The *l(2)03605 P* element induced lethality is not caused by disruption of the *Drosophila melanogaster* High Mobility Group genes D (*HMG-D*) or Z (*HMG-Z*)

by

Erick Robert James

B.Sc., The University of British Columbia, 1995

A thesis submitted in partial fulfillment of the requirements for the degree Master of Science

in

The Faculty of Graduate Studies

Department of Zoology

We accept this thesis as conforming to the required standard

The University of British Columbia

1999

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Zoology

The University of British Columbia
Vancouver, Canada

Date Feb 19/99

DE-6 (2/88)

Abstract

Chromatin is functionally defined as DNA, histones and non-histone chromosomal proteins (NHCP). The High Mobility Group (HMG) 1/2 proteins represent one well defined set of NHCP in eukaryotes . These proteins are thought to act as architectural components of chromatin and are well characterized biochemically but not biologically. HMG 1/2 proteins have been implicated in transcription, DNA replication, chromatin assembly and stability. The *Drosophila* HMG1 homologues are HMG-D and HMG-Z. The *Drosphila* HMGs are thought to play a role in establishing developmental specific regions of chromatin that are silent and easily replicatable, albeit at different developmental times and locations, but again their biological role is speculative. Unfortunately, no mutations exist in the HMG1/2 class of proteins and mutants would be extremely useful in assigning a biological function to these proteins. However, the *Drosophila* HMG gene are cytologically mapped to a precise region, banded region 57F8-11 on chromosome 2. I attempted to identify *P* element insert mutations in *HMG-D/Z* using the Berkeley *Drosophila* Genome Project *P* element insertion lines. The Berkeley *Drosophila* Genome Project has undertaken the daunting task of *P* element tagging essential genes in *Drosophila*. Their *P* element inserts are mapped cytologically and the expression pattern of enhancer trap expression are known. One such recessive lethal line, *l(2)03605/CyO* has a *P* element inserted at 57F8-10, the same cytological location as *HMG-D/Z*. In addition, the insert shows *β-galactosidase* expression in the nervous system, while *HMG-D/Z* mRNA is also found in embryonic neural tissues. This coincident expression pattern and the cytological location suggest that the *l(2)03605 P* element insertion may have disrupted either or both genes. Therefore, I set out to determine whether the *l(2)03605* P element is either inserted into and/or disrupting the *HMG-D/Z* genes. This work shows the *l(2)03605 P* element does not affect either *HMG-D* or *HMG-Z* and most likely disrupts another gene in the 57F8-10 region.

Table of Contents

List of Figures and Tables

Abbreviations

BDGP:	Berkely *Drosophila* Genome Project
CNS:	Central Nervous Sytem
HMG:	High Mobily Group
IPTG:	Isoppropyl-1-thio-β-D-galactoside
LEF-1:	Lymphoid Enhancer-binding Factor
NHCP:	Non-Histone Chromosomal Proteins
ORF:	Open Reading Frame
PCA:	perchloric acid
PCR:	Polymerase Chain Reaction
PNS:	Peripheral Nervous System
SRY:	Sex-determining factor
TCA:	trichloracetic acid
UTR:	Un-Translated Region
wt:	wild-type

Acknowledgment

I would like to thank Dr. Thomas A. Grigliatti for allowing me to pursue my scientific education in his laboratory. My Master's committee: Dr. Hugh W. Brock for all your constructive criticism and guidance, Dr. George Spiegelman. The "Back Lab": Michael O'Grady, James Whalen and Michael Harrington for all the technical and especially the moral support. The rest of the Grigliatti lab and finally my family and friends who put up with the "hermit" lifestyle science forces us into.

Introduction

High Mobility Group Chromatin Proteins

The formation and organization of chromatin, functionally defined as a complex of DNA, histones, and non-histone proteins (Alberts *et al.*, 1994) is an extremely important process in the life cycle of a eukaryotic cell. Not only must the DNA be condensed inside the small nucleus, but it must also remain a substrate for replication and transcription. Naked DNA is wrapped around a core of four basic proteins, collectively called histones: H2A, H2B, H3 and H4. This nucleosomal core is the basic structural unit of chromatin. Each nuclesome core consists of an octamer of histones associated with 146 nucleotide pairs of DNA. This octamer contains two each of histones H2A, H2B, H3 and H4. A fifth histone variant, H1 is a linker histone binds to the DNA between nucleosome cores, further condensing the DNA into a chromatosome. This form of DNA is the 30-nm fibre. Higher order compaction is poorly understood; it arises by further coiling and looping of the chromatin and is maximally condensed in the metaphase chromosome (Alberts *et al.*, 1994).

A significant component of chromatin are the non-histone chromosomal proteins (NHCP). These proteins include transcription factors, chromatin bound enzymes (replication, transcription and DNA repair machinery), scaffold proteins and the high mobility group proteins (Van Holde *et al.*, 1995).

The high mobility group (HMG) proteins are among the largest and best characterized group of non-histone chromosomal proteins (Bustin and Reeves, 1996). HMG proteins were originally identified in the early 1960s as acidic impurities in histone H1 preparations (Johns, 1964). All HMG proteins bind DNA, are extractable from nuclei with 0.35 M NaCl, are soluble in 5% perchloric acid (PCA) or 2 -5% trichloroacetic acid

(TCA) and have a high electrophoretic mobility in polyacrylamide gels (Johns, 1982). Hence, the name High Mobility Group proteins. Typically HMG proteins have a molecular mass <30 kDa, and are highly charged. Despite early purification and well defined protein structure and biochemical characterization, their cellular function remains an enigma (Bustin and Reeves, 1996; Grosschedl *et al.*, 1994). HMG proteins have been classified into three families based on their molecular mass, DNA-binding characteristics and amino acid sequence motifs. These families are the HMG-I(Y), HMG-14/-17 and the HMG-1/-2 class.

The HMG-I(Y) family of proteins are known as "architectural transcription factors" due to their ability to function both as components of chromatin and as secondary transcription factors (Reeves and Wolffe, 1996). HMG-I(Y) is involved in both the induction and negative regulation of transcription (Bustin and Reeves, 1996). HMG-I(Y) also interacts specifically with known transcription factors (for example NF- κB and ATF-2) and induces DNA structure changes at promoter/enhancer sequences, aiding in transcription initiation (Falvo *et al.*, 1995; Thanos and Maniatis, 1995). How HMG-I(Y) specifically acts as an architectural transcription factor is unknown.

Another class of HMG proteins is the HMG-14/-17 proteins, which are proposed to be transcriptional activators. By binding to nucleosome cores they change the rate of transcription elongation, but do not affect transcription initiation. HMG-14/17 are thought to be architectural components that help create an unfolded chromatin fibre that facilitates transcription by removing the repressive action of the histones (Bustin *et al.*, 1995).

The well studied HMG1/2 group can be further subdivided according to the number of HMG domains present in the protein and their specificity of sequence recognition. The HMG domain is a DNA-binding motif of approximately 80 amino acids (Ner, 1992). In general proteins with multiple HMG domains show little or no sequence specificity, (for example HMG-1 and-2 with two HMG-domains). Alternatively those

proteins with a single HMG domain show interactions with specific DNA sequences, (for example the transcription factors: sex-determining factor (SRY) (Sinclair *et al.*, 1990) and the lymphoid enhancer-binding factor (LEF1) (Travis *et al.*, 1991)). The cellular function of HMG1/2 proteins has yet to be clearly discovered. However, a role in transcription, replication, chromatin assembly and the stabilization of chromatin structure has been implicated (Landsman and Bustin, 1993). Recently, *Xenopus* HMG1 protein was found in early embryonic chromatin lacking H1 (Nightingale *et al.*, 1996).

With the stage set as HMG-proteins being architectural-chromatin proteins, a summary of our current understanding of two specific HMG1/2 proteins, HMG-D and HMG-Z will ensue followed by the key question of this thesis as it relates to these two proteins.

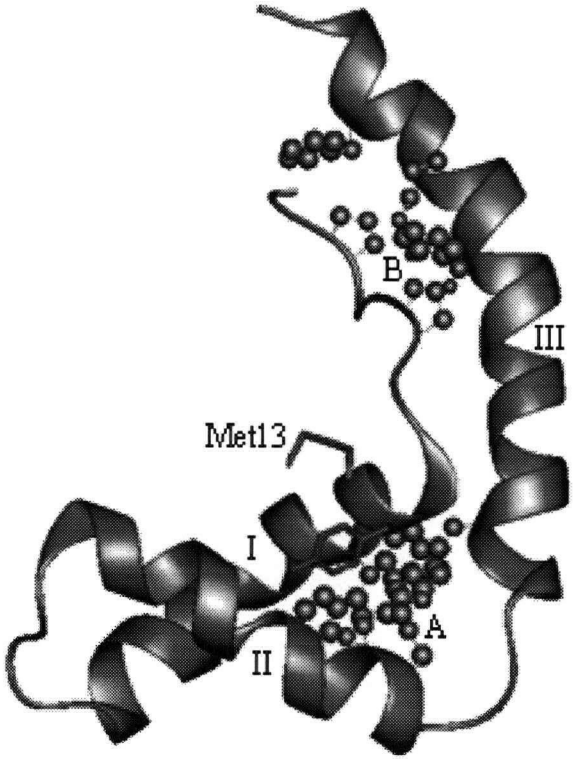
HMG-D and HMG-Z

The *Drosophila* HMG-D protein is a highly abundant chromosomal protein that is closely related by amino acid sequence to the mammalian HMG1 and the yeast *S. cerevisiae* NHP6 proteins (Wagner *et al.*, 1992). However, HMG-D only contains a single copy of the HMG-domain DNA binding motif (Landsman and Bustin, 1991, 1993). The HMG domain is followed by a basic region similar to the carboxyl terminal domain of histone H1 and a short carboxyl terminal acidic stretch similar to those in HMG1/2 (Churchill *et al.*, 1995). The NMR structure of the HMG domain of HMG-D is similar to the B-domain of HMG-1, which is an L-shaped fold formed by three α-helices (Jones *et al.*, 1994) held together by two hydrophobic cores (Balaeff *et al.*, 1998) (see Figure 1). Churchill *et al.* (1995) showed HMG-D preferentially binds to A/T rich DNA containing the dinucleotide sequence TG. This TG site is a readily deformable base combination and shows the preference of HMG-D for pre-bent DNA sites. Furthermore, HMG-D binding to DNA induces further bending in the DNA.

HMG-D is encoded by a single copy gene located at region 57F8-11 on the right arm of chromosome 2 (Wagner *et al.*, 1992). Northern analysis of *HMG-D* revealed a 1.3

Figure 1

The HMG-domain of HMG-D. The solution structure of HMG-D obtained by NMR. I, II and III represent α-helices, while A and B refer to hydrophobic cores (Balaeff *et al.*, 1998).


B
III
Met13
I
A
II

kb transcript and a 0.95 kb transcript. These transcripts are present at relatively equal and high levels in RNA prepared from 0-4 and 4-8 hr embryos with slightly lower amounts present in the 8-12 and 12-16 hr embryos. Trace amounts of *HMG-D* transcripts are found in later stages of development. *Neither HMG-D* mRNA is seen in the adult male, however both transcripts are present in the adult female (Ner *et al.*, 1993; Wagner *et al.*, 1992).

Stroumbakis and Tolias (1994) examined the distribution of *HMG-D* mRNA in the female germline and early embryo. In stage 1-6 egg chambers, *HMG-D* mRNA is found throughout the germarium, in the germline nurse cells and in the somatic follicle cells. At stage 7, *HMG-D* transcripts stop accumulating in follicle cells, however transcripts continue to accumulate in the nurse cells until stage 10. During stage 11, the nurse cells begin to transfer *HMG-D* mRNA into the oocyte where these transcripts accumulate around the periphery of the oocyte until the end of oogenesis. Zygotic *HMG-D* mRNA expression resumes in neuroblasts that are migrating from the neural ectoderm in the stage 9 embryo that is entering the slow phase of germband elongation. This zygotic expression continues throughout the remaining stages of embryogenesis, specifically in neuroblasts of the CNS. Most neuroblasts on both sides of the ventral cord and brain lobes and a subset of cells in the PNS express *HMG-D* mRNA.

The suggestion that later zygotic *HMG-D* expression may play a significant role in the development of the *Drosophila* embryonic nervous system (Stroumbakis and Tolias, 1994) is not without precedent. Another vertebrate HMG1 homologue, HMG-X of *Xenopus* is also abundantly expressed in neuroectoderm-derived tissues during embryogenesis (Kinoshita *et al.*, 1994). HMG-X is thought to participate in neural fate determination by regulating, through modulation of chromatin configuration, the ability of transcription factors to induce the expression of a specific set of genes.

Interestingly there is a second HMG-1 related protein in *Drosophila*, HMG-Z (Ner *et al.*, 1993). HMG-Z is 65% identical and 78% similar at the amino acid level to HMG-

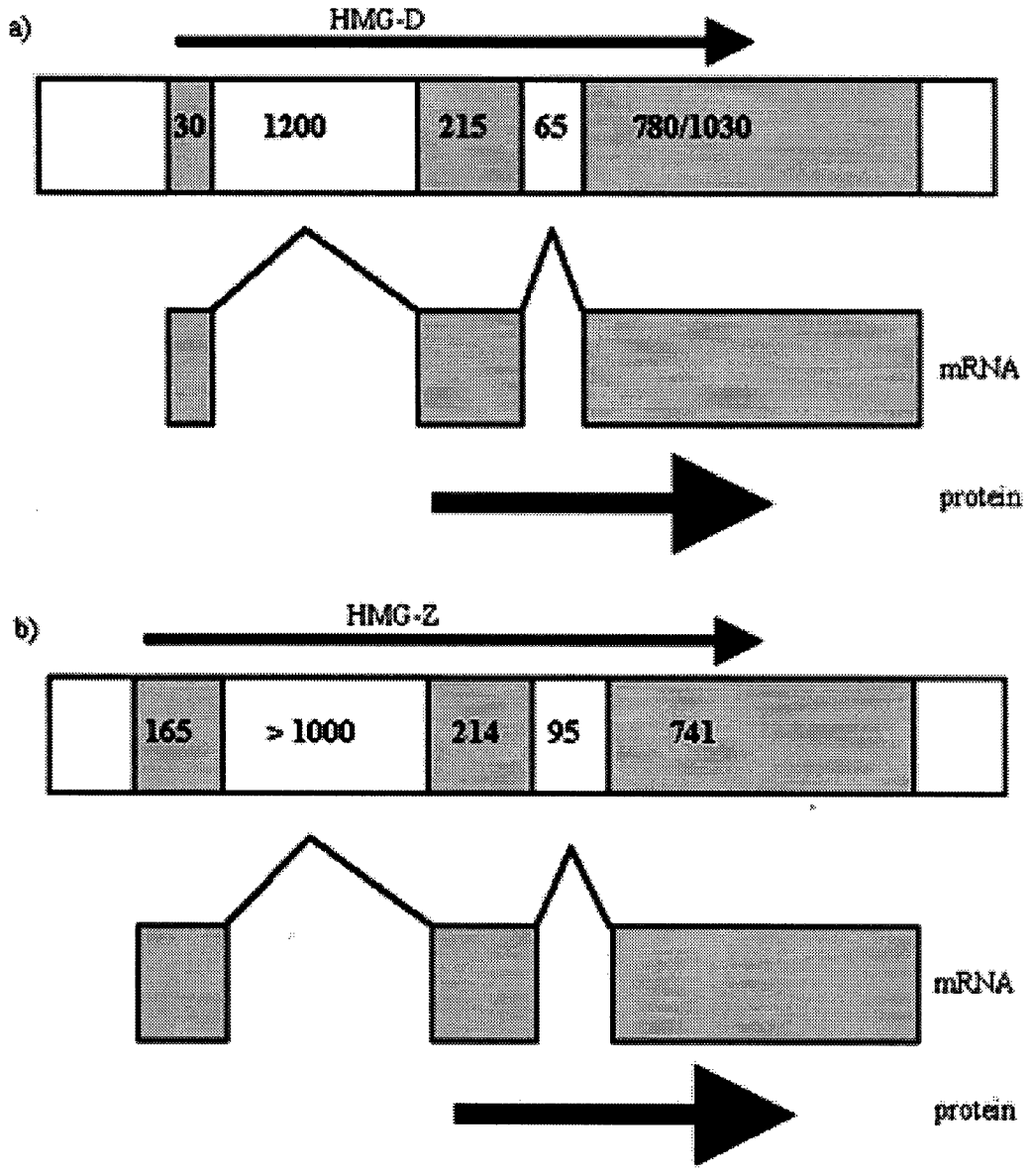
D. The genes encoding HMG-D and HMG-Z are adjacent on the chromosome and oriented head-to-head. *HMG-D* and *HMG-Z* each contain two introns, shared sequence and conserved intron/exon arrangements and their close proximity indicate a possible gene duplication (Ner *et al.*, 1993). *HMG-D*'s first intron named intervening sequence (IVS) one (IVS1) is 1.2 kb long and is within the 5′ UTR, resulting in a short exon of 30 bp. The second exon, encoding the 39 N-terminal residues of the HMG-D protein is 215 bp long. IVS2 is only 65 bp and is followed by exon3 which contains the rest of the coding region and the 3′ UTR. *HMG-Z* also contains three exons (exon1-165 bp, exon2-214 bp and exon3-741 bp) which are separated by two introns (IVS1 >1 kb and IVS2 95 bp) (Figure 2) (Ner *et al.*, 1993).

Northern analysis of *HMG-Z* reveals a 0.65 kb and a 1.2 kb transcript. A small amount of the 1.2 kb transcript appears in 0-4 hr embryos suggesting either that the maternal contribution of HMG-Z is minor or it arises from early zygotic expression. Both transcripts are seen in 4-8 hr and 8-12 hr embryos, with the 1.2 kb transcript being more abundant. Both transcripts are present in trace amounts in later embryos. There is a high level of the 0.65 kb transcript in 1st instar larvae, suggesting a requirement for HMG-Z protein during larval development (Ner *et al.*, 1993).

Functionally, HMG-D and HMG-Z may play similar roles but at different times during development. HMG-D is present at a constant level in all stages of development as determined by Western analysis (Ner and Travers, 1994). In embryos HMG-D is confined to the nuclei from the start of embryogenesis to the blastoderm stage as shown by *in situ* localizations with anti-HMG-D antibodies. The most intense initial staining is observed in mitotic chromosomes and the polar bodies. The polyploid yolk nuclei and the pole cell nuclei also contain HMG-D. Interestingly, histone H1 as detected by Western analysis first appears in the interphase and mitotic nuclei of some nuclear division cycle 7 embryos. During cycles 8 and 9, histone H1 staining becomes more intense and increases dramatically due to zygotic transcription. Histone H1 is an

Figure 2

Exon/intron organization of a) *HMG-D* and b) *HMG-Z* (Ner *et al.*, 1993).


a)
HMG-D
30
1200
215
65
780/1030
mRNA
protein
b)
HMG-Z
165
> 1000
214
95
741
mRNA
protein

architectural protein that binds to the linker DNA flanking the nucleosome core, aiding in the assembly of higher-order chromatin structures (Wolffe, 1997). During the early stages of development, there is a dramatic shift in the relative amounts of H1 and HMG-D, with H1 eventually becoming the more abundant protein (Ner and Travers, 1994). Initially, during early embryogenesis HMG-D is associated with a less condensed chromatin structure that lacks histone H1. As histone H1 increases it replaces HMG-D and the chromatin takes on a different, more compact chromatin structure. This is demonstrated by a decrease in the size of the mitotic chromatids as H1 displaces HMG-D.

The physiological significance of the differing states of chromatin during development is not known. However, there is the possibility that HMG-D containing chromatin is less compact, which would facilitate the rapid condensation and decondensation required during the very short early cleavage cycles (Ner and Travers, 1994). Also, there is a correlation of the appearance of histone H1 and the initial transcriptional competence of the cell after the mid-blastula transition (Anderson and Lenhyel, 1980). This suggests that chromatin containing HMG-D is transcriptionally silent, and transcriptional competence is associated with replacement of HMG-D by histone H1. Thus, HMG-D and HMG-Z may function in organizing silent chromatin (Ner and Travers, 1994). The *Xenopus* HMG1 protein also causes a silent chromatin conformation in pre mid-blastula transition embryos. HMG1 represses *in vitro* transcription of a dinucleosomal template to the same level as histone H1. There is a high concentration of HMG1 in the *Xenopus* egg and embryo that may contribute to transcriptional silencing of the early embryonic chromatin, in the same unknown manner as HMG-D (Ura *et al.*, 1996). However, a 6-fold higher concentration of HMG1 over H1 is required to repress transcription at the same level. In addition, histone H1 has a 40-fold higher affinity for nucleosomal DNA than HMG1.

Biochemical studies on HMG-D/Z protein structure (Jones *et al.*, 1994) and their DNA binding properties (Balaeff *et al.*, 1998; Churchill *et al.*, 1995) have shed little light

on these proteins biological significance. A recent study has discovered that there are two isoforms of an *E. coli* produced recombinant HMG-D. One isoform contains an oxidized methionine residue at position 13 of the protein while the other isoform is not oxidized at position 13. Furthermore, this oxidation of Met^{13} decreases the DNA binding affinity 3-fold over the non-oxidized isoform of HMG-D (Dow *et al.*, 1997). Unfortunately, this result only exists in *E.coli* and appears to be a result of the recombinantly produced protein. Unless the *Drosophila* HMG-D protein is also shown to be oxidized, the results bear little biological significance.

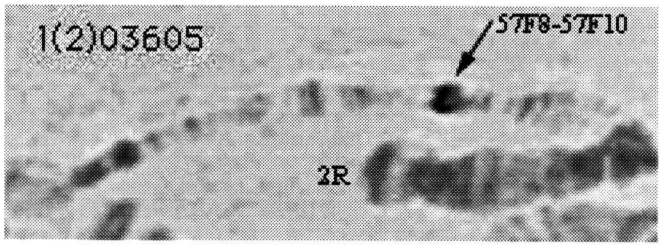
In the absence of biological data, we do not know exactly what roles HMG 1/2 proteins have during development and the cell cycle. The best way to get a handle on biological function is to perturb the system and find out what happens. Unfortunately, there are no known mutations in the HMG 1/2 proteins.

The Berkeley *Drosophila* Genome Project (BDGP) has undertaken an extremely important and productive task of *P* element tagging essential genes within the fruitfly. A *P* element is a *Drosophila* transposable element that has been used extensively for mutagenesis and germline transformation (Griffiths *et al.*, 1993). The *P* element Gene-disruption Project has produced many randomly induced *P* element lines, each containing a single *P* element insert that is recessive lethal. The cytological location of all inserts have been mapped by *in situ* hybridization to polytene chromosomes. In addition to cytological mapping, the project is also isolating and sequencing the DNA flanking the *P* element using either plasmid rescue or inverse PCR. With these two goals complete it will be possible to link the physical, cytological and genetic maps of the *Drosophila* genome (BDGP, 1996).

This large pool of *P* element tagged lines allows for the recovery of insertional mutations in essential genes and the cloning of those genes, such as HMG-D/Z. The *P*

Figure 3

In situ hybridization of polytene chromosomes using DNA flanking *l(2)03605 P* element as a probe. The figure indicates the *P* element is located on the right arm of chromosome 2 at cytological position 57F8-F10 (Spradling *et al.*, 1995)


1(2)03605
57F8-57F10
2R

element associated with the recessive lethal mutant *l(2)03605* was localized to the cytological location 57F8-F11 by *in situ* hybridization on polytene chromosomes (Figure 3), the same cytological location as the *HMG-D* and *HMG-Z* loci. The *PZ*-enhancer trap *P* element used to generate the recessive lethal lines contains the structural gene for *β-galactosidase*. Upon insertion into the genome, the *β-galactosidase* expression falls under the control of local promoters/enhancers (a classical enhancer trap experiment) (Mlodzik and Hiromi, 1992). The *β-galactosidase* expression pattern of the *l(2)03605 P* element is confined to the nervous system (Figure 4). This expression pattern is coincident with the zygotic expression pattern of *HMG-D* and *HMG-Z* (Ner, personal communication; Stroumbakis and Tolias, 1994)] Therefore, the *l(2)03605 P* element line is an excellent candidate for a mutant which disrupts *HMG-D, HMG-Z* or both to recessive lethality. This was the working hypothesis for this thesis. Unfortunately the *l(2)03605 P* element was shown not to disrupt either *HMG-D* or *HMG-Z*. Therefore, it must be affecting some other gene nearby.

Materials and Methods

Fly strains

The fly strains used in this study are described on Flybase (http://flybase.bio.indiana.edu/) and in Lindsley and Zimm (1992). All flies were obtained from the Bloomington stock centre.

P element recessive lethal line: Stock #: P1341, Cytological location: 057F08-10

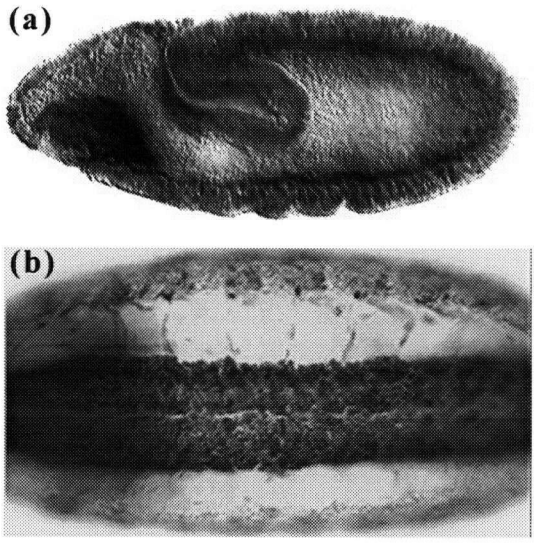
cn[1] *P*{*ry*$^{+}$=PZ}*l(2)03605*^{03605}/*CyO*; *ry*^{506}

Deficiency: Stock #: 2606, Breakpoints: 057B04;058B

Df(2R)Pu-D17, cn^{1} *bw*^{1} *sp*^{1}/*SM1*

Figure 4

LacZ expression pattern in *l(2)03605* late embryos is confined to the nervous system. (a) A lateral view of a 12 hour embryo with anterior to the left, posterior to the right, dorsal at the top and ventral at the bottom shows expression restricted to the CNS. (b) A dorsal view of a 18 hour embryo with anterior to the left and posterior to the right shows CNS expression but also expression in the PNS (Spradling *et al.*, 1995).


(a)
(b)

Genomic DNA isolation

Fifty flies were homogenized in 2 ml of a 0.2 M saccharose, 0.1 M Tris-HCl pH 9.2, 0.1 M NaCl, 0.05 M EDTA and 0.5% SDS solution at room temperature. The homegenate was then incubated at 65° C for 30 minutes, followed by the addition of 300 µl of 8M potassium acetate pH 10. This solution was then mixed on a vortex and placed on ice for 30 minutes, then centrifuged in a tabletop centrifuge. The supernatant was then treated with 5 µl of 1µg/µl RNase A for 30 minutes at room temperature, phenol/chloroform extracted with an equal volume of 1:1 phenol/chloroform , followed by ethanol precipitation with 95% ethanol, a wash with 70% ethanol and re-suspension in 50 µl of 1X TE.

P1 DNA isolation (P1 phage clone # DS08966-94-38 from the Berkeley Genome Project contains the *l(2)03605 P* element insertion site)

A single P1 containing colony was grown overnight in Luria-Bertani Medium (LB) with 25 µg/ml of kanamycin. 500 mls of LB with 25 µg/ml of kanamycin and 5% sucrose were inoculated with 500 µl of the overnight culture and shaken at ~ 250 rpm at 37° C for about 5 hours (OD_{550} = 0.15). Five mls of filter-sterilized 0.1 M IPTG were added to the culture and further grown with shaking at 37° C for another 5 hours (OD_{550} = 1.3 to 1.5). The cells were then harvested by centrifugation and the DNA isolated with a Qiagen maxi-prep kit according to the manufacturer's instructions. The resulting DNA was then re-suspended in 100 µl of TE.

LB medium

Ten grams of bacto-tryptone, 5 g bacto-yeast extract, 10 g NaCl were added to 950 mls deionized H_2O and shaken until the solutes had dissolved. The pH was adjusted to 7.0 with the addition of 5 N NaOH. The volume was adjusted to 1 liter with deionized H_2O and the medium was then sterilized by autoclaving for 20 minutes at 15 lb/sq. in. on the liquid cycle (Sambrook *et al.*, 1989).

Total RNA isolation

Fifty flies were homogenized in 1 ml of TRIzol® Reagent (GIBCOBRL). The insoluble material was removed by centrifugation at 12,000 x *g* for 10 minutes at 4° C. The cleared homogenate solution was then transferred to a fresh tube for phase separation. This homogenate solution was then incubated at room temperature for 5 minutes. 200 µl of chloroform was added, followed by incubation at room temperature for 2 to 3 minutes. The samples were then centrifuged at 12,000 x *g* for 15 minutes at 4° C. After centrifugation the aqueous phase was transferred to a fresh tube, and the RNA was precipitated by the addition of 500 µl of isopropyl alcohol. The samples were then incubated at room temperature for 10 minutes and centrifuged at 12,000 x *g* for 10 minutes at 4° C. The supernatant was subsequently removed and the RNA pellet washed with 1 ml of 75% ethanol. The RNA pellet was then briefly air-dried and redissolved in RNase-free distilled water.

Plasmid rescue

Two to 5 genomic equivalents of genomic DNA were digested with *Xba* I, *Xba* I/*Nhe* I, and *Xba* I/*Spe* I followed by either heat inactivation of the restriction enzyme or phenol/chloroform extraction. Digested DNA was resuspended in 20 - 50 µl of TE. Half

of the DNA was used for the subsequent ligation reaction. The ligation reaction was performed in a 200 µl volume to reduce intermolecular ligation events. Therefore 20 µl of 10X ligation buffer and 2 µl of 1 U/µl ligase was used and the remaining volume made up with distilled water. The ligation reaction was incubated at 16° C overnight. The DNA in the ligation mix was then precipitated by the addition of 20 µl of 3 M sodium acetate and 2 volumes of ethanol. The DNA pellet was then resuspended in 10 to 20 µl of TE. The resuspended DNA was then transformed into either Subcloning Efficiency DH5α™ Competent Cells (GIBCOBRL) following the manufacturer's instructions or lab made electro-competent *E.coli* (DH5α™) using the BTX Transporator™ Plus (see below for protocol).

Preparation of electro-competent cells

Five hundred mls of LB were innoculated with 1-5 mls of an overnight bacterial culture and incubated at 37° C until OD_{600} = 0.6. The culture was then chilled on ice for 30 minutes and the cells harvested by centrifugation at 4000 g for 15 minutes at 4° C. The resulting pellet of cells was then washed twice with 250 mls of ice-cold 10% glycerol (autoclaved), harvested by centrifugation at 4000 g for 15 minutes at 4° C and resuspended in 1 ml ice-cold GYT medium (10% glycerol, 0.125% yeast extract, 0.25% tryptone, autoclaved to sterilize). The competent cells were aliquoted into 1.5 ml ice-cold eppendorf tubes, flash frozen in liquid nitrogen and stored at -80° C.

Electroporation

BTX Disposable Cuvettes (P/N 610 (1 mm gap)) were used with a set charging voltage of 1.4 kV, field strength of 13.0 - 15.0 kV/cm and a pulse length of 5 - 6 msec. Forty µl of

electro-competent cells were placed in an ice-cold cuvette along with 1 ng/ul of DNA, mixed by 'flicking' and placed on ice for 1 minute. The electrical charge was applied and 960 µl of LB was added immediately, followed by gentle mixing with a pipette. The cuvette was then transferred to 37° C for 1 hour, followed by transfer to the appropriate agar plate.

Calculation of theoretical number of rescue plasmids

One can determine how many plasmids should be recovered for a set of rescue experiments based on the ratio of rescuable plasmid DNA to total diploid genomic DNA. The minimum size of the rescue plasmid would be 7 kb (Figure 5a) and the total diploid genome size of *Drosophila melanogaster* is 330 Mb (Rubin, 1998). Thus, the ratio of rescuable plasmid DNA to total genomic DNA is 7 kb:330 Mb or 2.12×10^{-5}. In the transformation portion of the rescue experiment 100 ng of DNA was transformed 100 independent times. Therefore 10,000 ng of DNA was transformed. To determine how much of this total DNA is plasmid DNA, we simply multiply 10,000 ng of DNA to the 2.12×10^{-5} ratio, coming up with 0.212 ng of plasmid DNA. The transformation efficiency (CFU) of the competent *E.coli* is 25 ng DNA giving 1×10^{4} colonies. Consequently 2.5×10^{-3} ng of DNA gives rise to a single colony. Simply dividing the total amount of plasmid DNA 0.212 ng by 2.5×10^{-3} ng yields 84.8 plasmids. Therefore, approximately 85 rescue plasmid should have been recovered.

DNA purification for labeled probes

DNA from a restriction enzyme digest or from PCR was subjected to electrophoresis and the desired band cut out of the gel and placed in a eppendorf tube at

-80° C for 15 minutes. The gel slice was allowed to thaw completely and centrifuged in a tabletop centrifuge at 14, 000 rpm for 10 minutes. The supernatant was transferred to a new tube with the addition of 1/10 volume of sodium acetate and 2 volumes of 95% ethanol. The precipitated DNA was again spun and washed with 70% ethanol and then air dried. The resulting DNA pellet was resuspended in 10 µl of TE. 1µl of the sample was then subjected to electrophoresis for quantification.

Southern and Northern analysis

Radioactively labeled (Amersham α^{32} PαdATP) DNA probes were made with the Boehringer-ManNheim random prime DNA labeling kit according to manufacturer's instructions and the probe purified with Pharmacia Biotoech MicroSpin™ S-300 HR columns. Transfer and hybridization procedures were followed according to standard protocols (Sambrook *et al.*, 1989). Pre-hybridization and hybridization was done in 25 mM KPO_4 pH 7.4, 5X SSC, 5X Denhardt's solution, 50 µg/ml sheared salmon sperm DNA, 50% formamide and 10% dextran sulfate at 42° C. Washes were as follows: once at room temperature to 60° C (room temperature wash was added and the solution brought to 60° C) in 1X SSC, 0.1% SDS, followed by a 15 minute wash at 60° C with the same wash solution. The blots were then washed twice with either 0.5X SSC, 0.1% SDS (low stringency) or 0.1X SSC, 0.1% SDS (high stringency) at 60° C for 15 minutes. Blots were exposed to Kodak XOMAT film at -70°C for 1-3 days.

Densitometic analysis

Analysis was carried out using NIH Image according to Chow and Evans (http://scrc.dcrt.nih.gov/imaging/tutorials/gel_density/table.html).

Polymerase Chain Reactions (PCR)

Reactions were completed in 50 µl reaction volumes. The final concentrations of the various components of the PCR are as follows: 1X of the 10X reaction buffer; 2.5 mM $MgCl_2$; 1.0 mM dNTPs; 50 ng DNA; 25 pmoles of primers; 1 U *Taq* DNA polymerase. The cycles were as follows: 1) 94° C for 1 minute; 2) 92° C for 1 minute; 3) 58° C for 1 minute; 4) 72° C for 2.5 minutes and 5) steps 2 through 4 were repeated 32 times.

Primers: P_{out}: 5′ CGG GAC CAC CTT ATG TTA TTT CAT C 3′

P1341B: 5′ GGA ATT CCG ACA GAA GAA TAC CGC AAA CTC 3′

Fly crosses

All crosses were carried out at 25° C on standard cornmeal sucrose media containing 0.04 % Tegosept as a mold inhibitor. In the reversion experiment 50 vials, each containing 1 (*l(2)03605*) male and 5 (Δ2-3) females were set up and allowed to lay eggs for 3 days, the adults were then transferred to fresh vials and allowed to lay for 3 more days. Dysgenic males were crossed to *l(2)03605* females and *ry*$^+$ males collected. These males were crossed to *CyO* females for balanced stock generation.

In the re-mobilization lines crossed to the deficiency experiment 5 re-mobilized males/females were crossed to 5 deficiency containing females/males. Offspring were then scored for straight vs curly wings.

Results

Failed plasmid rescue indicates the *l(2)03605* P-element may be deleted

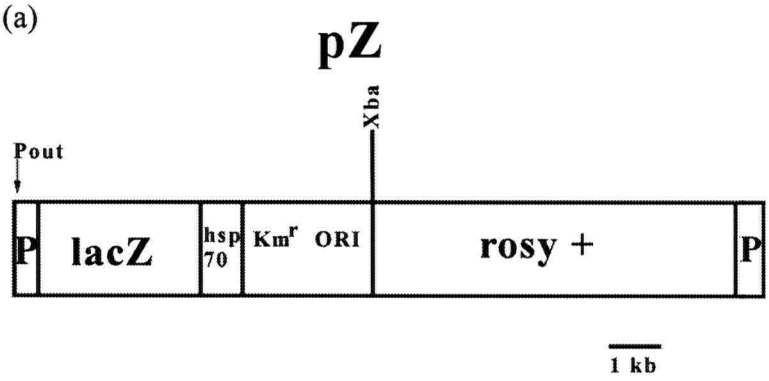
Plasmid rescue of the DNA flanking the *P* element insertion in strain *l(2)03605* was attempted in order to obtain DNA to fine localize the *P* element insertion site. Unfortunately the plasmid rescue experiment failed to yield any rescue plasmids. Genomic *l(2)03605/CyO* DNA was cut with enzymes suitable for rescuing the DNA flanking the *P* element, DNA was cut singly with *Xba* I and doubly with *Xba* I/*Nhe* I and *Xba* I/ *Spe* I. Cutting with *Xba* I/*Nhe* I or *Xba* I/ *Spe* I still allow for re-circularization of the plasmid, as the enzyme combinations produce compatible ends. Approximately 85 rescue plasmids should have been recovered (see Materials and Methods for calculation of this number), but a rescue plasmid was never recovered. This suggested that the *P* element had undergone some rearrangement or deletion that no longer allows for plasmid rescue.

The expected size of rescue plasmid should be > 7 kb as the unique *Xba* I site is ~ 7 kb from the 5′ end of the pZ element and the next restriction site is upstream of the *P* element within the genomic DNA (Figure 5a). To determine the actual size of the rescue plasmid, genomic Southern analysis was undertaken (Figure 5b). *l(2)03605* genomic DNA was cut with the same enzymes designed for the rescue: *Xba* I, *Xba* I/*Nhe* I and *Xba* I/*Spe* I, blotted and subsequently probed with a fragment of the *kanamycin* gene (343 bp *Pvu* I/ *Nru* I restriction fragment from Novagen's pET-28a(+)). The resulting bands in each digest are all less than 3 kb. Also, the *Xba* I/*Nhe* I digest shows two bands indicating a new *Nhe* I site within the pZ element. These results support the notion the *l(2)03605 P* element has undergone a rearrangement or deletion that disrupts plasmid rescue.

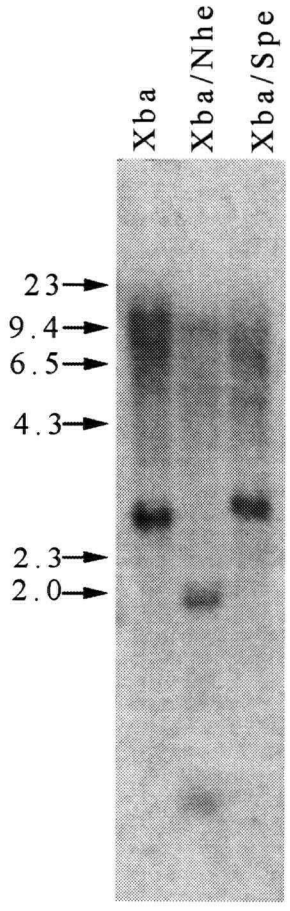
Figure 5

Determination of the size of rescue plasmid from *l(2)03605* flies. (a) Shows a map of the pZ plasmid with the unique *Xba* I restriction site ~ 7 kb from the 5´ end of the vector. (b) *l(2)03605* genomic Southern probed with a 343 bp *Pvu* I/ *Nru* I *kanamycin* gene fragment from Novagen's pET-28a(+). DNA was cut with enzymes designed for plasmid rescue: *Xba* I, *Xba* I/*Nhe* I and *Xba* I/*Spe* I. Southern shows bands less than the expected >7 kb, indicating the *P* element is deleted.

(a)


pZ
Xba
Pout
P
lacZ
hsp 70
Km^r ORI
rosy +
P
1 kb

(b)


Xba
Xba/Nhe
Xba/Spe
23
9.4
6.5
4.3
2.3
2.0

Polymerase chain reaction indicates that the *l(2)03605 P* element has not moved since its initial characterization

Since the plasmid rescue experiment was unsuccessful and the Genome Project managed to obtain the sequence of the DNA flanking the *l(2)03605 P* element, the incorrect fly line may have been used in my plasmid rescue experiments. To test this, a PCR experiment was undertaken to ensure the corrrect line had been obtained from the stock centre. A primer (P1342B) was designed from the Genome Project's sequence for use in a PCR experiment (Figure 6). Using this primer and a primer within the pZ element (Pout) should yield a product of 263 bp if the *P* element fly line is indeed the original *l(2)03605* line. Also, the experiment will confirm if the *P* element is in the original location determined by the Genome Project. The PCR experiment (Figure 7) yields the expected single 263 bp band (see lanes 2 and 3). The single primer controls failed to yield any products. This 263 bp product from a genomic primer and a *P* element specific primer indicates the *P* element has not moved and is in the Genome Project's suggested location and is therefore the correct *l(2)03605* Berkeley Genome Project characterized *P* element line.

The *l(2)03605 P* element induced recessive lethality is not caused by a disruption of the *HMG-Z* gene.

With both the enhancer trap *β-galactosidase* expression from the *P* element and the *HMG-Z* mRNA being confined to the nervous system the possibility the *P* element is disrupting the *HMG-Z* gene exists. A genomic Southern analysis of *l(2)03605/CyO* and Oregon-R (a wild-type strain) DNA using an *HMG-Z* cDNA as a probe was undertaken to

Figure 6

Sequence of DNA flanking the *P* element, including a primer (P1341B) designed for PCR. An *Eco* R1 restriction cut site was included in the primer for ease of future molecular manipulation of any PCR products. The 5´ end (nucleotide 1) is directly next to the *P* element. The sequencing was done by the Berkeley *Drosophila* Genome project from a successful plasmid rescue.

1 CTCCTGCCTT TNGTTATGTG TGTGAGGAGA TTGATATAGA ATGTGCTAAA

51 TAATACCAGG AATGACAGNC ACTCTACTTC CACTCTACAG ATCCGGCACT

101 CTAGCTCCTT NGTAGCCAAA TTAATTATCT ACCTCAATAT CAGNAGAAAA

151 CTAGGTATTT AAGAAGTTAT TTCAGTTGTT TGGANTCATT TTATTAATAT

201 GTTTNCGTTT CTTGAAGAGT TTGCGGTATT CTTCTGTCTT GAACAAATAN

251 TTGACAACTC GATTTAAGCG ACAACTTNGG TCTTTATAA CCTACATATGT

P1341B: 5´> GGA ATT CCG ACA GAA GAA TAC CGC AAA CTC < 3´

Figure 7

Polymerase chain reaction on *l(2)03605* genomic DNA using Pout and P1341B as primers. Lane 1 is a molecular weight marker, λ DNA cut with *Hind* III. Lanes 1,2,3,4 and 5 are the products of the PCR. Lanes 1 and 2 are identical and contain all the components required by the reaction. Lanes 3 and 4 are missing a single primer. An expected size band of 263 bp is shown in lanes 2 and 3.

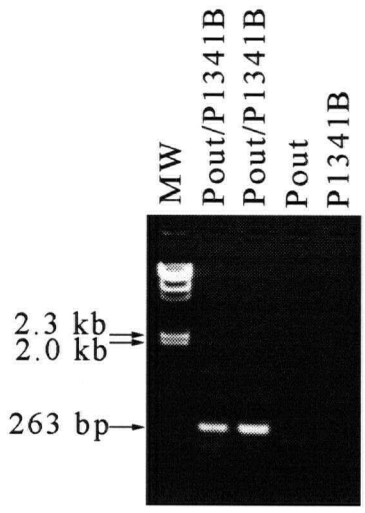
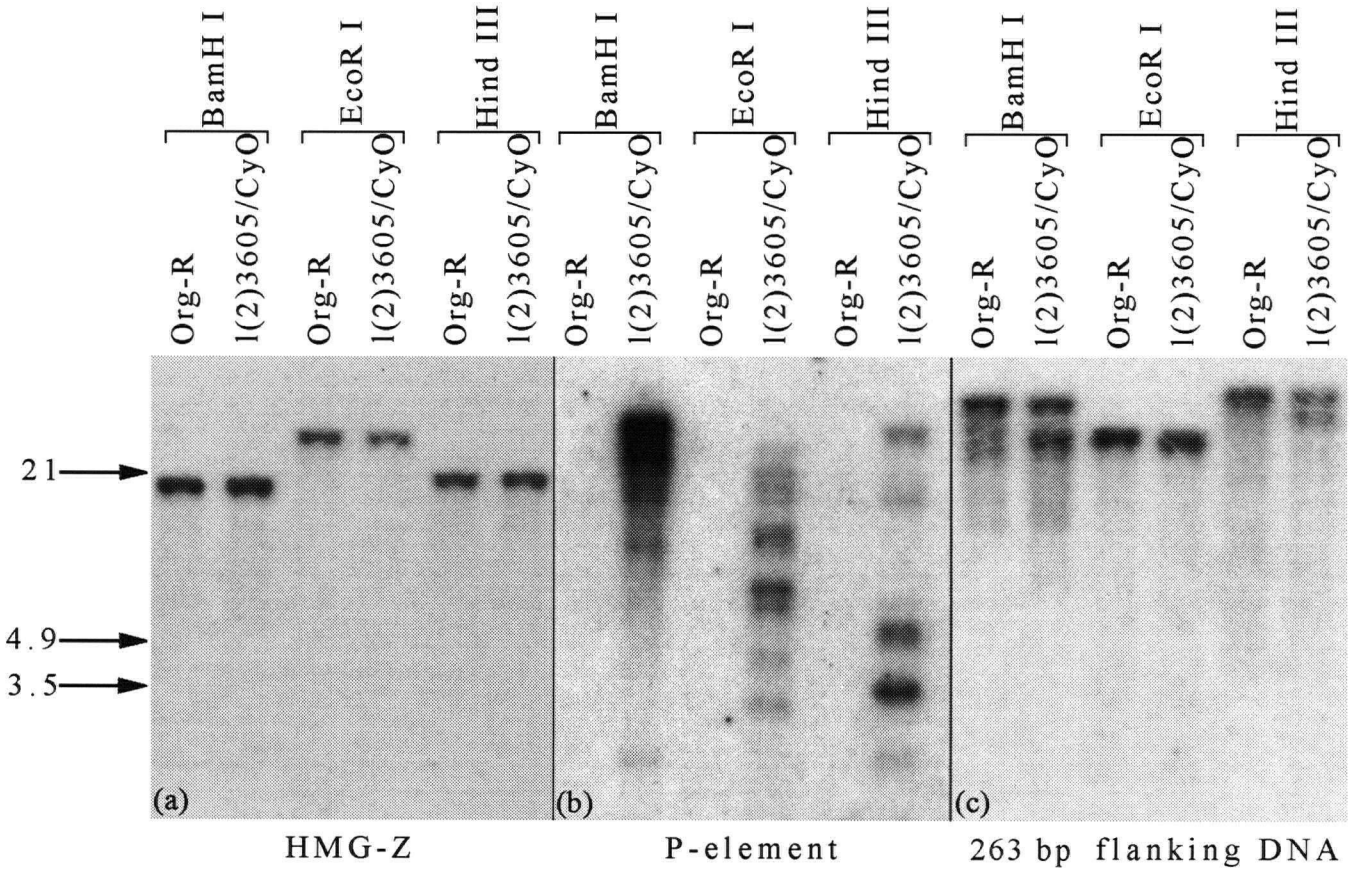

MW
Pout/P1341B
Pout/P1341B
Pout
P1341B
2.3 kb
2.0 kb
263 bp

Figure 8

Replicative genomic Southern of *l(2)03605* vs *wt* DNA cut with various restriction enzymes. (a) Blot probed with *HMG-Z* cDNA shows there is no difference between *l(2)03605* and *wt* DNA (b) Same blot probed with complete *P* element probe indicates the *l(2)03605* line contains *P* elements while the *wt* (Oregon-R) line does not. (c) Same blot probed with 263 bp flanking DNA indicates a polymorphism between the *l(2)03605* line and *wt*.


BamH I
EcoR I
Hind III
Org-R
l(2)3605/CyO
21
4.9
3.5
(a)
(b)
(c)
HMG-Z
P-element
263 bp flanking DNA

test this hypothesis (Figure 8a). Genomic DNA was cut with three restriction enzymes: *Bam* HI, *Eco* RI and *Hind* III, separated by electrophoresis, blotted and probed with a ^{32}P labeled *HMG-Z* cDNA. The autoradiograph indicates there is no polymorphism detected between the two fly lines. Therefore, it is unlikely the induced recessive lethality was caused by an insertion of a *P* element within the *HMG-Z* gene.

To confirm that the *l(2)03605* line contained *P* element sequences the same blot was stripped and re-probed with a probe made of the complete 2.9 kb *P* element (Figure 8b). The autoradiograph indicates as expected there are no *P* elements in the Oregon-R strain and *P* elements are present within the *l(2)03605* line. As a positive control to validate the 263 bp PCR product being composed of DNA flanking the *P* element, it was used as a probe on the previously probed and stripped blot (Figure 8c). This probe should show a polymorphism between the two fly lines. The blot indeed indicates a polymorphism is detected between the *l(2)03605* and wt lines. In both the *Bam* HI and *Hind* III digest the probe detected two bands in the *l(2)03605* line and only a single band in the wt (Oregon-R) fly. In both of these digests the upper band present in both the *wt* and *l(2)03605/CyO* line comes from the non-*P* element mutagenized chromosome. That is, in the *wt* case both chromosomes are "normal", but in the *l(2)03605/CyO* line the *CyO* chromosome is "normal" and the *l(2)03605* chromosome is mutant. Hence, the *CyO* and *wt* restriction cut patterns are the same. The novel band in the *l(2)03605/CyO* line is from the mutagenized chromosome. This polymorphism demonstrates that the 263 bp PCR product is next to the *l(2)03605 P* element, and that the 263 bp PCR product can detect a genomic polymorphism between *wt* and *l(2)03605* DNA.

The *P* element insertion site is contained within both a lambda clone for the *HMG-D/HMG-Z* genomic region and a P1 clone for the region.

To determine if the 263 bp flanking DNA and therefore the *P* element insertion site is contained within genomic clones for the region, Southern analysis was undertaken. A lambda clone (λG1) (Ner *et al.*, 1993) and various plasmid clones (Ner, unpublished) of the region were digested with *Eco* RI, blotted and probed with the 263 bp flanking DNA (Figure 9). The above clones for the 57F8-F11 genomic region were isolated in the initial cloning experiments of *HMG-D* and *HMG-Z* (Ner *et al.*, 1993; Ner and Travers, 1994). The autoradiograph indicates the 263 bp flanking DNA is not located within any of the plasmid clones, but is located within the lambda clone (λG1).

To further elucidate where the *P* element is inserted within the λG1 clone, a restriction digest analysis was completed. The λG1 clone was digested singly with *Bam*H I and *Sal* I, blotted and probed with the 263 bp flanking DNA. In conjunction a P1 phage clone (clone # DS086966-94-38) of the 57F8-11 region known to contain the *l(2)03605 P* element insertion site (mapped to this P1 clone by the Berkeley Genome Project) was digested with *Bam* HI and included on the blot (Figure 10). The autoradiograph indicates the *P* element insertion site is contained within a ~ 10 kb *Bam* HI fragment of the P1 clone, a ~ 9 kb *Bam* HI and a ~ 7-8 kb *Sal* I fragment. This places the *l(2)03605 P* element downstream of the *HMG-D* gene (Figure 11).

To confirm this location of the *P* element insertion site, genomic Southern analysis comparing *l(2)03605* and *wt* DNA cut with *Sal* I and hybridized with various probes was completed (Figure 12). The genomic blot was first probed with the 263 bp

Figure 9

Southern analysis of various λ and plasmid clones of the *HMG-D/HMG-Z* genomic region (Ner *et al.*, 1993) cut with *Eco* RI using the 263 bp flanking DNA as a probe. The blot indicates the flanking DNA and thus the *P* element insertion site is contained within the λG1 clone.

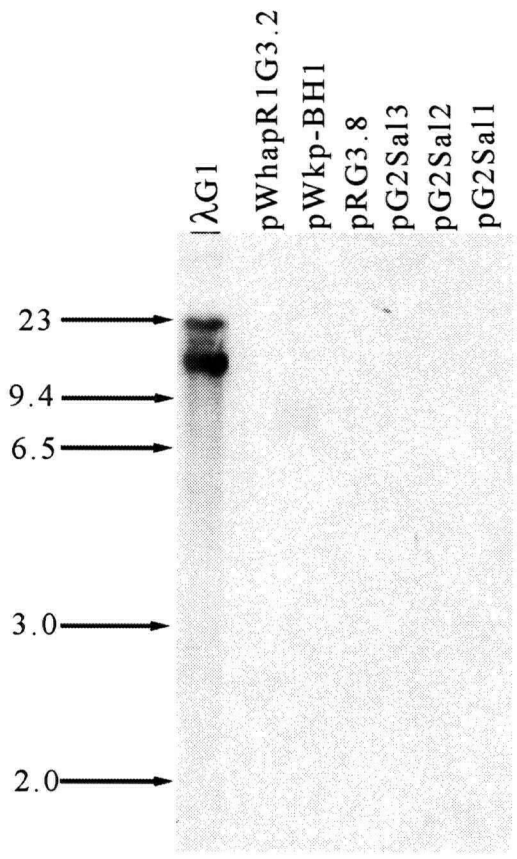

λG1
pWhapR1G3.2
pWkp-BH1
pRG3.8
pG2Sal3
pG2Sal2
pG2Sal1
23
9.4
6.5
3.0
2.0

Figure 10

Southern analysis of a P1 clone (DS086966-94-38) for the region cut with *Bam*H I and the λG1 (Ner *et al.*, 1993) clone cut with either *Bam*H I or *Sal* I using the 263 bp flanking DNA as a probe. The blot indicates the flanking DNA and thus the *P* element insertion site is contained within a ~ 10 kb *Bam*H I fragment of the P1 clone and a ~ 9 kb *Bam*H I and a ~ 7-8 kb *Sal* I fragment of the λG1 clone.

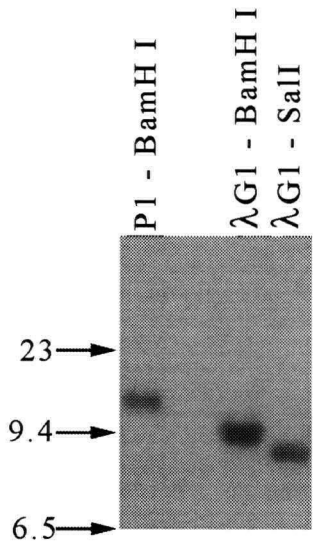
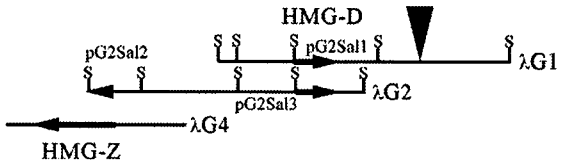

P1 - BamH I
λG1 - BamH I
λG1 - SalI
23
9.4
6.5

Figure 11

(a) A schematic diagram of the *HMG-D/HMG-Z* genomic region indicating the location of the *l(2)03605 P* element (Triangle), as well as some of the plasmid clones for the region from (Ner *et al.*, 1993). (b) A schematic diagram of the λG1 clone indicating the location of the *l(2)03605 P* element insertion site (Triangle), the *HMG-D* gene as well as the location of the plasmid clones used in the previous *P* element localization studies (Ner, unpublished). The map indicates the *l(2)03605 P* element insertion site is downstream of the *HMG-D* gene.

(a)


HMG-D
pG2Sal2
pG2Sal1
λG1
λG2
pG2Sal3
λG4
HMG-Z

(b)

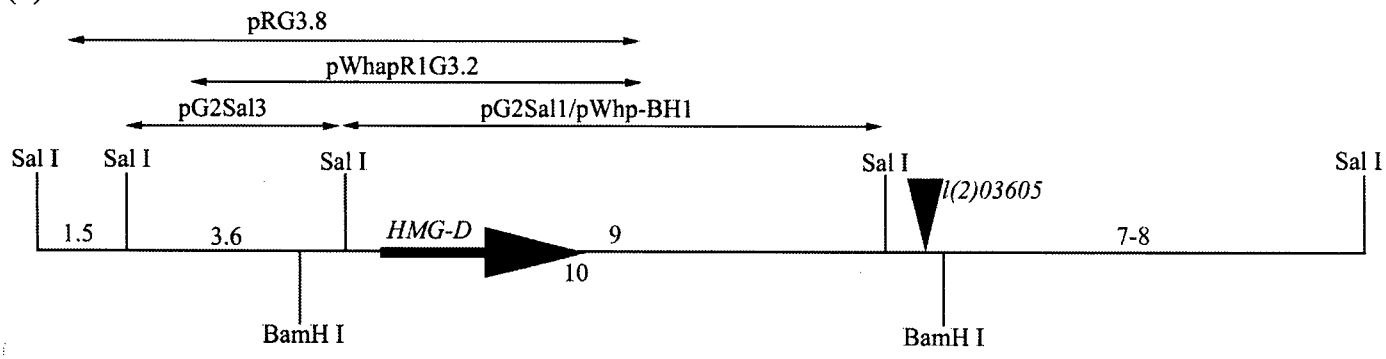
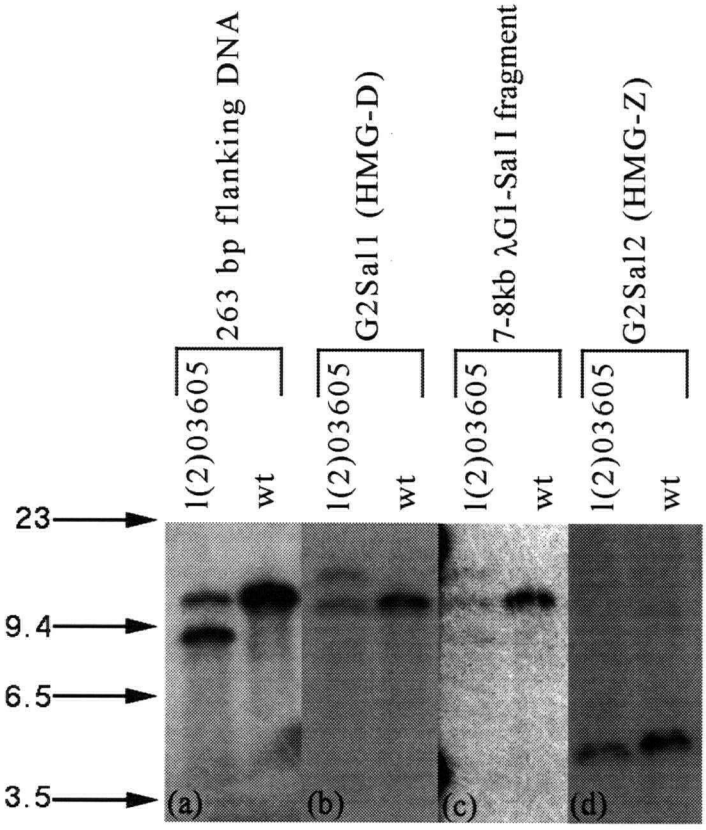

pRG3.8
pWhapR1G3.2
pG2Sal3
pG2Sal1/pWhp-BH1
Sal I
Sal I
Sal I
Sal I
Sal I
l(2)03605
1.5
3.6
HMG-D
9
10
7-8
BamH I
BamH I

Figure 12

Genomic southern analysis of *l(2)03605* vs *wt* DNA cut with *Sal* I using (a) 263 bp flanking DNA, (b) G2Sal1 (*HMG-D*), (c) 7-8 kb λG1-*Sal* I fragment and (d) G2Sal2 (*HMG-Z*) as probes. The blot indicates a polymorphism is uncovered between the *l(2)03605* and the *wt* line when using either the (a) 263 bp flanking DNA or (b) the G2Sal1 (*HMG-D*) clone as a probe. The blot also shows that when using the (c) 7-8 kb λG1-*Sal* I fragment as a probe the same bands light up as with the 263 bp flanking DNA and the G2Sal1 (*HMG-D*) clone probes. Thus, the 7-8 kb λG1-*Sal* I fragment spans both the 263 bp flanking DNA and the G2Sal1 (*HMG-D*) clone. Panel (d) indicates there is no difference between the *l(2)03605* and *wt* lines when using G2Sal2 (*HMG-Z*) as a probe.


l(2)03605
wt
263 bp flanking DNA
l(2)03605
wt
G2Sal1 (HMG-D)
l(2)03605
wt
7-8kb λG1-Sal I fragment
l(2)03605
wt
G2Sal2 (HMG-Z)
23
9.4
6.5
3.5
(a)
(b)
(c)
(d)

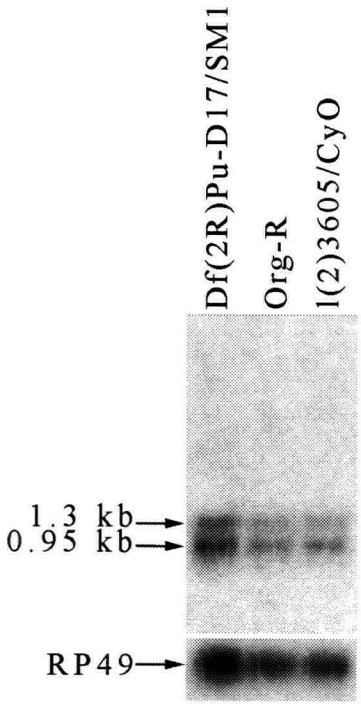
flanking DNA and this shows an expected polymorphism between the two fly lines (Figure 12a). The blot was stripped and re-probed with a plasmid clone of the region, G2Sal1 (see Figure 11). This clone contains the complete *HMG-D* ORF and also some downstream flanking DNA. The G2Sal1 clone likewise detects a polymorphism between the *l(2)03605* and *wt* lines. Since, the *P* element was localized to a 7-8 kb *Sal* I fragment of the λG1 clones (see Figure 11b) it was also used as a probe on the same blot (Figure 12c). This probe detects a polymorphism between the two lines and in fact detects the same bands as the 263 bp flanking DNA probe and the G2Sal1 probe. Therefore, this probe must span both of these DNA fragments. As a negative control the blot was re-probed with a plasmid clone that contains part of the HMG-Z gene (G2Sal2) (Figure 12d - see Figure 11 for location of plasmid clone). As expected this probe fails to detect a polymorphism between the two lines. This genomic Southern analysis confirms the *P* element insertion site is downstream of *HMG-D*.

The *l(2)03605 P* element does not affect *HMG-D* transcripts from adult females.

With the *P* element insertion site close to the *HMG-D* locus it might be disrupting expression of *HMG-D*. To test if the recessive lethality is associated with a reduction/elimination of *HMG-D* transcripts, total RNA from adult *l(2)03605/CyO*, Oregon-R and *Df(2R)Pu-D17/SM1* females was isolated and used in a Northern analysis with a ^{32}P labeled *HMG-D* cDNA as a probe (Figure 13). This same Northern blot was also probed with RP49 as a quantitative control for equal loading. The autoradiograph with the RP49 probe indicates the wt and *l(2)03605* RNA levels are approximately equal,

Figure 13

Northern analysis of total RNA isolated from adult females of *Df(2R)Pu-D17*, *Oregon-R* and *l(2)03605* using both a *HMG-D* cDNA and *RP49* as probes. The RP49 signal indicates that equal amounts of RNA were loaded from the *Oregon-R* and *l(2)03605* flies, however the *Df(2R)Pu-D17/SM1* lane is overloaded. The blot also indicates there is no difference in the two 1.3 kb and 0.95 kb *HMG-D* transcripts in RNA from *Oregon-R* and the *l(2)03605* line. Therefore, the *l(2)03605 P* element is not affecting the expression *HMG-D* transcripts.


Df(2R)Pu-D17/SM1
Org-R
l(2)3605/CyO
1.3 kb
0.95 kb
RP49

and thus there are no differences in the levels of either the 1.3 kb or 0.95 kb *HMG-D* transcripts. The deficiency lane is overloaded and therefore is not informative. Unfortunately, the Northern analysis indicates the *l(2)03605 P* element induced lethality is not caused by a reduction in *HMG-D* transcript levels.

Densitometric analysis was completed on the above Northern blot to ensure visual analysis of the band intensity was accurate (Figure 14). The numbers under the curves represent the area under that curve and is a measure of band intensity and thus RNA abundance. The densitometric analysis indicates the *wt* and *l(2)03605 HMG-D* RNA levels are equal with respect to both the 1.3 and 0.95 kb transcripts (Figure 14a,b). 4787 for *wt* vs 5185 for *l(2)03605* of the 1.3 kb transcript and 8458 for *wt* vs 8189 for *l(2)03605* of the 0.95 kb transcript. The RP49 quantitative control indicates there are equal levels of RNA in the *wt* and *l(2)03605* lanes, and that the deficiency lane is overloaded (24391 for *Df(2R)Pu-D17* vs 17169 for *wt* vs 15444 for *l(2)03605* of the RP49 transcript). Therefore, as stated previously there is no difference in the *HMG-D* RNA levels between *wt* and *l(2)03605* flies.

The *l(2)03605* recessive lethality is reversible and the newly generated re-mobilized lines fall into 3 categories: "precise" excisions, "imprecise" excisions and "local hops".

The close proximity of the *l(2)03605 P* element to the *HMG-D* gene, based on the mapping of the *P* element insertion site, may allow for the generation of new *HMG-D* alleles in a *P* element re-mobilization experiment. A mutant *HMG-D* allele could be generated if upon excision the *l(2)03605 P* element deleted flanking DNA into *the HMG-*

Figure 14

Densitometry analysis of Northern blot. Area plots are aligned as in the Northern from Figure 12. (a - b) Show the relationship between fly lines for the 1.3 kb and the 0.95 kb transcripts. The deficiency is twice the level of the *wt* and *l(2)03605* lines due to unequal loading (see RP49). The *wt* and *l(2)03605* lanes are essentially equal, indicating no difference between the mutant and wild-type. (c) Shows the RP49 control with the overloaded deficiency lane and the equally loaded *wt* and *l(2)03605* lanes.

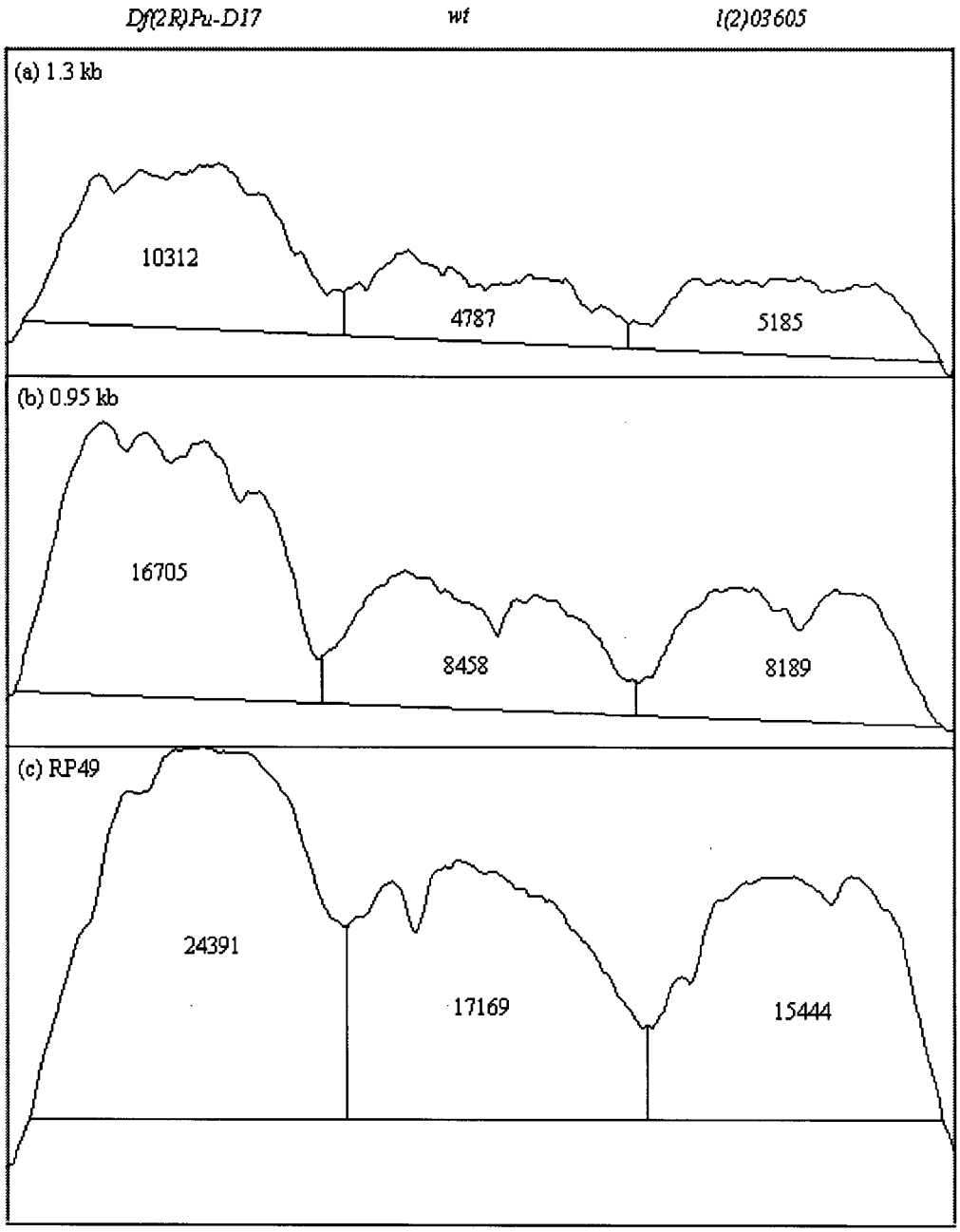
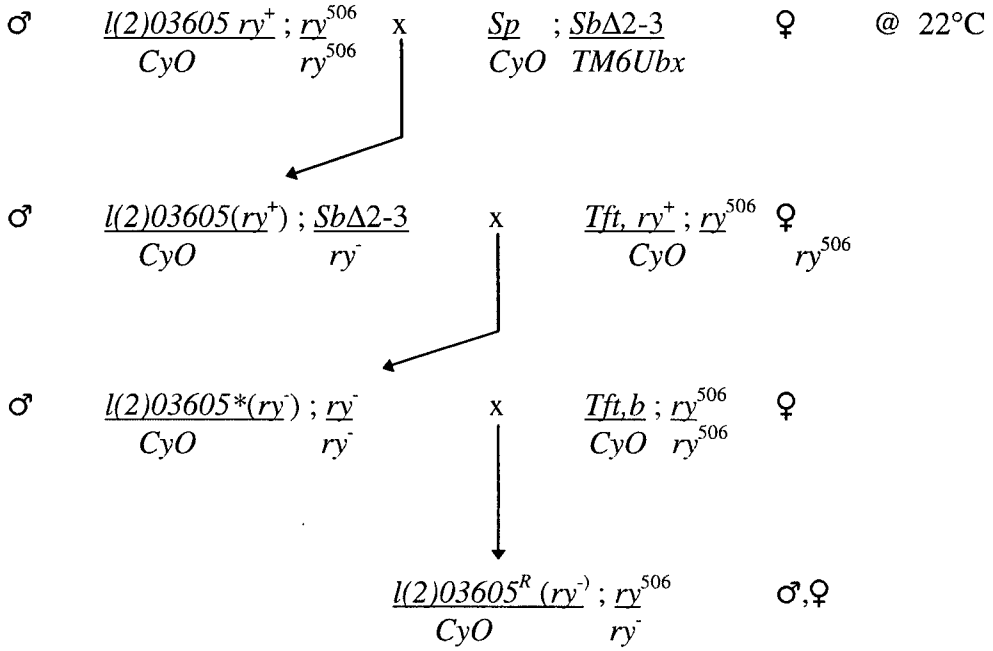

Df(2R)Pu-D17
wt
l(2)03605
(a) 1.3 kb
10312
4787
5185
(b) 0.95 kb
16705
8458
8189
(c) RP49
24391
17169
15444

Figure 15

Re-mobilization cross of *l(2)03605* line. *l(2)03605* males were crossed to a transposase source (Δ2-3) and dysgenic males were collected and then crossed to *l(2)03605* females. Male progeny were collected based on the loss of the *rosy* gene supplied by the pZ *P* element. *ry*⁻ males were then balanced over *CyO*.

Table 1

The re-mobilized lines were checked for recessive lethality. 4 lines are homozygous viable and therefore classified as "precise" excisions. 17 lines are still recessive lethal and have been classified as "imprecise" excisions that may have removed part of the flanking DNA surrounding the *P* element. * Homozygous flies were made of lines R15, R22 and R42. R46/R46 males are sterile and therefore a homozygous stock was not generated.



Homozygous viable lines*	Homozygous lethal lines
R15, R22, R42, R46	R1, R2, R8, R9, R10, R11, R16, R19, R20, R21, R28, R31, R33, R37, R38, R43, R47

D locus. This re-mobilization experiment began with the mobilization of the *l(2)03605 P* element and the generation of new lines. As shown is Figure 15, the *l(2)03605 P* element was re-mobilized as follows: *l(2)03605* males were crossed to female flies carrying a transposase source, *P*[Δ2-3*ry*$^{+}$](99B) abbreviated as Δ2-3 and single dysgenic male progeny were then crossed back to l*(2)03605* females. Females that had lost the *rosy*$^{+}$ phenotype supplied by pZ were collected and balanced over *CyO*. Twenty-one independently generated re-mobilization lines were established. Each line was then examined for straight (*l(2)03605***/l(2)03605**) and curly wing (*l(2)03605***/CyO*) phenotypes. Four homozygous viable lines were generated and 17 recessive lethal lines were recovered (Table 1). Those lines that contain straight winged flies are presumably revertants of the *l(2)03605 P* element induced lethality. To determine which *ry*$^{-}$ lines still had a disruption of the *l(2)03605* locus or were simply "local-hops" out of the *l(2)03605* locus and into a nearby essential gene, the lines were crossed to the deficiency *Df(2R)Pu-D17*, which is deleted for bands 57B4 - 58B (Table 2). The original *l(2)03605* line was crossed to the deficiency and as expected is lethal over *Df(2R)Pu-D17* (Table 2a). The 4 homozygous viable lines are all viable over the deficiency indicating they are "precise" excisions (Table 2b). Interestingly, the R46 line is only partially viable over the deficiency, depending on the direction of the cross. *R46/Df(2R)Pu-D17* flies are only recovered when the male parent contributed the deletion and the female contributes the re-mobilized *P* element. This line may be a partial revertant as homozygous R46 flies are rare and no progeny are recovered when homozygous R46 males are crossed to *wt* virgin females, indicating the homozygous R46 males are sterile. The re-mobilized lines that

Table 2

Re-mobilization lines were crossed to a deficiency for the region (*Df(2R)Pu-D17*). (a) Parental *l(2)03605* line crossed to the deficiency. As expected the transheterozygotes are not viable. (b) 4 of the lines are homozygous viable and viable over the deficiency and these are classed as "precise" excisions. (c) 5 lines that are recessive lethal but lethal over the deficiency were recovered, these lines are classed as "imprecise" excisions and may have deleted DNA flanking the *P* element. (d) 7 lines that are recessive lethal but viable over the deficiency are classed as "local-hops" and probably reflect the *P* element excising precisely and jumping into another essential gene not uncovered by the deficiency.

♂ *RXX* / *CyO* x *Df(2R)Pu-D17* / *SM1* ♀ & ♂ *Df(2R)Pu-D17* / *SM1* x *RXX* / *CyO* ♀

		Curly		Straight		Curly		Straight	
	Line	♂	♀	♂	♀	♂	♀	♂	♀
(a)	l(2)03605	22	26	0	0	21	18	0	0
(b)	R15	9	7	18	22	8	8	2	3
	R22	7	28	10	12	-	-	-	-
	R42	20	22	9	6	14	10	2	1
	R46*	13	14	0	0	22	20	1	3
(c)	R16	24	25	0	0	17	28	0	0
	R20	20	26	0	0	20	24	0	0
	R21	21	13	0	0	-	-	-	-
	R42	9	7	0	0	-	-	-	-
	R43	24	19	0	0	9	13	0	0
(d)	R8	20	20	8	10	17	21	14	15
	R10	26	24	20	7	6	5	3	0
	R28	29	24	15	10	8	12	2	5
	R31	21	31	11	22	16	8	4	4
	R33	20	19	26	22	20	15	3	7
	R37	13	24	19	13	17	15	3	4
	R47	19	18	4	8	12	19	15	14

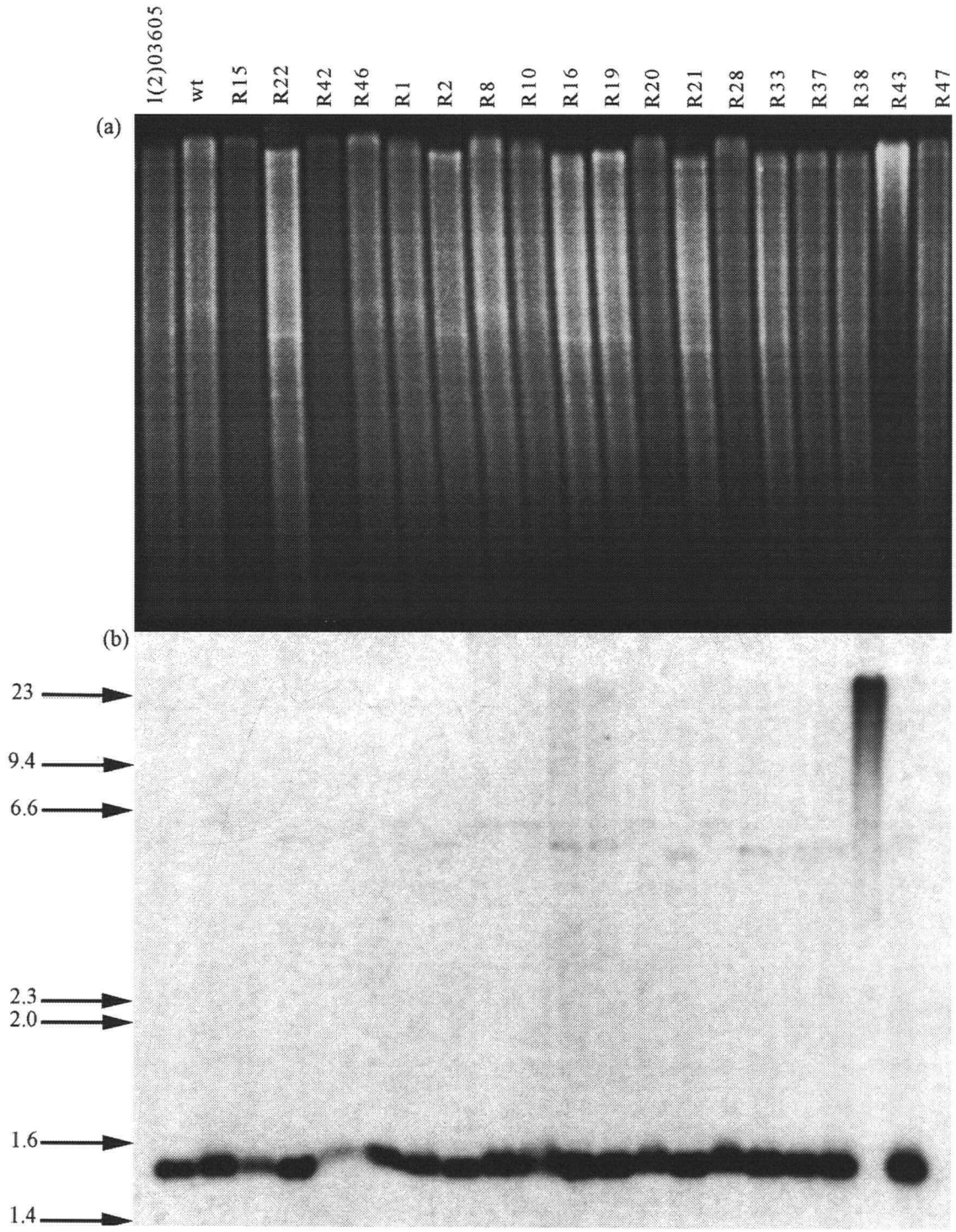
are still recessive lethal fall into two classes when crossed to the deficiency (Table 2c,d). Five of the lines are lethal over the deficiency and most likely represent an "imprecise" excision from the *l(2)03605* locus or simply a disruption of the ry^+ gene. The other 7 lines are all viable over the deficiency and probably represent a "local-hop" restoring the *l(2)03605* locus and disrupting another essential gene outside the boundaries of the deficiency. These "local-hops" are ry^-, showing that the mobilized element has suffered a partial deletion of the ry^+ gene.

The re-mobilized P-element fails to generate any new *HMG-D* alleles.

To determine if any of the re-mobilized lines deleted any of the *HMG-D* open reading frame, genomic Southern analysis of the original *l(2)03605*, *wt* and all the re-mobilization lines was undertaken. Genomic DNA was doubly digested with *Bgl* II and *Sal* I, which in wildtype flies produces a 1.5 kb fragment containing most of the *HMG-D* ORF, blotted and probed with a ^{32}P labeled cDNA for *HMG-D*. Probing with the *HMG-D* cDNA should reveal any polymorphisms between the parental l(2)03605 line and the re-mobilized lines, thus allowing detection of any deletions/re-arrangements created by mobilizing the *l(2)03605 P* element that disrupt the *HMG-D* coding region (Figure 16). The ethidium bromide stained gel is presented (Figure 16a) to show the wavy gel front which leads to a slight difference in mobility for a DNA of given size, which is also reflected in the autoradiograph (Figure 16b). The autoradiograph shows no polymorphisms between the original *l(2)03605*, *wt* and re-mobilized lines. However, re-mobilization line R43 was not digested completely and shows a higher band. This band

Figure 16

Genomic southern analysis of re-mobilization line DNA cut with *Bgl* II and *Sal* I which cuts out the *HMG-D* ORF. The blot is probed with an *HMG-D* cDNA. (a) Ethidium bromide stained gel prior to blotting indicates lines R15, R42 are under loaded and line R43 failed to cut. Note the gel front is wavy which leads to a slight difference in mobility for a DNA of given size, this is reflected in the autoradiogram. (b) The autoradiogram indicates there are no polymorphisms detected between the *wt*, *l(2)03605* and re-mobilization lines. Therefore, no new *HMG-D* alleles were created upon mobilization of the *l(2)03605 P* element. (Molecular weight values indicate migration of the molecular weight size markers and are thus not entirely accurate across the gel)


l(2)03605
wt
R15
R22
R42
R46
R1
R2
R8
R10
R16
R19
R20
R21
R28
R33
R37
R38
R43
R47
(a)
(b)
23
9.4
6.6
2.3
2.0
1.6
1.4

represents uncut DNA. Therefore no new *HMG-D* alleles were created in the *P* element re-mobilization experiment.

Discussion

Summary

A study was undertaken using a *P* element induced recessive lethal fly line obtained from the Berkeley *Drosophila* Genome Project to determine if this lethality was caused by disruption of either of the *HMG-D* or *HMG-Z* genes. In order to address this hypothesis the *l(2)03605 P* element insertion site had to be defined. The location of the *P* element insertion site, downstream of the *HMG-D* gene, was determined by mapping it to various genomic clones of the region. With the *P* elements close proximity to *HMG-D,* both (i) Northern analysis to determine the *P* element's affect on *HMG-D* transcript levels and (ii) re-mobilization of the P element in an attempt to generate novel mutant *HMG-D* alleles was undertaken. Unfortunately, the *l(2)03605 P* element does not disrupt either *HMG-D* or *HMG-Z* and no new *HMG-D* alleles were created.

The *l(2)03605* line is the same line initially characterized by the Genome Project

A plasmid rescue experiment using a Berkeley *Drosophila* Genome Project single *P* insert in the region where *HMG-D* and *HMG-Z* map was performed to obtain the genomic DNA next to the *P* element insertion. The failure to recover any rescue plasmids suggested that the *P* element may be deleted or re-arranged. Genomic Southern analysis (Figure 5b - *Kanamycin* gene probe) of the *l(2)03605 P* element fly line confirmed this.

In contrast, the BDGP had originally rescued the DNA flanking the *l(2)03605 P* element. The sequence was obtained in 1993 and the *l(2)03605 P* element stock was

maintained for a number of years until this project was started in 1997. Since *Drosophila* stocks must be maintained by replica/transfer rather than freezing, it is quite possible that a deletion or re-arrangement of the *l(2)03605 P* element occurred during that time interval.

The l(2)03605 P element does not disrupt either *HMG-D* or *HMG-Z*

The failure to recover *P* "rescue" plasmids from the l(2)03605 line not withstanding the goal of testing whether the *l(2)03605 P* element caused a knockout of *HMG-D* or *HMG-D* still remained. Evidence that led to this initial hypothesis was the coincident cytological location (57F8-11) of *HMG-D/Z* and the *P* element insertion site and the similar mRNA expression pattern of both in the developing nervous system. The hypothesis was a gamble, but one that was worth accepting. Unfortunately, the *l(2)03605 P* element does not disrupt either *HMG-D* or *HMG-Z* . What makes the primary premise of this hypothesis (a gene of interest mutated to lethality by a transposable element) such a gamble are outlined below: (i) the number of essential genes in *Drosophila* (ii) the susceptibility of these gene to disruption by insertional mutagenesis and finally (iii) the extent to which these essential genes have been mutated to saturation.

It is estimated that there are between 8000 to 15 000 genes in the *Drosophila* genome (Rubin, 1998). Approximately 3600 of these are essential loci based on the Miklos and Rubin (1996) estimate of 12 000 genes. The BDGP strain library has so far disrupted about one-third of the essential loci located on the autosomes. Specifically, 1182 *P* element tagged lines exist representing approximately 31% of the 3200 autosomal genes that can mutate to a distinguishable phenotype (Rubin, 1998). Thus, the probability that a particular gene of interest is disrupted by a *P* element is about 33%. However,

more and more lines are being generated and this will increase the feasibility of tagging susceptible loci.

The *l(2)03605 P* element is likely disrupting an unidentifed gene

Even though the *l(2)03605 P* element did not insert within *HMG-D/Z*, it did insert less than 10 kb downstream of *HMG-D* (see Figure 11). Since the *P* insert mutation has a recessive lethal phenotype, the *l(2)03605 P* element is likely affecting another gene downstream of *HMG-D*. There are two candidate genes within the 57F8-11 region, however neither of these appear to have expression confined to the embryonic nervous system. These genes are: *TATA Binding Protein* (*TBP*) at cytological map position 57F8-10 (Hoey *et al.*, 1990) and *Enhancer of eye roughness* (*ER2-5*) at cytological map position 57F-58A (Karim *et al.*, 1996). Therefore the *l(2)03605 P* element induced lethality is probably caused by disruption of a novel gene.

Re-mobilization of the *l(2)03605 P* element failed to generate new *HMG-D* alleles

The *l(2)03605 P* elements insertion site downstream of *HMG-D* may have allowed for the generation of new *HMG-D* alleles by *P* element imprecise excision. Unfortunately, no new *HMG-D* alleles were created as the re-mobilization screen was initially carried out as a reversion analysis and not enough flies were screened for imprecise excisions. The reversion screen was successful and four homozygous lines were generated (Figure 15, Table 1). If a more extensive screen is carried out, in the order of 7000 screened chromosomes (Salz and Schedl, 1987; Tsubota and Schedl, 1986), deleted *HMG-D* alleles should be recovered.

Conclusion

Since the lethality associated with the *l(2)03605 P* element insert does not result from a disruption of either *HMG-D* or *HMG-Z,* the question remains "what gene is the *P* insert affecting?". Using a P1 clone of the region as a probe on a comparative developmental Northern of *l(2)03605* vs *wild type* flies is a simple way to determine (i) if there is one or more transcribed genes in the region and (ii) if the *P* element is causing lethality by changing the expression pattern of some transcripts. Hopefully, with an observed difference it should be a simple matter of using fragments of the P1 clone as probes on the Northern to narrow down the location of the gene of interest. With the fragment identified, the DNA can be sequenced and compared with known genes to determine if the gene has been previously identified or is a novel one. Also, the P1 fragment can be used to screen a cDNA library to identify cDNA clones of the affected gene. Unfortunately this project did not provide any insight into HMG proteins, but with the original *l(2)03605 P* element induced mutation and the newly generated revertants, "partial" revertants and "imprecise" excisions this thesis may lead the way to characterizing an essential gene involved in the developing nervous system.

References

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. and Watson, J.D. (1994) *Molecular Biology of The Cell.* Garland Publishing, Inc.

Anderson, K. and Lenhyel, J.A. (1980) Changing Rates of Histone mRNA Synthesis and Turnover in *Drosophila* Embryos. *Cell*, **21**, 717-727.

Balaeff, A., Churchill, M.E. and Schulten, K. (1998) Structure prediction of a complex between the chromosomal protein HMG-D and DNA. *Proteins*, **30**, 113-35.

BDGP. (1996) *P* element Gene-disruption Project. *Berkeley Drosophila Genome Project.*

Bustin, M., Trieschmann, L. and Postnikov, Y.V. (1995) The HMG-14/-17 chromosomal protein family: architectural elements that enhance transcription from chromatin templates. *Semin Cell Biol*, **6**, 247-55.

Bustin, R. and Reeves, R. (1996) High-Mobility-Group Chromosal Proteins: Architectural Components the Facilitate Chromatin Function. *Progress in Nucleic Acid Reserach and Molecular Biology*, **54**, 35-100.

Churchill, M.E., Jones, D.N., Glaser, T., Hefner, H., Searles, M.A. and Travers, A.A. (1995) HMG-D is an architecture-specific protein that preferentially binds to DNA containing the dinucleotide TG. *Embo J*, **14**, 1264-75.

Dow, L.K., Changela, A., Hefner, H.E. and Churchill, M.E.A. (1997) Oxidation of a critical methionine modulates DNA binding of the *Drosophila melanogaster* high mobility group protein, HMG-D. *FEBS*, **414**, 514-520.

Falvo, J.V., Thanos, D. and Maniatis, T. (1995) Reversal of intrinsic DNA bends in the IFN beta gene enhancer by transcription factors and the architectural protein HMG I(Y). *Cell*, **83**, 1101-11.

Griffiths, A.J.F., Miller, J.H., Suzuki, D.T., Lewontin, R.C. and W.M., G. (1993) *An Introduction to Genetic Analysis*. W.H. Freeman and Company, New York, NY.

Grosschedl, R., Giese, K. and Pagel, J. (1994) HMG domain proteins: architectural elements in the assembly of nucleoprotein structures. *Trends Genet*, **10**, 94-100.

Hoey, T., Dynlacht, B.D., Peterson, M.G., Pugh, B.F. and Tjian, R. (1990) Isolation and Characterization of the Drosophila Gene Encoding the TATA Box Binding Protein, TFIID. *Cell*, **61**, 1179-1186.

Johns, E.E. (1964) Studies on histones. VII. Preparative methods for histone fractions from calf thymus. *Biochem. J.*, **92**, 55-59.

Johns, E.W. (1982) *The HMG Chromosomal Proteins.* Academic Press, London.

Jones, D.N., Searles, M.A., Shaw, G.L., Churchill, M.E., Ner, S.S., Keeler, J., Travers, A.A. and Neuhaus, D. (1994) The solution structure and dynamics of the DNA-binding domain of HMG-D from *Drosophila melanogaster*. *Structure*, **2**, 609-27.

Karim, F.D., Chang, H.C., Therrien, M., Wassarman, D.A., Laverty, T. and Rubin, G.M. (1996) A Screen for Genes That Function Downstream of Ras1 During *Drosophila* Eye Development. *Genetics*, **143**, 315-329.

Kinoshita, M., Hatada, S., Asashima, M. and Noda, M. (1994) HMG-X , a *Xenopus* gene encoding an HMG1 homolog, is abundantly expressed in the developing nervous system. *FEBS letters*, **352**, 191-196.

Landsman, D. and Bustin, M. (1991) Assessment of the transcriptional activation potential of the HMG chromosomal proteins. *Mol Cell Biol*, **11**, 4483-9.

Landsman, D. and Bustin, M. (1993) A signature for the HMG-1 box DNA-binding proteins. *Bioessays*, **15**, 539-46.

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

Miklos, G.L.G. and Rubin, G.M. (1996) The Role of the Genome Project in Determining Gene Function: Insights from Model Organisms. *Cell*, **86**, 521-529.

Mlodzik, M. and Hiromi, Y. (1992) Enhancer trap method in *Drosophila*: its application to neurobiology. In Conn, P.M. (ed.) *Methods in Neurosciences, Vol 9: Gene Expression in Neural Tissues*. Academic Press, New York, Vol. 9, pp. 397-414.

Ner, S. (1992) HMGs everywhere. *Current Biology*, **2**, 208-210.

Ner, S.S., Churchill, M.E., Searles, M.A. and Travers, A.A. (1993) dHMG-Z, a second HMG-1-related protein in *Drosophila melanogaster. Nucleic Acids Research*, **21**, 4369-71.

Ner, S.S. and Travers, A.A. (1994) HMG-D, the *Drosophila melanogaster* homologue of HMG 1 protein, is associated with early embryonic chromatin in the absence of histone H1. *Embo J*, **13**, 1817-22.

Nightingale, K., Dimitrov, S., Reeves, R. and Wolffe, A.P. (1996) Evidence for a shared structural role for HMG1 and linker histones B4 and H1 in organizing chromatin. *Embo J*, **15**, 548-61.

Reeves, R. and Wolffe, A.P. (1996) Substrate structure influences binding of the non-histone protein HMG- I(Y) to free nucleosomal DNA. *Biochemistry*, **35**, 5063-74.

Rubin, G.M. (1998) The *Drosophila* genome project: a progress report. *Trends in Genetics*, **14**, 340-342.

Salz, H.K., Cline, T.W. and Schedl, P. (1987) Functional Changes Associated with Structural Alterations Induced by Mobilization of a *P* element Inserted in the *Sex-lethal* Gene of *Drosophila. Genetics*, **117**, 221-231.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning - A laboratory manual.* Cold Spring Harbor Laboratory Press.

Sinclair, A.H., Berta, P., Palmer, M.S., Hawkins, J.R., Griffiths, B.L., Smith, M.J., Foster, J.W., Frischauf, A.M., Lovell-Badge, R. and Goodfellow, P.N. (1990) A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature*, **346**, 240-244.

Spradling, A.C., Stern, D., Kiss, I., Roote, J., Laverty, T. and Rubin, G.M. (1995) Gene disruptions using *P* transposable elements: an integral component of the *Drosophila* genome project. PNAS, **92**:10824-30.

Stroumbakis, N.D. and Tolias, P.P. (1994) Localized maternal and zygotic expression of the gene encoding *Drosophila* HMG D. *Biochim Biophys Acta*, **1218**, 245-9.

Thanos, D. and Maniatis, T. (1995) Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome. *Cell*, **83**, 1091-100.

Travis, A., Amsterdam, A., Belanger, C. and Grosschedl, R. (1991) LEF-1, a gene encoding a lymphoid-specific protein with an HMG domain, regulates T-cell receptor enhancer function. *Genes Dev*, **5**, 880-894.

Tsubota, S. and Schedl, P. (1986) Hybrid Dysgenesis-Induced Revertants of Insertions at the 5′ End of the *Rudimentary* gene in *Drosophila Melanogaster*: Transposon Induced Control Mutations. *Genetics*, **114**, 165-182.

Ura, K., Nightingale, K. and Wolffe, A.P. (1996) Differential association of HMG1 and linker histones B4 and H1 with dinucleosomal DNA: structural transitions and transcriptional repression. *Embo J*, **15**, 4959-69.

Van Holde, K., Zlatanova, J., Arents, G. and Moudrianakis, E. (1995) Elements of chromatin structure: histones, nucleosomes, and fibres. In Elgin, S.C.R. (ed.) *Chromatin Structure and Gene Expression.* Oxford University Press, New York.

Voelker, R.A., Greenleaf, A.L., Gyrurkovics, H., Wisely, G.B., Huang, S. and Searles, L.L. (1984) Frequent imprecise excisions among reversions of *P* element caused lethal mutations in *Drosophila. Genetics*, **107**, 279-294.

Wagner, C.R., Hamana, K. and Elgin, S.C. (1992) A high-mobility-group protein and its cDNAs from *Drosophila melanogaster. Mol Cell Biol*, **12**, 1915-23.

Wolffe, A.P. (1997) Histone H1. *Int. J. Biochem. Cell. Biol.*, **29**, 1463-1466.