Genetic and molecular characterization of chromosome three heterochromatin in *Drosophila melanogaster*

by

Kathleen Anne Fitzpatrick

B.Sc., University of British Columbia, 1980
M.Sc., University of British Columbia, 1985

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in the

Department of Molecular Biology and Biochemistry

SIMON FRASER UNIVERSITY

Summer 2005

Approval

Name: Kathleen Anne Fitzpatrick
Degree: Doctor of Philosophy

Title of Thesis: Genetic and molecular characterization of chromosome three heterochromatin in *Drosophila melanogaster.*

Examining Committee

Chair: Dr. G. Tibbits, Professor

Dr. B.M. Honda, Professor, Senior Supervisor
Department of Molecular Biology and Biochemistry, S.F.U.

Dr. D.A.R. Sinclair, Lecturer, Supervisor
Department of Molecular Biology and Biochemistry, S.F.U

Dr. N. Harden, Associate Professor, Supervisor
Department of Molecular Biology and Biochemistry, S.F.U.

Dr. D. Baillie, Professor, Internal Examiner
Department of Molecular Biology and Biochemistry, S.F.U.

Dr. Ross Hodgetts, Professor, External Examiner
Department of Biological Sciences, University of Alberta

Date Approved: April 29, 2005

SIMON FRASER UNIVERSITY

PARTIAL COPYRIGHT LICENCE

Abstract

Heterochromatin comprises approximately one third of the genome of the fruit fly, *Drosophila melanogaster*. It performs vital functions in chromosome segregation and cell division. One of the most interesting and unusual features of heterochromatin is its ability to silence genes artificially relocated near it, so the presence of active genes there is paradoxical. The presence of moderately and highly repeated sequences within heterochromatin and the absence of such features as meiotic recombination and polytene banding, make it difficult to study using standard genetic approaches. For example, although heterochromatin comprises approximately 60Mb of DNA, the earliest releases of the genome project included almost no heterochromatic sequence. More recent releases provide only about 21 Mb of sequence on a series of non-contiguous scaffolds, the majority of which remain unlocalized. My work has focused on two relatively poorly characterized heterochromatic regions: the right arm of chromosome 3 (3R) and a distal portion of 3L heterochromatin (distal h47). A large number of genetic mutations have been made via P-element insertion, excision and male recombination, and by X-ray mutagenesis. These mutations provide useful new tools for the identification and characterization of genes in these regions. In particular the deficiencies will allow a means for mapping computer generated gene models by PCR. The genetic and molecular tools developed for work with these previously unexplored regions will help advance our understanding of the organization and function of sequences in heterochromatin. They also provide a necessary foundation for a complete functional annotation of genes within these unusual and poorly understood regions of the chromosome.

Acknowledgements

I wish to thank my supervisor, Dr. Barry Honda, for providing me with the opportunity to work in this fascinating field, and for providing support and assistance throughout the project, especially as it drew to a close. I am also grateful to Drs. Nick Harden and Mike Smith, both members of my original committee, who sat through many meetings, providing encouragement and support throughout. Members of the Honda lab, in particular Monika Syrzycka and Sandra Schulze, have been enormously helpful, both scientifically and with coffee and emotional support. Dr. Vett Lloyd has been an inexhaustible source of support, ideas, suggestions, and has provided valuable feedback and encouragement throughout this very busy year. Dr. Don Sinclair, also a committee member, deserves special mention. He has been a bottomless well of information, encouragement, ideas, resources and nachos, for which I am eternally grateful.

One aspect of working in science is the opportunity to interact with other scientists. The following people generously provided strains and information: Dr. Alexei Tulin at the Carnegie Institute for some strains and talk about *Parp*; Dr. R. Levis, for P-element lines in chromosome 3 heterochromatin that he balanced before sending me; Dr. M. Leptin and Dr. Irion who shared alleles from their *abstrakt* screen; and Dr. A. Bejsovec who sent an allele of *zeppelin*. Also Dr. James Ehrman, who prepared the scanning EM images in chapter 4 and Alistair Coulthard, who examined some of the newly induced deficiencies cytologically.

Finally I want to thank my son, Christopher Devine, and husband, Bill Devine, who were supportive and helpful throughout.

Table of Contents

List of Tables

List of Figures

List of Abbreviations & Definitions

Abbreviation /Name	Description
[illegible]	Wild-type expression from a *white*$^{+}$ transgene.
3L	Left arm of third chromosome in *D. melanogaster*. Homologous to Muller's element D.
3R	Right arm of third chromosome in D. melanogaster. Homologous to Muller's element E.
ArmU	Database of genomic sequences from Release 3 of the *D. melanogaster* sequencing project, which cannot be mapped to any cytological position and therefore are called "unlocalized".
Chromo-domain	"Chromosome organization modifier": a protein sequence motif (~41 amino acids) found in a variety of chromosomally associated proteins.
BDGP	Berkeley *Drosophila* Genome Project (http://www.fruitfly.org/).
Flybase	Repository of *Drosophila* gene descriptions and stocks in Bloomington Indiana (http://flybase.bio.indiana.edu:82/).
HP1	Heterochromatin Protein 1.
Heteroallelic	See transheterozygote.
PEV	Position-Effect Variegation (described in Chapter One).
SNAP25	A large gene that maps to the middle-distal portion of 3L heterochromatin
Su(var)2-5	Suppressor of variegation 2-5 (encodes HP1)
Su(var)3-9	Suppressor of variegation 3-9 (encodes a methyl transferase)
Trans-Heterozygote	A genotype heterozygous for two different alleles; whether from the same gene or different genes is made clear from context. Sometimes used synonymously with *heteroallelic*, though this refers specifically to heterozygotes between different alleles from the same gene.

CHAPTER ONE: GENERAL INTRODUCTION

What is heterochromatin?

Cytological appearance

A genome is considered to be all of the DNA sequences carried by a cell or organism, but the total information in that genome extends beyond the DNA sequence. In eukaryotes, the DNA is organized into chromatin, a complex of nucleic acids and protein. The packaging of DNA into higher order structures accomplishes the organization of the extensive sequences into the more compact and manageable chromosomes (particularly during cell division) and also contributes to the interpretation and elaboration of the information from the DNA sequence into a functioning organism (i.e. epigenetics).

Chromatin is not uniform in appearance when examined microscopically. Much of the chromatin in a cell nucleus stains diffusely and appears loosely packed at interphase, but becomes condensed during mitosis. This chromatin is termed euchromatin. Some of the chromatin, however, remains highly condensed and darkly staining throughout the cell cycle and is named heterochromatin. Heterochromatin was first identified based upon this difference in appearance (Heitz 1928) and has since been found in most animal and plant genomes. Eukaryotic chromosomes are thus organized into structurally distinct components termed euchromatin and heterochromatin (reviewed in Grewal and Elgin 2002; Henikoff 2000). There is also considerable evidence that euchromatin and

heterochromatin are functionally distinct (see below). Heterochromatin in most organisms is found primarily in the pericentric (around the centromeres) and telomeric (at the chromosome tips) regions. Cytologically, heterochromatin itself is sometimes divided into α and β heterochromatin. β-heterochromatin is usually observed near the heterochromatin-euchromatin boundary and has a fuzzier, less condensed appearance relative to the more tightly packed α-heterochromatin (Gatti and Pimpinelli 1992). Heterochromatin is replicated late in S-phase and is very gene poor relative to euchromatin (Weiler and Wakimoto 1995).

The relative lack of genes in heterochromatin has led in the past, to its characterization of "junk" or even "ignorant" DNA (Ohno 1985). However, heterochromatin does contain some essential genes and carries out vital functions in a cell (see later).

Molecular organization

The different staining properties of euchromatin vs. heterochromatin may be attributed to differences in chromatin compaction as well as differing sequence composition. Euchromatin is composed primarily of single copy DNA and contains most of the genes; in contrast, heterochromatin is largely composed of repetitive DNA, consisting of middle repetitive (mainly transposable) elements (e.g. Pimpinelli et al. 1995; reviewed by Lohe and Hilliker 1995; Wallrath 1998;

Hennig 1999; Henikoff 2000) and highly repetitive satellite sequences (Lohe et al. 1993). β-heterochromatin may be composed mainly of middle repeats (transposable elements) whereas the α- heterochromatin may contain mostly the highly repeated satellite DNA. α- and β-heterochromatin may be interspersed in segments along the heterochromatin rather than defining two separate contiguous regions.

Chromatin structure

In eukaryotes, gene expression occurs in the context of chromatin structure (Jenuwein and Allis 2001). Molecular and biochemical studies have revealed the involvement of a "histone code", associated with pathways of de-acetylation, methylation and complex chromatin assembly, structure and function (e.g. Jenuwein and Allis 2001; Richards and Elgin 2002; Dernberg and Karpen 2002; Craig 2005). So, in addition to the broadly defined regions in chromosomes termed euchromatin or heterochromatin, there are regulated localized alterations in chromosome structure that lead to gene activation or repression. A reasonable hypothesis is that upstream regulatory sequences may be more or less easily accessible to transcription factors, depending on the state of the chromatin. This is illustrated by the alterations of *white* expression in transgenes inserted at various locations throughout the genome (for example, Sun et al. 2004). The fundamental unit of chromatin – the nucleosome – plays a critical role in how genes are regulated. Nucleosomes are composed of histones, which are subject to a wide range of post-translational modifications, including acetylation,

methylation and phosphorylation, as well as ubiquitination and ribosylation (Richards and Elgin 2002). Trans-acting gene regulatory factors not only recognize a particular regulatory sequence, but in addition, read a pattern of histone modification (the "histone code"), which can lead to a range of compacted states in the chromatin fibre, rendering the associated genes more or less open to the transcriptional machinery. This is known as chromatin remodeling, and the trans-acting factors involved often work in multimeric complexes and in combinatorial fashion, creating a range of variation in gene expression (Struhl 1999). Factors that initiate transcription form complexes that may include histone acetyltransferases, which acetylate histones at specific residues and create more open and accessible chromatin conformations. Transcriptional repressors may act in complexes that include histone deacetylases, which have an opposing effect.

Non-genic functions of heterochromatin

Despite its initial characterization as "inert" or "junk" DNA (Ohno 1985) heterochromatin nonetheless performs a variety of important functions. It contains the centromere in most higher organisms and plays important roles in meiotic pairing and sister chromatid cohesion and thus is essential for normal chromosome segregation (reviewed in Henikoff 2000; Sullivan et al. 2001; Blower and Karpen 2001; and Ahmad and Henikoff 2001).

Centromeric functions

In most organisms, centromeres are surrounded by heterochromatin and it has long been postulated that part of the essential role of heterochromatin is the maintenance of a functional centromere (e.g. Henikoff, 2000). In *Drosophila*, Murphy et al. (1995) (see also Sun et al. 1997; 2003) have constructed a minimal functional centromere for the *Dp1187* minichromosome (an X-chromosome derivative that segregates as a separate chromosome) using genetic techniques and characterized it by molecular mapping and partial sequencing. This 420kb functional centromere is composed of large blocks of simple repeat satellite DNA – 350 kb consisting primarily of two types of satellite sequence - interspersed with more complex sequence comprising transposons and a particular AT-rich sequence. Although these sequences define a functional centromere on the minichromosome, the sequences identified are found elsewhere on other *Drosophila* chromosomes but do not confer centromere –like properties. Other repeat sequences unrelated to those defined on the *Dp1187* chromosome have been implicated in centromere function on other chromosomes (Sun et al. 2003). Interestingly, recent refinements in the mapping and sequencing of the centromere associated sequences confirm that to date, there seems to be no particular satellite or repeat that uniquely defines the centromere (Sun et al. 2003). Instead, centromeres appear to be epigenetically determined in many eukaryotes and are defined in part by a specific histone H3 variant (CENP-A, or CID in *Drosophila*) that uniquely marks the centromeric

intervals. Recent work by both the Henikoff and Karpen labs has made great progress in dissecting the structure and properties of centromeres and the surrounding heterochromatin. Sullivan and Karpen (2004) have shown that centromeres themselves lack the specific modifications and unique chromatin structure of heterochromatin (such as H3K9; histone H3 lysine 9 modified by methylation), although they still require the surrounding heterochromatin for proper functioning, as evidenced by the studies defining the minimal centromere as an interval much larger than the actual region of kinetochore attachment. In addition, centromeric chromatin contains a dimethyl modification at histone H3, lysine 4, a modification that is usually associated with open but not yet active chromatin. This is in direct contrast to the structure of the surrounding heterochromatin in which H3K9 methylation is a common feature, but H3 lysine 4 is usually not modified. However, the centromeric chromatin also lacks numerous features of the more "active" euchromatin, such as specific acetylated histone modifications. The centromere thus has features of both euchromatin and heterochromatin in a unique combination. This suggests that surrounding heterochromatin, through its compact structure and unique histone modifications, as well as its replication timing, may act as a barrier to the spreading of centromeric chromatin beyond its limited range. Evidence for this comes from the work of Maggert and Karpen (2002), who used the mini chromosome, *Dp1187*, which was subjected to irradiation, followed by selection for the survival (i.e. stable transmission) of the resulting broken chromosomes. These were then sequenced in order to delimit a minimal centromere sequence. Surprisingly, they

found that some fragments missing the original centromeric sequences had acquired centromere function. Apparently centromere "identity" could shift or spread along the heterochromatin and allow the resulting chromosome fragments - having none of the original centromere sequences - to be maintained normally in the cells. Moreover, it was found that this shifting was apparently unidirectional. The *Dp1187* chromosome has euchromatic sequences juxtaposed to the centromere on one side and heterochromatic sequences on the other. In this experiment, centromere identity and function could apparently spread in the euchromatic direction but not towards the heterochromatic side. Thus one important function of heterochromatin may be defining and limiting the extent of the centromere. Heterochromatin may also define and limit the centromere by the sequestration of CID, the centromere specific histone H3 variant, leading to the restriction of CID-containing nucleosomes to the centromere region, as has been suggested by Henikoff et al. (2001).

Position-effect variegation

Heterochromatin displays a unique and interesting phenomenon in that euchromatic genes relocated into or near the heterochromatin can be variably inactivated. Either the inactivation of the gene or the "decision" to become inactivated will be clonally propagated, leading to groups of cells in which the gene is active, and others in which it is inactive. This creates a variegated

phenotype. This silencing property is epigenetic, since the sequence of the relocated gene is not altered, and it is termed position-effect variegation (PEV). *Drosophila* is an excellent model organism in which to study this phenomenon. PEV was first discovered in *Drosophila* (Muller 1930), and further work on PEV in this model organism has provided considerable insight into the nature of PEV (reviewed by: Spofford 1976; Karpen 1994; Lohe and Hilliker 1995; Weiler and Wakimoto 1995; Wallrath and Elgin 1995; Elgin, 1996; Henikoff 1998; Cryderman et al. 1998; Wallrath 1998; Grewal and Elgin 2002; Schotta et al. 2003; Craig et al. 2005; Dimitri et al. 2005).

Much of the research focusing on PEV, has centred on conditions (including temperature, and exposure to chemicals) as well as genes that enhance or suppress it. For example, modifiers of histone acetylation have been found to play a role in variegation; when histone deacetylation is inhibited with sodium butyrate, PEV is suppressed (Mottus et al. 1980). The fact that a chemical treatment known to alter histone modification, leads to greater expression from the variegating locus was an early indication of the role of histone modification in heterochromatin formation. Genetic screens for modifiers have identified many genes which when mutated, suppress or enhance PEV (Reuter and Wolff 1981; Sinclair et al. 1983, Wustmann et al. 1989). Some of these modifiers of variegation have been cloned and characterized, and a number have been found to encode chromatin proteins and in some cases, heterochromatin proteins (Cleard et al. 1997; Eissenberg et al. 1990; 1992; Reuter and Spierer 1992;

Sinclair et al. 1998; reviewed by Schotta et al. 2003). Thus, temperature and alterations in histone modification and the availability or dose of specific chromatin proteins, can modify PEV. PEV can also be altered by mutations in genes that affect transcription in general.

Heterochromatin proteins

The contrasting cytological appearance of heterochromatin and euchromatin correlates with different patterns of chromatin remodeling. Thus, euchromatin is more open and accessible to transcriptional regulators, whereas heterochromatin is not, and as described above can exhibit transcriptional silencing properties. The gene silencing properties of heterochromatin appear to be directly related to target accessibility. Relocation of an active gene into a region where the chromatin is "inaccessible" would create barriers to expression that may be only partly overcome by the transcriptional machinery. As describe above, many loci have been identified which when mutant, alter the expression of variegating loci. One of these is called *Suppressor of variegation 2-5* gene (*Su(var)2-5).* This gene encodes Heterochromatin Protein 1 (HP1) – a protein that binds modified histones as well as itself, causing the chromatin fibre to become even more compacted and transcriptionally repressed (Eissenberg and Elgin 2000). HP1 , which belongs to a class of chromatin components that may have a general

involvement in chromatin remodeling, has a specific protein motif called the chromo domain.

It is proposed that heterochromatin formation is based upon the methylation of one of the lysine residues on histone H3 (lysine 9, K9). Methylation of this residue is strongly correlated with a more compact chromatin conformation. HP1 is known to bind to methylated K9 histone H3 via its chromodomain (Nakayama et al. 2001). Via another domain, called the chromo shadow domain, it recruits the Suvar 3-9 protein, a methyltransferase that methylates K9H3 (Bannister et al. 2001; Lachner et al. 2001). This methylated residue then attracts more HP1 protein and so on (Jenuwein and Allis 2001; Peters et al. 2001; Sims et al. 2003). This "loop" propagates the spreading of heterochromatin, although as noted above (Sun et al. 2004), there is a limit to the distance of spreading. Recent evidence from the Wallrath lab (Danzer and Wallrath 2004) indicates that HP1 alone can account for limited spreading, approximately 2kb, even in the absence of Suvar3-9 protein activity. Longer range spreading – on the order of about 3.5 kb, requires the methyltransferase. The limited spreading based on HP1 alone is possible because HP1 can also bind its own chromo shadow domain, which contains a peptide motif – RVVVHPM – that strongly binds the chromodomain Smothers and Henikoff (2000). Thus HP1 can bind chromatin in limited regions and then further recruitment of HP1 can continue over short distances.

Although the precise role is unclear, it appears that small RNAs may be involved in the initial steps of heterochromatin formation by a method similar to RNAi (RNA interference; reviewed by Dernberg and Karpen 2002; Birchler et al. 2004). Interestingly, the chromodomain has been shown to depend on RNA for its binding to centric heterochromatin in mammalian cells. The association of HP1 with centric heterochromatin in these cells can be abolished by RNAse treatment (Maison et al. 2002). These authors speculated that the transcription of transposons within heterochromatin may be essential to its initiation and assembly. Double-stranded RNA produced from transposons near the centromere has been shown to direct a lysine 9-specific methylation activity that provides a binding site for HP1 (Reinhart and Bartel 2002; Volpe et al. 2002). Furthermore, HP1 can bind euchromatic sites where genes are actively transcribed , for example, the ecdysone response puffs and the heat shock puffs (Piacentini et al. 2003). That binding depends on RNA, is evidenced by the fact that RNAse treatment abolished it; moreover, the involvement of the chromodomain was shown by the failure of mutant HP1 lacking a chromodomain to bind these regions (Maison et al. 2002). All the evidence so far points to a complex process whereby RNA from transposons might recruit either Suvar3-9 or HP1 or both, to sites of potential heterochromatin, followed by a spreading phenomenon in which additional proteins are recruited until barriers or limits are reached. Although no specific sequences have been identified that serve as barriers, it is likely that such regions might also be epigenetically determined as

has been discussed above. This model could also explain the shifting of boundaries and centromeres over time.

Heterochromatin in Drosophila melanogaster

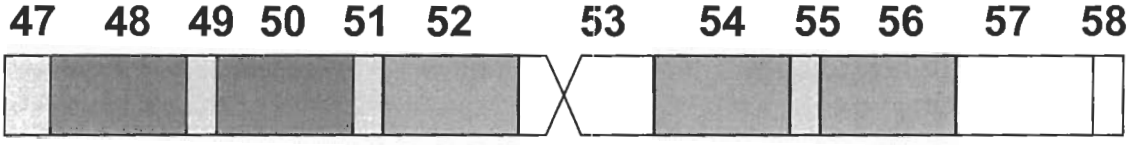
Heterochromatin has been well studied in the fruit fly, *Drosophila melanogaster* and some of its interesting properties have been best characterized in this organism. Arrangement of heterochromatin in *Drosophila* is similar to that of most eukaryotes: it is generally located at the pericentric (surrounding the centromeres) or telomeric (at the tips of the chromosomes) regions of the X chromosome and two major autosomes. The Y chromosome is entirely heterochromatic and the tiny fourth chromosome has an unusual arrangement: interspersed regions of heterochromatin and euchromatin (Sun et al. 2000).

Cytological mapping of heterochromatin is difficult because of the repetitive nature of the sequences within it, as well as other properties that make it interesting. Most physical chromosome mapping exploits the large, highly polytenized chromosomes in *Drosophila* salivary gland nuclei. However, these nuclei are in interphase, in which the heterochromatin of all the chromosomes remains under-replicated, and aggregates into an undifferentiated mass known as the chromocentre. To study heterochromatin organization and sequences, the condensed metaphase chromosomes ("mitotics") derived from larval brain tissue are used – here the chromosomes remain distinct instead of "clumping" at

the chromocentre, and the heterochromatic regions acquire a reproducible banding pattern in the presence of certain dyes. Maps can be generated by these approaches and they can be used to represent the orientation of specific heterochromatin intervals and for DNA sequence localization within those intervals. Thus, *Drosophila* heterochromatin has been divided into 61 regions: h1-h61. The heterochromatic Y-chromosome is composed of segments h1-h25; the X chromosome heterochromatin contains regions h26-h34; chromosomes 2 and 3 are each divided into 12 sections: h35-h46 and h47-58, respectively. The small fourth chromosome contains only 3 regions: h59-h61. A schematic representation of the map of chromosome 3 heterochromatin, is shown in Figure 1.1. Mitotic chromosomes pose some technical difficulties due to their small size and highly condensed structure. However they have been used to localize specific satellite sequences (Lohe et al. 1993) and transposons (Pimpinelli et al. 1995) to particular heterochromatic intervals. Various analyses reveal that heterochromatin is not a single entity but is rather composed of large blocks of highly repeated satellites, with some transposons interspersed and the regions of middle repetitive DNA (mainly transposons) themselves interspersed with small islands of unique sequence DNA (Weiler and Wakimoto 1995). It is likely that these differences, including the islands of unique sequence, are responsible for the different staining properties of each segment, although some proteins are known to bind specifically to certain regions of heterochromatin and this may also contribute to differential staining patterns.

Figure 1.1. Schematic representation of the cytological map of third chromosome heterochromatin.

Shadings on the heterochromatin indicate levels of fluorescence under DAPI staining, where the dark block, h48, fluoresces intensely, the medium grey blocks, h52, 54, and 56 show moderate fluorescence and the other intervals low fluorescence. Regions h53 and h57 are completely non-fluorescent. However, when an N-banding technique is used, only region h57 is stained. The combination of various staining patterns has led to the delineation of the map below, which is based upon work by Gatti et al. (1994), and includes the modifications of Koryakov et al. (2002), who differ in their designation of region h50 as showing lower fluorescence than h52 but more than regions h47, 49, 51, 55 and 58. Relative proportions of each segment shown are approximate.



The sequences at the euchromatin-heterochromatin junction are largely composed of transposable elements with single copy genes interspersed at a much lower density than found in standard euchromatin (one gene per 50 kb; approximately 6-fold lower than euchromatin; Hoskins et al. 2002). Heterochromatin/euchromatin boundaries are interesting and may have special properties. Genes in these regions function normally, despite residing in locations that would cause variegation of a more distal gene repositioned there. It is not yet clear whether the transition from proximal (towards the centromere) euchromatin into distal heterochromatin is abrupt, graded or mosaic, and what effect this may have on the genes in this region. In *Arabidopsis*, the transition from euchromatin to heterochromatin through the boundary region of an individual chromosome is gradual rather than abrupt in terms of the change in sequence composition (Haupt et al. 2001). The arrangement and properties of heterochromatin in *Arabidopsis* are similar to that of *Drosophila* (Sun et al. 1997).

Sun et al. (2004) have defined some limits to heterochromatic spreading with recent work on the heterochromatin of chromosome 4. The study of a set of P-element insertions along the fourth chromosome suggested that certain repeat elements, specifically the *hoppel* element, might serve as nucleation or initiation sites for heterochromatin formation, with subsequent spreading up to 10 kb away. It is not clear what defines this limit to spreading, nor whether any particular sequence element can account for it. One possibility is that the

boundary between euchromatin and heterochromatin, if it does exist, is an epigenetic mark. In fission yeast, *S. pombe,* the limits of heterochromatin are marked by specific intervals of chromatin having extremely high levels of histone modifications usually associated with active chromatin, i.e. certain residues on the histone tails are acetylated. The levels of acetylated histones along these boundary regions are much higher than in normal euchromatin and thus may serve as barriers to the spreading of heterochromatin beyond its limited range (Noma et al. 2001). An analogous study of *Drosophila* chromosomes has not been reported, but it is possible that similar arrangements exist in this organism.

Heterochromatic Genes in Drosophila

Identification of essential loci in heterochromatin

A puzzling feature of heterochromatin is the residence of essential loci in this apparently transcriptionally repressive environment. Data indicate that the heterochromatic environment is inhospitable to the functioning of euchromatic genes and yet a number of essential loci have been identified that reside there. In addition, they also exhibit PEV, but of a reciprocal kind – a heterochromatic gene can be shut down when translocated into distant euchromatin, and factors that enhance or suppress PEV for euchromatic genes have the opposite effect on variegating heterochromatic genes (Hearn et al. 1991; Eberl et al. 1993; Howe et al. 1995; Elgin 1996; Clegg et al. 1998). Some mutations that enhance expression of variegating euchromatic genes (i.e. *Su(var)s*) also enhance the

mutant phenotypes of some heterochromatic loci as well as reducing the wild type expression of a gene residing in heterochromatin (Clegg et al. 1998; Lu et al. 2000). So heterochromatic genes present something of a paradox in their apparent requirement for a heterochromatic environment for full function.

Classical genetic approaches have been used to identify a number of loci in *Drosophila* heterochromatin. Although standard genetics has been an effective approach when attempting to map genes in heterochromatin such approaches could be problematic since such aberrations could disrupt centromere function; and thus may not be recoverable. However, an elegant genetic method has been used to generate suitable deficiencies near the centromere in heterochromatin. This method was first described by Rasmussen (1960; as cited by Holm and Chovnick 1975). Using this approach, chromosomes are irradiated to produce compound left and right arms, and then irradiated again to restore the normal configuration. By this method, deficiencies surrounding the centromere can be recovered. A potential drawback of this approach is that complex rearrangements, including inversions, translocations and duplications can be produced, which can complicate subsequent analysis – both cytologically and genetically. Nonetheless, this procedure has been used successfully to create a set of heterochromatic lesions on chromosomes 2 and 3 heterochromatin. These have subsequently formed the basis for further genetic screens and more detailed studies of heterochromatic loci (Hilliker and Holm 1975, Hilliker 1976, Marchant and Holm 1988a, b). For example, the centric heterochromatin of the

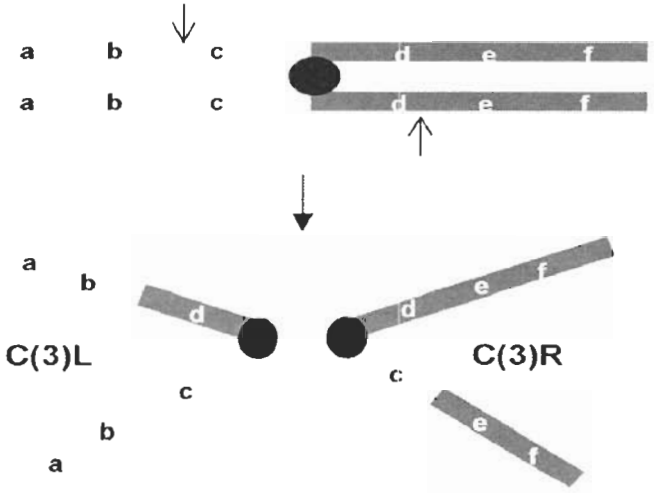
second chromosome has been the target of several genetic screens. Initial mapping studies involving the second chromosome placed two loci, *rolled* (*rl*) and *light* (*lt*), within or near the centromeric heterochromatin (Hessler 1958; Hannah 1951; Schultz 1937). Hilliker and Holm (1975) isolated a series of overlapping deficiencies in proximal chromosome 2 heterochromatin. These lesions allowed them to place both *lt* and *rl* within heterochromatin on the left and right arms, respectively. In the same work, 3 additional essential heterochromatic loci were identified. Subsequent work identified at least seven vital loci in 2L heterochromatin and at least six essential loci in 2R heterochromatin in subsequent EMS screens (Hilliker 1976). These data suggested that the gene density in the chromosome 2 heterochromatin was approximately 1% of that of euchromatin. Figure 1.2 shows a simplified schematic representation of a compound attachment and detachment experiment in which one of a large number of possible outcomes is depicted. The consequences of deletions or duplications of genetic material will depend upon whether or not essential genes are located within the regions affected.

Further results are still emerging from chromosome 2 (Yasuhara et al. 2003); and recently two additional loci - one in 2L heterochromatin and one in 2R - were reported (Coulthard et al. 2003) demonstrating that screens of the heterochromatin have not yet reached the saturation point (that point at which further mutagenesis will reveal no further loci). Recent work on chromosome 2R

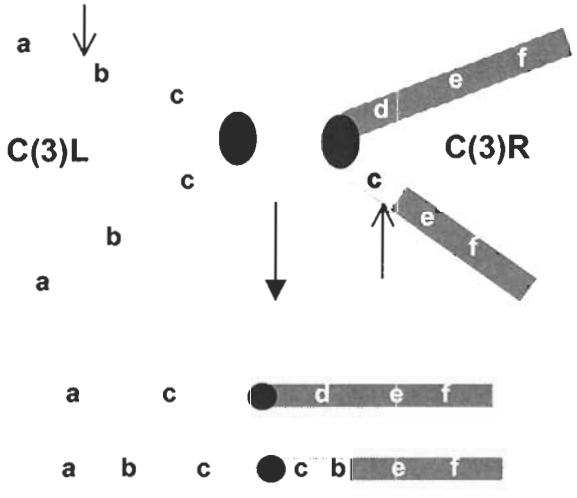
Figure 1.2. Schematic representation of the compound attachment-detachment method.

A. The generation of compound autosomes. Thin arrows depict irradiation breakpoints. A pair of possible outcomes is shown, in which the compound 3L bears a deficiency of 3L and a duplication of a portion of 3R and the reciprocal event. **B**. The detachment of compound autosomes could lead to numerous potential rearrangements. In the figure below, if the compound 3R autosome from A is irradiated, a number of breakpoints could be induced. One possible event (thin arrows) is shown. The resulting chromosome configurations are shown; both duplications and deficiencies may result.

A.



B.



heterochromatin has shown that there may be several previously unidentified loci in this region (Myster et al. 2004); this awaits further study.

Heterochromatic loci have also been identified on the X-chromosome; however some genetic studies suggest that *bobbed* is the only vital heterochromatic locus (Lindsley et al. 1960; Appels and Hilliker 1982). There are up to 16 genetic elements on the Y chromosome including six male fertility factors (Gatti and Pimpinelli 1992; Gvozdev et al. 2000; Carvalho et al. 2001; Carvalho 2002).

Drosophila was one of the first model organisms to have its genome sequenced. The first release was published in March 2000 (Adams et al. 2000) and has since undergone three revisions. The entire genome of the fruit fly is estimated to be about 176 megabases (Mb), roughly one third of which is heterochromatic (Celniker et al. 2002; Hoskins et al. 2002). However, due to its repetitive nature, the heterochromatic sequences have in general been poorly represented in the early releases. As an example, the first release of 120 Mb of euchromatic *Drosophila* genome sequence was missing most of the estimated 60Mb of heterochromatic DNA (Adams et al. 2000). However the recently released new whole genome shotgun sequence assembly (WGS3) greatly improved the assembly of the pericentric heterochromatin (Celniker et al. 2002; Hoskins et al. 2002).

As indicated above, gene annotation is particularly difficult in heterochromatin, where genes may span hundreds of kilobases. Gene finding programs are in general "trained" on euchromatic data sets in which heterochromatic genes would be poorly represented, if included at all. In particular, in the early releases of the *Drosophila* genome, the tendency was to separate a large transcription unit into smaller "genes". Most programs were unable to accurately identify heterochromatic genes because firstly, many of the genes spanned more than one "contig". The earliest release of the genome included about 700 small contigs of unlocalized sequence that contained some of the heterochromatic sequences – these contigs were small and most often did not contain entire genes, but occasionally only a single small exon. Secondly, even if an entire heterochromatic coding sequence was localized on one scaffold, the genes are often interrupted by numerous transposable elements within their introns (for example, the *Parp* locus; Tulin et al. 2002) and these are sometimes annotated as genes as well. A program that identified one or two exons from gene A would then discover "genes" B and C (actually transposons) and the remainder of gene A further along the contig would then be designated gene D and so on. Perhaps the best way to identify bonafide loci in heterochromatic scaffolds is to search for cDNA sequences – this laborious method allows the identification of very long transcription units regardless of the number of scaffolds they may span or the number of transposable elements within their introns. This method, of course, requires that a full-length cDNA sequence be available for each gene in heterochromatin, which is not the case for some heterochromatic loci. Release 3

of the genome provided improved gene annotation for heterochromatic as well as euchromatic genes (Misra et al. 2002). Release 3 also provided almost 21MB of heterochromatic sequences on a reduced number of larger scaffolds, for improved annotation (Hoskins et al. 2002) and a more thorough analysis of transposon content (Kaminker et al. 2002). However the number of loci identified genetically does not approach the number of predicted genes in heterochromatin (Hoskins et al. 2002) suggesting either that genetic screens have not yet identified all essential genes and/or that many predicted genes are either not expressed or not essential.

It is thought that roughly one third of all genes will mutate to a detectable phenotype (Miklos and Rubin 1996). Since the genome project predicts approximately 300 heterochromatic loci (Hoskins et al. 2002), this predicts that perhaps as many as 100 of these could be identified using classical genetic approaches. However, only 32 genes in *Drosophila* centric heterochromatin have been identified to date: (Hoskins et al. 2002). Of these, only a few have been characterized both molecularly and genetically: (Hilliker 1976; Devlin et al. 1990a,b; Biggs et al. 1994; Risinger et al. 1997; Hanai et al. 1998; Warren et al. 2000; Tulin at al. 2002, Schulze, 2003; Schulze et al. 2005).

Characteristics of heterochromatic genes in Drosophila

Since relatively few heterochromatic genes have been studied in detail, it is not possible to describe features of all heterochromatic loci definitively. However, for

those which have been characterized, the most striking organizational feature – with a few exceptions -is their very large size. Table 1.1 lists six heterochromatic genes that have been fairly thoroughly characterized. It can be seen from the table that a variety of types of loci reside in heterochromatin, with numerous functions. Cytological location is indicated along with the genomic scaffold identification number. The large size of most of these heterochromatic genes is due to their large introns, containing many repetitive sequences; the genes are surrounded by repetitive sequences also (Devlin et al. 1990a, b; Risinger et al. 1997; Warner et al. 1998; Tulin et al. 2002). As described previously, these genes appear to require this repetitive environment for full activity. So when a heterochromatic gene is relocated to a euchromatic environment, its expression is compromised in a variable manner (Wakimoto and Hearn 1990; Eberl et al. 1993; Howe et al. 1995) and this effect is enhanced in a *Suppressor of variegation 2-5* (*Su(va)r2-5*) mutant background (Eberl et al. 1993, Howe et al. 1995, Elgin 1996, Clegg et al. 1998). As mentioned, this is the converse of PEV of euchromatic loci.

The three loci in 3L heterochromatin that have been characterized (*Snap-25*, *Dbp-80* and *RpL15*) differ somewhat. *Snap-25* and *Dead box protein 80* (*Dbp80)* are both arranged much like a "typical" heterochromatic gene. Canonical heterochromatic genes such as *light* and *rolled* are large, spanning as much as 50kb or more; most of this expanse consists of large introns. *Dbp80* and *Snap-25* are even larger; in the case of *Snap-25*, the complete transcription unit,

Table 1.1. Some heterochromatic genes in *D.melanogaster.*
Examples of heterochromatic genes that have been characterized in some detail. The asterisk indicates cytology deduced by Schulze 2003, but not yet recorded in flybase.

GENE	FUNCTION	MAP	Genomic Scaffold ID	GENE SIZE (as annotated)	REFERENCE
light (CG18028)	Vacuolar assembly	2L, 40D3-4	AABU01002768	15,970bp	Devlin et al. 1990
rolled (CG12559)	MAP kinase	2R, h41	AABU01001947	50,341bp	Hoskins et al. 2002
Snap-25	Soluble NSF-attachment protein	3L, h47-h49*	AABU01002755	222,192bp	Risinger et al. 1997
Parp (CG40411)	Poly (ADP-ribose) polymerase	3R, 81F	AABU01002763	>200,000bp	Tulin et al. 2002
Dbp80	RNA helicase	3L, h49-h51*	AABU01002497	>140,000 bp	Schulze 2003
RpL 15	Large subunit ribosomal protein	3L, h49-h51*	AABU01002497	<2kb	Schulze 2003

approximately 200 kb in length, is more than 100 times longer than the mature message of approximately 2.1 kb. *RpL15* is a highly unusual heterochromatic gene in that it is very compact, with a single small intron. As ribosomal proteins are highly expressed, a long transcription unit may be transcribed too slowly to produce sufficient ribosomal protein. For example, some of the large Y-linked genes in *Drosophila hydei* contain large introns on the order of several megabases, that are predicted to take 2-3 days to transcribe completely (Reugels et al. 2000). It seems reasonable to speculate that there would be strong selection against any alteration that might lead to increased gene size for this locus so that its transcription could keep pace with that of rRNAs. To determine whether the type of organization observed with *Snap-25* and *Dbp80* is a feature of the gene or its heterochromatic location, examining a euchromatic orthologue in a closely related species is very useful. This has been done recently for both *RpL15* and *Dbp80* (Schulze 2003). When the euchromatic orthologue of *Dbp80* in *Drosophila virilis* is compared to the heterochromatic version *in Drosophila melanogaster*, the coding sequence of the two genes is highly conserved, but the gene organization is very different. There are many small exons and a corresponding number of very large introns in the heterochromatic *D. melanogaster Dbp80*; and relatively few, larger exons with a few small introns in the euchromatic *D. virilis* version. This would argue that large size is a feature of most heterochromatic loci, but this need not be a functionally necessary feature, as demonstrated by *RpL15.* Rather, it is likely

that these loci are large as a result of their placement into heterochromatin. As euchromatic genes are relocated into or near heterochromatin, they are closer to the insertion sites of numerous transposable elements, and over evolutionary time may begin to accumulate transposons between their exons. Any insertions that disrupted an exon and thus gene function would be selected against and those that interfered with splicing or processing of transcripts would also likely be lost relatively quickly. Transposons that did not interfere with processing and expression of a gene would be retained, and might provide increased target sites for further transposon insertions. Support for this idea comes from the work of Tulin et al. (2002). They mapped and sequenced a large stretch of 3R heterochromatin surrounding the *Parp* locus and found that the vast majority of transposons found in the introns, and up and downstream of the locus, were oriented in the opposite direction to *Parp* (Tulin et al. 2002). Thus any possible splice sites present in the transposons would not be expected to interfere with splicing and processing *Parp* because they would be not be read by the spliceosome. Transposons may not accumulate within heterochromatin by a slow or gradual process (eg. the P-element) but rather as a result of relatively rapid transposition "bursts" (Dimitri and Junakovic 1999).

Heterochromatic genes require heterochromatic proteins to function.

To date, no defining regulatory feature of heterochromatic genes has been identified; for example no universal promoter structure alterations or particular

upstream regulatory sequences have been found for heterochromatic genes. However, those that have been examined do appear to rely on certain heterochromatin proteins for their correct expression. The expression of the heterochromatic genes *light* and *rolled* in their native heterochromatic environment, is affected by a modifier of PEV (Clegg et al. 1998; Sinclair et al. 2000), and mRNA levels of these heterochromatic genes are reduced in a *Su(var)2-5* mutant background (Lu et al. 2000). Recent work on both *RpL15* and *Dbp80* (Schulze 2003; Schulze et al. 2005) showed that despite their differing organization and evolutionary history, both exhibited reduced mRNA expression in *Su(var)* mutant backgrounds. Thus heterochromatic loci appear to require their heterochromatic environment for correct functioning, which the more recent molecular experiments have confirmed. This suggests that specific heterochromatic chromatin complexes may be involved in the normal regulation of genes which reside in heterochromatin. The apparent dependence of heterochromatic genes upon heterochromatin proteins for full expression could be a defining feature of such loci, although very recent evidence indicates that HP1, encoded by *Su(var)2-5*, actually positively regulates a small number of euchromatic loci also (Cryderman et al. 2005)

Heterochromatic genes on chromosome 3

Cytologically, chromosome 3 heterochromatin consists of segments h47-52 on the left arm and h53-58 on the right arm (h53 contains the centromere). Mapping

of repeat elements has been performed, and numerous satellite and transposon sequences have been localized to specific heterochromatic segments (Lohe et al. 1993; Pimpinelli et al. 1995). Much of the centromeric heterochromatin in *D. melanogaster* is made up of a series of repeated satellite sequences, with some of chromosome 3 being composed of two simple satellite sequences that occupy at least part of h48 (1.686) and h57 (1.705) (Lohe et al. 1993). In addition, the 3R centromere region contains a large block of a 12 nucleotide repeat called the dodeca satellite sequence (Abad et al. 1992). Also, a substantial proportion of proximal heterochromatin is composed of clusters of transposable elements. Those on chromosome 3 include *blood*, *copia, gypsy*, and numerous others (Pimpinelli et al. 1995). More recent mapping of known deficiencies and presumed heterochromatic DNA sequences has localized certain genes in some of these segments (Dimitri et al. 2003).

The initial work to identify essential heterochromatic genes on chromosome three also involved the creation of compound autosomes and their subsequent detachment to generate potential heterochromatic deficiencies (Marchant and Holm 1988a). This work led to the isolation of a set of lesions in both 3L and 3R heterochromatin, some of which formed the basis for subsequent EMS mutagenesis (Marchant and Holm 1988b). The original estimate of 12 essential in the chromosome 3 heterochromatin has since been refined by a series of P-element, EMS and γ-irradiation screens which identified two additional complementation groups (Schulze et al. 2001; Kennison and Tamkun 1988;

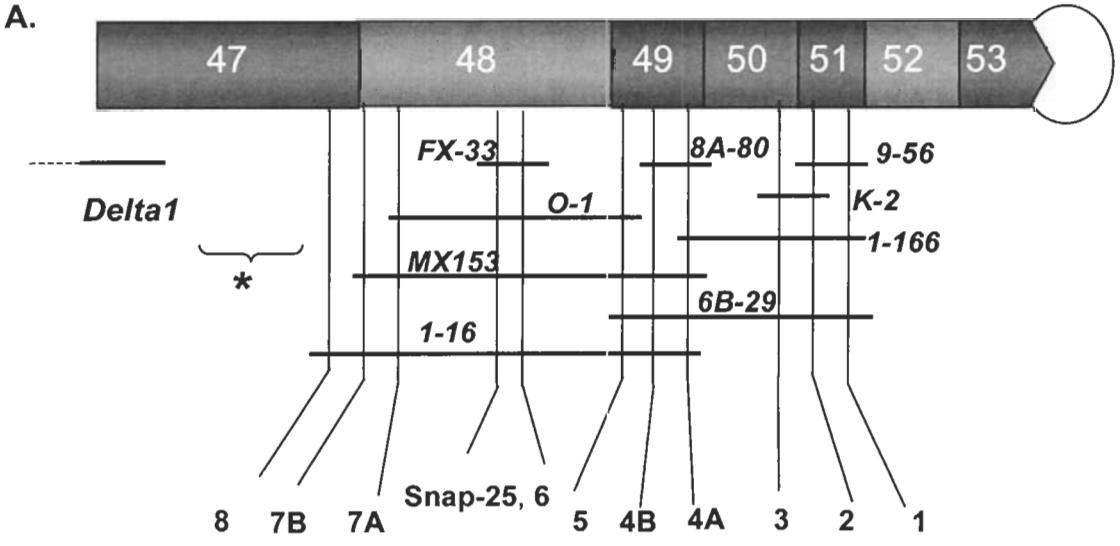
Vilinsky et al. 2002). It is possible that the region has not yet been saturated, as some of the loci are identified by single mutant alleles. The large number of available deficiencies and other lesions has, however, permitted a finer scale mapping of the lethal complementation groups originally defined by Marchant and Holm (1988b), leading to a new estimate for both their number and relative order. Figure 1.3 depicts a genetic map of 3L and 3R heterochromatin; note that the distances between loci are inferred from breakpoint frequency, and may differ substantially from the cytological arrangement of the genes.

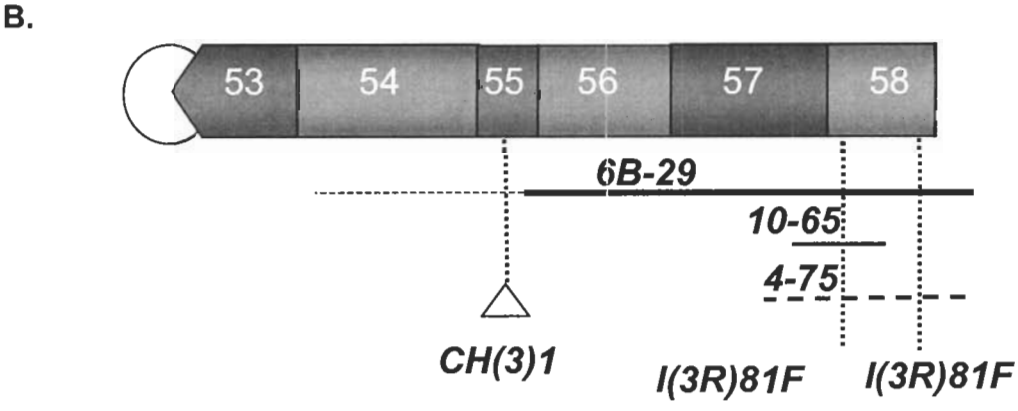
While a genetic map can be used to order loci, connecting the genetic information to the physical chromosome map is also very informative. As mentioned previously, heterochromatin can be characterized cytologically using mitotic chromosomes, and this has been done for some of the lesions in 3L and 3R heterochromatin by Koryakov et al. (2002; 2003) who mapped the breakpoints of many of the compound- attachment /detachment deficiencies to specific segments of mitotic chromosomes. From this work, combined with genetic analyses (Schulze 2003) a clearer picture of how heterochromatic genes are physically organized was obtained. Some of the lesions are not simple deficiencies, but more complex rearrangements. These initial findings placed the 3L heterochromatic genes into three large groups. *lethals 1-3* are located in h50 or 51, *lethals 4A-5* in h49 and *lethals 6-8* in h47-48. A finer resolution can be inferred, by comparing the cytological regions removed by overlapping deficiencies, with the lethal complementation groups they contain (Schulze

2003). The result is shown in Figure 1.3, which is a summary of the cytological and genetic data for this region.

The proposed arrangement of the loci in Figure 1.3 is interesting, because based on the information we have to date, they appear to be clustered in relatively small intervals of heterochromatin which show similarities in their staining properties. This suggests that genes are located in "islands" of more complex sequence within the bulk of the simpler, highly repetitive heterochromatic sequence. Currently, not enough is known about the precise locations of the genes because those intervals in which the loci probably reside tend to be small with relatively weak staining and are thus the most difficult to characterize with certainty. Therefore, the arrangement shown in Figure 1.3 must be considered hypothetical until each gene has been identified and physically mapped to the chromosomes. However in conjunction with recent genetic findings from the Honda lab, the cytological data of Koryakov et al. (2002) support a reasonably precise positioning of at least six of the loci. For example Koryakov et al. have mapped the heterochromatic breakpoint of a particular inversion (*In(3R)hb*^{bs23}) to region h49, and this same breakpoint has generated a lethal allele of *lethal 4A* (Sinclair, unpublished), placing this gene and the others in the group defined by Koryakov et al. (*lethal 4B* and *lethal 5*) within this quite small interval. A similar situation exists for the heterochromatic breakpoint of *In(3L)C90*, which is lethal in combination with existing lesions in *lethal 1* (Schulze 2003), placing this gene

Figure 1.3. Genetic map of (**A**) 3L and (**B**) 3R heterochromatin.
The shaded rectangle at the top of each figure represents the heterochromatic portion of chromosome 3. The circle represents the centromere, and the lethal complementation groups (genes) on the left arm are numbered *lethal 1 (1)* through *lethal 8 (8)* excluding *Snap-25*. Gene names are given for the right arm. Lines corresponding to the deficiencies delineate the regions of DNA missing from the respective deficiency chromosome and the minimal genetic extent of the deficiency. The asterisk in **A** indicates a region in h47 for which no lesions existed at the beginning of the study. *CH(3)1* in **B** indicates a P-element insertion (Zhang and Spradling 1994). *Df(3R)4-75* is depicted as a dashed line because Koryakoff et al. (2002) have not been able to detect a deficiency for this lesion.





(and *lethals 2* and *3*) in the cytological region h51. Further detailed genetic and cytological mapping will undoubtedly lead to improvement of this current picture.

In addition, Corradini et al. (2003) have also used a series of BACs to localize a set of 7 loci within the h47 segment on 3L; this is not shown in the figure, but some of these genes will be discussed in Chapter 4. Figure 1.3 also shows only a small subset of the lesions available that affect loci on 3L. A minimal set of deficiencies that define each locus is shown. For example, *lethal 4A* can be distinguished based on the lethality of its mutant alleles in combination with both the *8A-80* and *1-166* deficiencies, while *lethal 4B* alleles are viable in combination with *1-166*. The only loci between which no breakpoint appears to exist yet, are *Snap-25* and *lethal 6*, as shown in Figure 1.3.

Why study heterochromatin?

Heterochromatin can be challenging to work with, but possesses many fascinating properties. It is involved with vital functions in cells, defining centromere identity and function, for instance. And it can cause PEV of euchromatic genes. Despite this "silencing property" of heterochromatin, there are active and essential genes that reside there, and appear to require heterochromatin and its associated proteins for full function. There are still few reports of heterochromatic genes that have been thoroughly characterized. These must have developed unique mechanisms of escaping the gene silencing properties of heterochromatin. Conversely, they have apparently come to rely on

those very properties for their normal functioning. An in-depth study of these genes should therefore lead to a better understanding of the effects of chromatin structure on gene expression in general.

Project aims

In the work described in Chapter 3 I have undertaken a re-evaluation of the genetic and molecular characterization of the heterochromatin of 3R by the isolation of new deficiencies and other rearrangements. I have also initiated work to determine which coding sequences are missing from the deficiencies. In Chapter 4, I summarize my work on the characterization of 3L heterochromatin, in which I defined a distal region between the euchromatin and heterochromatin, which has hitherto been absent from the deficiency coverage and within which no mutations had been isolated at the onset of the study. Further, I have been able to use a single embryo PCR method to connect some coding sequences to particular deficiency intervals throughout the 3L heterochromatin, and have thus assisted in the placement of loci within the region. The isolation of new lesions in these previously uncharacterized regions of chromosome 3 heterochromatin should lead to a greatly improved understanding of the organization and structure of genes therein.

CHAPTER TWO: MATERIALS AND METHODS

Genetics

Drosophila culture conditions, stocks and strains

Flies were grown on standard cornmeal-sucrose medium with either Tegosept or propionic acid as a mould inhibitor. Stocks were routinely maintained and crosses performed at room temperature (22°C), unless otherwise indicated. For embryo collections, apple juice agar plates were made according to Ashburner (1989).

Descriptions of most mutations, special chromosomes, and deficiencies used in this work can be found at the flybase website (http://flybase.bio.indiana.edu:82/). A table of strains used and their sources is listed in Appendix A. The genes, and extent of various third chromosome heterochromatic deficiencies, are shown in Figure 1.3 (Marchant and Holm 1988a,b; Schulze et al. 2001). To simplify nomenclature, 3L genes *l(3)80Fj* through to *a* (proximal to distal), are numbered outwards from the centromere i.e. *j = lethal 1 or l1,* through to *a=l8.*

Mutagenesis Screens

P-element mobilization screens

To isolate P-element induced lethal mutations in 3R heterochromatin, p{*UAS-GFP.Y}B1* homozygotes were crossed to *SbΔ2.3/TM3,* a strain containing a

transposase source. Δ2.3 is a P-element which lacks the terminal repeats. Such an element produces transposase but is itself unable to mobilize. The resulting p{*UAS-GFP.Y}B1/SbΔ2.3* males were crossed *en masse* to *TM3 Sb/TM6 Hu Tb* virgin females. Individual p{*UAS-GFP.Y}B1/TM6* males were selected and pair-mated to 3-5 *Df(3L+R)6B-29/TM3Ser* virgins. Lines which yielded no p{*UAS-GFP.Y}B1/Df(3L+R)6B-29* progeny were kept for further characterization. The crossing scheme is shown in Figure 2.1.

A P-element inserted in distal h47 heterochromatin was mobilized as shown in Figure 2.2. Transheterozygous *w/Y; KG01176/SbΔ2.3* males were crossed *en masse* to *w/w; TM3 Sb/TM6 Hu Tb* virgin females with about 6-8 males and 25-40 females in each bottle. Progeny were selected based on alteration in *white*$^{+}$ expression (the *KG01176* P-element includes a *white*$^{+}$ reporter gene, which displays variable expression of eye pigments depending upon the chromatin environment into which the element is inserted). In addition, any flies with *Minute*-like characteristics (short, fine bristles and delayed development) were selected.

Screen for radiation-induced lethal mutations in 3R and 3L heterochromatin

e^{s}/e^{s} males were treated with 4000 Rads of X-irradiation, allowed to recover for 24 hours, and mated to *TM3 Sb /TM6 Hu Tb*, females. Single F1 e^{s} **/TM3, Sb Ser* or e^{s} **/TM6* males (where * indicates a

Figure 2.1. Schematic diagram of crosses performed to induce lethal P-element insertions into 3R heterochromatin.
Lines in which *p{UAS-GFP.Y}/Df(3L+R)6B-29* were lethal were retained for further characterization.

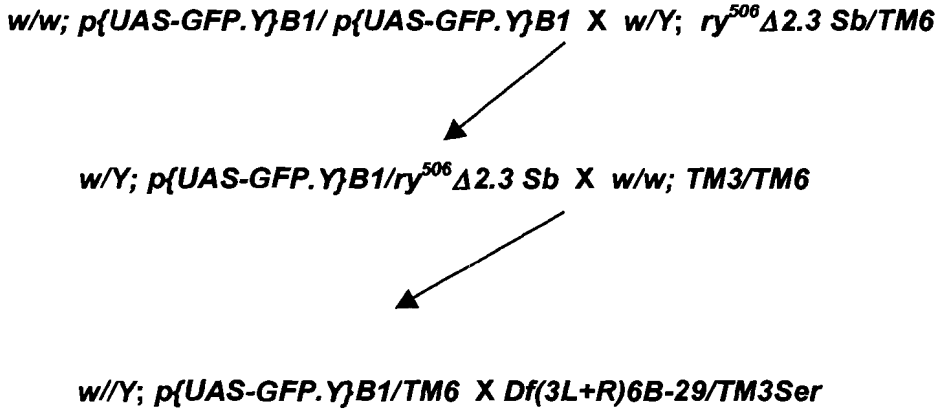
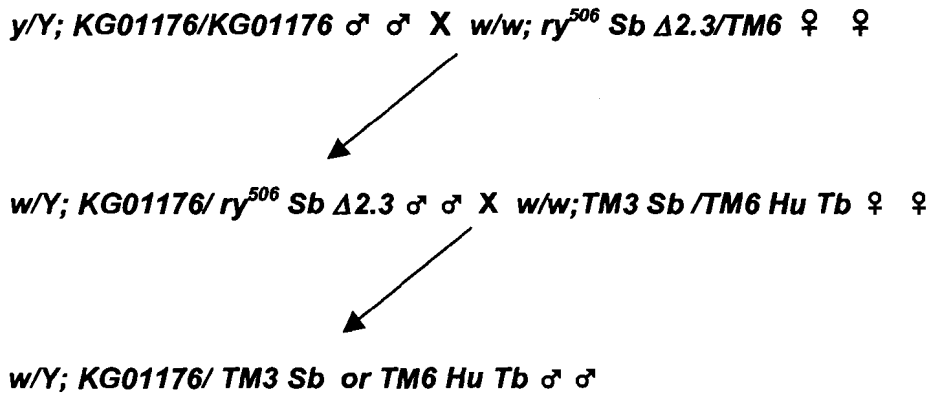


Figure 2.2. Schematic depiction of crosses to mobilize a P-element in distal h47.
KG01176/TM3Sb or *TM6 Hu Tb* were selected for altered *w+* expression and/or a *Minute* phenotype.



mutagenized third chromosome) were then crossed to *Df(3L+R)6B-29/TM3,Ser* (3-5 females per vial), and stocks of putative lethal mutations (i.e. lines in which no e^s */ *Df(3L+R)6B-29* were obtained) were established from e^s */*TM3, Ser* flies. Lethal lines obtained were crossed to *Df(3L)FX3/TM3Sb* and *Df(3L)1-166/TM3Sb* or *Df(3R)XM3*, to determine whether the lethal lesion had been induced in 3L (by a combination of the first two deficiencies) or 3R (the third deficiency). 3L lines were crosses to a subset of deficiencies in 3L to more precisely define their extents, and then to selected alleles of all lethal complementation groups within that interval. 3R lines were further delineated by complementation crosses to all known 3R lesions, and balanced with a *TM3SerGFP* balancer for further analysis (single embryo PCR, as described below), with the exception of the *13B-688* lesion. This mutation was lethal in combination with the *TM3Ser* and *TM3Sb* balancers. As the original male parent was still alive when this was discovered, it was crossed to *TM3/TM6* virgins and the line subsequently kept as a *TM6* balanced stock. The crossing scheme for the X-ray screen is depicted in Figure 2.3.

Screen for X-ray induced lesions in distal h47

Males bearing a P-element insertion, *KG03264*, near the *nrm* locus, were irradiated with 4000 Rads, allowed to recover for 24 hours and then crossed to unmated females of the genotype *w/w; TM3Sb/TM6 Hu Tb*. F1 male progeny were to be selected for loss of *w+* expression from the reporter gene on the

KG03264 P-element. However no flies were obtained that had discernable alterations in white expression. So, males collected (about 1000) were crossed to females of genotypes *Df(3L)FX3* or *Df(3L)MX18* to attempt to isolate deletions of distal 3L heterochromatin. Any chromosomes that showed lethality in combination with one of these deficiencies was balanced with TM3 and retained for further characterization. The crossing protocol is shown in Figure 2.4.

Male Recombination Screens

Male recombinants were obtained by crosses following the method of Preston et al. (1996). The generalized crossing scheme is shown in Figure 2.5. P-element/TM3SbSer males were crossed to *Bc/CyOΔ2.3; ri p^p/ri p^p* females. *+/ CyOΔ2.3; P-element/ri p^p* males were collected and mated en masse to *Df(3R)XM3/TM3 Sb Ser* females with 6-8 males and 30-50 females per bottle. Parents were removed to fresh bottles every three days and new virgin females were added as needed. *ri* p^{p+} */TM3SbSer* or *ri^+ p^p/TM3 Sb Ser* males were collected from the progeny. These were pair mated to 3-5 *Df(3R)XM3/TM3 Sb Ser* females and individual ri p^{p+} or *ri^+* p^p/TM3 lines established. Because the starter lines, *EP(3)3632* or *KG08740*, were originally viable in combination with *XM3*, lethality in combination with the *XM3* deficiency was taken as the first indication that a lethal lesion had been produced. The line *CH(3)1* was presumed to be distant from the proximal limit of the XM3 deficiency, so lethality was not

Figure 2.3. Schematic depiction of X-ray screen to produce new lesions in 3L and 3R heterochromatin.

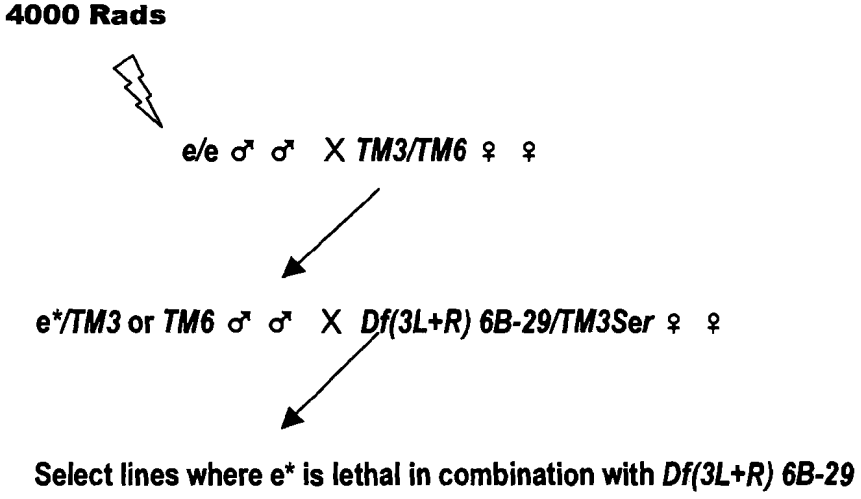


Figure 2.4. Schematic depiction of irradiation of P-element line *KG03264*.

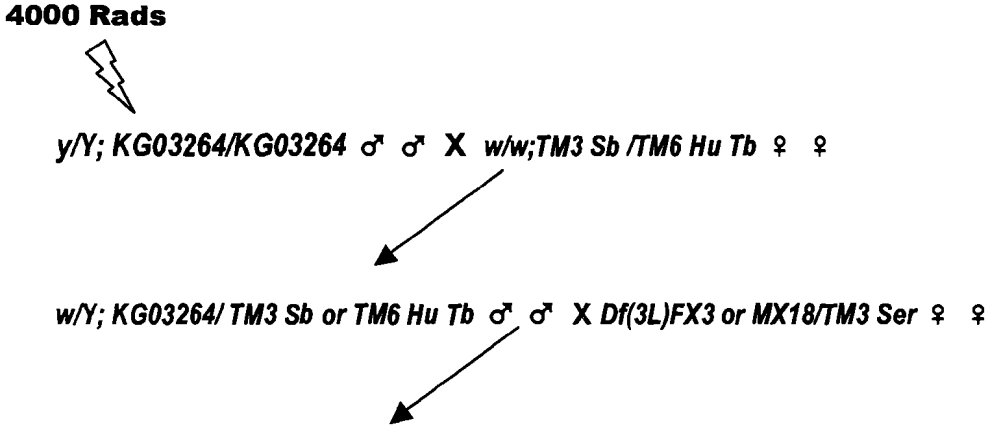


Figure 2.5. Generalized Crossing scheme for male recombination experiments.
In this schematic, individual P-elements used are not shown. See text for a description of the lines used.

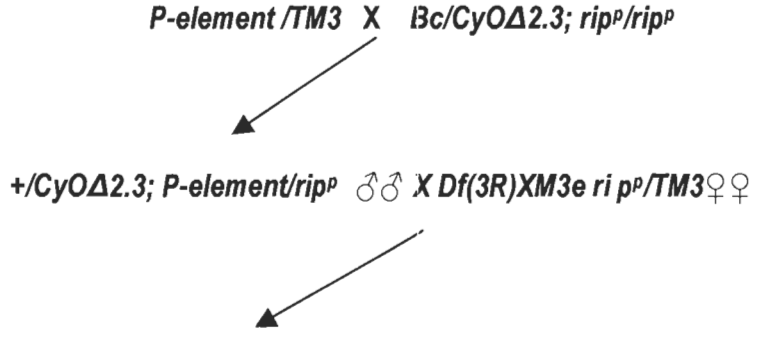
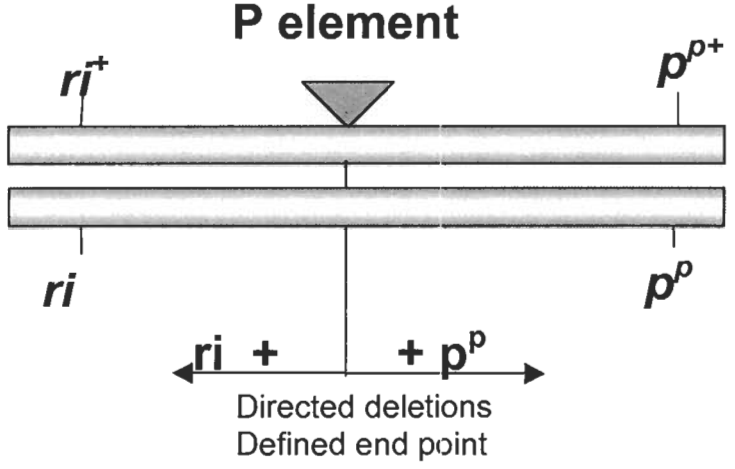


Figure 2.6. Isolation of potential deficiencies via male recombination.

The figure below depicts the site of a P-element with a triangle. Male recombination can yield deletions starting at the site of the P-element and extending either to the right (*ri*$^{+}$ *p*^{p}) or to the left (*ri p*$^{p+}$) as shown.



initially used as an assay to select lesions. The *P-GFP* lesion (generated in the screen depicted in Figure 2.1) was lethal in combination with the *XM3* deficiency, so the intention was that any recombinants that were lethal with the *XM3* deficiency would be retained and analyzed further. For distal 3L lesions, one of the starting lines was named *Rox80C*, a transgene inserted into the distal 3L heterochromatin, and whose flanking sequence indicates that the element is inserted into one of 3 hobo elements in this region. The other was a lethal excision derivative of *KG06133*, a line with two P-elements inserted into 3L heterochromatin, one near the *SNAP-25* locus and the other near the *n-Acetylcholine Receptor* gene in the h47 heterochromatin segment. The line had already been subjected to imprecise excision attempts and a number of lethal derivatives characterized by complementation. According to Preston et al. (1996), male recombination occurs as a result of an incorrect transposition by the P-element to a nearby site on the homologue. The markers on the recombinant chromosomes will indicate the direction of potential deletions: *ri* p^{p+} recombinants would be expected to generate deletions that extend to the left of the P-element, while ri^{+} p^{p} recombinants could generate deletions that would extend to the right of the element (see Figure 2.6 for a simple depiction). All male recombination experiments were performed at 25°C, as this was determined by Preston et al. to yield the greatest frequency of recombination. The lines were crossed to all known lesions in 3R or 3L heterochromatin – as appropriate - to determine whether a deletion had been isolated, and if so, the

extent of the deletion. New lesions were crossed *inter se* to determine which deleted common regions.

EMS screen for lethal mutations in chromosome 3 heterochromatin

An EMS screen has been initiated for lethal mutations in chromosome three heterochromatin. The procedure involves examining the previously mutagenized "Zuker" lines (Koundakjian et al. 2004) for mutations in the region of interest. Homozygous lethal lines are selected and then crossed to the following deficiencies: *Df(3L)Delta1*^{AK}, *Df(3L)FX53*, *Df(3LR)6B-29*, *Df(3R)XM3*. The deficiencies were selected to provide coverage for as much of the heterochromatin as possible. This screen is currently in progress.

PCR characterization of male recombination lines

Lines which were homozygous viable were collected as homozygotes and their DNA prepared (see below for method). Flanking primers were designed against genomic sequence on either side of the EP3632 line (EP1 and EP4, listed in Appendix B) and distal to the KG08740 line (EX1R, also listed in Appendix B); as the element is located 29 nucleotides from the limit of known sequence on 3R, it was not possible to design primers for the proximal side of the element. Primers were designed for both the second site of the *KG06133* line and for the region

near the *n-AChR-α* insertion site. PCR was performed where possible using flanking primers and a primer designed against the terminal repeats of the P-element (P-out) to determine whether the P element remained intact in the recombinant line, and whether or not a proximal or distal lesion had been created. The *Rox80C* lines could not be characterized in this way, given that the element was in the midst of a large repeat. Any rearrangement of the sequence to one side or another of the P element or a partial transposition of the element, would be expected to result in the lack of a PCR product to one side of the element, while leaving the other side unaffected.

Microscopy and Photography

Light microscopy

Wings were dissected from adult flies, rinsed in 95% Ethanol and mounted in aquamount. Bright field images were acquired using a Zeiss Axioplan microscope, with Northern Eclipse (version 5.0) software.

Scanning Electron Microscopy

Whole flies for scanning EM were killed by freezing and then stored in 70% ethanol. Flies in 70% EtOH were placed in microporous specimen capsules (Electron Microscopy Sciences, Fort Washington PA) and further dehydrated in 80 and 90% EtOH (minimum 10 min. at each step), followed by 4 x 10 min. in anhydrous EtOH. Specimens were then transferred to a Denton DCP-1 critical

point dryer (Denton Vacuum, Moorestown NJ) for drying. Dried specimens were kept in a desiccator until examination. Individual flies were attached at the posterior of the abdomen to aluminum foil points affixed to 10 mm aluminum SEM stubs with conductive colloidal carbon. This mounting allowed imaging the ventral thorax and lateral head portion of the same specimen, using a stage rotation converter described in Ehrman and Kaczmarska (2002). Specimens were coated from two sides with approximately 20 nm gold in a Hummer 6.2 sputtering unit (Anatech Ltd., Springfield VA) with argon as the source gas. Images were acquired using a JEOL JSM-5600 SEM (JEOL USA, Boston MA) operating at 10 kV and 39 mm working distance.

Molecular Biology

Characterization of P-element derived lines

Plasmid rescue

Plasmid rescue was attempted with some of the PZ and EP derived recombinant lines in an effort to obtain flanking genomic sequence. 1ug of genomic DNA was digested with the appropriate enzymes in a 20 ul total volume for 2 hours. 5ul of each reaction was run on an agarose gel to check the quantity and completeness of digestion. 10 ul of each digest was then used in a ligation reaction consisting of 60ul 5X buffer (supplied with enzyme) 2 ul 100mM ATP, 2ul T4DNA ligase (Pharmacia) and autoclaved distilled, deionized water (ddH_20) to 300 ul. Ligation

took place overnight at 4°C. Ligations were precipitated, and resuspended in 50 ul ddH_20. 2ul was used to transform DH5α cells, which were plated on kanamycin plates (25ug/mL). The enzymes used to attempt to obtain flanking genomic sequence were XbaI and Spe I in the case of PZ lines, and EcoRI and BamHI, in the case of the EP lines.

Inverse PCR

Flanking sequence from SuPorP lines could not be obtained by plasmid rescue, but by inverse PCR, according to the protocol described on the Baylor gene disruption website (http://flypush.imgen.bcm.tmc.edu/pscreen/) with the exception that PCR products were not sequenced directly but first cloned into the pCR2.1 vector, and sequenced using M13 forward and reverse primers. PCR was as described below. Inverse PCR was also attempted with lines derived from EP and PZ elements. The primers used for each element, along with the Tm for each primer, are given in Table 2.1. Note that for this approach, enzymes were selected that cut within the P-element and different pairs of primers could be used after ligation, for specific PCR amplification of either the 3' or 5' flanking sequence of the element (as specified in Table 2.1).

iPCR for the P-GFP mutant utilized a different approach. 1-5ug genomic DNA was digested with XhoI, KpnI or Not I, according to manufacturer's specifications, in a total volume of 50 ul for at least 4 hours at 37°C. The enzyme was

Table 2.1. Restriction enzymes, primer names and sequences and Tms used for inverse PCR.

The 5' and 3' beside each element indicates that either 5' or 3' flanking sequences should be specifically amplified by using the indicated primers.

Element	Rest. Enz	primer	primer sequence	Tm
EP 5'	HinPl I	Plac1	CACCCAAGGCTCTGCTCCCACAAT	60°C
		Pwht1	GTAACGCTATCACTCCGAACAGGTCACA	
EP 3'	HinPl I	Pry1	CCTTAGCATGTCCGTGGGGTTTGAAT	55°C
		Pry4	CAATCATATCGCTGTCTCACTCA	
P{PZ} 5'	HinPl I	Plac1	shown above	60°C
		Plac4	ACTGTGCGTTAGGTCCTGTTCATTGTT	
P{PZ} 3'	HinPl I	Pry1	shown above	55°C
		Pry4	shown above	
SuPorP 5'	HpaII	Plac1	CAC CCA AGG CTC TGC TCC CAC AAT	60°C
		Pwht1	shown above	
SuPorP 3'	HpaII	3.rev.hpa2	TTG CCA CTT GCT CAT ACG TC	55°C
		Pry4	shown above	

inactivated at 65°C for 20 minutes and 5-15 ul of digested DNA was used in ligations of 100 -300ul overnight at 16°C. PCR reactions utilized the P-out primer and a Tm of 55°C. The cycling profiles were similar to those for single embryo PCR, described below.

Bacterial and phage strains, vectors and culture conditions

The *E.coli* bacterial strains XL1-Blue and DH5α were used to propagate plasmid DNA (pBluescript, pCR2.1, plasmid rescue products) by electroporation (XL1-Blue) and heat shock (DH5α) transformation. Bacterial cultures were routinely grown in 2YT medium (Sambrook et al. 1989) with appropriate concentrations of antibiotics where necessary (100ug/mL ampicillin , 50 ug/mL kanamycin, or 25ug/mL chloramphenicol).

Isolation of Plasmid DNA

Plasmid DNA for general use (restriction digests, generation of probes, subcloning of genomic DNA and stock maintenance) was isolated using standard alkaline lysis procedures as described in Sambrook et al. (1989). Where necessary, quantitation was performed using an Ultrospec III (Pharmacia).

Isolation of genomic DNA

Isolation of Adult DNA

Drosophila genomic DNA was isolated using the method outlined by Jowett (1998). The procedure outlined was designed and used to extract DNA from a sample of 200 flies, but was often scaled down for smaller numbers, so the whole procedure could be carried out in 1.5mL eppendorf tubes. In this variation, up to 30 flies were homogenized in 750ul Fly Lysis Buffer (100mM Tris-HCl pH 8, 50mM NaCl, 50mM EDTA, 1%SDS, 0.15mM spermine, 0.5mM spermidine). Two ul of a 10mg/mL stock of Proteinase K were added, and the preps were incubated in a 37°C water bath for up to four hours. The preps were then phenol extracted once, phenol:chloroform extracted twice, and chloroform extracted once (equal volume). After ethanol precipitation, the pellets were washed in 70% ethanol, dried in a vacuum dessicator and resuspended in 70ul of deionized, distilled autoclaved water or TE8. RNAse was added to a concentration of 100ug/mL, and the samples were quantitated by spectroscopy, and run on an agarose gel to check integrity. This DNA was used for Southern analysis and PCR.

Isolation of DNA from single embryos

A modified version of the procedure described by Hatton and O'Hare (1999), was used to extract DNA from single embryos homozygous for specific mutations or deficiencies. Individual embryos from parents heterozygous for the mutation of interest transheterozygous with balancer chromosomes, each bearing a GFP transgene, were collected from apple juice agar plates, placed in 0.6mL

eppendorf tubes and homogenized in 13ul of embryo lysis buffer. 2ul of a 20ug/ul stock of proteinase K were added, and the preps were incubated in a 37°C water bath for 30 minutes. Thereafter the preps were heated to 99°C in a PCR machine for 10 minutes, and the DNA was used within 1 -2 weeks in PCR reactions to identify homozygous mutant embryos. Originally, visual inspection of embryos was used to identify those which expressed GFP, and thus carried the balancer chromosome. However, subsequent PCR of embryos selected in this way, indicated that a proportion of the "non-GFP" embryos did, in fact, generate a PCR band with GFP specific primers. Thus homozygous deficiency embryos were selected by PCR rather than by microscopic inspection. Embyros the DNA of which failed to amplify a GFP band, but did yield a control PCR band, were selected for genome amplification and further PCR characterization.

To obtain enough genomic DNA from homozygous deficiency embryos, genomic DNA was amplified with the Genomi-phi kit, according to the manufacturer's specifications (Amersham). The DNA obtained was quantified by spectroscopy and run on an agarose gel for confirmation; it was then diluted to approximately 200 ng/ul with TE8 for use in PCR reactions.

Restriction digests

All restriction digests were carried out with enzymes purchased from Invitrogen, Roche, or NEB, according to their protocols. For plasmid digests, 200-1000 ng of

DNA was digested with 1-3 units of enzyme at 37°C, for minimally 1½ hours. For genomic DNA digests used in Southern analysis, 10-20ug of DNA was digested in a volume of 300ul, using up to 120 units of enzyme (Invitrogen high concentration enzymes, 40u/ul), and these digests were carried out for minimally 4 hours. In all cases, reactions were stopped at 65°C for ~20 minutes, and the genomic digests were precipitated and resuspended in a volume suitable for gel electrophoresis (usually 20 ul).

Agarose Gel Electrophoresis

DNA gels

DNA samples were loaded on 0.7%-1.5% agarose gels, depending on the size of the DNA fragments. Genomic digests were generally run in 0.7% gels overnight. Ethidium bromide was added to a final concentration of 500ng/ul for visualization of the DNA. All gels were run in 0.5X TBE (Sambrook et al. 1989). Gels were photographed with a UVP documentation system.

Isolation of DNA from agarose gels

DNA was digested with restriction enzymes and electrophoresed through agarose gels in clean 0.5X TBE as described above. If the digestion products were to be used for cloning, no pictures under short wave UV light (302nm) were

taken. The band(s) were excised from the gel under long wave UV light (360 nm), cut into pieces, and extracted from the agarose according to the GFX PCR DNA and Gel band purification kit (cat. no. 27-9602-01)

Cloning

Cloning PCR products

PCR amplified DNA was gel extracted as described and cloned into the T-tailed vector pPCR 2.1 using the TA cloning kit (Invitrogen). The protocol was followed in general, except that the One-Shot™ competent cells were sometimes replaced with DH5α, and 2YT medium was often used in place of SOC.

Southern Analysis

DNA transfer

Genomic DNA, isolated as described above, was cut with appropriate restriction enzymes. Gels were treated as described in Sambrook et al. (1989), with the following variations. The denaturation step was carried out using a solution of 1.5N NaCl, 0.5N NaOH, for 20-30 mnutes on a shaking table, followed by 2 neutralization washes (0.5MTris-HCl pH7.5, 1.5N NaCl) of 20 minutes each. Gels were rinsed with deionized water between washes and the depurination

step was eliminated. Transfer to positively charged nylon membrane (Hybond N^+, Amersham) was carried out overnight, using 10X SSC as transfer buffer. The membrane was rinsed briefly in 6X SSC and then the DNA covalently linked to the membrane using a UV Stratalinker (Stratagene).

Probe labeling and hybridization conditions

Template insert DNA was digested and purified away from vector sequences as described above (GFX). 20-100ng was labeled with ^{32}P, following instructions from a Boehringer Mannheim random-primed DNA labeling kit (cat. no.1-004-760). Blots were prehybridized at 65°C for at least 1 hour in FSB-7% SDS (50mM Sodium pyrophosphate, 100mM Sodium dihydrogen phosphate, 7% SDS), with 100 ug/mL denatured herring testis or salmon sperm DNA used as a blocking agent. Hybridization took place overnight in the same buffer at the same temperature, and the blots were washed 2-3X in FSB-1%SDS, also at 68°C. After washing, the blots were drained, wrapped in Saran, and exposed to X-ray film, (Amersham Hyperfilm™ MP) in a light-tight cassette with an intensifying screen, usually at -70°C.

PCR amplification

Taq polymerase from Invitrogen was used to amplify DNA for cloning, preparation of template DNAs for probes and general diagnostic tests.

Templates were either cloned or genomic and conditions used for amplification were as described in the manufacturer's protocols, with some minor changes. Reactions were almost always 25ul, but if the volume increased, all amounts were scaled up accordingly. For cloned template, 10-100 ng was used per reaction, and for genomic, 100-500ng was used.

Reaction mix

For single band amplification with Taq, the reaction mix was made as follows: 2.5ul 10X buffer (supplied with enzyme), 1ul 25mM Magnesium, 2ul 2.5mM NTPs, 1ul each of 10uM primer working stocks, 1-2.5UTaq polymerase, and deionized distilled autoclaved water to a final volume of 25ul. For multiplexing, the amount of dNTPs added was increased to 4ul. For inverse PCR, the NTPs were reduced to 1 ul.

Cycling profiles

All PCR reactions were carried out using a PCR Sprint (Thermo Hybaid) machine. For single embryo PCRs the cycling profile used was as follows. A denaturation step of 2 minutes at 95C was followed by 30-35 cycles of 95°C for 30s, Tm for 40s and 68°C for 50s. A final elongation step was performed at 72°C for 5-7 minutes. The Tm depended on the melting temperature of the primers

employed, and the Tm used for primers other than the inverse PCR primers can be found in Appendix B.

The inverse PCR cycling profile was as follows. A 5-minute 95°C denaturation step was followed by 35 cycles of 95°C for 30s, Tm for 1 min, 68°C for 2 min. A 10-minute elongation step was performed at 72°C. 5-10 ul of the PCR reaction was run on an agarose gel. If a single band was apparent, with no background, the PCR reaction was used to clone the fragment into pCR2.1. If any suggestion of background or multiple bands was apparent, the bands were gel purified and used for cloning.

DNA Sequencing

Templates that required sequencing were cloned first, and purified as described above (Plasmid DNA preparation). All samples were prepared for automated sequencing as follows: 100ng of template per kilobase of DNA (including vector), primer added to an approximate final concentration of 300 picomoles, and deionized distilled water to a final volume of 12ul. Samples were then sent to the University of Calgary Core DNA and Protein services (www.ucalgary.ca/~dnalab) for sequencing.

CHAPTER 3: CHARACTERIZATION OF 3R HETEROCHROMATIN

Introduction

The heterochromatin of 3R has been much less thoroughly studied than that of other chromosomes. At the onset of this work, only three deficiencies in 3R heterochromatin were available. These had been generated by Marchant and Holm (1988a) and were all that remained from that study. Cytologically, 3R heterochromatin has a somewhat different appearance than other heterochromatic regions in that it appears to have a more abrupt heterochromatin/ euchromatin boundary (Gatti and Pimpinelli 1992). This may imply that 3R heterochromatin is more likely to contain highly repeated satellite sequences and less likely to house unique sequence interspersed with transposons. Similarly, the sequence across the boundary between the heterochromatin and euchromatin is not very extensive, and the recent updates by the genome project have provided surprisingly few inroads into the heterochromatic sequence. For instance in Release 1, the proximal-most 3R genomic scaffold was AE003607 and the proximal-most gene was a putative protein kinase, with Celera Genome number: *CG12581*. In subsequent releases, no further genes have been identified and the sequence has been extended towards the centromere by a scant 751 nucleotides (Hoskins et al. 2002). By contrast, during the same period of time, the sequence on the 3L heterochromatin/euchromatin boundary in Release 1 ended with the *neuromusculin (nrm)* locus and genomic scaffold AE003599, and Release 3 has

extended the sequence by approximately 300 kilobases and an additional seven loci have been identified in the region (Hoskins et al. 2002).

The heterochromatin of the right arm of chromosome 3 has been divided into segments h53 through h58. These segments are distinguished based on banding patterns under different staining protocols. The 3R heterochromatin is home to a number of satellites and transposable elements. Specifically, a 12 nucleotide repeat called the dodeca satellite (Abad et al. 1992) is found in a large block on segment h53, near the centromere. Although this satellite is very close to the centromere, reduction of the size of the satellite (which is normally around 1Mb) does not appear to interfere with chromosome stability (Carmena et al. 1993). The dodeca satellite is one of the few with a single known location in *Drosophila* chromosomes, and thus can be considered a marker for the third chromosome. A large block of (AAGAG)n satellite is found at segment h57 (Lohe et al. 1993). It is possible that this satellite block is responsible for the difficulties encountered in extending the sequence information further inward on 3R.

Prior to the onset of this study, there were only two essential loci identified in 3R heterochromatin: *l(3R)81Fa* and *l(3R)81Fb* (Marchant and Holm 1988b). Several deficiencies had been created through a compound autosome attachment-detachment method. Figure 1.2 shows a schematic representation of how this method works. The available deficiencies were: *Df (3L+R)6B-29*, which appeared to delete both 3R loci as well as several loci on 3L; *Df(3R)10-65*, which

reportedly deleted only *l(3R)81Fa*; and *Df(3R)4-75*, which apparently deleted both loci. No original alleles of *l(3R)81Fa* or *b* remain, so that it has not been possible to determine whether any newly isolated loci are alleles of these genes. Although a number of difficulties may result from using the attachment-detachment method to generate lesions (including the likelihood that many such lesions are quite complex rearrangements and thus difficult to characterize fully), this work nonetheless was instrumental in providing the first entry points into the heterochromatin of chromosome 3.

Two P-element induced mutations were also available: *CH(3)1* and *CH(3)4*. These were isolated by Zhang and Spradling (1994) in a large-scale attempt to isolate P-mutations in heterochromatic loci. Both mutants are homozygous lethal and lethal in combination. *CH(3)1* was determined to reside at segment h55 by *in situ* hybridization (Zhang and Spradling 1994). Although *CH(3)4* failed to complement *CH(3)1*, no P-element was detected on 3R; rather the P-element for this mutant resides in 3L and is a lesion in *lethal 6* there. The 3R second site lethality on the *CH(3)4* chromosome may have arisen from a so-called "hit and run" event, in which a P-element inserts into a region and then excises from it, leaving behind a slightly altered sequence (Tulin et al. 2002). Though *CH(3)1* is homozygous lethal, it is viable in combination with all three of the deficiencies in 3R. Since the P-element had been localized cytologically to segment h55, it appeared that the locus in question was fairly deep in 3R heterochromatin. At that time, the affected locus was not known, but the viability of the mutant in

combination with our 3R deficiencies indicated that there may well be further loci to be discovered in the 3R heterochromatin.

The paucity of genetic resources with which to study 3R heterochromatin presented a major challenge. A first step was to survey known deficiencies and mutations near the heterochromatin boundary, with the aim of locating further lesions in the region. The next step undertaken was to try to isolate more reagents in order to increase the resolution at which we could study the resident genes. The third step was to attempt to use these resources to identify at least some of the coding sequences absent under the deletions available and perhaps to connect known coding sequences with lethal phenotypes.

P-element insertion was employed to isolate mutations in additional loci in 3R. I also utilized P-element mediated male recombination in an attempt to create directed deletions extending away from the original site of the P-element (Preston et al. 1996). The advantage of using P-element mediated male recombination is that (in most cases), one terminus of the deletion created will be known on a molecular level, and since many lesions created by this method retain the original P-element, it is often possible to define both endpoints molecularly, as well as to obtain sequence "tags" from flanking DNA (Preston et al. 1996). Using this method, a series of polar deficiencies could in theory be isolated that would yield information about the number and order of loci in a given direction. Preston et al. (1996) utilized a well-characterized P-element insertion

(*pCaSpeR*) on the second chromosome. 23,188 males were examined, of which 1.05% were recombinants. Of these recombinants, approximately one third were associated with deletions, another third with duplications and those remaining were associated with no rearrangements. In approximately 53% of the resulting recombinant lines, the P-element was retained, thereby providing a possible means for examining the flanking sequences and identifying both breakpoints of the lesion in question.

The disadvantage of P-element mobilization is that one is confined to the P-element lines available and of those known at the time of the onset of this work, very few were placed in convenient locations. The frequency of P-element insertion into heterochromatin appears to be low, either because insertion into the more compact heterochromatin occurs less frequently relative to euchromatin, or because insertions into heterochromatin are less easily detected due to inactivation of reporter genes on the elements. Evidence from the Karpen lab (Konev et al. 2003) would argue for both. The use of an attached X (X^X/Y) background in a series of mobilization experiments apparently increased the frequency of insertions into heterochromatin, as assayed by the variegation of the reporter gene. The presence of two X chromosomes and one Y would introduce additional heterochromatin into the nuclei of such individuals. This is thought to have a "titrating" effect, whereby additional heterochromatin reduces the availability of heterochromatin proteins to other heterochromatic regions of the genome. The overall effect could be somewhat less condensed

heterochromatin. This is the basis upon which the presence of an additional Y chromosome is thought to suppress PEV, for example (Gowen and Gay 1934). This implies that by "loosening" the heterochromatin, P-element insertion can be enhanced as has been shown by Konev et al. (2003). In addition, using a *yellow*$^{+}$ marker instead of the more traditional *white*$^{+}$, increased the frequency of detectable insertions, implying that a marker less susceptible to position-effects could allow identification of insertion events that might otherwise have been missed. Finally, some P-elements inserted in heterochromatin appear to be located somewhat distantly from the genes they are influencing, so that in some cases, obtaining flanking sequence may not yield useable information regarding the genes affected (D. Sinclair, personal communication).

P-element insertions into 3R heterochromatin might be expected to present similar challenges. Of the elements available for these studies, none was as well characterized as the euchromatic insertion used by Preston at al. At the onset of this work, the generation of deletions using male recombination had been attempted with heterochromatic insertions, on a small scale, using a P-element near the boundary of chromosome 2L to generate some rearrangements there (Howe et al. 1995), so it was not certain that it would work well on chromosome 3, nor whether it would work at all for insertions deeper in heterochromatin. Subsequently, Sun et al. (2004) studied the effect of alterations in the chromosome sequence surrounding particular chromosome 4 P-element insertions on the expression of the reporter gene on the element. They used a

variety of methods to alter sequences near the reporter genes, and one pair of screens used male recombination to try to create duplications or deletions of flanking sequence on either side of the P-insert. Theirs was the first demonstration that this method was feasible with respect to the chromatin of the fourth chromosome, which has a somewhat unusual arrangement. The heterochromatin and euchromatin on chromosome 4 are interspersed along the banded region of the chromosome, with a block of "typical" heterochromatin surrounding the centromere, as with other chromosomes.

An X-ray screen was also employed to avoid the potential drawbacks of using P-elements – including the possibility that the method might not work efficiently. The disadvantages of using X-irradiation as a mutagen include the fact that one cannot control the start and end point of the lesions isolated and that there are no convenient/reliable means to identify sequence at the endpoints of the lesions. Also, the frequency of more complex rearrangements such as inversions (even multiple inversions) and translocations is much higher. In addition, some evidence from studies on the Y-chromosome suggests that not all regions of the heterochromatin are equally susceptible to X-rays (Gatti and Pimpinelli 1983), so that the method has its own bias. However, X-ray mutagenesis does not depend on the sites of pre-existing elements and as such, may be less biased than the P-element method. The combination of both methods was expected to yield a reasonable number of useful deficiencies; the aim was to isolate perhaps 10 -15 new lesions, of varying extents. Both extensive and small lesions would be of

value. The small lesions would be helpful in creating a more detailed genetic map. The largest deficiencies could be used as a basis for EMS screens to attempt saturation mutagenesis of the 3R heterochromatin.

Previous screens for EMS mutations (Marchant and Holm 1988b) have yielded alleles of only two loci, but it may be that this was because the lesions used for selection deleted only a small subset of the 3R heterochromatin. Marchant and Holm (1988b) obtained a number of lethals in combination with *Df(3R)10-65* and *Df(3R) 4-75*. Both of these deficiencies were subsequently examined cytologically. Koryakov et al. (2002; 2003) have shown that *Df(3R)10-65* deletes much of segment h58, but none of the more proximal 3R heterochromatin whereas *Df(3R)4-75* is in fact a multi-point inversion with no cytologically detectable deletion. *Df(3L+R)6B-29* was observed to delete much of the 3R heterochromatin, but part of the 3R heterochromatin appeared abnormal and this may be due to the fact that the chromosome contains rearranged heterochromatin originating from any number of segments. In any case, this apparently larger deletion was not used in the original screens of Marchant and Holm.

More recently, a screen was reported in which *Df(3L+R)6B-29* was used as one of the deficiencies over which to isolate new radiation or P-element induced mutations (Schulze et al. 2001). Over 7,000 chromosomes were tested in combination with the *6B-29* deficiency in the γ-radiation screen, and 70 mutations

isolated; however none of these identified new loci in 3R (or 3L). 5 new alleles of *l(3R)h1* (otherwise known as *l(3R)81Fa*) were isolated and 3 alleles were identified that did not map into any known smaller deficiencies across the heterochromatin. Unfortunately, these alleles were not characterized further and no longer exist. The P-element screen yielded 8 mutants, from 8400 cultures tested; all of these were alleles of previously identified loci on 3L. The fact that no 3R loci were identified in these screens supported the notion that new lesions would be needed to distinguish between two possibilities: that 3R is decidedly gene poor compared to other heterochromatic regions or, alternatively, that there is something unusual in the particular configuration of the *6B-29* chromosome which precludes the identification of essential loci in the region. *CH(3)1* was a case in point; clearly a P screen yielded that lethal mutation in 3R heterochromatin and yet the mutation was viable in combination with the *Df(3L+R)6B-29* chromosome – despite cytological analysis that indicated that *Df(3L+R)6B-29* deleted almost all of the 3R heterochromatin. It was of prime importance, therefore, to try to obtain larger, more useful lesions to utilize in mutagenesis screens for essential loci.

Thus the aims of this phase of the research were: to create new lesions in 3R by P-mutagenesis, P-element mediated male recombination and X-irradiation, to compare the results of the two latter approaches; and to begin to correlate coding sequences with deficiencies under which they were missing. A longer-term goal, not encompassed by this project is the correlation of coding sequences with

mutant alleles of 3R loci. The work has resulted in the isolation of a new set of genetic resources, which should substantially improve the picture we have of this rather difficult region of the genome and raised questions about the enigmatic nature of at least one locus in this region.

Results and Discussion

Complementation with newly obtained deficiencies and EMS alleles.

At the onset of the study, I obtained a set of deficiencies from the Bloomington Drosophila stock center; these mapped in or near 3R heterochromatin. I tested them for complementation with the pre-existing deficiencies and mutations. Several deficiencies and two P-element induced mutations were tested and two deficiencies: *Df(3R)XM3* and *Df(3R)A321R,* both failed to complement the 3R heterochromatic lesions already available. Simultaneously I had the opportunity to test a series of EMS induced alleles produced during a screen for *abstrakt (abs)* alleles (Irion and Leptin 1999). *abs* is located in 81F/82A, quite near the heterochromatin/euchromatin boundary on 3R. One of the complementation groups identified – designated at the time *L-1/8* (for Leptin #1 and #8) – failed to complement some of the deletions. These could be alleles of the original *l(3R)81Fb* based upon their failure to complement *Df(3R)4-75,* while complementing *Df(3R)10-65.* Figure 3.1 shows the complementation map produced by these initial tests. At the time these were performed, no cytological data were available against which to position deficiencies. However, the figure incorporates some cytological information that was generated after these experiments were completed. Placement of *CH(3)1* at h55 is based on the report of Zhang and Spradling (1994). Koryakov et al. (2002) suggested that *Df(3L+R)6B-29* deletes most of h54-h58 on 3R; however complementation of

Df(3L+R)6B-29 with *CH(3)1* indicated that some of the DNA might still be present; thus this deficiency is represented by a dashed line just distal to the *CH(3)1* insertion point, and extending towards the centromere. Koryakov et al. (2002; 2003) showed that *Df(3R)4-75* was a multiple inversion rather than a simple polar deficiency. Based upon complementation results, one breakpoint must reside within the *Df(3R)10-65* interval, and another at *L-1/8* as indicated in Figure 3.1; however there are other breakpoints associated with this chromosome (for instance, at h50 on 3L). Though there is no cytologically detectable deficiency in *Df(3R)4-75*, it is possible that there is a small deficiency which overlaps with that of *Df(3R)10-65*, but is too small to be detected by standard cytology.

Isolation of a P-mutant, P-GFP, in heterochromatin

A P-element induced mutation, lethal in combination with *Df(3L+R)6B-29*, was obtained from the screen depicted in Figure 2.1. The starting strain was *p{UAS-GFP.Y}B1* (Yeh et al. 1995), a homozygous viable insert in the euchromatin of chromosome 3R. The P-element was constructed by insertion of a GFP reporter gene into the pPUAST vector. A schematic map of the element based on information in flybase, is shown in Figure 3.2. A total of 768 males that were *p{UAS-GFP.Y}B1/TM6* were selected and individually crossed to *Df(3L+R)6B-29/TM3Ser* females. Of these, 256 cultures gave no progeny due to sterility, mould or bacteria. From the remaining 512 lines, one was obtained that failed to

Figure 3.1. Initial genetic map based upon complementation data. The filled rectangle represents the 3R heterochromatin. Lines below the rectangle indicate presumed deficiency intervals. Thicker lines indicate postulated extents of lesions generated by Marchant and Holm (1988b). Thinner lines indicate the more recently identified deficiencies. A dashed line represents a region of uncertainty regarding the limits of *Df(3L+R)6B-29* (though it complements *CH(3)1*, cytological analysis indicates much of the heterochromatin of 3R is removed by this deficiency). *Df(3R)4-75* is shown as an inversion with two arrows indicating postulated breakpoints. *L-1/8* indicates the postulated position of the essential complementation group isolated by Irion and Leptin (1999). This map is not to scale: *Df(3R)10-65* deletes most or all of cytological region h58 (Koryakov et al. 2002), which represents just a small portion of the 3R heterochromatin. However, *Df(3R)10-65* is disproportionately expanded in this figure.

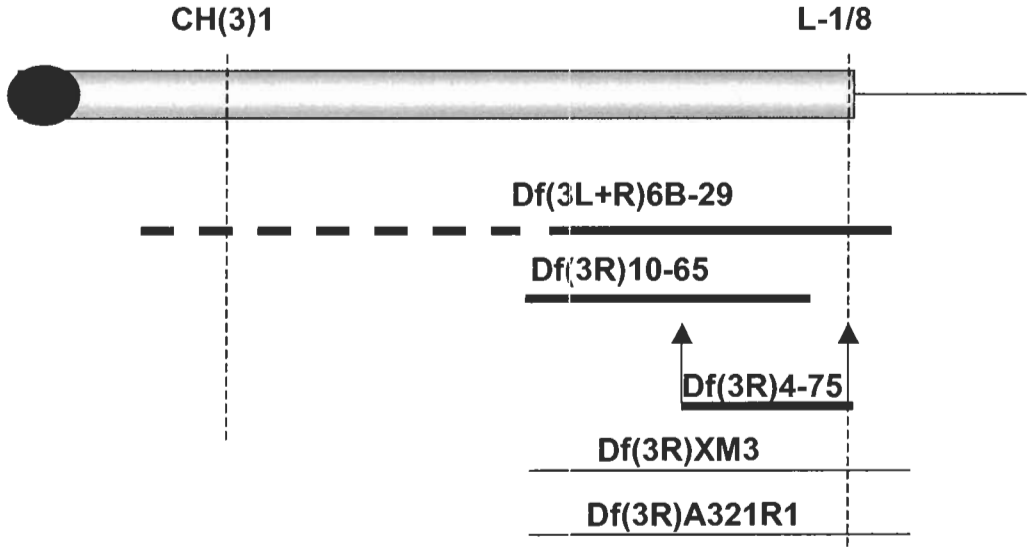


Figure 3.2. Schematic diagram depicting element used to generate the *PGFP* mutant.

The original vector of the P-element line, p{*UAS-GFP.Y}B1*. Thickened black bars represent small (less than 40 nt) linker regions containing restriction enzyme sites added to facilitate vector construction. The total length is approximately 6.5 kb.

complement *Df(3L+R)6B-29*. This mutant was named *P-GFP*. It is lethal in combination with *Df(3R)4-75*, *Df(3R)XM3*, *Df(3R)A321R1* and *Df(3R)10-65* and in all cases *P-GFP/Df* transheterozygotes are lethal during the pharate adult stage. Pharate adults show no obvious morphological defects (not shown). In combination with *Df(3R)A321R1*, a small number of visibly normal (except for a raised wing phenotype) transheterozygous adults eclose. These die within the first day of eclosion.

It is interesting that this rather small-scale screen yielded a mutation in 3R heterochromatin when a much larger scale P-element screen (Schulze et al. 2001) did not. The latter screen involved the examination of more than 10 times as many mutagenized chromosomes as this one, but yielded only alleles of *lethal 2* and *lethal 3* on 3L heterochromatin. Many variations between the two experiments might explain this difference but the most significant may be the type of element used. The Schulze et al. screen utilized wild type P-elements from a *Birmingham 2* strain. The results of Zhang and Spradling (1994) provide additional evidence for the possible influence of P-element type on insertion site preference; their work involved the mobilization of pPZ elements and yielded a different pattern of insertions across the 3L and 3R heterochromatin.

Reversion experiments indicated that the P-element was intact; or at least enough of the P element termini were retained to revert the lethality in combination with *Df(3R)10-65*. Revertant lines were obtained that were both

white^{+} and *white*$^{-}$, indicating the likelihood of a second site insertion; most likely the original insertion site. The line *P-GFP-15,* referred to below, was isolated in this experiment. The *white*$^{+}$ expression from the original *P-GFP* mutant was moderate, with no evidence of variegation. To confirm that this *white*$^{+}$ expression was associated with a second site insertion, a simple recombination experiment was undertaken. *w/Y P-GFP/TM3* males were crossed to *Oregon R* females and *non-Sb* transheterozygous virgin females collected. These were crossed to *w/Y TM3Sb/TM6HuTb* males. None of the resulting *white*$^{+}$ progeny showed any variegation of *white*$^{+}$. 50 *w/Y TM3* and *TM6* (*w*$^{+}$ and *w*$^{-}$) males were collected. Each of the males was crossed to *Df(3R)10-65/TM3Ser* virgin females. Among the first 10 crosses examined, two were found to be both *white* $^{-}$ and pupal lethal in combination with the deficiency (the remaining lines were discarded, as the hypothesis had been confirmed). The line *P-GFP-4*, referred to below, was isolated in this experiment. The fact that *white*$^{-}$ lines retaining the lethality of the original *P-GFP* mutant could be generated, suggested that the heterochromatic insert in the *P-GFP* mutation was not expressing *white*$^{+}$; rather a second site P-element had been responsible for the pigmentation originally observed.

The fact that the *white*$^{+}$ reporter in the heterochromatin does not express any detectable eye pigment, could be due to either position-effects, or an internal rearrangement in the P-element. In the former case, the element would retain the *white* coding sequence, but its expression would be repressed due to the chromatin structure surrounding the element. In the latter case, deletions or

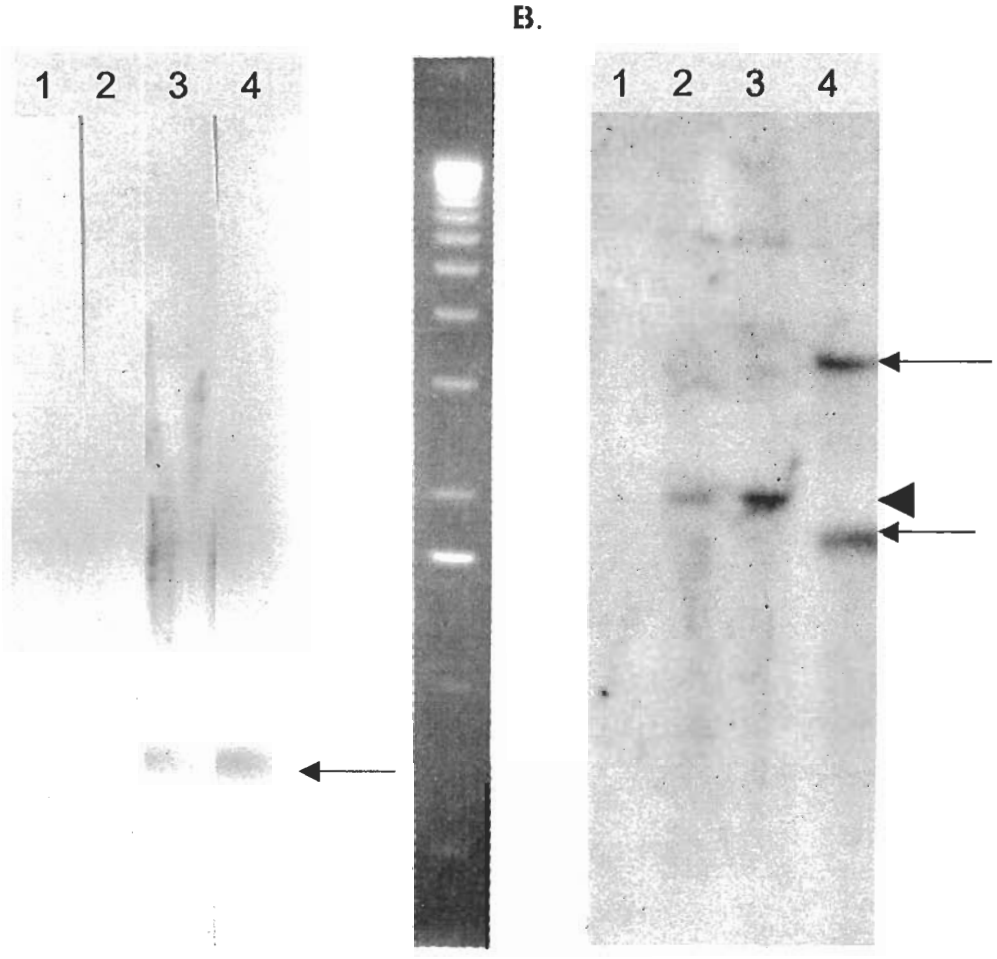
other rearrangements within the P-element would prevent expression of the reporter, regardless of position. A very simple set of experiments was performed to distinguish between these two alternatives. First, the *P-GFP* strain was grown at 29°C. It is known that position-effect variegation is affected by temperature. However, when grown at the higher temperature, the line did not show any *white*$^{+}$ expression. Secondly, the *P-GFP* strain was crossed into a background bearing the *Su(var)2-5*^{01} mutation, which encodes a non-functional form of HP1. This mutation strongly suppresses PEV, yet the *P-GFP* line still did not express any pigment in this background. Finally the line was crossed to a Δ2.3 strain in order to mobilize the element. Several bottles of *P-GFP/Δ2.3* males crossed to a large number of *w/w; TM3/TM6* females were grown at 25°C. The progeny were examined for expression of *white*$^{+}$, which would indicate that the P-element had transposed into a more amenable environment for *white* expression. Well over 1000 progeny were examined, many more than in the reversion experiment described previously. No progeny were observed with any *white* pigment expression. The most reasonable conclusion from these experiments is that the P-element in the *P-GFP* line has been rearranged and lost at least part of the *white* reporter gene sequence.

A genomic Southern that shows the apparent lack of GFP sequences in the heterochromatic insertion, supports this conclusion. This experiment was performed using the *P-GFP* line, the original line from which *P-GFP* was derived (p{*UAS-GFP.Y}B1*), and two lines isolated in the reversion and recombination

experiments as described above. One was designated *P-GFP-4*; in this line a *white*$^+$ expression had been lost, but the line continued to exhibit pupal lethality in combination with the *Df(3L)10-65*. The other line, designated *P-GFP-15*, had no *white*$^+$ expression and was viable in combination with both *Df(3L)10-65* and *P-GFP,* suggesting that the element responsible for the *P-GFP* mutation had excised. Genomic DNA for the Southern blot was cut with *PstI* which was expected to yield a small fragment containing only the GFP fragment, and two flanking fragments containing either side of the P-element plus flanking DNA. The Southern blot was first probed with a 700bp GFP probe, generously provided by Dr. S. Ner. After 8 weeks, to allow the signal to diminish, the same blot was re-probed with the first third of the P-element, which should have hybridized to only one of the two flanking P-element bands. The results are shown in Figure 3.3. The *P-GFP* and original p{*UAS-GFP.Y}B1* line both showed a 700 bp GFP band, but neither of the other two lines showed this, strongly suggesting that the GFP coding sequence had been lost from the P-element inserted into the heterochromatic site, whereas the second site P-element was intact. On the blot probed with the P-element fragment, the p{*UAS-GFP.Y}B1* line shows two bands approximately 1.7 and 3.1 kb in size. The *P-GFP* and *P-GFP-4* lines share a band in common, approximately 2kb in size, presumably the heterochromatic insert. However, there is no clear second site band shown for *P-GFP*, so unless the fragment associated with the second site P-element is fortuitously also 2kb, the data are puzzling. Certainly the fact that the *white*$^+$ expression could be removed from the chromosome by recombination without eliminating the *P-GFP*

Figure 3.3. Genomic Southern blot, probed with GFP (**A**) and the first third of the P-element (**B**).

Genotypes are as follows: Lane 1: *w; PGFP-15/PGFP-15*, a homozygous viable revertant line; Lane 2: *w; PGFP-4/TM3Sb*, a line which has lost *white*$^{+}$ expression but retains pupal lethality in combination with *Df(3R)10-65*; Lane 3 *P-GFP/TM3Sb*, which is thought to have a second site P-insertion; lane 4: p{*UAS-GFP.Y}B1/* p{*UAS-GFP.Y}B1*; the homozygous viable P-element line from which *P-GFP* was derived. The arrow in A indicates a 700 bp band, while the arrows in B indicate bands of 1.7 and 3.1 kb. The arrowhead in B indicates a band of 2kb. The 1kb DNA ladder from the gel is shown between the two blots.



mutant, strongly argues that a second site element had been present on the *P-GFP* third chromosome.

It is possible that *P-GFP* is an allele of the original *l(3R)81Fa* of Marchant and Holm (1988b); however this cannot be tested. *P-GFP* has a complementation pattern identical to that of *l(3R)81Fa* as far as can be determined based on lethality with the same deficiencies. However, the time of lethality differs; the original *l(3R)81Fa* alleles isolated were embryonic or larval lethal, whereas the *P-GFP* mutant flies reach the pharate adult stage. This mutant is likely not a null mutation, which could explain its later lethal phase.

Attempts were made to use inverse PCR (iPCR) to isolate flanking sequence from this mutant in order to determine with what locus it is associated. To date, these attempts have been unsuccessful. The restriction enzymes used to perform the iPCR were: *Kpn1*, *XhoI*, *NotI* and *BamHI*; these were expected to cut outside of the P-element, such that a product might be generated using the P-out primer. Since the rate of male recombination is extremely low relative to all other elements tested (see below), it is possible that the particular sequence composition of the region into which this element has inserted is not amenable to iPCR, which relies on the use of appropriate restriction sites flanking the element. These sites must be close enough to generate a product by PCR.

Table 3.1. Results of X-ray screen.

genotype	# of males tested	# (%) sterile	#(%) fertile	# (%) lethal with *Df(3L+R)6B-29*	# (%) located on 3L	# (%) located on 3R
e*/TM6	1870	472 (25.2%)	1398 (74.8%)	17 (0.91%)	14 (0.75%)	3 (0.16%)
e*/TM3	1959	546 (27.9%)	1413 (72.1%)	14 (0.71%)	11 (0.56%)	3 (0.15%)
Total	3829	1018 (26.6%)	2811 (73.4%)	31 (0.81%)	25 (0.65%)	6 (0.16%)

X-ray screen generation of new lesions on 3R

In order to obtain additional lesions in the chromosome 3 heterochromatin, I undertook a mutagenesis screen using X-rays as a mutagen. The advantage of this over P-element mutagenesis is that it is more random. Thus we might expect to be able to identify loci not previously known in the region and extend the coverage of our deficiencies across a greater interval. The screening protocol is described in the Materials and Methods (Figure 2.3). Table 3.1 summarizes the results of the screen.

Once lethal mutations were identified, they were crossed to a set of deficiencies on either 3L or 3R to assign the lesions to one arm or the other. A total of 3829 males were tested; of these, approximately 26% were sterile and the crosses discarded. The results for *TM6* males and *TM3* males were very similar, both in frequency of sterility and the frequency and proportion of lethal lesions on 3L and 3R (Table 3.1). Of the remaining cultures 31, or 0.8% yielded putative lethals with *Df(3L+R)6B-29*. This deficiency was chosen for the screen because we could simultaneously screen both 3L and 3R heterochromatin and because to our knowledge, it deleted more of 3R heterochromatin than any of the other 3R lesions (Koryakov et al. 2002). The intention was to identify lesions that would overlap with *Df(3L+R)6B-29* and perhaps extend further into 3R heterochromatin. Not surprisingly, based on the results of previous screens such as that of Schulze et al. (2001), there was a much lower recovery of lesions in 3R than 3L;

approximately 4-fold more lethals were isolated that mapped to 3L than to 3R. This is significant considering that *Df(3L+R)6B-29* deletes only part of the 3L heterochromatin while apparently deleting nearly all of the 3R heterochromatin. Despite the relative paucity of X-ray induced mutations on 3R compared to 3L, it is nonetheless significant that approximately one fifth of the lesions isolated did map to 3R. This contrasts with the results of Schulze et al. (2001). A γ- screen for which approximately twice as many chromosomes were screened and twice as many lethals isolated, led to the identification of only 5 lesions from the 3R region – this is less than one tenth of the total number of lesions isolated. It is difficult to know how to reconcile the differences in the results of the two screens. The radiation dose used was somewhat higher in the current experiment – 4000 Rads compared to 3000-3700 used by Schulze et al. However, this is probably not a sufficient difference to account for the distinct results. Also, the current experiment utilized X-rays whereas the former used γ--rays. However it isn't clear why this would affect the results. Culture conditions, the chromosome irradiated, and other small differences may all contribute to the differences observed in the two experiments. It is likely that insufficient numbers of chromosomes were screened in either experiment to draw meaningful conclusions about the rates of mutation observed. It is worth noting that one of the lethals isolated in this work – *13B-688* - was viable with only the *TM6* balancer, presumably due to a breakpoint of the multiply inverted *TM3* chromosome coinciding with a lesion on the *13B-688* chromosome. The *13B-688* lesion would not have been isolated if *TM3* was the only balancer used in the

screen. Thus, the slightly higher rate of recovery of 3R lethals in the current study, might be in part attributable to the fact that I utilized both balancers.

The current radiation screen also generated one non-lethal line in which flies exhibited a phenotype including awkward locomotion, outspread legs and melanized patches on the legs, particularly around the joints. This may be an allele of *Splayed* (*Spl*), a locus for which no alleles currently exist. The *Spl* locus was first identified as a dominant phenotype, like that described above, associated with a deficiency of polytene bands 81F5-82A (Lindsley et al. 1972). Later, two temperature sensitive alleles were isolated (Tasaka and Suzuki 1973). However, these have been lost, and thus new alleles of *Spl* would be of potential use, as they would allow a characterization of this gene, and perhaps eventually the identification of the coding region associated with it. The mutant has been tentatively mapped in the distal 3R region, as the phenotype is observed in combination with both the *Df(3L+R)6B-29* and the *Df(3R)XM3* deficiencies. In addition, the line is semi-lethal with both of these deficiencies. The placement of the locus in a region defined by both the above mentioned deficiencies is in agreement with the original localization of the phenotype to 81F/82A by Lindsley et al. (1972) with the exception that the phenotype of this line is observed to be recessive, rather than dominant. It is possible that the similarity of phenotype is coincidental, or that the mutants are different in nature than those originally observed.

Characterization of 3R X-ray induced lesions

The 6 newly induced lethal lesions in 3R were characterized by *inter se* complementation (Table 3.2) and by complementation with other 3R lesions as follows (Table 3.3). The lesions were crossed to all the known/presumed 3R deficiencies: *Df(3L+R)6B-29, Df(3R)XM3, Df(3R)10-65, Df(3R)A321R1* and *Df(3R)4-75;* as well as *Df(3R)EP-167R,* a lesion made from the P-element line, *EP(3)3632,* by male recombination (described below). In addition, the lesions were crossed to *PGFP*, the aforementioned P-element insertion lethal, the *L-1/8* alleles and *CH(3)1*, the homozygous lethal P-element insertion at h55 (Tulin et al. 2002).

As mentioned, one of the lesions produced, *13B-688*, was lethal in combination with *TM3Sb* and *TM3Ser* balancers and was therefore maintained by balancing it with *TM6* (as alluded to above). This aberration likely has a breakpoint in common with the multiply inverted *TM3* balancer chromosome. The breakpoints of *In(3L+R)TM3* are known and reported in flybase (www.flybase.bio.indiana.edu) and none of those listed resides in 3R heterochromatin, implying that *13B-688* is a complex rearrangement with at least one breakpoint outside of the 3R heterochromatin. Interestingly, *13B-688* is semi-lethal in combination with some of the other lesions tested (see also later discussion of MR-induced lesions). Though semi-lethality is still a failure to complement, and might indicate allelism, these results imply that *13B-688* might not delete the coding region of the 3R

heterochromatic gene it is affecting. The complementation pattern indicates that *13B-688* could be allelic with *l(3R)81Fa*, the gene which is deleted by *Df(3R)10-65* (see introduction of this chapter). *13B-688* could be an inversion with at least one breakpoint elsewhere on the chromosome and another which affects *l(3R)81Fa*. The breakpoint need not be altering the coding sequence of the locus – it is possible that the inversion has juxtaposed sequence nearby which could alter its expression sufficiently to produce lethal and semi-lethal phenotypes in combinations between *13B-688* and the various alleles of *l(3R)81Fa*. Figure 3.4 shows a speculative model depicting this possibility. *13B-688* has not yet been examined cytologically, so this idea awaits further evidence.

The five 3R heterochromatin mutations isolated by the screen reported by Schulze et al. (2001) were all alleles of a single locus, *l(3R)81Fa*. All six lesions isolated in this screen showed lethality in combination with *Df(3R)10-65*, implying that they too are allelic with *l(3R)81Fa*. However, two of the new lesions,

Df(3R)7B-90 and *Df(3R)9B-919*, appear to be deletions which affect multiple loci. For example *Df(3R)7B-90*, failed to complement every known lesion in 3R with the exception of *CH(3)1* (which, remarkably, complements every other known lesion in 3R) and *13B-688*, with which it is semi-lethal. *Df(3R)9B-919* appears to be less extensive, showing lethality in combination with *Df(3R)EP-167* and *Df(3R)10-65*, but viability with two of the other new X-ray induced lesions. This implies that *Df(3R)9B-919* may not extend proximally as far as does *Df(3R)10-65*, with which all of the X-ray induced lesions are lethal. Two other lesions may be associated with deletions. *5A-388* is lethal in combination with both *Df(3R)10-*

Table 3.2. *Inter se* complementation among 6 new X-ray induced lesions in 3R heterochromatin.

For simplicity, the prefixes: "Df(3R)" or "I(3R)" have been omitted.

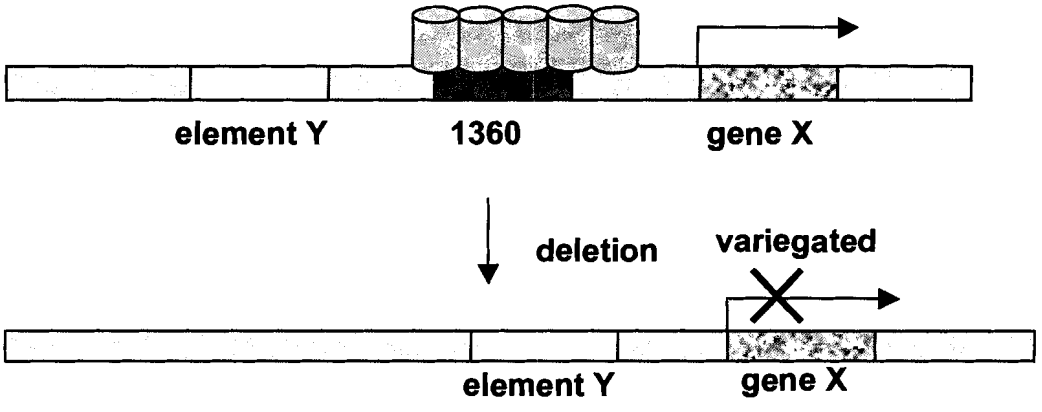
Line	7B-90	3A-204	5A-388	13B-688	9B-919	18A-1676
7B-90	L	L	L	SL	L	L
3A-204		L	V	V	V	L
5A-388			L	SL	L	L
13B-688				L	SL	V
9B-919					L	V
18A-1676						L

Table 3.3. Complementation test results between putative deficiency lines in 3R and all pre-existing lesions in 3R.

Line	6B-29	XM3	A321	10-65	EP-167	4-75	zep	PGFP	CH(3)1
7B-90	L	L	L	L	L	L	L	L	V
3A-204	L	L	L	L	V	V	V	V	V
5A-388	L	L	L	SL	V	SL	V	V	V
13B-688	L	L	L	L	V	L	V	V	V
9B-919	L	L	L	L	L	L	L	L	V
18A-1676	L	L	L	L	V	V	V	V	V

Figure 3.4. Schematic depiction of an hypothetical basis for lesions to fail to complement a locus without affecting its coding sequence.

The image below shows a theoretical heterochromatic gene X. This gene is dependent upon the presence of HP1 (cylinders) for correct expression. The 1360 element, shown near gene X, recruits HP1 and/or other elements of heterochromatin, allowing normal expression of gene X. A deletion could remove sequences near gene X, including the 1360 element (Sun et al. 2004). A different repetitive element, Y, is now juxtaposed to gene X but because it does not actively recruit the heterochromatin "machinery", expression of gene X is variably compromised. The survival of an individual bearing this chromosome in combination with a deficiency of the locus would depend on the amount of expression of the locus. If expression is greater than some required threshold, the individual will survive, and if below that threshold the individual would die. In this way, a deficiency could display semi-lethality with some mutations without actually altering the coding sequence.



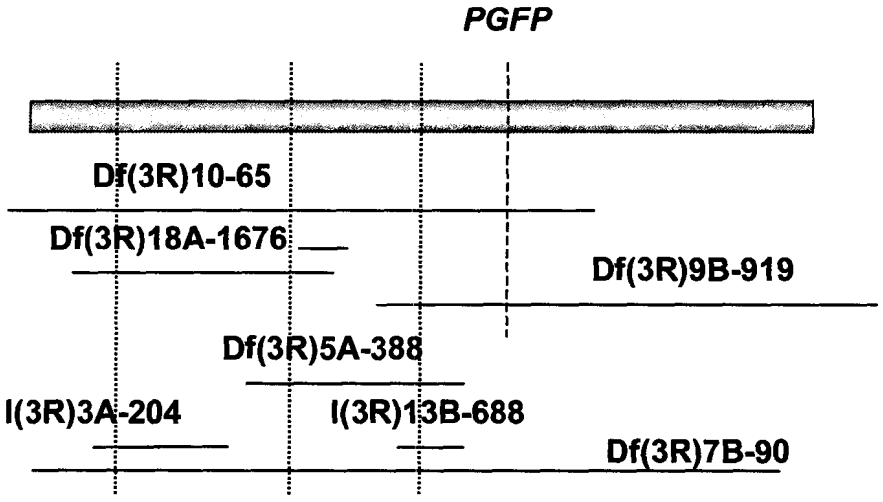
65 and *Df(3R)9B-919* as well as with *18A-1676*. On the other hand *18A-1676,* while lethal in combination with *Df(3R)10-65*, is viable in combination with *Df(3R)9B-919*. Moreover,*18A-1676* is lethal in combination with *3A-204*, which is viable with *5A-388* and *Df(3R)9B-919* but lethal with *Df(3R)10-65.*

Figure 3.5 shows a depiction of this arrangement; *l(3R)13B-688* is included as a single gene lesion. However, as discussed above, it is likely a complex rearrangement. Thus, as many as three new lethal complementation groups might be defined by these lesions as shown in Figure 3.5: (i) one defined by the lethal overlap between *18A-1676* and *3A-204* but not encompassed by *5A-388*; (ii) a second defined by a lethal overlap between *18A-1676* and *5A-388*, encompassed by *Df(3R)10-65* but not by *Df(3R)9B-919*; and (iii) a third defined by lethality with *Df(3R)9B-919*, *5A-388* and *Df (3R)10-65*, but not *PGFP*, with which *5A-388* and *18A-1676* are both viable. If the above interpretation is correct, then only *3A-204* affects a single locus; the other two would also be multi-locus deletions and would thus be named *Df(3R)18A-1676* and *Df(3R)5A-388*. *3A-204* would be named: *l(3R)3A-204.*

Viability of some of the smaller X-ray induced lesions with *P-GFP* might be instances of inter-allelic complementation specifically in the cases of the smaller lesions that are lethal with *Df(3R)10-65* but not *P-GFP.* Inter-allelic complementation occurs when different lethal alleles of a single locus complement, leading to potential overestimation of the number of loci residing in

Figure 3.5. Complementation pattern among the 3R lesions that fail to complement *Df(3R)10-65.*

Shown is h58, solid bar, and the *10-65* deficiency, as indicated. The heavy dashed vertical line indicates the position of the essential locus defined by the PGFP insertion line. Finer dashed lines indicate 3 new complementation groups that could in theory be defined by these lesions.



a genetic interval. The likelihood and frequency of inter-allelic complementation is probably related to the type of product encoded by a gene. A large and complex protein with multiple functions might be more likely to lead to a situation where mutations in different functional regions allowed sufficient wild type function for the survival of the trans-heterozygote. Inter-allelic complementation might also be expected when a gene product acts as a dimer. It is therefore possible that none of the aforementioned putative loci are actually new complementation groups. In fact, Marchant and Holm (1988b) isolated many EMS alleles lethal in combination with *Df(3R)10-65*, but assigned all of them to a single complex locus by complementation analysis. Although one must be cautious about concluding that the current study has defined novel loci within this interval, it is possible that different screening protocols, mutagens and approaches have yielded hitherto undetected loci.

Df(3R)4-75 is a complex rearrangement, with a putative breakpoint at the site of the *P-GFP* mutation. In principle, complementation analysis involving all four single site lesions and *Df(3R) 4-75* should clarify whether any of the lesions define new complementation groups. However, there is no consistent relationship between these lesions and *Df(3R)4-75*; *13B-688/Df(3R)4-75* is completely lethal while *5A-388/Df(3R)4-75* exhibits semi-lethality. On the other hand, both *3A-204* and *18A-1676* are fully viable when transheterozygous with *Df(3R) 4-75.* Although it is possible that these lesions define separate loci, I cannot rule out the possibility of inter-allelic complementation. An examination of

the lethal phases and phenotypes of the lesions might be informative in this regard.

Homozygous *Df(3R)10-65* flies die in the first larval instar. The doomed larvae remain small and move sluggishly, persisting for several days before dying. When *3A-204*, *5A-388* and *18A-1676* and *Df(3R)9B-919* flies are crossed to *Df(3R)10-65/+* flies, the resulting transheterozygous mutant progeny show the same phenotype for each of the lesions; small, sluggish first instar larvae that die after several days. While this is not conclusive evidence, it certainly strongly suggests that these apparently separate complementation groups are in fact alleles of a single locus, which display inter-allelic complementation. Of the 6 lesions generated, only *Df(3R)7B-90* and *Df(3R)9B-919* are lethal in combination with *P-GFP*. The lethality in these cases is during the pharate adult stage as is seen with *P-GFP* in combination with the original deficiencies. Thus, *P-GFP* might define a separate, and later acting locus. In this case the other lesions would respectively be alleles of *l(3R)81Fa* (based on a similar lethal phase to that reported by Marchant and Holm 1988b), and *P-GFP* a novel 3R heterochromatic locus. Alternatively, *P-GFP* might be another allele of *l(3R)81Fa*, with an altered lethal phase and phenotype. The recalcitrance of *P-GFP* to molecular analysis presents a challenge in distinguishing between the two hypotheses. The most conservative interpretation would make them all alleles of a single locus.

Each of the new X-ray induced mutations was crossed to an EMS allele of the *abstrakt* (*abs*) locus. This locus resides just outside the 3R heterochromatin and thus may serve as an "anchor point" on the genetic map. Of all lesions tested, only *Df(3R)XM3*, *Df(3R)A321* and *Df(3R)7B-90* failed to complement *abs*. A composite complementation and deficiency map that combines all of the information obtained from the original lesions and the newly generated X-ray mutants, is shown in Figure 3.6. The figure portrays the four "single site" lesions as alleles of the locus defined by *P-GFP* because the bulk of the evidence points to this interpretation.

Male Recombination Experiments

The aim of this work was to create a set of overlapping deficiencies, most of which still retained the original P element so that molecular information regarding the breakpoints could be generated. The male recombination experiments were conducted as outlined in the methods section. Four different starting P-elements were used; these were selected because their insertion sites spanned as much of the 3R heterochromatin as possible (Figure 3.7).

Male recombination frequencies vary greatly among different types of P elements and different locations.

Table 3.4 shows the results of the male recombination experiments using P elements in 3R as starting points. As both the location and the type of element used were different for each experiment, differences in frequencies of recombination and production of lesions may be ascribed to one or both of these factors. It seems reasonable that different sequence compositions surrounding a P-element would influence its mobility. Variations in chromatin structure might compromise transposition and insertion attempts. The possible significance of the original position of the P-element is shown in the work of Sun et al. (2004; described in the Introduction) in which small-scale male recombination experiments (among other methods) were used to generate deletions or duplications next to P- elements. Two screens were performed, with two different elements. The first tested 836 lines and found 1 line with altered flanking sequence, a rate of .12%, whereas the second screen examined 542 lines and isolated 9 lines with altered flanking sequences – a rate over 10-fold higher (1.6%). The two elements used to initiate these screens differed only in their original positions on the small fourth chromosome.

Figure 3.6. Probable genetic map of all new X-ray induced mutations relative to pre-existing lesions.

Figure is not to scale. Heavier lines indicate original lesions, thinner lines, the X-ray lesions produced in this study. The mutants, *3A-204*, *5A-388*, *13B-688*, and *18A-1676*, are shown as probable single site mutations allelic with the *P-GFP* mutant (see text).

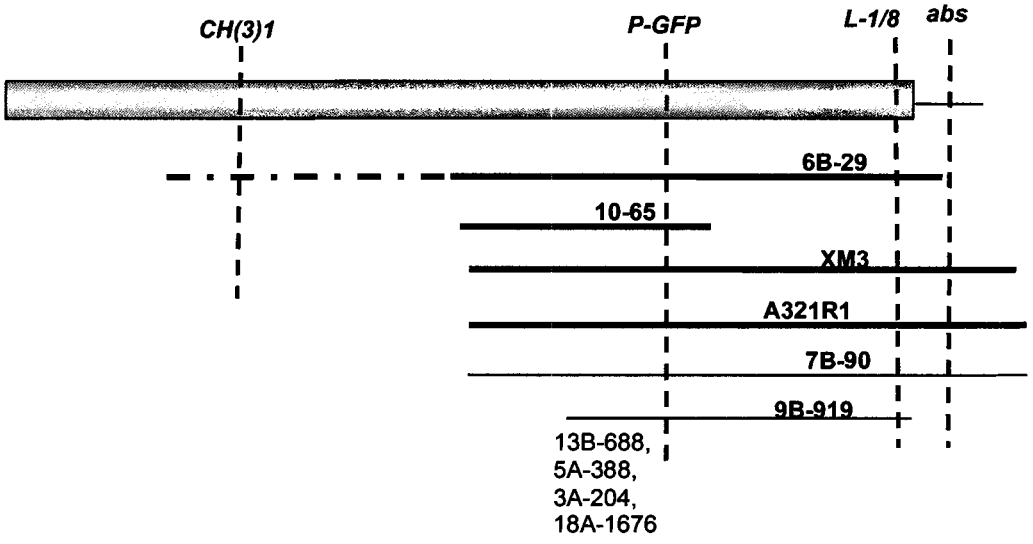
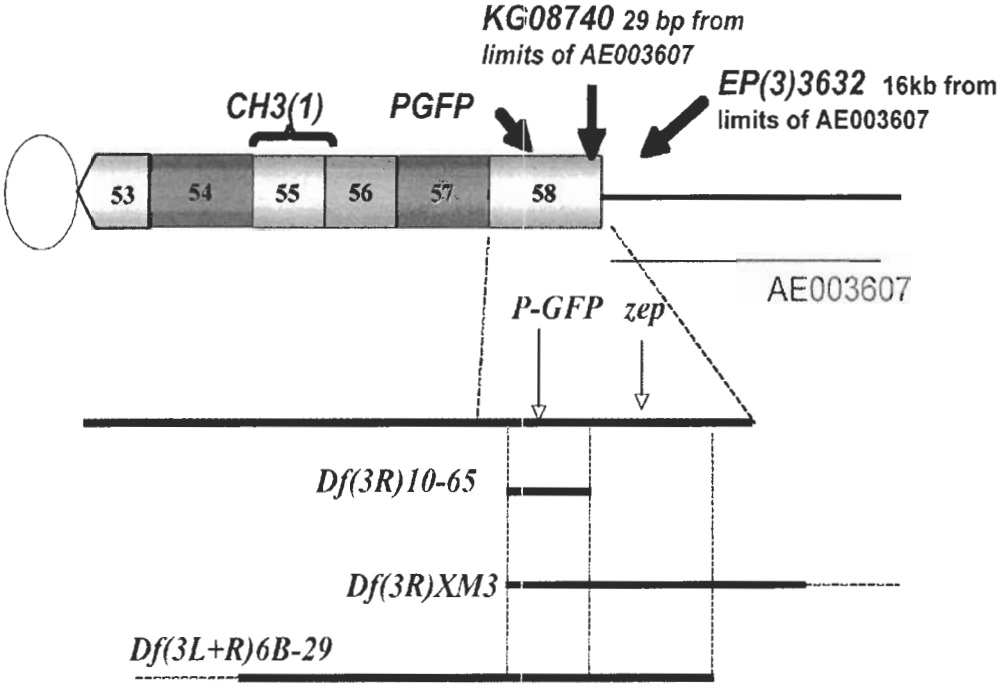


Figure 3.7. Original positions of P elements used for 3R male recombination experiments.

Positions depicted are not exact, and are based on published results in the case of *CH(3)1*, which has been placed in h55 (Zhang and Spradling 1994); database information in the case of *KG08740* and *EP(3)3632*; 29 base pairs and 15.6 kilobases from the edge of the annotated 3R sequence (AE003607) respectively, and an estimate in the case of *PGFP*, which is lethal in combination with *Df(3R)10-65,* which in turn has been shown by Koryakov et al. (2002) to lack most of the h58 interval. The thick rectangle represents 3R heterochromatin segments h53 to h58; the circle, the centromere; and the arrows (or bracket in the case of *CH(3)1*) the approximate insertion points of the P-elements used.



The highest recombination frequency was obtained with the EP element (column 2, Table 3.4); note that only ri p^{p+} recombinants were selected in this experiment, so the frequency obtained represents theoretically only half of the total recombination frequency. Thus, this is the only male recombination experiment that yielded frequencies comparable to that of the original study by Preston et al. (1996). From over 10,000 males scored in this experiment and 45 lines generated and tested, however, only one – *Df(3R)EP-167* – turned out to be a useful lesion. As the EP element used was viable in combination with *Df(3R)XM3*, a lesion which we knew deleted loci both distal and proximal to the insertion site, lethality with *Df(3R)XM3* deficiency was used to determine which lesions were good candidates for deletions of essential loci. The frequency of deficiencies induced in all of these experiments was much lower than that predicted by Preston et al. (1996). There are likely a number of reasons for this, but the principal one may be the relative paucity of essential loci in heterochromatin, and the large distances between genes. By definition, recombination-generated deletions in heterochromatin must be rather large in order to affect one or more essential genes and thus be detectable. However this work confirms that the male recombination method does work in centric heterochromatin and indicates that the method has utility for generating useful lesions.

Table 3.4. Results of male recombination experiments with 4 different P-element strains in 3R heterochromatin.

For the *EP(3)3632* experiment, only half of the recombinants (*ri* p^{p+}) were selected. "Unique events" (row 5) refers to the fact that recombinant events occur prior to meiosis and thus may be clustered. Any set of similar recombinants that arose from the same set of parents were considered to have arisen from a single recombinant event. Note that this could therefore be an underestimate of the actual recombinant frequency.

Experiment	EP3632 3R (EU/HET)	KG08740 3R (EU/HET)	PGFP 3R (HET)	CH(3)1 3R (HET)
Type of P-Element	pEP	SuPorP	P[GFP]	pPZ
Males scored	10,269	8,666	4,165	3,535
Recombinants (%)	45 (.44%)	42 (.48)	1 (.02)	8 (.23)
Unique events (%)	24 (.24%)	21 (.24)	1 (.02)	6 (.17)
Lethal lesions	1 (.01%)	6 (.06%)	0	1 (.03%)

All recombinant lines from the *EP(3)3632* experiment were tested for the presence of the original P-element by PCR of genomic DNA templates with P-out and sequence specific primers, complementary to sequences flanking the insertion point (Table 3.5). The lines could be categorized as those from which the P-element had cleanly excised, those for which the element was still intact and those for which one or the other end bore some sort of lesion. In the first case, homozygous genomic DNA could be collected from those lines which were homozygous viable and PCR attempted with only the flanking primers, which should yield a wild type band if the P element has cleanly excised.

The single putative deficiency line generated from this experiment was tested with all other lesions in 3R and determined to be lethal in combination with *L-1/8*, *Df(3R)XM3*, *Df(3R)A321*, *Df(3R)4-75* and *Df(3R)6B-29*, but not *Df(3R)10-65* or *PGFP*. It has been designated *Df(3R)EP-167*. The rate of deficiency production is much lower than that reported by Preston et al. (1996), as mentioned. This may reflect the nature of heterochromatic sequence, either in the relative paucity of nearby sites into which to transpose, or the fact that the genes in heterochromatin are less densely arranged such that deletions must extend very far in order to remove an essential gene.

Table 3.5. Characterization of the 24 recombinant events from *EP(3)3632*.

The lines were characterized using P-out primer and primers specific for proximal and distal flanking genomic sequence. The fourth category describes two events for which proximal products could not be produced, but for which distal products could. These were not lethal in combination with the EP-167 deficiency, nor with Df(3R)XM3. It is postulated that small rearrangements may have taken place in either flanking sequence or the P-element itself.

Category	P-out + Proximal primer	P-out + Distal primer	Proximal + Distal primer	Number (%)
Proximal deletion, P-element retained	-	+	ND	1 (4.2%)
No rearrangement, P-element retained	+	+	-	6 (25%)
P-element excised	-	-	+	15 (62.5%)
non-lethal rearrangement or defective P-element	-	+	-	2 (8.3%)

Overall, the frequency of recombination from the *KG08740* insertion site was quite high, but only about half that predicted by Preston et al. (1996), and the number of apparent deletions is somewhat higher than for *EP(3)3632.* Again, whether this is due to the particular type of element, or its surrounding environment - or both – is difficult to assess in the case of this element, for which flanking sequence is known only on the distal side of the element. However, it is reasonable to suggest that since the sequence on the proximal side has been so recalcitrant to analysis, it is likely to be fairly repetitive. We would not necessarily expect this to lead to either greater recombination frequency or a greater induction of deletions. It is interesting to note that even though the *KG08740* MR experiment gave a lower overall rate of male recombination (recall that both recombinant classes were collected in this experiment), there was a higher frequency of mutants generated. The *KG08740* lesions in all cases are lethal in combination with the *EP-167* lesion. This might reflect the fact that the *KG08740* P-element was closer to an essential locus and that even small deletions would lead to lethality. However, the lack of available sequence data proximal to the element would rather imply that some highly repetitive sequence was located just proximal to *KG08740.* Definitive evidence for this awaits more molecular and sequence data.

Just as with element *EP(3)3632*, the original *KG08740* P-element was viable with *Df(3)XM3*, so the best way to identify potential deficiencies was to test for

lethality with that deficiency. Interestingly, all 9 recombinants lethal with *Df(3R)XM3* turned out to be ri p^{p+} even though in this experiment – and all subsequent experiments – I selected and characterized both recombinant classes generated. *Df(3R)XM3* certainly does delete vital loci distal to the *KG08740* insertion site, as it is lethal in combination with an *abs* allele, for instance. It seems quite possible that the loci immediately distal to *KG08740* might not be essential, although this awaits further analysis.

The lethal lines were tested with PCR using P-out primer and a primer designed against exon 1 of gene *CG12581*, which is located just distal to the original site of insertion. The PCR experiment tested for the presence of the P element in its original location. 9 recombinant lines were obtained; of these, #33-36 were all generated from the same set of parents, but #33 and #36 have slightly different complementation patterns. #34 and 35 have identical complementation patterns and thus may not represent separate events. #44 and #46 were also generated by the same set of parents, but having different complementation patterns, are likely to be unique events. The main differences between all of the lesions is in their ability to complement *l(3R)13B-688*, and/or *l(3R)5A-388*. While it is true that these display an apparent inter-allelic complementation, it would be expected that this complementation pattern would be consistent – i.e. the same in combination with a particular deficiency. Thus if #44 and #46 were two instances of the same event, then it would be expected that *l(3R)13B-688*, and/or *l(3R)5A-388* would behave similarly in combination with both lines. This is the main basis

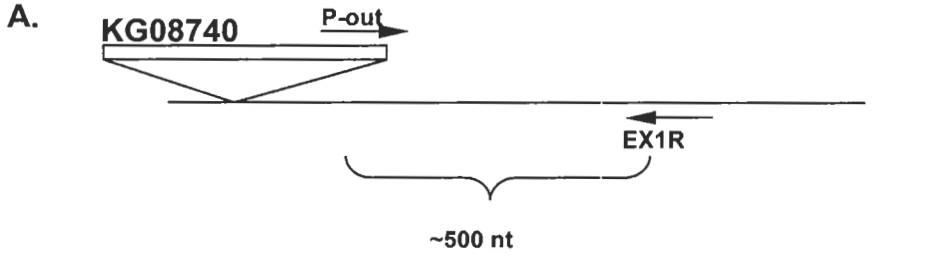
for delineating #33, #34/35, and #36, as separate lesions, as well as #44 and #46. This awaits further molecular analysis for confirmation.

Figure 3.8A shows a schematic depiction of the insertion point of the original P-element in the *KG08740* line at position 29 on the genomic scaffold AE003607. Also shown is the approximate position of the primer used in this experiment, EX1R, which is complementary to sequence about 500 bp from the site of insertion (the first exon of a gene called CG12581). All of the lines tested (so far) show a band of the expected size with no background bands. An *Oregon R* or *Canton S* wild type genomic DNA control was routinely included in all tests. Figure 3.8B shows the PCR results for the lines that were tested.

Recombinants from the *CH(3)1* insert were obtained at a lower frequency than the two aforementioned, which may indicate that sites deeper in heterochromatin are less able to undergo transposition, and thus less likely to generate recombinants. 8 recombinants were obtained, only 6 of which were deemed to be independent events. All but one of the recombinants were homozygous viable and thus not useful lesions. Complementation tests were performed with *CH(3)1-1* (the name given to the homozygous lethal recombinant line) and all other 3R lesions (Tables 3.6 and 3.7). The recombinant is lethal in combination with all lesions in 3R heterochromatin proximal to *Df(3R)EP-167,* with the exception of *P-GFP*. This is not the pattern expected if the *CH(3)1* P- element

Figure 3.8. PCR results for KG08740 lines.

A. A schematic representation of the DNA near the insertion site of *KG08740* and the primers used for the PCR. **B.** PCRs used genomic template from *KG08740* recombinant lines balanced over TM3Sb. (lanes 1-6; lines 18, 20, 22,33,35,and 36, respectively), as well as the homozygous parent line KG08740/KG08740, lane 7, Canton S as a non-P containing control, lane 8 , and a no template control, lane 9. The recombinant and original lines all generate a band of approximately 500bp, indicating that the distal part of the P-element and the distal flanking sequence are both intact.



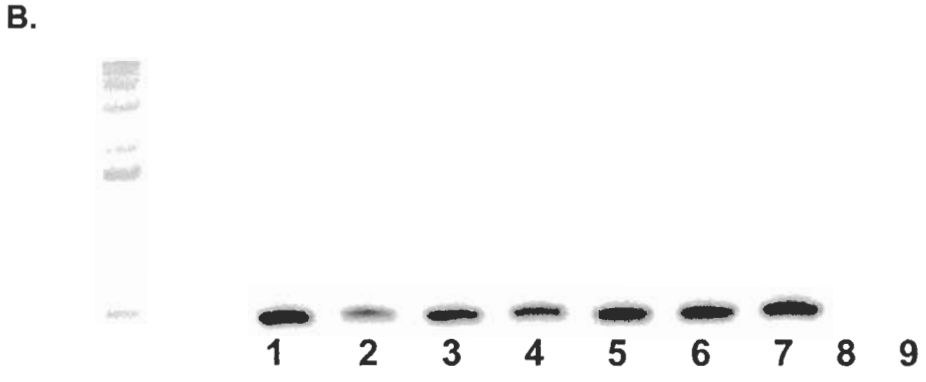


Table 3.6. Male recombination lines crossed to all original 3R lesions. L= lethal (<1% of expected transheterozygous progeny) V=viable (> 25% of expected transheterozygous progeny).

Expt.	Line #	6B-29	XM3	A321	10-65	4-75	L-1/8	PGFP	CH(3)1
CH(3)1	1	L	L	L	L	L	V	V	V
EP(3)3632	167	L	L	L	V	L	L	V	V
KG08740	18	L	L	L	L	L	L	V	V
	20	L	L	L	V	L	L	V	V
	22	L	L	L	V	L	L	V	V
	33	L	L	L	L	L	L	L	V
	34	L	L	L	L	L	L	L	V
	35	L	L	L	L	L	L	L	V
	36	L	L	L	L	L	L	L	V
	44	L	L	L	L	L	L	L	V
	46	L	L	L	L	L	L	L	V

Table 3.7. Complementation results for crosses of all MR-induced lines against the new *X-ray*-induced lesions.
A number of lines showed semi-lethality (SL) in combination with either the *5A-388* or the *13B-688* mutations.

Expt.	Line #	7B-90	3A-204	5A-388	13B-688	9B-919	18A-1676
CH(3)1	1	L	L	L	L	L	L
EP(3)3632	167	L	V	V	V	L	V
KG08740	18	L	L	L	SL	L	L
	20	L	V	V	V	L	V
	22	L	V	V	SL	L	V
	33	L	L	L	L	L	L
	34	L	L	SL	SL	L	L
	35	L	L	SL	SL	L	L
	36	L	L	L	SL	L	L
	44	L	L	V	L	L	L
	46	L	L	SL	V	L	L

was originally located at h55. The flanking marker exchange on the chromosome would predict that a lesion had begun at the site of *CH(3)1* and deleted proximally from that point. The results suggest one of two possibilities: that the *CH(3)1* element is not located where it was thought to reside but is rather located in h58, distal to *Df(3R)10-65*, but proximal to *Df(3R)EP-167*; or that the male recombination event was accompanied by a local hop and a subsequent additional event (recombination or imprecise excision) that created the more distal lesion. This will be discussed more fully in a later section.

Of all the P element strains on 3R used for these experiments, the lowest recombination frequency was obtained with the *PGFP* mutant. About 500 more males were examined than in the *CH(3)1* experiment with 8-fold fewer recombinants and 6-fold fewer independent recombinant events. The extremely low frequency of male recombination may also reflect the that the *PGFP* element is inserted into a somewhat hostile environment, from which transposition may be hampered either by very compact chromatin structure, and/or a lack of appropriate nearby sequences into which the element might transpose. The single recombinant obtained showed viability when transheterozygous with *Df(3R)XM3*, and so was not a useful line.

All of the recombinant lines that were lethal in combination with *Df(3R)XM3* were tested for complementation with all other known 3R lesions. Tables 3.6 and 3.7 summarize the results of these tests. Note that a number of these lesions show

semi-lethality with *13B-688*. However, several of these: lines *8740-34*, *35* and *46* show semi-lethality with the *5A-388* deficiency as well. In addition, line *8740-44* is fully viable in combination with *l(3R)5A-388*, while line *8740-46* is fully viable with *l(3R)13B-688*. This cannot be explained simply as different extents of the two deficiencies, since the larger of the two should be lethal with both lesions. It indicates that the locus or loci deleted by *Df(3R)10-65* must show a very unusual pattern of inter-allelic complementation, and that it is likely that there are complexities associated with the lesions generated. Further explanations await a thorough cytological analysis of these deficiencies.

Inverse PCR was attempted with several lesions from these experiments; *EP-167*, *8740-18, 20, 22*, and *CH(3)1-1* (see Chapter 2 for description of the method). From these, only *CH(3)1-1* yielded a product, which was cloned, sequenced and found to be composed entirely of sequence homologous to the P-element. This indicates the fortuitous presence of a HinPI I site immediately outside of the element such that no flanking sequence could be obtained.

X-ray and male recombination induced mutagenesis gave comparable results.

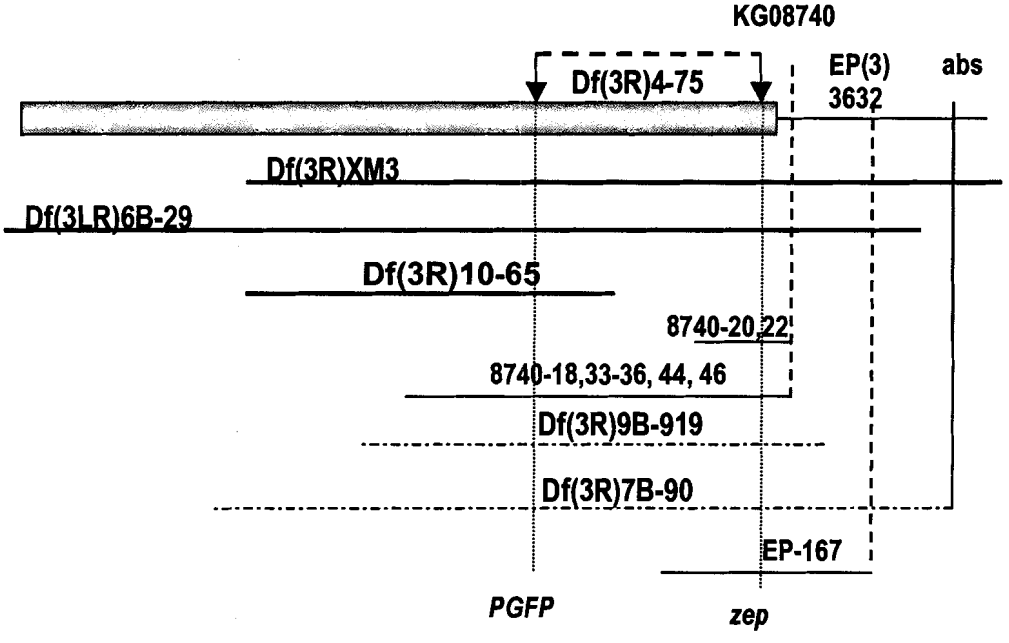
It would appear that the two methods do yield somewhat different lesions. Obviously there are differences in site specificity because the male recombination method utilizes pre-existing P-element insertion sites, nevertheless, both methods have yielded lesions that are confined mainly to the

h58 region. In the case of the X-ray lesions, there should be no preference for this interval unless this is the only region deleted by *Df(3LR)6B-29* that contains essential genes. The unusual complementation patterns among these lesions test the assumption that simple polar deficiencies are being created by these methods. Certain lesions display complex complementation patterns that are not easily explained by polar deficiencies. The instances of semi-lethality could indicate that at least some of these lesions are associated with inversion breakpoints which may be causing the repression of nearby heterochromatic loci such that some individuals make enough gene product for full viability, while most do not. It is clear for instance, that *l(3R)13B-688* affects more than one site, since it shows lethality with theTM3 balancer chromosome.

Figure 3.9 shows a genetic map including all but one of the new lesions. In some cases, lesions that may be distinct, such as *8740-18, 33, 34, 35, 36, 44,* and *46* are depicted together, since the differences between them are based on their complementation behaviour with a set of X-ray lesions that most likely define one locus. Some of these lesions may well extend more proximally than is shown on the figure, but there are currently no genetic mutations that would allow this to be determined. The *CH(3)1-1* recombinant is not shown on this map, as its complementation pattern cannot be reconciled with the apparent localization of the *CH(3)1* P-element to the much more proximal h55 region (see below).

Figure 3.9. Complementation map of MR-induced lesions relative to original and X-ray induced lesions.

Here *Df(3R)4-75* is shown as an inversion above the solid bar (which represents h58 only), with arrows to mark the lethal breakpoints. Original sites of P-elements are represented by dashed vertical lines. The solid vertical line represents the *abs* locus as a reference point. Original lesions are shown as thicker horizontal lines. MR-derived lesions are represented by the thinner lines. X-ray induced lesions are shown as dashed horizontal lines. Two lethal complementation groups are represented by the vertical dotted lines. Where several recombinant lines have the same complementation pattern, they are represented by a single line labeled with all lesions that displayed that pattern. This map is based upon the assumption that the remaining four X-ray induced lesions, and the P-GFP mutant, are all alleles of a single locus.



The zeppelin locus

The *zeppelin* (*zep*) locus has been localized to 3R heterochromatin. *zeppelin* is one of several loci originally identified by Ostrowski et al. (2002) in an EMS screen for genes involved in chitin synthesis. Homozygous mutant embryos display a characteristic "blimp" phenotype, in which devitellinized embryos swell to approximately 3 times their normal size. The chitin produced in these mutants may be more elastic than normal. Although the embryos can move, sometimes displaying a strong hyperactivity, they are unable to hatch and eventually die. The *zeppelin* allele was crossed to *Df(3R)4-75* and the two lesions failed to complement (Ostrowski et al. 2002). I obtained the allele, zep^{LP13}, and mapped the mutant against all the lesions I had produced or collected and determined that the locus fell under the *Df(3R)EP-167* deficiency. A pair of EMS alleles from the Leptin lab, *L-1* and *L-8,* also failed to complement this deficiency, and failed to complement the *zeppelin* allele. Thus there are now three known EMS alleles of the locus: zep^{LP13}, zep^{L1} and zep^{L8}. A large number of deficiencies fail to complement *zep*: *Df(3R)XM3*, *Df(3R)A321*, *Df(3R)7B-90* and *Df(3R)9B-919*, as well as *Df(3R)4-75* and *Df(3R)EP-167* as noted above. In addition, every *ri* p^{p+} recombinant from the KG08740 male recombination experiment which displayed lethality with *Df(3R)XM3* (and thus had likely produced a deficiency) also failed to complement *zep*, indicating that the locus may be relatively close to the edge of the known sequence in the region. One pair of recombinants, *8740#20* and

8740#22 appear to delete only the *zep* locus as they show no lethality in combination with any of the more proximal lesions (*Df(3R)10-65*, for example, or the *PGFP* mutant). As these two lesions *(#20* and *#22*) arose in different bottles, from different parents, they represent independent lesions, presumably with different breakpoints. All of the other lesions identified in the *KG08740* MR screen – *8740 #18, 33, 34, 35, 36, 44*, and *46* - extend further toward the centromere, as they are also lethal in combination with other proximal lesions. *zep* is likely to be the original *l(3R)81Fb*, based on its similar complementation pattern.

Single embryo PCR experiments.

All lesions that were thought to be possible deletions were balanced with GFP-containing balancers so that homozygous deficiency embryos could be identified (see material and methods). A subset of the putative deficiencies (*Df(3R)XM3*, *Df(3R)7B-90* and *Df(3R)EP-167*) showed lethality with the *TM3-GFP* balancer which was mutant for *Sb*, and so the *TM3-pAct-GFP* balancer, marked with *Ser* was used instead. Once homozygous deficiency embryos were obtained, PCR could be performed to determine which sequences, if any, are missing from the deficiency. This can provide useful indications of which loci are likely to be responsible for mutant phenotypes of the point (and other) mutants identified in the heterochromatin. The difficulty is that very little DNA can be obtained from a single embryo and only one or two PCRs can be performed with it. An alternate approach involved the use of a genome amplification kit (Genomi-phi; see

materials and methods), which was designed for potential forensic use on very tiny samples (for example small amounts of blood). This gave the potential for amplifying the whole genome of homozygous deficiency embryos and then screening for the presence or absence of many genes from a single embryo. Control experiments indicated that at least in the case of heterochromatic genes, the quality of DNA amplified made a large difference to the quality of the resulting DNA. Single embryos do not yield much DNA; moreover homozygous deficiency embryos may also be dead or dying when collected, so that DNA quality is not always ideal. Homozygous embryos were selected and their DNA amplified. PCRs were performed on amplified DNA templates of several embryos of each genotype (where possible). In some cases, the results were clear-cut, but in others, there were difficulties with background bands, and variability in the results. Only tests that have been repeated several times and give reproducible results are reported here. Sample positive controls are shown in Appendix C.

3R heterochromatic genes tested by PCR

Glutamine:fructose-6-phosphate aminotransferase1

The *Glutamine:fructose-6-phosphate aminotransferase1* (*Gfat1*) locus encodes an aminotransferase which catalyzes the rate-limiting step of the UDP-N-acetylglucosamine synthesis pathway. In insects, the Gfat protein is involved in the synthesis of chitin, which is a homopolymer of N-acetylglucosamine. The production of chitin is regulated to coincide with a period just prior to larval

moults, which are ecdysone-regulated events. Graack et al. (2001) first identified the *Gfat* locus in Drosophila, by homology cloning using the mouse homologue. They identified two alternately polyadenylated forms of the cDNA, one which was 2406 nt in length and the other 2413. *In situ* hybridization to polytene chromosomes placed the locus at position 81F on chromosome 3R. No mutations have ever been identified in this locus. A recent release of the genome project, Release 3.2, places the *Gfat1* locus in 3R, in heterochromatic segment h54-55. *Gfat1* resides on genomic scaffold AABU01002542. The gene has as many as 10 exons, depending upon the splicing, and 5 different splice variants are annotated on the scaffold. The locus spans only about 7.5 kb –relatively small, for a heterochromatic gene - with the largest intron being 2,245 nt long. A depiction of the organization of the locus is shown in Figure 3.10, and described in detail in Table 3.8, which describes the genomic organization of each of the loci described in this section. Table 3.9 summarizes the PCR results for all the genes tested.

Primers were designed against a region of exon 1, and are listed appendix B, along with all other primers used for this work. It was of interest to determine the deficiencies from which *Gfat1* might be absent, not only because it is a heterochromatic locus and therefore of interest, but also because of its potential role in chitin synthesis, which makes it an excellent candidate for the *zeppelin* locus. However, its placement at h54 or h55 by the genome project, if accurate, would make it quite unlikely that *Gfat1* corresponds to *zep*, since most of the data

Figure 3.10. Organization of the *Gfat1* locus on the genomic scaffold AABU01002542.
Also shown is the arrangement of a second gene, CG40198. The scaffold is 48.5 kb in total. The diagram is not to scale.

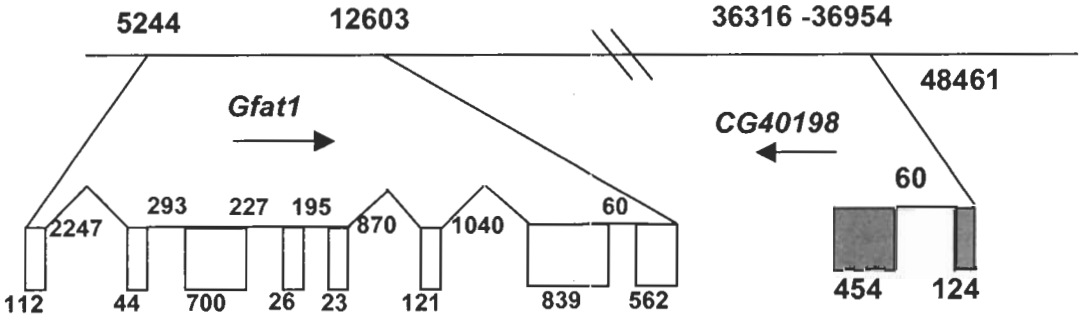
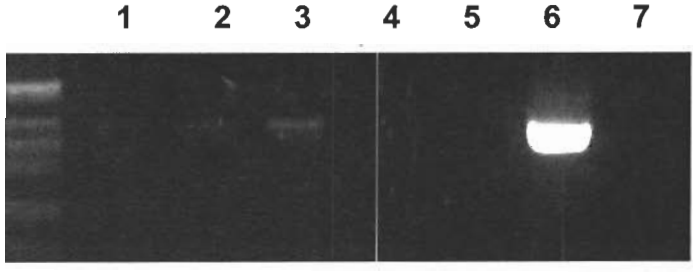


Figure 3.11. Preliminary PCR results indicating a possible link between the *zep* locus and the *Gfat* gene.
A. Lanes 1,2 *Df(3LR)6B-29*; Lane 3, *Df(3R)10-65*; Lanes 4, 5, *Df(3R)EP-167*; Lane 6, *Canton S* control; Lane 7, no template control. **B**. Lane 1, Df(3R)10-65; Lanes 2, 3, Df(3LR)6B-29; Lane 4, Oregon R; Lane 5, no template control.

A.



B.

Table 3.8. Organization of 3R genes studied on genomic scaffolds. The scaffold number for each gene is given in the text.

Gfat1	CG40197				
exon#	Start	End	Size	Intron #	Size
1	5244	5356	112		
2	7603	7647	44	1	2247
3	7940	8640	700	2	293
4	8867	8893	26	3	227
5	9088	9111	23	4	195
6	9981	10102	121	5	870
7	11142	11981	839	6	1040
8	12041	12603	562	7	60
Parp	CG40411				
exon #	Start	End	Size	intron #	Size
1	300082	299650	432		
2	114139	113959	180	1	185511
3	112603	112462	141	2	1356
4	112406	112073	333	3	56
5	82845	82312	533	4	29228
6	28680	28111	569	5	53632
7	21895	21777	118	6	6216
8	20854	20156	698	7	923
CG40451					
Exon #	Start	End	Size	Intron #	Size
1	244769	245013	244		
2	245116	245293	177	1	103
3	254683	254882	199	2	9390
CG40159					
Exon #	Start	End	Size	Intron #	Size
1	123808	123508	300		
2	123391	122997	394	1	117
CG12581					
exon #	Start	End	Size	intron #	Size
1	380	509	129		
2	578	1913	1335	1	69
3	7784	8649	865	2	5871
4	9439	10200	761	3	790
Dsk	CG18090				
Exon #	Start	End	Size	Intron #	Size
1	15388	16170	782	n/a	

indicate that the locus lies closer to the heterochromatin border (see previous discussion of the *zep* locus). Figure 3.11 shows the preliminary PCR results, which appear to shown a PCR product for the *Gfat1* gene in *Df(3R)10-65*, but none of the other lesions tested, all of which show lethality in combination with *zep* alleles. This strengthens the possibility that the *zeppelin* locus encodes *Gfat1*, and that it resides more distally than predicted by the genome project.

CG12581

The *CG12581* locus is the most proximal on the euchromatic scaffold of 3R (AE003607). It encodes a putative protein kinase gene and is expressed during numerous stages of development, based upon the cDNA libraries (ovary, adult head, larval/pupal and 0-22 hour embryonic) in which it is represented. It is not known if the locus is essential, as no mutations have been identified in *CG12581*, and there are no deficiencies whose overlap would delete only this locus.

The gene is small for a gene in/near heterochromatin, spanning only 6.9 kb, with 2 or 3 introns and 2 alternatively spliced transcripts. A figure of the organization of *CG12581* is shown in Figure 3.12. Repeatable results have not yet been obtained for the *CG12581*; although PCR data do indicate that no product is amplified from *Df(3R)EP-167* homozygous embryos, results with some other

genotypes have yielded background bands and unclear results. Therefore the data are not presented here and further work needs to be done to complete the characterization of this distal region.

Drosulfakinin

The *drosulfakinin* (*Dsk*) locus (also depicted in Figure 3.12, and Table 3.8) resides on genomic scaffold AE003607, very near CG12581, and encodes a small protein with neuropeptide activity (Nichols et al. 1988). Since the locus was first identified and mapped to the 81F region (Nichols et al. 1988), no mutants have been isolated. Several lesions in 3R, such as *Df(3R)XM3*, and *Df(3R)EP-167*, as well as others, are expected to delete this gene; therefore a screen utilizing one of these lesions could allow the isolation of many mutations for genes at this boundary region, leading to a fuller characterization of the region.

Preliminary PCR results, shown in Figure 3.13 indicate that of the lesions tested, only *Df(3R)10-65* yields a product for *Dsk*.

CG40159

CG40159 is found on genomic scaffold AABU01002691. This genomic scaffold is probably located in 3R heterochromatin, based on the work of Konev et al.

Figure 3.12. Map of Genomic scaffold AE003607.
The figure indicates the known loci proximal to the *EP(3)3632* insertion point and shows the position of the *KG08740* element as well. *CG12581* has two alternative transcripts, one of which does not utilize the first intron shown.

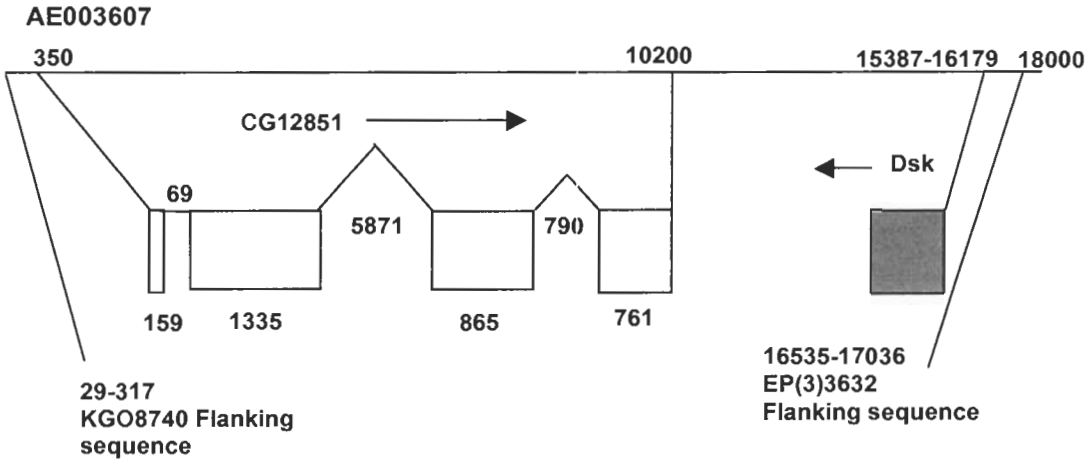
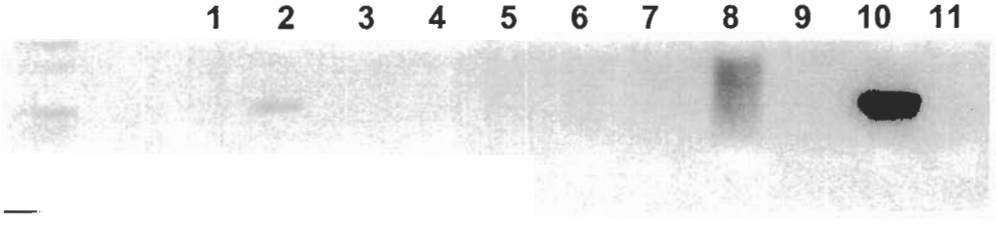


Figure 3.13. Preliminary PCR results for the *Dsk* gene and selected homozygous deficiency embryos.
The lanes indicated are as follows. Lanes 1 and 3, *Df(3LR)6B-29*; Lane 2, *Df(3R)10-65* (loaded out of sequence); Lanes 4 and 5, *Df(3R)7B-90*; which was loaded out of sequence; Lanes 6 and 7, *Df(3R)EP-167*; Lanes 8 and 9, *Df(3R)XM3* (note that lane 8 corresponds to lane 6 in the control PCRs shown in Appendix C (part A) and therefore should have been removed from the image); Lane 10, *Oregon R* control; Lane 11, no template control.



(2003) who identified one P-element insertion near this locus and mapped it cytologically to segment h56. *CG40159* is predicted to encode a protein of unknown function, but does produce transcripts and associated protein (Drysdale et al. 2005). Figure 3.14 shows the genomic organization of the *CG40159* locus on genomic scaffold AABU01002691, with the locations of the other annotated loci indicated as well.

The PCR results depicted in Figure 3.15, suggest that *CG40159* is absent from *Df(3LR) 6B-29,* but none of the others tested: *Df(3R)10-65*, *Df(3R)7B-90*, *Df(3R)XM3*, and *Df(3R)EP-167*. This confirms that *Df(3LR)6B-29* extends further proximally than the other lesions, at least to the h56 segment. This means that the *6B-29* deficiency is still the most extensive available in 3R heterochromatin. However, the male recombination lines derived from the *KG08740* strain and the *CH(3)1* insertion, have not yet been characterized by single embryo PCR. There is a possibility that one of these might extend further proximally as well. It would be very useful to obtain another deficiency that extended as far as *Df(3LR)6B-29*, but which has not been produced by compound autosome detachment.

The Poly (ADP-ribose) polymerase locus

The *Poly (ADP-ribose) polymerase (Parp)* gene has been localized to 3R heterochromatin by at least two different labs. The gene was first cloned and

Figure 3.14. Genomic organization of heterochromatic scaffold AABU01002691.

Gene *CG40159* is shown on the scaffold, which is 151,240 nucleotides long.

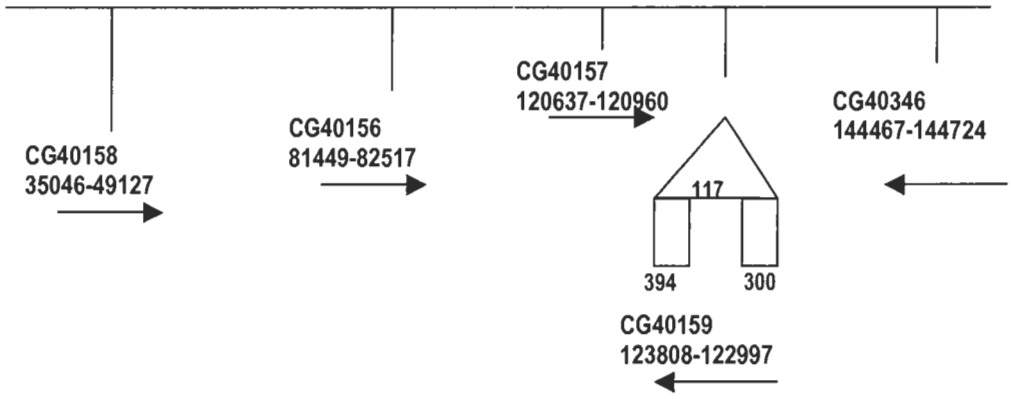
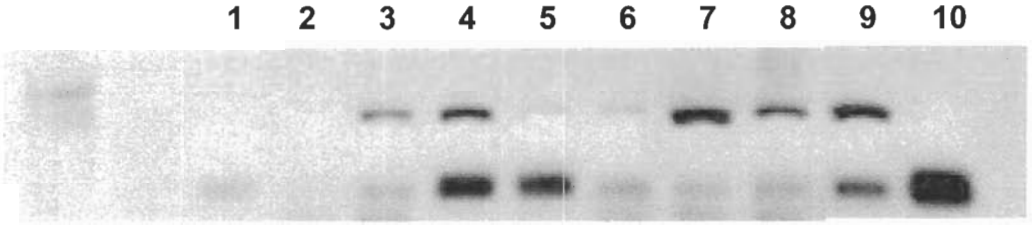


Figure 3.15. Preliminary PCR data for CG40159 and selected homozygous deficiency embryos.

Lanes 1 and 2; *Df(3LR)6B-29*; Lane 3 *Df(3R)10-65*; Lanes 4 and 5, *Df(3R)7B-90*; Lanes 6 and 7, *Df(3R)EP-167*; Lane 8, *Df(3R)XM3*; Lane 9, *Oregon R*, positive control; Lane 10, no template control.



characterized by Uchida et al. (1993) using primers to conserved domains from other species. Later the locus was mapped to region 81F, suggesting that it was a good candidate for one of the two essential complementation groups already identified in 3R heterochromatin (Hanai et al. 1998). Parp is involved in numerous functions, binding to DNA breaks (Hanai et al. 1998) and performing roles in tissue polarity (Uchida et al. 2001), chromatin structure and gene regulation (Tulin et al. 2002; Tulin and Spradling 2003). In 1994, a series of PZ insertions into heterochromatin was isolated, and a lethal insert, *CH(3)1*, was localized to the h55 interval (Zhang and Spradling, 1994). Tulin et al. (2002) later mapped the PZ element to a region near an alternative promoter site in the *Parp* locus. Excising the P element restored viability of *CH(3)1.* A second lethal mutation of *Parp*, *CH(3)4* was not fully characterized from a molecular perspective although it was known that there was no P-element inserted in 3R heterochromatin in the mutant. The *Parp* gene produces three differentially spliced transcripts: ParpI, ParpII and Parp-e. Parp-e is a transcript that initiates at an alternate first exon, 185 kb from the rest of the locus, and encodes a catalytically inactive form of Parp. The *CH(3)1* insertion has been shown to reduce the amount of the ParpI transcript, and mutant larvae arrest at the onset of ecdysis during the second instar, and subsequently die. Although ectopic expression of the ParpI variant transcript allows larvae to survive up to the third instar, no adults develop. However, the Parp-e transcript variant, when ectopically expressed rescues the mutation more fully, allowing some fertile

adults to eclose. A stock has been produced containing both the *UAS Parp-e* transgene and a ubiquitous driver, that allows the *CH(3)1* strain to be maintained in a homozygous state (Tulin et al. 2002). As mentioned above, the Parp-e variant leads to the formation of a catalytically inactive isoform of the Parp protein, and Tulin et al. (2002) have proposed a model whereby the inactive form of the protein is potentially required to regulate expression of the active form, most likely by altering the chromatin structure upstream of the *Parp* locus. Figure 3.16 shows a schematic map of the organization of the Parp locus and the other loci on the genomic scaffold AABU01002763

Initially, I tested the *CH(3)1* mutant for its ability to complement the deficiencies available and it showed complete viability and fertility with all of the following: *Df(3L+R)6B-29*, *Df(3R)10-65*, *Df(3R)4-75*, *Df(3R)A321*, and *Df(3R)XM3*. Thus, it did not seem to be a candidate for either of the two loci already identified in 3R heterochromatin: *l(3R)81Fa* or *l(3R)81Fb* (see figure 1.3B). Moreover, as other mutations and male recombinants were produced, complementation tests were performed and each lesion tested fully complemented *CH(3)1*. Thus, when performing single embryo PCR, primers to the *Parp* coding sequence were often used as controls to test for the presence of good quality DNA or for use In identifying homozygous deficiency embryos. However, early on during these experiments, it was evident that *Df(3L+R)6B-29*, and *Df(3R)XM3* homozygous deficiency embryos could not be identified using these primers. Many sets of embryos were tested and all gave either 2 bands (control and GFP) or no bands.

Since the *Parp* primers had been used successfully to identify a number of other homozygous deficiency embryos, it seemed unlikely that the primers were at fault. The homozygous deficiency embryos were thus identified using a different set of control primers, designed against a gene called *grp84* on the X chromosome. Several sets of exon-specific primers for *Parp* have been used and the results are consistent and repeatable: although the *CH(3)1* mutation is viable with these two deficiencies, they nonetheless, appear to lack at least exons 2-6 of the coding sequence. PCR analysis shows that the *Parp* gene is absent from: *Df(3L+R)6B29*, *Df(3R)7B90* (not tested with all primers yet), *Df(3R)XM3*, and present in *Df(3R)EP-167*, and *Df(3R)10-65* (Figure 3.17). The absence of *Parp* bands from homozygous *Df(3L+R)6B-29*, *Df(3R)XM3* and *Df(3R)7B-90* is unexpected, given that the *CH(3)1* and *CH(3)4* mutations are fully viable when transheterozygous with these deficiencies. PCR results alone might be insufficient to conclude that the gene is absent from these deficiencies; however the failure to generate *Parp* product in hundreds of putative homozygous embryos from several attempts, strengthens the assertion.

CG40451

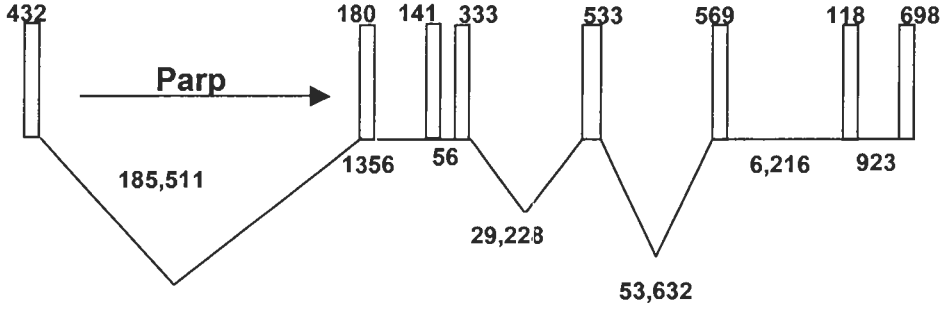
The gene *CG40451* resides on the same genomic scaffold as *Parp* (AABU01002763) nested within the first intron of *Parp* but oriented in the opposite direction. The gene organization is illustrated in Figure 3.16; 2 other

genes are nested in the same intron. *CG40451* encodes a putative mitochondrial inner membrane transport protein, with predicted functions as a carrier in intercellular protein transport. The gene produces transcripts and associated proteins (Drysdale et al. 2005). Since *CG40451* resides within the very large first intron of the *Parp* locus, it can be used to help determine whether the deficiencies examined still retain some sequence upstream of the exons tested. If those deficiencies that show viability with *CH(3)1* and apparently delete most of the coding sequence, still retained some upstream regulatory sequence, then the unusual complementation behaviour could perhaps be explained by a transvection- like mechanism, whereby the regulatory sequences on one homologue can initiate some transcription on the other homologue. This phenomenon requires the pairing of homologous chromosomes during interphase, which does occur in *Drosophila*, unlike many other organisms. This could potentially explain the data, although transvection does not usually restore full function, but rather leads to partial complementation (Duncan 2002).

Preliminary results indicate that *CG40451* is missing from all the same deficiencies as was *Parp*. In addition, the locus appears to be absent from *Df(3R)10-65* as well (Figure 3.18). This will need further confirmation, since DNA from only one *Df(3R)10-65* homozygous embryo could be tested, but if correct, this allows the placement of the genomic scaffold containing *Parp* in the h58 interval, much more distally than had been thought. In addition, it would make the possibility of a transvection-like phenomenon, suggested above, unlikely.

Figure 3.16. Schematic representation of *Parp* and nearby loci.
A. Genomic organization of the *Parp* locus on scaffold AABU01002763. The entire scaffold length is 336,328 bp and *Parp* extends from nucleotide 300,082 - 20156, a gene length of just under 280 kb. Thus, *Parp* extends nearly across the entire scaffold. **B**. The 3 loci nested within intron 1. For a more detailed description of the coordinates of *CG40451* within the genomic scaffold, see Table 3.8

A.



B.

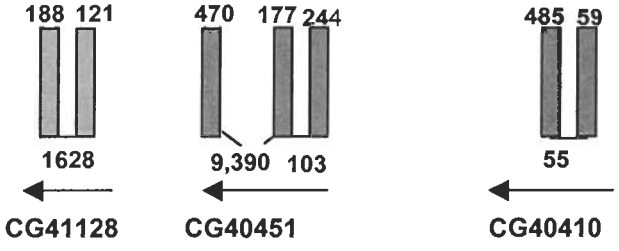
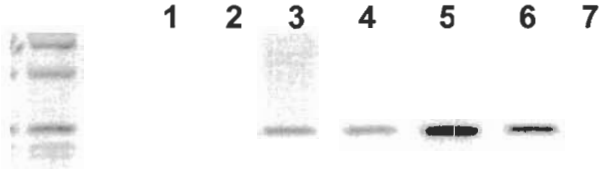
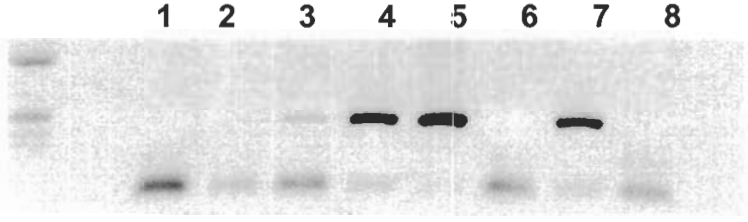


Figure 3.17. Preliminary PCR results for *Parp* and selected homozygous deficiency embryos.

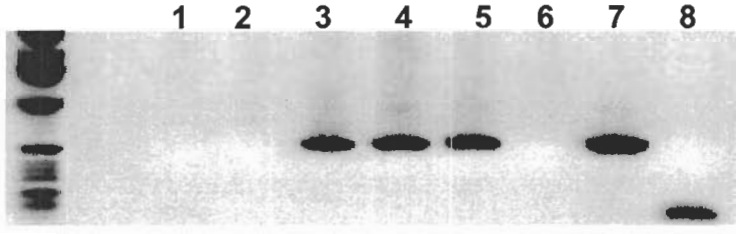
A. Parp 1F and 1R primers. Lanes 1 and 2, *Df(3LR)6B-29*; Lane 3, *Df(3R)10-65*; Lanes 4 and 5, *Df(3R)EP-167*; Lane 6, *Oregon R*; Lane 7, no template control



B. Parp 2F and 3R primers. Lanes 1-5 as above; Lane 6, *Df(3R)XM3*; Lane 7 *Oregon R*; Lane 8, no template control.



C. Parp 4F and 4R primers. Lanes 1-8, as above.



D. Parp 5F and 5R primers. Lanes 1-4, *Df(3R)6B-29*; Lanes 5 and 6, *Df(3R)7B-90*; Lanes 7 and 8, *Df(3R)EP-167*; Lane 9, *Df(3R)XM3*; Lane 10, *Oregon R*; Lane 11, no template control.

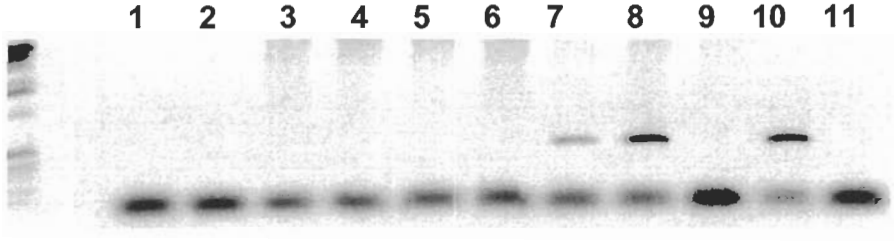


Figure 3.18. Preliminary PCR data for *CG40451* and selected homozygous deficiency embryos.

Lanes 1 and 2, *Df(3LR)6B-29;* Lane 3, *Df(3R)10-65*; Lanes 4 and 5, *Df(3R)EP-167*; Lane 6, *Df(3R)XM3*; Lane 7, *Oregon R*; Lane 8, no template control.

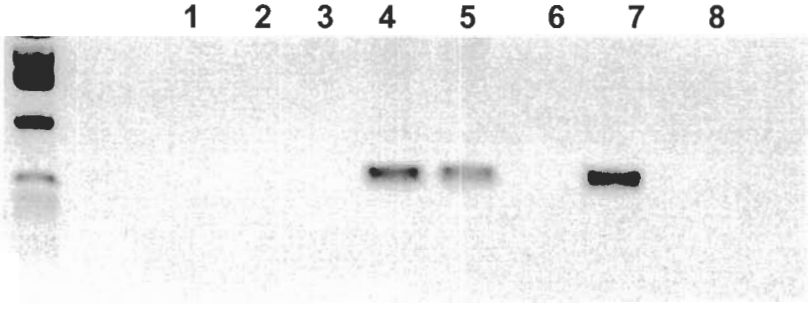


Table 3.9. Summary of PCR results to date: products present or absent for genes tested by single embryo PCR with deficiency homozygote embryos of various 3R deficiencies.
The Df(3R) prefixes have been left off the deficiency names. ND – not done; NC not complete.

Lesion	Parp	CG40451	CG40159	Dsk	Gfat
6B-29	-	-	-	-	-
XM3	-	-	+	-	-
10-65	+	-	+	+	+
EP-167	+	+	+	-	-
7B-90	NC	ND	+	-	-

However, these results may help to explain the complementation pattern of the *CH(3)1-1* male recombinant line, which showed lethality in combination with *Df(3R)10-65* and all of the other lesions that fail to complement *Df(3R)10-65*. These complementation data can be explained if *Parp* is in fact located at h58, between *Df(3R)10-65* and *Df(3R)EP-167*. This is supported by the fact that deficiencies generated in *ri* p^{p+} male recombinants would be expected to delete proximally (Preston et al. 1996).

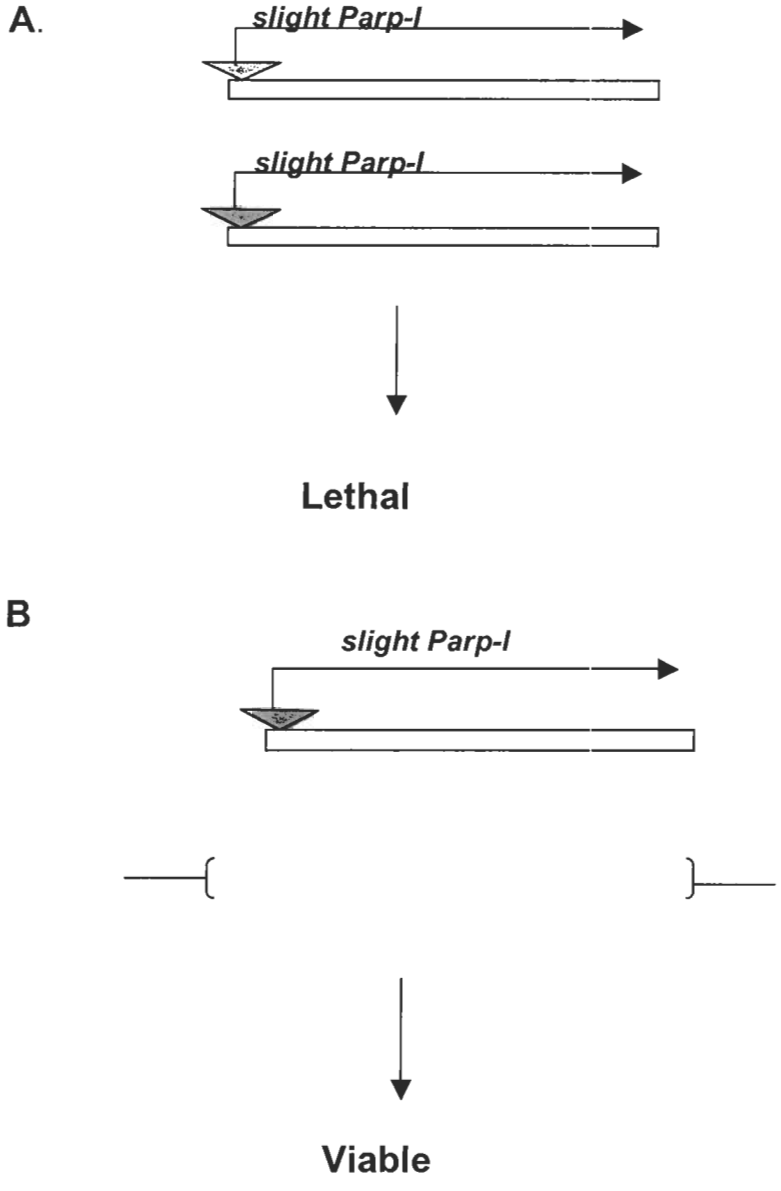
Revising the location of the Parp locus, and the unusual nature of Parp

Some recent work by the Karpen lab has identified a series of P-alleles in heterochromatin and three of these have inserted near the gene model *CG40451* (Konev et al. 2003). *CG40451* has been localized to the same genomic scaffold as the *Parp* locus by the heterochromatin gene project. All three of the P element inserts near *CG40451* have been localized by mitotic *in situ* hybridization and all three apparently map to heterochromatin segment h58 (Konev et al. 2003). This is in conflict with the previous results of Zhang and Spradling. It is not likely that the scaffold extends 3 segments across the heterochromatin, firstly because it is too small to extend so far (only about 336 kb in length) and secondly because the h57 segment is known to contain a very large block of satellite sequence that would preclude the generation of a cohesive genomic scaffold across it. Single embryo PCR with 3R lesions showed that no band was amplified for the

CG40451 locus from the *Df(3L+R)6B29*, *Df(3R)XM3*, and *Df(3R)10-65* deficiencies. Interestingly, *Df(3R)10-65* has been shown to lack most of the h58 interval of 3R heterochromatin (Koryakov et al. 2002). If correct, these results support the possibility that *Parp* is located more distally than originally thought. However, if correct, they also point to an anomalous characteristic of the locus – or of the two known mutations, as has been mentioned previously. How could a P element insertion be homozygous lethal but viable in combination with a deficiency? One would expect this only if the lethality of the P element was in fact from a second site elsewhere on the chromosome. Some mutations in heterochromatin are undoubtedly associated with second site lethals, as evidenced by the fact that the *CH(3)4* mutation on 3L apparently has a second site mutation on 3R that is allelic with *CH(3)1* (Tulin et al. 2002). However, Tulin et al. (2002) have shown by rescue experiments that the lethality associated with *CH(3)1* is indeed due to perturbation of the *Parp* locus. The *Parp-l* transcript can partially rescue lethality, while the *Parp-e* transcript can fully rescue the mutant. They have also shown that a low level of the *Parpl* variant transcript was produced in a variegated manner in some tissues of the *CH(3)1* mutant. The unusual complementation behavior of the Parp mutants is shown in Figure 3.19 in schematic form.

To explain this behavior, I hypothesized that the presence of deficiencies in the heterochromatin may somehow increase pools of needed heterochromatin proteins that might allow a "leaky" *Parp* mutant to survive. Normal

Figure 3.19. Schematic representation of the enigmatic complementation displayed by the *CH(3)1* allele of *Parp*.
The triangle depicts the insertion point of the P-element and the filled rectangle represents the *Parp* gene. A low level of Parp-I transcript is reported by Tulin et al. (2002). The homozygous mutant situation is shown in **A**; the transheterozygous *CH(3)1/Df(3L+R)6B-29* situation in **B**. In the lower panel, the deficiency is represented by the brackets.



heterochromatic genes require certain heterochromatic proteins for proper expression, and their expression is reduced in the presence of mutants that produce reduced levels of such proteins (Clegg et al. 1998; Lu et al. 2000; Schulze et al. 2005), Perhaps large heterochromatic deletions free up additional amounts of limiting heterochromatic proteins, such that a mutant phenotype is rendered less severe. This proposal is consistent with the observation of Tulin et al. (2002) that the mutant *CH(3)1/CH(3)4* larvae do produce small amounts of the *Parp-I* transcript.

To test this theory, *CH(3)1/Df(3L+R)6B-29* and *CH(3)1/Df(3R)XM3* transheterozygotes were placed in an *X^X/Y* background to determine if the addition of extra heterochromatin from the attached X might counteract the effect of the deletions. Several other deficiencies were used as well; *Df(3R)10-65*, which does not delete the *Parp* locus, and should thus serve as a negative control; *CH(3)1-1*, the recombinant line derived from *CH(3)1*; *Df(3R) 9B-919*; and *Df(3R)8740-33*, one of the newly induced male recombination lesions. If *CH(3)1/6B-29* survives because the large *6B29* deficiency removes a substantial amount of heterochromatin thus making more heterochromatin protein(s) available, then reducing the dose of the heterochromatin proteins might be expected to counteract this effect. Previous experiments (Clegg et al. 1998; Sinclair et al. 2000) have shown that reduced HP1 dosage could enhance the severity of phenotype of heterochromatic mutants; this model proposes that an

excess of heterochromatic proteins could have the opposite effect. Figure 3.20 shows a schematic depiction of the model.

Figure 3.21 shows the crosses made for the experiment described above. Table 3.10 shows the results of the crosses. The numbers are low but there are clearly as many surviving female transheterozygotes as male in all cases. Thus there is no evidence that extra heterochromatin in the background has any affect on survival of the *Parp* allele, *CH(3)1*, in combination with any of the deficiencies tested. The failure of these experiments to demonstrate an effect of chromatin structure on viability of *CH(3)1* with the deficiencies make it less likely that the hypothesis is correct, but it is still possible that something more specific is taking place, perhaps with one particular heterochromatic protein that is associated with the third chromosome, the dose of which was not sufficiently increased by this method. Additionally, there could be something unusual about *Parp,* which is itself involved in chromatin structure modification and exhibits a form of self-regulation as previously described. It is perhaps worth noting that the lethal phase of the X-ray induced lesions *l(3L)3A-204*, *l(3L)18A-1676*, and *l(3L)5A-388* in combination with *Df(3R)10-65* was during the first larval instar, the same lethal phase observed for RNAi-induced knockdown of *Parp* (Tulin et al. 2002). It is tempting to speculate that these may all be alleles of the *Parp* locus. This interesting possibility awaits further analysis, but if correct, would provide an opportunity to study *Parp* regulation and its many functions in greater detail.

Figure 3.20. Proposed model to explain the unusual complementation behaviour of *Parp* mutants.
The triangle represents the *CH(3)1* insertion; the filled rectangle, the *Parp* gene; and the textured circles, the excess heterochromatin protein available as a result of the large deletion, *Df(3LR)6B-29*. The model proposes that the deficiency makes excess HP1 protein available which can increase expression of the needed *Parp-e* transcript from the *Parp* locus. In this way, the combination of the mutant and a deficiency for the locus could be viable.

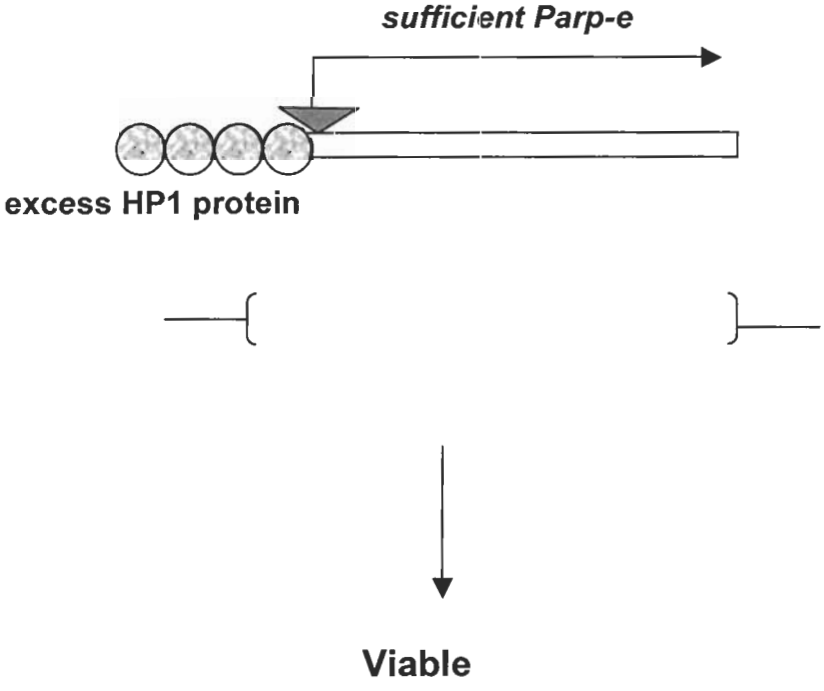


Figure 3.21. Crossing scheme for combining *Parp* allele, *CH(3)1* and various deficiencies with an attached X (*X^X*).

Here the deficiencies used are referred to as *Df(3R)*. Those tested were: *Df(3R)6B-29 ri p^p^*, *Df(3R)XM3 e ri p^p^*, *Df(3R)10-65 ri p^p^*, *Df(3R)9B-919 e*, *Df(3R)CH(3)1-1 ri* and male recombinant *8740#33 ri*.

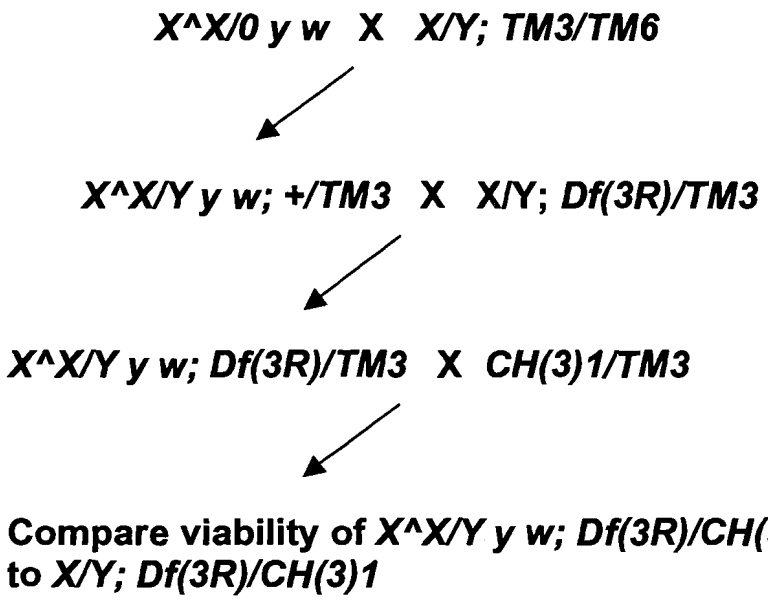


Table 3.10. Results of the X^X crosses outlined in Figure 3.21.
Number of progeny of each genotype are shown. CH(3)1/TM3 males were crossed to X^X/Y females of the genotypes indicated in the first column.

X^X/Y; Def/TM3:	X^X/Y;			X/Y;		
	Df/TM3	CH(3)1/TM3	Df/CH(3)1	Df/TM3	CH(3)1/TM3	Df/CH(3)1
6B-29 ri p^p^	13	12	15	15	11	15
XM3 e	14	16	13	16	14	19
10-65 ri p^p^	10	13	15	18	11	13
CH(3)1-1 ri	13	15	15	11	11	14
9B-919 e	10	14	17	19	12	18
8740-33 ri	10	13	9	8	7	12

EMS screen for new mutations

An EMS screen for new mutations in chromosome three heterochromatin is currently in progress in the lab. It involves screening the previously mutagenized "Zuker lines" (Koundakjiana et al. 2004), for mutations that are lethal in combination with one of four deficiencies, selected to cover as much of chromosome 3 heterochromatin as possible (see Chapter 2 for a description of the screening protocol). Currently, the screen of approximately 6000 mutant lines is between one third and one half completed. Preliminary results show that the screen is yielding lesions at the rate of about one 3R lesion for every 5 in 3L, the same rate observed in the X-ray screen. Several mutations have been identified which are lethal in combination with *Df(3L+R)6B-29* but not the most distal, *Df(3R)XM3*, nor any of the 3L deficiencies. These are good candidates for genes that may be located in more proximal regions of the 3R heterochromatin. These are being mapped by recombination to ensure that they reside in heterochromatin, and are not second site lethals.

Interestingly, one mutation has been identified which shows a *Spl* – like phenotype, as was observed in one of the X-ray lesions, described previously. The generation of putative new alleles of a gene for which there are no longer any mutant alleles, means that there is a new opportunity to identify and characterize this interesting locus.

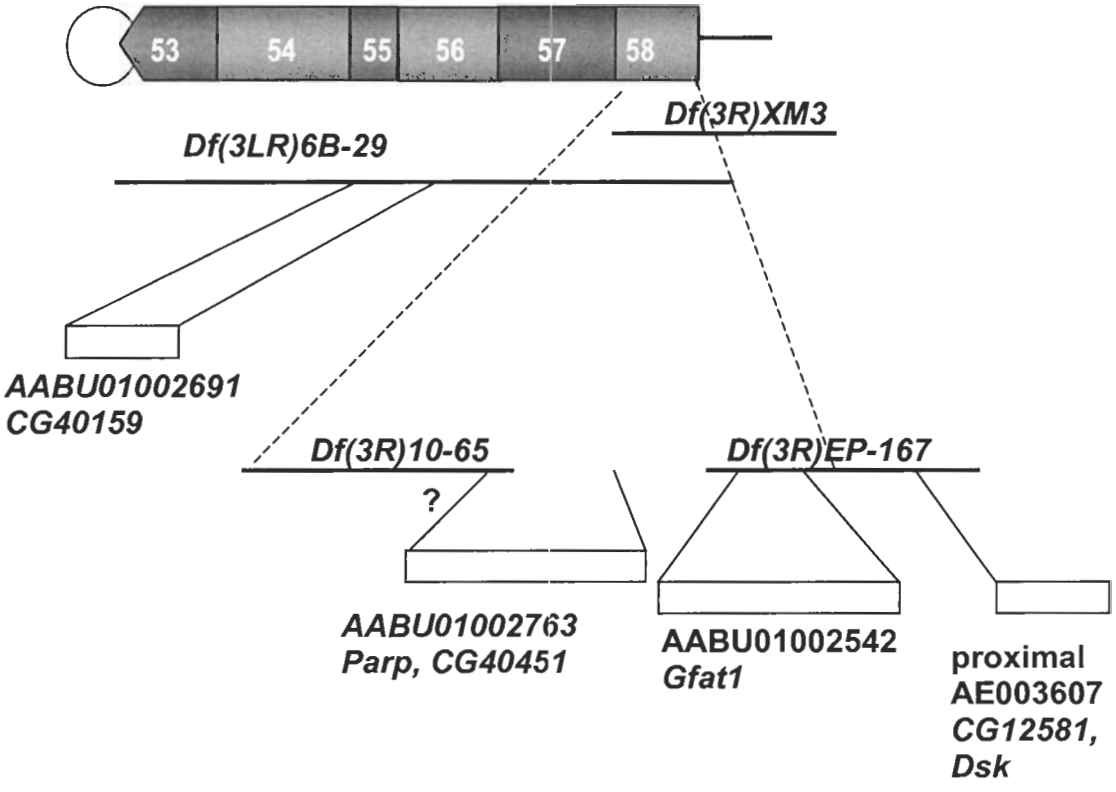
Conclusions and Future directions

The aim of this project was to generate new lesions in 3R heterochromatin and improve the genetic and molecular characterization of the region. The use of both X-rays and P-element induced male recombination has yielded a number of potentially useful lesions. Figure 3.22 shows a simple summary schematic depicting my proposed locations for 4 genes in 3R heterochromatin. There appears to be three genetic elements residing in 3R heterochromatin: *l(3R)81Fa*, with several new alleles identified which may help to identify the gene (*l(3R)3A-204*, *l(3R)5A-388*, *l(3R)13B-688* and *l(3R)18A-1676*); *Parp*, which may reside distal to *l(3R)81Fa*, a finding supported both by my own results and recent data generated in other labs (Konev et al. 2003), and the *zeppelin* locus which may encode *Gfat1* and most likely corresponds to *l(3R)81Fb*. 3 EMS induced alleles of *zep* (from the Bejsovec and Leptin labs) have been identified, along with numerous deficiencies which delete it, produced through both X-ray and male recombination.

P-GFP may define a novel fourth complementation group associated with *Df(3R)10-65*, based on its later lethal phase, a possibility which will need molecular analysis to confirm. However, it could also be a less severe allele of the same locus defined by the four new X-ray lesions: *l(3R)13B-688*, *l(3R)5A-388*, *l(3R)3A-204* and *l(3R)18A-1676*.

Figure 3.22. Summary map depicting proposed location of genes mapped by PCR.

The h58 interval is expanded, as indicated by the dashed lines, to show more detail. Scaffold AABU01002691 is placed at h56, based on PCR data from this work, and the cytology of Konev et al. (2003). The scaffolds containing *Parp* and *Gfat1* have both been placed in h55 by other researchers (see text), but PCR data suggest the relative locations in h58. A question mark is placed beside one line connecting the AABU01002763 scaffold with *Df(3R)10-65* because more data is required to be certain about whether/where the deficiency breaks within the scaffold.



It is important to note that to date, no lesion has been isolated which seems to delete further into 3R heterochromatin than h58, except *Df(3LR)6B-29*; however there are some new lesions, such *as Df(3R)8740-33*, which could be used in screens and which might yield mutations in previously unidentified loci. The question of whether there is something unusual in the configuration of the *Df(3L+R)6B-29* lesion is still open – but some new deficiencies have been generated that might allow this question to be addressed. Such deficiencies, produced by both X-ray and male recombination, can be used to screen for further, more extensive deficiencies, perhaps extending as far as *Df(3L+R)6B-29* appears to, but without the complexities that could exist for this lesion. The 3R heterochromatin is extensive, and several genomic scaffolds have been localized there, containing more gene models than can be accounted for by the small number of lethal complementation groups identified.

The Drosophila Heterochromatin Genome Project predicts many more gene models than can be accounted for by genetic analysis. In some cases, such as 3L heterochromatin (see next chapter), a number of genes have been to date inaccessible to genetic analysis due to the difficulty of obtaining and maintaining appropriate lesions with which to study them. In addition, it has been estimated that only about one third of all euchromatic genes are essential, and there is no *a priori* reason to assume that heterochromatic genes would be any different in this respect. So, by screening for lethal heterochromatic mutations, we are potentially overlooking perhaps two thirds of the loci therein. Finally, as

described in the Introduction, computer based annotation of heterochromatic genes is particularly difficult (though improving all the time) and there is a possibility that some of the predicted gene models may not be expressed genes, or that some are in fact, as yet unrecognized transposons.

The *Parp* locus has emerged as an enigmatic gene. Several lines of evidence would appear to place it in h58, between *Df(3R)10-65* and *Df(3R)EP-167.* Its placement here, along with the viability of both P-element induced alleles in combination with a number of deletions that appear to lack most if not all of the *Parp* coding sequence, suggest that it possesses unusual properties, perhaps a form of self-regulation that is complex. It is also significant, given the location of *Parp* between two deficiencies which delete *l(3R)81Fa* and *l(3R)81Fb,* that despite several screens, in addition to those reported here, no new alleles of *Parp* have been reported aside from the two alleles generated in the Zhang and Spradling (1994) screen. This may reflect the fact that the only means of identifying new alleles of *Parp* would have been through failure to complement *CH(3)1,* which, as demonstrated here, fully complements deletions which appear to remove most or all of the *Parp* coding sequence. The novel lesions generated here will allow a fuller and more detailed characterization of this fascinating gene.

The generation of a large number of new lesions and some new mapping information provides numerous potential future research directions (Figure 3.22). All of the new deficiencies should be examined cytologically, both to confirm the

presence of deficiencies and via use of genomic DNA probes or relevant BACs, to connect scaffolding information and the genetic maps with cytological data. The presence or absence of certain genes from particular deficiency intervals could also be assessed by using genomic Southern blots, since there are now many lesions to test, as well as using quantitative techniques such as real time PCR. An advantage of real time PCR is that one can test for relative doses of a particular DNA sequence, so that heterozygotes could be tested, rather than homozygous embryos. Several loci near the heterochromatin boundary have been identified which can be studied in greater detail – *CG12581* and *Dsk* for example – to determine whether or not they are essential and if so, the lethal phase and phenotype.

Df(3R) EP-167 might be used in screens for alleles of a small number of boundary genes, *Dsk*, *CG12581* and the *zeppelin* locus. The alleles of *zeppelin* should be sequenced to determine if they are mutant in the *Gfat1* locus, which seems like the most likely candidate gene for *zeppelin*, based on the preliminary PCR data, the genetic data, and its known role in chitin synthesis. Thus, it is unlikely to be positioned at h55 as indicated by the genome project, but more likely resides in h58. As no mutations of *Gfat1* have ever been isolated, the characterization of *zep* alleles should lead to a greater understanding of the functioning of this locus.

The *P-GFP* mutation should be studied further – imprecise excisions and/or much larger-scale male recombination experiments might help to determine what the affected gene is. Since it has been demonstrated that this insertion will undergo male recombination, at a low frequency, larger scale experiments could yield more recombinants. Finally, iPCR could also be performed, to try to obtain flanking sequence.

All of the deficiencies generated can be examined using single embryo amplified PCR and the data obtained confirmed using independent methods, either via real time PCR, Southern analysis, and/or cytological/FISH examination. Finally, the largest 3R deficiencies can be used to continue screening for EMS mutations in 3R. It remains possible that 3R heterochromatin has not been fully saturated by mutagenesis, since *Df(3LR)6B-29* was created by compound autosome detachment and could potentially be more complex than it appears.

The genetic resources reported here represent a significant step towards the ultimate goal of a complete functional annotation of the 3R heterochromatin and open up some intriguing questions about the nature and identities of the loci therein.

CHAPTER 4: CHARACTERIZATION OF 3L HETEROCHROMATIN

Introduction

3L heterochromatin has been much more thoroughly studied than 3R and many loci have been identified therein. Original genetic screens by Marchant and Holm (1988a, b) identified 10 essential loci resident within 3L heterochromatin (*l(3L)h1, 2 , 3, 4A, 4B, 5A, 5B, 6, 7,* and *8;* for convenience, referred to here as *lethal 1, 2,* etc.). Since that time, two loci have been shown to be a single gene with many alleles, some combinations of which produce small numbers of viable but sterile transheterozygotes with a variety of cuticular defects (*l(3L)h 5A* and *5B*, now referred to as simply *lethal 5*; Sinclair, unpublished). However, additional screens in the Deitcher lab (Vilinsky et al. 2002) have yielded two additional essential loci, *Snap-25* which is located very near *lethal 6*, and *lethal 7A*, which is apparently a new locus proximal to the original *lethal 7* (which has been renamed *lethal 7B*). Further screens, involving EMS, radiation and P-element mobilization (Schulze et al. 2001) did not yield alleles of any new loci.

For the most part, the 3L heterochromatic loci have been identified genetically, with no clear-cut molecular characterization (with the exception of the *Snap25* locus; Vilinsky et al. 2002). However, recently one of these was thoroughly characterized (Schulze 2003; Schulze et al. 2005). The gene in question is one of the most proximal of the 3L loci and is known as *l(3L)h2*, or *RpL15*, a ribosomal protein gene, deletions of which are associated with a weak *Minute* phenotype as heterozygotes. Some allelic combinations produce varying

numbers of homozygous escapers, which are substantially delayed in development and have a strong *Minute* phenotype, consisting of fine, thin bristles, sex comb gaps and misrotated genitalia. A number of P-element induced and EMS alleles were characterized and a partial rescue was obtained with an *RpL15* cDNA construct under the influence of a ubiquitous driver. *RpL15* expression was also demonstrated to decrease when the dosage of HP1 was reduced, a characteristic of two other heterochromatic loci examined (see introduction). A gene encoding a putative DEAD-box helicase, *Dbp80* – located very close to *RpL15* - was similarly characterized and responds to altered dose of HP1 in the same way. To date, no mutations of *Dbp80* have been isolated, despite a number of screens. It is likely to be a non-essential locus.

The other locus that has been characterized is *Snap-25* (*Synapse protein 25*; Vilinsky et al. 2002). This is a very large locus (more typical of the few heterochromatic loci that have been characterized to date), spanning over 200 kb and with a large number of mutant alleles generated (see Introduction for a description of *Snap-25*). *Snap-25* is located close to the distal region of the heterochromatin, near *l(3L)h6* (referred to as *lethal 6).* As no deficiencies exist which break between the two loci, their relative order has not yet been determined.

Many gene models exist on the BDGP database; these are unlocalized, and a subset of these must correspond to the various loci that are currently identified in

3L heterochromatin based only upon genetics. Thus one goal of this part of the project was to try to correlate coding sequences from the unlocalized scaffolds with functional genes.

In the course of this work, however, it became apparent that some portion of the heterochromatin had still not been uncovered by deficiencies. Our distal-most heterochromatic and proximal-most euchromatic deficiencies complemented; thus there was an interval between them that might contain essential loci. At the same time, I aligned a set of BAC sequences from the database and determined that several new loci could be directly connected to existing proximal euchromatic scaffolds (*Qm*, *CkII-α*, *n-AChR-α*, and *α-Cat*). Subsequently the BDGP released the alignment of the same BACs and three additional loci were identified through improved annotation approaches (*CG32230*, *CG32231* and *CG32350*). In addition, work by Corradini et al. (2003) utilized BACs for in situ hybridization and confirmed that the above-mentioned 7 genes were located at heterochromatic segment h47, along with the *nrm* (*neuromusculin*) locus, which was thought to be the most distal locus within h47. We expected that some of these loci would correspond to our most distal essential loci and began to test this using single embryo PCR techniques described in chapters 2 and 3. However it became clear that none of *Qm*, *CkII-α*, *n-AChR-α*, or *α-Cat* are absent from even the distal-most heterochromatic lesions (data not shown). Thus another and perhaps more pressing need was to try to obtain deficiencies in this interval and to use them to isolate and characterize mutants in the interval.

Very recently, an updated annotation of the region has been released, in which the gene order is altered relative to previous releases, and an eighth locus identified: *CG33217*. This change underlines the difficulty in working with these regions because of the repetitive nature of the sequence and incomplete information on the organization and makeup of heterochromatin. These difficulties also clearly demonstrate the need for thorough genetic and molecular characterizations of the interval, including the generation of deficiencies and other rearrangements which could help to answer questions of gene order and genomic scaffold overlap.

During the course of the work, two mutant alleles of the *casein kinase II alpha* (*CkII-α*) locus were isolated (Lin et al. 2002). The EMS-induced *Timekeeper* (*Tik*) allele shows dominant long period circadian rhythm defects; the strongest mutant of its kind to affect circadian rhythm. *Tik* is also recessive lethal, and can thus be used for mapping deficiencies and other lesions in the interval. A spontaneous partial revertant, *TikR,* is a strong loss of function mutant, which retains the recessive lethality but apparently lacks the dominant effects. Both of these alleles have been used to characterize newly induced mutations in distal h47.

The project utilized several P-elements that were inserted in and around this distal heterochromatic interval, h47. More became available as the project progressed. Any lesions generated were characterized as fully as possible. A haplo-insufficient locus, *Qm*, can account for the lack of deletions for this

segment; however, this same locus makes the recovery, maintenance and characterization of the lesions somewhat difficult.

The goals of this facet of my project were to attempt to generate lesions in the distal region of 3L heterochromatin which had apparently not previously been uncovered by mutagenesis, and to attempt to use PCR based methods to map the large number of heterochromatic gene models into deficiency intervals within the heterochromatin of 3L, including some more proximal intervals. Several lesions have been isolated in the distal region of 3L heterochromatin and several CG gene models have been tentatively placed into known deficiency intervals. In the course of the research, a mobilization experiment yielded apparent deficiencies in/near the *Snap-25* gene and these were genetically characterized. There is some evidence that a new locus has been identified in this interval. This is significant given the large amount of mutagenesis that has been devoted to 3L perhaps without yet saturating the region. Alternatively, as demonstrated in chapter 3, there is also some possibility of complex genetics within heterochromatin that may not be easily deciphered using traditional complementation analysis.

Results and Discussion

Generation of lesions in the uncharacterized distal heterochromatin region

P-element containing strains utilized

Several approaches were taken to generate new lesions in distal 3L heterochromatin, some of which made use of P-element insertion lines created by the gene disruption group (http://flypush.imgen.bcm.tmc.edu/pscreen/). 3 of the lines contain a P-element called *SuPorP*, which contains *su(hw)* (*suppressor of hairy wing*) sites around the *white*$^+$ reporter gene to minimize some position-effects (Roseman et al. 1995). The element also uses both *white*$^+$ and *yellow*$^+$ reporter genes; there is some evidence (Konev et al. 2003) that the *yellow*$^+$ gene is less susceptible to position-effects than the *white*$^+$ gene. Figure 4.1 shows a map of the distal heterochromatin region and the locations of P elements that were used for the work.

One of the lines was a modified pUAST construct, which had been engineered to contain a *Rox1* transgene (Kelley et al. 1999). The research addressed a difficulty prevalent in P-element mobilization experiments, namely the fact that many sites exist where transgenes have inserted but which cannot be detected

due to chromatin conformation in the surrounding sequences causing reporter gene suppression. These effects have been demonstrated in heterochromatin; w^+- containing transgenes inserted into or near heterochromatin can display very pale w^+ expression, no expression or variegated expression. However a number of euchromatic sites display this phenomenon as well. The goal of the work of Kelley et al. was to study the ability of the *rox1* transgene to recruit components of the dosage compensation machinery to ectopic, autosomal sites. The dosage compensation system is the mechanism by which male and female flies achieve equivalent expression levels of X-linked genes despite having different numbers of X chromosomes. The proteins and RNAs involved in dosage compensation act upon the male X to acetylate histones (histone H4 on lysine 20; reviewed by Gilfillan et al. 2004) and thus create a more "open" chromatin structure, and higher transcription levels. Kelley et al. (1999) predicted that chromatin structure near the transgene insertion sites would be altered – made more accessible – such that the w^+ reporter gene would be expressed in males, even if the transgene was inserted into a region where the reporter would normally be suppressed. The female siblings of the w^+ males would have white eyes due to lack of alteration in chromatin structure, since there should be no alteration in chromatin structure near the transgene in females.

Many such lines were isolated by Kelley and his colleagues, in which males had w^+ eyes but females did not. One of these lines was cytologically mapped to chromosome 3L at approximately 80C. iPCR flanking DNA from this line

indicated that it had inserted into a *hoppel* element (Kelley et al. 1999). Note the work of Sun et al. 2004, discussed in chapters 1 and 3; the presence of a *hoppel* element, at least on chromosome 4, appears to be a predisposing factor for reporter gene suppression due to recruitment of heterochromatin forming proteins. This supports the argument that, without the *rox1* gene on the P-element, this insertion line would never have been detected. In the uncharacterized region of distal 3L heterochromatin, there are 3 *hoppel* elements that are more or less intact and each contains an exact match to the flanking sequence obtained from the transgenic line. Thus it was not possible to determine the exact insertion site in the region. The *Rox80C* line showed viability with all of the deficiencies in the region, which indicated that it was likely present in the uncharacterized distal 3L heterochromatin. I utilized this insertion line for male recombination experiments with the prediction that such a construct would alter chromatin structure around the insertion site and thus might enhance the frequency of male recombination, since a more "open" chromatin structure might be expected to facilitate transposition (results shown later; see Table 4.8).

Imprecise excision of a P-element near the Qm locus

The strain *KG01176* contains an inserted *SuPorP* element near the *Qm* locus (see Figure 4.1). Because *Qm* encodes a ribosomal protein, it was presumed that imprecise excisions that deleted all or part of *Qm* might result in a classic *Minute* phenotype, consisting of delayed development and fine, thin bristles as

well as some other defects. Therefore a screen was undertaken to mobilize the element and select for *Minute* and/or w^- progeny. The scheme used is shown in Chapter 2. 4,866 males were scored for altered w^+ expression and/or *Minute*-like characteristics. 44 w^- or variegating w^+ lines were isolated, only one of which showed homozygous lethality. This one was further characterized. The line, called *KG01176-Y*, was lethal when transheterozygous with *Df(3L)Delta1*^{AK}, which implied that a lesion had been created that extended proximally from the site of the P-insertion. However, the lesion was not lethal in combination with a lethal *nrm* allele (*nrm*^{165}), indicating that the deficiency may not extend very far. The original parent line, *KG01176* was homozygous viable and viable in combination with *Df(3L)Delta1*^{AK}. The only potential overlap between *Df(3L)Delta1*^{AK} and *KG01176-Y* is the locus named *CG32458*, a putative gene whose function and properties are currently unknown and which is nested within the *nrm* locus. Alternatively, *KG01176-Y* could be the result of a transposition into a nearby locus. As no other lethal alleles of other loci deleted by *Df(3L)Delta1*^{AK} are available, it was not possible to test whether this had occurred.

Figure 4.1. Map of P-element positions for lines used in the mobilization experiments.

The solid rectangle represents a region of the distal h47 region, which consists of just over 350kb. Gene designations are given below the rectangle; P elements used in these experiments are depicted as filled triangles.

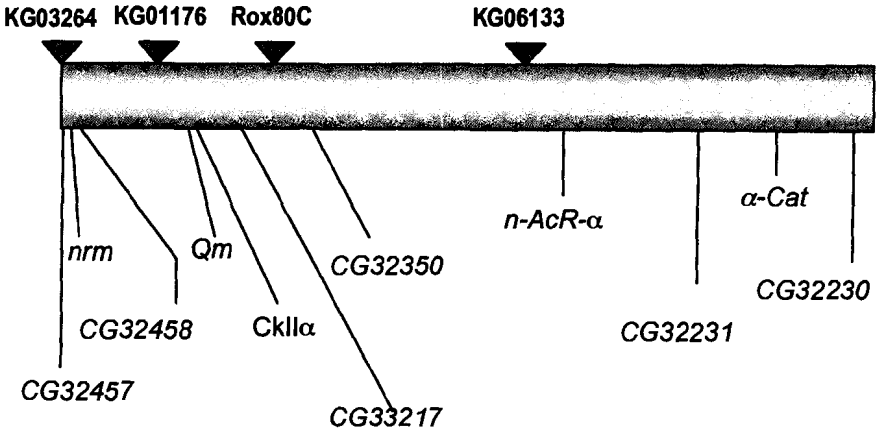
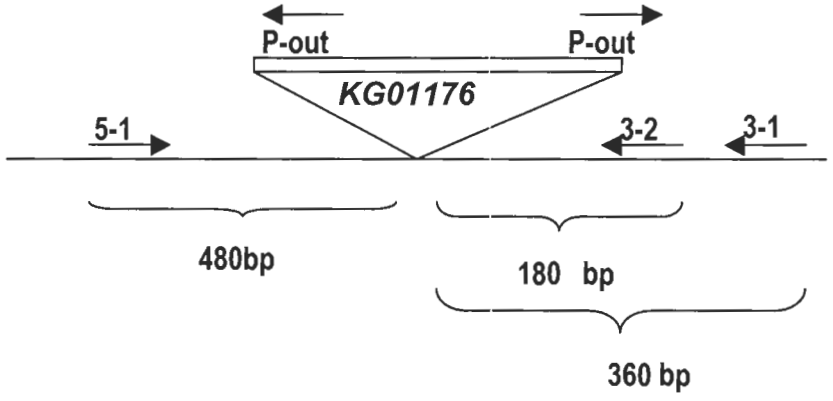


Table 4.1. Coordinates of the genes depicted in Figure 4.1.
Coordinates are shown for the entire arm of 3L (column 2) and the individual scaffolds on which the genes were originally located (column 3). Individual scaffold numbers are shown above the genes they contain. AE003599, 2986 and 2656 represent sections 80, 81 and 82 of 83 contiguous segments that together make up the 3L arm. Sizes of each scaffold are indicated in brackets next to the "AE" number. Section 83 (Scaffold AE002665) is just over 490 kb and contains only transposable elements. The asterisk indicates a gene that has only recently been identified and thus was not studied in this work.

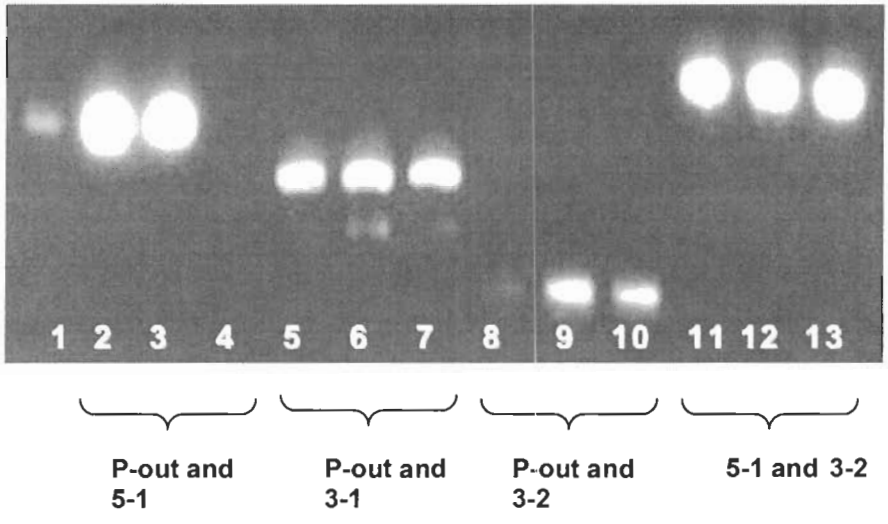
Scaffold: gene:	**3L genomic coordinates:**	**AE003599** (250,029) **coordinates**	**P element**	**3L Position:**
CG32457	22,921,351 - 22,922,000	194,984 - 195, 634	*KG03264*	22,894,357
nrm	22,932,282 - 22,971,879	205,915 - 245,512		
CG32458	22,974,864 - 22,968,282	238,497 - 241,915		
Scaffold: gene:	**3L genomic coordinates:**	**AE002986** (100,825) **coordinates:**	**P-element:**	**3L Position:**
Qm	23,027,471 - 23,029,315	51,135 - 52979	*KG01176*	22,999,228
CkII-α	23,033,421 - 23,037,303	57,085 - 60965		
*CG33217**	23,059,872 - 23,066,052	83,536 - 89,716		
Scaffold: gene:	**3L genomic coordinates:**	**AE002656** (203,636) **coordinates:**	**P-element**	**3L Position:**
CG32350	23,081,494 - 23,084,330	4630 - 7229	*Rox80C*	5' (unknown)
n-AChR-α	23,182,870 - 23,218,386	105,769 - 141,282	*KG06133*	23,175,671
CG32231	23,234,819 - 23,235,472	157,718 - 158,371		
α-Cat	23,252,515 - 23,271,946	175414 - 194845		
CG32230	23,273,167 - 23,275,447	196,066 - 198,346		

Figure 4.2. PCR characterization of the *KG01176-Y* lesion.
A. Schematic depiction of the P-element insert and relative positioning of primers used in PCRs. Expected band sizes are indicated below the brackets. **B**. PCR results. Lane 1 – 500bp band from the 1kb ladder, for reference; lanes 2, 3, 5, 6, 8, 9, 11, 12 - *KG01176/TM6* parent line as positive controls; lanes 4, 7, 10, 13 – *KG01176-Y/TM6*. Primers used for each set of reactions are indicated below the corresponding lanes of the gel. Primers 5-1 and 3-1 generated a band for all lines due to the presence of uninterrupted genomic template from the balancer chromosome; this set of reactions was included as an additional control for template and primer integrity. The *KG01176-Y* line fails to amplify a band on the 5' side, indicating the likely presence of a lesion on the distal side of the element.

A.



B.



In addition to the lethal line, 45 putative *Minute* lines were isolated in the screen. One, *KG01176-A1,* was strongly *Minute* in phenotype, and dominantly female sterile. It was found to be lethal in combination with an allele of *CkII-α,* named *TikR* (described in the introduction to this chapter). This strongly suggested that a lesion had been created which extended across the *Qm* locus and into the *CkII-α* locus. The lesion may extend further, but it hasn't been possible to test this. The fact that an apparent deletion of *Qm* created severely compromised flies and female sterility might explain the lack of genetic reagents for this interval – such lesions would not be easily obtained, due to delayed development, and then would be exceedingly difficult to maintain. *KG01176-A1* was crossed to several strains in this region to try to pinpoint how far it extended. A summary of the complementation results is shown in Table 4.2.

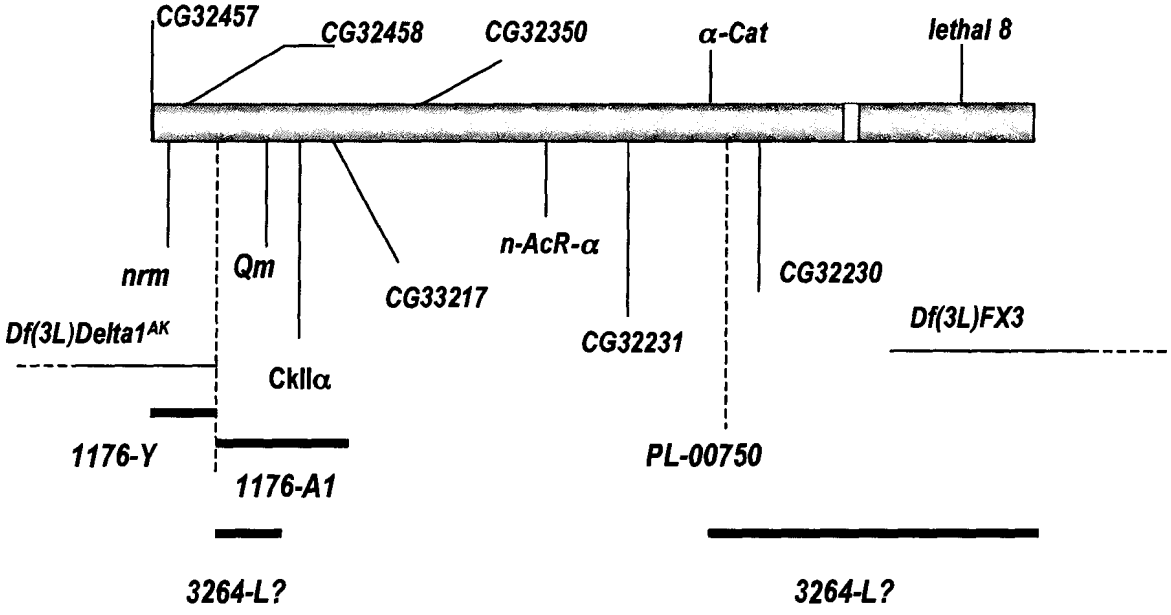
A screen for deletions using X-ray induced P-element loss

The strain *KG03264* contains an inserted *SuPorP* element distal to the *nrm* (see Figure 4.1) locus. It was X-irradiated in an attempt to create lesions that overlapped with the current proximal most deficiency in euchromatin, *Df(3L)Delta1*^{AK}, and then extended proximally toward the heterochromatin. The scheme used is shown in the Materials and Methods Section (Chapter 2). After 1,672 F1 males were scored, it became apparent that the screen was not working as originally planned. At that point only two males had been isolated with slightly altered w^+ expression and none that entirely lacked w^+ expression. It was likely that second site P-elements were present in the stock, a possibility that is

Table 4.2. Complementation summary for *KG01176-Y*, *KG01176-A1* and *KG03264-L*.
The asterisk indicates a noticeable wing phenotype in transheterozygous progeny.

mutant line	*Df(3L)Delta1AK*	*nrm165*	*Tik*	*Df(3L)FX3*	*G18*	*PL-00750*
KG01176-Y	L	V	V	V	ND	ND
KG01176-A1	V	ND	L	V	V	ND
KG03264-L	V	ND	V*	SL(?)	SL?	SL

Figure 4.3. Genetic map based upon complementation data from Table 4.2.
The horizontal bar represents the distal h47 region, with the white bar representing a gap of unknown length between *CG32230* and *lethal 8.* The vertical dashed line between lesions 1176-Y and 1176-A1 represents the original insertion site of the *KG01176* P-element; the vertical dashed line between *α-Cat* and *CG32230* represents a lethal insertion mutant named *PL-00750* (Hacker et al. 2003). Thin horizontal lines show two pre-existing deficiencies and the thicker horizontal lines indicate the presumed extents of the newly induced lesions listed in Table 4.2. Two potential locations are indicated for *3264-L*; it is not yet possible to determine if either location is correct.



consistent with the fact that P-element reporter gene expression in heterochromatin is generally very weak, or non-existent; and the selection of heterochromatic P-insertion lines based on w^+ expression is very frequently associated with second site insertions. This has been observed with several lines in our lab.

In any case, rather than discard the lines, the 1000 that were still available were crossed to 3L heterochromatin deficiencies (*Df(3L)FX3* and *Df(3L)MX18*; see Appendix A) to assay them for any mutations that may have been generated in 3L heterochromatin independently of the presence of the P-element insertion. The hope was that lesions might have been generated which overlapped with these two deficiencies and extended distally into the previously uncharacterized region of h47. 12 lines were isolated that showed lethality across the 3L heterochromatin; upon retesting, 9 of them were mapped into specific 3L heterochromatic intervals. The lesions are named "*l(3L)3264-* " with a single letter designation for those which appear to involve a single locus, and those which appear to affect multiple loci are designated "*Df(3L)3264-* ", again with a single letter designation. Table 4.3 shows a summary of *inter se* complementation and complementation with lethal alleles of loci within the 3L heterochromatin.

Figure 4.4 shows a map of these new lesions. Table 4.4 lists the alleles used for each of the loci in these experiments. 4 of the lesions appear to be multi-locus

deficiencies, with the largest being *Df(3L)3264-T*. This deficiency deletes all loci from *lethal 3* through to *lethal 8.* It is lethal in combination with the *lethal 8* allele, *G18*, only when the cross is performed with *G18* females and *Df(3L)3264-T* males. If the cross is performed in the opposite way, transheterozygous *G18/Df(3L)3264-T* survivors exhibit a *Minute*-like phenotype of fine, thin bristles, but are otherwise viable and fertile. This phenomenon has been observed with other lesions that apparently break between *lethal 7B* and *lethal 8. Df(3L)3264-E* deletes *lethal 2* through *4B* and *Df(3L)3264-A* deletes just *lethal 2* and *3. Df(3L)3264-D* is a discontinuous lesion, apparently deleting *lethal 4A* through *lethal 5* and then *lethal 7A* and *7B* more distally. This discontinuity of loci affected has been observed with numerous lesions in the 3L heterochromatin and may reflect that some are actually associated with inversions or other types of rearrangements in heterochromatin rather than simple deletions.

Three of the lesions apparently affect only single loci. *l(3L)3264-R* is lethal only with *lethal 3* alleles; *l(3L)3264-J* with *lethal 4A* alleles, and *l(3L)3264-G* with *lethal 6* alleles.

Two of the lesions have complementation patterns that involve semi-lethality, implying that they may not delete loci, but are perhaps affecting their expression in some other way. *Df(3L)3264-F* is lethal in combination with *Snap-25* alleles but semi-lethal with *lethal 6* alleles. Perhaps this lesion deletes *Snap-25* but does not delete *lethal 6.* Alternatively, it may have altered some regulatory sequence or other important binding sites near the locus that influence its

expression without completely compromising the gene. *l(3L)3264-W* is also semi-lethal in combination with both *lethal 6* and *lethal 7A* alleles. If this is a single site lesion, this might imply that some sequence has been deleted that affects the expression of both of these genes. If this were the case, it would mean that *lethal 6* lies distal to *Snap-25* and thus we would be able to order these two genes. *l(3L)3264-W* is also lethal in combination with *l(3L)32364-G*, which affects only *lethal 6* and semi-lethal in combination with *Df(3L)3264-F* which is also semi-lethal with *lethal 6*. The overlap of all these lesions points to a distal position for *lethal 6*, and will require single embryo PCR, and other analyses to determine which coding sequences are absent from each lesion.

One lesion shows complexity in its complementation pattern. *l(3L)3264-C* shows lethality in combination with *Df(3L)3264-T* and *D*, and with two previously identified deficiencies (*Df(3L)O-1* and *Df(3L)γ-28*; see Appendix A) which in combination appear delete only lethal 5. However *l(3L)3264-C* is viable and fertile in combination with all *lethal 5* alleles tested. It is possible that another locus lies in this region, which has not previously been identified, that lies near *lethal 5*, to one side or the other. However, the *lethal 5* locus does display numerous cases of inter-allelic complementation at least for viability (Sinclair, unpublished, reported in Fitzpatrick et al. 2005). Although *l(3L)3265-C* is viable and fertile in combination with members of both groups of *lethal 5* alleles, it is possible that it may define a third class of *lethal 5* alleles which complements the other two. A molecular analysis would be required to rule out this possibility

Table 4.3. Complementation data summary for new X-ray induced lesions in 3L heterochromatin.

A. *Inter se* complementation. The prefix 3264- has been left off the top row. **B.** Complementation with mutations in essential complementation groups in 3L heterochromatin. L= lethal, V=viable, SL=semi-lethal. Underlined regions indicate multi-locus deletions.

A.

lesion	A	C	D	E	F	G	J	R	T	W	X
3264-A	L										
3264-C	L	L									
3264-D	V	L	L								
3264-E	L	V	L	L							
3264-F	V	V	V	V	L						
3264-G	V	V	V	V	V	L					
3264-J	V	V	L	L	V	V	L				
3264-R	L	V	V	L	V	V	V	L			
3264-T	L	L	L	L	L	L	L	L	L		
3264-W	V	V	V	V	SL	L	V	V	L	L	
3264-X	L	V	V	L	V	V	V	SL	L	V	L

B.

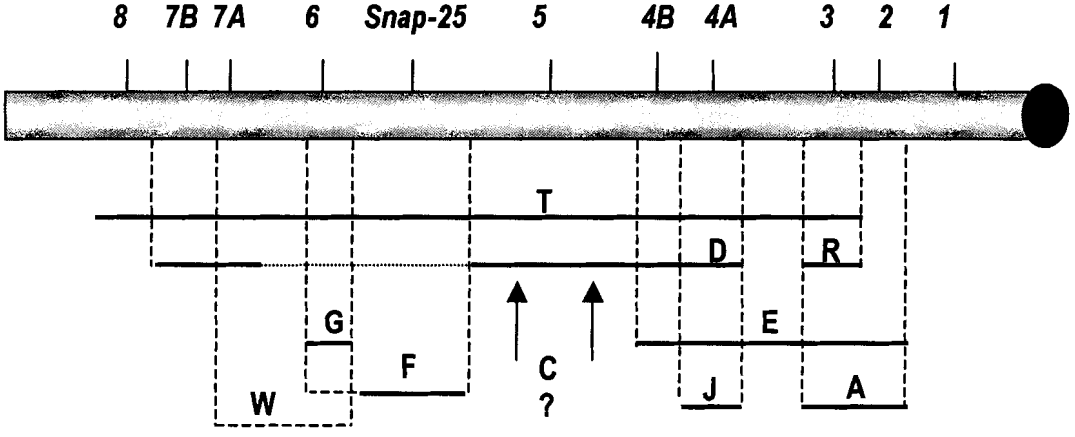
line	*leth 1*	*leth 2*	*leth 3*	*leth 4A*	*leth 4B*	*leth 5*	*Snap-25*	*leth 6*	*leth 7A*	*leth 7B*	*leth 8*
3264-A	V	L	L	V	V	V	V	V	V	V	V
3264-C	V	V	V	V	V	V	V	V	V	V	V
3264-D	V	V	V	L	L	L	V	V	L	L	V
3264-E	V	L	L	L	L	V	V	V	V	V	V
3264-F	V	V	V	V	V	V	L	SL	V	V	V
3264-G	V	V	V	V	V	V	V	L	V	V	V
3264-J	V	V	V	L	V	V	V	V	V	V	V
3264-R	V	V	L	V	V	V	V	V	V	V	V
3264-T	V	V	L	L	L	L	L	L	L	L	L
3264-W	V	V	V	V	V	V	V	SL	SL	V	V

Table 4.4. Alleles used to determine minimal genetic extents of new KG03264-derived lines.

Included are the gene names, allele names, the methods by which the alleles were induced, the source of the mutants and any other pertinent information.

Gene name	allele	induced by:	source/ref	notes/other info
l(3L)h1	*In (3L)C90*	X-ray	Robert 1970	second breakpoint at 62D on 3L
	Df(3L)2-30	γ-irradiation	Marchant and Holm 1988a	duplication of part of h52 (Koryakov et al. 2002)
l(3L)h2	*72*		Schulze et al. 2001	second site at *l(3L)h5* (unpublished observation)
	1-166-37	EMS	Marchant and Holm 1988b	
l(3L)h3	*1-166-38*	EMS	Marchant and Holm 1988b	
	3-1[29]	P-element	Schulze et al. 2001	wild type P-elements: all other background P-elements removed (Schulze 2003)
	vtd4	EMS(?)	Schulze et al. 2001	
l(3L)h4A	*CH(3)53*	P-element	Zhang and Spradling 1994	
	CH(3)4d	P-element	Zhang and Spradling 1994	
l(3L)h4B	*G10*	EMS	Vilinsky et al. 2002	
	G30	EMS	Vilinsky et al. 2002	
l(3L)h5	*G5*	EMS	Vilinsky et al. 2002	
	G12	EMS	Vilinsky et al. 2002	
	1-16-I	γ-irradiation	Schulze et al. 2001	
	1-16-N	γ-irradiation	Schulze et al. 2001	
l(3L)h6	*CH(3)4*	P-element	Zhang and Spradling 1994	second site lethal at *Parp* locus, on 3R (Zhang and Spradling 1994; Tulin et al. 2002)
	G4	EMS	Vilinsky et al. 2002	
	G59	EMS	Vilinsky et al. 2002	
Snap-25	*CH(3)7*	P-element	Zhang and Spradling 1994	
	1-16-O	γ-irradiation	Schulze et al. 2001	
l(3L)h7A	*l(3L) fsa2*	EMS	Felsenfeld and Kennison 1995	only extant allele of this locus
l(3L)h7B	*G43*	EMS	Vilinsky et al. 2002	
	G1	EMS	Vilinsky et al. 2002	second site lethal between *l(3L)4B* and *l(3L)5* (unpublished observations)
l(3L)h8	G18	EMS	Vilinsky et al. 2002	only extant allele of this locus unusual parent specific complementation - see text

Figure 4.4. Complementation map for the new X-ray induced lesions. Diagram is not to scale. Essential loci are indicated above the bar, which represents the 3L heterochromatin. Solid lines indicate those loci with which the lesions are fully lethal, dashed lines indicate semi-lethality. Lesions are referred to by letter designation for simplicity. Lesion "D" (*Df(3L)3264-D*) contains a dotted line which indicates the discontinuity of the deficiency. The dotted line indicates the loci with which the lesion is viable. Arrows beside *l(3L)3264-C* indicate possible positions for this lesion. See text for further discussion.



(which would itself require the identification of the coding sequence that corresponds to *lethal 5*).

There is no evidence that any of the newly isolated lesions extend further than the *FX3* or *MX18* deficiencies, since the most extensive and distal of the deletions, *Df(3L)3264-T*, shows complete lethality with *G18*, the *lethal 8* allele, only when the cross is made with *G18* females and *Df(3L)3264-T* males. This strongly suggests that the lesion extends no further than the other lesions which show a similar complementation pattern, and not as far as *Df(3L)FX53*, which *G18* fails to complement in both maternal and paternal crosses.

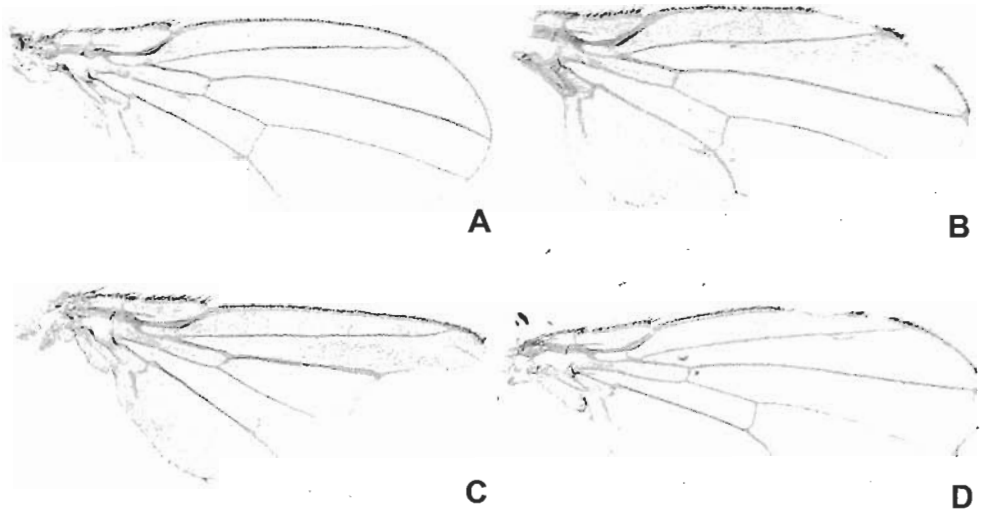
In addition to the deficiencies described above, 3 *Minute* lines were isolated in the course of the screen. Preliminary complementation tests using one of the lines indicated a possible connection with the h47 distal heterochromatin. Flies bearing the *3264-L* lesion are very debilitated; dominantly female sterile and severely *Minute*. The stock must be maintained in combination with a wild type chromosome and selected every generation. The phenotype of the heterozygous *KG03264-L* flies includes thin, fine bristles, rough and reduced eyes, wing notches and wing vein defects. The similarity of *KG03264-L* defects to the *KG01176-A1* line prompted further tests.

A summary of the complementation data for *KG03264-L* is shown in Table 4.5. Heterozygous *KG03264-L* males were crossed to *TikR*, an allele of *CkII-α*, and

Table 4.5. Complementation data for *KG03264-L*.
Shown are those crosses for which some semi-lethality was observed and a description of any phenotypes observed.

Male parent:	female parent:	Progeny:			
		+/TM3	*+/G18e*	*3264-L/TM3*	*3264-L/G18e*
3264-L/+	*G18e/TM3 (F)*	22	26	6	4 (males)
		+/TM3	*+/PL-00750*	*3264-L/TM3*	*3264-L/PL0075*
3264-L/+	*PL-00750/TM3Ser*	15	44	30	1 (male)
		QIII/TM3	*3264-L/TM3*	*3264-L/QIII*	
3264-L/TM3	*QIII/TM3 (29C)*	74	5	7 (more severe M)	
3264-L/TM3	*QIII/TM3 (22C)*	48	16	10 (more severe M)	

Figure 4.5. Wing defects observed with *KG03264-L* in transheterozygotes.
Genotypes: **A,** *l(3L)KG03264-L/QIII*; **B**, *l(3L)KG03264-L/TikR*; **C,** *l(3L)KG03264-L/TM3Sb*; **D,** *l(3L)KG03264-L/TM6Tb Hu*.



the only lesion known in the region. Transheterozygous progeny were viable and fertile, though with some wing vein and wing margin defects (Figure 4.5). This indicates that the lesion associated with *KG03264-L* does not affect the *CkII-α* locus. Similarly, *KG03264-L* is viable in combination with the *neuromusculin* allele, nrm^{165}. Since the P-element in the original *KG03264* line was inserted distal to the *nrm* locus, these data suggest that there has not been a deficiency produced at that position. Because both *KG01176-A1* and *KG03264-L* were female sterile, there was no opportunity to test for complementation between them.

However, there is some evidence that the *KG03264-L* lesion may be located in or near h47. Tests with *Df(3L)FX3* showed that the *KG03264-L* chromosome is semi-lethal in combination with the *FX3* deficiency. Thus, the *Minute* phenotype might be the consequence of a lesion in or near more proximal heterochromatin. The most compelling result supporting this assertion is that of the cross between *KG03264-L* and a lethal *piggyBac* insertion between the *α-Catenin* and *CG32230* loci, *PL-00750* (Hacker et al. 2003). Only 1 transheterozygous fly – a male - was observed of 90 in total (Table 4.5). This fly was severely *Minute.* This suggests that the lesion in the *KG03264-L* strain may encompass or otherwise affect the *α-Cat* (or *CG32230*) locus.

It was of interest to determine whether the *KG03264-L* mutant might coincide with the *QIII* locus, also known as *M(3)80A*, a temperature-sensitive *Minute* locus

isolated, characterized and mapped near the chromosome 3L heterochromatin by Sinclair et al. (1981). *QIII* is homozygous viable at low temperature with moderate *Minute* characteristics, but heterozygotes at room temperature are wild type. At 29 ^{o}C, the heterozygotes exhibit all of the dominant traits of the *Minute* phenotype – fine thin bristles, rough eyes and reduced viability. The homozygotes die during the first larval instar at 29^{o}C (Sinclair et al. 1981). *QIII* has been shown to be distinct from the *Qm* locus through work by Kay et al. (1988) who were unable to rescue *QIII* with constructs containing the genomic region of *Qm*, which encodes ribosomal protein Rp21 (Sinclair, unpublished). If *KG03264-L* is not associated with *Qm* it might still be somewhere nearby on the chromosome based on the complementation data described above. Therefore complementation tests were performed to try to ascertain whether the two lesions defined a single locus.

KG03264-L males crossed to *QIII* females at 29^{o}C, yielded 74 QIII/TM3 progeny, and 5 *KG03264-L/TM3* progeny with strong *Minute* phenotypes, and 7 *KG03264-L/QIII* flies with malformed thoraces, very reduced and rough eyes, very few and very thin bristles. These died soon after eclosion and were never tested for fertility. Although at the restrictive temperature, *KG03264-L* appears to exhibit semi-lethality with *QIII,* the *KG03264-L/TM3* flies were also strongly compromised. Crosses between *KG03264-L* and *Oregon R* demonstrated that the viable transheterozygotes are usually more plentiful than the *KG03264-L*/balancer flies, sometimes with less severe phenotypes, although in the case of

TikR, strong wing phenotypes were observed in transheterozygous flies. For the present, it is not possible to determine unequivocally whether *KG03264-L* might be associated with *QIII*. Only complete lethality would have confirmed that.

*KG03264-L/TM3*males crossed with *QIII/TM3* females yielded viable *KG03264-L/QIII* offspring at room temperature (22 $\pm$ 2 ^{o}C), which were less severe in phenotype than *KG03264-L* in combination with some other mutations (such as *TikR*) or balancers. For instance, the wings lacked the notched wing margins and extra wing vein material (Figure 4.5). A hallmark of *Minute* loci is that mutations in different *Minute* loci are non-additive; that is the phenotype of two combined mutant loci should be the same as the most severe of the two, since both mutations affect the same final product: the ribosome (Sinclair et al. 1981). The *QIII* mutant is wild type in the heterozygous state at 22^{o}C and *KG03264-L/+* males exhibit variable wing defects. The data do not allow us to determine whether or not *QIII* and *KG03264-L* are allelic, but do demonstrate the importance of genetic background in modulating phenotype as the *KG03264-L* chromosome exhibits quite variable defects depending upon the chromosome with which it is combined (Figure 4.5).

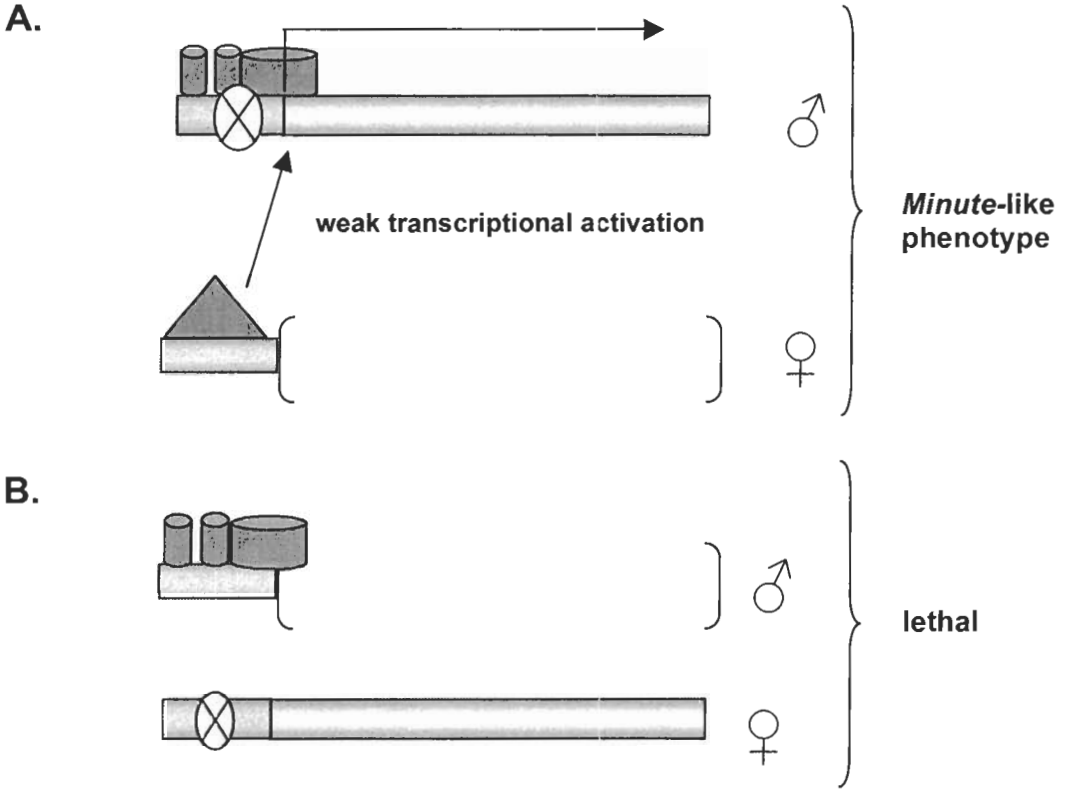
Another gene of interest in the 3L region is *lethal 8*, an allele of which (*G18*) displays unusual complementation behaviour, as has been described above. This could be the result of two interacting phenomena – imprinting and transvection. It is possible that the gene, *lethal 8,* is paternally imprinted at the

regulatory region. Suppose also that the mutant allele *G18* is mutant in the regulatory sequences such that transcription cannot be initiated. Finally, it is possible that the deficiencies that show the unusual complementation pattern with G18 delete coding but not regulatory sequences. If the *G18* allele was derived from a male parent, the regulatory region would be mutant and imprinted and thus inactive. However, the regulatory sequences from the female-derived deficiency could – if the chromosomes pair appropriately – initiate low levels of transcription of the intact coding sequence on the male-derived homologue. This phenomenon has been observed in several loci (reviewed by Duncan 2002). The reciprocal cross would give different results, however. If the G18 mutant allele is derived from the female parent and the deficiency chromosome from the male, now the intact regulatory sequence from the deficiency chromosome would be inactive due to imprinting as a result of transmission through the male germline. A schematic depiction of this model is shown in Figure 4.6; this highly speculative model requires further testing and analysis for confirmation. Other explanations are also possible; for instance, a simpler model might involve a dominant maternal effect of the *G18* mutant. Further alleles of *lethal 8* might help to address this question.

Crosses between the *KG03264-L* mutant and *G18* were performed in the only configuration possible, between *G18* females and *KG03264-L* males. The data are difficult to interpret given the poor viability of *KG03264-L* in combination with the balancer chromosome; however the *G18* and *KG03264-L* chromosomes may

Figure 4.6. Schematic depiction of possible explanation for unusual complementation between *lethal 8* allele, *G18*, and certain deficiencies.

The filled rectangles depict the gene, *lethal 8*, with regulatory and coding sequences shaded distinctly. The triangle represents a transcription factor, which may be capable of initiating transcription on a suitably paired homologue. Dark cylinders represent an imprint on the paternal chromosome at the regulatory region. The white circles indicate a regulatory mutation in the *G18* allele and the brackets depict the coding sequence missing from the deficiency chromosome. **A.** According to the model, through transvection, intact regulatory sequence from the deficiency chromosome may promote a low level of transcription from the coding sequence on the homologue, but only if the deficiency chromosome is maternally derived. **B.** In the converse situation, the deficiency chromosome is paternally derived and the regulatory region imprinted such that no transvection is possible.



be semi-lethal in combination. If *KG03264-L* failed to complement *G18*, then the cross, as performed, would be expected to yield no transheterozygotes as the *G18* allele was introduced through the female parent. The most likely explanation for the apparent semi-lethality is the effect of genetic background. However, if this explanation is adopted here, it must be applied to all instances of semi-lethality observed, including that with *Df(3L)FX3*.

Though the possible failure to complement the *PL-00750* lesion near *α-Cat* is promising, it would be premature to draw any conclusions until more mutants are isolated with which to test this. It is unfortunate that the *KG03264-L* line is so weak and difficult to work with; complementation with the *Minute* mutant is difficult because of the relatively poor viability of heterozygotes alone. Homozygous deficiency (or mutant) embryos cannot be obtained as very few females are observed, but those that survive are invariably sterile. Thus, although *KG03264-L* has been kept and characterized, it is not a convenient lesion to use for studying the region. *KG01176-A1* does appear to delete some of the h47 interval, but the dominant female sterility and poor viability make it difficult to work with also. It was therefore of interest to try to obtain further lesions.

Mobilization of KG06133.

The *KG06133 SuPorP* element is inserted near the *n-AChR-α* locus in distal h47 (Figure 4.1). A second insertion site associated with sequence contained in an unlocalized scaffold (AE002583) was reported for this line. The line displayed a very small amount of variegated w^+ expression. Initially, the low level and variegated pattern of w^+ expression characteristic of this line suggested that the second site might be on another chromosome, and lost from the balanced stock. An initial experiment (by D. Sinclair) involved the mobilization of the *KG06133* line. Flies were selected for lack of w^+ expression and then tested in combination with *Df(3L)FX53* (see Appendix A). 12 lines were generated and 11 of these were lethal in combination with *Df(3L)FX53*. Tests with individual mutant alleles of genes deleted by the *FX53* deficiency showed that all of the excision lines were lethal with both *lethal 6* and *Snap-25* alleles (see Figure 1.3). The second site flanking sequence associated with the *KG06133* element can be found on the large genomic scaffold: AABU01002775; the one on which the *Snap-25* locus resides. The flanking sequence matches part of a repetitive element, which is found in several locations throughout the genome, but two perfect matches exist, one on either side of the *Snap-25* locus itself. Genetic analysis showed that a mutation, *G8*, in the *lethal 4A* locus (see Figure 1.3 for map; *lethal 4A* is separated from *Snap25* by at least two essential loci and maps more proximally than *Snap25*) was lethal in combination with all of the excision lines although it was viable with *Snap25* and *lethal 6*. It appears that the *G8* mutant line bears a

Table 4.6. Complementation of *KG06133* excision lines with lesions near *Snap-25*.

For *4A* and *4B*, all mutant alleles were used, but results for the *G8* allele of *lethal 4A* are shown separately (column 6) because this allele bears a putative second site near *Snap-25*. The following mutant alleles were used: *lethal 6: CH(3)4, CH(3)119*; *Snap25:* CH(3)7, *1-16-0*; *lethal 7A: fsa*^{2} . J, G and D in the last three columns refer to *Df(3L)3264-J*, *Df(3L)3264-G* and *Df(3L)3264-D,* respectively.

line#	*l4A*	*l4B*	*l6*	*Snap-25*	*G8*	*l7a*	*J*	*G*	*D*
1	V	V	L	L	L	V	L	L	V
3	V	V	L	L	L	V	L	L	V
4	V	V	L	L	L	V	ND	ND	ND
5-1	V	V	L	L	L	V	L	L	V
5-2	V	V	L	L	L	V	ND	ND	ND
5-3	V	V	L	L	L	V	ND	ND	ND
6-1	V	V	L	L	L	V	L	L	V
6-2	V	V	L	L	L	V	ND	ND	ND
7	V	V	L	L	L	V	L	L	V
8	V	V	L	L	L	V	L	L	V
10	V	V	L	L	L	V	L	L	V

Figure 4.7. Schematic map of scaffold AABU01002755 with two possible P-element insertion sites for KG06133.
The scaffold is 712,331 bp in length and represented by the large rectangle. Two potential insertion sites are shown as dark triangles above the scaffold. Only 4 genes are shown on the map; these are genes for which cDNAs exist. *CG40050* is probably the *lethal 6* locus (Syrzycka, unpublished).

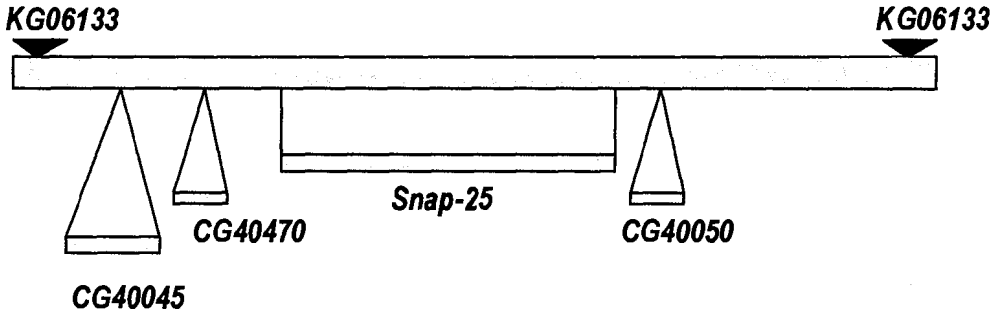


Table 4.7. Candidate genes for a novel lethal lesion near *Snap-25*.
Genes included are shown on Figure 4.7, above and the positions of the flanking sequence of the KG06133 second site insertion are also shown. The genes with unknown functions produce transcripts and proteins, according to the genome annotations.

candidate gene	possible function	location on scaffold
second site flanking sequence	n/a	23,390-23,770
CG40045	ubiquitin ligase	99,114-151,125
CG40470	unknown	192,633-199,972
Snap-25	synapse associated protein	238,366-460,663
CG40050	unknown	486,325-509,801
second site flanking sequence	n/a	688,335-688,715

second site lethal mutation affecting a gene that lies close to the Snap 25 and lethal 6 genes.

Alternatively it could be an allele of one of the two, which displays inter-allelic complementation. The excision lines tested apparently do not break between the two loci (see Table 4.6 for a summary of the complementation data). This is significant because *lethal 6* appears to be associated with *CG40050*, on the same scaffold as the *Snap-25* locus and only approximately 25kb away from it (M. Syrzycka, unpublished results). See Figure 4.7 for a schematic map of the *Snap-25* scaffold, the two potential sites of the second site P-insertion and several CG gene models which appear to be possible candidates for the second site lethal on the *G8* chromosome. Table 4.7 lists the candidates and their putative functions, and homologies – the second site lethal is predicted to lie between the insertion site on the scaffold and the *Snap-25* locus.

MR-generated deficiencies.

Despite the fact that the excised lines carried lethal lesions elsewhere on the chromosome, it was decided to use one (line 1, called *KG06133-ex1*) of the excision lines as a starting point to perform male recombination experiments. These lines still were, at the time, the only suitable ones available (i.e. in position) for mobilization experiments to isolate lesions in the distal

uncharacterized region. A schematic representation of the crosses performed in these experiments is shown in the Materials and Methods section.

The male recombination experiment with *KG06133-ex1* yielded 27 recombinant lines, of which 16 were deemed to be unique recombinant events (Table 4.8). Of these, 11 were recessive lethal and could be used in complementation tests. The *ri* p^{p+} lines were tested with *Df(3L)Delta1*^{AK}, *TikR*, *PL-00750*, as well as *Df(3L)FX3* and *G8* to test for the second site lethality associated with the original line. The results are shown in Table 4.9A. None of the lines has generated a lethal lesion except line *6133-15*, which appears to be associated with a deficiency that extends distally at least as far *as Df(3L)Delta1*^{AK} (see discussion of characterization of this lesion, below). All ri^{+} p^{p} lines were crossed to *G18*, the *lethal 8* allele, to identify lesions extending proximally. They were also tested for the presence of the second site lethal near *Snap-25* by crosses with both *Df(3L)FX3*, which deletes all loci from *lethal 3* to *lethal 8*; and with the *G8* allele to check for the second site lethality more directly. The results are shown in Table 4.9B. Two of the three lines appear to have retained the second site lethality; however one line, *6133- 6*, is viable in combination with *Df(3L)FX3.* This line is also lethal in combination with *6133-15*, the deficiency which has apparently been induced in distal h47. Therefore, the line needs to be characterized further. It could be a deficiency that extends proximally, but not as far as *lethal 8*, as evidenced by the viability and lack of phenotype of *6133-6* in combination with *G18.* At the time the experiments were in progress, the *α-Cat* locus in distal h47

was thought to be distal to the site of insertion of the *KG06133* P-element. So, the *ri* p^{p+} lines were tested for lethality in combination with the *PL-00750* line, but the others were not. In light of the recent re-ordering of the loci, this must be tested for all of the *ri*+p^{p} lines.

A subset of the *KG06133* – derived lines were subjected to PCR using P-out and a pair of primers specific to the genomic sequence flanking the insertion site near the *n-AcR-α* locus. Results of the PCRs are shown in Table 4.10. Note that in most cases the P-element has apparently excised from the site.

The most important line generated is *KG06133-15*; flies from this line have *Minute* bristles and the *6133-15/TM3* or *TM6* progeny from the balancing cross were delayed in development by at least 4 days at room temperature, with a strong *Minute* phenotype. Attempts to create a balanced stock failed due to the fact that the lines were dominantly female sterile, just as were the two *Minute* lesions previously described. Thus, complementation analysis could not be performed with these other two *Minutes*. Although I have been unable to perform the definitive test, the failure of both *KG06133-15* and *KG01176-A1* to complement *TikR*, an allele of the *CkII-α* locus, strongly indicates that these two both delete the nearby *Qm* locus, which encodes a putative ribosomal protein.

Genetic characterization of KG06133-15

The *KG06133-15* lesion failed to complement *Df(3L)Delta1*^{AK} as well as *TikR* (Table 4.9A). Crosses to the *G18* mutant line produced viable transheterozygous progeny that were female sterile and *Minute*, but more or less wild type in other respects. These data together imply that the lesion created by male recombination extends from the site of the P-element, distally, deleting the loci from *n-AChR-α* to the *Delta1*^{AK} deficiency. In addition to the crosses shown in Table 4.9A, for initial characterization of all the recombinant lines, two additional crosses were performed with *KG06133-15*. *nrm*^{165} allele was not lethal in combination with the *KG06133-15* lesion; nor was *chro*^{3258}, an allele of *chromator*, the next locus for which we have a mutant allele within the *Delta1*^{AK} deficiency. This shows that the *KG06133-15* lesion does not extend very far into the *Delta1*^{AK} deficiency region. So, the distal breakpoint of the lesion must lie somewhere between *nrm* and *Qm*. A technique such as real time PCR, once optimized, might allow us to more definitively determine the exact loci and sequences absent under this lesion. As well, most of the loci removed by the *Delta1*^{AK} deficiency have not been studied, and few lethal mutations exist. The isolation and mapping of more lethal alleles within this deficiency interval would be useful for characterizing *Df(3L)KG06133-15* and other deficiencies in the region.

Table 4.8. Summary of male recombination results for KG06133-ex1 and Rox80C elements.
The asterisk indicates that there may be one other lesion in this region, but this remains to be tested (see text).

Experiment	KG06133	Rox80C
Type of element	SuPorP	GMRox1
Number of males examined	12,530	3,920
Number recombinants	27	8
% recombinants	0.22%	0.20%
Number unique events	16	4
% unique events	0.13%	0.10%
# lethal lesions in distal h47	1*	1
% lesions	0.01%	0.03%

Table 4.9. Complementation summary for homozygous lethal male recombinant lines from KG06133-ex1.

A. *ri* p^{p+} lines

Line #	*Delta1AK*	*TikR*	*PL-00750*	*Df(3L)FX3*	*G8*
2	V	V	V	L	ND
12	V	V	V	ND	V
15	L	L	V	L	L
20	V	V	V	L	L
22	V	V	V	L	L
30	V	V	V	L	L
33	V	V	V	L	L
34	V	V	V	L	L

B. ri^{+} p^{p} lines

Line #	*TikR*	*Df(3L)FX3*	*G8*	*G18*
6	V	V	ND	V
10	N	L	L	V
18	N	L	L	V

Table 4.10. Summary of PCR characterization to test for presence of original P-element in male recombinant lines.

Primers used are described in Appendix B. + indicates the generation of a product with the primers specified, - indicates the absence of a product.

line #	KGN5-1+ P-out	KGN3-1+P-out
6133-1	-	+
6133-2	-	-
6133-6	-	-
6133-10	-	+
6133-12	+	-
6133-15	-	-
6133-18	-	-
6133-20	-	-
6133-22	-	-
6133-30	-	-
6133-33	-	-

An examination of mitotic chromosomes from *Df(3L)KG06133-15* indicates that the lesion is not large enough to be detected cytologically. (A. Coulthard, personal communication). Given that the limits of cytological analysis on orcein stained mitotic chromosomes is in the range of perhaps several hundred kb, this indicates that the lesion must be smaller than this. This is not unlikely, given that the loci thought to be deleted collectively cover only approximately 200 kb.

Phenotypic Characterization of Df(3L)6133-15

Df(3L)6133-15 causes a strong *Minute* phenotype with variable defects in combination with specific balancers; in combination with the *TM3SerGFP* balancer, deficiency-bearing males and females were sterile, the males having non-everted genitalia. In combination with *TM3SbGFP* or *TM6,* deficiency-bearing males are fertile enough to keep a "balanced" stock (see description below).

Females bearing *Df(3L)KG06133-15* were sterile in combination with all of the balancers tested. However, examination of the ovaries showed a difference in phenotype between the *TM3SbGFP* and *TM6* balancers. In *Df(3L)KG06133-15/TM3SbGFP* females, the ovaries were small and poorly developed. In contrast, *Df(3L)KG06133-15 /TM6* females had ovaries that looked fairly normal, with apparently mature eggs. Nonetheless, these flies laid only a few eggs that failed to hatch.

In combination with the *TM3SbGFP* balancer, the *Minute* bristles of *6133-15* were readily apparent, even in a *Stubble* background, but in addition certain bristle defects and a variable rough eye phenotype could be seen. Missing and duplicated dorsocentral (DC) bristles, and posterior supraalar (PSA) bristles were observed, but these were observed in less than 10% of the flies examined. The most common bristle defects were duplications of the anterior scutellar (ASC) bristles, and even some cases of tripled ASCs. The next most common defect was missing ASC bristles. Some sex differences were noted but these may reflect the small sample sizes more than real sex-specific effects. The *Minute* phenotype of *Df(3L)6133-15* may have these pleiotropic effects through altering doses of proteins needed in strict amounts for correct function (such as *Notch*, *Serrate* etc.), which then interact with lesions on the balancer chromosome. This would explain the balancer influence on the phenotypes observed – these were generally more severe in combination with the *TM3* balancer than the *TM6* balancer. Interactions with unrelated loci have been documented for some *Minute* loci examined in the past. (for example, Sinclair et al. 1981; 1984; Saeboe-Larssen et al. 1998). *Minute*-induced compromised expression of loci involved in signaling, such as the *Notch* pathway (involved in eye, wing and bristle formation) could be responsible for the pleiotropy of this *Minute* lesion. Indeed, mutant alleles of the *RpL15* locus have been shown to enhance a weak Notch allele, resulting in enhanced wing notching and in some combinations, thickened wing veins (Schulze 2003; Schulze et al. 2005).

One wing defect was quite unusual and fairly frequent in combination with the *TM3SbGFP* balancer: an additional anterior crossvein, or the beginnings of one were frequently visible between L2 and L3, just anterior to the real crossvein between L3 and L4. Because this phenotype was variable in severity and not apparent in all flies, it might be attributable to a partially penetrant effect of the *Minute* mutation. These seems likely, since the defect could be occasionally seen in *Df(3L)KG06133-15/TM6* flies also. A summary of the wing, bristle and eye defects is shown in Table 4.11. Thoracic defects are shown in Figure 4.8; wings in Figures 4.9 and 4.10; and eye defects in Figure 4.11.

Scanning electron microscopy shows the severity of the *Minute* phenotype of *Df(3L)6133-15* over both balancers tested. Fine, thin bristles and delayed development are two traits that did not vary among the *6133-15* flies, no matter which balancer was used. Figure. 4.7 shows sample images of the thoracic defects along with controls.

In combination with the *TM3SbGFP* balancer, the deficiency flies have some regions on the eye where the ommatidia look normal. The regular hexagonal array and the bristle number and placement are also wild type in places. However, there are patches of the eye where the shape of an ommatidium is pentagonal, more rounded, or sometimes nearly square, with some bristles missing or misplaced. Some bristles are duplicated, and a few even triplicated.

Some apparently emerge from the centre of the ommatidium rather than at a corner where three ommatidia meet (not shown). In other cases the bristle is placed at the interface between two ommatidia rather than the corner (not shown). A small number of fused ommatidia are also visible. Control flies do not show these defects.

Because of the female sterility of the *Df(3L)KG06133-15* chromosome in combination with the *TM3* and *TM6* balancers, it was kept balanced with a second deficiency; *Df(3L)FX3/TM3SbGFP*. The second site lethal on the *6133-15* deficiency is also lethal with the *FX3* deficiency so that it was possible to keep this stock reasonably well balanced, though still with the necessity of selecting the *6133-15* males each generation (as they were far less numerous than their *FX3* siblings). For examination of wing defects, eye and bristles phenotypes, the *FX3* siblings from the same vials were used as controls.

Figure 4.8. Scanning EM showing thoracic defects of the *Df(3L)6133-15* lesion in combination with two different balancer chromosomes.

A. P-GFP/TM6; **B**. *Df(3L)6133-15/TM6*; **C**. *Df(3L)FX3/TM3SbGFP*; **D**. *Df(3L)6133-15/TM3SbGFP*. All flies used for EM were females. Arrows indicate missing bristles, while arrowheads indicate duplicated bristles.

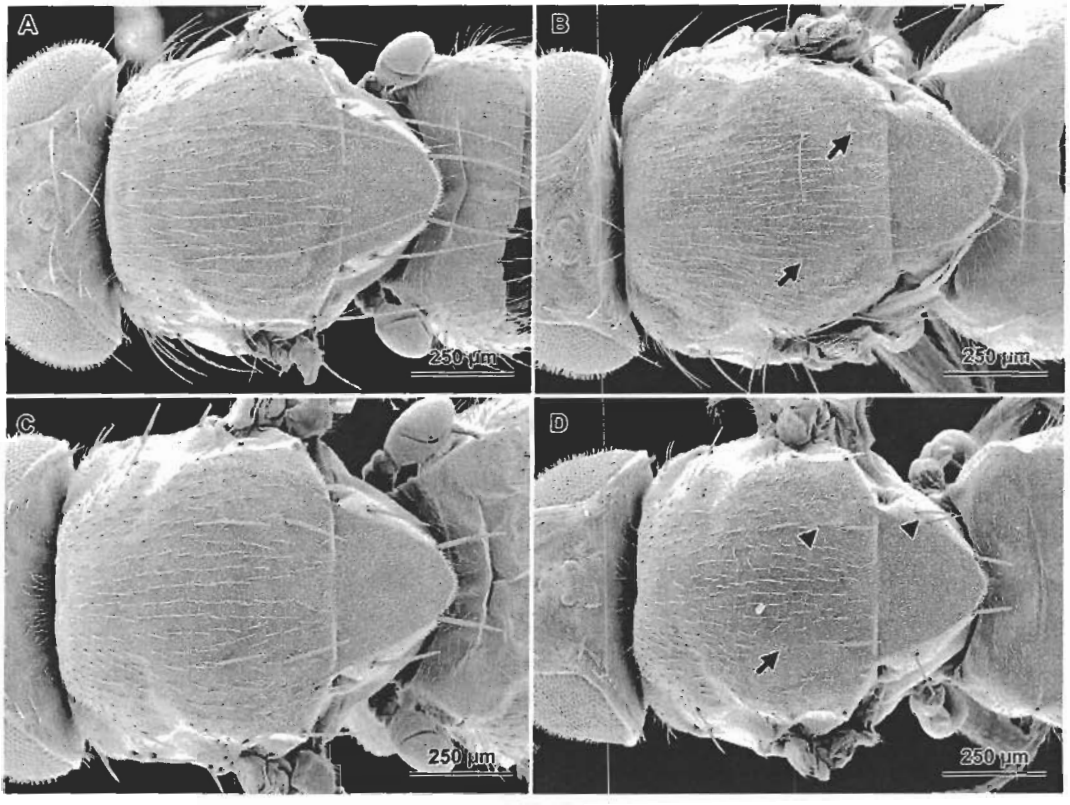


Figure 4.9. Wing defects observed in *Df(3L)6133-15/TM6* flies.

A-D: females**; E-H**; males. **A-C**; **E-G**: *Df(3L)6133-15/TM6*. **D**, **H**: *KG06133/TM6* as a control. *KG06133* is the chromosome from which *Df(3L)6133-15* was derived. Arrows in **A** and **G** indicate blisters in the wings, the arrowheads in **B** and **F**, notches and the arrows in **C** and **F**, an ectopic anterior crossvein. None of these defects were observed in the control flies.

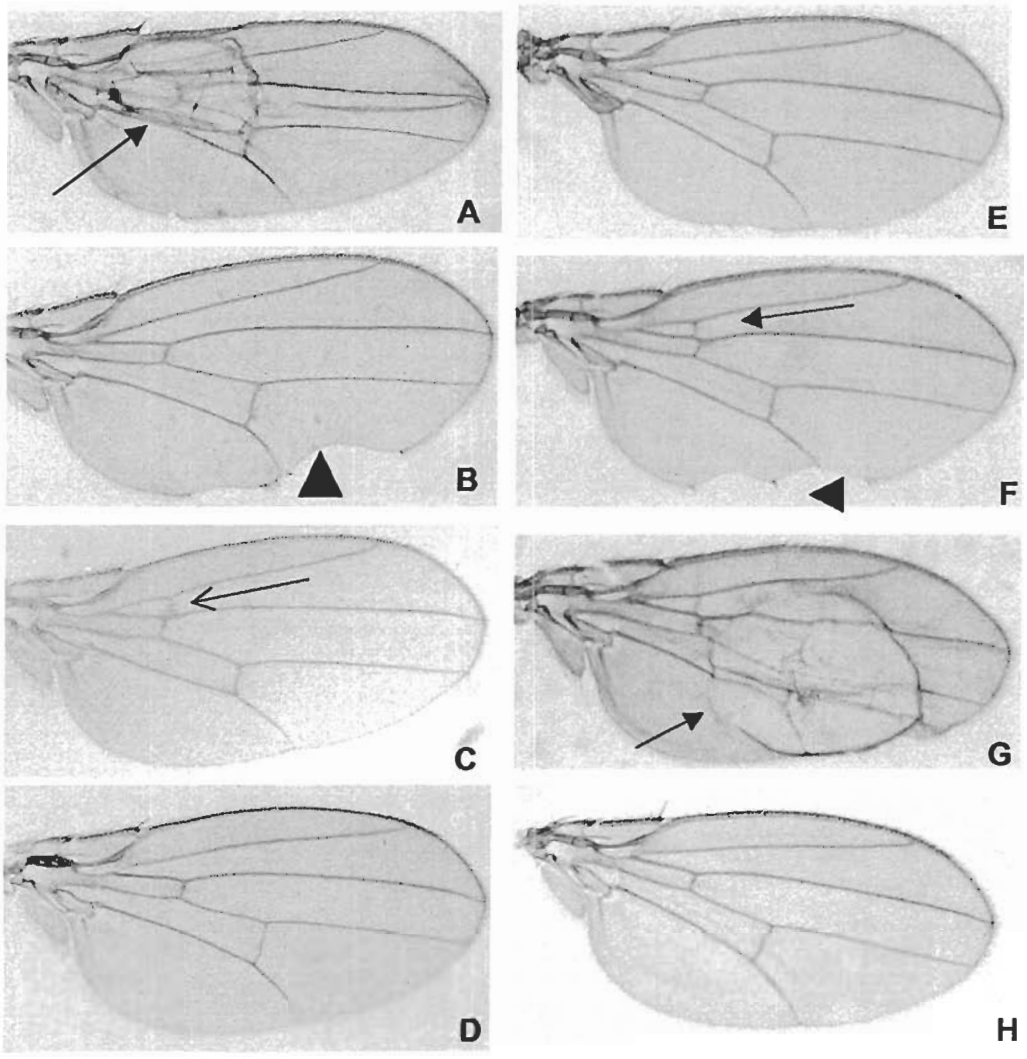


Figure 4.10. Wing phenotypes observed in *Df(3L)6133-15/TM3SbGFP.*

3 wings shown for each sex, to show variation in phenotype. **A-C**, females; **D-F**, males. Defects in wing venation are indicated by arrows. Degree of wing notching was also highly variable, as shown.

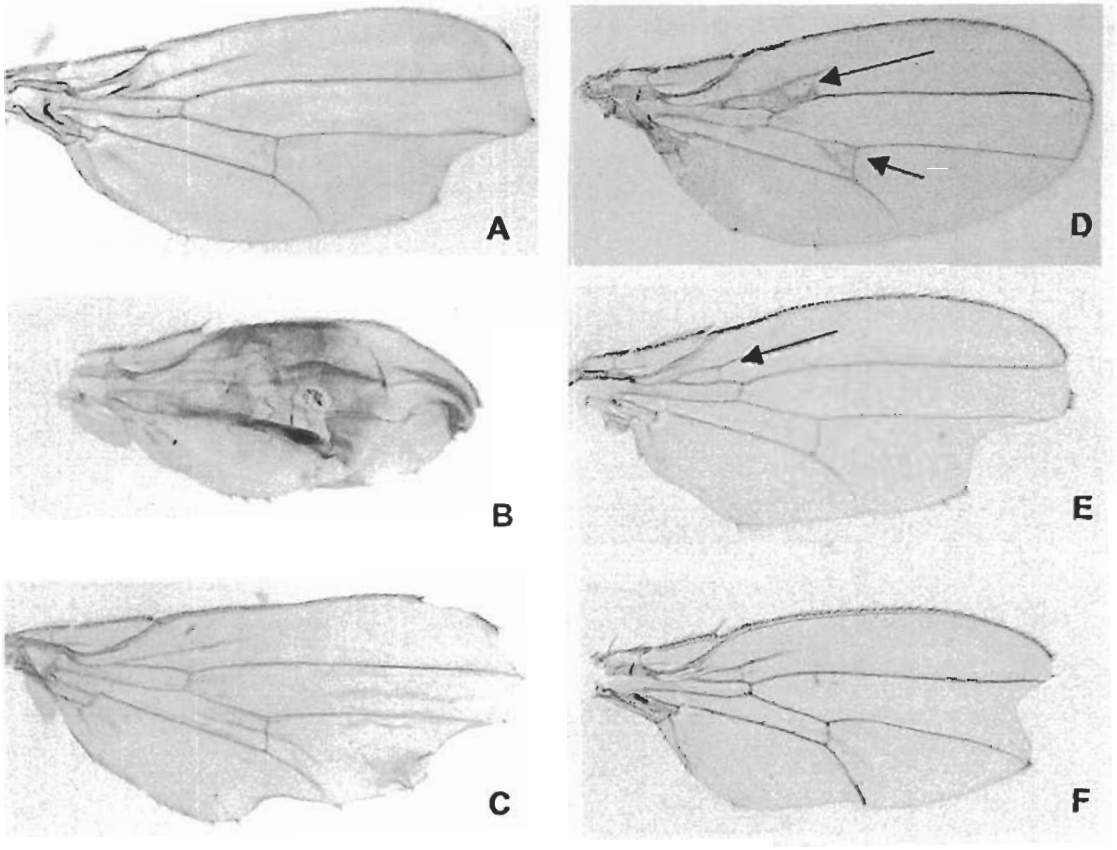


Figure 4.11. Scanning EM of eye phenotypes of *Df(3L)6133-15* in combination with the*TM6* or *TM3SbGFP* balancer chromosomes.

A, **B**: *P-GFP/TM6* control. **C**, **D**: *Df(3L)6133-15/TM6*; in **D**, the asterisk indicates a pentagonal ommatidium, the arrow a missing bristle and the arrow head a duplicated bristle. A slight disarray of the ommatidia and bristles is visible throughout. **E**, **F**: *Df(3L)FX3/TM3SbGFP*. **G, H**; *Df(3L)6133-15/TM3SbGFP*; arrowheads in **H** point to duplicated bristles, and the arrow indicates a fused ommatidium. Most ommatidia in this image show altered shapes and bristle pattern disruptions.

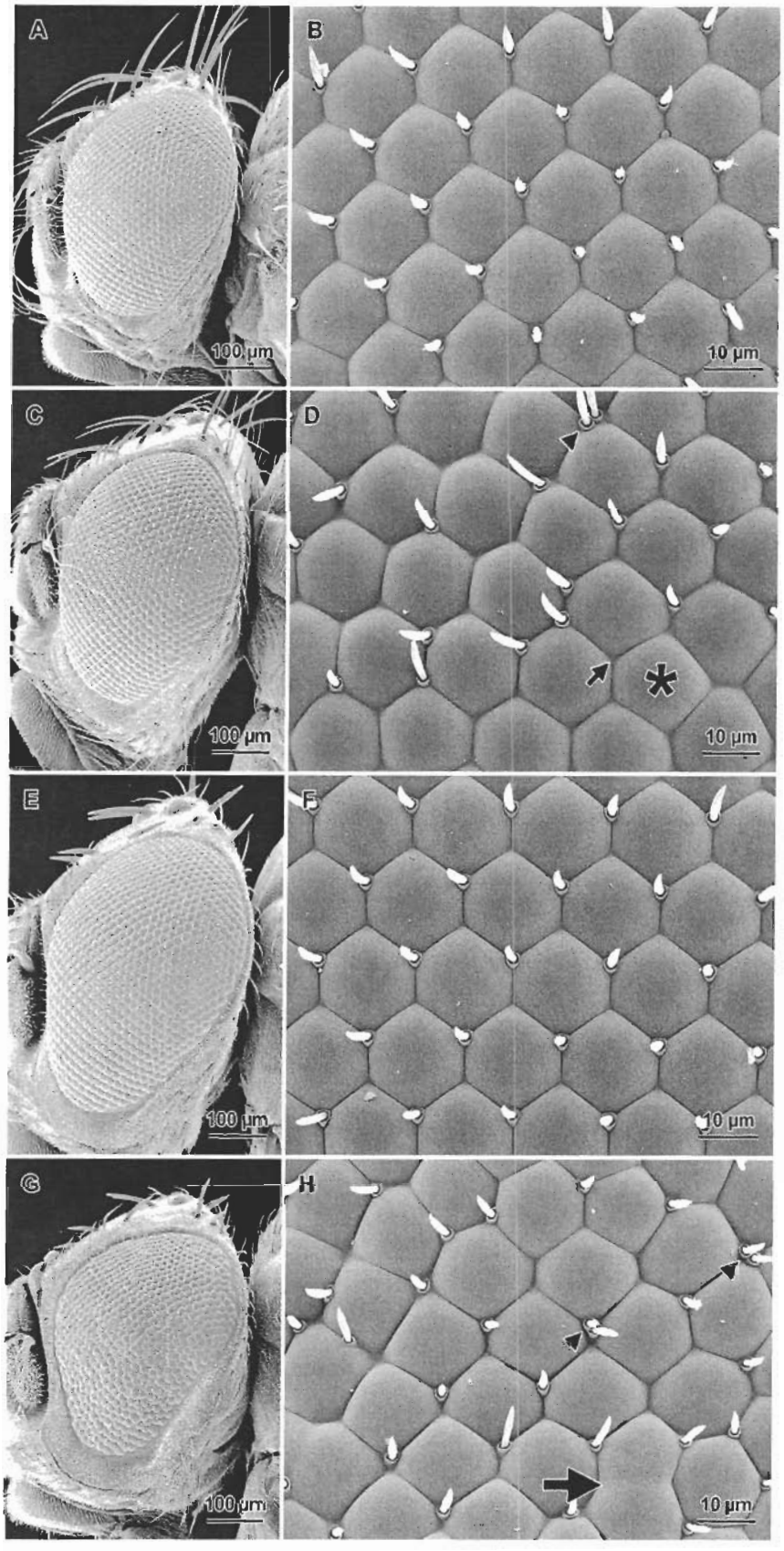

A
100 µm
B
10 µm
C
100 µm
D
10 µm
E
100 µm
F
10 µm
G
100 µm
H
10 µm

Table 4.11. Quantification of phenotypes of *Df(3L)6133-15* in combination with *TM6*, or *TM3SbGFP* balancers.

A. Frequency of notches on wings, blisters and extra anterior crossveins (ACV; see Figure 4.9. and Figure 4.10). **B.** Frequency of anterior scutellar (ASC) bristles either missing, duplicated, or triplicated on thorax. *Df(3L)FX3/TM3SbGFP* control flies are siblings from the *TM3SbGFP* stock. *KG06133/TM6* is the original line from which *6133-15* was derived, in combination with the *TM6* balancer as a control. "Total ASC", column 4, refers to the total number of anterior scutellar bristles expected, 2 per fly.

A.

genotype	sex	# wings	# notched	%	# blistered	%	# extra ACV	%
6133-15/TM3SbGFP	F	58	42	72.4	4	6.9	12	20.7
Df(3L)FX3/TM3SbGFP	F	50	4	8.0	0	0.0	0	0.0
6133-15/TM3SbGFP	M	78	40	51.3	5	6.4	21	26.9
Df(3L)FX3/TM3SbGFP	M	50	2	4.0	0	0.0	0	0.0
6133-15/TM6	F	76	35	46.1	21	27.6	5	6.6
KG06133/TM6	F	50	0	0.0	0	0.0	0	0.0
6133-15/TM6	M	60	25	41.7	11	18.3	5	8.3
KG06133/TM6	M	50	0	0.0	0	0.0	0	0.0

B.

genotype	sex	# flies	total ASC	# dup ASC	%	#trp ASC	%	# missing ASC	%
6133-15/TM3SbGFP	F	31	62	37	59.7	11	17.7	0	0.0
Df(3L)FX3/TM3SbGFP	F	25	50	0	0.0	0	0.0	0	0.0
6133-15/TM3SbGFP	M	40	80	23	28.8	1	1.3	0	0.0
Df(3L)FX3/TM3SbGFP	M	25	50	0	0.0	0	0.0	0	0.0
6133-15/TM6	F	34	68	6	8.8	0	0.0	8	11.8
KG06133/TM6	F	25	50	0	0.0	0	0.0	0	0.0
6133-15/TM6	M	24	48	7	14.6	0	0.0	11	22.9
KG06133/TM6	M	24	48	0	0.0	0	0.0	0	0.0

Df(3L)KG06133-15 frequently showed a number of bristle duplications on the thorax in combination with this balancer, an effect that was never observed in the control siblings (Table 4.8; Figure 4.8). Again, these were not observed in every fly but were somewhat variable.

The variety of defects observed have been seen in some other *Minute* mutations, and may stem from a global reduction in protein synthesis due to a lack of sufficient numbers of ribosomes. Certain dose-sensitive proteins such as Notch might be compromised as a result. The phenotype includes both smaller, thinner bristles, which may directly stem from reduced protein synthesis, and duplicated bristles, which could perhaps result from a compromise in the lateral inhibition function of *Notch*. *Notch* signaling is required in developing epithelial cells to inhibit neural (i.e. bristle) fates in the cells surrounding a presumptive bristle. Compromising *Notch* dose by reducing the levels of protein synthesis could perhaps lead to occasional extra bristles where this inhibition has failed. Similarly, *Notch* expression when compromised, can lead to wing notching due to failure of the presumptive wing margin to be properly specified in the developing imaginal discs.

In combination with *TM6* balancer, the *6133-15* deficiency shows a strong and invariant *Minute* phenotype, with some variable duplications of some bristles and

variable absences of others (Table 4.11). The control flies for this experiment consisted of the *PGFP* mutant in combination with *TM6*.

Attempts to characterize *6133-15* DNA have been hampered by the fact that the P-element appears to have excised from the region. PCR with flanking primers and P-out have been performed and these show that neither the 3' nor 5' end is intact in this line (Table 4.10). iPCR could be used to molecularly define the breakpoints, which will strengthen the argument that a multi-locus deficiency has been generated, and would provide better information on which loci have been deleted. However, this is only possible when the P-element is still intact.

Male recombination with the Rox80C line

Because *Df(3L)KG06133-15* was difficult to work with, and it appeared that any large deletion in this interval which affected *Qm* was likely to be female sterile, it was of interest to try to generate some smaller lesions within distal h47 that would be lethal in combination with the *Df(3L)KG06133-15* lesion without causing *Minute* or female sterile defects. Thus it would be possible to begin identifying and characterizing individual loci within distal h47. The *Rox80C* line, which has been described previously, was used for this experiment. The mobilization of the element was performed as described in the methods section. Because the parent line was viable in combination with the *Df(3L)KG06133-15*, lethality in

combination with *Df(3L)KG06133-15* could be used as an assay for deletions or local hops into essential loci in the interval.

A summary of the male recombination experiment is shown in Table 4.8. Of the 8 recombinant lines generated in the experiment, only two were lethal when transheterozygous with *Df(3L)KG06133-15.* As both were derived from the same original bottle of parents, it is likely that they represent the same event. It is noteworthy that despite the presence of the *Rox1* transgene on the P-construct, the recombination frequency observed did not differ greatly from any of the other experiments performed using P elements in this region. Thus the expectation that the frequency would be higher and thus should yield more lesions turned out to be incorrect. On the other hand, this element was in the midst of a repeat region, which might have mitigated the effect of the *Rox1* transgene. More experiments using a similar element in a variety of locations, would be required to draw significant conclusions about this question.

The lethal recombinant line was named *Rox-2* and is an ri^{+} p^{p} recombinant. The *Rox-2* line was crossed to every lesion known in the region, including the *TikR* alleles. The line was lethal only in combination with the *6133-15*, and it was viable in combination with *TikR*, *G18*, and *PL-00750.* The two *Minute* lines *KG01176-A1* and *KG03264-L* are very debilitated and so the line has not yet been crossed to these. *Rox-2* was characterized by single embryo PCR. Homozygous embryos were obtained and PCR performed with specific primers

to the following loci: *Qm*, *CG32350*, *α-Cat*, *CkII- α*. The results are shown in a later section (Figure 4.15). It appears that a single locus has been deleted by this lesion: *CG32350*. This locus is interesting as it shows homology with the *vps11p* locus in *S. cerevisiae*. *vps* genes are vesicular protein sorting genes and are involved with the secretory traffic and vacuole biogenesis. The genes are divided into 6 classes (A-F) based on the phenotypes they produce when mutant. 85 mutations have been isolated in *Drosophila* that affect eye colour; a subset of these is thought to be involved in pigment granule biosynthesis (called the granule group; Lloyd et al. 1998). Pigment granules appear to be a special class of late endosomes, or lysosomes, so when endosome function is compromised by weak mutants in such loci, pigment may not be properly delivered to the granules (Lloyd et al. 1998). *vps11p* belongs to the C class of vps proteins, as do a number of other *Drosophila* eye colour genes. These include *light* (*vps41p*; Warner et al. 1998),*deep orange* (*vps18p,* Shestopal et al. 1997, Mullins and Bonifacino 2001) and *carnation* (vps33, Mullins and Bonifacino 2001). All of these are involved in late endosome functions (Piper and Luzio 2001).

The subsequent analysis of this allele is complicated by the fact that the recombinant now bears the mutation p^{p} , which is an eye colour mutation that encodes a transport molecule. Since *vps11* is a vesicular protein sorting homologue, it would be important to determine whether mutant *vps11* alleles caused a visible eye phenotype similar to other eye colour mutations, which have been subsequently found to encode vps homologues. As the deletion is lethal in

combination with the *KG06133-15* deletion, it appears that the locus is essential. Weaker alleles might display eye colour phenotypes. Alternatively one could test for dominant interactions with other eye colour loci as have been demonstrated with other such loci. Dominant lethal, semi-lethal, eye colour and sterile interactions have been observed in various combinations of so-called granule group genes (Lloyd et al. 1998). This can be done using the *ri* p^p stock as a control to ensure that observed effects are due to *vps11* and not p^p (*pink* is another granule group gene; Lloyd et al. 1998).

Importantly, this work provides the first evidence that *CG32350* encodes an essential gene. Further work is now possible to characterize the gene more fully; this would include using X-ray mutagenesis to isolate lesions which are lethal with the *Rox-2* deletion, the *Tik* alleles and possibly the *PL00075* insertion line also. A subsequent, thorough EMS mutagenesis of the region may yield numerous new lethal, visible and sterile mutations of the loci in h47 to facilitate a more complete functional annotation of this previously uncharacterized region.

In the interval since the PCRs were performed, the gene order reported for loci in h47 has been altered (as has been mentioned previously). As discussed previously with complementation tests involving the *KG06133* male recombination lines, some tests did not appear necessary based on the published gene order at the time. However, since the order of the genes in distal h47 has recently been revised, further complementation tests will be necessary

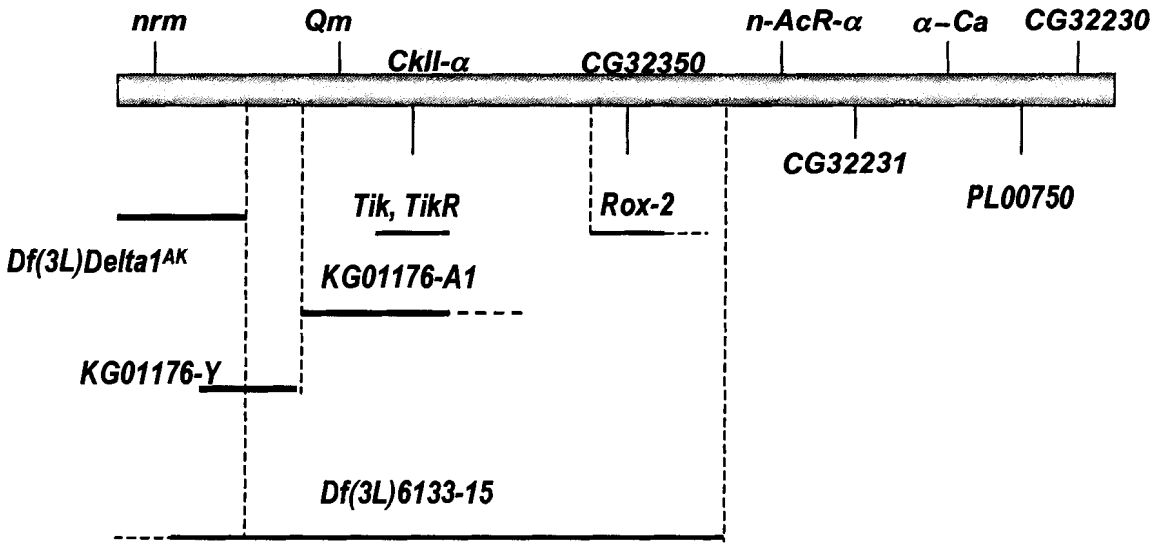
for complete characterization of the lines generated. Similarly, the PCR results which indicated that *α-Cat* was intact in the *Rox-2* homozygous deficiency embryos (Figure 4.15) made it appear unnecessary to test for the presence of *n-AChR-α*, in these embryos, since this gene was located proximal to *α-Cat*. However, given the revised gene order, one should retest the *Rox-2* homozygous embryos for the presence of the *n-AChR-α* locus. The lesion could extend farther than is currently thought.

Figure 4.12 shows the best current genetic map for the h47 interval. It takes into account all of the complementation and PCR data generated in this work, combined with the recent (Nov. 2004) re-annotation of the interval in which the gene order of several loci has been altered. This alteration underlines the extreme difficulty of working with these genomic intervals, which are full of repeats and very complex to overlap and align.

Molecular characterization of newly induced and previously existing lesions.

It was of interest to attempt to determine which coding sequences were absent under particular deficiencies. This would be the first step in identifying loci removed by the deficiencies and pinpointing candidate coding sequences for particular mutant loci.

Figure 4.12. Proposed genetic map of distal h47, including all known lesions in the region thus far.
Df(3L)Delta1^{AK} is a pre-existing lesion in proximal 3L euchromatin. *Tik* and *TikR* were generated by Lin et al. (2002), and *PL00750* by Hacker et al. (2003). The remainder were generated in this work.



Initial work was based on the assumption that the four loci identified (at that time) in h47: *Qm, CkII-α, α-cat* and *n-AChR-α* were potential candidates for some of our most distal heterochromatic loci. Therefore, single embryo PCR was performed with DNA from single homozygous deficiency embryos for a variety of genotypes, most importantly, *Df(3L)Delta1*^{AK} and *Df(3L)FX3,* and with a series of primers designed against the loci known at the time: *Qm, CkIIα, α-cat, n-AChR-α.* Once the protocol was optimized, it became apparent that all 4 loci were present in all the deficiency embryos tested. Thus, I concluded that a substantial number of interesting loci were located in segments outside of the regions delineated by the deficiencies. Therefore, these segments have never been subjected to mutagenesis screens. During the time that lesions were being isolated within the distal h47 region, a concurrent study was undertaken to try and determine candidate genes for 3L heterochromatin from among the many unlocalized scaffolds in the database. A number were selected for further study, in some cases because of known phenotypes of mutant loci.

All of the loci chosen, the scaffold on which they reside, homology information where known, and the gene(s) to which it was thought they might correlate, are listed in Table 4.12. Appendix B lists the primers, their sequences and the Tms used. Some of the data were clear and repeatable whereas the results for some genes were highly variable and thus suspect. Therefore, I will present data for which repeatable results have been obtained.

Table 4.12. List of genes tested with single embryo PCR, using deficiency homozygote embryos from various 3L heterochromatin deletions.

Putative functions listed based on flybase annotations.

gene name	CG #	putative function/domains	scaffold	reason for selection
Qm	CG17521	ribosomal protein	AE002898	h47 (Koryakov et al. 2002)
CkII-a	CG17520	casein kinase subunit	AE002898	h47 (Koryakov et al. 2002)
CG32350	CG32350	vps 11	AE002656	h47 (Koryakov et al. 2002)
n-AChR-a	CG12414	acetyl choline receptor subunit	AE002656	h47 (Koryakov et al. 2002)
	CG17552			
CG32231	CG32231	unknown	AE002656	h47 (Koryakov et al. 2002)
α-Catenin	CG17947	α-Catenin/cytoskeleton	AE002656	h47 (Koryakov et al. 2002)
CG32230	CG32230	NADH dehydrogenase	AE002656	h47 (Koryakov et al. 2002)
CG17920		not real gene/ unique seq tag	AE003139	Same scaffold as CG17454
CG17454		not real gene,/seq tag	AE003139	Splicing factor in 3het: flybase
Camlik	CG17698	calmodulin related kinase	3L_wgs_centromere extension	lethal 4A?
CG40300	CG40300	piwi/PAZ domian	3L_wgs_centromere extension	same scaffold as Camlik
scarecrow	CG17595	homeodomain transcription factor	AABU01002775	lethal 5?
	CG18452			
CG17374	CG17374	fatty acyl synthase	AABU01002764	lethal 4A flanking sequence
CG40351	CG40351	methyltransferase, RNA binding	AABU01002694	lethal 3?
CG41099	CG41099	Zn finger/Ankyrin domain	AABU01002742	lethal 4A?
CG12460	CG12460	spliceosome factor	AABU01002755	Snap-25 scaffold
CG40050	CG40050	rieske domain	AABU01002755	Snap-25 scaffold
CG40470	CG40470	unknown	AABU01002755	Snap-25 scaffold
CG40045	CG40045	ubiquitin conjugating enzyme	AABU01002755	Snap-25 scaffold
CG15831		ATPase, myosin related	AABU01002710	lethal 4A?
CG17514	CG17514	transcriptional regulator	AABU01002710	lethal 3?
eIF4B	CG10837	translational initiation factor	AABU01002740	3 het, flybase

Single embryo PCR results

Of the many gene models tested, only a small number yielded repeatable results. Thus only these few will be presented here. The others remain as possible loci resident in 3L or proximal 3R heterochromatin.

Genomic region near lethal 7A.

CG17698 and *CG40300* both appear to map to the genomic interval occupied by *lethal 7A.* For both genes, products failed to amplify from *Df(3L)O-1*, but were generated from *Df(3L)γ-28*, and *Df(3L)FX33.* The PCR results and a proposed map for these two loci are shown in Figure 4.13. *CG17689* encodes a putative calcium and calmodulin related protein kinase and is expressed in numerous tissues (Drysdale et al. 2005). *CG40300* resides on the same genomic scaffold as *CG17698* (AABU01002770) and encodes a protein of unknown function, which contains 2 conserved piwi domains. Either of these genes might be correlated with the *lethal7A* locus. Sequencing the one EMS allele available would be needed to determine whether either *CG17698* or *CG40300* are encoded by *lethal7A.* As fsa^2, the *lethal7A* allele, is lethal during late embryogenesis, with no striking cuticular defects, the phenotype gives no clues regarding which gene is likely to be involved.

CG17920, *CG17454*, originally annotated as genes on genomic scaffold AE003139, have been removed in more recent releases (Drysdale et al. 2005). Both mapped to this interval as well (data not shown) and can thus serve as unique sequence tags for this genomic region also.

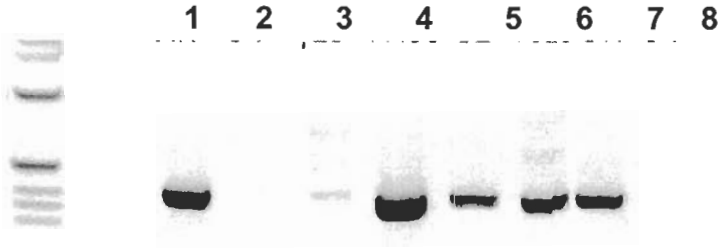
scarecrow

The *scarecrow (scro; CG17594)* gene encodes a homeobox type transcription factor that is expressed in a subset of the nervous system (Zaffran et al. 2000). PCR data indicate that this gene is absent from *Df(3L)O-1* and *Df(3L)γ-28*. The only genetically defined locus deleted by both of these is *lethal 5*. The PCR results and a map are shown in Figure 4.14. *scro* seems an unlikely candidate for the *lethal 5* locus, as its expression is confined to the nervous system, and a small number of surviving transheterozygotes for alleles of *lethal 5* show defects including sex-comb gaps, thickened aristae and abdominal tergite defects. It is not clear how disruptions in a gene expressed only in neurons could cause these diverse defects. However, the X-ray generated lesion, *l(3L)3264-C* seems to be localized to the same interval and complements all known *lethal 5* alleles, and so could define a novel essential gene in the vicinity of *lethal 5*. Preliminary PCR results (shown in Figure 4.14C) indicate that embryos homozygous for *l(3L)3264-C* fail to yield a product for *scro*, supporting the suggestion that *scro* may be another essential gene in the region of *lethal 5*. Recently, several P-elements inserted near *scro* became available (Konev et al. 2003). These help to confirm

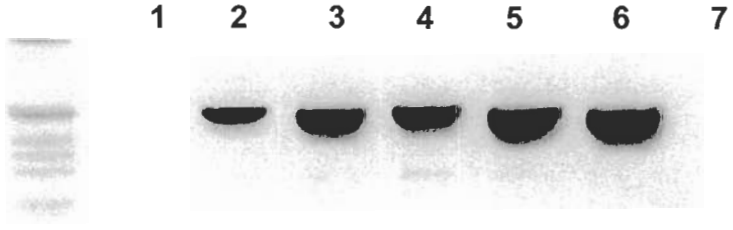
Figure 4.13. Preliminary PCR data for (**A**) *CG17698* and (**B**) *CG40300.*

A. Lane 1, *Df(3L)Delta1*^{AK}; Lane 2, *Df(3L)O-1*; Lane 3, *Df(3L)γ-28*; Lane 4, *Df(3L)FX33*; Lane 5, *Df(3L)C-2*; Lane 6, *Df(3L) vtd*^{2}; Lane 7, *Oregon R*; Lane 8, no template control. **B.** Lane 1, *Df(3L)O-1*; Lane 2, *Df(3L)γ-28*; Lane 3, *Df(3L)FX33*; Lane 4, *Df(3L)C-2*; Lane 5, Df(3L) *vtd*^{2}; Lane 6, *Oregon R*; Lane 7, no template control. **C** shows a map of the deficiencies used and the predicted position of the two genes, as indicated by the asterisk and arrow. Gene numbers and names are shown along the top of the figure.

A.



B.



C.

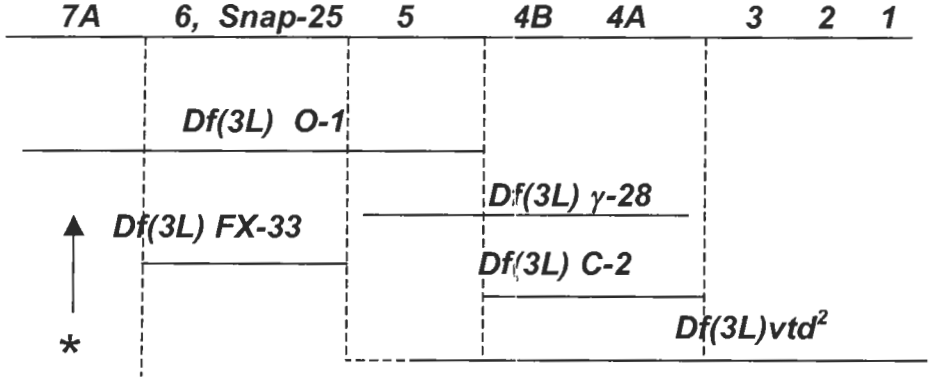
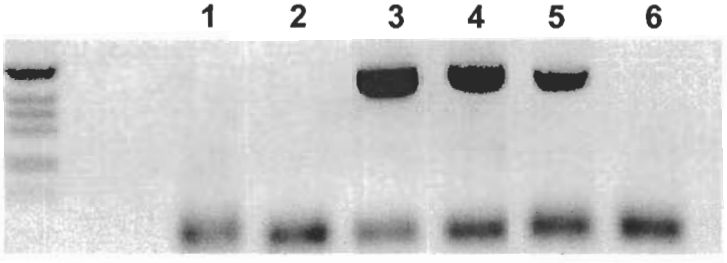
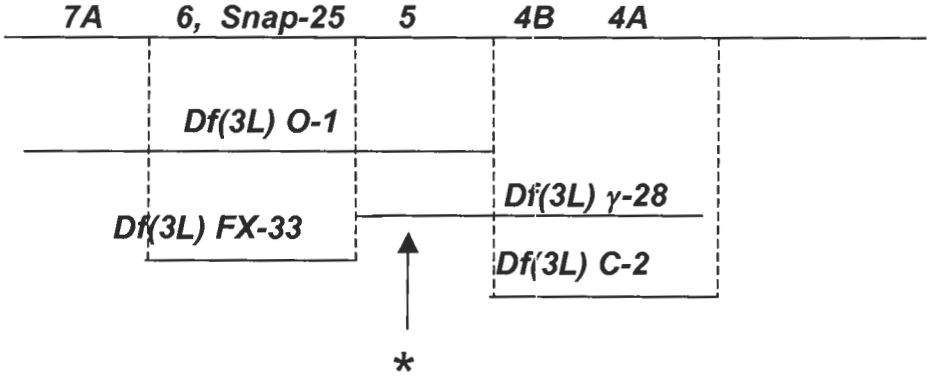


Figure 4.14. Preliminary PCR data for the *scarecrow* gene.
A. Lane 1, *Df(3L)O-1*; Lane 2, *Df(3L)γ-28*; Lane 3, *Df(3L)FX33*; Lane 4, *Df(3L)C-2*; Lane 5, *Oregon R*; Lane 6, no template control. **B**. Genetic map; the asterisk indicates the presumed location of *scarecrow*. **C**. Preliminary PCR data results from two homozygous *l(3L)3264-C* embryos. Lanes 1 and 2, *scro* primers; Lane 3 Oregon R control with *scro* primers; Lanes 5 and 6, the same two *l(3L)3264-C* embryos using primers to *CG17514* as a control.

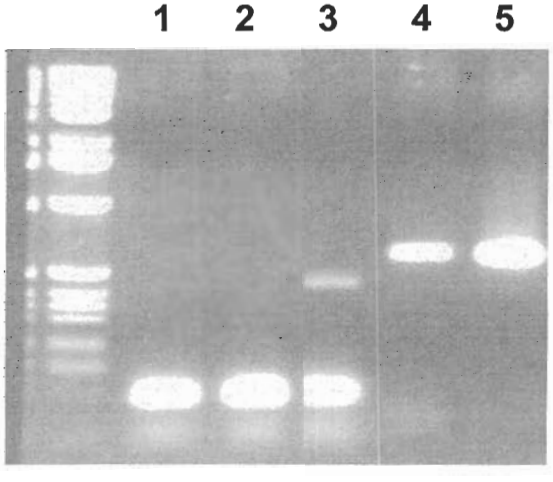
A.



B.



C.



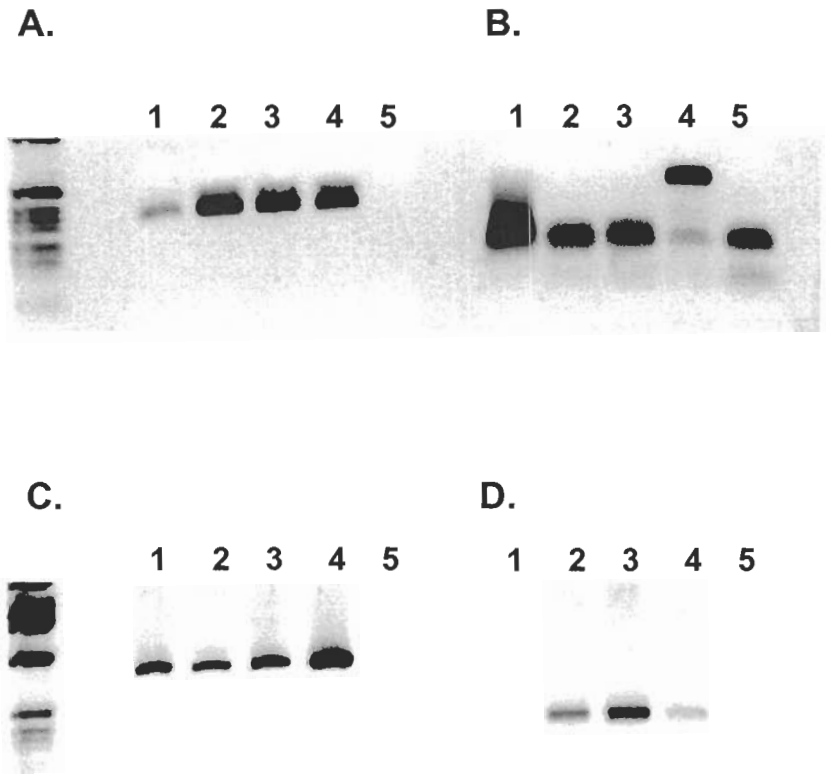
the placement of the *scro* scaffold near *lethal 5*, as Konev et al. have mapped their inserts to h47-50. These can now be used to characterize the region near *lethal 5* in more detail. As most of the inserts are homozygous viable, lethality in combination with *Df(3L)-O-1* and *Df(3L)γ-28* can be used to assay for the production of lethal derivatives.

Distal h47

The *Rox-2* putative deficiency line, lethal in combination with *Df(3L)6133-15,* was tested for the presence of *α-Cat*, *Qm*, *CkII-α* and *CG32350*, in amplified genomic DNA of homozygous deficiency embryos. Three such embryos were tested, and the results, shown in Figure 4.15, indicate that the only gene missing from this lesion is *CG32350*. As described earlier, further characterization is needed, as there is a possibility that the lesion could extend further proximally, and might delete *n-AcR-α- 80B* and *CG32231* as well. Deletions of either of these two loci would be of value, as neither has been extensively studied and no mutations have been isolated for either locus. *CG32231* may not be expressed, as no cDNAs have been reported in the data base (Drysdale et al. 2005), thus a deletion of this locus might allow us to determine whether or not it is an essential gene. *n-AcR-α- 80B* function has been studied using *in vitro* constructs (Lansdell and Millar 2000) but has never been shown to be an essential locus. A deletion of the locus would provide an opportunity for further studies of the gene including an assessment of the null phenotype.

Figure 4.15. PCR analysis of *Rox-2* homozygous embryos.

Lanes 1-3, *Rox-2/Rox-2*; Lane 4, *Oregon R*; Lane 5, no template control. Primers used: **A.** *Qm* **B.** *CG32350*; only lane 4 shows a band of the predicted size (upper band; the heavy lower bands are primer dimers); **C**. *CkII-α*; **D**. *α-Cat*. The PCR reactions with different primers serve as internal controls for these experiments.



Conclusions and Future Directions

The work presented here is a significant step toward a more complete characterization of a previously inaccessible region, the distal h47 region of 3L heterochromatin. Several lethal mutations have been collected from various sources, and five lesions have been generated in this work (see Figure 4.12). These mutations can now be used as a genetic resource to initiate a more comprehensive functional annotation of the interval. Numerous approaches have been used to attempt to characterize the region, and it has been shown that *Qm* is most likely the haplo-insufficient locus responsible for the paucity of deletions within this interval. Lesions involving this locus are dominantly female sterile and severely *Minute*. The most useful lesions that I have generated are *Df(3L)KG06133-15* and the smaller *Df(3L)Rox-2.* When transheterozygous, these two deficiencies delete only one locus in common, *CG32350*, a gene with *vps11* homology that has not been studied previously. This work provides the first demonstration that *CG32350* encodes an essential gene, and opens up the possibilities for further characterization of this locus (see below). It should be possible to generate more deficiencies, using male recombination from the *Rox80C* insertion as a basis for screening on larger scale to try to generate more lesions extending proximally. Selection of only proximal lesions would avoid the difficulties associated with deletions of the *Qm* locus, nearby. It should be possible to screen for lesions that extend as far as *lethal 8,* allowing us to "close the gap" across the heterochromatin/ euchromatin boundary and to identify and

characterize all essential genes within this region. It would also be possible to utilize further X-ray mutagenesis, using lethality in combination with *Rox-2* as the basis for screening.

Aside from the generation of new deficiencies in this region, the means now exist to begin to characterize the essential locus, *CG32350*. As part of such an analysis, *Df(3L)Rox-2* can be tested for dominant interactions with mutant alleles of a number of loci the gene products of which encode related functions, such as *light*, *carnation*, *deep-orange*, and others for which dominant interactions have been reported (Lloyd et al. 1998). New alleles can be generated via EMS mutagenesis, using *Df(3L)Rox-2* for screening. Alternatively, new lesions that extend further proximally might be used also. Perhaps the greatest strength of this work is that the lesions generated here provide a basis upon which many subsequent projects can be built, all leading to the ultimate goal of a full functional annotation of all of the heterochromatic genes on chromosome 3.

The use of single embryo PCR has led to the localization of some genomic scaffolds into particular deficiency intervals across the 3L heterochromatin. The use of this technique can complement what is known genetically, and provide additional information regarding the locations of certain genomic scaffolds. Specifically, the *CG17698* and *CG40300* loci, have been localized to a genetic interval occupied by the *lethal 7A* locus (Figure 4.13C). Either of these genes might be candidates for *lethal 7A*, and this can now be tested by sequencing the

lethal 7A allele (and any additional EMS alleles that are produced by ongoing EMS screens). In addition, the *scarecrow* locus has been localized to the genomic interval in which *lethal 5* resides (Figure 4.14). It seems unlikely that *scro* corresponds to *lethal 5,* mainly because its expression is confined to the nervous system; however preliminary data indicates that there may be at least one additional essential locus within this region, defined by *l(3L)3264-C*, which might be *scro*. Single embryo PCR and genomic Southern analysis with all γ-irradiation induced alleles of *lethal 5* could be used to confirm this. Several lines with P elements inserted near the *scro* locus are now available (Konev et al. 2003). Male recombination or imprecise excision experiments should allow the generation of lethal lesions in this locus. No mutant alleles of the *scro* locus have ever been isolated, but the localization of the gene in the region of overlap between *Df(3L)O-1* and *Df(3L)γ-28* now provides the means for generating and characterizing mutants of this potentially very interesting gene.

While single embryo PCR is a useful tool for identifying candidate sequences for deficiency intervals, it should be emphasized that not all mutations of a given gene are confined to protein coding regions. One illustration of this comes from the work of Robert et al. (2001), studying a gene, *flamenco*, in the X-chromosome heterochromatin. A P-element insertion line was generated that created a *flamenco* mutant; however, within 100 kb surrounding the site of insertion, no affected coding sequence was identified. Moreover, chromosome rearrangements that altered *flamenco* function were mapped to a position more

than 130 kb from the site of P-element insertion. The authors concluded that the P-element was acting at a significant distance to alter the expression of an as yet unidentified sequence. In addition, flanking genomic sequence of two P-element induced alleles of *lethal 4A* reside on genomic scaffold AABU01002764, at a position 40kb from the nearest gene, *CG17374* (Fitzpatrick, unpublished observation). If P-elements could disrupt expression of a gene from such a distance, so might deletions that remove even fairly distant regulatory sequences without altering the coding sequence of a locus. Thus, results from PCR analysis of deficiencies must be complemented with genetic analysis and the molecular characterization of mutant alleles of each locus. For this reason, the ongoing EMS screen will be of great value in providing mutant alleles 3L heterochromatic genes, for sequencing and further phenotypic analysis. This is of particular importance for the distal loci, *lethal 7A- lethal 8*, for which only one or two alleles exist at present.

The work described here provides a platform upon which a great deal of future work may be built, has generated some crucial tools to allow access to previously inaccessible regions of the heterochromatin, and contributes to the ongoing effort to fully characterize these interesting regions of the genome.

Appendix A. Genetic mutants and deficiencies used in this work.

Mutant name	mutagen	gene(s) affected	source/reference
CH(3)1	P-element	*Parp*	Tulin et al. 2002
CH(3)4	P-element	*lethal 6*; second site at Parp	Tulin et al. 2002
abs^{9}	P-element excision	*abstrakt*	Irion and Leptin 1999
nrm^{165}	EMS	*neuromusculin*	Kania and Bellen 1995
Tik	Spontaneous derivative of *Tik*	*Casein kinase II-α*	Lin et al. 2002
TikR	piggyBac element	*Casein kinase II-α*	Lin et al. 2002
PL-00750		unknown; between *α-Cat* and *CG32230*; lethal	Hacker et al. 2003
Df(3L)FX3	γ-irradiation	*lethal 3-8*	Vilinsky et al. 2002
Df(3L) γ -28	γ-irradiation	*lethal 4A-5*	Schulze et al. 2001
Df(3L)O-1		*lethal 5 – 7A*	Schulze et al. 2001
Df(3L)MX18		*lethal 1-7B*	Vilinsky et al. 2002
Df(3L)FX53	x-ray		Vilinsky et al. 2002
In (3L)C90	γ-irradiation	lethal 1	Roberts 1970
Df(3L)2-30	*EMS*	lethal 1	Marchant and Holm, 1988a
72	*EMS*	*lethal 2*	Schulze et al. 2001
1-166-37	EMS	*lethal 2*	Schulze et al. 2001
1-166-38	P-element	*lethal 3*	Schulze et al. 2001
3-1^{29}	EMS	*lethal 3*	Schulze et al. 2001
vtd^{4}	P-element	*lethal 3*	Schulze et al. 2001
CH(3)53	P-element	*lethal 4A*	Zhang and Spradling 1994
CH(3)4d	EMS	*lethal 4A*	Zhang and Spradling 1994
G10	EMS	*lethal 4B*	Vilinsky et al. 2002
G30	EMS	*lethal 4B*	Vilinsky et al. 2002
G5	EMS	*lethal 5*	Vilinsky et al. 2002
G12	EMS	*lethal 5*	Vilinsky et al. 2002
1-16-I	γ-irradiation	*lethal 5*	Schulze et al. 2001
1-16-N	γ-irradiation	*lethal 5*	Schulze et al. 2001
CH(3)4	P-element	*lethal 6*	Zhang and Spradling 1994
G4	EMS	*lethal 6*	Vilinsky et al. 2002
G59	EMS	*lethal 6*	Vilinsky et al. 2002
CH(3)7	P-element	*Snap-25*	Zhang and Spradling 1994
1-16-O	γ-irradiation	*Snap-25*	Schulze et al. 2001
l(3L) fsa2	EMS	*lethal 7A*	Felsenfeld and Kennison 1995
G43	EMS	*lethal 7B*	Vilinsky et al. 2002
G1	EMS	*lethal 7B*	Vilinsky et al. 2002
G18	EMS	*lethal 8*	Vilinsky et al. 2002

Appendix B. Primers used, their sequences and Tms.

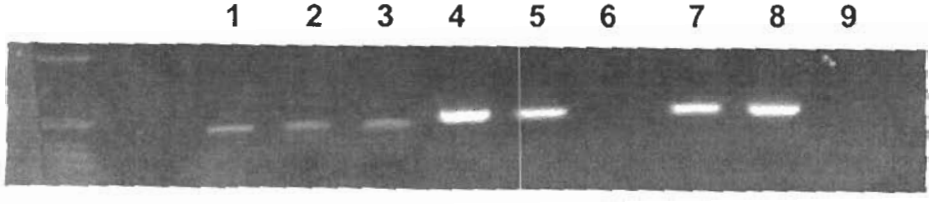
name	gene	sequence	tm
parp 2F	*Parp*	aaaaaccagcgtcccagctcagta	54°C
parp3R	*Parp*	catcatctgctaagctctgaaatc	
parp4F	*Parp*	gaagaactaccagatacaaaaaga	49°C
parp4R	*Parp*	catcattgtttttattcgttttca	
parp5F	*Parp*	tataggcttaagaaccagcataa	50°C
parp5R	*Parp*	aggagttcttatttctcttgaatgt	
Gfat7F	*Gfat1*	ctggtaggacgacgatgttgc	60°C
Gfat7R	*Gfat1*	ggtgtatgccttggtcgagg	
159F	*CG40159*	cgaatcctcaaacaatagtgcccaag	60°C
159R	*CG40159*	gagtgtggattatcaagtgggccactc	
451F	*CG40451*	agcttctattggacttatttcaacg	60°C
451R	*CG40451*	ttacaagacgtaccatttcgagc	
dskF	*Dsk*	ccatcatagcttactctaagaattcg	60°C
dskR	*Dsk*	tagatgttcaccagtctgccaagg	
Grp84-1	*Grp84*	acgcttctcgctgatggac	50-65°C
Grp84-2	*Grp84*	gtcgcagtaactggattgagt	
screxF	*scarecrow*	cctgcaagctcaccaaaggcgag	60°C
screxR	*scarecrow*	cacttacataaccgaaaggagtacacttagg	
CexF	*CG17698*	tcattgtttacttctagggatccta	60°C
CexR	*CG17698*	acaatattccaatccaagtaaagactc	
PexF	*CG40300*	tgacattcttccctgatatcctgag	60°C
PexR	*CG40300*	agcattgtccgggtacaagaagac	
920F	*CG17920*	tgagactgatgagcttatcatacc	60°C
920R	*CG17920*	attggtggttgtaactgctacggtc	
GFPF	*GFP*	caagagtgccatgcccgaag	55°C
GFPR	*GFP*	gacagggccatcgccaattg	
M13F	N/a	gtaaaacgacggccagtg	60°C
M13R	N/a	ggaaacagctaatgacccat	
Qm1	*Qm*	gtatatttgatttgggaagaa	55°C
Qm2	*Qm*	agtgcaaaagaggtcataaga	
acat1	*α-Cat*	aatttaatgaatgccgttgta	55°C
acat2	*α-Cat*	ttaaacagcgtcagcaggact	
ck2-1	*CkII-α*	atattgttaggaccaccaacc	57°C
ck2-2	*CkII-α*	catatatccagtgagtaatca	
SET1	*CG40350*	ttgacacagatgcaggacgttc	55°C
SET 2	*CG40350*	ctatacgtgcaatgcccaagtg	

Appendix B. Continued. Primer sequences

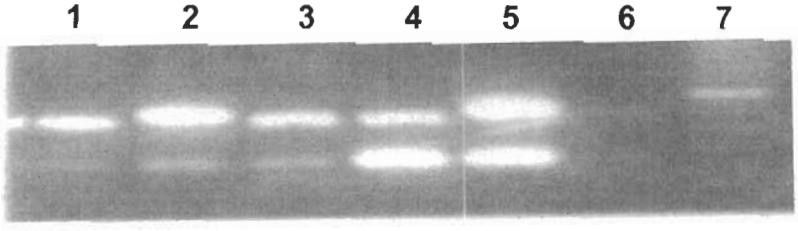
Name	Gene	Sequence	Tm
vpsF	*CG32350*	caaactctagagattagataaag	60°C
vpsR	*CG32350*	ctgcaatagtatctagagaatc	
KGN5-1	6133 flank	agtcacttcgaatggacgatcagc	60°C
KGN3-1	6133 flank	catatatggagaccggatattgtcc	
P-out 31	N/a	cgacgggaccaccttatgttatttcatcatg	60°C
EX1R	*CG12581* 5'flank	cagcatacgcacacacaaagcgctg	60°C
EP1	EP(3)3632 3'flank	gcagcaccaccaggaagc	60°C
EP4	EP(3)3632 KG01176	ggtgtcaccgtcacactg	60°C
11765-1	5' flank KG01176	ccttataatggcagcagccgac	60°C
11763-1	3' flank KG01176	gtaaacggtagctaattcgagcgg	60°C
11763-2	3' flank	tcgtcttctagcacaacacgcac	60°C
454F	*CG17454*	gtaaaatagtggaatctgacagtac	55°C
454R	*CG17454*	tctggatgsgtttgaagagctgc	
129F	*CG40129*	tcagctaactatgaacgatttcagc	55°C
129R	*CG40129*	caacgtttatgcatgtgttccaatc	

Appendix C: Selected positive controls for single embryo PCRs.

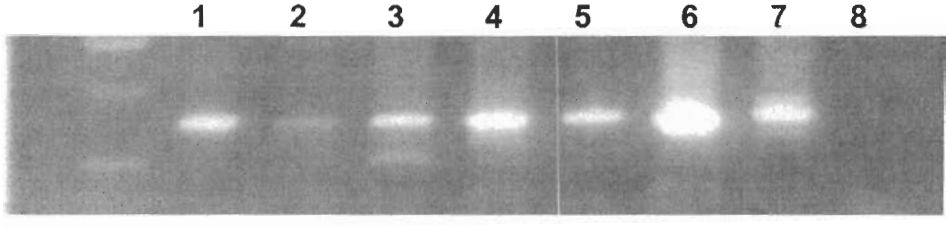
A. Control for 3R experiment. Primers: 129F and 129R. Primers were designed against *CG40129,* a heterochromatic locus on the second chromosome (Appendix B).
Lanes 1 and 2: *Df(3LR)6B-29*, Lane 3: *Df(3R)10-65*, Lanes 4 and 5: *Df(3R)EP-167*, Lanes 6 and 7: *Df(3R)XM3* (note that the lane 6 template failed, and thus this lane was removed from images of most of the 3R PCR experiments since a failure to generate a product with this template would have been meaningless), Lane 8: *Oregon R* control, Lane 9: no template negative control.



B. Control for 3L PCR experiments. Primers used: grp841 and 2, lower band; SET1 and 2, upper band. Lane 1: *Df(3L)O-1*, Lane 2: *Df(3L)FX33*, Lane 3: *Df(3L)C-2*, Lane 4: *Df(3L)γ-28*, Lane 5: *Df(3L)vtd*^{2}, Lane 6: no template control. Ladder in this gel (Lane 7) and the gel below (C), is the 100 bp ladder, rather than the 1 kb ladder in the others.



C. Control for 3L PCR experiments. Primers used: CkII-α 1 and 2. Lanes 3: *Df(3L)O-1*, Lane 4: *Df(3L)γ-28*, Lane 5: *Df(3L)C-2*; Lane 6: *Df(3L)FX33*, Lane 7: *Oregon R*, Lane 8: no template control. Lanes 1 and 2 show bands for embryos collected for use in other experiments, not reported in this work.



References

Abad, J.P., Carmena, M., Baars, S., Saunders, R.D.C., Glover, D.M., Ludena, P., Sentis, C., Tyler-Smith C., and Villasante, A. 1992. Dodeca Satellite: A conserved G+C-Rich satellite from the centromeric heterochromatin of *Drosophila melanogaster. Proc. Natl. Acad. Sci.* 89: 4663-4667.

Adams et al. 2000. The genome sequence of *Drosophila melanogaster. Science* 287: 2185-2195.

Ahmad, K., and Henikoff, S. 2001. Centromeres are specialized replication domains in heterochromatin. *J. Cell Biol.* 153: 101-110.

Appels R. and Hilliker A.J. 1982. The cytogenetic boundaries of the rDNA region within heterochromatin in the X chromosome of Drosophila melanogaster and their relation to male meiotic pairing sites. *Genet Res.* 39:149-156.

Ashburner, M. 1989. Drosophila: A laboratory manual. Cold Spring Harbour Laboratory Press.

Baldwin M.C. and Suzuki, D.T. 1971. A screening procedure for detection of putative deletions in proximal heterochromatin of Drosophila. *Mutat Res.* 11: 203-213.

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

Biggs, W.G.III., Zavitz, K.H., Dickson, B., var der Straten, A., Brunner, D., Hafen, E., and Zipursky, S.L. 1994. The *Drosophila rolled* locus encodes a MAP kinase required in the sevenless signal transduction pathway. *Embo J.* 13: 1628-1635.

Birchler, J.A., Kavi, H.H., and Fernandez. H.R. 2004. Heterochromatin: RNA Points the Way. *Current Biology* 14: R759–R761.

Blower, M.D., and Karpen, G.H. 2001. The role of Drosophila CID in kinetochore formation, cell-cycle progression and heterochromatin interactions. *Nat. Cell Biol.* 8:730-9.

Bronner, G. and Jackle, H. 1996. Regulation and function of the terminal gap gene huckebein in the Drosophila blastoderm. *Int. J. Dev. Biol.* 40: 157-165.

Carmena, M., Abad, J.P., Villasante, A., and Gonzalez, C. 1993. The *Drosophila melanogaster* dodecasatellite sequence is closely linked to the centromere and can form connections between sister chromatids during mitosis. *J Cell Sci.* 105: 41-50.

Carvalho, A.B. 2002. Origin and evolution of the Drosophila Y chromosome. *Curr. Opin. Genet. and Dev.* 12: 664-668

Carvalho, A.B., Dobo, B.A., Vibranovski, M.D., and Clark, A.G. 2001. Identification of five new genes on the Y chromosome of *Drosophila melanogaster. Proc. Natl. Acad. Sci. USA* 98: 13225-13230.

Carvalho, A.B., Vibranovski, M.D., Carlson, J.W., Celniker, S.E., Hoskins, R.A., Rubin, G.M., Sutton, G.G., Adams, M.D., Myers, E.W., and Clark, A.G. 2003. Y chromosome and other heterochromatic sequences of the Drosophila melanogaster genome: how far can we go? *Genetica* 117: 227-237.

Celniker, S.E., Wheeler, D.A., Kronmiller B., Carlson, J.W., Halpern, A., Patel, S., Adams, M., Champe, M., Dugan, S.P., Frise, E., et al. 2002. Finishing a whole-genome shotgun: release 3 of the *Drosophila melanogaster* euchromatic genome sequence. *Genome Biol*: 3:research 0079.1-79.14.

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

Clegg, N.J., Honda, B.M., Whitehead, I.P., Grigliatti, T.A., Wakimoto, B., Brock, H.W., Lloyd, V.K., and Sinclair, D.A.R. 1998. Suppressors of position-effect variegation in *Drosophila melanogaster* affect expression of the heterochromatic gene light in the absence of a chromosome rearrangement. *Genome* 41: 495-503.

Corradini, N., Rossi, F., Verni, F., and Dimitri, P. 2003. FISH analysis of Drosophila melanogaster heterochromatin using BACs and P elements. *Chromosoma* 112: 26-37.

Coulthard, A.B., Eberl, D.F., Sharp, C.B., and Hilliker, A.J. 2003. Genetic analysis of the second chromosome centromeric heterochromatin of *Drosophila melanogaster. Genome* 46: 343-352.

Craig, J.M. 2005. Heterochromatin – many flavours, common themes. *BioEssays 27: 17-28.*

Cryderman, D.E., Cuaycong, M.H., Elgin, S.C.R., and Wallrath, L.L. 1998. Characterization of sequences associated with position-effect variegation at pericentric sites in *Drosophila heterochromatin*. *Chromosoma* 107: 277-285.

Cryderman, D.E., Grade, S.K., Li, Y., Fanti, L., Pimpinelli, S., and Wallrath, L. 2005. Role of HP1 in euchromatic gene expression. *Developmental Dynamics* 232: 767-774.

Danzer J.R., and Wallrath L.L. 2004. Mechanisms of HP1-mediated gene silencing in Drosophila. *Development.* 131: 3571-80.

Dernburg A.F., and Karpen G.H. 2002. A chromosome RNAissance. *Cell* 111:159-162.

Devlin, R.H., Holm, D.G., Morin, K.R. and Honda, B.M. 1990a. Identifying a single copy DNA sequence associated with the expresision of a heterochromatic gene, the light locus of *Drosophila melanogaster*. *Genome* 33: 405-415.

Devlin, R.H., Bingham, B. and Wakimoto, B.T. 1990b. The organization and expression of the light gene, a heterochromatic gene of *Drosophila melanogaster*. *Genetics* 125: 129-150.

Dimitri, P., and Junakovic, N. 1999. Revising the selfish DNA hypothesis: new evidence on accumulation of transposable elements in heterochromatin. *Trends in Genetics* 15: 123-124.

Dimitri, P., Corradini, N., Rossi, F., Verni, F., Cenci, G., Belloni, G., Zhimulev, I.F., and Koryakov, D.E. 2003. Vital genes in the heterochromatin of chromosomes 2 and 3 of *Drosophila melanogaster*. *Genetica* 117: 209-215.

Dimitri, P., Junakovic, N., and Arcà, B. 2003. Colonization of Heterochromatic Genes by Transposable Elements in *Drosophila*. *Mol. Biol. Evol.* 20: 503-512.

Dimitri, P., Corradini, N., Rossi, F. and Verni, F. 2005. The paradox of functional heterochromatin. *BioEssays* 27: 29-41.

Drysdale, R.A., Crosby, M.A. and The FlyBase Consortium. 2005. FlyBase: genes and gene models. Nucleic Acids Research 33:D390-D395. http://flybase.org/

Duncan, I.W. 2002. Transvection effects in *Drosophila*. *A. Rev. Genet.* 36:521-556.

Eberl, D.F., Duyf, B.J. and Hilliker, A.J. 1993. The role of heterochromatin in the expression of a heterochromatic gene, the rolled locus of *Drosophila melanogaster*. *Genetics.* 134: 277-292.

Ehrman, J.M. and Kaczmarska, I. 2002. A 90-degree-tilt rotary adapter for SEM. *Microscopy Today* 2: 24-26.

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

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

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

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

Felsenfeld, A.L., and Kennison, J.A. 1995. Positional signalling by *hedgehog* in *Drosophila* imaginal disc development. *Development* 121:1-10.

Fitzpatrick, K.A., Sinclair, D.A., Schulze, S.R., Syrzycka, M., Honda, B.M. 2005. A genetic and molecular profile of chromosome 3 centric heterochromatin in *Drosophila melanogaster.* Genome: in press.

Gatti, M., Pimpinelli, S. 1983. Cytological and genetic analysis of the Y chromosome of Drosophila melanogaster. Chromosoma 88:349-373.

Gatti, M. and Pimpinelli, S. 1992. Functional elements in *Drosophila melanogaster* heterochromatin. *Annu Rev Genet.* 26: 239- 275.

Gatti, M., Bonaccorsi, S. and Pimpinelli S. 1994. Looking at Drosophila mitotic chromosomes. *Methods. Cell Biol.* 44: 371-391.

Gilfillan, G.D., Dahlsveen, I.K. and Becker PB. 2004.Lifting a chromosome: dosage compensation in Drosophila melanogaster. FEBS Lett. 567: 8-14.

Graack, H.R., Cinque, U., and Kress, H. 2001. Functional regulation of glutamine:fructose-6-phosphate aminotransferase 1 (GFAT1) of *Drosophila melanogaster* in a UDP-N-acetylglucosamine and cAMP-dependent manner. *Biochem J.* 360: 401-412.

Grewal S.I. and Elgin S.C. 2002. Heterochromatin: new possibilities for the inheritance of structure. *Curr Opin Genet Dev.* 12:178-87.

Gvozdev, V.A., Kogan, G.L., Tulin, A.A., Aravin, A.A., Naumova, N.M., Shevelyov,Y.Y. 2000. Paralogous stellate and Su(Ste) repeats: evolution and ability to silence a reporter gene. *Genetica.* 109:131-40

Hacker, U., Nystedt, S., Barmchi, M.P., Horn, C., and Wimmer, E.A. 2003. piggyBac-based insertional mutagenesis in the presence of stably integrated P elements in Drosophila. *Proc. Natl. Acad. Sci* 100: 7720-7725.

Hanai, A., Uchida, M., Kobayashi, S., Miwa, M. and Uchida, K. 1998. Genomic organization of *Drosophila* poly (ADP-ribose) polymerase and distribution of its mRNA during development. *J. Biol. Chem.* 273: 11881-11886.

Hannah, A. 1951. Localization and function of heterochromatin in *Drosophila melanogaster. Adv. Genet.* 4: 87-125.

Hatton, L.S., and O'Hare, K. 1999. Deficiency mapping of genes in Drosophila using single embryo PCR. *Tech. Tips Online* 1:TO1816.

Haupt W., Fischer T.C., Winderl, S., Fransz, P. and Torres-Ruiz, R.A. 2001. The centromere1 (CEN1) region of Arabidopsis thaliana: architecture and functional impact of chromatin. *Plant J.* 27: 285-296.

Hearn, M.G., Hedrick, A., Grigliatti, T.A. and Wakimoto, B.T. 1991. The effect of modifiers of position-effect variegation on the variegation of heterochromatic genes of Drosophila melanogaster. *Genetics.* 128:785-797.

Heitz, E. 1928. Das Heterochromatin der Moose *I. Hahrbeucher Wiss. Bot.* 69: 762-818.

Hennig, W. 1999. Heterochromatin. *Chromosoma* 108:1-9.

Henikoff, S. 1998. Conspiracy of silence among repeated transgenes. *BioEssays* 20:532-535.

Henikoff, S. 2000. Heterochromatin function in complex genomes. *Biochimica et Biophysica Acta* 1470:O1-O8.

Henikoff, S., Eissenberg, J.C., Hilliker, A.J., Schmidt, E.R., Wallrath, L.L. 2000. Reaching for new heitz. Genetica 109: 7-8.

Henikoff, S., Ahmad, K., and Malik, H.S. 2001. The centromere paradox: stable inheritance with rapidly evolving DNA. *Science* 293: 1098-1102.

Hessler, A.Y. 1958. V-type position effects at the light locus in *Drosophila melanogaster. Genetics,* 43: 395-403.

Hilliker, A.J. 1976. Genetic analysis of the centromeric heterochromatin of chromosome 2 of *Drosophila melanogaster*: Deficiency mapping of EMS-induced lethal complementation groups. *Genetics* 83: 765-782.

Hilliker, A.J., Holm, D.G. 1975. Genetic analysis of the proximal region of chromosome 2 of *Drosophila melanogaster*. I. Detachment products of compound autosomes. *Genetics* 81: 705-721.

Holm, D.G., Chovnick, A. 1975. Compound autosomes in Drosophila melanogaster: the meiotic behavior of compound thirds. Genetics 81: 293-311.

Hoskins, R.A., Smith, C.D., Carlson, J.W., Carvalho, A.B., Halpern, A., Kaminker, J.S., Kennedy, C., Mungall, C.J., Sullivan, B.Z., Sutton, G.G., Yasuhara, J.C., Wakimoto, B.T., Myers, E.W., Celniker, S.E., Rubin, G.M., and Karpen, G.H. 2002. Heterochromatic sequences in a Drosophila whole-genome shotgun assembly. *Genome Biology* 3:research0085.1-0085.16.

Howe, M., Dimitri, P., Berloco, M., and Wakimoto, B.T. 1995. Cis-effects of heterochromatin on heterochromatic and euchromatic gene activity in *Drosophila melanogaster. Genetics* 140: 1033-1045.

Irion, U., and Leptin, M. 1999. Developmental and cell biological functions of the *Drosophila* DEAD-box protein abstrakt. *Curr. Biol.* 9:1373-1381.

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

Jowett, T. 1998. In: Drosophila: a practical approach. D. Roberts, Ed. Second edition. Oxford University Press.

Kaminker, J.S., Bergman, C.M., Kronmiller, B., Carlson, J., Svirskas, R., Patel, S., Frise, E., Wheeler, D.A., Lewis, S.E., Rubin, G.M., Ashburner, M., and Celniker, S.E. 2002. The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective. *Genome Biol.* 3: research 0084.

Kania, A., Han, P.L., Kim, Y.T., and Bellen, H. 1993. *neuromusculin*, a *Drosophila* gene expressed in peripheral neuronal precursors and muscles, encodes a cell adhesion molecule. *Neuron* 11: 673-687.

Kania, A., Bellen, H.J. 1995. Mutations in *neuromusculin*, a gene encoding a cell adhesion molecule, cause nervous system defects. *Roux Arch. Dev. Biol.* 204: 259-270.

Karpen, G.H. 1994. Position-effect variegation and the new biology of heterochromatin. *Curr. Opin. Genet. Dev.* 4: 281-291.

Kay, M.A., Zhang, J.Y., and Jacobs-Lorena, M. 1988. Identification and germline transformation of the ribosomal protein Rp21 gene of Drosophila: Complementation analysis with the Minute QIII locus reveals nonidentity. *Molec. gen. Genet.* 213: 354-358

Kelley, R.L., Meller, V.H., Gordadze, P.R., Roman, G., Davis, R.L., and Kuroda, M.I. 1999. Epigenetic spreading of the Drosophila dosage compensation complex from roX RNA genes into flanking chromatin. *Cell* 98: 513-522.

Kennison, J.A., and Tamkun, J.W. 1988. Dosage-dependent modifiers of *Polycomb* and antennapedia mutations in *Drosophila.* Proc. Natl. Acad. Sci. U S A. **85**:8136-40.

Konev, A.Y., Yan, C.M., Acevedo, D., Kennedy, C., Ward, E., Lim, A., Tickoo, S., Karpen, G.H. 2003. Genetics of P-element transposition into Drosophila melanogaster centric heterochromatin. *Genetics* 165: 2039-2053.

Koryakov, D.E., Zhimulev, I.F., and Dimitri, P. 2002. Cytogenetic analysis of the third chromosome heterochromatin of *Drosophila melanogaster. Genetics* 160: 506-517.

Koryakov, D.E., Domanitskaya, E.V., Belyakin, S.N., and Zhimulev, I.F. 2003. Abnormal tissue-dependent polytenization of a block of chromosome 3 pericentric heterochromatin in *Drosophila melanogaster. J. Cell Sci.* 116:1035-1044.

Koundakjiana, E.J., Cowana, D.M., Hardya, R.W. and. Beckera, A.H. 2004. The Zuker Collection: A Resource for the Analysis of Autosomal Gene Function in *Drosophila melanogaster. Genetics* 167: 203-206.

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

Lansdell, S.J., and Millar, N.S. 2000. Cloning and heterologous expression of Dalpha4, a Drosophila neuronal nicotinic acetylcholine receptor subunit: identification of an alternative exon influencing the efficiency of subunit assembly. *Neuropharmacology* 39: 2604-2614.

Letsou, A., Alexander, S., Orth, K., and Wasserman, S.A. 1991. Genetic and molecular characterization of tube, a Drosophila gene maternally required for embryonic dorsoventral polarity. *Proc. Natl. Acad. Sci. USA* 88:810-814.

Lin, J.M., Kilman, V.L., Keegan, K., Paddock, B., Emery-Le, M., Rosbash. M., and Allada, R. 2002. A role for casein kinase 2alpha in the *Drosophila* circadian clock. *Nature* 420: 816-820.

Lindsley, D.L., Edington, C.W., von Halle, E.S. 1960. Sex-linked recessive lethals in Drosophila whose expression is suppressed by the Y chromosome. *Genetics* 45:1649-1670.

Lindsley, D.L., Sandler, L., Baker, B.S., Carpenter, A.T.C., Denell, R.E., Hall, J.C., Jacobs, P.A., Miklos, G.L.G., Davis, B.K., Gethmann, R.C., Hardy, R.W., Hessler, A.Y., Miller, S.M., Nozawa, H., Parry, D.M. and Gould-Somero, M. 1972. Segmental aneuploidy and the genetic gross structure of the Drosophila genome. *Genetics* 71:157-184.

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

Locke, J. and M^{c}Dermid, H.E. 1993. Analysis of Drosophila chromosome 4 using pulsed field gel electrophoresis. *Chromosoma* 102: 718-723.

Locke J., Podemski L., Aippersbach N., Kemp H., and Hodgetts R. 2000. A physical map of the polytenized region (101EF-102F) of chromosome 4 in *Drosophila melanogaster. Genetics.* 155: 175-183.

Lohe, A. R. and Hilliker, A.J. 1995. Return of the H-word (heterochromatin). *Curr. Opin. Genet. Dev.* 5: 746-755.

Lohe, A. R., Hilliker, A.J., and Roberts, P.A. 1993. Mapping simple repeated DNA sequences in heterochromatin of *Drosophila melanogaster. Genetics* 134: 1149-1174.

Lu, B.Y., Emtage, P.C., Duyf, B.J., Hilliker, A.J., and Eissenberg, J.C. 2000. Heterochromatin protein 1 is required for the normal expression of two heterochromatin genes in *Drosophila*. *Genetics* 155: 699-708.

Maggert, K.A. and Karpen, G.H. 2001. The activation of a neocentromere in Drosophila requires proximity to an endogenous centromere. *Genetics* 158: 1615-1628.

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

Marchant G.E. and Holm, D.G. 1988a. Genetic analysis of the heterochromatin of chromosome 3 in *Drosophila melanogaster*. I. Products of compound autosome detachment. *Genetics* 120: 503-517.

Marchant G.E. and Holm, D.G. 1988b. Genetic analysis of the heterochromatin of chromosome 3 in *Drosophila melanogaster*. II Vital loci identified through EMS mutagenesis. *Genetics* 120: 519-532.

Miklos, G.L., and Rubin, G.M. 1996. The role of the genome project in determining gene function: insights from model organisms. *Cell* 86:521-529.

Misra, S., Crosby, M.A. et al. 2002. Annotation of the *Drosophila melanogaster* euchromatic genome: a systematic review. *Genome Biol* 3(12): research0083.1–83.22.

Moazed, D., 2001. Common themes in mechanisms of gene silencing. *Mol. Cell* 8: 489-498

Moschetti, R., Caizzi, R., and Pimpinelli, S. 1996. Segregation distortion in *Drosophila melanogaster*: genomic organization of Responder sequences. *Genetics* 144:1665-1671.

Mottus, R., Reeves, R. and Grigliatti, T.A. 1980. Butyrate suppression of position-effect variegation in Drosophila melanogaster. *Mol Gen Genet.* 178: 465-469.

Muller, H. J. 1930. Types of visible variations induced by X-rays in *Drosophila. J. Genet.* 22: 299-234.

Mullins, C., and Bonifacino, J.S. 2001. The molecular machinery for lysosome biogenesis. *BioEssays* 23: 333-343

Murphy, T.D., and Karpen, G.H. 1995. Localization of centromere function in a Drosophila minichromosome. *Cell* 82:599-609.

Myster S.H., Wang F., Cavallo R., Christian W., Bhotika S., Anderson C.T., and Peifer, M. 2004. Genetic and bioinformatic analysis of 41C and the 2R heterochromatin of Drosophila melanogaster: a window on the heterochromatin-euchromatin junction. *Genetics* 166:807-822.

Nakayama, J., Rice, J.C., Strahl, B.D., Allis, C.D. and Grewal, S.I., 2001. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* 292: 110-113.

Nichols, R., Schneuwly, S.A., and Dixon, J.E. 1988. Identification and characterization of a Drosophila homologue to the vertebrate neuropeptide cholecystokinin. *J. Biol. Chem.* 263: 12167-12170.

Noma, K., Allis, C.D. and Grewal, S.I. 2001. Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. Science 293:1150-1155.

Ohno, S. 1985. Dispensible genes. *Trends Genet.* 1:160-164.

Ostrowski, S., Dierick, H.A., and Bejsovec, A. 2002. Genetic control of cuticle formation during embryonic development in *Drosophila melanogaster*. *Genetics* 161: 171-182.

Peters, A.H., O'Carroll, D. Scherthan, H., Sauer, S. et al. 2001. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell 107: 323-337.

Piacentini, L., Fanti, L., Berloco, M., Perrini, B., Pimpinelli, S. 2003. Heterochromatin protein 1 (HP1) is associated with induced gene expression in Drosophila euchromatin. *J Cell Biol.*161:707-714.

Pimpinelli, S., Berloco, M., Fanti, L., Dimitri, P., Bonaccorsi, S., Marchetti, E., Caizzi, R., Caggese, C., and Gatti, M. 1995. Transposable elements are stable structural components of *Drosophila* heterochromatin. *Proc. Natl. Acad. Sci. USA* 92: 3804-3808.

Piper, R.C., and Luzio, J.P. 2001. Late endosomes: sorting and partitioning in multivesicular bodies. *Traffic* 2: 612-621.

Preston, C.R., Sved, J.A., and Engels, W.R. 1996 Flanking duplications and deletions associated with P-induced male recombination in Drosophila. *Genetics* 144:1623-1638.

Rao, S.S., Stewart, B.A., Rivlin, P.K., Vilinsky, I., Watson, B.O., Lang, C., Boulianne, G., Salpeter, M.M., and Deitcher, D.L. 2001. Two distinct effects on neurotransmission in a temperature-sensitive SNAP-25 mutant. *EMBO J.* 20: 6761-6771.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D. and Jenuwein, T., 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406: 593–599.

Reinhart, B.J., and Bartel, D.P. 2002. Small RNAs correspond to centromere heterochromatic repeats. *Science* 297:1831.

Reugels, A.M., Kurekc, R., Lammermanna, U. and Bünemanna, B. 2000. Mega-introns in the dynein gene DhDhc7(Y) on the heterochromatic Y chromosome give rise to the giant threads loops in primary spermatocytes of *Drosophila hydei*. *Genetics* 154: 759-769

Reuter, G., and Wolff, I. 1981. Isolation of dominant suppressor mutations for position effect variegation in *Drosophila*. *Mol. Gen. Genet.* 182: 516-519.

Reuter, G. and Spierer, P. 1992. Position-effect variegation and chromatin proteins. *BioEssays*. 14: 605-612.

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

Risinger, C., Deitcher, D.L., Lundell, I., Schwarz, T.L., and Larhammar, D. 1997. Complex gene organization of synaptic protein SNAP-25 in *Drosophila melanogaster*. *Gene* 194:169-177.

Roberts, P.A. 1970. Screening for X-ray induced crossover suppressors in *Drosophila melanogaster*: Prevalence and effectiveness of translocations. *Genetics* 65: 429-448.

Roseman, R.R., Johnson, E.A., Rodesch, C.K., Bjerke, M., Nagoshi, R.N., and Geyer, P.K. 1995. A P element containing suppressor of hairy-wing binding regions has novel properties for mutagenesis in *Drosophila melanogaster*. *Genetics* 141: 1061-1074.

Saeboe-Larssen, S., Lyamouri, M., Merriam, J., Oksvold, M.P., and Lambertsson, A. 1998. Ribosomal protein insufficiency and the Minute Syndrome in *Drosophila*: A dose-response relationship. *Genetics* 148: 1215-1224.

Sambrook, J., Fritsch, E.F., and Maniatis, T. 1998. Molecular Cloning: a laboratory manual. Second Edition. Cold Spring Harbour Laboratory Press.

Schotta, G., Ebert, A., Dorn, R. and Reuter, G. 2003. Position-effect variegation and the genetic dissection of chromatin regulation in *Drosophila*. *Seminars in Cell & Developmental Biology* 14: 67-75.

Schultz, J. 1936. Variegation in *Drosophila* and the inert chromosome regions. *Proc. Natl. Acad. Sci. U.S.A.* 22:27-33.

Schulze, S. 2003. A genetic and molecular characterization of two heterochromatic genes in *Drosophila melanogaster* and their homologues in *Drosophila virilis*. Ph.D. thesis, Simon Fraser University, Canada.

Schulze, S., Sinclair, D.A.R., Silva, E., Fitzpatrick, K.A., Singh, M., Lloyd, V.K., Morin, K.A., Kim. J., Holm, D.G., Kennison, J.A., and Honda, B.M. 2001. Essential genes in proximal 3l heterochromatin of *Drosophila melanogaster*. *Mol. Gen. Genet.* 264: 782-789.

Schulze, S.R., Sinclair, D.A.R., Fitzpatrick, K.A. and Honda, B.M. 2005. A Genetic and Molecular Characterization of two proximal heterochromatic genes on chromosome 3 of *Drosophila melanogaster*. Genetics: in press

Shestopal, S.A., Makunin, I.V., Belyaeva, E.S., Ashburner, M., Zhimulev, I.F. 1997. Molecular characterization of the deep orange (dor) gene of *Drosophila melanogaster*. *Molec. gen. Genet.* 253: 642-648.

Sims, R.J. 3rd, Nishioka, K., and Reinberg, D. 2003. Histone lysine methylation: a signature for chromatin function. Trends Genet. 19: 629-639.

Sinclair, D.A.R., Suzuki, D.T., and Grigliatti, T.A. 1981. Genetic and developmental analysis of a temperature-sensitive Minute mutation of *Drosophila melanogaster*. *Genetics* 97: 581-606.

Sinclair, D.A.R., Mottus, R.C. and Grigliatti, T.A. 1983. Genes which suppress position-effect variegation in Drosophila melanogaster are clustered. *Mol. Gen. Genet.* 191:326-333.

Sinclair, D.A.R., Grigliatti, T.A., and Kaufman, T.C. 1984. Effects of a temperature-sensitive *Minute* mutation on gene expression in *Drosophila melanogaster*. *Genet. Res.* 43: 257-275.

Sinclair, D.A.R., Clegg, N.J., Antonchuk, J., Milne, T.A., Stankunas, K., Ruse, C., Grigliatti, T. A., Kassis, J.A., and Brock, H.W. 1998. *Enhancer of Polycomb* is a suppressor of position-effect variegation in *Drosophila melanogaster*. Genetics 148: 211-220.

Sinclair, D.A.R., Schulze, S., Silva, E., Fitzpatrick, K.A., and Honda, B.M. 2000. Essential genes in autosomal heterochromatin of *Drosophila melanogaster*. *Genetica* 109:9-18.

Smothers, J.F., and Henikoff, S. 2000. The HP1 chromo shadow domain binds a consensus peptide pentamer. *Curr Biol.* 10:27-30.

Snay-Hodge, C.A., Colot, H.V., Goldstein, A.L., and Cole, C.N. 1998. Dbp5p/Rat8p is a yeast nuclear pore-associated DEAD-box protein essential for RNA export. *EMBO J.* 17:2663-2676.

Spofford, J.B. 1976. Position-effect variegation in *Drosophila*. In: Ashburner M, Novitski E (eds) The genetics and biology of *Drosophila*, vol 1c. Academic Press, London, pp 955-1018.

Spradling, A.C., Stern, D., Beaton, A., Rhem, E.J., Laverty, T., Mozden, N, Misra, S., and Rubin, G.M. 1999. The Berkley Drosophila Genome Project gene disruption project: Single P-element insertions mutating 25% of vital Drosophila genes. *Genetics* 153: 135-177.

Struhl, K. 1999. Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. *Cell* 98:1-4.

Sullivan, B.A., Blower, M.D., and Karpen, G.H. 2001. Determining centromere identity: cyclical stories and forking paths. *Nat Rev Genet.* 2: 584-596.

Sullivan B.A., and Karpen, G.H. 2004. Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nat Struct Mol Biol.* 11:1076-83.

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

Sun, F.L., Cuaycong, M.H., and Elgin, S.C.R. 2001. Long-range nucleosome ordering is associated with gene silencing in Drosophila melanogaster pericentric heterochromatin. *Molec. Cell. Biol.* 21: 2867-2879.

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

Sun, X., Wahlstrom, J., and Karpen, G. 1997. Molecular structure of a functional Drosophila centromere. *Cell* 91:1007-1019.

Sun, X., Le, H.D., Wahlstrom, J.M., and Karpen, G.H. 2003. Sequence analysis of a functional Drosophila centromere. *Genome Res.*13:182-194.

Tasaka, S.E., Suzuki, D.T. 1973. Temperature-sensitive mutations in *Drosophila melanogaster. Genetics* 74:509-520.

Tulin, A., Stewart, D., and Spradling, A.C. 2002. The Drosophila heterochromatic gene encoding poly(ADP-ribose) polymerase (PARP) is required to modulate chromatin structure during development. *Genes and Dev.* 16:2108-2119.

Tulin, A. and Spradling, A.C. 2003. Chromatin loosening by Poly(ADP)-Ribose Polymerase (PARP) at Drosophila puff loci. *Science* 299: 560-562

Uchida, K., Hanai, S., Ishikawa, K., Ozawa, Y.I., Uchida, M., Sugimura, T., and Miwa, M. 1993. Cloning of cDNA encoding *Drosophila* poly(ADP-ribose) polymerase: leucine zipper in the auto-modification domain. *Proc. Natl. Acad. Sci. USA* 90: 3481-3485.

Uchida, M., Hanai, S., Uematsu, N., Sawamoto, K., Okano, H., Miwa, M., and Uchida, K. 2001. Genetic and functional analysis of PARP, a DNA strand break-binding enzyme. *Mutat. Res.* 477: 89-96.

Vilinsky, I., Stewart, B.A., Drummond, J., Robinson, I. and Deitcher, D.L. 2002. A Drosophila SNAP-25 null mutant reveals context dependent redundancy with SNAP-24 in neurotransmission. *Genetics* 122: 625-642.

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

Wakimoto, B.T., and Hearn, M.G. 1990. The effects of chromosome rearrangements on the expression of heterochromatic genes in chromosome 2L of Drosophila melanogaster. *Genetics* 125: 141-154.

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

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

Warner, T.S., Sinclair, D.A., Fitzpatrick, K.A., Singh, M., Devlin, R.H., and Honda, B.M. 1998. The light gene of *Drosophila melanogaster* encodes a homologue of *VPS41*, a yeast gene involved in cellular protein trafficking. *Genome* 41: 236-243.

Warren, W.D., Lin, E., Nheu, T.V., Hime, G.R., and M^{c}Kay, M.J. 2000. *Drad21*, a *Drosophila rad21* homologue expressed in S-phase cells.. Gene 250: 77-84.

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

Wustmann, G., Szidonya, J., Taubert, H., and Reuter, G. 1989.
The genetics of position-effect variegation modifying loci in Drosophila melanogaster. *Molec. gen. Genet.* 217:520-527.

Yan, C.M., Dobie, K.W., Le, H.D., Konev, A.Y., and Karpen, G. H. 2002. Efficient recovery of centric heterochromatin P-element insertions in *Drosophila melanogaster*. *Genetics* 161: 217-229.

Yasuhara, J.C., Marchetti, M., Fanti, L., Pimpinelli, S., and Wakimoto, B.T. 2003. A strategy for mapping the heterochromatin of chromosome 2 of *Drosophila melanogaster*. *Genetica* 117: 217-226.

Yeh, E., Gustafson, K., and Boulianne, G.L. 1995. Green fluorescent protein as a vital marker and reporter of gene expression in *Drosophila*. *Proc. Natl. Acad. Sci. USA* 92: 7036-7040.

Zaffran, S., Das, G., and Frasch, M. 2000. The NK-2 homeobox gene *scarecrow* (*scro*) is expressed in pharynx, ventral nerve cord and brain of *Drosophila* embryos. *Mech. Dev.* 94: 237-241

Zhang, P., and Spradling, A.C. 1994. Insertional mutagenesis of *Drosophila* heterochromatin with single P elements. *Proc. Natl. Acad. Sci. USA* 91: 3539-3543.