashburner M, Zhimulev IF (1997) Molecular characterization of the deep orange (dor) gene of Drosophila melanogaster. *Mol Gen Genet* **253**: 642-648.

Shpiz S, Kwon D, Rozovsky Y, Kalmykova A (2009) rasiRNA pathway controls antisense expression of Drosophila telomeric retrotransposons in the nucleus. *Nucleic Acids Res* **37**: 268-278.

Sidorenko L, Dorweiler JE, Cigan AM, Arteaga-Vazquez M, Vyas M, Kermicle J, Jurcin D, Brzeski J, Cai Y, Chandler VL (2009) A dominant mutation in mediator of paramutation2, one of three second-largest subunits of a plant-specific RNA polymerase, disrupts multiple siRNA silencing processes. *PLoS Genet* **5**: e1000725.

Simon JA, Tamkun JW (2002) Programming off and on states in chromatin: mechanisms of Polycomb and trithorax group complexes. *Current Opinion in Genetics & Development* **12**: 210-218.

Simpson F, Peden AA, Christopoulou L, Robinson MS (1997) Characterization of the adaptor-related protein complex, AP-3. *J Cell Biol* **137**: 835-845.

Sinclair DA, Clegg NJ, Antonchuk J, Milne TA, Stankunas K, Ruse C, Grigliatti TA, Kassis JA, Brock HW (1998) Enhancer of Polycomb is a suppressor of position-effect variegation in Drosophila melanogaster. *Genetics* **148**: 211-220.

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

Soppe WJ, Jacobsen SE, Alonso-Blanco C, Jackson JP, Kakutani T, Koornneef M, Peeters AJ (2000) The late flowering phenotype of fwa mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. *Mol Cell* **6**: 791-802.

Spradling AC (1986) P element-mediated transformation. In *Drosophila: A Practical Approach*, Roberts DB (ed), pp 175-197. Oxford: IRL Press.

Spradling AC, Rubin GM (1982) Transposition of cloned P elements into Drosophila germ line chromosomes. *Science* **218**: 341-347.

Srivastava M, Hsieh S, Grinberg A, Williams-Simons L, Huang SP, Pfeifer K (2000) H19 and Igf2 monoallelic expression is regulated in two distinct ways by a shared cis acting regulatory region upstream of H19. *Genes Dev* **14**: 1186-1195.

Stam M (2009) Paramutation: a heritable change in gene expression by allelic interactions in trans. *Mol Plant* **2**: 578-588.

Stam M, Belele C, Dorweiler JE, Chandler VL (2002) Differential chromatin structure within a tandem array 100 kb upstream of the maize *b1* locus is associated with paramutation. *Genes Dev* **16**: 1906-1918.

Starcevic M, Dell'Angelica EC (2004) Identification of snapin and three novel proteins (BLOS1, BLOS2, and BLOS3/reduced pigmentation) as subunits of biogenesis of lysosome-related organelles complex-1 (BLOC-1). *J Biol Chem* **279**: 28393-28401.

Sullivan DT, Sullivan MC (1975) Transport defects as the physiological basis for eye color mutants of Drosophila melanogaster. *Biochem Genet* **13**: 603-613.

Sun FL, Cuaycong MH, Craig CA, Wallrath LL, Locke J, Elgin SC (2000) The fourth chromosome of Drosophila melanogaster: interspersed euchromatic and heterochromatic domains. *Proc Natl Acad Sci U S A* **97**: 5340-5345.

Surani MA, Barton SC, Norris ML (1984) Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature* **308**: 548-550.

Suzuki S, Renfree MB, Pask AJ, Shaw G, Kobayashi S, Kohda T, Kaneko-Ishino T, Ishino F (2005) Genomic imprinting of IGF2, p57(KIP2) and PEG1/MEST in a marsupial, the tammar wallaby. *Mech Dev* **122**: 213-222.

Suzuki T, Oiso N, Gautam R, Novak EK, Panthier JJ, Suprabha PG, Vida T, Swank RT, Spritz RA (2003) The mouse organellar biogenesis mutant buff results from a mutation in Vps33a, a homologue of yeast vps33 and Drosophila carnation. *Proc Natl Acad Sci U S A* **100**: 1146-1150.

Sved JA, Eggleston WB, Engels WR (1990) Germ-line and somatic recombination induced by in vitro modified P elements in Drosophila melanogaster. *Genetics* **124**: 331-337.

Swaminathan J, Baxter EM, Corces VG (2005) The role of histone H2Av variant replacement and histone H4 acetylation in the establishment of Drosophila heterochromatin. *Genes Dev* **19**: 65-76.

Szabo PE, Tang SH, Silva FJ, Tsark WM, Mann JR (2004) Role of CTCF binding sites in the Igf2/H19 imprinting control region. *Mol Cell Biol* **24**: 4791-4800.

Takagi N, Sasaki M (1975) Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature* **256**: 640-642.

Teixeira FK, Colot V (2010) Repeat elements and the Arabidopsis DNA methylation landscape. *Heredity* **105**: 14-23.

Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, Diemer K, Muruganujan A, Narechania A (2003) PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res* **13**: 2129-2141.

Thorvaldsen JL, Duran KL, Bartolomei MS (1998) Deletion of the H19 differentially methylated domain results in loss of imprinted expression of H19 and Igf2. *Genes Dev* **12**: 3693-3702.

Thummel CS, Boulet AM, Lipshitz HD (1988) Vectors for Drosophila P-element-mediated transformation and tissue culture transfection. *Gene* **74**: 445-456.

Tillib S, Petruk S, Sedkov Y, Kuzin A, Fujioka M, Goto T, 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.

Tiwari S, Schulz R, Ikeda Y, Dytham L, Bravo J, Mathers L, Spielman M, Guzman P, Oakey RJ, Kinoshita T, Scott RJ (2008) MATERNALLY EXPRESSED PAB C-TERMINAL, a Novel Imprinted Gene in Arabidopsis, Encodes the Conserved C-Terminal Domain of Polyadenylate Binding Proteins. *Plant Cell* **20**: 2387-2398.

Tremblay KD, Duran KL, Bartolomei MS (1997) A 5' 2-kilobase-pair region of the imprinted mouse H19 gene exhibits exclusive paternal methylation throughout development. *Mol Cell Biol* **17**: 4322-4329.

Tremblay KD, Saam JR, Ingram RS, Tilghman SM, Bartolomei MS (1995) A paternal-specific methylation imprint marks the alleles of the mouse H19 gene. *Nature Genetics* **9**: 407-413.

Tweedie S, Ashburner M, Falls K, Leyland P, McQuilton P, Marygold S, Millburn G, Osumi-Sutherland D, Schroeder A, Seal R, Zhang H (2009) FlyBase: enhancing Drosophila Gene Ontology annotations. *Nucleic Acids Res* **37**: D555-559.

Umlauf D, Goto Y, Cao R, Cerqueira F, Wagschal A, Zhang Y, Feil R (2004) Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nature Genetics* **36**: 1296-1300.

Vagin VV, Klenov MS, Kalmykova AI, Stolyarenko AD, Kotelnikov RN, Gvozdev VA (2004) The RNA interference proteins and vasa locus are involved in the silencing of retrotransposons in the female germline of Drosophila melanogaster. *RNA Biol* **1**: 54-58.

Vaissiere T, Sawan C, Herceg Z (2008) Epigenetic interplay between histone modifications and DNA methylation in gene silencing. *Mutat Res* **659**: 40-48.

VandeBerg JL, Robinson ES, Samollow PB, Johnston PG (1987) X-linked gene expression and X-chromosome inactivation: marsupials, mouse, and man compared. *Isozymes Curr Top Biol Med Res* **15**: 225-253.

Varmuza S, Mann M (1994) Genomic imprinting--defusing the ovarian time bomb. *Trends Genet* **10**: 118-123.

Vazquez J, Muller M, Pirrotta V, Sedat JW (2006) The mcp element mediates stable long-range chromosome-chromosome interactions in Drosophila. *Mol Biol Cell* **17**: 2158-2165.

Verona RI, Thorvaldsen JL, Reese KJ, Bartolomei MS (2008) The transcriptional status but not the imprinting control region determines allele-specific histone modifications at the imprinted H19 locus. *Mol Cell Biol* **28**: 71-82.

Vielle-Calzada JP, Baskar R, Grossniklaus U (2000) Delayed activation of the paternal genome during seed development. *Nature* **404**: 91-94.

Waddington CH (1942) The Epigenotype. *Endeavour* **1**: 18-20.

Wakefield MJ, Keohane AM, Turner BM, Graves JA (1997) Histone underacetylation is an ancient component of mammalian X chromosome inactivation. *Proc Natl Acad Sci U S A* **94**: 9665-9668.

Wallrath LL, Elgin SC (1995) Position effect variegation in Drosophila is associated with an altered chromatin structure. *Genes Dev* **9**: 1263-1277.

Wang L, Brown JL, Cao R, Zhang Y, Kassis JA, Jones RS (2004) Hierarchical recruitment of polycomb group silencing complexes. *Mol Cell* **14**: 637-646.

Warner TS, Sinclair DA, Fitzpatrick KA, Singh M, Devlin RH, Honda BM (1998) The light gene of Drosophila melanogaster encodes a homologue of VPS41, a yeast gene involved in cellular-protein trafficking. *Genome* **41**: 236-243.

Wei ML (2006) Hermansky-Pudlak syndrome: a disease of protein trafficking and organelle function. *Pigment Cell Res* **19**: 19-42.

Wierzbicki AT, Haag JR, Pikaard CS (2008) Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. *Cell* **135**: 635-648.

Williams RW, Rubin GM (2002) ARGONAUTE1 is required for efficient RNA interference in Drosophila embryos. *Proc Natl Acad Sci U S A* **99**: 6889-6894.

Wood AJ, Oakey RJ (2006) Genomic imprinting in mammals: emerging themes and established theories. *PLoS Genet* **2**: e147.

Wu CT, Howe M (1995) A genetic analysis of the Suppressor 2 of zeste complex of Drosophila melanogaster. *Genetics* **140**: 139-181.

Xiao W, Gehring M, Choi Y, Margossian L, Pu H, Harada JJ, Goldberg RB, Pennell RI, Fischer RL (2003) Imprinting of the MEA Polycomb gene is controlled by antagonism between MET1 methyltransferase and DME glycosylase. *Dev Cell* **5**: 891-901.

Yasuda GK, Schubiger G, Wakimoto BT (1995) Genetic characterization of ms (3) K81, a paternal effect gene of Drosophila melanogaster. *Genetics* **140**: 219-229.

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

Yoon YS, Jeong S, Rong Q, Park KY, Chung JH, Pfeifer K (2007) Analysis of the H19ICR insulator. *Mol Cell Biol* **27**: 3499-3510.

Zaratiegui M, Irvine DV, Martienssen RA (2007) Noncoding RNAs and gene silencing. *Cell* **128**: 763-776.

Zhang Q, Zhao B, Li W, Oiso N, Novak EK, Rusiniak ME, Gautam R, Chintala S, O'Brien EP, Zhang Y, Roe BA, Elliott RW, Eicher EM, Liang P, Kratz C, Legius E, Spritz RA, O'Sullivan TN, Copeland NG, Jenkins NA, Swank RT (2003) Ru2 and Ru encode mouse orthologs of the genes mutated in human Hermansky-Pudlak syndrome types 5 and 6. *Nat Genet* **33**: 145-153.

Zhang SD, Odenwald WF (1995) Misexpression of the white (w) gene triggers male-male courtship in Drosophila. *Proc Natl Acad Sci U S A* **92**: 5525-5529.

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

APPENDIX A COPYRIGHT PERMISSIONS

NRC RESEARCH PRESS LICENSE
TERMS AND CONDITIONS

Jul 17, 2010

This is a License Agreement between Lori A McEachern ("You") and NRC Research Press ("NRC Research Press") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by NRC Research Press, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	2471531065967
License date	Jul 17, 2010
Licensed content publisher	NRC Research Press
Licensed content publication	Genome
Licensed content title	The *pink* gene encodes the *Drosophila* orthologue of the human Hermansky–Pudlak syndrome 5 (*HPS5*) gene
Licensed content author	Monika Syrzycka, Lori A. McEachern, Jennifer Kinneard, et al.
Licensed content date	Jun 1, 2007
Volume number	50
Issue number	6
Type of Use	Thesis/Dissertation
Requestor type	Author (original work)
Format	Print and electronic
Portion	Full article
Order reference number	
Title of your thesis / dissertation	INTER-KINGDOM EPIGENETICS: CHARACTERIZATION OF MAIZE B1 TANDEM REPEAT-MEDIATED SILENCING IN DROSOPHILA
Expected completion date	Sep 2010
Estimated size(pages)	300
Total	0.00 USD
Terms and Conditions	

General Terms & Conditions

Permission is granted upon the requester's compliance with the following terms and conditions:

1. A credit line will be prominently placed in your product(s) and include: for books the author, book title, editor, copyright holder, year of publication; for journals the author, title of article, title of journal, volume number, issue number, and the inclusive pages. The credit line must include the following wording: "© 2008 NRC Canada or its licensors. Reproduced with

permission," except when an author of an original article published in 2009 or later is reproducing his/her own work.

2. The requester warrants that the material shall not be used in any manner that may be derogatory to the title, content, or authors of the material or to National Research Council Canada, including but not limited to an association with conduct that is fraudulent or otherwise illegal.

3. Permission is granted for the term (for Books/CDs-Shelf Life; for Internet/Intranet-In perpetuity; for all other forms of print-the life of the title) and purpose specified in your request. Once term has expired, permission to renew must be made in writing.

4. Permission granted is nonexclusive, and is valid throughout the world in English and the languages specified in your original request. A new permission must be requested for revisions of the publication under current consideration.

5. National Research Council Canada cannot supply the requester with the original artwork or a "clean copy."

6. If the National Research Council Canada material is to be translated, the following lines must be included: The authors, editors, and National Research Council Canada are not responsible for errors or omissions in translations.

v1.3

Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK10816854. Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

If you find copyrighted material related to this license will not be used and wish to cancel, please contact us referencing this license number 2471531065967 and noting the reason for cancellation.

Questions? customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

July 17, 2010

University of California Press
2120 Berkeley Way
Berkeley, CA 94704-1012

I am preparing my Ph.D. thesis for submission to the Faculty of Graduate Studies at Dalhousie University, Halifax, Nova Scotia, Canada. I am seeking your permission to include a manuscript version of the following book chapter as a chapter in the thesis:

The epigenetics of genomic imprinting: Core epigenetic processes are conserved in mammals, insects and plants. Lori A. McEachern and Vett Lloyd. In: Epigenetics: Linking Genotype and Phenotype in Development and Evolution. Edited by B. Hallgrímsson and B.K. Hall. In Press.

Canadian graduate theses are reproduced by the Library and Archives of Canada (formerly National Library of Canada) through a non-exclusive, world-wide license to reproduce, loan, distribute, or sell theses. I am also seeking your permission for the material described above to be reproduced and distributed by the LAC(NLC). Further details about the LAC(NLC) thesis program are available on the LAC(NLC) website (www.nlc-bnc.ca).

Full publication details and a copy of this permission letter will be included in the thesis.

Yours sincerely,

Lori McEachern

Permission is granted for:

a) the inclusion of the material described above in your thesis.

b) for the material described above to be included in the copy of your thesis that is sent to the Library and Archives of Canada (formerly National Library of Canada) for reproduction and distribution.

Name: _____ Title: _____

Signature: _____ Date: _____