

To the Graduate Council:

I am submitting herewith a dissertation written by Rajesh Krishnan entitled "Molecular Characterization of an X1R-linked, Retina-specific and Dosage-compensated Gene of *Drosophila miranda*". I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Life Sciences.

Mary Ann Handel

Dr. Mary Ann Handel, Major Professor

We have read this dissertation
and recommend its acceptance:

Ranjan Garguly

David Brian

John W. Kowitz

Accepted for the Council:

C. W. Mink

Associate Vice Chancellor
and Dean of The Graduate School

MOLECULAR CHARACTERIZATION OF AN X1R-LINKED,
RETINA-SPECIFIC AND DOSAGE-COMPENSATED GENE
OF *DROSOPHILA MIRANDA*

A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Rajesh Krishnan

May 1992

ACKNOWLEDGEMENTS

I wish to express my utmost appreciation to my major professor, Dr. Ranjan Ganguly, for his support, guidance and friendship.

I also wish to thank the members of my advisory committee, Dr. Mary Ann Handel, Dr. John Koontz and Dr. David Brian for guidance and encouragement. I have derived much comfort from the unfailing friendship of Drs. Phiroze and Manju Sethna and Dr. Sivadasan Kanangat.

Finally, I would also wish to thank Dr. Nivedita Ganguly, Kausik and Sheila Ray, Ken Swanson, Susan Dombrowski, Mark Shannon, Cynthia Park, Guy and Kim Caldwell, Vivek Korde and P. Joshi for their friendship, discussions and jokes.

ABSTRACT

Dosage compensation is a gene regulatory mechanism that equalizes the expression of X-linked genes in the two sexes despite a two-fold higher X-linked gene dosage in the females relative to males. In *Drosophila*, dosage compensation is accomplished through the transcriptional hyperactivation of X-linked genes in males to approximately twice that in females. The transcriptional hyperactivation is believed to be mediated through the interaction between cis-acting sequences of the X chromosome and trans-acting factors that appear to act in a male-limited manner. One of the major questions arising from this model has been the identity and properties of cis-acting dosage-compensatory sequences. So far no discrete sequence that mediates male-specific transcriptional hyperactivation has been identified. The study reported in this dissertation is the initiation of a project, the long term goal of which is to identify and characterize cis-acting sequences that mediate male-specific transcriptional hyperactivation. The focus of this study has been the X1R chromosome of *D. miranda*, since the homolog of this chromosome in *D. melanogaster* is autosomal (3L) and therefore does not face the problem of dosage effect in the two sexes. This lends an evolutionary perspective to this study, since it can be expected that the nature of modifications that a gene is subjected to due to its X-linked and dosage compensated status may be highlighted by a comparison with a homolog that is not X-linked and under no pressure to overcome a dosage difference in the two sexes. This dissertation describes the characterization of an X1R-linked gene of *D. miranda* with the major objective of establishing it as a system which could be subsequently exploited in the detailed characterization of cis-acting sequences that mediate dosage compensation in *Drosophila*. The results of this study establish that a retina-specific gene, 507, which is 3L-linked in *D. melanogaster*, is a potentially good system with which to address some specific questions regarding dosage compensatory cis-acting sequences: What are they? What is their physical and functional relationship with other cis-acting transcriptional control elements that influence gene transcription? In addition it is expected that a comparative study of gene 507 of *D. miranda* and *D. melanogaster* may address the question of how dosage-compensatory cis-acting sequences evolve. In this study, gene

507, was shown to be X1R-linked and dosage-compensated in *D. miranda*. In order to initiate the mapping of its transcriptional control regions, gene 507 of *D. miranda*, *507mr*, was introduced into the *D. melanogaster* genome at autosomal sites. The transgenic *507mr* is expressed in the retina-specific manner appropriate for this gene and furthermore, the steady-state levels of transgene-derived mRNA were 1.4-1.6 fold higher in males, indicating a partial response to the dosage-compensatory machinery of *D. melanogaster*. A nucleotide sequence comparison of the transgene with that of the *D. melanogaster* homolog has highlighted some interesting features which may be involved in male-specific transcriptional hyperactivation of this gene. In addition, the conceptual translational product of this gene was found to be a homolog of the vertebrate visual system protein, arrestin, which plays an important role in visual transduction.

TABLE OF CONTENTS

CHAPTER 1. GENERAL INTRODUCTION	1
Sex Chromosomes and Dosage Compensation	1
Mechanisms of Dosage Compensation	3
Dosage compensation in <i>Drosophila</i> : Transcriptional hyperactivation of the male	
X chromosome relative to the female X chromosome	5
Models For Sex Specific Modulation of Transcription of X-Linked Genes	6
X-chromatin conformation in the two sexes	11
Cis-acting Elements and Trans-acting Factors in Dosage Compensation	12
The X1R Chromosome of <i>Drosophila miranda</i> : a model to study the mechanism of dosage	
compensation in <i>Drosophila</i>	14
Project Objectives and Significance	17
 CHAPTER 2. DOSAGE COMPENSATION OF AN X1R-LINKED RETINA-SPECIFIC GENE IN	
<i>DROSOPHILA MIRANDA</i>	20
INTRODUCTION	20
MATERIALS AND METHODS	22
RESULTS	23
Isolation of clones 507ml and 507mr	23
Transcription mapping of gene 507	25
Chromosomal locations of genes 507ml and 507mr	28
Dosage compensation of gene 507 in <i>D.miranda</i>	28
Tissue-specific expression of gene 507	33
Analysis of homology between clones 507ml and 507mr	37
DISCUSSION	46
 CHAPTER 3. RETINA-SPECIFIC EXPRESSION AND MALE-SPECIFIC	
TRANSCRIPTIONAL HYPERACTIVATION OF GENE 507 OF <i>D. MIRANDA</i>	
WHEN EXPRESSED FROM AUTOSOMAL SITES IN <i>D. MELANOGASTER</i> ..	49
INTRODUCTION	49
MATERIALS AND METHODS	54
RESULTS	57
Transgene-specific probe does not cross-hybridize with 507 mRNA of host	
strain	57
Transformant lines 1A, 6A and 10D have a single autosomal insertion of the	
intact transgene.	57
Transgene 507mr is transcribed only in the retina of the adult compound	
eye	60
Quantitation of steady-state level of transgene derived 507 mRNA in the two	
sexes	65
DISCUSSION	69
 CHAPTER 4. SEQUENCE COMPARISON OF GENES 507MR AND 507ML HIGHLIGHTS	
DOMAINS POTENTIALLY SIGNIFICANT IN REGULATION AND FUNCTION	74

INTRODUCTION	74
MATERIALS AND METHODS	76
RESULTS	77
Alignment of nucleotide sequences of gene <i>507ml</i> and <i>507mr</i>	77
Comparison of 5' non-coding sequences of gene <i>507ml</i> and <i>507mr</i>	79
Comparison of coding regions and introns	87
Comparison of 3' non-coding regions	92
DISCUSSION	95
CHAPTER 5. GENERAL DISCUSSION	97
BIBLIOGRAPHY	103
APPENDIX I	118
I. Screening of genomic and cDNA libraries	118
II. Nucleic acids isolation	119
III. Gel electrophoresis and blot transfer of nucleic acids	122
IV. Radiolabelling of nucleic acid probes	124
V. In situ hybridization to tissue sections	125
VI. In situ hybridization to polytene chromosomes	125
VII. P-element mediated transformation	126

LIST OF TABLES

TABLE	PAGE
1. Dot blot quantitation* of Adh mRNA . I	32
2. Dot blot quantitation* of Adh mRNA . II	32
3. Dosage compensation of gene 507 in <i>D. miranda</i> . I.	35
4. Dosage compensation of gene 507 in <i>D. miranda</i> . II.	36
5. Chromosomal locations of transgene 507mr	59
6. Steady-state levels of 507mr mRNA in transformants	67
7. Steady-state levels of endogenous 507ml mRNA in transformants	68

LIST OF FIGURES

FIGURE	PAGE
1. Quantitation of X-linked gene activity in heteroploids	8
2. Quantitation of X-linked gene activity in segmental aneuploids	10
3. Origin of karyotypes of <i>D. melanogaster</i> and <i>D. miranda</i>	15
4. Restriction maps of clones 507, 507ml and 507mr	24
5. Northern analysis of clone 507ml and 507mr	26
6. Chromosomal locations of genes 507ml and 507mr	29
7. Dot-blot analysis of Adh mRNA levels in male and female <i>D. miranda</i>	31
8. Quantitation of steady-state level of 507 mRNA in male and female <i>D. miranda</i>	34
9. Northern analysis of <i>eym</i> (eye missing) and <i>glass</i> flies	38
10. Localization of 507 mRNA to the retina of the adult compound eye	39
11. Interspecies Southern-blot hybridization A.	40
12. Interspecies Southern-blot hybridization B.	43
13. Interspecies northern-blot hybridization	44
14. Summary of interspecies cross-hybridization experiments	45
15. The genomic fragments containing the gene 507 of <i>D. melanogaster</i> and <i>D.</i> <i>miranda</i>	55
16. Northern-blot analysis of <i>ry</i> ⁵⁰⁶ flies with transgene specific probe	58
17. Southern-blot analysis of transformant lines	61
18. Qualitative northern-blot analysis of transformant lines	62
19. Localization of transgene-derived mRNA to retina of transformants	63
20. Quantitative sex-specific northern-blot analysis of transformant lines	66
21. Genes 507ml and 507mr	78
22. Comparison of 5' non-coding sequences of genes 507mr and 507ml	80

FIGURE		PAGE
23.	Potential stem and loop forming sequence in the putative promoter region of gene <i>507mr</i>	85
24.	Comparison of coding regions and introns	88
25.	Comparison of regions downstream of the translational stop codon	93

CHAPTER 1

GENERAL INTRODUCTION

SEX CHROMOSOMES AND DOSAGE COMPENSATION

In many sexually reproducing organisms, certain chromosomes are either limited to a particular sex or are present in varying numbers in the two sexes. Such chromosomes for which the two sexes are dimorphic are referred to as sex chromosomes. Examples are the Z and W chromosomes in birds and the X and Y chromosomes in mammals and *Drosophila*. In the XY system, the X and Y chromosomes are morphologically distinguishable from each other and are referred to as heteromorphic chromosomes. In mammals and in *Drosophila* the female carries two homomorphic sex chromosomes, the X chromosomes, and the male carries two heteromorphic sex chromosomes, the X and the Y. Models describing the evolution and establishment of heteromorphic sex chromosomes are based on the idea that the chromosomes have evolved from a single ancestral element (Bull 1983 and references therein). In the XY system, for example, the proto Y chromosome is believed to have lost, through a process of degenerative evolution, alleles of most X-linked genes and is now a largely heterochromatic chromosome bearing a very small number of active genes (reviewed by Bull 1983 ; Jablonka and Lamb 1990; Charlesworth 1991). This loss of alleles of X-linked genes from the Y chromosome in turn results in an imbalance of X linked gene dosage in the two sexes, because the males have one dose and females have two doses of the X chromosome. However, as early as the 1920s, many *Drosophila* geneticists were attaching considerable significance to the observations that, despite this dosage difference, the phenotypic expression of most X linked genes were similar in the two sexes (reviewed by Stern 1960 ; Luchesi and Manning 1987). Muller (1931) examined the expression of an X-linked hypomorphic mutation *white apricot* (w^a) which affects eye pigmentation and found that homozygous ($X^{w^a} X^{w^a}$) females and hemizygous ($X^{w^a} Y$) males had similar eye pigmentation. On the

other hand a female with one dose of the w^a allele and a deficiency for this locus on the other chromosome showed a lower level of pigmentation than the male hemizygous for w^a . The implication that the single w^a allele in the males represented the equivalent of more than a single w^a allele in the female led Muller to suggest the existence of a mechanism that accounts for the phenotypic equivalence of X-linked genes which are in one dose in males and in two doses in females (Muller 1931 ; 1950). He called this mechanism dosage compensation and predicted that it must operate at most X-linked loci and that although studies of this mechanism were necessarily limited to mutant alleles, it must have evolved to operate on wild type loci. Stern (1929 ; Stern 1960 for review) proposed a hypothesis to explain the evolution of such a mechanism which acts to neutralize X linked gene dosage difference in the two sexes. Stern (1929) invoked the principles of natural selection in his proposal that a consequence of having different doses of X-linked genes would be differences in selectiveness between the two sexes which in turn could be disastrous for a species. A simplistic interpretation of this would be that an extra dose of X-linked genes could make females "more fit" than males, resulting in a preferential "survival" of females, a situation that would affect the propagation of a species. Interestingly, he successfully tested one of the predictions of this hypothesis. The dosage-compensatory mechanism need not evolve to operate at an X-linked locus that is in equivalent dosage in the two sexes. His studies (Stern 1929) of a locus, *bobbed* (*bb*), that is present both on the X and the Y chromosome, showed that the expression of bobbed in two dose females (X^bX^b) was twice that in one dose males (X^bO), implying that the dosage compensatory mechanism had not evolved at this locus. The evolution of a dosage compensatory mechanism is thus seen to be an adaptation to the establishment of heteromorphic sex chromosomes, and in the XY system it can be said, albeit arguably, that the dosage-compensation mechanism operating on the X chromosomes has co-evolved with the functional degeneration of its homologue, the Y chromosome. Dosage compensation, first demonstrated in *Drosophila*, has since been demonstrated in mammals (Russell 1961; Lyon 1961), in fishes (Anders and Klinke 1966), in the nematode *Caenorhabditis elegans* and in coccids (Berlowitz et al.1968).

MECHANISMS OF DOSAGE COMPENSATION

Dosage compensation serves to override the imbalance in X-linked gene dosage and ensures equivalent amounts of X-linked gene products in the two sexes. An obvious requirement is that each sex possess mechanisms both, for sensing the number of X chromosomes and for modulating X-linked gene activity to some level common to both sexes. The mechanism of counting chromosomes, despite a steady flow of information now starting to emerge from studies in *Drosophila* (Torrez and Sanchez 1989 ; Parkhurst et al.1990 ; Duffy and Gergen 1991), is a relatively uncharted terrain and is not considered here. On the other hand, the mechanisms of modulation of X chromosomal activity to effect dosage compensation are relatively well understood.

In mammalian somatic cells one of the two female X chromosomes is heteropycnotic or highly condensed (Ohno et al. 1957) and late in replicating (Yerganian 1960). However, the non-heteropycnotic X chromosome resembles the single male X chromosome both in morphology and in terms of replication pattern (Yerganian 1960). The speculative connections that could be made between these observations and dosage compensation were strengthened by an interesting set of observations made by Russell and Bangham (1959) (Russell 1961 for review). Mice homozygous for the recessive allele of the autosomal *brown* (*b*) locus have a uniformly brown coat color while those heterozygous (*b⁺b*) have a uniformly wild-type (agouti) coat color. However, in heterozygous mice, when the *b⁺* allele was placed on one of the two X chromosomes via a translocation event, the coat color was found to be mosaic for brown and wild-type (Russell and Bangham 1959). Furthermore, this mosaicism was restricted to females and males carrying the same translocation had a uniform wild-type coat color. The interpretation of this observation led to the hypothesis (Lyon 1961) which proposed that one of the two female X chromosomes was randomly inactivated so that each cell has a single active X chromosome, thereby equalizing the dosage of functional X-linked genes in males and females. Thus in the case of the *brown* locus translocations, cells differed with respect to the activity status of the X chromosomal segment adjacent to *b⁺*. Those in which *b⁺* lay next to inactive X chromosomal DNA did not express *b⁺*. This was

presumably due to a "position effect" whereby the inactive status of the adjacent X chromosomal DNA caused the inactivation of the b^+ allele. While X chromosome inactivation has now been well established as the mechanism of dosage compensation in mammals, by the time of the formulation of this hypothesis, it had become clear that in *Drosophila* both X chromosomes in the females are active.

In *Drosophila*, the formation of giant chromosomes in larval tissues through the process of polytenization greatly simplifies analysis of chromosome morphology and function. As early as 1936, Offerman observed that the width of a single polytene X chromosome in males is comparable to that of the paired X chromosomes of the female suggesting that the single male X chromosome may be twice as wide as each female X chromosome. Dobzhansky (1957) used interspecific hybrids to be able to compare the single X chromosome in males with unpaired, individually distinguishable X chromosomes of the females and found that the male X chromosome was twice as wide as each of the female X chromosomes. There was no visible difference in morphology between the two X chromosomes in the female. Moreover, *Drosophila* females heterozygous for X-linked loci were not known to exhibit the mosaicism characteristic of random X inactivation. Kazazian et al. (1965) observed that female *Drosophila* heterozygous for the alleles of the X-linked *6PGD* gene encoding "fast" and "slow" electrophoretic variants of the enzyme, have enzymes of "fast", "slow" and "intermediate" mobility, the latter presumably being heterodimers including both the "fast" and "slow" monomers. Barring a case of escape from inactivation, these observations indicated that both X chromosomes within a female cell are active. These observations and transcription autoradiography experiments (Mukherjee and Beerman 1965, 1966) to be discussed later in this chapter resulted in X inactivation being ruled out as a mechanism of dosage compensation in *Drosophila*. It follows then that a sex-specific modulation of the rate of activity of X-linked genes is responsible for dosage compensation in this organism.

If unequal dosage of X-linked genes is to be compensated for through the sex-specific modulation of X-linked gene activity, the mechanism of modulation can be expressed in terms of the basal level of activity of an X-linked gene. This would in turn lead to either hyperactivation to account for the single X chromosome dose in males or

hypoactivation to account for the double X chromosome dose in females. Nevertheless, since it is the males who have the lower dose, either kind of modulatory activity described is bound to manifest itself as an observation that an X-linked gene in males is twice as active as each of its counterparts in female. Then, of course, there is the question of the level in the hierarchy of gene expression at which gene activity is modulated, bringing along with it the currently fashionable questions of the identity and functions of cis-acting elements and trans-acting factors that must be involved.

A variety of kinds of evidence has accumulated that has significantly enhanced our understanding of the mechanism of dosage compensation in *Drosophila* (reviewed by Stern 1960 ; Lucchesi 1973 ; Lucchesi 1978 ; Baker and Belote 1983 ; Mukherjee 1986; Lucchesi and Manning 1987). This evidence is reviewed here in order to present the various aspects of dosage compensation in *Drosophila* as they are understood today. Most of the studies of this phenomenon in this genus have centered around the species *D. melanogaster* and so unless mentioned otherwise all subsequent references are to this species.

Dosage compensation in *Drosophila*: Transcriptional hyperactivation of the male X chromosome relative to the female X chromosome

The observation that the male X chromosome had twice the width and diffusiveness of each of the single female X chromosomes had prompted a prophetic comment from Dobzhansky (1957), who wrote, " the bloated appearance of the male X indicates accumulation in this chromosome of some products of gene activity. A male has only one X, but this single X works apparently twice as hard as does each of the two X's in the female". Aronson et al. (1954) and Rudkin (1964) used UV microspectrophotometry to show that the single X chromosome in the males has half the amount of DNA but approximately 10% more protein than the paired X chromosomes in the females. Mukherjee and Beerman (1965, 1966) measured ³H-uridine uptake by X chromosomes in male and female nuclei and demonstrated that the single X chromosome in males incorporates twice as much ³H-uridine as each of the two X chromosomes in the

females. These results are consistent with the hypothesis that the X chromosome in the male is twice as active as each of the X chromosomes in the female. Furthermore, they strongly suggest that the modulation of X chromosomal activity takes place at the level of transcription.

Ganguly et al. (1985) measured the steady-state level of the transcript encoded by the X-linked gene *G6PD*, (glucose-6-phosphate dehydrogenase), and showed it to be equal in the two sexes. This was then observed to be the case with transcripts from other X linked genes like those encoding *Sgs4*, and *6PGD* (Breen and Lucchesi 1986 ; Kaiser et al. 1986 ; Gutierrez et al. 1989). While it could still be argued that the equivalence of steady state levels of transcripts is established post transcriptionally, the possibility does seem remote in view of the results of transcriptional autoradiography mentioned earlier, and it has come to be accepted that sex-specific modulation of X chromosomal activity is at the level of transcription.

Models For Sex-Specific Modulation of Transcription of X-Linked Genes

The sex-specific modulation of sex-linked characters in the two sexes had been attracting a lot of attention even in the early 1920s. Bridges (1922), Stern (1960 for review), proposed that for every gene there exist plus and minus modifier genes randomly distributed over the genome. As such the sum of the effects of all the modifier genes would then result in the observed level of activity of the target gene. The X chromosome was also hypothesized to carry modifier genes and it was proposed that any gene, whether X-linked or autosomal, whose plus or minus modifiers are disproportionately represented on the X chromosome, would show a different activity in males relative to females. On the other hand X linked or autosomal genes that were not controlled to a significant extent by modifiers on the X chromosome would be unaffected by the X chromosome dosage difference in the two sexes. Muller (1950) built on this modifier hypothesis to suggest that the X chromosome encoded minus modifiers that are not dosage-compensated and that they served as repressors of X-linked gene activity (reviewed by Lucchesi 1973). This would result in lower X-linked activity in females relative to males because the

female would have two doses of minus modifiers while males would have one. A testable aspect of such a model is the existence of distinct X-linked loci that repress X-linked gene activity.

This has been attempted through the use of heteroploids and X-segmental aneuploids. Heteroploids are flies and larvae that have an abnormal number of X chromosomes or of sets of autosomes while segmental aneuploids are those that have an abnormal number of any segment of a single chromosome. Fig. 1., adapted from Lucchesi and Manning (1987), shows the summary of observations of X linked gene activity in heteroploids. On a per-gene dose basis, the rate of X chromosomal activity in a triploid (3X:3A) female was found to be equivalent to that of a normal (2X:2A) female, based on the observation that the triploids had 1.5 times more X chromosomally encoded products than did the normal diploid females (Lucchesi and Rawls 1973a ; Lucchesi and Rawls 1973b ; Faizullin and Gvozdev 1973 ; Maroni and Plaut 1973b). On a per-gene dose basis, the X-linked gene activity, in triploid intersexes (2X:3A) was found to be greater than that in diploid (2X:2A) or triploid (3X;3A) females, but less than that in diploid males (1X:2A) (Lucchesi and Rawls 1973b ; Maroni and Plaut 1973a ; Maroni and Plaut 1973b ; Faizullin and Gvozdev 1973). Similarly, in the case of metamales (1X;3A), X-linked gene activity was found to be substantially higher than that of a normal diploid male and that of a diploid or triploid female (Lucchesi et al. 1977). A generalization that emerges from these results (Fig. 1) is that X-linked gene activity, per gene dose, increases with a decrease in the ratio of the number of X chromosomes to the number of sets of autosomes, the X:A ratio. It needs to be mentioned here that in *Drosophila*, it is the X:A ratio that determines sex in a cell-autonomous fashion; male cells have an X:A ratio of 0.5 while female cells have a ratio of 1.0 (reviewed by Cline 1985). The results from experiments in hyperploids are also suggestive in that when the X:A ratio is less than that in a normal diploid male, X-linked gene activity is higher than that in normal males and normal females. This suggests that an X:A ratio of less than one enhances X linked gene activity. Moreover, these results raise the possibility that autosomes play a role in dosage compensation, and as such they are inconsistent with the hypothesis that negative regulators of X-linked activity encoded by the X chromosome are the mediators of

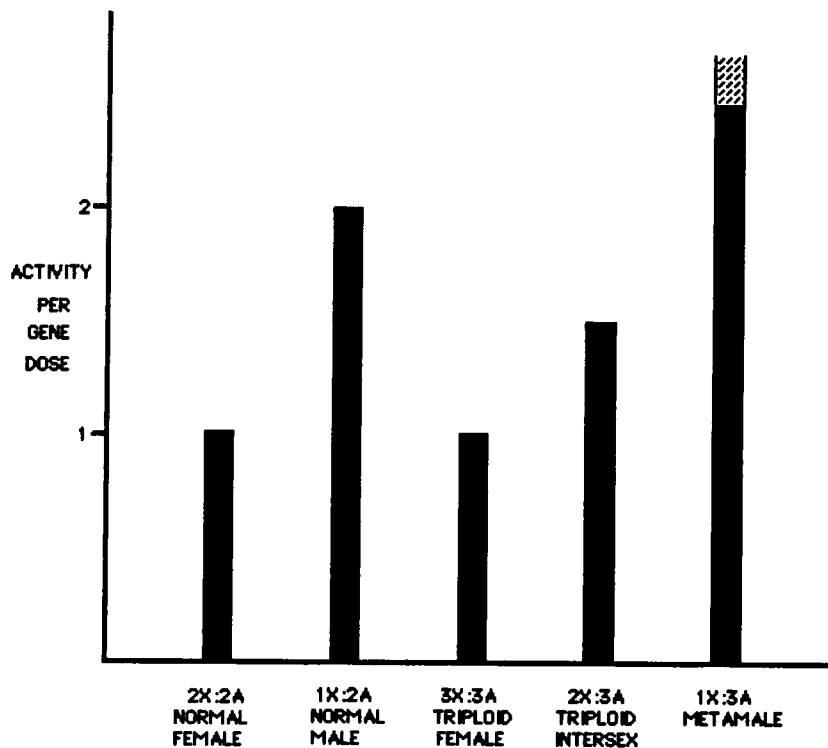


Figure 1. Quantitation of X-linked gene activity in heteroploids.

The illustration shown is adapted from Lucchesi and Manning (1987). Activity per gene dose is plotted against karyotypes representing normal or various heteroploid situations. In the case of metamales (1X:3A), X-linked activity per gene dose has been reported as being substantially higher than that of a normal male and so the representative bar has an open, hatched end. X-linked activity was determined as the level of X-linked gene encoded enzyme activity as well as by quantitative transcription autoradiography analysis of X chromosomes.

dosage compensation. The above observations are consistent, however, with a model proposed by Maroni and Plaut (1973b) in which the X chromosome titers a factor(s) of autosomal origin that enhances X chromosomal activity in the males. This model was further strengthened by Maroni and Lucchesi (1980), who used X-Y translocations to produce X-segmental aneuploids with one single whole X chromosome along with different segments of the X chromosome so that they had 1.25, 1.5, 1.62 and 1.85 X chromosomes relative to the normal complement of autosomes (Fig. 2). The striking observation was that, relative to autosomal transcription, the total transcriptional activity of the X chromosome remained unchanged in all the aneuploids. Thus with an increase in X chromosomal material, the amount of transcription per locus was seen to be decreased so that the overall X linked activity is equivalent. No difference was observed between regions of the X that were duplicated and those that were present in single copies. The apparent suggestion is that X chromosomal DNA competes for a limiting factor that ordinarily serves to enhance X chromosomal transcription. However this particular interpretation would convey the impression that there are no distinct loci on the X chromosome that participate as numerator elements in the calculation of the X:A ratio unless one assumes that such loci were fortuitously either present or absent in each of the aneuploids studied. The term numerator elements (see Cline 1988) designates X-linked loci that are thought to participate in the mechanism by which chromosomes are counted to determine the X:A ratio. At this time there is some information on hand about the identity of numerator elements of the X:A ratio; one such element has been shown to be the T4 gene within the *achaete-scute* complex at 1A on the X chromosome (Torrez and Sanchez 1989 ; Parkhurst et al. 1990). In light of this information a reasonably accurate explanation of the experiments of Maroni and Lucchesi (1980) would be that despite the increase in X chromosomal material in each of the aneuploids tested, the X:A ratio as such did not change and remained at 0.5, triggering the male mode of dosage compensation (note that the *achaete-scute* complex (*As-c*) is unduplicated in all aneuploids studied, Fig. 2). Perhaps the increasing amounts of X chromosomal DNA titrated a fixed amount of autosomally encoded trans-acting factors so that at any locus transcription decreased with an increase in X chromosomal DNA. The results of these

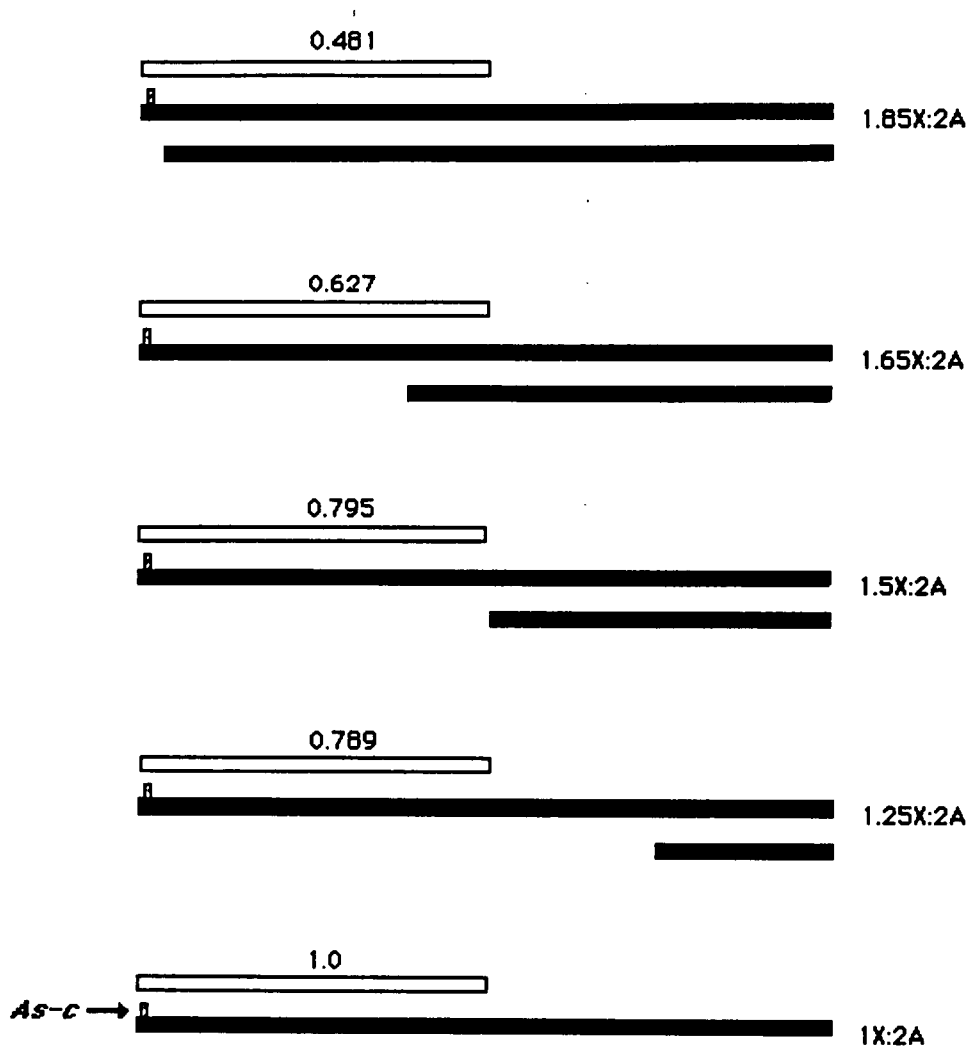


Figure 2. Quantitation of X-linked gene activity in segmental aneuploids.

The illustration shown is adapted from Lucchesi and Manning (1987). Solid black bars represent X chromosomal DNA while open bars designate the region of the X chromosome immediately below them that were analysed by transcription autoradiography. At the bottom of the figure is the normal 1X:2A male X chromosome. The level of activity in the region designated by the open box in the normal male is 1.0. Quantitation of X-linked activity of the same region in each of the aneuploids is reported as percentage of activity in the normal male. At the extreme left of each full length X chromosome, a small rectangle indicates the *As-c* complex which is unduplicated in all the aneuploids used in this study.

experiments with segmental aneuploid would then be consistent with the Maroni and Plaut (1973b) model for a male mode of dosage compensation in which the X:A ratio of 0.5 causes the expression of autosomally encoded factors that mediate hyperactivation of X-linked transcription. Such factors would then have to necessarily be superimposed upon the basic transcriptional machinery of each X-linked gene that needs to be hyperactivated. This indeed is the currently prevailing model and has been strengthened by the discovery of autosomal mutations, to be discussed later, that exert a lethality on males due to a failure of X chromosome hyperactivation.

X-chromatin conformation in the two sexes

The higher order structures of transcription units and neighboring regions are thought to play an important role in transcriptional activity. Thus differences in transcriptional activity of the male and female X chromosome could be caused by differences in X-chromatin conformation in the two sexes. The probe of choice for chromatin conformation has been relative nuclease sensitivity of the regions being compared. Hypersensitivity to nucleases is correlated with increased accessibility, a parameter taken to be directly proportional to the transcriptional activity of the template (Weintraub and Groudine 1976). Eisenberg and Lucchesi (1976 ; 1983) compared nuclease sensitivity of the X-linked, dosage-compensated heat shock gene, *Hsp82*, of *D. pseudobscura* and failed to detect any sex-specific differences. This however, does not rule out the possibility that sex specific differences in conformation do exist but lie outside the region tested in their study. Another attempt at probing X-chromatin conformation in the two sexes was the use of a semi in-vitro assay in which a heterologous (*E. coli*) polymerase was used to transcribe preparations of fixed polytene chromosomes (Umeil and Plaut 1973 ; Sederoff et al. 1973). Utilizing this strategy, Chatterjee and Mukherjee (1981) observed a higher transcriptional activity of the male X chromosome relative to the female X chromosome, presumably indicating a more accessible X-chromatin conformation in males. Chatterjee (1981) further demonstrated that the male and female X chromosomes displayed a differential susceptibility to high salt

treatments prior to the heterologous polymerase assays. This has led to the suggestion that a difference in the nucleoprotein organization of the X chromosome exists in the two sexes (Chatterjee 1981). However, it needs to be emphasized here that a central caveat in the interpretation of differences in chromatin conformation is the degree to which the observations reflect a cause rather than an effect of functional differences. At this time therefore it must be said that no correlation has been made between transcriptional hyperactivity of the male X chromosome and any definable sex-specific differences in chromatin structure of the X chromosome.

Cis-acting Elements and Trans-acting Factors in Dosage Compensation

An essential aspect of a regulatory mechanism acting at the level of a chromosome is the interaction between trans-acting "effector" factors and cis-acting "identifier" elements. In a model in which a single X chromosome in males is to be hyperactivated by autosomally encoded factors, the immediate considerations are the sequences that confer a "X chromosomal identity", and the nature and mode of action of the factors that actually cause hyperactivity. On both these fronts some information is on hand.

The existence of cis-acting sequences that regulate male specific hyperactivity has been demonstrated in a number of cases (reviewed by Lucchesi and Manning 1987). An X-chromosomal segment translocated to an autosome is transcriptionally hyperactivated in males (Seecof et al. 1969; Tobler et al. 1971). Ectopic autosomal expression of specific X-linked genes rather than chromosomal segments has also been examined using the technique of P-element mediated transformation (Rubin and Spradling 1982) Such studies have suggested that cis-acting sequences mediating male-specific transcriptional hyperactivity are closely linked to the genes they regulate (Hazelrigg et al. 1984 ; Pirrotta et al. 1985; Hoffmann et al. 1987; Gutierrez et al. 1989). The observation that an autosomal gene, *Xdh*, (Spradling and Rubin 1983) or a normally non-dosage compensated X-linked gene, *Lsp*, (Ghosh et al. 1989), can be hyperactivated when at ectopic sites on the male X chromosome suggests that the cis-acting sequences mediating male-specific transcriptional hyperactivation have the properties of orientation and position

independence characteristic of enhancer-like sequences. On the other hand, however, Roehrdanz et al. (1977) showed that a large autosomal fragment, does not appear to be dosage compensated when X-linked. However, it has not been possible to identify a discrete sequence that mediates male-specific hyperactivation of an X-linked gene or even to draw a consistent picture with regards the organization of such sequences on the X chromosome.

The function of a cis-acting sequence is almost always expressed in terms of the function of trans-acting factors that directly or indirectly bind to it. A set of autosomal mutations collectively called "male-specific lethals" (*msl*) have been shown to exert their lethality through the reduction of X linked gene transcription by two-folds in the males (Golubovsky and Ivanov 1972 ; Tanaka et al. 1976 ; Belote and Lucchesi 1980a ; Belote and Lucchesi 1980 ; reviewed by Lucchesi and Manning 1987). These mutations do not appear to have any effect on X chromosome transcription in the female somatic cells (Belote and Lucchesi 1980a) and are therefore thought to be involved as trans-acting determinants of male-specific hyperactivation of X-linked gene transcription (reviewed by Baker and Belote 1983 ; Lucchesi and Manning 1987). Recently, one of the male-specific lethal genes, *maleless* (*mle*) (Golubovsky and Ivanov 1972 ; Fukunaga et al.1975; Tanaka et al. 1976), has been cloned and its product shown to interact specifically with the X chromosome in a male-limited manner (Kuroda et al. 1991). The predicted *mle* gene product contains regions of similarity with a family of RNA and DNA helicases and especially resembles the former (Kuroda et al. 1991). A speculative generalization with respect to the products of the male-specific lethal genes would be that they interact with specific X-linked cis-acting sequences that mediate male specific hyperactivation.

Given our current understanding of the involvement of cis-acting and trans-acting elements in dosage compensation, some aspects of the problem on hand become readily apparent. It can be expected that the hyperactivation machinery is superimposed on the other transcriptional control elements of a gene, like those that confer tissue or developmental stage specificity of transcription. As such this can be visualized as occurring independently at every transcription unit, or at the level of discrete domains, microenvironments of hyperactivity, each of which encompasses a few genes. At any of

these levels the hyperactivation machinery could act by either generating and stabilizing a chromatin conformation promoting hyperactivity, or by interacting directly with the transcription apparatus, or both! As such the molecular mechanisms involved in male-specific hyperactivation can serve as a system for the study some very relevant questions asked in the field of control of gene transcription. However, what does appear to be a unique feature of this hyperactivation mechanism is the apparent precision with which a two-fold increase of transcription occurs in males.

THE XIR CHROMOSOME OF *DROSOPHILA MIRANDA*: A MODEL TO STUDY THE MECHANISMS OF DOSAGE COMPENSATION IN *DROSOPHILA*

This dissertation describes the characterization of an X-linked gene in *Drosophila miranda* as an initiation of a project, the long-term goal of which is to understand the molecular mechanisms involved in X-linked dosage compensation. The choice of the system described offers the additional perspective of evolution of the cis-acting sequences involved in this mechanism. The ancestral chromosomal arrangement of the genus *Drosophila* is believed to have consisted of five acrocentric rod shaped elements, A - E, and a small dot-shaped element, F, (Patterson and Stone, 1952). During the course of evolution, these chromosomal elements have undergone several structural rearrangements to give rise to the karyotypes of the present day species (Patterson and Stone 1952). Fig. 3 is an illustration depicting the origin of the karyotypes of *D. melanogaster* and *D. miranda*, species of the *melanogaster* and *obscura* groups respectively. Element A is the ancestor of the telocentric X chromosome while element D is the ancestor of the autosomal arm 3L of *D. melanogaster*. A centric fusion between these two elements gave rise to a metacentric A-D or X-D element in *D. miranda*. While females are homozygous, males are hemizygous for this metacentric chromosome (Patterson and Stone 1952). This metacentric X chromosome of *D. miranda* is also referred to as XL.XR (Patterson and Stone 1952) in which XL and XR have originated from ancestral elements A and D

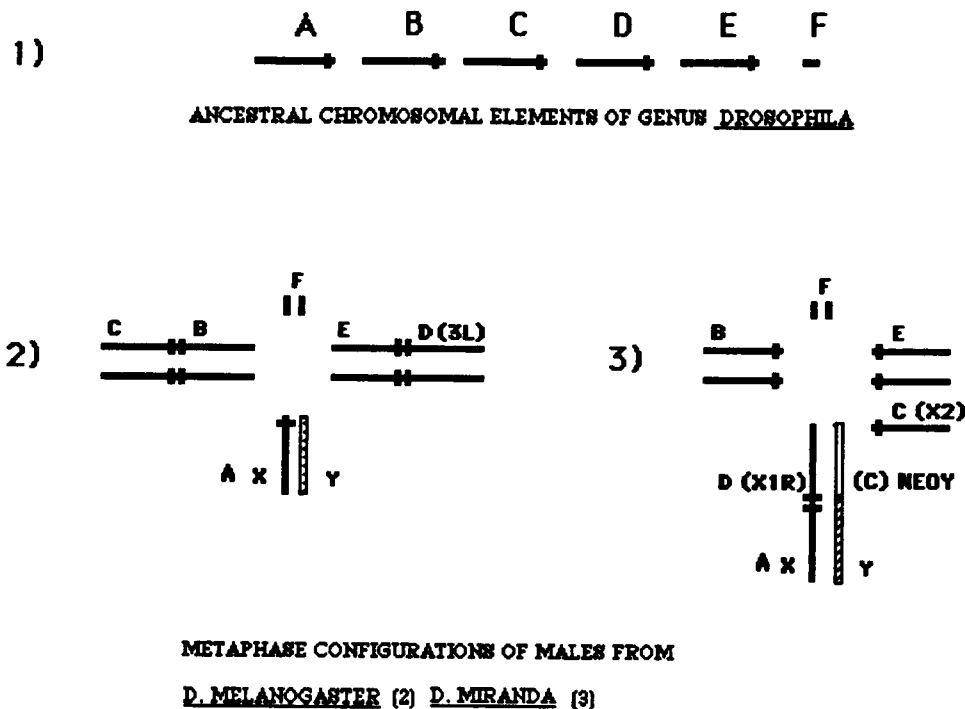


Figure 3. Origin of karyotypes of *D. melanogaster* and *D. miranda*.

The ancestral chromosomal elements of genus *Drosophila* are depicted in panel 1. Fusion of elements D and E, and elements B and C gave rise to the karyotype of *D. melanogaster* shown in panel 2. Panel 3 shows the karyotype of *D. miranda*. Fusion of element A and D and subsequent loss of the free-standing D element has given rise to the X1R chromosome. The X1R chromosome is in two copies in females and a single copy in the males. Fusion of element C to the Y chromosome has given rise to a neo-Y chromosome. The free standing C element is in two copies in the females and a single copy in the male and has been designated as the X2 chromosome.

respectively. Through an unrelated translocation event, *D. miranda* also possesses another X chromosome referred to as X2 (Fig. 3). The right arm of the metacentric X chromosome resulting from the A-D fusion is therefore referred to as the X1R chromosome. Since autosome 3L of *D. melanogaster* and the X1R chromosome of *D. miranda* have the same origin, they are expected to be homologous to each other. Homology between the X1R of *D. miranda* and autosome 3L of *D. melanogaster*, first observed through cytogenetic analysis, has now been established at the level of interspecific hybridization of cloned sequences to polytene chromosomes (Steinmann 1982 ; Steinmann et al. 1984). The D element has therefore evolved as the autosome 3L in *D. melanogaster* and the X1R chromosome in *D. miranda*. In the latter species, X1R-linked genes face the problem of unequal dosage in the two sexes and they are expected to have evolved to respond to the dosage-compensatory machinery. Transcriptional autoradiographic analysis of the distal 20% of the X1R chromosome has demonstrated that this chromosome is dosage compensated (Strobel et al. 1978). Also, in *D. pseudobscura*, a sibling species of *D. miranda*, a comparable situation exists in the form of an XR chromosome representing the A-D fusion. This chromosome too, has been analyzed by transcriptional autoradiography and shown to be dosage compensated (Pierce and Lucchesi, 1980). Further, the steady-state level of an XR linked gene *Hsp82* has been shown to be equal in the two sexes, suggesting that XR linked genes are dosage compensated (Pierce and Lucchesi 1980; Dey and Lucchesi 1983). However, no study of dosage compensation at the level of an individual X1R-linked transcription unit of *D. miranda* has been reported so far.

Given our understanding of the mechanism of dosage compensation it is reasonable to assume that the X1R chromosome of *D. miranda* has acquired cis-acting sequences that mediate dosage compensation. The X1R chromosome of *D. miranda* thus represents a chromosome that is dosage-compensated in one species and not another. At the level of an individual X1R-linked gene which is dosage compensated it becomes interesting to seek the modifications within the transcription unit relative to its 3L-linked homologue in *D. melanogaster* that enable it to be hyperactivated in a male-specific manner. The nature of such modifications can be expected to shed some light on the mode of evolution

of cis-acting transcriptional control elements in general and dosage compensatory-sequences in particular.

PROJECT OBJECTIVES AND SIGNIFICANCE

It is the objective of this study to characterize an X1R-linked gene of *Drosophila miranda* in order to establish it as a system that can be exploited to study the mechanism of dosage compensation at the level of a single transcription unit. The primary interest lies in characterizing the cis-acting sequences that mediate male specific hyperactivation in terms of their organization with respect to other cis-acting sequences that regulate tissue or developmental stage-specificity of transcription of a gene, a first step towards analysis of their function. As detailed in section C, an X1R-linked gene is expected to have an autosomal homologue in *D. melanogaster*. It can be assumed that while the basic transcriptional control elements of the gene, like those that confer tissue and developmental stage-specificity would be conserved in the two species, the X1R-linked gene would have additional regulatory elements mediating male-specific transcriptional hyperactivation. The identity, organization, and function of such sequences will therefore be of considerable importance both in terms of elucidating the molecular mechanisms of transcriptional hyperactivation as well as in terms of the mode of evolution of cis-acting sequences that regulate transcriptional hyperactivation.

The specific objectives of this study are described below:

Objective 1

Establishment of the appropriateness of an X1R-linked genomic clone of *D. miranda* for use as a system to study the molecular mechanisms of dosage compensation in *Drosophila* with an emphasis on the characteristics and organization of cis-acting sequences that mediate male-specific transcriptional hyperactivation.

An ideal choice is that of a gene which is dosage-compensated, and transcribed in a tissue, and developmental stage-specific manner and whose mRNA is abundant enough to be easily quantitated.

Objective 2A

Characterization of the X1R-linked gene of *D. miranda* and its 3L-linked homologue in *D. melanogaster* in terms of the extent of transcription units and homology between them.

The usefulness of carrying out an analysis of homology between the X1R linked gene being studied and its homologue in *D. melanogaster* is that information with respect to highly conserved regions in non-coding DNA could highlight putative regulatory sequences that have to do with transcriptional control of this gene in terms of tissue and developmental stage specificity. While such conclusions are necessarily speculative they are also testable and would aid the design of constructs to be tested in P-element mediated transformation experiments.

Objective 2B

Localization of the cis-acting sequences involved in transcriptional control of an X1R linked gene of *D. miranda* using P-element mediated transformation. This will involve the construction of transgenic flies carrying the X1R-linked gene at autosomal sites in *D. melanogaster*. Male-specific hyperactivity of the transgene is expected to be reflected by a higher steady state level of transgene derived mRNA in the males relative to females.

Expression of the X1R-linked gene of *D. miranda* in *D. melanogaster* is the method of choice for localization of cis acting transcriptional control elements at this locus. An XR linked gene, HSP82, of *D. pseudobscura* has been successfully expressed at autosomal as well as X-linked sites in *D. melanogaster* and has been reported to respond to the dosage-compensatory machinery (Jaffe and Laird 1986). This precedent as well as the lack of a suitable method for germ line transformation of *D. miranda*, makes the choice of P-element mediated transformation of *D. melanogaster*, a valid one.

It is essential to emphasize here that, with the exception of its "enhancer-like" properties, there is at present no established characteristic shown by such sequences that would aid this search. Thus it may be possible that localization of the basic transcriptional regulatory elements of this gene, ones that enable it to be expressed in its characteristic

tissue-specific pattern, can be accomplished, but that cis-acting elements mediating male-specific transcriptional hyperactivation lie outside the regions tested during the duration of this study. In any case, the appropriate (in terms of its basic expression pattern) expression of this gene from a defined transgenic fragment will serve, in future studies, as the reference point in the search for the more significant element(s) that mediate male-specific transcriptional hyperactivation.

CHAPTER 2

DOSAGE COMPENSATION OF AN X1R-LINKED RETINA-SPECIFIC GENE IN *DROSOPHILA MIRANDA*

INTRODUCTION

Dosage compensation refers to the mechanism which equalizes the amounts of X-linked gene products in females with two and males with one dose of X-linked genes. In *Drosophila* this is mediated by the hyperactivation of transcription of the male X chromosome to approximately twice the level of that of the female X chromosome. This transcriptional hyperactivation is in turn thought to be mediated by trans-acting factors that, in a male-limited manner, interact with specific cis-acting sequences on the X chromosome. No discrete cis-acting sequence(s) that confers male-specific transcriptional hyperactivation has been identified yet. As such therefore, very little is known about the physical and functional relationships between such sequences and those that mediate the basic transcriptional patterns of a gene in terms of its tissue or temporal specificity.

With the long-term goal of addressing the above questions, the work described in this dissertation is focussed on the X1R chromosome of *D. miranda*. The X1R chromosome is the right arm of the metacentric X chromosome in this species and X1R-linked genes are present in two copies in females and in a single copy in males. This chromosome therefore faces the problem of dosage compensation. Transcription autoradiographic studies of 20% of this chromosome have suggested that it is transcriptionally hyperactivated in males (Strobel et al. 1978). An X1R-linked gene which is dosage compensated and expressed in a tissue specific manner would be a good system with which to characterize cis-acting sequences mediating dosage compensation. The fundamental questions to be asked would be the with regards the physical and functional relationships between such sequences and those that confer the basic expression pattern of the gene. They could be represented by a single sequence motif or by dispersed

elements that collectively function to cause the optimal level of hyperactivation. In either case such an element(s) could be completely separable from other transcriptional control elements of a gene or could be tightly linked. These questions have been asked of some X-linked genes in *D. melanogaster* (see chapter 1) and of an XR-linked gene, Hsp82, of *D. pseudobscura* (Jaffe and Laird 1986; Sass and Meselson 1991 ; see chapter 1). At this stage the only consensus property that has emerged is that the cis-acting sequences mediating dosage compensation appear to be "enhancer-like" in their function.

There is also an additional attraction in choosing the X1R chromosome of *D. miranda* for the purpose of characterization of dosage compensatory cis-acting sequences. Studies of chromosome evolution in the genus *Drosophila* have suggested that the X1R chromosome of *D. miranda* and the autosomal arm 3L of *D. melanogaster* have evolved from the same ancestral element, D (Paterson and Stone 1952). Homology between these two chromosomes has since been established by cross-hybridization of cloned DNA probes to polytene chromosomes (Steinmann 1982 ; Steinmann et al. 1984). This therefore presents a unique situation wherein a gene that is X-linked and dosage-compensated in one species would be autosomal and non-dosage compensated in another. A comparison of the transcriptional control regions of the gene in these two species can therefore be expected to shed some light on the mode of evolution of cis-acting sequences that mediate dosage compensation.

In this chapter the initial characterization of an X1R-linked gene of *D. miranda* is described. It was isolated as a homologue of one of twenty genomic clones of *D. melanogaster* isolated by Levy et al. (1982). The twenty genomic clones were isolated using the criterion that they hybridized with end-labelled mRNA isolated from tissues of the adult head, and not with end-labelled mRNA isolated from the rest of the body. One of these, clone 507, was mapped to the autosome 3L of *D. melanogaster*, and appeared to contain a simple transcription unit encoding an 1.8 kb mRNA which was extremely abundant, constituting 0.66% of the head poly-adenylated mRNA (Levy et al. 1982 ; Levy and Manning 1982). It was expected that the *D. miranda* homologue of this clone would be X1R-linked and as such could be a suitable candidate for this study. A transcribed fragment of clone 507 of *D. melanogaster* (507ml) (Levy et al. 1982) was used to isolate

the *D. miranda* homologue, clone 507mr, from a *D. miranda* genomic library (Krishnan et al. 1991). The results described in this chapter demonstrate that clone 507mr is X1R-linked in *D. miranda* and like in *D. melanogaster* encodes a 1.8 kb mRNA. In both species this mRNA (507 mRNA) is detected only in the retinal photoreceptor cells of the adult compound eye. In *D. miranda*, the steady state level of 507 mRNA in the two sexes is equal, demonstrating that it is dosage compensated. Clones 507ml and 507mr show considerable restriction polymorphism and lack of sequence similarity, as judged by nucleic acid hybridization, in some regions including the 3'-transcribed region of the gene. The 5'-portion of the transcribed region as well as some regions outside the transcribed regions appear to be highly conserved in the two species, and these latter regions may have some regulatory significance at this locus. A major portion of this chapter has been published in a modified form in Krishnan et al. (1991).

MATERIALS AND METHODS

Drosophila strains and cultures

Wild type *D. melanogaster* and *D. miranda* (strain S204) were obtained from Dr. J. Lucchesi, University of North Carolina and Dr. R. Norman, Arizona State University, respectively. *Eye missing* (*eym*) flies were received from Y. Inoue, Mishima, Japan and *glass* flies were received from the Bloomington stock center, Indiana. Flies were raised on cornmeal-agar-molasses medium at 23 C and 60% humidity with a 12 h dark/light cycle.

Genomic clones 507, 507ml and 507mr

The isolation of clone 507 has been described by Levy et al. (1982). A 2.3- Kb EcoRI fragment from this clone was used to isolate an overlapping clone 507ml and clone 507mr from EMBL 3 genomic libraries of *D. melanogaster* and *D. miranda* respectively. These procedures are detailed in Appendix I.

Nucleic acid isolation, labeling and blot hybridization

Total cellular RNA from whole adult flies, heads only and bodies only, were isolated by a hot phenol method (Helms et al. 1985). Oligo (dT)-cellulose chromatography of RNA was performed according to Maniatis et al. (1982). High-molecular-weight nuclear DNA from adult flies and DNA from the recombinant lambda clones were prepared according to Ganguly et al. (1985) and Helms et al. (1982), respectively. DNA fragments were labeled by the nick-translation procedure using a commercially available kit (Bethesda Research Laboratories, Md.). The specific activity of the ^3H and ^{32}P labeled DNA was 2×10^7 and 1×10^8 cpm/ug respectively. The poly (A)+ RNA was end-labeled using T4 polynucleotide kinase to a specific activity of 2×10^9 cpm/ug according to Levy et al. (1982). Random hexamer-primed, single-stranded heterogeneous cDNA probes of specific activity 2×10^9 cpm/ug were made against adult head poly (A)+ RNA according to Maniatis et al. (1982). DNA and RNA blot hybridizations were performed according to Maniatis et al. (1982). These procedures are detailed in Appendix I.

***In situ* hybridization**

In situ hybridization to the polytene chromosomes was performed according to Levy et al. (1982). *In situ* hybridization of 507 gene probe to 507 mRNA within cryostat sections of heads of *D. melanogaster* and *D. miranda* were performed by Ken Swanson as described by Hafen and Levine (1986). The DNA probe was labelled with ^3H to a specific activity of 2×10^7 cpm/ug according to Hafen and Levine (1986). The procedures are detailed in the Appendix.

RESULTS

Isolation of clones 507ml and 507mr

Clone 507 was originally isolated by Levy et al. (1982) as a genomic clone that hybridized to end-labeled mRNA from heads but not bodies of *D. melanogaster*. The terminal 2.3-kb EcoRI fragment within clone 507 (Fig. 4) was shown to hybridize to a 1.8 kb mRNA (R.Ganguly unpublished observations). Since the 2.3 Kb EcoRI fragment

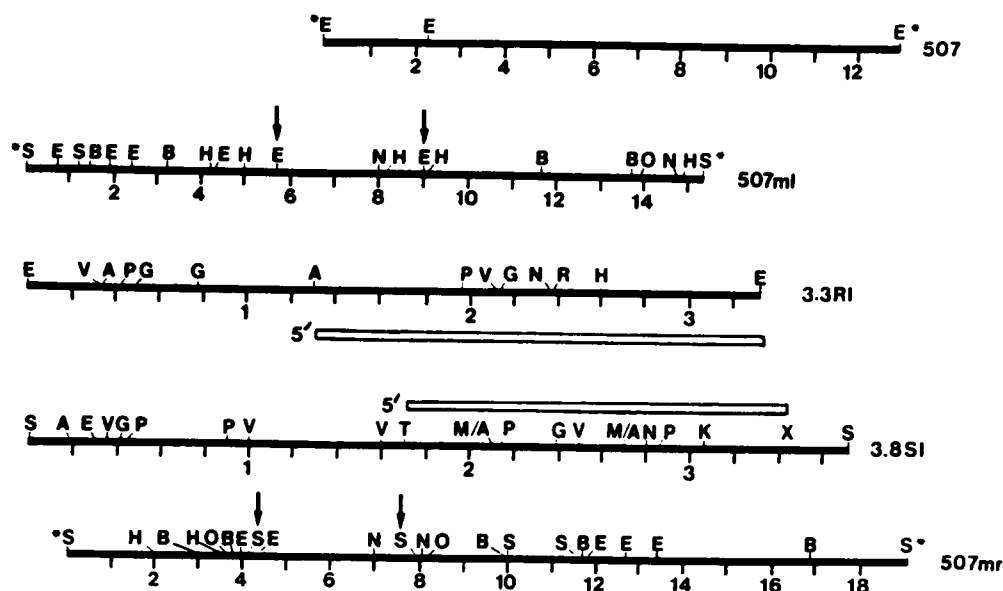


Figure 4. Restriction maps of clones 507, 507ml and 507mr.

On top is shown clone 507 originally isolated by by Levy et al (1982). Clone 507ml and clone 507mr were isolated from genomic libraries of *D. melanogaster* and *D. miranda* respectively using the 2.3 kb Eco RI fragment of 507ml as probe. Restriction sites with asterisks belong to the lambda vector arms. Arrows indicate the 3.3 Kb EcoRI and 3.8 Kb SalI fragments of the 507ml and 507mr clones respectively which are the only fragments detected by the 2.3 Kb EcoRI fragment of clone 507. Also these fragments appear to contain the entire transcribed portion of gene 507. Detailed restriction maps of these two clones are also shown (3.3RI and 3.8 SI).

E - EcoRI; B - BamHI; S - SalI; P - PstI; H - HindIII; A - Aval; G - BglII; V - PvuII; N - NcoI; O - XhoI; T - SstI; M - SmaI; R - NarI; X - XbaI; K - KpnI

represents a transcribed portion of the gene (gene 507) encoding the 1.8-kb mRNA it was chosen as the probe used to isolate the homologous clone, 507mr, from a genomic library of *D. miranda* (Fig. 4). The 2.3-kb EcoRI fragment of clone 507 hybridizes (not shown) to a 3.8-kb SalI fragment, which is approximately centered within clone 507mr (Fig. 4). Since one of the EcoRI sites of the terminal EcoRI fragment of clone 507 could belong to the vector (Fig. 4), this fragment was used as a probe on a Southern blot of EcoRI-digested genomic DNA of *D. melanogaster* and hybridization was seen with a 3.3-kb fragment (not shown). In order to isolate a clone in which this transcribed region of gene 507 could be accompanied by flanking sequences it was necessary to isolate an overlapping clone containing sequences flanking this region. Clone 507ml is one such overlapping clone in which the 3.3 Kb EcoRI fragment is approximately centered within the insert (Fig. 4). The detailed restriction maps of clones 507ml and 507mr are shown in Fig. 4.

Transcription mapping of gene 507

To determine the extent of the transcribed regions within clones 507ml and 507mr, various restriction fragments from these clones were probed with labeled random-primed cDNA made against adult head mRNA and also separately with end-labeled poly (A)+ mRNA from adult heads. The 3.3 Kb EcoRI fragment of clone 507ml and the 3.8 Kb SalI fragment of clone 507mr (see Fig. 4) were the only fragments within these clones that hybridized with these heterogeneous cDNA or end-labeled mRNA probes. These two fragments were also the only ones to hybridize to a 1.8 Kb mRNA in the adult head poly (A)+ mRNA on Northern blots (Fig. 5). The heterogeneous cDNA probes, described above, were then used to further localize the transcribed region within the 3.3 kb EcoRI and 3.8 Kb SalI fragments. Detailed restriction maps of these fragments and the extents of the transcribed region within them are shown in Fig. 4. The precise boundaries of the transcribed regions and the locations of introns, if any, could not be detected by this analysis.

To determine the direction of transcription (Fig. 4) the 3.3 Kb EcoRI and the 3.8 Kb SalI fragments were cloned into the vector pGEM3. Single stranded probes

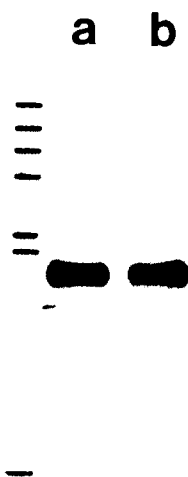


Figure 5. Northern analysis of clone 507ml and 507mr.

Equivalent amount (0.1 ug) of head poly (A+) mRNA of *D. miranda* (lane a) and *D. melanogaster* (lane b) was electrophoresed on a formaldehyde agarose gel, blotted onto nylon membrane and hybridized to ^{32}P -labelled 3.8 Sall fragment of 507mr (lane a) and 3.3 Kb EcoRI fragment of 507ml (lane b). Lines on the margin represent migration of (from the top) 23.1, 9.4, 6.6, 4.4, 2.3, 2.0, and 0.6 Kb marker DNA fragments.

complementary to one or the other strand were generated by transcription from either SP6 or T₇ promoters, and hybridized to northern blots containing mRNA from adult heads (not shown).

Two additional observations made during these transcription mapping experiments need be mentioned here. The 0.7-kb KpnI-SalI fragment of *D. miranda* (Fig. 4) did not hybridize to heterogeneous head cDNA or end-labelled head mRNA probes from *D. melanogaster* and similarly the 0.65-kb HindIII-EcoRI fragment of *D. melanogaster* (Fig. 4) did not cross hybridize with heterogeneous cDNA probes from *D. miranda*. This suggested an apparent lack of homology between the 3' transcribed regions of gene 507 in the two species and this is discussed later. There is also an as-yet unconfirmed possibility that the 5' most fragments of the 3.3 Kb EcoRI and 3.8 Kb SalI fragments have some transcribed region of a 1.8 Kb mRNA. The 0.45-kb EcoRI-BglII fragment of *D. melanogaster* (Fig. 4) shows faint hybridization with a 1.8 Kb mRNA when used as a probe on northern blots containing head poly (A)+ mRNA of *D. melanogaster* (not shown). Also, when the 0.4-kb SalI-PstI fragment of *D. miranda* (Fig. 4) was used as a probe on northern blots containing head total RNA of *D. miranda*, a faint hybridization at 1.8 Kb was observed after prolonged exposure (not shown). In later studies discussed below it appears that this region, relative to other regions mapping outside the major transcribed portion of this gene, is highly conserved in the two species. It is therefore tempting to speculate that an exon of gene 507 or one that encodes a 1.8-kb mRNA exists within the 0.45-kb EcoRI-BglII fragment of *D. melanogaster* (Fig. 4) and the 0.4-kb SalI-PstI fragment of *D. miranda* (Fig. 4).

Barring the possible presence of small exons which escaped detection by the transcription mapping methods discussed above, the entire transcribed region of gene 507ml and 507mr appears to be centered within the 3.3-kb EcoRI and the 3.8-kb SalI fragments of clones 507ml and 507mr respectively. As such therefore clones 507ml and 507mr appear to contain a simple transcription unit and considerable amount of flanking sequence at either end. This latter characteristic makes them suitable for use in the search for the transcriptional control elements of gene 507, a major objective of this study.

Chromosomal locations of genes *507ml* AND *507mr*

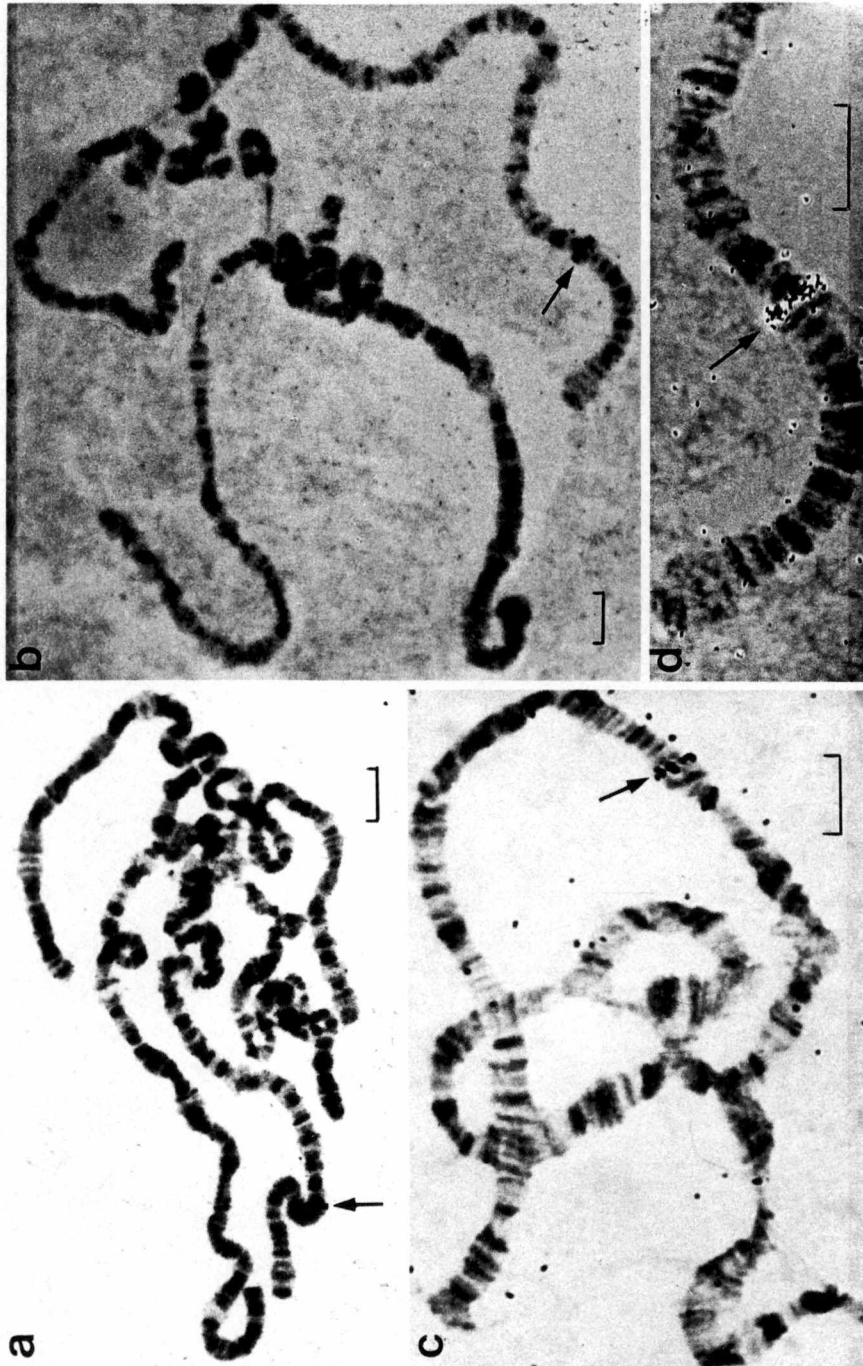
Polytene chromosome preparations of *D. melanogaster* and *D. miranda* were hybridized with ³H-labeled *507ml* and *507mr* DNA respectively. Fig. 6a and 6b show that in female as well as male *D. miranda* a single site of hybridization is seen on the X1R chromosome. The site of hybridization was mapped as 12A according to the published chromosome map of Das et al. (1982). Fig. 6c shows that in the case of *D. melanogaster* hybridization of the *507ml* probe is seen only at 66D on the 3L chromosome. This matches the observation of Levy et al. (1982) who mapped clone 507 to the same site in *D. melanogaster*. The chromosomal location of gene *507mr* is thus in agreement with the established homology between the chromosome 3L of *D. melanogaster* and the X1R chromosome of *D. miranda*. Furthermore, in the latter species the females have two copies of gene 507 while males have one, necessitating that this gene be dosage-compensated.

Dosage compensation of gene *507* in *D. miranda*

The presence of gene *507mr* in two copies in the females and a single copy in the males raised the question of whether gene *507mr* is dosage-compensated. The steady state level of 507 mRNA in the two sexes was determined by northern-blot analysis. In order to eliminate loading differences between sexes so as to ensure that equal amount of mRNA from each sex was being compared, the blots had to be hybridized with a gene probe whose mRNA is known to be in equal amounts in the two sexes. The hybridization intensity of this internal control was then used to normalize the hybridization intensity of 507 mRNA in each sex. The *Adh* gene which is autosomal in *D. miranda* (R.Ganguly, unpublished observations) and is expected to be in equal amounts in the two sexes was used as the internal control in these experiments. *Adh* has served as an internal control in sex-specific quantitation of mRNA in *D. melanogaster* (Ganguly et al. 1985 ; Guitierrez et al. 1989), but no previous report on the level of *Adh* mRNA in the two sexes of *D. miranda* existed. Therefore a quantitative RNA dot-blot analysis was carried out (Fig.7) and the levels of *Adh* mRNA in the two sexes of *D. miranda* were shown to be equal (Tables 1 and 2). The quantitative northern analysis for sex-specific levels of 507

Figure 6. Chromosomal locations of genes *507ml* and *507mr*.

In situ hybridization of the ^3H -labelled *507mr* DNA to *D. miranda* female (a) and male (b) polytene chromosomes; and ^3H -labelled *507ml* DNA to *D. melanogaster* (c) polytene chromosomes. The higher magnification of the X1R chromosome of male *D. miranda* is shown in panel (d). Arrows indicate the sites of hybridization. Bars = 10 microns.



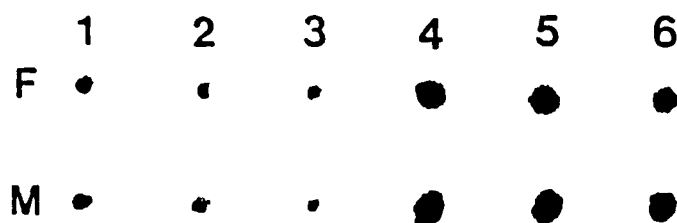


Figure 7. Dot-blot analysis of *Adh* mRNA levels in male and female *D. miranda*.

For each dot, total RNA isolated from adult males (M) and females (F) were denatured, spotted onto nylon membrane and hybridized to ^{32}P -labelled *Adh* gene probe of *D. melanogaster*. The autoradiogram was scanned with a laser densitometer (LKB 2222-010 Ultrosan XL) and the scans analysed with the LKB 2400 Gelscan XL software. The values for each dot are shown in Tables IA and IB.

TABLE 1

Dot blot quantitation* of *Adh* mRNA . I

<u>SEX</u>	<u>DOT1</u>	<u>DOT2</u>	<u>DOT3</u>	<u>mean $\pm$ SD</u>	<u>M/F Ratio</u>
Male	0.76	0.87	0.80	0.81 $\pm$ 0.05	1.05
Female	0.74	0.82	0.76	0.77 $\pm$ 0.03	

* Laser Densitometric units ; 0.5 micrograms poly A+ RNA / Dot

TABLE 2

Dot blot quantitation* of *Adh* mRNA . II.

<u>SEX</u>	<u>DOT4</u>	<u>DOT5</u>	<u>DOT6</u>	<u>Mean $\pm$ SD</u>	<u>M/F Ratio</u>
Male	1.63	1.61	1.52	1.58 $\pm$ 0.05	1.02
Female	1.53	1.60	1.51	1.55 $\pm$ 0.04	

* Laser Densitometric units ; 1.0 micrograms poly A+ RNA / Dot

mRNA in the two sexes was performed with head poly (A)+ mRNA (Fig. 8), and whole adult poly (A)+ mRNA (Fig. 8). Hybridization intensity of a band to be quantitated was determined by laser densitometric scanning of autoradiograms with non-saturated exposure. The normalized level of 507mRNA in each sex is expressed in terms of the ratio between levels of 507 mRNA and *Adh* mRNA (Table 3). Two additional Northern blots of whole adult poly (A)+ mRNA (not shown) were hybridized as for the above experiment but were quantitated differently. In this case bands to be quantitated were excised from the blot and radioactivity in each excised band determined by liquid scintillation counting. The normalized level of 507 mRNA in each sex is once again expressed as the ratio between levels of 507 mRNA and *Adh* mRNA (Table 4). The normalized male / female ratios of 507 mRNA (Tables 3, 4) clearly show that the steady state level of 507 mRNA is equal in the two sexes of *D. miranda*. Since 507mr is X1R-linked and present in two doses in the females and a single dose in males, on a per-gene-dose basis, males have twice the amount of 507 mRNA relative to the females indicating that gene 507mr is dosage-compensated.

Tissue-specific expression of gene 507

Levy et al.(1982) showed that the 1.8 Kb 507 mRNA is approximately 177 times more abundant in the head tissues relative to body tissues of *D. melanogaster*. This apparent tissue specificity of 507 gene expression was a significant aspect of the rationale for choosing clone 507 for this study. Since it can be assumed that discrete cis-acting sequences mediate tissue specificity of transcription of this gene it is expected that such sequences will be important reference points for the characterization of the cis-acting sequences that mediate dosage compensation of this gene in *D. miranda*. In *D. miranda*, as in *D. melanogaster*, the 507 mRNA is undetectable in tissues other than those of the head (not shown). The head of an adult *Drosophila* comprises mostly of tissues of the two large compound eyes and of the brain. To localize the 507mRNA to either the brain or the eye, the mutation *eym* (*eye missing*) was exploited. These flies lack both the compound eyes (Inoue 1980 ; Lindsey and Zimm 1985) and do not have any detectable retinal or laminal tissues (Ganguly et al. unpublished observations). The 507 mRNA was

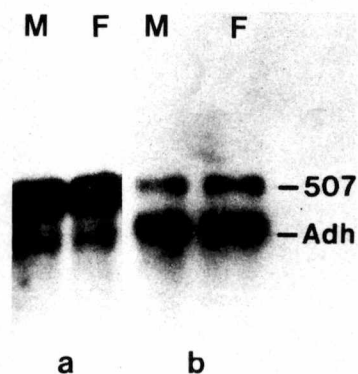


Figure 8. Quantitation of steady-state level of 507 mRNA in male and female *D. miranda*.

Equivalent amount (0.5 ug) of male (M) and female (F) head and whole-fly poly (A)+ mRNA were electrophoresed on a formaldehyde agarose gel, blotted and hybridized to a mixture of ^{32}P -labelled *507mr* and *Adh* coding fragments. After exposure the autoradiograms were scanned with a laser densitometer (LKB 2222-010 Ultrosan XL) and the scans were analysed with the LKB 2400 Gelscan XL software. Lighter exposures than the one shown in the Figure were used scanned to ensure linearity. The values for each band are given Table 3.

Blot a.- head poly (A+) RNA ; Blot b.- whole fly poly (A+) RNA.

TABLE 3
Dosage compensation of gene 507 in *D. miranda*. I.

<u>RNA</u>	<u>MALE</u>		<u>FEMALE</u>		<u>M/F</u>
	<u>507</u>	<u>Adh</u>	<u>507/Adh (a)</u>	<u>507</u> <u>Adh</u> <u>507/Adh (b)</u>	<u>(a)/(b)**</u>
HEAD	10.0*	2.1	4.76	9.3 1.9 4.89	0.97
WHOLE ADULT	4.5	28.7	0.16	5.7 33.8 0.17	0.94

* LASER DENSITOMETRIC UNITS **MEAN (a) / (b) : 0.96 ± 0.015

TABLE 4

Dosage compensation of gene 507 in *D. miranda*. II.

<u>RNA</u>	<u>MALE</u>			<u>FEMALE</u>			<u>M / F</u> <u>(a) / (b) **</u>
	<u>507</u>	<u>Adh</u>	<u>507/Adh (a)</u>	<u>507</u>	<u>Adh</u>	<u>507/Adh (b)</u>	
WH1	610*	985	0.62	980	1712	0.57	1.09
WH2	270	334	0.81	419	456	0.92	0.88
<u>* RADIOACTIVITY IN CPM</u>							<u>**MEAN (a) / (b) : 0.985 ± 0.11</u>

undetectable in the heads of *eym* flies (Fig. 9 b,c) suggesting that gene 507 is transcribed only in tissues of the adult compound eye. To further localize the expression of gene 507 to specific cells within the eye the mutation *glass* which removes all the photoreceptor cells of the adult compound eye was used. As shown in (Fig. 9 d,e), 507 mRNA is undetectable in head tissues of *glass* flies suggesting that gene 507 is transcribed only within the photoreceptor cells. 507mRNA was also localized to the retina of the adult compound eye by in situ hybridization of ³H labelled gene 507 DNA to sections of heads of *D. melanogaster* and *D. miranda* (Fig. 10). These results collectively establish that gene 507 is expressed only in the retinal photoreceptor cells of the adult compound eye.

Analysis of homology between clones 507ml and 507mr

The considerable restriction pattern polymorphism between clones 507ml and 507mr is evident in Fig. 4. To determine the pattern of divergence in nucleotide sequences of these two clones, a series of inter-species cross hybridizations were carried out. Such a comparison was expected to be useful in several ways. Defined cloned loci of chromosome 3L of *D. melanogaster* and their homologues in *D. miranda* have not been compared in terms of degree of sequence divergence. *D. melanogaster* and *D. miranda* are thought to have diverged approximately 45 million years ago (Pardue et al.1987). During the course of this divergence gene 507 of *D. miranda* has been subjected to additional selection pressures brought on by its X-linked and dosage-compensated status. In general, functionally constrained sequences at this transcription unit could be expected to be conserved in the two species. Regions that are found to have diverged considerably could either be functionally unimportant or could reflect changes that accomodate the X-linked and dosage-compensated status of this gene in *D. miranda*. Comparisons were made of the entire genomic lambda clones 507ml and 507mr, followed by a more detailed comparison of the transcribed regions of gene 507 within these clones and the results of this analysis, presented in Figs. 11, 12 and 13, are summarized in Fig. 14.

To compare the two genomic clones 507ml and 507mr, five different fragments (A - E) of clone 507mr (Fig. 11, top panel) were individually hybridized to DNA blots

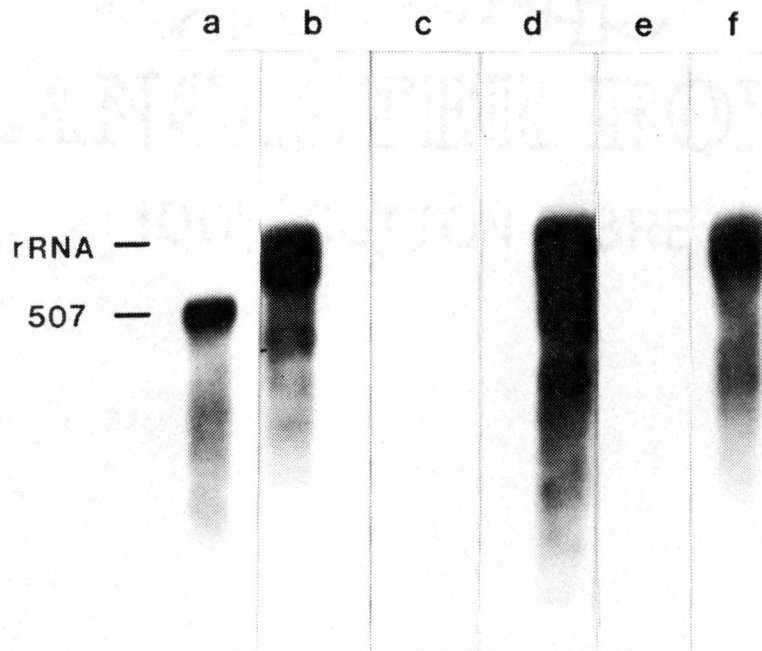


Figure 9. Northern analysis of *eym* (eye missing) and *glass* flies.

Total RNA was extracted from heads of flies as described in Appendix I. Ten micrograms of head total RNA was electrophoresed in formaldehyde agarose gel as follows: Blot a: wild type *D. melanogaster* strain oregon R ; Blot c: *eym* ; Blot e: *glass*. These were hybridized with radiolabelled 3.8 Kb SalI fragment of gene 507mr. To demonstrate the integrity of the RNA in blots not showing hybridization with the 507 gene probe, these blots were rehybridized with a ribosomal gene probe. Lanes b, d and f represent blots a, c and e respectively.

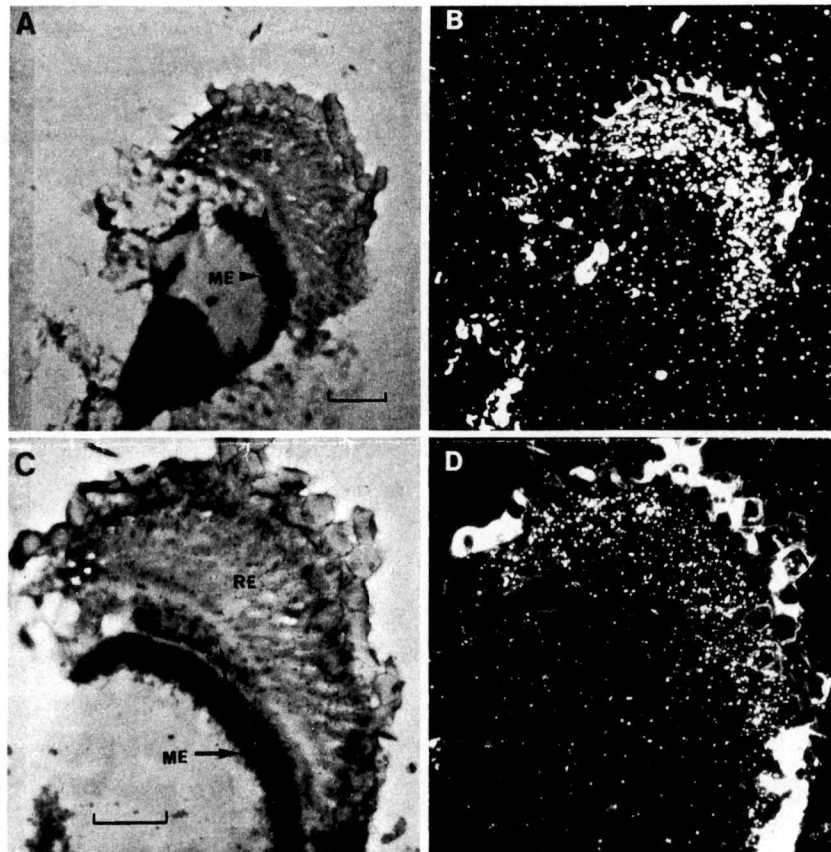


Figure 10. Localization of 507 mRNA to the retina of the adult compound eye.

In situ hybridization of the ^3H -labelled 3.3 EcoRI (Fig. 4) and 3.8 SalI (Fig. 4) DNA fragments to cryostat sections of the heads of *D. melanogaster* (A, B) and *D. miranda* (C, D) respectively. A, C - bright field; B, D - dark field. RE - retina; LA - lamina; ME - medulla. Bars = 100 microns.

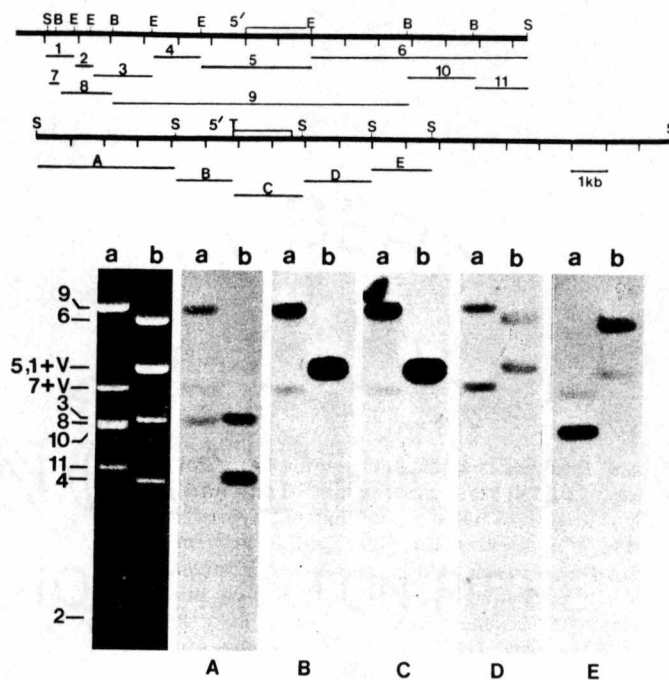


Figure 11. Interspecies Southern-blot hybridization A.

The restriction maps shown have been aligned arbitrarily. A sub-clone of the *Sall* DNA fragment of the 507ml clone was digested separately with *Bam*HI and *Eco*RI to generate target fragments 1 - 11. Five Southern blots, each containing *Bam*HI (lane a) and *Eco*RI (lane b) fragments, were produced by electrophoresing all the restriction digests on the same gel in parallel. For each lane an equivalent amount of DNA was digested. A photograph of the ethidium-bromide stained gel is included to show the migration of each of the targeted DNA fragments. After Southern-transfer, each blot containing lanes a and b was hybridized individually to 10.0×10^6 cpm of ^{32}P -labelled probes A - E from clone 507mr. The designations of the probes used are shown below each corresponding blot. The specific activity of the probes ranged between $80 - 90 \times 10^7$ cpm/microgram. All blots were hybridized in 6X SSC, 5X Denhardt's solution at 68 C, and washed in 0.1X SSC, 0.1% SDS at 60⁰ C. Some probes were contaminated with plasmid vector sequences and therefore hybridize faintly to the vector DNA (V) on the blots.

of 11 different restriction fragments (1-11, Fig. 11, top panel) of clone 507ml. The hybridization conditions were stringent (washes at 0.1X SSC, 0.1% SDS at 60 C). Each probe hybridized only to the corresponding DNA fragments of the 507ml clone as follows: A to 3, 4, 8 and 9 ; B to 5 and 9 ; C to 5 and 9 ; D to 6 and 9 ; and E to 6 and 10 (Fig. 11, bottom panel). Faint hybridization of certain probes to the vector DNA (V, Fig. 11) was the result of a minor contamination of the probe with vector DNA. Fragment 9 shows a reduced intensity of hybridization in all blots presumably because of the inefficiency of transfer of such a large fragment onto the nylon membrane. The pattern of hybridization suggests that the homology between 507mr and 507ml is conserved in a collinear fashion and no major sequence rearrangement involving gene 507 has occurred since the divergence of the *obscura* and the *melanogaster* groups. These data do not allow the quantitation of the absolute level of homology between conserved regions. Nevertheless, since all the probes used in the above experiment had a similar specific activity ($8-9 \times 10^7$ cpm/ μ g) and an equal amount of each probe was hybridized to identical DNA blots, the relative difference in hybridization signal (Fig. 10) can be used as a measure of the relative level of homology in various regions of the 507ml and 507mr DNA. The results shown in Fig. 11 suggest that the highest level of homology lie within the transcribed regions (compare blots B and C with others, Fig. 11). The next level of homology is found between fragments 4 and A (blot A, Fig. 11) and fragments 10 and E (blot E, Fig. 11). Hybridization intensity appears to be lowest between probe D and fragment 6 (blot D, Fig. 11). Interestingly hybridization of probe D with fragment 9 is stronger than that with fragment 6 (blot D, Fig. 11) and in fact probe D does hybridize strongly with fragment 5 (not shown). This suggests that sequences within the 5' end of probe D are highly conserved in the two species and are present within fragment 5 of 507ml. Since fragment E hybridizes strongly with fragment 10 (blot E, Fig. 11), it can be said that the region represented by the 5' half of fragment 6 of 507ml and that by the 3' portion of probe D of 507mr represents an island of divergence between these two clones. In the region upstream of the 5' end of the transcribed region, probe A hybridizes only weakly to fragment 8 (blot A, Fig. 11) although the length of probe A must include the region represented by fragment 8 on a colinearity basis.

The 3.3-kb EcoRI fragment of clone 507ml and the 3.8-kb SalI fragment of clone 507mr appear to have all of the transcribed portion of gene 507, and presumably fragment B of 507mr (Fig. 11) and the 5' half of fragment 5 of 507ml (Fig. 11) represent the immediate 5' non-transcribed portion of this gene. In order to compare these regions by cross hybridization, the 3.3-kb EcoRI fragment of 507ml was digested with BglII plus EcoRI and PstI plus Aval to generate nine different fragments (Fig. 12, top panel). The resulting DNA blots of these fragments were then probed with fragment B of Fig.11. In the case of EcoRI + BglII digested DNA, only fragments 1, 2 and 3 hybridized to fragment B, with fragment 1 showing the highest signal intensity (Fig. 12.lane a). In the case of the PstI+Aval digested DNA fragments 5, 7, and 8 hybridized to fragment B with fragment 7 showing the highest signal intensity (Fig. 12, lane b). Comparison of the hybridization intensities with various fragments shown in Fig.12 suggests that strong homology exists in the area overlapped by fragments 1 and 7 (shown between arrows, Fig. 12) of the 507ml gene and downstream from this region the homology between 507ml and 507mr decreases until the transcribed regions are reached.

The homology between transcribed regions of this gene was also tested by inter-species northern blot hybridization (Fig. 13). When a DNA fragment of one species containing the entire transcribed portion of gene 507 was used as a probe, the expected 1.8 Kb mRNA was detected in both species (not shown). Under similar conditions of hybridization however, the 3' transcribed portions of gene 507 in the two species do not cross hybridize, a feature that had emerged during the transcription mapping experiments described earlier. These portions are represented by the KpnI-SalI fragment within the 3.8-kb SalI fragment of *D. miranda* (Fig. 13B) and the HindIII-EcoRI fragment within the 3.3-kb EcoRI fragment of *D. melanogaster* (Fig. 13C). When the KpnI-SalI fragment of *D. miranda* is used as the probe, it hybridizes to a 1.8-kb mRNA from *D. miranda* but fails to detect any mRNA in *D. melanogaster* (Fig. 13A, lanes 1, 2). However when lane 2 was reprobed with the 5' portion (SstI-KpnI fragment of the 3.8 Kb SalI fragment, Fig. 13B) of the *D. miranda* gene, hybridization is seen with a 1.8 Kb mRNA (Fig. 13A, lane 2'). Similarly the HindIII-EcoRI fragment of gene 507 from *D. melanogaster* (Fig. 13C) did not hybridize with mRNA from *D. miranda* but hybridized to a 1.8 Kb mRNA from

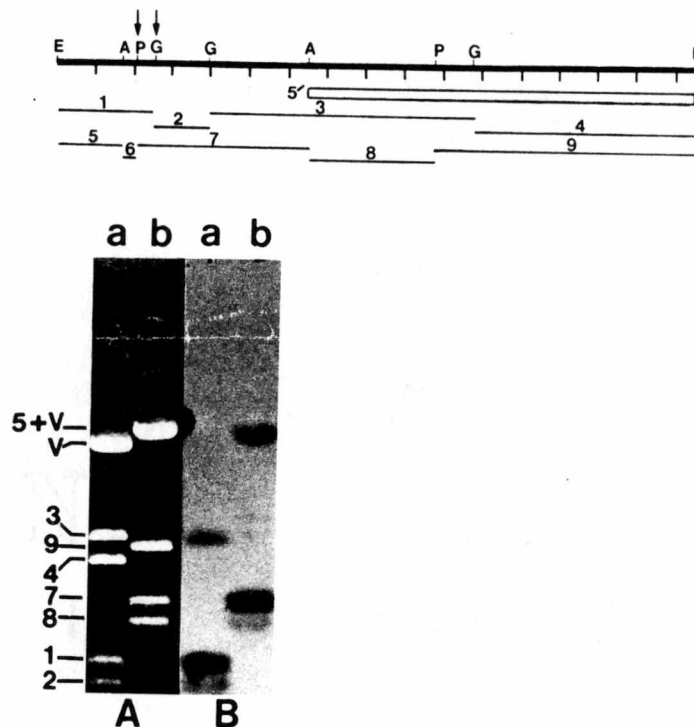


Figure 12. Interspecies Southern-blot hybridization B.

The 3.3 KB EcoRI fragment (Fig. 4) of clone 507ml (Fig. 4) was sub-cloned in the plasmid pBSK (Stratagene) and digested with BglII plus EcoRI and AvaI plus PstI. Products of each digestion were electrophoresed in lanes a and b, respectively, blotted and hybridized with the 1.8 Kb Sall - SstI fragment (fragment B, Fig. 11) representing the 5' non-transcribed DNA of gene 507mr. Conditions for hybridization and washing were the same as in Figure 11. The photograph of the ethidium bromide stained gel showing migration of each fragment is shown in (A). The autoradiogram of the blot is shown in (B). V represents plasmid vector DNA. The arrows indicate the region shared by fragments 1 and 7. The scale shown on the map equals 100 bp.



Figure 13. Interspecies northern-blot hybridization

(A). Poly (A⁺) RNA isolated from adult heads were fractionated on formaldehyde agarose gel, blotted onto nylon membrane and hybridized to ³²P-labelled fragments representing transcribed regions of gene 507ml and 507mr. Lane 1, 3, 5 - *D. miranda* RNA; Lanes 2, 4, 6 - *D. melanogaster* RNA.

(B). Probes used were : Lanes 1, 2 - 3' portion of the 507mr gene (KpnI-SalI fragment of gene 507mr) ; Lanes 3, 4 - 3' portion of the 507ml gene (HindIII-EcoRI fragment of gene 507ml).

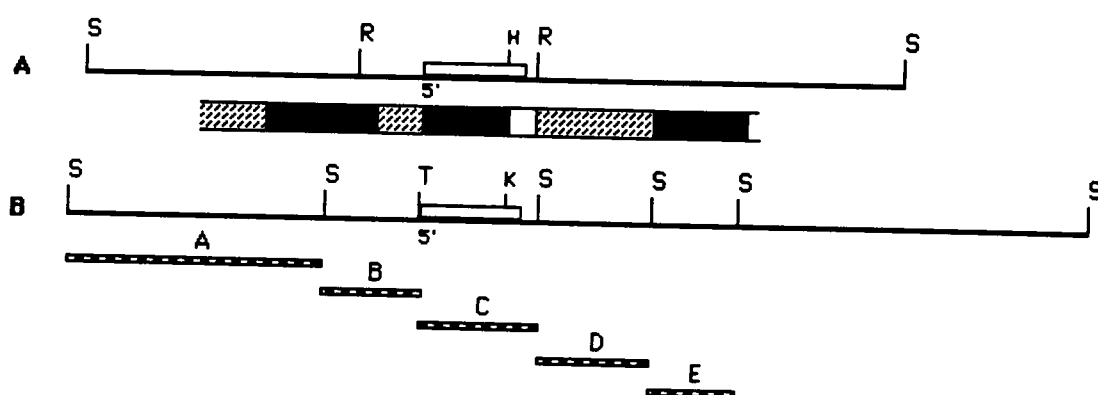


Figure 14. Summary of interspecies cross-hybridization experiments.

The clones 507ml and 507mr are shown as A and B respectively. The open rectangle on each map represents the extent of the transcribed region. The alignment of the two clones is only approximate and reflects the results of cross hybridization shown in Fig.11. The bar between the two clones illustrates degree of homology. Solid regions represent high degree of homology, hatched regions indicate relatively less homology and the sole open region indicates absence of homology. A, B, C, D and E were the fragments of clone 507ml used as probes in the cross hybridization experiments shown in Fig. 11.

S - SalI; R - EcoRI; T - SstI; H - HindIII; K - KpnI.

D. melanogaster (Fig. 13A, lanes 3, 4). However when lane 3 was reprobed with the 5' portion of the 507 gene of (AvaI-HindIII fragment of 3.3 Kb EcoRI, Fig. 13C) of *D. melanogaster*, hybridization was seen with a 1.8 Kb mRNA of *D. miranda* (Fig. 13A, lane 3'). These results suggest that while the 5' portions of gene 507 in these two species is highly conserved, the 3' portions appear to have very little homology. However, at this stage the 3' portions tested in these hybridizations have not been characterized in terms of distribution of coding and non-coding sequences.

DISCUSSION

The clones 507ml and 507mr, from *D. melanogaster* and *D. miranda* respectively, appear to contain the entire transcribed region and some flanking, non-transcribed regions of gene 507. While clone 507ml maps to a single site at 66D on the autosome 3L of the *D. melanogaster* genome, the homologous clone 507mr in *D. miranda* maps to a single site, 12A, on the X1R chromosome in accordance with the homology between autosome 3L of *D. melanogaster* and the X1R chromosome of *D. miranda*. Thus in *D. miranda* there are two doses of gene 507 in the female and one dose in the male. However, the steady state level of 507 mRNA in the two sexes is equivalent indicating that, on a per-gene dose basis, males have twice the amount of 507 mRNA relative to females. These findings establish that gene 507 is dosage compensated in *D. miranda*. Earlier enzymatic and transcriptional autoradiographic studies have demonstrated that the X1R chromosome of *D. miranda* and the XR chromosome of its sibling species *D. pseudobscura*, are dosage compensated (Abraham and Lucchesi 1974 ; Mukherjee and Chatterjee 1975 ; Strobel et al. 1978). Gene 507, represents the first cloned X1R linked gene from *D. miranda* to be shown to be dosage compensated at the steady state mRNA level. Thus, gene 507mr has come under the influence of the cis-acting sequences that mediate male-specific transcriptional hyperactivation.

The absence of 507 mRNA in *eym* (eye missing) and *glass* (photoreceptor cells absent) flies suggests that gene 507 may be transcribed only in the retinal photoreceptor cells. In accordance with this, is the *in situ* localization of 507 mRNA to the retina of the

compound eye of both *D. melanogaster* and *D. miranda*. According to the currently prevailing model of transcriptional control in eukaryotes, it can be assumed that specific cis-acting sequences of the 507 transcription unit mediate the appropriate developmental stage and tissue specificity of transcription of this gene, and further that such sequences would be conserved between the two species. In *D. miranda*, however, the 507 transcription unit has come under the additional influence of a cis-acting element(s) that mediates male-specific transcriptional hyperactivation, and whose actions have to be functionally superimposed on those mediating developmental and tissue specificity of transcription.

The comparison of nucleic-acid sequence homology, as judged by restriction mapping and cross-species hybridization, between clones 507mr and 507ml has yielded some interesting features of these clones (Fig. 14). Within the transcribed regions of gene 507, the 5' ends are highly conserved while the 3' ends appear to have diverged completely. Within non-transcribed regions the sequences approximately 0.8-kb upstream, and 3.0-kb downstream of the transcribed region show strong sequence homology. Functionally constrained sequences are believed to be conserved in evolution and it is possible that these highly conserved regions play a regulatory role in the 507 transcription unit. On the other hand these highly conserved regions are separated from the transcribed regions by relatively less conserved regions. These latter regions may represent those that are either not of functional significance or that represent changes to accommodate the X linked status and dosage compensation of gene 507mr.

The localization of the transcriptional control elements of gene 507mr is therefore of considerable significance to the basic concerns of the molecular mechanisms involved in dosage compensation in *Drosophila*. Moreover, it is possible that the comparison of the organization of cis-acting transcriptional control elements of gene 507mr and 507ml would shed some light on the mode of evolution of the cis-acting sequences that mediate male-specific transcriptional hyperactivation of gene 507mr. Towards the further characterization of gene 507 on these fronts transgenic lines of *D. melanogaster* carrying the gene 507 of *D. miranda* at single autosomal sites, were synthesized to facilitate mapping of transcriptional control elements of this gene. The results of these experiments

are presented in the next chapter. Also, simultaneously, the nucleotide sequences of the 3.3-kb EcoRI fragment of gene *507ml* and the 3.8-kb SalI fragment of gene *507mr* were determined and compared. The results of this analysis are presented in Chapter 4.

CHAPTER 3

RETINA-SPECIFIC EXPRESSION AND MALE-SPECIFIC TRANSCRIPTIONAL HYPERACTIVATION OF GENE 507 OF *D. MIRANDA* WHEN EXPRESSED FROM AUTOSOMAL SITES IN *D. MELANOGASTER*

INTRODUCTION

X-linked genes in *Drosophila* are in unequal dosage in the two sexes and are therefore subject to dosage compensation, a mechanism that ensures equivalence of X-linked gene products in the two sexes despite the gene dosage difference between them (Muller 1931 ; also see Chapter 1). In *Drosophila*, dosage compensation is mediated through the hyperactivation of transcription of X-linked genes in the males to approximately twice the level in females (Mukherjee and Beerman 1961 ; Chapter 1). This male specific transcriptional hyperactivation is in turn thought to be mediated by cis-acting sequences that interact with trans-acting factors in a male-limited manner (Reviewed by Luchessi and Manning 1987; see chapter 1). However, no such cis-acting sequence(s) that mediates male-specific hyperactivation has been identified so far and as such, considerable confusion exists with respect to the physical and functional relationship of such sequences with the genes they influence. A strategy for the investigation of the properties of the proposed dosage compensatory cis-acting sequences has been the relocation of an X-linked gene to an autosome with the help of P-element mediated transformation (Rubin and Spradling 1982). Also, although not initially carried out in the context of dosage compensation, a few autosomal genes have been relocated to the X chromosome and have subsequently added to the information on hand regarding cis-acting dosage compensatory elements. The activity of the relocated gene in either of these two kinds of relocation is assayed for sex-specific differences, and hyperactivity in males

relative to females is indicative of the effect of cis-acting sequences that ordinarily serve to mediate dosage compensation. In such studies autosomal genes when inserted into the X chromosome have often shown male-specific hyperactivity. When the autosomal *Xdh* gene was transposed to the X chromosome and assayed at the level of enzyme activity, it was found to be an average 1.6 fold more active relative to females (Spradling and Rubin 1983 ; Laurie-Ahlberg and Stam 1987). Similarly assays of enzyme activity in each sex showed that an X-linked insertion of the autosomal *Ddc* gene showed a 1.4-1.6 fold higher activity in males than in females. Interestingly, this hyperactivity was restricted to older adults and was absent in pupae and freshly eclosed adults (Scholnick et al. 1983 ; Marsh et al. 1985). The autosomal *Adh* gene showed the typical two-fold hyperactivation in male larvae when expressed from X-linked sites (Sass and Meselson 1991). Interestingly, this contrasted with earlier observations that in some cases X-linked insertions of the *Adh* gene which had considerably larger stretches of flanking DNA sequences, were only partially compensated in larvae (Laurie-Ahlberg and Stam 1987) while in others they failed to be hyperactivated in larvae and in adults (Goldberg et al. 1983). A possible explanation proposed by Sass and Meselson (1991) is that the increased amounts of flanking sequences could have "insulated" the *Adh* gene from the long-range action of the dosage compensatory sequences. There is also the possibility that insertions at certain sites on the X chromosome are not dosage compensated as is the case for an insertion of the dosage-compensated X-linked gene *Sgs-4* at an ectopic site on the X chromosome. Hiebert and Birchler (1992) have demonstrated that the promoter of a retrotransposon, *copia*, when integrated into the X-linked gene *white*, exhibits dosage compensation.

The above observations have collectively suggested that the cis-acting sequences mediating male-specific hyperactivation have a long range action that could influence any promoter that falls within this range, a characteristic of many enhancer sequences. Further they suggest that such sequences may not be linked to individual transcription units but are distributed over the X chromosome, functioning at the level of the chromosome in organizing regions of transcriptional hyperactivity. Thus, of interest are the effects of relocation of an X-linked gene that normally is not dosage compensated, to other regions

on the X chromosome. Normally not dosage-compensated, the X-linked gene, *Lsp1* did show male-specific hyperactivation when at ectopic sites on the X chromosome (Ghosh et al. 1989). This suggests that in its normal location *Lsp1* is, in some way shielded from the effects of cis-acting sequences mediating dosage compensation despite the capability of its promoter to interact successfully with such sequences. This observation is contrary to the idea that dosage compensatory cis-acting sequences are tightly linked to the promoters they influence and that each gene is hyperactivated by its own individual dosage compensatory cis-acting sequence. On the other hand, each dosage-compensated X-linked gene could have its own dosage-compensatory cis-acting sequences in proximity to its regulatory region, but these sequences could also be capable of a long-range effect. In this context it becomes interesting to observe the effects of autosomal relocation of a dosage compensated X-linked transcription unit. On this front, each dosage compensated X-linked gene to be analyzed by autosomal relocation has been shown to retain the property of male-specific hyperactivation when relocated to autosomal sites, even when accompanied by only small lengths of flanking DNA. Studies of the expression of the X-linked *white* gene from autosomal positions, have demonstrated that the region extending from the transcription start site to -216 bp upstream causes male-specific hyperactivation of autosomal inserts (Pirrotta et al. 1985). In these studies, however, gene activity was assayed in terms of level of eye pigmentation and no measurements were made of the steady state mRNA levels. Sequences mediating male specific transcriptional hyperactivation of another X-linked gene *Sgs-4*, have been localized to within -838 bp upstream of the transcriptional start site and 130 bp downstream of the polyadenylation site (Krumm et al. 1985 ; McNabb and Beckendorf 1986 ; Hofmann et al. 1987). In these studies both protein as well as mRNA levels were quantitated and expressed in terms of the endogenous autosomal *Sgs-3* gene. A "housekeeping" X-linked gene *6Pgd* carried as a 7.4 kb genomic fragment has been shown to be male-specifically hyperactivated 1.5-2.0 fold at the steady-state mRNA as well as protein level when expressed from autosomal sites (Gutierrez et al. 1989). All these studies involved genes of the species *D. melanogaster*. In the only such study of an X-linked gene from another species, the XR linked *Hsp82* gene of *D. pseudobscura* has been shown to be hyperactivated in adult

males relative to females when relocated to autosomal sites in *D. melanogaster* accompanied by 5.0-kb of 5' upstream and 3.0-kb of 3' downstream sequences (Jaffe and Laird 1986). The same transgenic lines however failed to be hyperactivated when either heat-shocked or non-heat shocked larvae were assayed (Sass and Messelson 1991). In every case described a certain degree of variability with respect to level of hyperactivation existed between lines differing in the site of insertion. The range of variability extended from partial 1.5-fold to complete, 2.0-fold, hyperactivation for each gene. In the case of *white* and *Sgs-4* genes some lines also showed over compensation, seen as a more-than-two-fold hyperactivation in males. It is possible that the response of an autosomally inserted X-linked gene to the dosage-compensation machinery, is influenced by its new genomic location so that transcriptional hyperactivity is not the ideal two-fold and / or is subject to tissue- or developmental-stage-specific effects. This influence may be compounded by increasing complexity of the regulatory interactions that, independently of dosage compensation, are responsible for the expression patterns of the gene in question. This latter problem has perhaps complicated the further analysis of the dosage compensatory sequences of the *white* and *sgs-4* genes despite the localization of the regions in which their dosage compensatory sequences may reside. In the case of the former, multiple and separable cis-acting sequences direct its expression different tissues at different developmental stages (Hazelrigg et al. 1984 ; Levis et al 1985 ; Pirodda et al. 1985). Moreover, perhaps due to a difficulty in quantitating transcripts, no analysis of autosomal insertions of the *white* gene has been carried out at the steady-state mRNA level and the reliance on phenotypic pigmentation levels quantitation of transcriptional activity may have introduced an additional degree of complexity. The *Sgs-4* gene on the other hand, is transiently expressed in the late third instar larval period and its expression is induced by the steroid hormone ecdysone (Korge 1975). Autosomal insertions show a variability in level of expression and dosage compensation even between individuals of the same line (McNabb and Beckendorf 1986). Alleles of this gene exist that show impaired levels of activity and dosage compensation even at its resident site on the X chromosome (Korge 1981). While these problems may have confounded efforts to map the dosage-compensatory sequences of these genes, the P-element mediated relocation

studies involving these genes, along with those involving *6Pgd* and *Hsp82*, mentioned above, do suggest that essential cis-acting elements that mediate dosage compensation may reside close to transcription units and that they function even when removed from the X chromosome into an autosomal background.

Relocation of individual X-linked genes to autosomal positions therefore continues to be of promise in the identification and characterization of dosage-compensatory cis-acting sequences, the major long term goal arising from this dissertation. It is useful to focus on a gene that is transcribed in a single kind of tissue in adults and whose transcript is easily quantitated. These features are significant improvements over the use of the *white* and *sgs-4* genes of *D. melanogaster* and also different from the "housekeeping" gene *Pgd*, since it is possible to study the relationship between the dosage compensatory sequences and those that mediate tissue specificity. Furthermore, this study has focussed on the X1R chromosome of *D. miranda* in order to include an additional perspective of the mode of evolution of dosage compensatory sequences. The X1R chromosome of *D. miranda* is homologous to an autosome, 3L, of *D. melanogaster* giving rise to the interesting situation of a gene X-linked and dosage-compensated in one species while being autosomal in another. The isolation and characterization of a retina-specific gene 507, which is 3L-linked in *D. melanogaster* while being X1R-linked, and dosage-compensated in *D. miranda* are described in Chapters 2. It has been shown that a 3.8-kb *SalI* fragment of gene 507 of *D. miranda* (507mr) appears to have the entire transcribed region of the gene 507 in this species (Chapter 2). This 3.8-kb *SalI* fragment was introduced into autosomal positions of *D. melanogaster* by P-element mediated germline transformation. The lack of a suitable system for germline transformation of *D. miranda* necessitates this cross-species approach but it was expected that the dosage-compensatory mechanism is highly conserved in this genus, as is the case among the mammals. The results show that the cis-acting elements necessary for correct retina specific expression of gene 507mr are present within the 3.8-kb *SalI* fragment carrying the intact transcription unit. The steady state level of transgene derived 507 mRNA is 1.4 - 1.6 fold greater in males than in females suggesting a partial response to the dosage-compensatory machinery of *D. melanogaster*.

MATERIALS AND METHODS

Drosophila strains and cultures

The host strain for P-element mediated transformation, *ry*⁵⁰⁶, was received from Dr.J. Lucchesi, University of North Carolina. Flies were raised on cornmeal-agar-molasses medium at 23° C and 60% humidity. Balancer stocks in a *ry* background were obtained from the Drosophila stock center in Bloomington Indiana.

P-element vector, helper plasmid, and transformation constructs

The P-element plasmid vector pDM30 (Mismar and Rubin 1987), and the transposase encoding helper plasmid pn25.7wc (Karess and Rubin 1984) were received from Dr. Jerry Rubin, University of California, Berkeley. pDM30 vector carries the ampicillin resistance gene and bacterial replication origin of pUC to facilitate cloning and propagation in *E.coli* and the *Drosophila* marker gene *ry*⁺ to allow easy identification of transformants among *ry* / *ry* hosts. The 3.8-kb Sall fragment of gene *507mr* (Fig.15) was cloned into the unique Sall site of pDM30, to yield the transgene construct pDM/3.8Smr. shown in Fig.15.

Germline Transformation

The germline transformation procedure used was described by Karess and Rubin (1984). The transgene construct pDM/3.8Smr and the helper plasmid pn25.7wc were co-introduced into pre-blastoderm embryos by microinjection as described by Karess (1984). The details are supplied in the Appendix. Injected individuals that proceeded to adulthood, were individually mated with host strain flies. The progeny of these matings, the G1 flies, were screened for the presence of flies with *ry*⁺ eye colour. G1 flies showing *ry*⁺ eye colour were then individually mated with host strain flies of the opposite sex to establish G2 isolines. Each G2 isoline was propagated by back crossing with the host strain. The transformants were crossed to balancer stocks to determine chromosomal linkage, and homozygous flies were then derived from balanced stocks (Appendix). Chromosomal locations of the inserts were determined by *in situ* hybridization of a transgene-specific probe (Fig. 15) to polytene chromosomes of homozygous larvae. In all six different single

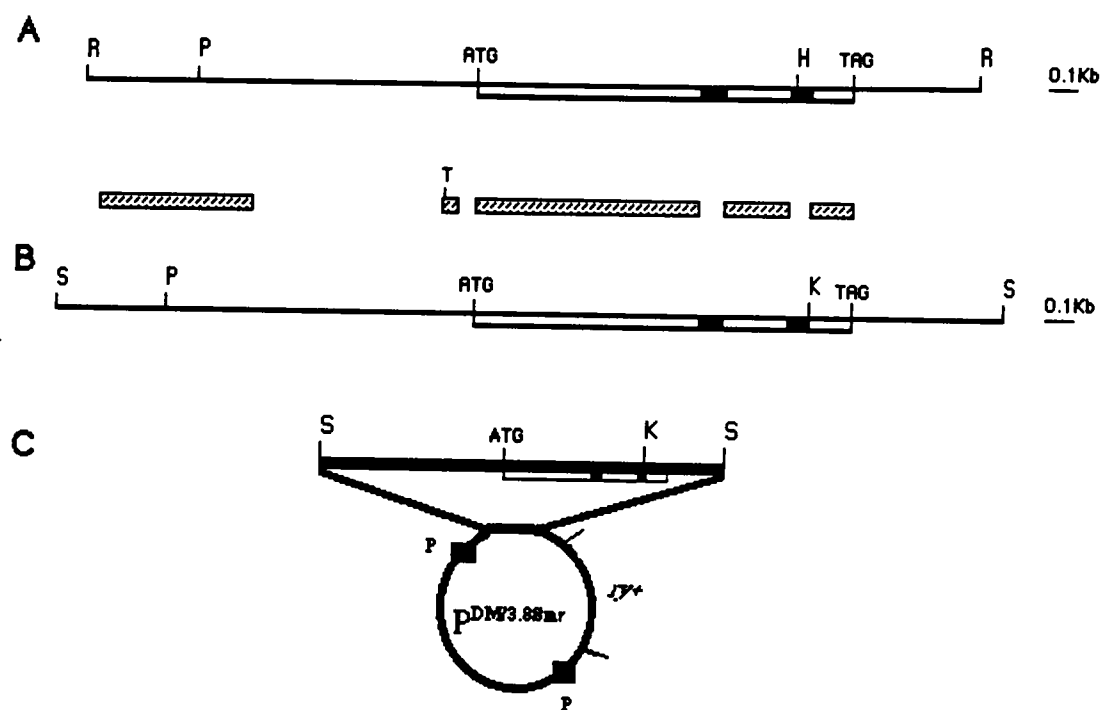


Figure 15. The genomic fragments containing the gene 507 of *D. melanogaster* and *D. miranda*.

A) The 3.3-kb EcoRI fragment containing the gene 507 of *D. melanogaster*.; B) The 3.8-kb SalI fragment containing gene 507 of *D. miranda*.; C) The P-element mediated transformation plasmid pDM/3.8Snr carrying the 3.8-kb SalI fragment containing gene 507 of *D. miranda* along with the marker gene *ry*.

The features shown on maps A and B are derived from the nucleotide sequences of the two fragments and of the 507 cDNA in *D. melanogaster* (presented in Chapter 4). Maps A and B are aligned with respect to their translational start sites. The rectangle associated with each map designates the extent of the coding region of the gene. Hatched boxes between the two maps show regions of high nucleotide sequence homology between the two species. The small box just upstream of the ATG codon is a highly conserved seventy base pair stretch. Within it an identically conserved 24 bp. stretch contains a TATA element and is marked as T.

insert lines were obtained, of which three have been characterized and assayed in this study. The genetic crosses are described in the Appendix.

***In situ* hybridization to polytene chromosome**

Polytene chromosome preparations were made from third instar larvae of homozygous transformed lines as described in Ashburner (1989a). The transgene specific KpnI-Sall fragment of gene 507mr was labelled with ^{35}S to a specific activity of 2×10^7 cpm / microgram by random-hexamer priming. Hybridization and washing were carried out as described in Ashburner (1989b). The procedures are detailed in the Appendix.

Northern-blot analysis of transformed lines

Each of the three transformed lines was assayed for tissue-specificity of expression and steady-state level of transgene-derived mRNA in the two sexes. Total RNA was isolated independently from heads and bodies of adult flies using the hot phenol method (Helms et al. 1985). In the case of assays for male-specific hyperactivation, flies were first sorted into sexes before decapitation and RNA isolation. Total RNA was electrophoresed on formaldehyde agarose gels and blotted onto nylon membranes according to standard procedures (Maniatis et al. 1982). Appropriate ^{32}P labelled probes (see Results) were made by random-hexamer priming. Hybridization and washing was carried out as described in Maniatis et al. (1982) (see Appendix). Quantitation of hybridization intensity was done by laser densitometric scanning of autoradiographic bands. Multiple non-saturating exposures and duplicate scans of each exposure were used to ensure linearity and to eliminate possible scanning surface aberrations. The procedures are detailed in the Appendix.

RESULTS

Transgene specific probe does not cross-hybridize with 507 mRNA of host strain.

It has already been demonstrated in chapter 2 that the KpnI-SalI fragment representing the 3' portion of the transcribed region of gene *507mr* does not cross hybridize with gene *507* of *D. melanogaster* strain Oregon R. Though the recipient, *ry*⁵⁰⁶ flies used in this study are of the species *D. melanogaster*, they have a predominantly *Canton S* genetic background (J. C. Luchessi, pers. comm.). It was necessary therefore to confirm that the transgene specific probe did not cross hybridize with 507 mRNA of this recipient strain. Fig. 16 shows the results of a northern blot analysis of head total mRNA from the host *ry*⁵⁰⁶ flies. As expected the KpnI-SalI fragment of gene *507mr* failed to hybridize to any mRNA from these flies (lane a, Fig. 16). This lane was reprobbed with the *Adh* gene probe to demonstrate integrity of the RNA (lane b, Fig. 16). The endogenous *507ml* gene is expressed in these flies as seen by the hybridization of the HindIII-EcoRI fragment of gene *507ml* with the appropriate 1.8 Kb mRNA (lane b, Fig.16). Thus transcript derived from the endogenous gene *507* of the host *ry*⁵⁰⁶ flies, though present, will not interfere with the analysis of transgenic *507mr* derived transcripts.

Transformant lines 1A, 6A and 10D have a single autosomal insertion of the intact transgene.

Through segregation analysis of balanced lines the chromosomal linkage group of each line was determined (described in Appendix). To confirm the autosomal linkage and to map the location of the transgene, the KpnI-SalI fragment of gene *507mr* was hybridized to polytene chromosome preparations from larvae of homozygous transformant lines. Hybridization was seen only at single autosomal sites in each of the three lines, and no cross hybridization was seen with the endogenous *507ml* gene at location 66D on chromosome 3L (not shown). The autosomal locations of the transgene in the three lines are shown in Table 5.

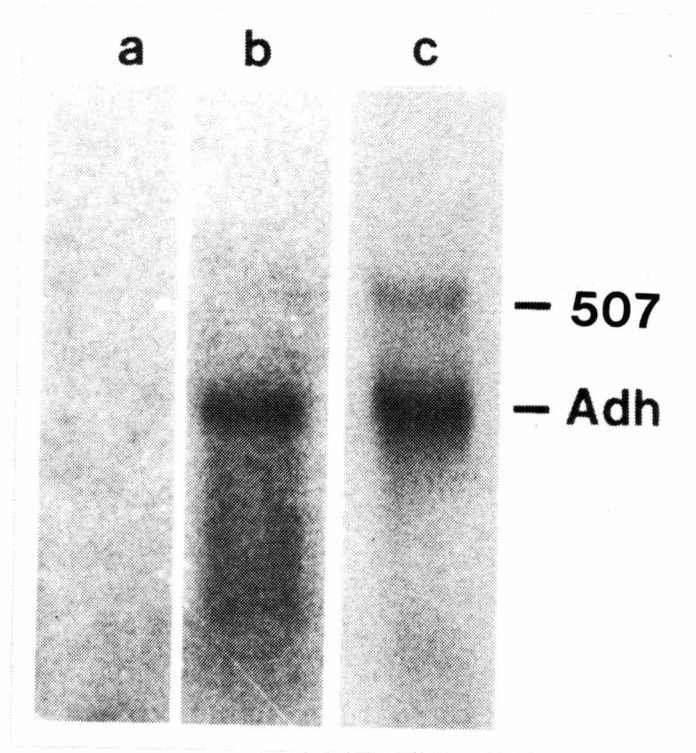


Figure 16. Northern-blot analysis of *ry*⁵⁰⁶ flies with transgene specific probe

Total RNA was extracted from heads of adult *ry*⁵⁰⁶ flies, electrophoresed on formaldehyde gels and blotted on to nylon membrane. Lane (a), hybridization with the ³²P labelled KpnI-SalI fragment of *D. miranda*; Lane (b) rehybridization of lane (a) with ³²P labeled *Adh* gene probe to ensure integrity of mRNA ; Lane (c) Hybridization with ³²P-labeled HindIII-EcoRI fragment of gene *507ml* (Fig. 15) and *Adh* gene probe.

TABLE 5

CHROMOSOMAL LOCATIONS OF TRANSGENE 507MR

<u>Lines</u>	<u>Chromosomes</u>	<u>Locations</u>
1A	2R	44A-C
6A	2R	51D
10D	3R	100A

To demonstrate if the 3.8 Kb *Sall* fragment carrying the transgene in each of the three lines, genomic DNA from homozygous flies was digested with *Sall* and hybridized with the radiolabelled *KpnI-Sall* fragment of gene 507mr (Lanes a, b and c, Fig. 17). Also electrophoresed and hybridized in parallel was the *Sall* digested genomic DNA of the host strain *ry*⁵⁰⁶ (lane d, Fig. 17) and of *D. miranda* (lane e, Fig. 17). In each transformant line hybridization is seen with the appropriate 3.8 Kb fragment (Lanes a, b and c, Fig. 17), identical to the hybridization seen with *Sall* restricted *D. miranda* DNA (Lane e, Fig. 17). No hybridization is seen with DNA from the host strain in which the gene 507ml is expected to be within a 14.0 Kb *Sall* fragment (Krishnan 1991) further confirming the specificity of the probe for the transgene. The three transgenic lines 1A, 6A and 10D therefore carry the gene 507mr at a single autosomal site in their genomes, but at different locations among the three. Studying the behaviour of the transgene in the context of three different chromosomal locations could indicate its sensitivity to position effects.

Transgene 507mr is transcribed only in the retina of the adult compound eye.

Gene 507 in both species has been shown to be transcribed only in the retinal photoreceptor cells of the adult compound eye (Chapter 2). To assay for the appropriate tissue specificity of expression of the transgenic gene 507mr, northern blots of RNA isolated separately from heads and bodies of transformants were hybridized with the *KpnI-Sall* fragment of gene 507mr (Fig. 18). Hybridization is seen with the appropriately sized (1.8 Kb) mRNA in lanes with RNA from heads (left panel, Fig. 18) and not in lanes with RNA from bodies (right panel, Fig. 18). All the lanes were counterprobed with the *Adh* gene probe to demonstrate integrity of the RNA (Fig. 18).

To further test whether the expression of the transgene is restricted to the retina of the compound eye, cryostat sections of the heads of transformants were hybridized *in situ* to radiolabelled *KpnI-Sall* fragment of gene 507mr (Fig. 19). All three lines, 1A, 6A (Fig. 19) and 10D (not shown), showed localization of transgene-derived 507 mRNA to the retina. To confirm that hybridization seen is specific to the transgene derived mRNA, control head sections of the host *ry*⁵⁰⁶ were hybridized in parallel. No hybridization above background levels was seen in such controls (Fig. 19), confirming that the transgenic

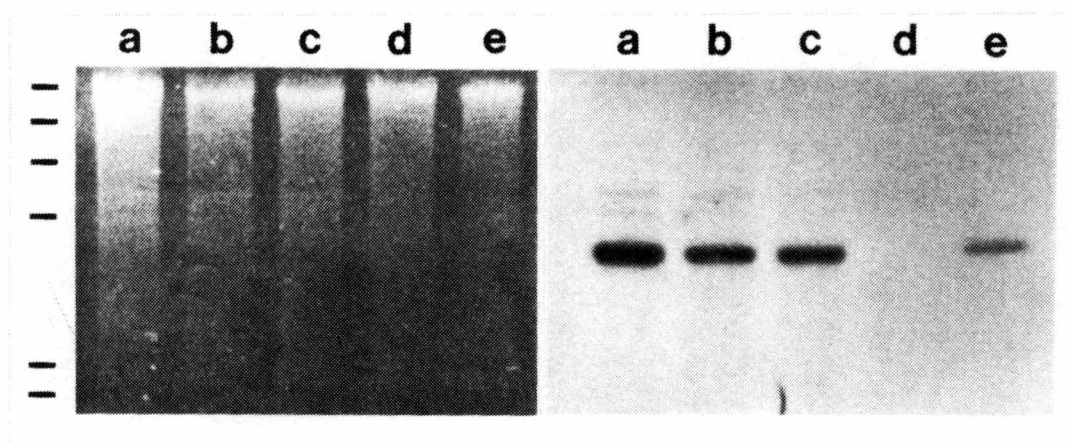


Figure 17. Southern-blot analysis of transformant lines

Genomic DNA was isolated from homozygous flies of the three transformant lines , the host *ry*⁵⁰⁶, and *D. miranda* , digested with Sall and the resulting Southern-blot hybridized with transgene specific KpnI-Sall fragment (Fig. 15). **a** - Line 1A ; **b** - Line 6A ; **c** - Line 10D ; **d** - *ry*⁵⁰⁶ ; **e** - *D. miranda*. The left panel shows the ethidium bromide-stained gel prior to blotting and the right panel shows the autoradiogram. Sizes of the marker DNA indicated on left are from top: 23.1, 9.4, 6.6, 4.4, 2.3 and 2.0 kb.

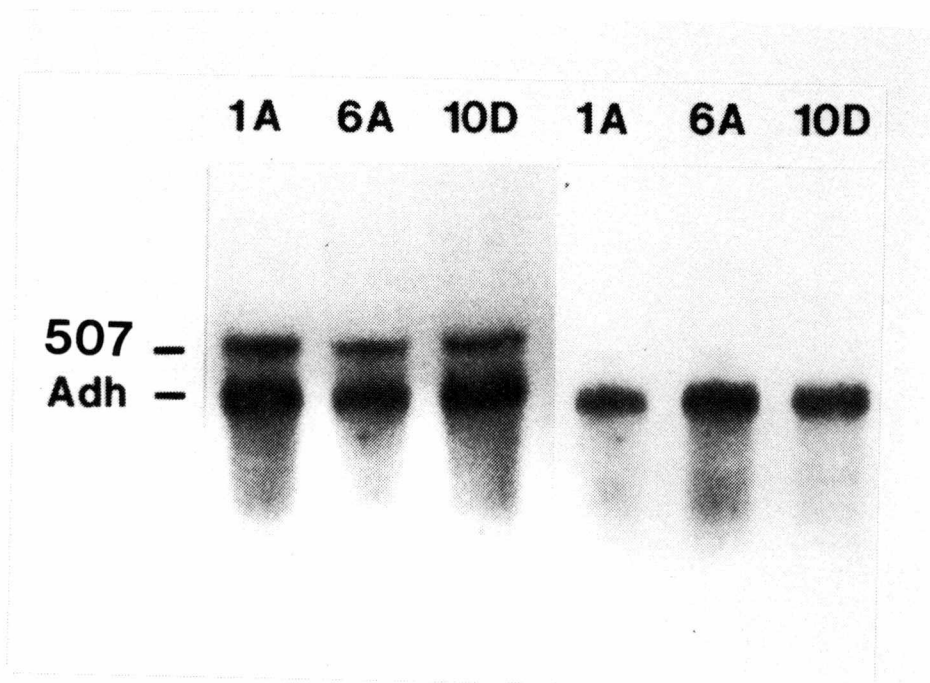
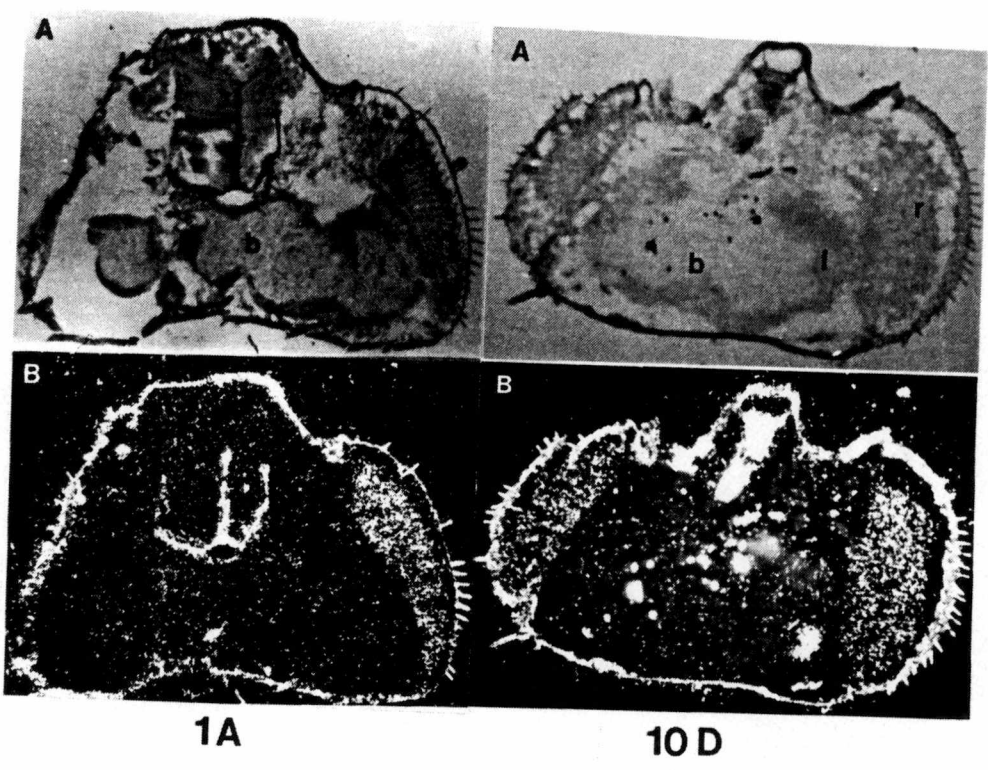
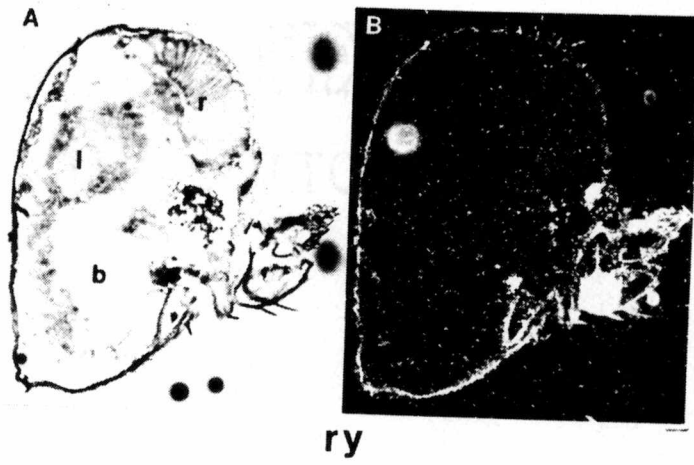


Figure 18. Qualitative northern-blot analysis of transformant lines.

Total RNA isolated from the **head** (left panel) and **body** (right panel) of each line was probed with a mixture of the transgene specific KpnI-SalI fragment (Fig. 15) and the *Adh* gene probe. Since the concentration of *Adh* mRNA in the body is very high, three times less *Adh* gene probe was used to hybridize the blot with body mRNA.

Figure 19. Localization of transgene-derived mRNA to retina of transformants.

³⁵S-labeled transgene-specific KpnI-SalI fragment was hybridized with cryostat sections of heads of transformant lines 1A and 10D and of *ry*⁵⁰⁶. A - bright field; B - dark field; r - retina; l - lobula; b - brain. No hybridization is seen with sections of *ry*⁵⁰⁶.



507mr in these transformant lines mimics its wild-type expression pattern, strongly suggesting that all the cis-acting sequences needed for correct tissue specific expression of gene *507mr* are present within the 3.8 Kb *Sall* fragment used in this study.

Quantitation of steady-state level of transgene derived 507 mRNA in the two sexes.

In order to test whether sequences mediating male-specific hyperactivation in response to the dosage-compensatory machinery of *D. miranda* are present within the 3.8-kb *Sall* fragment used as the transgene, the steady-state level of transgene derived 507 mRNA was determined in each sex of the three transformed lines. Northern-blot analysis of total RNA from heads of adult flies was carried out as described in Materials and Methods. Blots with equal amounts of head total RNA from adult flies were first hybridized with the transgene specific *KpnI-Sall* fragment. To ensure specificity of the probe for transgene derived mRNA, head total mRNA from the host *ry506* flies were hybridized in parallel. After hybridization with the gene *507mr* specific probe and exposure, the blots were rehybridized with the autosomal *Adh* gene probe whose transcript is known to be in equal amounts in the two sexes of both *D. melanogaster* and *D. miranda* (Ganguly et al. 1985 ; Guitierrez et al. 1989; Krishnan et al.1991; Sass and Meselson 1991). The use of *Adh* as an internal standard thus normalizes the hybridization intensities of 507 mRNA which may otherwise remain uncorrected for quantitation and loading differences between samples being compared. A representative northern blot is shown in Fig. 20. For lines 1A and 6A the normalized male / female ratios of transgene derived 507 mRNA are the mean of four independent RNA extractions and quantitations while the male / female ratio for line 10D is from a sample size of two (Table 6). One of the RNA samples used in the above experiment (Sample C, Table 6) was also used to quantitate the level of endogenous 507ml mRNA in the two sexes of each transformed line (Table 7). The quantitation was done identically as for the transgene-derived 507mr mRNA except that the probe used was the *HindIII-EcoRI* fragment of gene *507ml* (Fig. 15) which is specific for the endogenous 507 and does not cross hybridize with 507 mRNA of *D. miranda* (Krishnan et al. 1991 ; Chapter 2) the normalized ratios of the endogenous 507 mRNA obtained from this experiment are shown in Table 7. Previously

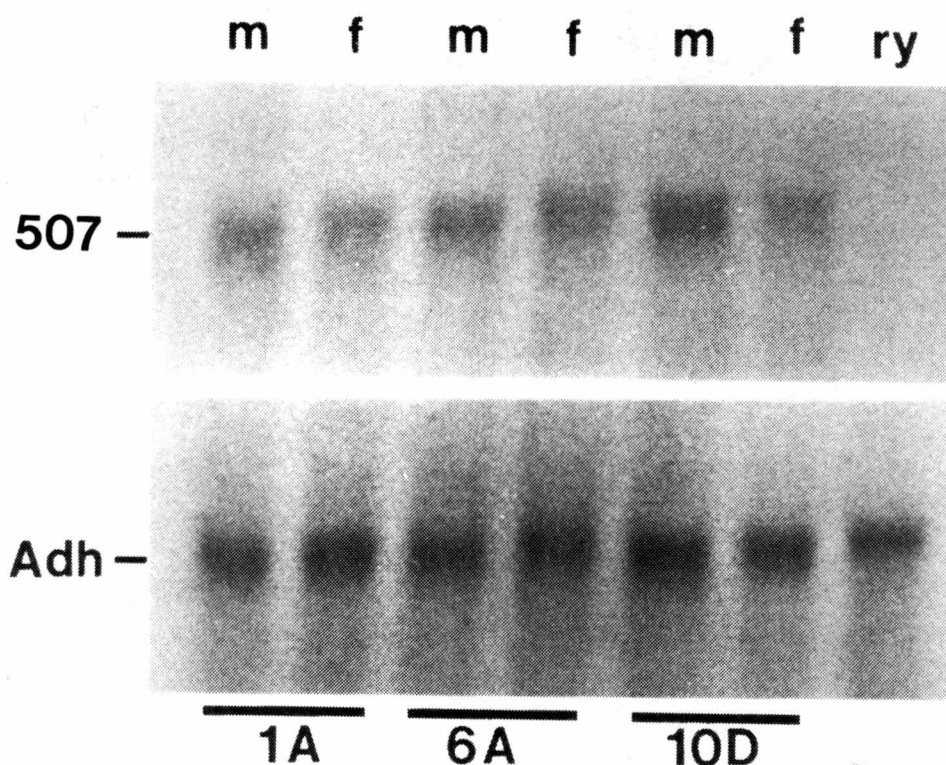


Figure 20. Quantitative sex-specific northern-blot analysis of transformant lines

A representative northern blot used to determine the steady state level of transgene derived mRNA in the two sexes of each transformant line, 1A, 6A and 10D. 15 ug of head total RNA from males (**m**) and females (**f**) of transformant and control *ry*⁵⁰⁶ flies were electrophoresed in parallel on formaldehyde agarose gels. Northern blots were first hybridized with the ³²P-labeled transgene specific KpnI-SalI fragment (Fig. 15). After exposure (**top panel**) the blots were rehybridized to the *Adh* gene probe (**bottom panel**). Radioactive cpm of the *Adh* probe used adjusted to get appropriate exposures of *Adh* bands in a much shorter exposure time than was required for 507 bands. This ensured non-interference from previous 507 derived signals in the quantitation of *Adh* derived signals. Appropriate non-saturated exposures were scanned by laser densitometry using the LKB 2222-010 Ultrascan XL densitometer and scans were analysed with the LKB 2400 Gelscan XL software. The values of each band are shown in Table 6.

TABLE 6
STEADY STATE LEVELS OF 507MR mRNA IN TRANSFORMANTS

Lines	RNA samples *	Male				Female				a/b	Mean±S.D
		50Z	Adh	507/Adh (a)	50Z	Adh	507/Adh(b)				
6A	J6	4.40	8.57	0.514	1.59	5.03	0.317			1.62	
	C6	5.86	12.64	0.463	4.45	13.62	0.327			1.42	1.56 ± 0.089
	N6	9.86	2.16	4.56	5.89	2.14	2.75			1.65	
	D6	4.96	0.78	6.35	3.80	0.93	4.11			1.54	
1A	J1	1.61	2.04	0.79	1.48	2.69	0.55			1.43	
	C1	5.02	12.65	0.39	3.94	14.00	0.28			1.41	1.42 ± 0.043
	N1	6.42	1.85	3.48	8.25	3.49	2.36			1.47	
	D1	11.33	6.01	1.89	4.67	3.33	1.40			1.35	
10D	J10	2.03	4.86	0.42	1.81	6.00	0.30			1.39	1.46 ± 0.07
	C10	6.77	16.56	0.41	3.55	13.25	0.27			1.53	

*RNA samples with the same alphabetical codes were electrophoresed on the same gel, hybridized together and exposed on the same autoradiogram. Values represent intensity of hybridization as determined by laser densitometric scanning of non-saturated exposures and are adjusted for background. Each value represents a mean of two independent scans.

TABLE 7
STEADY STATE LEVELS OF ENDOGENEOUS 507ML mRNA IN TRANSFORMANTS

Line	RNA samples	Male			Female			a/b	Mean S.D
		50Z	Adh	50Z/Adh (a)	50Z	Adh	50Z/Adh(b)		
6A	C6	6.332	5.4	1.173	4.029	3.154	1.277	0.919	
1A	C1	2.518	2.655	0.948	2.759	3.391	0.814	1.165	1.053 ± 0.11
10D	C10	4.776	5.368	0.89	4.185	5.054	0.828	1.075	

* RNA samples are the same as those of code C shown in Table 6. The three samples were electrophoresed on the same gel, hybridized together, and exposed on the same autoradiogram. Values represent intensity of hybridization as determined by laser densitometric scanning of non-saturated exposures and are adjusted for background. Each value represents a mean of two independent scans.

the 507mr mRNA of has been shown to be in equal amounts in the two sexes in the two sexes of *D. miranda* using the same strategy for quantitation as described here (Krishnan et al.1991; Chapter 2). It is unreasonable therefore to suggest that a sex-specific difference in 507 mRNA levels relative to *Adh* mRNA levels represents a sexually dimorphic characteristic of either gene 507 or the gene *Adh*. Nevertheless, the level of 507ml mRNA relative to the *Adh* mRNA was quantitated in the two sexes of the host *ry506* strain in the same way as described for the transformed lines and found to be equal (not shown). These results therefore establish that in each of the three transformed lines the steady-state level of transgene derived 507 mRNA is 1.4 to 1.6 fold higher in males than in females. The typical male specific hyperactivation in response to the dosage-compensatory machinery would result in a two-fold difference of mRNA level between males and females. These results therefore could be interpreted as a partial response of the transgenic gene 507mr to the dosage-compensatory machinery of *D. melanogaster*. These results strongly suggest that the dosage-compensatory cis-acting sequences of gene 507mr may reside within the 3.8-kb *Sall* fragment used in this study. Precedents do exist for partial dosage compensation of X-linked genes, both at resident as well as at autosomal sites and these are described in the discussion.

DISCUSSION

The X1R-linked gene 507 of *D. miranda* was expressed from three different autosomal sites in *D. melanogaster* using P-element mediated transformation. A 3.8 kb *Sall* fragment containing the intact gene with 1.637 Kb of DNA upstream of the translational start site and 0.517 Kb of DNA downstream of the translational stop codon, appears to have all the cis-acting regulatory sequences necessary for appropriate retina specific expression of gene 507. The interactions that confer tissue-specificity of transcription of gene 507 are therefore well conserved in the two species. The results of this study also demonstrate the transgene 507mr is hyperactivated between 1.4 to 1.6 fold relative to females suggesting a partial response to the dosage compensatory, transcriptional hyperactivation machinery of *D. melanogaster*.

While it is possible that dosage compensatory cis-acting sequences of *D. miranda* do not function optimally in *D. melanogaster*, the partial dosage compensation seen in this study is not without precedent, even for studies entirely within the species *D. melanogaster*. Autosomal insertions of the *white* gene have predominantly shown complete dosage compensation but exceptions showing partial as well as over compensation also existed (Hazelrigg et al. 1984). In the case of the X-linked gene *6Pgd* again autosomal insertions averaged a 1.8-fold hyperactivation with three lines showing a 1.6 fold hyperactivation (Guitierrez et al. 1989). For autosomal insertions of the X-linked *Sgs-4* gene partial dosage compensation, averaging 1.5 fold hyperactivation in males, was the norm though exceptions that showed complete dosage compensation did exist (Krumm et al. 1985 ; Hofmann et al. 1987 ; McNabb and Beckendorf 1986). As mentioned in the Introduction, X-linked insertions of autosomal genes also often show only partial dosage compensation. A change in the level of expression of a relocated gene is a general and common feature of gene relocation experiments and has been ascribed to a position effect exerted on the gene by its new chromosomal environment. Different genes could be expected to have a differential sensitivity to position effects depending on peculiarities of individual promoters. It is reasonable therefore to suggest that the partial dosage compensation seen in the three autosomal insertions of gene *507mr* in this study represent suppressive position effects. It is also possible that the dosage compensatory element(s) is itself, independent of any effect on promoter elements, is susceptible to chromosomal position effects. Perhaps for optimal efficiency of the dosage compensatory element(s) a characteristic environment provided on the X-chromosome is required and autosomal insertions that do show full dosage compensation are in regions that supply this requirement.

The sensitivity of dosage compensation to the chromosomal location and activity status of the promoter and/or other regulatory elements of the gene is also evidenced by the effect of tissue-specificity and developmental stage-specificity on dosage compensation. An example is the X-linked *Hsp82* gene of *D. pseudobscura*, autosomal insertions of which are dosage-compensated in adults (Jaffe and Laird 1986) but not in the larval stages (Sass and Meselson 1991). Other such examples involving dosage

compensation of autosomal gene inserted into the X chromosome have been covered in the introduction. In this study, however, in each line the tissue and developmental stage specificity of transcription exhibited by the transgenic 507mr was the known normal pattern of expression of this gene (Krishnan et al. 1991). However, since the endogenous 507 gene is intact and functional there are four doses of the promoter elements of gene 507 in the transformants, which is different from the normal situation for this gene. The level of expression of the transgene with respect to normal levels of expression of gene 507 was not quantitated, however, and therefore any effects of the presence of functional endogenous 507ml gene on the expression of the transgene have not been addressed in this study.

The partial dosage-compensatory response of the autosomal insertions in this study could also result from a lack of representation of all the required cis-acting dosage compensatory sequences within the construct used. This is to be conceptually differentiated from the possible need for an appropriate chromosomal environment for efficient functioning of a dosage-compensatory element(s). Thus, there is possibly a multipartite arrangement of the dosage-compensatory cis-acting sequences so that a lack of complete representation of these sequences does not necessarily have an "all or none" effect but has an effect on efficiency of hyperactivation. An example of such a situation is provided by studies on the cis-acting control elements that confer appropriate tissue specificity, and optimal level of transcription on the autosomal *NinaE* gene of *D. melanogaster*. Removal of an enhancer-like control element (Control element I, Mismar and Rubin 1987) results in only a two-fold reduction of transcriptional activity while removal of the adjacent control element II causes a 30-fold reduction in activity (Mismar and Rubin 1987). Removal of both control elements results in a 30-fold reduction in activity and yet control element I efficiently substitutes for control element II (Mismar and Rubin 1987). Neither element affects the appropriate tissue-specificity of transcription, which in turn is mediated by a third adjacent control element (Mismar and Rubin 1989). Interestingly, for the X-linked *Sgs-4* gene, partially dosage compensated alleles have been reported (Korge 1981 ; Kaiser et al. 1986). Such partially dosage-compensated alleles were found to differ from the wild-type fully dosage-compensated

alleles in their 5' regulatory regions.

While the results from this study do not convincingly resolve the central question of the organization of cis-acting dosage compensatory sequences with respect to individual transcription units, they do strongly suggest that at least some the basis components of such sequences are closely linked to the transcription unit. On-going experiments using constructs with more and less flanking sequences should further resolve some of the questions raised by this study.

However the ideal attempt to resolve the prevailing confusion regarding dosage compensatory sequences would require the demonstration that the removal of specific sequences from a dosage-compensated gene results in the loss of dosage compensation and that this loss is independent of chromosomal location and of the normal transcriptional activity of the gene. The results of the study presented in this chapter suggest that gene *507mr* could be analyzed in this manner and is therefore potentially a good system with which to identify and characterize cis-acting sequences mediating dosage compensation.

The simultaneous nucleotide sequence analysis of this gene in the two species along with the corresponding cDNA from *D. melanogaster* (see following chapter, Ch.4) has resulted in a detailed characterization of the 3.8-kb fragment used in this study. With the help of this information certain aspects of the quantitative analysis performed in this study could be improved. An interesting change would be to mark the transgenic *507mr* in order to be able to distinguish it from the endogenous transcript. This would enable the comparison of steady state levels of transgene derived *507* mRNA in the two sexes relative to the levels of the endogenous *507*mRNA with a single detection and quantitation step. Moreover the comparison of the nucleotide sequence of the non-coding DNA within the 3.8-kb *SalI* fragment of gene *507mr* with that of the corresponding region from *D. melanogaster* could now be of use in identification of potential transcriptional control elements. This sequence comparison, although handicapped by a lack of information regarding the transcriptional start site of the gene in either species, is presented in the next chapter. On-going experiments with transgenic constructs designed to further localize the cis-acting transcriptional control sequences are expected to identify

the basic regulatory elements of this gene. Given that these promoter elements do interact with the dosage compensatory machinery, it is possible that in addition to experiments designed to localize dosage-compensatory cis-acting sequences, trans-acting interactions at the promoter elements could be studied for sex-specific differences.

CHAPTER 4

SEQUENCE COMPARISON OF GENES *507MR* AND *507ML* HIGHLIGHTS DOMAINS POTENTIALLY SIGNIFICANT IN REGULATION AND FUNCTION

INTRODUCTION

The head of the adult fruitfly *Drosophila melanogaster* predominantly consists of tissues of two large compound eyes, and of the brain. From a genomic library of *D. melanogaster*, Levy et al. (1982) isolated twenty clones whose transcripts were highly abundant (3-177 times) in the adult head relative to the adult body. One of these clones, 507, was mapped to locus 66D on the autosome 3L of *D. melanogaster* (Levy et al 1982). The homologue of this clone in *Drosophila miranda* has been the focus of this dissertation given the objective of establishing gene 507 in this species as a model system for the study of molecular mechanisms in dosage compensation. The autosome 3L of *D. melanogaster* has evolved as an X (X1R) chromosome in *D. miranda*. As is typical for the X chromosome, in the latter species the X1R chromosome is present in two copies in the females and a single copy in the males. Despite this dosage difference, the steady state level of 507 mRNA is equal in the two sexes of *D. miranda* indicating that this gene is dosage compensated (Chapter 2 ; Krishnan et al. 1991). In *D. miranda* therefore gene 507 is expected to have acquired cis-acting sequences that mediate dosage compensation through male-specific transcriptional hyperactivation, and a major long term objective of this project is the identification and characterization of such sequences.

The initial characterization of clones 507ml and 507mr has been discussed previously. In both species clone 507 encodes a 1.8 Kb mRNA that is expressed only in the retinal photoreceptor cells of the adult eye (Chapter 2 ; Krishnan et al. 1991). A 3.3-kb EcoRI fragment of clone 507ml and a 3.8-kb Sall fragment of clone 507mr appeared

to have the entire transcribed region of this gene within the two species (Chapter 2 ; Krishnan et al. 1991). When the 3.8 kb SalI fragment of clone *507mr* was introduced into the genome of *D. melanogaster* by *P-element*-mediated transformation, the characteristic 1.8-kb mRNA was transcribed from it in the retina-specific manner appropriate for gene *507*. In addition the steady-state level of transgene derived *507* mRNA was 1.4 to 1.6 fold higher in males relative to females indicating that the transgene was responding to the dosage-compensatory machinery (Chapter 3). Therefore, for gene *507* of *D. miranda*, all the cis-acting sequences necessary for correct retina-specific expression and at least partially functional components of those necessary for dosage compensation, are present within the 3.8-kb SalI fragment containing the gene. In order to facilitate the further localization of transcriptional control elements of gene *507mr* and the characterization of this gene in terms of the organization of coding and non-coding regions, its nucleotide sequence was determined. Earlier cross-hybridization experiments have shown that genes *507ml* and *507mr* are highly conserved within the transcribed regions and further, that the 3.3-kb EcoRI fragment of gene *507ml* is the approximate equivalent of the 3.8kb SalI fragment of gene *507mr* (Chapter 2 ; Krishnan et al. 1991). A cross-species comparison of the flanking sequences of gene *507* in the two species could serve to highlight potential transcriptional control elements as highly conserved regions. The reliability of cross-species comparison of nucleotide sequences in the identification of potential regulatory elements of genes in *Drosophila* has been demonstrated by many investigators (Blackman and Messelson 1986 ; Scholnick et al. 1986 ; Martin et al. 1988 ; Henikof and Eghtedarzadeh 1987 ; Fortini and Rubin 1990). The nucleotide sequence of the 3.3-kb EcoRI fragment containing gene *507ml* in *D. melanogaster* was therefore determined for comparison with that of gene *507mr*. To complement these efforts, a *507* cDNA was isolated from a head cDNA library of *D. melanogaster* and its sequence used to characterize the genomic sequences of gene *507ml* and *507mr*. It was found that the conceptual translation product of gene *507* is homologous to the vertebrate protein, arrestin, which is involved in the visual phototransduction pathway (Kuhn and Wilden 1987). Yamada et al. (1990) independently and simultaneously isolated a cDNA from an expression cDNA library of *D. melanogaster*, using monoclonal antibodies specific for

a 49 Kd retina-specific protein that undergoes light induced phosphorylation. The cDNA was found to be homologous to clone 507 isolated by Levy et al. (1982) and to share a 42 % amino acid sequence identity with vertebrate arrestin. My nucleotide sequence analysis of the 507 cDNA is in complete agreement with the report of Yamada et al. (1990).

The detailed analysis of the nucleotide sequences of genes *507mr*, *507ml* and the 507 cDNA reveals that in both species, gene 507 is interrupted by at least two introns. The exons are remarkably conserved, showing an 85% homology at nucleic acid level and a 98% similarity at amino acid level. None of the amino acid changes occur at domains thought to be similar to the vertebrate homologue (Krishnan and Ganguly 1990 ; Yamada et al 1990; Kuhn and Wilden 1987). The nucleotide sequences upstream of the translation start codon and downstream of the translation stop codon have diverged considerably. In the midst of this divergence however, there are conserved sequences resembling basal promoter elements that could be the promoter elements of gene 507. Also, there are two significant changes in the vicinity of these putative basal promoter elements that could play a role in male-specific transcriptional hyperactivation of gene *507mr*. Many features of interest from the standpoint of transcriptional control arising from this cross-species comparison are handicapped by the absence of information regarding the start site of transcription of gene 507 in either species. The nucleotide sequence of gene *507mr* has been published (Krishnan and Ganguly 1990) and has been deposited in the EMBL nucleotide sequence data bank with accession number X54084.

MATERIALS AND METHODS

Head cDNA library

A *D. melanogaster*, strain Oregon R, head cDNA library made in the *lambda* SWAJ2 vector was a gift from Seymour Benzer. The cDNAs in this library are cloned as EcoRI-XbaI inserts, with EcoRI marking the 5' end and XbaI the 3' end of each insert.

Isolation of 507 cDNA

Six cDNAs complementary to gene 507 were isolated from a *D. melanogaster*

head cDNA library using the plaque lift method of Benton and Davis (1977) (detailed in Appendix). Restriction analysis confirmed that all six clones were identical except for being truncated to different lengths at their 5' termini. The longest cDNA was 1.57 Kb in length and this was subcloned as an EcoRI-XbaI fragment into the plasmid pBSK+ (Stratagene).

Nucleotide sequencing

The 3.8-kb SalI fragment of gene *507mr*, the 3.3-kb fragment of gene *507ml* and the 507 cDNA were subcloned into the plasmid vector pBSKII+ (Stratagene). Overlapping deletions were generated by use of restriction sites as well as by unidirectional Exonuclease III digestion (Hennikoff 1984). Sequencing was by the dideoxy method of Sanger et al. (1977) using the commercially available Sequenase kit (United States Biochemicals). Either ^{32}P or ^{35}S dATP was used as label for the sequencing reactions. In a few cases synthetic oligonucleotide primers were used to sequence regions not made accessible by enzymatic deletions. The Beckmann Microgenie nucleic acids analysis software was used to analyse DNA sequences.

RESULTS

Alignment of nucleotide sequences of gene *507ml* and *507mr*

The nucleotide sequences of the 3.8-kb SalI fragment of *D. miranda* and the 3.3-kb EcoRI fragment of *D. melanogaster* were compared in three different parts. The first part (Box 1 ; Fig. 21) is upstream of the translational start codon and consists of 1637 bp for *D. miranda* and 1549 bp for *D. melanogaster*. The translational start site of gene *507ml* has been convincingly established by Yamada et al (1990) who determined the sequence of the first ten amino acids of the *in-vivo* translational product of the 507 cDNA. The translational start codon of *D. miranda* was inferred from the high degree of conservation of the open reading frame downstream of it.

The second region (Box 2 ; Fig. 21) extends between the translational start codon and the stop codon and includes the two introns. The positions of the introns in gene

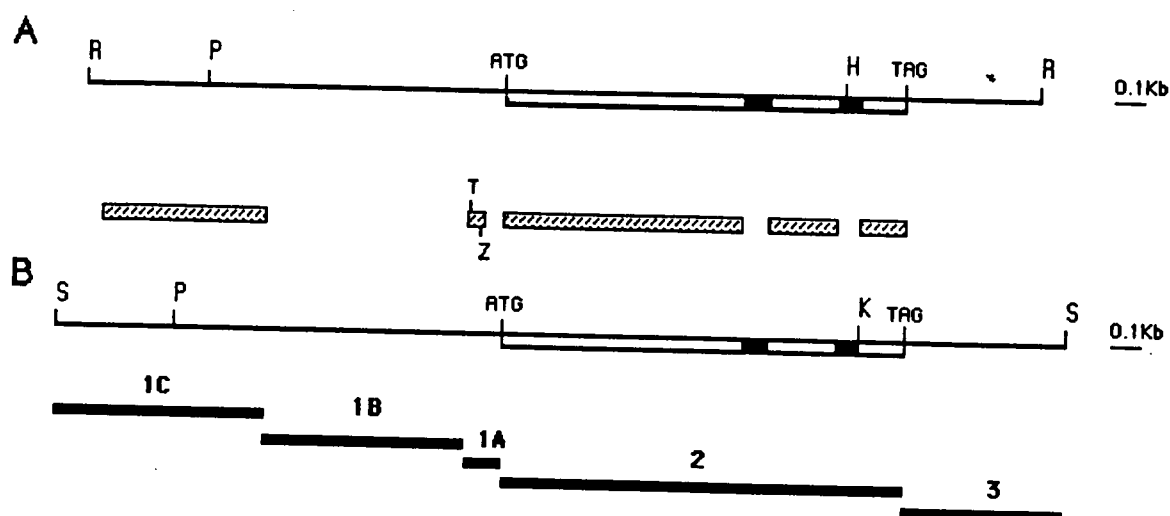


Figure 21. Genes 507ml and 507mr

The 3.3-kb EcoRI fragment containing gene 507ml (A) and the 3.8-kb SalI fragment containing gene 507mr (B) have been aligned at their translation start sites, identified in this study. The open areas of the rectangle associated with each map designates exons, while solid areas designate introns. The hatched boxes between the two genes indicate regions of sequence conservation. Intervening unmarked regions are those that show, less than 50% homology. The hatched box just upstream of the translational start codon is marked with T and Z. T indicates the position of the putative TATA element that is identically conserved in the two species. Z indicates the site at which the genomic sequence of gene 507ml determined in this study and the sequence of the 507 cDNA reported by LeVine et al. (1990) cease to match. Solid boxes indicate the three parts of division of the nucleotide sequence comparison in this study. Box 1 is further divided into three regions, 1A, 1B and 1C.

Restriction sites shown are: R: EcoRI; P: PstI; H: HindIII; S: SalI; K: KpnI.

507ml were determined by alignment of genomic and cDNA nucleotide sequences. The positions of the introns in gene *507mr* are inferred from similarity with gene *507ml*.

The third region (Box 3 ; Fig. 21) is downstream of the translational stop codon extending to 517 bp for gene *507mr* and 401 bp for gene *507ml*. This region includes the polyadenylation signal and site for gene *507ml* and, by inference from similarity, those for gene *507mr*.

Comparison of 5' non-coding sequences of gene *507ml* and *507mr*

The alignment of the sequences upstream of the translation start site (Fig. 22) highlighted three distinct regions in terms of degree of homology (Boxes 1A, 1B, 1C, Fig. 21). The first region (1A, Fig. 21) immediately upstream of the translational start site and extending further upstream to approximately to the nucleotides numbered 1464 and 1367 in genes *507mr* and *507ml* respectively, has the most likely candidate for the TATA box of this gene. The latter is seen within an identically conserved 25 base pair stretch, 5'CAGGGCAAAGGTATATAAGCAGGT, which in turn is part of a highly conserved 70 bp stretch ending at nucleotides 1576 and 1465 for *507mr* and *507ml* respectively (Fig. 22). In eukaryotes, the transcription start site is almost invariably located between 20-30 bp downstream of the TATA box (see Chambon 1984 ; Serfling et al. 1985). I have not been able to demonstrate experimentally that this is the case for either of the two genes.

The second region (1B, Fig. 21) immediately upstream of region 1A is relatively poorly conserved in the two species. This region extends upstream of region 1A, from nucleotide 1464 to 865 in gene *507mr* and from nucleotide 1367 to 844 in gene *507ml* (Fig. 22). Despite the overall poor conservation of region 1B, three interesting features are highlighted within it. The first is a highly conserved 32 bp stretch containing a putative GCAAT motif in the reverse orientation, and a GCAAAT motif resembling a major portion of the octamer binding sequence, both of which are commonly encountered features of eukaryotic promoters (reviewed by Serfling et al. 1985 ; Dynan 1987 ; Maniatis et al. 1987 ; Lewin 1990a ; Lewin 1990b). The GCAAT motif is known to be

Figure 22. Comparison of 5' non-coding sequences of genes 507mr and 507ml.

The sequence of gene 507mr (mr-) and gene 507ml (ml-) have been aligned for maximum homology using the Beckman Microgenie nucleic acids analysis software. The numbering begins at the SalI site and the EcoRI site for mr- and ml- respectively. This region extends 1638 and 1549 nucleotides upto the translational start codons for mr and ml respectively. Region 1A (see text): 1638 to 1464 in mr and 1549 to 1367 in ml. The putative TATA element beginning at nucleotide 1514 for mr and 1405 for ml is shown in bold and underlined. Region 1B (see text): 1464 to 865 in mr and 1367 to 844 in ml. The GCAAT, GCAAAT motifs are shown in bold and underlined at nucleotides 1456 for mr and 1340 for ml. The tandem heptad repeat found only in mr is shown bold and underlined at nucleotide 1277. Region 1C (see text): 865 upstream to 1 in mr and 844 to 1 in ml. Shown in lowercase and underlined is the sequence upstream of ATG upto the 5' terminus of the 507ml cDNA isolated in this study.

B. comparison of the 5' portion of the 507 cDNA of LeVine et al. (1990), marked Zu with genomic sequences of corresponding regions of 507mr and 507ml. The nucleotide, G, at which the cDNA sequence differs from the mr and ml genomic sequences is shown in bold and underlined.

A.

```

mr-  GTCGACGAGTGC GGGGAACAGGCCAAACGTCGGACGGCAAACCTCTCCCGAGTACAGCTCA 60
      :: : :      :: :      :: :      :: : : : :
ml-  GAATTC      GCCGA      GGA      TACAGCTCA 23

mr-  AAGTCGTGCGTCCGCATCTTGCGCAGATCCCAGCCTCGGATGACCGGATCCGATCCCCCA 120
      :: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ml-  AAGACGTGCGTCCTCATCTTGCGCAGATCCCAGCCGCGGATCAGACCGTCCGAACCTCCA 83

mr-  GTCACCAGGATATTCCGGTCAAAGTGGCTCCAGTCGCAG      G      CGGTGGCATGC 171
      :: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ml-  GTGACCAGGACATTGCGATCGAAGTGGACCAAGTCGCAGCATAGAGCCTCACTCGCATGC 143

mr-  TGCCTCGATGCTCATCAGGGGTTTACCTGCAAAATCAAGAGAATCCATAAGTTCAACTG 231
      :: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ml-  GCCTCGATGCTCATCAGGGGTTTACCTGCAAAAGTCAAGAGAGTTCCATAAGTTAAGATG 202

mr-  CCCGTCGGTACTCAGCCTTGCAAAGAGATTGCGGATCAGCGGGGAAAACCTTTGCGCAATA 291
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ml-  TCATCGGTACTTACG TTCGAAACAGATTGGCGATCA GGGAGAGAC TTGCCCCGTA 257

mr-  GATCAAATCACTGTGCCCCACAAAGGT GTGATGGAGTGCTGTCTGTGTCAGTCCCACAG 350
      :: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ml-  GATCAGATCGTTGTGCCCCACAAATGTGGTGATGGAGTTCTGTCTGTTGCAGTCCCACAA 317

mr-  CTTCAGAGTGCGGTCCAGCTGGCGGAGAGGAGCGTATGGTAGTTCCACTGCTCG CCCC 409
      :: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ml-  TTTCAGAGTCCAATCCCAGCTGCCCCGAGAGCAG GTGTGGTAGTTCCACTTCTTGACCCC 376

mr-  AGTCCAGGCTGTAGATCTCGTTCTTGCTCCTGCAGGCAGATCAGCGGCTGTTTGGGCG 469
      :: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ml-  AGTCCAGGCTGTACACCTCGTTCTTGCTCCTGCAGGCAGATCAGCGGCTGTTT GCGG 435

mr-  TCTCCTG CCCTCGGCGCCTCGGCGTCCAAACCGGACCAGATCTGAAGCAGACCATCGCC 528
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ml-  TAGCTGGTTGCTCGCCG ATCGCCATCCAATCCGCACCAGATCTGAAGCGATCCATCACC 494

mr-  AGAGGCTGTGGCTGCCACATCGGCGGCAGAGGGGCACCAGGCGACGTCG AC AGCCATC 586
      :: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ml-  GGA GCAGTGGCGGCGATATCGGCGGCATACGGACACCAGGCGACGTGGAACAATCCATC 553

mr-  CGACCACTCCAGACGGCACAGTTCACTCAG T TGGC TCTGCCCCGT CTGAACTCG 639
      :: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ml-  GGACCACTCTAGGCGGCACAGTTCTCCGAGACTCTGGCCATCTGTGGATGAGGAATTTGT 613

mr-  AGAAAAC TG CAG AG AAGAG ACCACCACCGGC CACCGTACAATTG GC 685
      :: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ml-  GTTGGAGTTCTGTTCCAGCAGAAAGAGGGAACCACTCCAGCAAGACCGTACAGTTGCTG 673

mr-  TGCAGGCAGGAACAGGGTTAATGGTTAATGGTGGGCTGGATTTTGGGTGTGGACGTGTGA 745
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ml-  GGTATCGAAAAAATGTTTCTATTAGTATATTGTTGGTTTTAGAGATT AAGATTGA 731

```



```

mr-  GGTTCGTACGGCACTGGGGTTTCGCATTTTCTCCGGGCAAACGGAAACTCAGGAGCTCC 1584
      ::: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ml-  GGTTCGTACGGCACTTGGGTTTCGCATTTTCTCCGGCCGAACGGAAACTCTATAACGCG 1475

mr-  CAGTCGTGGTCTTAG  GAAAAGGAAACAAGCAAAGAAAAGAAAACCGCTAGT..... 1637
      :: : : : : : : : : : : : : : : : : : : : : : :
ml-  CATCGCCCGACTTAGCCTCGAATCCAGTTGCATTGCGCAGGAATCAAAGAAATCCACTAA 1535

mr-  .....
ml-  TCCACTAATCAACA

```

1549

B.

```

mr-  GGTTCGTACGGCACTGGGGTTTCGCATTTTCCGGCCGGGCAAACGGAAACTCAGGAGCTCC 1584
      ::: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ml-  GGTTCGTACGGCACTTGGGTTTCGCATTTTCTCCGGCCGAACGGAAACTCTATAACGCG 1475
      : : : : : : : : : : : : : : : : : : : : : : :
zu-  TTCCTTATCTTGGGTCTTCTCCTTGGGCGCCTTCTTGTCGCAGGAAACTCTATAACGCG

```

able to function in either orientation (Lewin 1990b) and both these motifs are identically conserved in the two species. However these motifs are not aligned together by this analysis and this may be due to the overall dissimilarity including deletions and additions of the sequences around this 32 bp stretch. This stretch is 54 and 59 bp upstream of the putative TATA element in *507mr* and *507ml* respectively (Fig. 22), which also is an organizational characteristic often displayed by these elements (reviewed by Serfling et al. 1985 ; Dynan 1987 ; Lewin 1990).

On inspection of the sequences just upstream of this putative CAAT motif in the two species, the second interesting feature, a possible rearrangement of sequence is noticed. In *507mr*, a potential stem and loop structure of the sequence, 5'TGTTTGCAGTTTTTCTTTCTGCATAACG 3', is associated with the putative CAAT box (Fig. 22, 23). This potential stem and loop motif is not only absent from *D. melanogaster* but appears to be disrupted by a rearrangement so that the sequence 5' TGTTTGCAGTTTT 3' is found intact, but far upstream of the putative CAAT box, 5' TGCATAACG 3' (Fig. 22, 23). This sequence rearrangement appears significant to gene *507mr* in that it may have resulted in the formation of a motif that could potentially form a stem and loop structure that is often implicated in protein binding. The AT richness of the region between the putative stems could promote transient single strandedness or "breathing". The latter property is thought to be a desirable characteristic of promoter regions and has been implicated in the increase of transcription rates (Bode et al. 1992). It is tempting to speculate further that this stem and loop structure has been selected for in gene *507mr* by the requirement for male-specific transcriptional hyperactivation. A sex-specific trans-acting factor(s) may cause the differential function of this stem and loop sequence in the two sexes.

The third feature of interest in region 1B the one seen upstream of the putative CAAT motif just discussed. Starting at nucleotide 1278 in *507mr*, five tandem repeats of a heptad (T)(G)(G)(G)(C)(G/A/T)(A/G) are seen and these are absent from *507ml* (Fig. 22). Such repeat motifs are often found in enhancer-like sequences (reviewed by Serfling et al. 1985 ; Dynan 1987) and perhaps this sequence may have a role in male-specific transcriptional hyperactivation.

507ML

TGTTTGCA**GT****TTTCCCTTAGTTCCTATGTTCTATCTTATTGTCGTGTTGGACAAGCCAAACAAGTGCA****TAA****CG**

507MR

TGTTTGCA**GT****TTTCTTTTCTGCA****TAA****CG**

Figure 23. Potential stem and loop forming sequence in the putative promoter region of gene 507mr.

The sequences shown are from the junctions of regions 1A and 1B (Fig. 22). The nucleotides conserved in the two species are shown in bold. The five nucleotides that could be involved in stem formation are underlined. In the case of 507ml, the stem towards the 3' end has a G instead of the C seen in 507mr.

The next region (Box 1C, Fig. 21) is at the most 5' ends of the two genomic fragments, extending approximately between nucleotides 1 and 865 in gene *507mr* and between nucleotides 1 and 844 of gene *507ml* (Fig. 22). This region shows a 72% homology which contrasts greatly with the very low homology between the sequences of region 1B immediately downstream of it (Fig.22). This high degree of homology is of unknown significance. This region, especially if representing non-transcribed DNA, could have sequences necessary for transcriptional control of gene *507* in terms of tissue- and developmental-stage specificity. I have previously mentioned (Chapter 2) that there exists a possibility that this region represents a transcribed portion of gene *507* or an mRNA similar in size to *507* (Chapter 2). In case of the former possibility, an untranslated leader exon may be present there, although it is not readily apparent why an untranslated portion of the gene would show such a high degree of conservation. This unconfirmed possibility of gene *507* having a transcribed leader that is separated from the first translated exon by an intron, is consistent with a recent report of the nucleotide sequence of a *D. melanogaster* *507* cDNA longer, at the 5' end, than the one used in this study (LeVine III, et al. 1990). At a certain site (Z, Fig. 21 ; Fig. 22), downstream of the putative TATA box, the cDNA sequence of LeVine III et al. (1990) ceases to match with my report of the genomic sequence of *507ml*. The site of divergence is within a stretch of twelve nucleotides which is identically conserved in the *D. miranda* (Fig. 22). This divergence of cDNA from genomic sequence would imply an exon-intron junction and the cDNA sequence upstream of this junction would be part of the untranslated exon. However, this putative leader exon is not present within the 3.3-kb *EcoRI* fragment of gene *507ml* and neither is a sequence homologous to it to be found in the 3.8-kb fragment of gene *507mr*. It needs to be recalled here that for gene *507mr* the 3.8-kb *SalI* fragment has all the requirements to be transcribed into a 1.8-kb mRNA, the correct size of *507* mRNA. Therefore, it is reasonable to assume that in the case of gene *507mr* we have the sequence of the entire transcribed portion of this gene. Also, from the putative exon-intron junction and the genomic sequence (see legend Fig. 22) it can be inferred that the terminus of the putative intron is not the consensus 5'-AG-3' known to be a stringent requirement at 3' acceptor splice sites (Shapiro and Senapathy 1987). Also the terminus of the putative

untranslated exon, the 5' donor splice site, does not conform to the consensus 5' donor splice site sequence where again a universally conserved and stringent requirement for cleavage following 5'-GT-3', is known (Shapiro and Senapathy 1990). These requirements are complied with at the other two introns within gene 507 in both species (Fig. 24). Non-conformist splice junctions do exist, however, and the splice junction of this putative leader and the first translated exon matches a splice junction seen in Human IG heavy chain gene (see Shapiro and Senapathy 1987).

On the other hand, it would be difficult to reconcile this possibility of an undetected exon with the observation of the highly conserved promoter elements, and potential transcriptional hyperactivation elements discussed in the context of regions 1A and 1B, which are downstream of region 1C (Fig. 22). While enhancer-like transcriptional control elements could be found downstream of the transcriptional initiation site (Gillies 1983 ; Banerji et al. 1983), the positioning of the TATA box upstream of the start site (see Chambon et al. 1984) is known, in all cases tested, to be a stringent requirement of RNA polymerase II-transcribed eukaryotic genes with a functional TATA box. At this time, however, the exceptionally high degree of conservation of the putative basal promoter elements of this gene is clearly one of the striking and unambiguous feature of the 5' upstream sequences of gene 507*mr* and 507*ml*. The further interpretation of the various suggestive features emerging from this comparison will have to await the identification of the transcriptional start site of gene 507.

Comparison of coding regions and introns

The nucleotide sequences of gene 507*mr* and 507*ml* that extend between the translational start codon and the stop codon are shown in Fig. 24. The conceptual translation of gene 507 yields a 401 amino acid protein with a predicted molecular weight of 49-Kd. This is in complete agreement with the reports of Yamada et al. (1990) and LeVine et. al. 1990). Within coding sequences there are 144 nucleotide substitutions of which 137 are silent, resulting in synonymous codons. Of the seven amino acid substitutions (Fig. 24) five are conservative and none of the seven changes affects domains thought to be of functional importance through comparison with the vertebrate

Figure 24. Comparison of coding regions and introns

The sequence is arranged in four rows. At the top is the amino acid sequence of the 507mr protein, below which is the nucleotide sequence of gene 507mr (mr-). The nucleotides of gene 507ml (ml-) are entered only when different from those of mr. Similarly amino acids of 507ml are entered in the fourth row only where they differ from mr. Numbers refer to amino acids. Introns are shown in the lower case. The putative branch site, see text, are shown in bold and underlined.

170 180
 M V S L V I K K L Q Y A P L N R G Q R L
 mr- ATGGTCAGCCTGGTGATCAAGAAGCTGCAGTATGCCCCCTGAACCGGGGCCAGCGGCTG
 * * * * *
 ml- T A C T C T T T

190 200
 P S S L V S K G F T F S N G K I S L E V
 mr- CCCAGCTCGCTGGTGAGCAAGGGATTACCTTCTCGAACGGCAAGATCAGTCTGGAGGTC
 * *
 ml- C T

210 220
 T L D R E I Y Y H G E K T A A T V Q V S
 mr- ACGCTCGACAGGGAGATCTACTATCAGGCGAGAAGACGGCCGCCACGGTGCAGGTGTCC
 * * * * *
 ml- T G T T C C G

230 240
 N N S K K S V K S I K C F I V Q H T E I
 mr- AACAACTCGAAGAAATCGGTGAAGAGCATCAAGTGCTTCATTGTGCAGCACACGGAGATC
 * * * * *
 ml- C G C T C

250 260
 T M V N A Q F S K H V A Q L E T K E G C
 mr- ACGATGGTGAATGCGCAGTTCAGCAAGCATGTGGCCAGCTGGAGACCAAGAGGGCTGC
 * * * * *
 ml- C C C T G

270 280
 P I T P G A N L T K T F Y L I P L A A N
 mr- CCCATCACCCCGGGCGCCAATCTGACCAAGACCTTCTACCTTATCCCCTGGCAGCCAAC
 * * * * *
 ml- T C T G G G T

285
 N K D R H
 mr- AACAAAGGATCGTCATGgtaacgaaagagcctacagcgggcggtggcagtg gagtctcta
 * : : : : : : : : : : : : : : : :
 ml- C gta ggatttgatgtgatagttgtgattctggatagtttcta

286 290
 G I A L D G H L K D E D
 mr- at ctctgtgtgggttttctttaagGCATTGCCCTGGATGGTCATCTCAAGGACGAGGAT
 : : : : : : : : : : : : : : : :
 ml- atggattatttaattgtgaactcag T C CT G

300 310
 V N L A S S T M V Q E G K N T G D A C G
 mr- GTCAATCTGGCCTCCTCCACCATGGTCCAGGAGGGCAAGAACACTGGCGATGCCTGCGGT
 * * * * *
 ml- C T A G TC G A T
 S

320 330
 I V I S Y S V R I K L N C G T L G G E M
 mr- ATTGTCATCTCCTATTTCGGTTCGGATCAAGTTGAACTGCGGCACCCTTGCGGGCGAGATG
 * * * * *
 ml- A C G C T C C T G A T A

340 350
 Q T D V P F K L L Q P A P
 mr- CAAACGGATGTGCCCTTCAAGCTGCTGCAGCCAGCTCCTGgtaaggct ccaggacaccc
 * * * * *
 ml- G C A G A G gtaagcttgcaagattact

351
 G T I E K K
 mr- cccgacagaacacccg ctctaattggttctctatctcatagGTACCATTGAAAAGAAAC
 : : : : : : : : : : : : : : : :
 ml- tgaacttgacttctagtttgtaattcctttgatgtggtatag A G G

360 370
 R S N A M K K M K S I E Q H R N V K G Y
 mr- GCTCGAATGCCATGAAGAAGATGAAATCGATTGAGCAGCATCGCAATGTCAAGGGCTACT
 * * * * *
 ml- A G C C C G T

380 390
 Y Q D D D D N I V F E D F A K M R M N N
 mr- ACCAGGACGATGACGATAATATTGTCTTTGAGGACTTTGCCAAGATGCCGAATGAACAATG
 * * * * *
 ml- T A C C C

401
 V N M A D
 mr- TGAATATGGCCGATTAG
 * * *
 ml- C G C

homologue (Krishnan and Ganguly 1990 ; Yamada et al. 1990). Thus there appears to be a strong selection against changes in the amino acid sequence of gene 507.

As expected, in contrast with the high degree of conservation of exonic sequences, the sequences within the two introns have diverged considerably. The only conserved motif seen is the CTAAT sequence, thought to be the consensus branch site is well conserved in both the introns (Keller and Noon 1985). The GT-AT rule of 5'-donor 3'-acceptor splice sites is conserved at both the intron-exon junctions in both the species (Fig. 24).

Previously, it was demonstrated that the KpnI-SalI fragment containing the 3' of gene 507mr (Fig. 21) does not cross hybridize with the corresponding HindIII-EcoRI fragment of gene 507ml (Fig. 21). The third exon of gene 507 (Fig. 24) is included within this fragment but is highly conserved, and it is the sequences downstream of the stop codon that have diverged considerably. The complete lack of cross hybridization of these two fragments may therefore have to do with peculiarities of the radiolabelling process considering that the conserved regions are small terminal portions of these fragments.

Comparison of 3' non-coding regions

Fig. 25 shows a comparison of the sequences downstream of the translational stop codon in the two species. For *D. melanogaster* the site of poly-adenylation (Fig. 25) was identified by aligning the genomic and cDNA sequences. Twenty nucleotides upstream of this site is a putative poly-adenylation signal, UAUAAA, which differs from the consensus, AAUAAA, in the first position (Fig.24). Such a sequence has however been shown to direct poly-adenylation of the mRNA encoding the Hepatitis B virus surface antigen (Simonsen and Levinson 1983) and is also the most likely polyadenylation signal for the mammalian *CREM* gene (Foulkes et al. 1991). This motif is identically conserved in *D. miranda* where too, it may function as the poly-adenylation signal. This comparison of 3' untranslated regions also highlights patches of conservation within an overall sequence homology of only 57.5% (Fig. 24) the functional significance, if any, of which are not apparent at present.

Figure 25. Comparison of regions downstream of the translational stop codon.

Sequences of gene 507mr (mr-) and gene 507ml (ml-) are aligned as in Fig. 22. The site of polyadenylation of gene 507ml is shown in bold and underlined at nucleotide 3114 of ml. The polyadenylation signal for ml is seen 25 bases upstream at 3089 (bold and underlined) and an identical sequence is seen at 3187 for mr.

```

mr- AGAAACCTCTTCTGATATATATCTACGTTTGTCCACGTTTTTGTGTAAAACTGCGTCT 3031
    :  ::  :  :  :  ::  :  ::  ::  ::  ::
ml- GCCACCGATCTTCAATCTCTTCTAGCTCTGTTGTGAACCGAATTG  AACTAAATC 2938

mr- CGAAGCACTGAAAACTTGTTCAGCAGCAGAGGAGGGTCTCTCTCTAAGGGCCACAGAAT 3091
    :  ::  :  :  :  ::  :  ::  ::  ::  ::
ml- GAAAGCACTGAAAATTTGTTCAGCAG  TCTCTCGTGGTGGCGCACAA 2985

mr- CTCTTTGTATTATAATCTCGTATACTCTCATATAGAAAACTTTGGTGGATCTAGAGTTA 3151
    :  ::  :  :  :  ::  :  ::  ::  ::  ::
ml- TCTTTGTATTATAA  TCGTAT  CTATAATGT  AACTTTGGTGGATCTGGAGT 3034

mr- TATCCTATCCGGTGGGACCGCGAGTCCCCACATATTATAAATGTGTGTATTGGCTATC 3211
    :  ::  :  :  :  ::  :  ::  :  :  :  :  :  :
ml- TTGTA  CCCGTACATATGCATGT  TGATGAGTTGAACAGACCTCGTGTGCTAT 3086

mr- TAGACTAGAAGCACTTGTGTATCTGTGTATCTTTGTATCCAACTGTGCGGTAATCCTC 3271
    :  ::  :  :  :  ::  :  ::  :  :  :  :  :  :
ml- ATTATAA  A  TG  TGT  TATGGCTATCAAAGCACTCACCTTTTCTCTCGTTTGC 3137

mr- AGTGGGTTATGGTTTGGT  TTTGGTGTGGTTTTGGTTTTGGTTTGGGTTTGGGAGCTTC 3330
    :  ::  :  :  :  :  :  :  :  :  :  :  :  :  :  :
ml- GATGTAGTATCCTTTTGTGATTGCGTAAGGATTCGGTTTTTGGTAGTGGAT  TGTCTT 3195

mr- GGAGATGCTTAGGCTTTCAACTTCAGAAGTCGAATGATGGACAGAGCTGCGCGGCAGAAC 3390
    :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
ml- GCGGGTGTGTGGGTTTCAACTTCAGAAGTCGAATGATCTAAACAACCTGCGCAGCTCGAC 3255

mr- CAGGAAGGCTCT  CT  CTGGCCTTGTGCCTGCCCTTGGGGTGGCTCTAATGGGGTGGCAA 3448
    :  :  :  :  :  :  :  :  :  :  :  :  :
ml- GACGAAGGATCTGCTGCTGGCAGGAATTC..... 3284

mr- AAGAAAGGCTCGCATAGGTCCACCTTTCAACCCCGTCGAC 3488
ml- .....

```

DISCUSSION

In order to facilitate the mapping of the transcriptional control elements of gene 507, the nucleotide sequences of gene 507*mr* of *D. miranda* and gene 507*ml* of *D. melanogaster* were determined and aligned for comparison. The organization of coding and non-coding DNA of the gene was determined by alignment of genomic nucleotide sequences with that of the 507 cDNA isolated from *D. melanogaster*. The sequences were compared in three different parts: the region upstream of the translation start site (box 1, Fig. 21), those between the translational start site and the translational stop site including the introns (box 2, Fig. 21) and the sequences downstream of the translational start site which include the polyadenylation site (box 3, Fig. 21). In the region upstream of the translational start site, (box 1, Fig. 22), within an overall similarity of 60.5 % there exist pockets of highly conserved regions among which are the putative TATA and GCAAT motifs are identically conserved. In addition there are two interesting sequence motifs with features that may be correlated with male-specific transcriptional hyperactivity, a requirement for the dosage compensation of gene 507 in *D. miranda*. One of these motifs could potentially form a stem and loop structure facilitating transient single strandedness which is a desired property for transcriptional activation. The other is a tandem repeat motif of the kind often associated with enhancers (Serfling et al. 1985 ; Dynan, 1987) although it is not known if its core sequence is similar to any known enhancer sequence. Both these motifs are observed only in *D. miranda* and in the case of the stem and loop structure a rearrangement of existing sequences, which are present in *D. melanogaster*, appears to have given rise to the potential for loop formation. The absence of information regarding the transcriptional start site does handicap this analysis and the possibility of it being located upstream of the putative promoter elements identified here, considerably weakens the arguments forwarded for the putative dosage compensatory motifs of gene 507*mr*. The speculation regarding the latter could be excused in light of the accompanying evidence that sequences within this 3.8-kb SalI fragment of gene 507*mr* do cause, male specific hyperactivation of transcription (Chapter 3). Also, the results of analysis of expression of the transgenic 507 support the contention that sequences that mediate

photoreceptor cell-specific transcription of gene *507mr* must be present within the 3.8 kb SalI fragment (Chapter 3). The regulatory sequences of a few *Drosophila* the photoreceptor cell specific genes *Rh3* and *Rh4* (Fortini and Rubin 1984) and *Nina E* (Mismer and Rubin 1987) have been identified. No significant match exists between these sequences and those present in the 3.8-kb SalI fragment of gene *507mr*. The identification of the transcription start site of gene 507 in the two species would aid the characterization of the regulatory sequences of these genes and experiments towards those ends are in progress. It is likely that with the identification of the transcriptional start site the very sequences presented in this chapter could be more illuminating than they are at present. This information could then be used in specific mutagenesis of suspected transcriptional control elements, in order to characterize the cis-acting transcriptional control region of gene 507 and especially, the dosage compensated gene *507mr*.

The conceptual translation product of gene 507 is a 401 amino acid protein that shares a 42% amino acid identity with the vertebrate visual phototransduction protein arrestin. The comparison of the amino acid sequences of the protein in the two species shows only two non-conservative amino acid exchanges, neither of which lies in a domain thought to be of functional significance as judged by comparison with the vertebrate homologue (Yamada et al. 1990 ; LeVine III, et al. 1990). Gene 507 thus shows a strong selection against changes in the amino acid sequence. Interestingly, there is another *Drosophila* gene having a greater than 42% amino acid-similarity with vertebrate arrestin has been identified (Hyde et al. 1990 ; Smith et al. 1990). Gene 507, now named *ArrB*, however shares only 47% homology with this second *Drosophila* arrestin gene, *ArrA*. As is suspected for the different kinds of opsin genes, a family of arrestin-like protein genes may exists, whose members are differentially expressed within the different photoreceptor cells of the retina. The characterization of the transcriptional control elements of this gene, significant as it is to the problem of dosage compensation, may also add considerably to the characterization of visual system genes of *Drosophila*.

CHAPTER 5

GENERAL DISCUSSION

The biological problem addressed in this dissertation is that of male-specific transcriptional hyperactivation of X-linked genes in *Drosophila*. This is thought to be mediated by trans-acting factors that interact with cis-acting sequences in a male-limited manner. While at least four such trans-acting factors have been genetically identified (reviewed by Luchessi and Manning 1987), no discrete cis-acting sequence mediating dosage compensation has been identified so far. An essential question has been whether each X-linked gene has its own dosage-compensatory sequences, or whether transcriptional hyperactivation of the X-linked genes in males is effected at the larger level of higher order chromosomal conformation, and thus is not a gene-specific mechanism. Among the participants in the dosage compensatory mechanism the only molecular entity that has been characterized to significant detail is the gene product of the *mle* (maleless) gene). This gene has been implicated by genetic and physiological evidence (Belote and Lucchesi 1986) as a trans-acting factor that appears to be essential for male-specific transcriptional hyperactivity. It is interesting therefore that its gene product was found predominantly associated with the male X chromosome at multiple discrete sites along its length (Kuroda et al 1991). However this gene product is also found in females, though it does not exhibit a preference for the X chromosome in them. The conceptual translational product of this gene bears similarity to a class of helicases and is especially similar to RNA helicases (Kuroda et al. 1991). This information in itself can be accommodated both at the level of individual transcription units or at the level of higher order organization of chromatin.

With respect to the cis-acting sequences mediating dosage compensation, a majority of studies including the one described in this dissertation suggest that individual genes appear to have closely associated dosage-compensatory cis-acting sequences. These studies have relied on the ability to assay the activity of an X-linked gene when inserted

at autosomal sites accompanied by defined lengths of flanking sequences. A hyperactivity of transcription of such autosomal insertions in males relative to females is ascribed to the presence of dosage compensatory cis-acting sequences. Though in most cases these studies have been handicapped by a failure to observe the full two fold-hyperactivation characteristic of normal dosage compensation, the general observation that dosage compensatory sequences could be tightly linked to gene they control decries the role of higher order chromosome conformation in mediating male specific transcriptional hyperactivity. Yet these very sequences that are being implicated in dosage compensation may be acting through their effect on chromosomal structure rather than on RNA synthesis itself.

The problem of transcriptional hyperactivation is discussed here in the context of what is understood of transcriptional control in general and of the correlation of such control with higher order packaging of chromatin. The prevailing view is that effects of higher order chromosomal conformations must be mediated through discrete cis-acting elements that bind trans-acting factors which either directly affect promoter interactions or which participate in the adoption of the chromosomal conformation that promotes hyperactivity. The two kinds of interactions may co-exist and in such a case from a mechanistic point of view, it would be necessary to differentiate between them. Despite these complications, or perhaps because of them, the most efficient approach towards characterization of dosage compensatory interactions would still have to be at the level of individual dosage compensated transcription units such as the one used in this study.

Transcriptional control in eukaryotes involves a complex interplay between cis-acting DNA sequences and trans-acting protein factors that interact with them. Both the cis-acting sequences involved as well as the trans-acting factors involved have been traditionally grouped into two classes. One class of cis-acting sequences and the protein factors that interact with them are global, are essential participants in the initiation of transcription, and are conserved throughout evolution (reviewed by Guarente et al. 1988). The cis-acting elements in this class serve as guide points for the step-by-step assembly of the trans-acting factors that constitute the multicomponent transcriptional initiation complex (Hawley and Roeder 1985 ; Buratowski et al.1989). This complex requirement

at initiation is therefore an obvious rate-limiting step for RNA synthesis and in this context the second class of regulatory factors becomes relevant. This second class is represented by cis-acting sequences that attract trans-acting factors which help or interfere with the assembly of the initiation complex (see Guarente 1987; Ptashne 1988 ; Ptashne and Gann 1990). The availability or unavailability of these latter factors therefore significantly affects transcription initiation. Due to technical considerations, the identification of such trans-acting factors has invariably relied upon the identification of the cis-acting sequences with which they interact. For a great number of eukaryotic genes there is now a correlation between its activity status and the discrete cis-acting sequences and trans-acting factors involved. However, this apparent simplicity of the control mechanism in such cases results from the isolation of events at the initiation of transcription of a gene, from the context of the higher order organization of the nucleus. It has been suspected since 1885 (Rabl 1885; see review by Gasser and Laemmli 1987) and now considerably established that the higher order packaging of chromosomes greatly influences transcriptional activity (see review by Gasser and Laemmli 1987 ; Phi-Van and Stratling 1990). However, the only convincing and direct correlation between control of transcriptional activity and DNA packaging has been the inhibitory effect of histone proteins on gene transcription (Felsenfeld 1992 for review).

On the other hand, the correlation between a state of decondensation and increased transcriptional activity (Weintraub and Groudine 1975) has often served as the experimental approach to determine the involvement of higher order packaging in transcriptional control. However, these studies do not distinguish between a cause of a transcriptional event and an effect of it. Therefore a general attempt to study sex-specific differences in X-chromosomal packaging, as detected by probes of chromatin conformation that measure accessibility to DNA, may not be of use to the concept of transcriptional hyperactivation in the context of dosage compensation. On the other hand, if conformational changes are localized to specific sites at the level of specific transcription units, they could then be altered to test for an actual participating role in hyperactivity. Through the study described in this dissertation, gene 507 has been characterized to the level necessary for attempting such experimentation, although it is

possible that the kind of interaction referred to is at a great distance from the basic transcription unit.

Most other correlations between transcriptional activity and chromosomal order have resulted from the observation that a gene when transposed to a new chromosomal location often fails to display its characteristic level of activity displaying instead, a "position effect" (Brinster et al. 1981; Spradling and Rubin 1983). The most testable and therefore appealing hypothesis to explain a position effect has been the suggestion of the existence of boundary elements on the chromosome. Such elements are defined by their ability to confer on a gene the ability to display its characteristic expression independent of chromosomal position (Kullum and Schedl 1991 ; reviewed by Bonifer et al. 1991). This arrangement is perhaps indispensable given the long-range effects shown by many transcriptional control elements, since they serve to delimit a domain beyond the confines of which these control elements cannot act (Stief et al. 1989; Bonifer et al. 1990). In itself therefore the domain arrangement appears to merely ensure appropriate activity that is regulated in the classical manner of promoter-regulator interactions discussed above. At this level of definition, a boundary element should perhaps not be considered a transcriptional control element in the context of interactions at the promoter of a gene. However, if a difference between activity of a gene in two situations is an altering of the boundary element, so that the confines of the domain housing this gene are altered, the boundary element may appear to be a transcriptional control element in currently used assays for mapping cis-acting sequences. The domain organization as understood today has little to offer by way of the mechanism by which nearly every gene on the male X chromosome is transcriptionally hyperactivated to an apparently precise two-fold that of the female. On the other hand, a few X-linked genes "escape" transcriptional hyperactivation in males and in these situations the concept of domain organization can be applied to explain the shielding of such genes from the effects of neighboring dosage compensatory elements. It is of significance instead, to the ability of a few X-linked genes to escape transcriptional hyperactivation. In the same vein, the phenomenon of position effect on relocated genes cannot therefore be taken to imply a direct correlation between transcriptional activity and the importance of a particular "chromosomal configuration,"

at least in the context of the required interactions at a promoter.

On another front, in what is now an exciting frontier in transcriptional control and DNA replication mechanisms, a role has been suggested for higher order chromosomal conformations in terms of reducing the search volume of the nucleus so that trans-acting factors could find their targets more efficiently (Gasser and Laemmli 1987). The working model for this hypothesis is that cis-acting sequences which are to bind trans-acting factors are preferentially located at a proteinaceous nuclear scaffold at which the trans-acting factors in turn would be sequestered (Gasser and Laemmli 1987). Though these models are heavily reliant on an ill defined component referred to as the nuclear matrix (Gasser and Laemmli 1987 ; Phi-Van and Stratling 1990), the proximity of matrix attachment regions (MAR) with cis-acting transcriptional control sequences of a variety of genes (reviewed by Phi-Van and Stratling 1990) makes this largely speculative concept attractive. Here again, in the context of interactions at a promoter the effect of chromosomal organization could be mediated through the classical set of interactions between trans-acting activator proteins and the cis-acting sequences they recognize. In relocation studies a matrix attachment region may be mapped as a transcriptional control element but in the context of interactions occurring at the promoter it may or may not be one, in the strict sense of the term. However a correlation between higher order conformation and gene activity would be reasonable if matrix attachment is a variable between the two situations in which gene activity is being compared. If this is the only variable, then the essential trans-acting factors involved in this transcriptional control mechanism will be those that facilitate this differential matrix attachment. Then a case could be made for the sole and direct effect of "chromosomal configuration" on gene activity. Of relevance to this discussion is once again that gene 507, at its present level of characterization, is amenable to experiments designed to analyze the role of matrix attachment in dosage compensation. For a variety of genes MARs have mapped close to the cis-acting transcriptional control elements (reviewed by Phi-Van and Stratling 1988) and one could expect this to be the case for gene 507.

Thus in experimental terms, a focus on an individual dosage-compensated gene could still accommodate the characterization of a mechanism that may be exerted at the

larger level of chromosomal order. The observation, in the study described in this dissertation that gene 507 is partially dosage-compensated with the help of sequences residing within a 3.8-kb fragment is therefore of considerable significance in the long term to the elucidation of the mechanism of male-specific transcriptional hyperactivation in *Drosophila*.

It has also been stressed that the exceptional karyotype of *D. miranda*, offers an additional perspective of the mode of evolution of cis-acting transcriptional control elements since the X1R chromosome has an autosomal homolog in *D. melanogaster*. Yet another and perhaps more attractive system to approach the problem of evolution of cis-acting transcriptional control elements is provided by *D. miranda* through its X2 and Neo-Y chromosomes. This is an exceptional situation in which the same chromosomal element, C, is evolving both as an X and a Y chromosome in the same species (Fig 3. Chapter 1). Evolution of the Y chromosome is characterized by a loss of genetic activity and it is expected that many neo-Y linked genes are inactive. In such cases while the females, with two X2 chromosomes will have two doses of element C linked genes, males with one X2 and one neo-Y chromosome may have one or two doses depending on whether the gene on the neo-Y is active or inactive. An X2-linked gene whose neo-Y linked homolog has been inactivated, will face the problem of dosage compensation and indeed transcription autoradiography experiments have suggested that portions of the X2 chromosome are transcriptionally hyperactivated in males. To characterize both the mechanism involved in the inactivation of neo-Y linked genes and those involved in the resulting dosage compensation of the corresponding X2-linked genes is a major interest of Dr. Ranjan Ganguly in whose laboratory this dissertation research was carried out. The demonstration that a gene from *D. miranda* can be successfully expressed in *D. melanogaster* and that the dosage-compensatory factors are at least partially conserved in the two species, sets the stage for such studies on dosage-compensated X2-linked genes as and when they are identified.

BIBLIOGRAPHY

- Abraham, I. and Lucchesi, J. C. (1974) Dosage compensation in *Drosophila pseudobscura* and *Drosophila willistonii*. *Genetics*, 79, pp. 11-19.
- Aronson, J. F., Rudkin, G. T. and Schultz, J. (1954) A comparison of giant X-chromosomes in male and female *Drosophila melanogaster* by cytophotometry in the ultraviolet. *J. Histochem. Cytochem.* 2, pp. 458-459.
- Ashburner, M. (1989) in: *Drosophila: A laboratory handbook*. CSH press.
- Baker, B. S. and Belote, J. M. (1983) Sex determination and dosage compensation in *Drosophila melanogaster*. *Ann. Rev. Genet.*, 17, 345-393.
- Banerji, J., Olson, L. and Schaffner, W. (1983) A lymphocyte specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell*, 33, pp.729-740.
- Belote, J. M. and Lucchesi, J. C. (1980) Control of X chromosome transcription by *maleless* gene in *Drosophila*. *Nature*, 285, pp. 573-575.
- Benton, W. D. and Davis, R. W. (1977) *Science*, 196, pp. 180-182.
- Blackman, R. K. and Meselson, M. (1986) Interspecific nucleotide sequence comparisons used to identify regulatory and structural features of the *Drosophila Hsp82* gene. *J. Mol. Biol.* 188, pp. 499-515.
- Bode, J., Kohwi, L., Dickinson, T., Klehr, J. D., Mielke, T. and Kohwi-Shigematsu, T. (1992) Biological significance of unwinding capability of nuclear matrix-associating DNA. *Science*, 255, pp. 195-197.

- Bonifer, C. et al. (1991) Dynamic Chromatin: The regulatory domain organization of Eukaryotic gene loci. *J. Cell. Biochem.*, 47, pp. 99-108.
- Breen, T. R. and Lucchesi, J. C. (1986) Analysis of dosage compensation of a specific transcript in *Drosophila melanogaster*. *Genetics*, 112, 483-491.
- Bridges, C. B. (1922) The origin of variations in sexual and sex-limited characters. *Amer. Naturalist*, 56, pp. 51-63.
- Brinster, R. L. et al. (1981) Somatic expression of Herpes thymidine kinase gene in mice following injection of a fused gene into eggs. *Cell*, 27, pp. 223-231.
- Bull, J. J. (1983) Evolution of sex determination mechanisms. Benjamin/ Cummings, Menlo Park, California.
- Buratowski, S., Hahn, S., Guarente, L. and Sharp, P. A. (1989) Five intermediate complexes in transcription initiation by RNA Polymerase II. *Cell*, 56, pp. 549-561.
- Chambon, P. et al. (1984) Promoter elements of genes coding for proteins and modulation of transcription by estrogens and progesterone. In "Recent progress in Hormone Research, 40, pp. 1-39.
- Charlesworth, B. (1991) The evolution of sex chromosomes. *Science* 251, pp. 1030-1033.
- Chatterjee, R. N. and Mukherjee, A. S. (1981) In situ transcription analysis of chromatin template activity of the X chromosome of *Drosophila* following high molar NaCl treatment. *Chromosoma*, 82, 515-523.
- Cline, T. W. (1988) Evidence that *sisterless-a* and *sisterless-b* are two of several discrete "numerator elements" of the X:A sex determination signal in *Drosophila* that switch *Sxl*

- between two alternative stable expression states. *Genetics*, 119, pp. 829-862.
- Cline, T. W. (1985) Primary events in the determination of sex in *Drosophila melanogaster*. in "The Origin and Evolution of Sex, Alan R. Liss, Inc., pp. 301-327.
- Dobzhansky, T. (1957) The X chromosome in the larval salivary glands of hybrids of *Drosophila insularis* x *Drosophila tropicalis*. *Chromosoma*, 8, pp. 691-698.
- Duffy, J. and Peter Gergen, J. (1991) The *Drosophila* segmentation gene *runt* acts as a position specific numerator element necessary for the uniform expression of the sex-determining gene *Sex-lethal*. *Genes and Development*, 5, pp. 2176-2187.
- Dynan, W.S. (1987) What mechanisms underlie tissue-specific gene transcription? *Trends in Genetics*, 3, pp. 121-122.
- Eisenberg, J. C. and Lucchesi, J. C. (1976) Chromatin structure and dosage compensated heat shock genes in *Drosophila pseudobscura*. In: From Bacteria to Man. M. J. Schlesinger, M. Ashburner and A. Tissieres, eds., Cold Spring Harbour Laboratory Publ., pp. 109-113.
- Eisenberg, J. C. and Lucchesi, J. C. (1983) Chromatin structure and transcriptional activity of an X-linked heat-shock gene. *J. Biol. Chem.*, 258, pp. 13986-13991.
- Faizullin, L. Z. and Gvozdev, V. A. (1973) Dosage compensation of sex-linked genes in *Drosophila melanogaster*. *Mol. Gen. Genet.* 126, pp. 233-245.
- Felsenfeld, G. (1992) Chromatin as an essential part of the transcriptional mechanism. *Nature*, 355, pp. 219-224.

- Fortini, M. E. and Rubin, G. M. (1990) Analysis of cis-acting requirements of the Rh3 and Rh4 genes reveals a bipartite organization to rhodopsin promoters in *Drosophila melanogaster*. *Genes and Development*, 44, pp. 444-463.
- Foulkes, N. S., Borrelli, E. and Sassone-Corsi, P. (1991) CREM Gene: Use of alternative DNA-binding domains generates multiple antagonists of cAMP-induced transcription. *Cell*. 64, pp. 739-749.
- Fukunaga, A., Tanaka, A. and Oishi, K. (1975) *Maleless*, a recessive autosomal mutant of *Drosophila melanogaster* that specifically kills male zygotes. *Genetics*, 81, pp. 135-141.
- Ganguly, R., Ganguly, N. and Manning, J. E. (1985) Isolation and characterization of the glucose-6-phosphate dehydrogenase gene of *Drosophila melanogaster*. *Gene*, 35, pp. 91-101.
- Gasser, S. M. and Laemli, U. K. (1987) A glimpse at chromosomal order. *T.I.G.* 3, pp. 16-22.
- Ghosh, S., Chatterjee, R. N., Bunick, D., Manning, J. E. and Lucchesi, J.C. (1989) The *LSP1-alpha* gene of *Drosophila melanogaster* exhibits dosage compensation when it is relocated to a different site on the X-chromosome. *EMBO J.* 8, pp. 1191-1196.
- Gillies, S. D. (1983) A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. *Cell*, 33, pp. 717-728.
- Goldberg, D. A., Posakony, J. W. and Maniatis, T. (1983) Correct developmental expression of a cloned *Alcohol dehydrogenase* gene transduced into the *Drosophila* germ line. *Cell*, 34, pp. 59-73.

Golubovsky, M. D., and Ivanov, Y. N. (1972) Autosomal mutation in *Drosophila melanogaster* killing the males and connected with female sterility. *Drosophila Info. Serv.*, 49, pp.117.

Grunstein, M. (1990) *Ann. Rev. Cell Biol.*, 6, pp. 643-678.

Guarente, L. (1988) UASs and Enhancers: Common mechanism of transcriptional activation in yeast and mammals. *Cell*, 52, pp. 303-305.

Gutierrez, A. G., Christensen, A. C., Manning, J. E. and Lucchesi, J. C. (1989) Cloning and dosage compensation of the 6-phosphogluconate dehydrogenase gene (*Pgd*⁺) of *Drosophila melanogaster*.

Hafen, E. and Levine, M. (1986) The localization of RNAs in *Drosophila* tissue sections by *in situ* hybridization. In: Roberts, D. B. ed. *Drosophila: A Practical Approach*. IRL Press, Oxford, pp. 139-158.

Hazelrigg, T., Levis, R., and Rubin, G. M. (1984). Transformation of *white* locus DNA in *Drosophila*: Dosage compensation, *zeste* interaction and position effects. *Cell*, 36, pp. 469-481.

Helms, C., Graham, M. Y., Dutchik, J. E. and Olson, M. (1982) A new method for purifying Lambda DNA from phage lysates. *DNA*, 4, pp. 39-49.

Henikoff, S. and Eghtedarzadeh, M. K. (1987) Conserved arrangement of nested genes at the *Drosophila* Gart locus. *Genetics* 117, pp. 711-725.

Heschl, M. F. P. and Baille, D. L. (1990) Functional elements and domains inferred from sequence comparisons of a heat shock gene in two nematodes. *J. Mol. Evol.* 31, pp. 3-9.

Hiebert, J. C. and Birchler, J. A. (1992) Dosage compensation of the *copia* retrotransposon in *Drosophila melanogaster*. *Genetics*, 130, pp. 539-545.

Hofmann, A., Keinhorst, A., Krumm, A. and Korge, G. (1987) Regulatory sequences of the *Sgs-4* gene of *Drosophila melanogaster* analyzed by P-element mediated transformation. *Chromosoma*, 196, pp. 8-17.

Hyde, D. R., Mecklenburg, K. L., Pollock, J. A., Vihtelic, T. S. and Benzer S. Twenty visual system cDNA clones: One is a homolog of human arrestin. *Proc. Natl. Acad. Sci. USA.*, 87, pp. 1008-1012.

Inoue, Y. (1980) *eym: eyemissing*. *Dros. Inf. Service.*, 55, pp. 206.

Jablonka, E. and Lamb, M. J. (1990) The evolution of heteromorphic sex chromosomes. *Biol. Rev.*, 65, pp. 249-276.

Jaffe, E. and Laird, C. (1986) Dosage compensation in *Drosophila*. *TIG*, 6, 316-321.

Kaiser, K., Furia, M. and Glover, D. M. (1986) Dosage compensation at the *sgs4* locus of *Drosophila melanogaster*. *J. Mol. Biol.*, 187, 529-536.

Karess, R. E. and Rubin, G. M. (1984) Analysis of P transposable element functions in *Drosophila*. *Cell*, 38, pp. 135.

Karess, R. E. (1985) P Element mediated germ line transformation of *Drosophila*. In: *DNA cloning. A Practical Approach*. Vol. II. ed. D. M. Glover, IRL Press. pp. 121-141.

Kazazian, H. H. Jr., Young, W. J. and Childs, B. (1965) X-linked *6-phosphogluconate dehydrogenase* in *Drosophila*: subunit associations. *Science*, 150, pp. 1601-1602.

Korge, G. (1981) Genetic analysis of the larval secretion gene *Sgs-4* and its regulatory chromosomal sites in *Drosophila melanogaster*. *Chromosoma*, 62, pp. 155-174.

Korge, G. (1975) Chromosome puff activity and protein synthesis in larval salivary glands of a larval saliva protein in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA*, 72, pp. 4550-4554.

Krishnan, R. and Ganguly, R. (1990). Nucleotide sequence of the arrestin-like 49-kd protein gene of *Drosophila miranda*.

Krishnan, R., Swanson, K. and Ganguly, R. (1991) Dosage compensation of a retina specific gene in *Drosophila miranda*. *Chromosoma*, 100, pp. 125-133.

Krumm, A., Roth, G. E. and Korge, G. Transformation of salivary gland secretion protein gene *Sgs-4* in *Drosophila*: stage- and tissue-specific regulation, dosage compensation and position effect. *Proc. Natl. Acad. Sci. USA*, 82, pp. 5055-5059.

Kuroda, M. I., Kernan, M. J. Kreber, R. Ganetzky, B. and Baker, B. S. (1991) The *maleless* protein associates with the X chromosome to regulate dosage compensation in *Drosophila*. *Cell*, 66, pp. 935-947.

Laurie-Ahlberg, C. C. and Stam, L. F. (1987) Use of P-element mediated transformation to identify the molecular basis of naturally occurring variants affecting *Adh* expression in *Drosophila melanogaster*. *Genetics*, 115, pp. 129-140.

Le Vine III, H., Smith, D. P., Whitney, M., Malicki, D. M., Dolph, P. J., Smith, G. F. H., Burkhart, W. and Zuker, C. S. (1990) Isolation of a novel visual-system-specific arrestin: an *in-vivo* substrate for light dependant phosphorylation. *Mechanisms in Development*, 33, pp. 1-96.

Levis, R., Hazelrigg, T. and Rubin, G. M. (1985) Seperable cis-acting control elements for expression of the white gene of *Drosophila*. EMBO J. 4, pp. 3489-3499.

Levy, L. S., Ganguly, R., Ganguly, N. and Manning, J. E. (1982) The selection, expression, and organization of a set of head-specific genes in *Drosophila*. Dev. Biol. 94, pp. 451-464.

Lewin, B. (1990) Commitment and activation at Pol II promoters: A tail of protein-protein interactions. Cell, 61, pp. 1161-1164.

Losa, R. and Brown, D. D. (1987) Cell, 50, pp. 801-808.

Lucchesi, J. C., Belote, J. M. and Maroni, G. (1977) X-linked gene activity in metamales (XY;3A) of *Drosophila*. Chromosoma, 65, pp. 1-7.

Lucchesi, J. C. and Rawls, J. M. Jr. (1973b) Regulation of gene function: a comparison of enzyme activity levels in normal and intersexual triploids of *Drosophila melanogaster*. Genetics, 73, pp. 459-464.

Lucchesi, J. C. and Rawls, J. M. Jr. (1973a) Regulation of gene function: a comparison of enzyme activity levels in relation to gene dosage in diploids and triploids of *Drosophila melanogaster*. Biochem. Genetics, 9, pp. 41-51.

Lucchesi, J. C. (1978) Gene dosage compensation and the evolution of sex chromosomes. Science, 202, 711-716.

Lucchesi, J. C. (1973) Dosage compensation in *Drosophila*. Ann. Rev. Gen. 7, pp. 225-237.

- Lucchesi, J. C. and Manning, J. E. (1987) Gene dosage compensation in *Drosophila melanogaster*. *Adv. Genetics* 24, pp. 371-429.
- Lyon, M. F. (1961) Gene action in the X chromosome of the mouse (*Mus musculus*) *Nature* 190, pp. 372-373.
- Maniatis, T., Fritsch, D. F. and Sambrook, J. (1982) Molecular cloning. A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Maroni, G. and Plaut, W. (1973b) Dosage compensation in *Drosophila*. II. Glucose-6-phosphate dehydrogenase activity. *Genetics*, 74, pp. 331-342.
- Maroni, G. and Plaut, W. (1973a) Dosage compensation in *Drosophila melanogaster* triploids. I. Autoradiographic study. *Chromosoma*, 40, pp. 361-377.
- Maroni, G. and Lucchesi, J. C. (1980) X chromosome transcription in *Drosophila*. *Chromosoma*, 77, pp. 253-261.
- Marsh, J. L., Gibbs, P. D. and Timmons, P. M. (1985) *Mol. Gen. Genet.*, 198, pp. 393-403.
- Martin, C. H., Mayeda, C. A. and Meyerowitz, E. M. (1988) Evolution and expression of the *Sgs-3* glue gene of *Drosophila*. *J. Mol. Evol.* 201, pp.273-287.
- McNabb, S. L. and Beckendorf, S. K. (1986) Cis-acting sequences which regulate expression of the *Sgs-4* glue protein gene of *Drosophila*. *EMBO J.* 5, pp. 2331-2340.
- Mismer, D. and Rubin, G.M. (1987) Analysis of the promoter of the *NinaE* opsin gene in *Drosophila melanogaster*. *Genetics*, 116, pp. 565-578.

- Mukherjee, A. S. (1986) Genetics of dosage compensation: An old concept in new perspective. Presidential address, Seventythird Indian Science Congress, Delhi, India.
- Mukherjee, A. S. and Beerman, W. (1965) Synthesis of ribonucleic acid by the X chromosomes of *Drosophila melanogaster* and the problem of dosage compensation. *Nature*, 207, pp. 785-786.
- Mukherjee, A. S. and Chatterjee, S. N. (1975) *Chromosoma*, 53, pp. 91-96.
- Mukherjee, A. S. (1966) Dosage compensation in *Drosophila*: an autoradiographic study. *The Nucleus*, 9, pp. 83-96.
- Muller, H. J., League, B. B. and Offerman, C. A. (1931) Effects of dosage changes of sex-linked genes, and the compensatory effects of other gene differences between male and female. (Abstr.) *Anat. Rec.* 51 (suppl.) pp. 110.
- Muller, H. J. (1932) Further studies on the nature and causes of gene mutations. *Proc. 6th Int. Congr. Genet.* 1, pp. 213-255.
- Muller, H. J. (1950) Evidence of the precision of genetic adaptation. Harvey Lecture Series XLIII, pp. 165-229.
- Offerman, C. A. (1936) Branched chromosomes as symmetrical duplications. *J. Genet.*, 32, pp. 103-116.
- Ohno, S., Kaplan, W. D. and Kinosita, R. (1957) Conjugation of the heteropycnotic X and Y chromosomes of the rat spermatocyte. *Exptl. Cell Research*, 13, pp. 395-426.
- Parkhurst, S. M., Bopp, D., Ish-Horowicz, D. (1990) X:A ratio, the primary sex determining signal in *Drosophila* is transduced by helix-loop-helix proteins. *Cell*, 63,

1179-1191.

Patterson, J. I. and Stone, W. S. (1952) Evolution in the Genus *Drosophila*. MacMillan Press, New York.

Phi-Van, L. and Stratling, W. H. (1990) Association of DNA with Nuclear Matrix. in: Jeanteur, P. et al. ed. Progress in Molecular and Subcellular Biology 11, pp. 1-9.

Pierce, D. A. and Lucchesi, J. C. (1980) Dosage compensation of X-linked heat-shock puffs in *Drosophila pseudobscura*. Chromosoma, 76, pp. 245-254.

Pirrotta, V., Steller, H. and Bozetti, M. P. (1985) Multiple upstream regulatory elements control the expression of the *Drosophila white* gene. EMBO J., 4, pp. 3501-3508.

Ptashne, M. and Gann, A.A.F. (1990) Activators and Targets. Nature, 346, pp. 329-331.

Roehrdanz, R. L., Kitchens, J. M. and Lucchesi, J. C. (1977) Lack of dosage compensation for an autosomal gene relocated to the X chromosome in *Drosophila melanogaster*. Genetics, 85, pp. 489-496.

Rubin, G. M. and Spradling, A. C. (1982) Genetic transformation of *Drosophila* with transposable element vectors. Science, 218, pp. 348-353.

Rudkin, G. T. (1964) The proteins of polytene chromosomes. "The nucleohistones," eds. Bonner and Ts'o, pp. 184-192. Holden Day, San Fransisco.

Russell, L. B. (1963) Evidence from six X-autosome translocations bearing on the single-active-X hypothesis. Proceedings of the XIth International Congress of Genetics, The Hague, Netherlands, September 1963.

Russell, L. B. and Bangham, J. W. (1959) Variegated type position effects in the mouse. *Genetics* 44, pp. 532-.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74, pp. 5463-5467.

Sass, H. and Meselson, M. (1991) Dosage compensation of the *D. pseudobscura* HSP82 gene and the *melanogaster* *Adh* gene at ectopic sites in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA*, 88, pp. 6795-6799.

Scholnick, S. B., Morgan, B. A. and Hirsh, J. (1983) The cloned *Dopa decarboxylase* gene is developmentally regulated when reintegrated into the *Drosophila* genome. *Cell*, 34, pp. 37-45.

Sederoff, R., Clynes, R., Poncz, M. and Hachtel, S. (1973) RNA synthesis by exogenous RNA polymerase on cytological preparations of chromosomes. *J. Cell Biol.*, 57, pp. 538-550.

Seecof, R. L., Kaplan, W. D. and Futch, D. G. (1969) Dosage compensation for enzyme activities in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA.*, 62, 528-535.

Serfling, E., Jasin, M. and Schaffner, W. (1985) Enhancers and eukaryotic gene transcription. *Trends in Genetics*, 5, pp. 224-230.

Shapiro, M. B. and Senapathy, P. (1987) RNA splice junctions of different classes of Eukaryotes: sequence statistics and functional implications in gene expression. *Nucl. Acids. Res.*, 15, pp. 7155-7174.

Simonsen, C. C. and Levinson, A. D. (1983) Analysis of processing and polyadenylation

- signals of the hepatitis B virus surface antigen gene by using simian virus 40-hepatitis B virus chimeric plasmids. *Mol. Cell. Biol.* 3, pp. 2250-2258.
- Smith, D. P., Shieh, B. H. and Zuker, C. (1990) Isolation and structure of an arrestin gene from *Drosophila*. *Proc. Natl. Acad. Sci. USA.* pp. 1003-1007.
- Spradling, A. C. and Rubin, G. M. (1983) The effect of chromosomal position on the expression of the *Drosophila* xanthine dehydrogenase gene. *Cell*, 34, pp. 47-57.
- Steinmann, M. (1982) Analysis of chromosomal homologies between two species of the subgenus *sophophora*: *D. miranda* and *D. melanogaster* using clone DNA segments. *Chromosoma*, 87, pp. 77-88.
- Steinmann, M., Pinsker, W. and Sperlich, D. (1984) Chromosomal homologies within the *Drosophila obscura* group probed by *in situ* hybridization. *Chromosoma*, 91, pp. 46-53.
- Stern, C. (1960) Dosage compensation - development of a concept and new facts. *Can. J. Genet. Cytol.* 2, 105-118.
- Stern, C. (1929) Uber die additive Wirkung multipler Allele. *Biolog. Zbl.* 49, pp. 261-290.
- Stief, A., Winter, D. M., Stratling, W. H. and Sippel, A. E. (1989) A Nuclear DNA attachment element mediates elevated and position-independent gene activity. *Nature*, 341, pp. 343-345.
- Strobel, E., Pelling, C. and Arnheim, N. (1978) Incomplete dosage compensation in an evolving *Drosophila* sex chromosome. *Proc. Natl. Acad. Sci. USA*, 75, 931-935.

- Tanaka, A., Fukunaga, A. and Oishi, K. (1976) Studies on the sex-specific lethals of *Drosophila melanogaster*. II. Further studies on the male-specific lethal gene, *maleless*. *Genetics*, 84, 257-266.
- Tobler, J., Bowman, J. T. and Simmons, J. R. (1971) Gene modification in *Drosophila*: dosage compensation and relocated v^+ genes. *Biochem. Genet.*, 5, pp. 111-117.
- Torrez, M. and Sanchez, L. (1989) The *scute* (T4) gene acts as a numerator element of the X:A signal that determines the state of activity of *Sex-lethal* in *Drosophila*. *EMBO J.*, 8, pp. 3079-3086.
- Urmeil, N. and Plaut, W. (1973) Interaction of poly-L-lysine with chromatin. *J. Cell Biol.*, 56, pp. 139-144.
- Weintraub, H. and Groudine, M. (1976) Chromosomal subunits in active genes have an altered conformation. *Science*, 193, pp. 848-856.
- Yamada, T. et al. (1990) A 49-kd phosphoprotein in the *Drosophila* photoreceptor is an *arrestin* homolog. *Science*, 248, pp. 483-486.
- Yerganian, G. (1960) "*Brittle-bristle*", genetic evidence for the triheterosomic scheme of sex determination. *Genetics*, 45, pp. 1018.

APPENDIX

METHODS

I. Screening of genomic and cDNA libraries

The genomic libraries of *D. miranda* (Krishnan et al 1991) or *D. melanogaster* (Levy et al 1982) genomic libraries were previously constructed in the lambda vector EMBL3. The *D. melanogaster* cDNA library was in lambda swaj2 vector and was received as a gift (Palazollo et al.). Host KH802 cells, grown aerobically for 16 hours were infected with 1×10^5 pfu per 100 mm plate and grown to confluence. Plaques were blotted on to nitrocellulose as follows: Nitrocellulose filters were placed on the surface of the plates for thirty seconds and three holes punched asymmetrically to indicate orientation. Filters were then removed into 0.1 M NaOH and 1.5 M NaCl for five minutes, lightly dabbed on whatmann filter paper and placed in 0.2 M Tris.HCl pH 8.0 and 2X SSCP (10X SSCP is 1.2 M NaCl, 0.15 M sodium citrate, 0.13 M K_2HPO_4 and 10 mM EDTA). The blots were air dried and baked in vacuo at 80° C for three hours. Blots were pre-hybridized in 6X SSC, 5X Denhardt's solution with 200 ug/ml denatured sonicated salmon sperm DNA for four hours at 68° C. Hybridization was carried out under the same conditions with $10\text{--}15 \times 10^6$ cpm of probe. Plaques from sites of hybridization were picked and placed in SM (100 mM NaCl, 8 mM $MgSO_4$, 50 mM Tris.HCl pH 7.5, 0.01 % gelatin) at 4° C. they were then titred and appropriately replated with KH802 cells at a lower density to get well isolated plaques which were then subjected to a second round of screening as described above. The well isolated and therefore cleanly picked plaques from this round of screening were subjected to a third round of screening to establish purity. Pure plaques were then stored in 0.5 ml of SM at 4° C.

II. Nucleic acids isolation

Isolation of genomic DNA

Whole adult flies were frozen at -80° C until use. Frozen flies were homogenized

in Nuclei Isolation Buffer (NIB) (50mM Tris HCl pH 7.6, 25 mM KCl, 5 mM magnesium acetate, 350 mM sucrose) using a Dounce homogenizer. The nuclear pellet was obtained by centrifugation at 1800 x g and was washed twice with NIB. The washed pellet was then resuspended in DNA Extraction Buffer (DEB) (0.15 M NaCl, 100 mM Tris HCl pH 8.0, 10 mM EDTA). Sodium perchlorate was added to a final concentration of 0.5 M and incubated for five minutes with gentle mixing. SDS was added to a final concentration of 0.1 % with gentle mixing, and further incubated for five minutes. This was then extracted thrice with equal volume of a 1:1 mixture of DEB saturated phenol and chloroform and then twice with equal volume of chloroform. To the final aqueous phase, 1/10th volume of 2M NaOAc pH 7.5 added and nucleic acids precipitated in 2.5 volumes of 95 % ethanol at -20° C. DNA was dissolved in TE (10 mM Tris HCl, 1mM EDTA) and quantitated spectrophotometrically at 260 nm.

Isolation of plasmid DNA

Plasmid DNA were isolated by the alkaline lysis method. Cells were grown aerobically in Luria broth (1% Tryptone, 0.5% yeast extract, 1% NaCl, 0.2% $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ and 1% Casamino acids) containing 50 ug/ml ampicillin. Cells were harvested by centrifugation at 1800 x g, resuspended in Solution I (50 mM glucose, 10 mM EDTA, 25 mM Tris.HCl) and incubated at room temperature for 5 minutes. Two volumes of freshly prepared solution II (0.2 N NaOH and 1% SDS) were added with gentle mixing and the mixture incubated on ice for 5 minutes. One and a half volumes of Solution III were added with gentle mixing and incubation continued for further five minutes. Cell debris was removed by centrifugation at 15000 x g and the resulting supernatant filtered through 4 layers of cheese cloth. Nucleic acids were precipitated by addition of chilled 1.2 volumes of isopropanol at -20° C for 10 mins. Nucleic acids collected by centrifugation at 15000 x g were dissolved in TE (10 mM Tris.HCl pH 7.6, 1 mM EDTA). Half volume of 7.5 M ammonium acetate was added and the mixture incubated on ice for 10 minutes. The resulting precipitate or RNA was separated by centrifugation at 15000 x g and nucleic acids precipitated from the supernatant with 2.5 volumes of chilled ethanol at -80° C for 30 mins. The nucleic acid pellet obtained was resuspended in TE and treated with

DNase-free RNase at a concentration of 10 ug/ml for 30 minutes at 37° C. This was extracted twice with equal volume of a 1:1 mixture of DEB saturated phenol and Chloroform and twice with chloroform. To final aqueous phase 1/10th volume 2 M sodium acetate pH 7.5 and 2.5 volumes ethanol added for precipitation at -20° C. Plasmid DNA was dissolved in TE or in water if to be used for sequencing reactions.

Isolation of phage DNA

5 x 10⁵ pfus of phage were incubated with 100 ul of host bacteria (18 hr culture) for fifteen minutes at 37 C and bulk seeded with 3 ml of LD top agar (5% tryptone, 2.5 % yeast extract, 2.5 % NaCl, 0.7% bactoagar) on 100 mm LD bottom agar (1.5 % agar) plates. Incubation was for approximately seven hours until just post confluence. The plates were then cooled, overlayed with 6 ml lambda diluent (10 mM Tris, pH 8.0, 2 mM MgCl₂) and kept overnight at 4° C. The phage containing diluent from each plate was loaded onto a 2 ml DE52 cellulose column and washed with 3-5 ml chase buffer (10 mM Tris HCl pH 8.0, 10 mM magnesium acetate, 60 mM sodium acetate). The phage were concentrated at the bottom of the column with 1.2 ml of elution buffer (10 mM Tris HCl pH 8.0, 50 mM magnesium acetate) and collected in 0.9 ml of the same buffer. Proteinase K and SDS were added to the eluent to a concentration of 0.8 mg/ml and 0.4%, respectively and this mixture incubated at room temperature for five minutes. 3M potassium acetate (ph 8.8) was then added to a final concentration of 1 M and the resulting suspension was incubated at 88° C for 15 minutes. After cooling on ice for 15 minutes, the resulting, the resulting precipitate was removed by centrifugation at 15000 x g and phage DNA precipitated from the supernatant by addition of 2.5 volumes of chilled ethanol. The DNA was dissolved in TE and quantitated spectrophotometrically at 260 nm.

Isolation of total cellular RNA

RNA was isolated either from fresh flies (for quantitative northern blots) or from flies frozen at -80° C (for qualitative northern blots). For quantitation fresh flies of the same age were used. Heads and bodies were separated by freezing flies in liquid nitrogen

and shaking them vigorously for decapitation following which heads were separated with the help of a sieve.

Flies were homogenized in hot (68° C) phenol saturated with 0.2 M sodium acetate pH 5.0 and 4% SDS. To the homogenate equal volume of 0.2 M sodium acetate pH 5.0 was added and after thorough mixing was incubated at 68° C for ten minutes. This was then cooled to room temperature mixed with half volume of chloroform, mixed and the aqueous phase obtained by centrifugation at 1800 x g. This was then extracted twice with two volumes of a 1:1 mixture of 0.2 M sodium acetate saturated phenol and chloroform and then once with two volumes of chloroform. RNA was precipitated with 2.5 volumes of ethanol. RNA pellets obtained by centrifugation at 15000 x g were dissolved in water pretreated with diethylpyrocarbonate. Quantitation was done spectrophotometrically at 260 nm.

Isolation of poly A+ RNA

Oligo dT cellulose column was equilibrated with 1X TES and 0.5 M NaCl (1X TES is 10 mM Tris.HCl pH 7.5, 1 mM EDTA, 0.1 % SDS) until column eluent showed a pH of 7.0. The total cellular RNA to be fractionated was diluted to 0.5 mg/ml, 1/10th volume of 10X TES added and the sample incubated at 68° C for 2 mins and then placed in ice for one minute. 1/10th volume of 5 M NaCl was added and the sample loaded onto column. The column was then washed with 10 ml of 1X TES and 0.5 M NaCl and then with 10 ml of 1X TES. The eluent was collected in six fractions and RNA in each fraction quantitated by spectrophotometry at 260 nm. Peak fractions were pooled and RNA precipitated by addition of 1/10th volume 2M sodium acetate and 2.5 volumes ethanol at -20° C.

III. Gel electrophoresis and blot transfer of nucleic acids

Southern blotting

DNA to be blotted were electrophoresed at 5 volts/ cm in agarose gels of appropriate concentration, containing 1 ug/ml ethidium bromide. Gel and chamber buffer

was TBE (90 mM Tris pH 8.0, boric acid and 2 mM EDTA). After electrophoresis, gels were treated with denaturing solution (0.5 N NaOH, 1.5 M NaCl) for one hour followed by two changes of neutralizing solution (3M NaCl, 0.5 M Tris.HCl pH 7.6) for 2-3 hours. The DNA was blotted onto reusable nylon membrane using 6X SSC (1X SSC is 150 mM NaCl, 15 mM sodium citrate) as the transfer buffer. The DNA was fixed onto the nylon membrane by baking in vacuo at 80° C for 4 hours.

Southern hybridization

The blots were prehybridized at 68° C in 6X SSC and 5X Denhardt's (1X Denhardt's is 0.02% ficoll, 0.02% Polyvinyl pyrrolidone, and 0.02% bovine serum albumin fraction V) and 100 ug/ml denatured sonicated salmon sperm DNA for four hours. Hybridization was then carried out under the same conditions for 16-18 hours. After hybridization, blots were washed in two changes of 0.1X SSC, 0.1 % SDS at 60° C for 30 minutes and then in two changes of 0.1X SSC at 60° C for 30 minutes. wet blots were covered with Saran wrap and exposed to X-ray film (Kodak) at -20° C for appropriate lengths of time.

Northern blotting

RNA samples were denatured in 50% deionized formamide at 68° C and electrophoresed at 5 volts/cm in 1.1% agarose gels containing 6.6% formaldehyde. Gel and chamber buffer were MOPS (40 mM 3-(N-Morpholino) propanesulfonic acid, 10 mM sodium acetate, 1 mM EDTA). After electrophoresis, gels were soaked in water for an hour to remove formaldehyde and stained in 2 ug/ml ethidium bromide prior to blotting. The RNA were blotted onto reusable nylon membranes by capillary transfer using 20X SSC as the transfer medium and were fixed onto the blots by baking in vacuo at 80 C for four hours.

Northern hybridization

The blots were prehybridized at 42° C in 50 % Formamide, 0.8 M NaCl, 0.1 M PIPES pH 7.0, 5X Denhardt's solution and 100 ug/ml denatured sonicated salmon sperm

DNA for three hours. Hybridization was carried out under the same conditions for 16-18 hours. After hybridization, blots were washed in 4 changes of 2x SSC, 0.05% sodium pyrophosphate, 0.05% SDS at room temperature for 40 minutes followed by two changes of 0.2X SSC, 0.05% sodium pyrophosphate, 0.05% SDS at 60° C for 30 minutes. Wet blots were covered with saran wrap and exposed to X-ray film (Kodak) at -20° C for appropriate lengths of time. Blots to be rehybridized were prehybridized for an hour as described above before the next round of hybridization.

IV. Radiolabeling of nucleic acid probes

Nick translation

DNA was nick translated using commercially available reaction kit from Bethesda research laboratories (BRL) according to the instructions provide by the manufacturer. The probe DNA was separated from unincorporated nucleotides by exclusion chromatography using sephadex G-50 columns.

Random primed labeling

DNA to be labeled was denature by incubation for three minutes at 90° C with 3 ng/ul random hexanucleotide primers. The mixture was then cooled on ice and the primer extension reaction set up in 1X oligo labeling buffer (200 mM HEPES pH 6.6, 50 mM B-mercaptoethanol, 50 mM Tris pH 8.0, 5 mM MgCl₂, 0.6 mM EDTA), 50 uCi of ³²P-labeled dATP or ³⁵S-labeled dATP, 2.5 mg/ml BSA, and 0.5 mM each of dGTP, dCTP, dTTP, and 5 units of the Klenow fragment of DNA polymerase I. The mixture was incubated for two hours at 25° C prior to purification of probe by sephadex G-50 column chromatography.

Heterogeneous cDNA probe

Poly A+ RNA were mixed with 50 uCi alpha ³²P dCTP, 2.4 ng random hexanucleotide primers and methyl mercury hydroxide (10 mM) and incubated for 10 minutes at room temperature. B-mercaptoethanol to 100 mM and 60 units of RNasin were

added and incubated for further five minutes. This was then added to the reverse transcriptase reaction buffer (50mM Tris.HCl pH 7.5, 75 mM KCl, 6 mM MgCl₂, 200 uM dATP, 200 uM dGTP, 20 uM dCTP, 20 mM DTT). 30 units of AMV reverse transcriptase were then added and incubated for one hour prior to purification of probe column chromatography using sephadex G-50.

V. *In situ* hybridization to tissue sections

Heads of freshly anaesthetized adult flies were manually dissected, embedded in OCT (Tissue-Tek), and frozen in liquid nitrogen. The resulting blocks were then trimmed and equilibrated to the cryostat cutting temperature of -15° C. 6 um sections were placed on gelatinized slides. The slides were plated on a warm plate for five minutes and then fixed in 4% paraformaldehyde for 20 minutes. The sections were then incubated serially for five minutes each in 3X PBS and 1X PBS and then were dehydrated through graded alcohol. Dehydrated sections were stored in a dry box at room temperature until hybridization.

The probe to be used was radiolabeled with ³⁵S-dATP by the random primer method. Denatured probe was added to the hybridization mix consisting of 50% deionized formamide, 0.6% NaCl, 10 mM Tris.HCl pH 7.6, 1 mM EDTA, 1X Denhardt's solution, 27 ug/ml tRNA). Seventeen ul of hybridization mix was added to each slide under a siliconized cover slip. The slides were stored in a humid chamber for 18 hours at 37 C. After hybridization, slides were washed at 37° C in four changes of wash solution (50 % technical grade formamide, 0.6 M NaCl, 10 mM Tris-HCl pH7.6, 1 mM EDTA) within the first hour followed by 3 more changes at four hour intervals. The slides were then covered with Kodak (NTB2) X-ray emulsion and stored in a dry dark box at 4° C till development. After development slide were stained in Giemsa, and a cover slip affixed on the sections using permamount.

VI. *In situ* hybridization to polytene chromosomes

Third instar larvae were rinsed in saline and placed in a drop of 45% acetic acid on gelatinized slides. Salivary glands were dissected out and freed of associated fatty

tissue, and soaked in a fresh drop of 45% acetic acid for three minutes. A siliconized coverslip was placed on the slide and the glands squashed firmly. The salivary gland squashes were incubated overnight at 4° C. The cover slips were dipped in liquid N₂ and the frozen coverslips flipped off with a razor blade. Slides were incubated in 95 % ethanol for 15 minutes, air dried and stored in a dry box until hybridization.

Prior to hybridization slides were incubated in 2X SSC at 68° C for 10 minutes. The chromosomes were transferred to 0.1 M triethanolamine (pH 8.0) at room temperature. Acetic anhydride was added gradually with continuous stirring till a concentration of 5ml/litre of triethanolamine. After fifteen minutes, the slides were washed in twice in 2X SSC at room temperature for five minutes each, dehydrated through an alcohol gradient and air dried. The chromosomes were then denatured in 0.07N NaOH for five minutes at room temperature and then treated twice with 2X SSC for five minutes each after which they were dehydrated through the alcohol gradient again. Hybridization and washings were carried out as described for head tissue sections.

VII. P-element mediated transformation

Microinjection of DNA into embryos

The P element plasmid construct and helper plasmid DNA were mixed in a 3:1 ratio in a final concentration of 1 mg/ml. Needles were fashioned from glass microcapillary pipettes using a vertical pipette puller (David Kopf instruments), to get a fine point of < 1.0 micron. Staged embryos were collected by inducing flies to shed stored eggs for a period of two hours and thereafter collecting eggs laid at 15 minute intervals. The eggs were dechorionated by rolling them gently over double sided scotch tape, and were aligned on the edge of the tape so that the posterior 1/3rd of the each embryo was overhanging the edge. The embryos were dechorionated and then desiccated for appropriate lengths of time and then covered with halocarbon oil (Series 700, Halocarbon products). They were then placed on the micromanipulator stage for microinjection. DNA was deposited in the pocket just inside the vitelline membrane, where the pole cells form. Any embryo past the cellularization stage was killed as they would be unlikely to yield germline transformants. Injected embryos were incubated in

oil, till they hatched approximately 50 hours later. Early larvae were gently transferred to vials containing standard *Drosophila* medium (Cornmeal, molasses, agar).

Go Flies

Survivors of the microinjection stage are referred to as the Go generation. In many cases since the *ry* gene product is able to diffuse through the body and confer *ry*⁺ eye colour, many Go flies were *ry*⁺. However whether *ry* or *ry*⁺, adults of this generation were then individually mated to host (*ry*) flies of the opposite sex. These matings were carried out in a manner that would maximize the progeny generated each *ry*⁺ Go fly. The progeny of these matings were referred to as G1 flies.

G1 Flies

G1 flies were screened for *ry*⁺ eye colour and such flies were identified as the germ line transformants. Less than 50% of *ry*⁺ Go flies yielded *ry*⁺ G1 progeny. The *ry*⁺ G1 flies are the germ line transformants and were individually propagated by mating with host *ry* flies of the opposite sex to establish G2 isolines.

Balancing G2 Isolines

G2 progeny from each G1 line were crossed to balancer stocks with *ry* background to identify chromosome linkage prior to synthesis of homozygous stocks. *Cyo;ry*, a second chromosome balancer marked with curly wings or *Sb,ry* a third chromosome balancer marked with stubble bristles were used. Markers on balancers are dominant and balancer chromosomes are homozygous lethal. Only virgin females were used in these crosses.

example: transgene on third chromosome

Go	<i>ry</i> / <i>ry</i>		microinject <i>P(507)ry</i> ⁺
	Go adult	X	<i>ry</i> / <i>ry</i>
G1	<i>P(507)ry</i> ⁺ , <i>ry</i> / <i>ry</i>		transgene insertion on third chromosome

P(507)ry+,ry / ry X *Sb,ry / Sb⁺ ry*

G2 *P(507)ry+,ry / Sb,ry* wild type eye colour and stubble bristles

P(507)ry+,ry / Sb,ry X *P(507)ry+,ry / Sb,ry*

G3 *P(507)ry+,ry / P(507)ry+,ry* homozygous for transgene, wild type
eye colour and bristles

AND

P(507)ry+,ry / Sb,ry Heterozygous (balanced) wildtype eye colour and
stubble bristles

VITA

Rajesh Krishnan was born on the fifth of April 1964 to Lalitha and A.V Krishnan in Coimbatore, India. After early education in a variety of schools, he entered St. Xavier's college, University of Bombay, India, and received a Bachelor's degree in Microbiology and Biochemistry, in April 1985. This was followed by Master's degree in Biotechnology with a specialization in Molecular biology from the Department of Zoology, University of Poona, India, in April 1987. Rajesh entered the Cell, Mol. and Dev. Program at the University of Tennessee, Knoxville, in January 1988 and worked towards a doctoral degree in the molecular genetics laboratory of Dr. Ranjan Ganguly.