The importance of the promoter in *Drosophila* dosage compensation

A thesis presented in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Genetics

at Massey University, Palmerston North,
New Zealand.

Corey Laverty

2009

Abstract

Dosage compensation is the equalisation of gene expression from unequal doses of genes. *Drosophila* males up-regulate transcription from their single X chromosome to equal that from the two female X chromosomes. Five *male-specific lethal* (*msl*) genes are required in males, and encode the main agents of the up-regulation. At least these proteins, together with either or both of two non-coding RNAs, form the MSL chromatin-modifying complex. Female-specific translational repression of a key component, *msl2*, limits the complex to males. The MSL complex binds to the X chromosome at hundreds of distinct loci, acetylates nucleosomes, and de-condenses the chromatin. Together with possibly many co-factors, the transcriptional up-regulation caused by MSL complex appears to counteract repressive factors to achieve an average effect of transcriptional doubling.

Here, I have studied the initiation of MSL regulation on the X chromosome with a variety of approaches. In order to study early events, dosage compensation was induced in females with ectopic expression of *msl2* from the tetracycline system. However, low background expression without activation prohibited further studies. To identify novel factors that affect dosage compensation, a reporter gene system based on variable eye size was evaluated. The system provided a dose-dependent phenotype, but could not report additional up-regulation by the MSL complex, and was thus unsuitable for the proposed mutational screen.

The quantifiable *lacZ* gene was measured in a strict comparison of expression from an eye-specfic (GMR) or a constitutive (*armadillo*) promoter. At defined locations on the X chromsome, *armadillo-lacZ* acquired local compensation, but *GMR-lacZ* did not. Further modifications upstream of *GMR-lacZ* increased the response, and confirmed the importance of the promoter in attraction of dosage compensation. To corroborate this with the established importance of genic sequences in MSL attraction, a combinatorial model of attraction is proposed. The relative importance of early or constitutive expression was also tested, by providing *GMR-lacZ* with extra expression through the tetracycline system. A burst of embryonic expression, and constitutive expression, were both insufficient to increase dosage compensation of the transgene. Finally, the compensation of GMR-mediated transgenes was confounded by 'transvection' effects of chromosome pairing. This effect may have wider implications on the study of compensation at individual genes.

Acknowledgements

Thank you, Max Scott, my supervisor and mentor. Your ideas and advice throughout were invaluable. If I have done well, it is because of you. You've taught me how to work, how to research, how to think. Thank you, Kathryn Stowell for encouragement at all stages, and especially for your review of the figures. Al Rowland and David Penny, for tuning my mind to science philosophy, and evolution. Alasdair Noble, for arguing statistics.

To all who helped, more thanks than I can muster: Fang Li for your many hours giving critical help with western blots, polytene preparations, and cloning. Your advice was always accurate and valued. Esther Belikoff for expert advice on micro-injection, and for tirelessly tending my flies when I couldn't. Anja Schiemann for β-galactosidase assays. Carolina Concha for RNA advice. Vikki Weake for getting me started, and shared micro-injections. Helen Fitzsimons for discussions on the draft. All 'Flyspot' members, past and present, especially Abhimanyu Sarkar and Charles Ellen, for a co-operative and considerate work ethic, and a friendly, productive environment. Emma Brasell for making fly food, and for instant help at the spec. when the printer failed. Ava Handley, for countless helpful discussions and thoughts throughout the write-up. Ava, Sarah Lobb, and Jessika Charlton, for food and flywork. Sara Sutherland, for proofing my introduction.

For supplying flies, plasmids, and antibodies: Andreas Bergmann, Johannes Bischof, Michele Calos, Mitzi Kuroda, Herman Steller, Ernst Wimmer. J. Bischof for adice on the ϕC31 system. Art Alekseyenko for advice on the MRE site in armadillo. For financial support: Massey University, the Institute of Molecular Biosciences, and the New Zealand Freemasons.

To those that personally supported me: thank you, thank you, thank you. I could not do this without any of you. Most especially, my mother and father. Your love, generosity, and unquestioning, continual support is breathtaking. Jared and Gemma for your love and help, especially in Ireland. Ava Handley, for your love and encouragement, your strength when I had none, and for listening, advising, laughing. Sophie Borchert for your encouragement and friendship, and for support through the middle years. Lucy Pearse for the start. Angela Woodley for always being there. Emma Brasell, Charles Ellen, for distraction. Amanda Staddon-Smith, Sara Sutherland, Melle Steringa, for tolerance.

Now it is done. I've learnt so much.

Contents

List of Figures

List of Tables

List of abbrv

Aside from standard molecular biology abbreviations, and *le Système international d'unités* (SI) conventions, the following abbreviations are used:

~	approximately
6PGD	6-phosphogluconate dehydrogenase
abbrv	abbreviations
a.k.a.	also known as
arm	armadillo
BAD	biotinylation activation domain
β-gal	β-galactosidase
ChIP	Chromatin Immuno-precipitation
DRE	DNA replication-related element
FUM	fumarase
G	generation
G6PD	glucose-6-phosphate dehydrogenase
GFP	green fluorescent protein
GMR	Glass multimer reporter
H3K36	histone 3, lysine 36
H3K36me3	tri-methylated H3K36
H3S10	histone 3, serine 10
H4K16	histone 4, lysine 16
HA	hemaglutanin
HAT	histone acetyl transferase
hid	head involution defective
HIS	6 histidine residues
hsp	heat-shock protein
IRES	internal ribosome entry site
LB	Luria-Bertani
mle	maleless
mof	males-absent on the first
MRE	MSL recognition element
msl	male-specific lethal
NURF	nucleosome remodeling factor
p	plasmid
PCNA	proliferating cell nuclear antigen
RFP	red fluorescent protein
roX	RNA on the X
RT	reverse transcriptase
rtTA	reverse tTA
SCF	DNA supercoiling factor
SV40	Simian vacuolating virus 40
Sxl	Sex-Lethal
tet	tetracycline
tetO	tetracycline operator (TRE)
TetR	tetracycline repressor
TRE	tetracycline response element
TRF2	TBP-related factor
tsp	transcription start point
tTA	tetracycline-controlled transcriptional activator
tTS	tetracycline-controlled transcriptional silencer
UAS	upstream activation sequence
UTR	un-translated region
X-gal	5-bromo-4chloro-3indolyl-β-D-galactopyranoside
y	yellow
YFP	yellow fluorescent protein
w	white
WT	wild-type

Chapter 1

Introduction

1.1 Dosage Compensation in Drosophila melanogaster

A serious problem of gene deficiency exists in eukaryotic species with sex chromosomes. Where individuals of different sexes differ in their complement of chromosomes, compensation must be made for the accompanying gene dosage problems. The term "Dosage Compensation" is applied to a collection of mechanisms that have evolved to solve this problem, and has been most extensively studied in mammals, *Caenorhabditis elegans*, and *Drosophila melanogaster*.

Essentially two problems must be solved: equilisation of the sex to autosome difference, and equilisation of the sex chromosomes between the sexes. Whereas all three systems appear to up-regulate the expression of X-linked genes to provide equality between the single male X chromosome and paired autosomes, they differ somewhat in their solutions to the second problem. The now over-expressed X chromosomes in mammalian females and *C. elegans* hermaphrodites (both XX) are down-regulated in a female-specific manner to achieve the correct expression level (Lyon, 1961; Meyer and Casson, 1986). The system in *Drosophila*, described below, has no need of the added complexity of female-specific down-regulation, as the up-regulation of the X is male-specific.

Interesting biological questions thus arise: What mediates the up-regulation in *Drosophila* males? How is the process male-specific? How is the X chromosome targeted? What is the mechanism by which a precise two-fold regulation is attained?

1.1.1 Compensation of X chromosome genes

X-linked expression is equal between sexes

Male *Drosophila* posses only one X chromosome, and yet mutations in X-linked genes display phenotypes equivalent to those observed in females, which have two X chromosomes. This was first observed with the apricot eye-colour produced by the mutant w^a allele of the X-linked *white* gene (Muller et al., 1931). Males with only one w^a allele, and females with two, both display the same level of eye pigmentation. Yet the eye colour is inherently capable of responding to differences in gene dose, as levels of pigmentation alter in response to gene duplications and deletions within a sex (Muller, 1932). Muller's results suggest that males and females make equal amounts of X-linked gene products. This is also observed for X-linked genes that encode enzymes with easily measurable activities, such as fumarase (FUM), tryptophan 2,3-dioxygenase, glucose-6-phosphate dehydrogenase (G6PD) and 6-phosphogluconate dehydrogenase (6PGD) (Whitney and Lucchesi, 1972; Tobler et al., 1971; Seecof et al., 1969). More directly, amounts of protein are observed to be equal between the sexes, such as the peptide from the X-linked *Sgs-4* gene (Korge, 1975).

The equality in protein levels is a direct result of modifications to the process of transcrip-

tion. Microscopically-visible polytene chromosomes of the salivary glands were incubated in (tritiated) [^{3}H]uridine and examined by autoradiography (Mukherjee and Beermann, 1965). The amount of radio-labelled uridine incorporated by the single male X was equal to the total incorporated by *both* female X chromosomes, implying an equivalent level of RNA synthesis. Similar results are observed in the quantification of specific RNA transcripts of X-linked genes, including the serine-4 tRNA, and the mRNAs from *Sgs-4* and the G6PD gene *Zw* (Birchler et al., 1982; Breen and Lucchesi, 1986; Kaiser et al., 1986; Ganguly et al., 1985).

The increased level of transcription observed from male X chromosomes could theoretically be a result of more X-linked DNA in males. However, measurements confirming twice the amount of X-linked chromosomal DNA in females as in males (Aronson et al., 1954; Rudkin, 1964) are convincing arguments against the possibilities of earlier replication or unequal replication of the male X (Roehrdanz and Lucchesi, 1977). Dosage compensation therefore acts at the level of transcription, to equalise the amount of X-linked RNA transcribed between the sexes.

Transcription of male X chromosome is doubled

Although the simplest mechanism for the equalisation between sexes is an up-regulation of the male X chromosomal genes, the same result could be achieved through a down-regulation of the female X chromosomes. However, both female X chromosomes are active, as occasional observations of unpaired female X chromosomes in autoradiography experiments reveal that both chromosomes incorporate equal levels of tritiated uridine (Mukherjee and Beermann, 1965). Furthermore, the level of uridine incorporation to the X chromosomes by either sex is also equal to that incorporated by a pair of autosomes, suggesting that both female X chromosomes are transcribed at the expected level, and that two copies of all genes provide the required level of gene expression.

The theory of up-regulation in males is also supported by the identification of male-specific factors that are implicated in dosage compensation (Section 1.1.2). Even the simple cytological observation that the width of the male X chromosome appears equal to that of the paired female X chromosomes (Offermann, 1936) now suggests a “hyperinflation” of the male X (Lakhotia and Mukherjee, 1969), given that no extra chromosomal material is present in the male X (Aronson et al., 1954; Rudkin, 1964).

A modified male-specific model is also possible. Large, aneuploid deficiencies in chromosomal dose often lead to up-regulation of other genomic regions, presumably due to the loss of more negative than positive gene regulators (Birchler et al., 2001). As the single X chromosome in males resembles aneuploidy, we might expect all other autosomes to increase in expression, and the main mechanism of dosage compensation would therefore be to *reduce the male autosomal* level of expression, to restore the balance (Birchler et al., 2003). However,

significant reductions in the amounts of known compensation machinery through the use of RNAi reduce the level of mRNAs produced from X-linked genes, but do not alter levels from autosomes (Hamada et al., 2005; Straub et al., 2005*a*). The change in the level of expression from specific X-linked genes is clear even when normalised to specific genomic DNA, thus removing the doubts present with normalisation to autosomal transcripts. Perhaps the gradual evolution of the sex chromosomes in *Drosophila* did not present equivalent biological dosage problems to those observed in spontaneous aneuploidy (Straub et al., 2005*a*). The simplest explanation remains the most likely: dosage compensation in *Drosophila* is achieved through a male-specific up-regulation of transcription from the single X chromosome.

1.1.2 Known factors linked to compensation

Male-specific lethal proteins

Five proteins that are essential for male survival have been linked to dosage compensation, and are collectively termed the *male-specific lethal* or MSL proteins. Extensive screens of ethyl methane sulphonate-treated chromosomes identified male-specific lethal alleles of the genes *msl-1*, *msl-2*, and *maleless* (*mle*) on the second chromosome (Belote and Lucchesi, 1980*b*); *msl-3* on the third (Lucchesi et al., 1982); and *males-absent on the first* (*mof*) on the X chromosome (Hilfiker et al., 1997). The male-specific lethality caused by loss of these genes does not necessarily mean that they have no general, or even female-specific (but less critical) roles.

Basic characteristics of the five proteins can be outlined, following the molecular cloning of the respective genes (Kuroda et al., 1991; Palmer et al., 1993; Bashaw and Baker, 1995; Gorman et al., 1995; Kelley et al., 1995; Zhou et al., 1995; Hilfiker et al., 1997). The roles of MSL1 and MSL2 appear to be largely structural, although MSL2 contains a RING finger domain that may have E3-ubiquitin ligase activity (Bashaw and Baker, 1995; Kelley et al., 1995; Zhou et al., 1995; Fang et al., 2003). MSL3 contains a chromodomain capable of binding methylated nucleosomes (Buscaino et al., 2006; Larschan et al., 2007; Sural et al., 2008). MOF is a histone acetyl transferase (HAT), but also has a zinc-finger motif important for nucleosome binding, and a chromo-barrel domain known to bind RNA (Hilfiker et al., 1997; Smith et al., 2000; Akhtar and Becker, 2000; Akhtar et al., 2000; Akhtar and Becker, 2001).

Proteins with male-specific roles are good candidates for the effectors of dosage compensation. Indeed, MSL1, MSL2, and MLE are necessary for the unusually wide male X chromosome, and MLE for the increased tritiated uridine uptake by the male X, and the doubling of *Sgs-4* mRNA (Gorman et al., 1993; Belote and Lucchesi, 1980*a*; Breen and Lucchesi, 1986). Likewise, mutations in MSL1, MSL2, and MLE eliminate the compensation of FUM, G6PD, and 6GPD (Belote and Lucchesi, 1980*a*). Perhaps most convincingly, fluorescent antibodies against all five MSL proteins reveal hundreds of binding sites specifically on the male X chro-

mosome, although both MLE and MOF appear to bind autosomes and female X chromosomes to a lesser degree (Kuroda et al., 1991; Palmer et al., 1993; Bashaw and Baker, 1995; Kelley et al., 1995; Zhou et al., 1995; Gorman et al., 1995; Gu et al., 1998). Taken together, these observations suggest the involvement of the MSL proteins in dosage compensation.

A male-specific chromatin-modifying complex

The sub-nuclear localisation of the MSL proteins to the male X chromosome suggests synergy of action, or even physical interactions between the proteins. Indeed, dual labelling of the male X chromosome by fluorescent antibodies against pairs of MSLs shows that they bind to the same locations (Bone et al., 1994; Bashaw and Baker, 1995; Gorman et al., 1995). Furthermore, the wild-type binding pattern of any one protein depends on the presence of all other MSLs (Gorman et al., 1993; Palmer et al., 1994; Gorman et al., 1995; Bashaw and Baker, 1995; Hilfiker et al., 1997; Gu et al., 1998; Buscaino et al., 2003). As amounts of protein also seem interdependent (Gorman et al., 1995; Chang and Kuroda, 1998; Buscaino et al., 2003), the stabilising interactions of a protein complex seem likely.

A basic description of the interactions within the complex is possible. MSL1, 2, 3, and MOF co-immunoprecipitate, but MLE is more loosely associated (Kelley et al., 1995; Copps et al., 1998; Smith et al., 2000). A subset of X chromosomal sites is bound by MSL1 or MSL2 in the absence of MSL3, MLE, or MOF, suggesting the possibility of partial complexes consisting of just MSL1 and MSL2 (Palmer et al., 1994; Bashaw and Baker, 1995; Hilfiker et al., 1997; Lyman et al., 1997). Nevertheless, the structural core of the complex is MSL1, as direct, independent, *in vitro* associations are observed with MSL2, MSL3, and MOF, yet not amongst most other MSL proteins (Copps et al., 1998; Scott et al., 2000; Morales et al., 2004; Li et al., 2005). Recombinant MOF and MSL3 also bind weakly *in vitro*, but MSL1 and MSL3 together as a pair are necessary for MOF association, and specificity of MOF acetyl-transferase activity (Buscaino et al., 2003; Morales et al., 2004).

An RNA component to the complex is suggested by the observation that MSL3, MLE and MOF (but not MSL1 or MSL2) do not localise to the X chromosome upon treatment with RNase (Richter et al., 1996; Akhtar et al., 2000; Buscaino et al., 2003). Accordingly, the two X-linked, non-coding RNAs *roX1* and *roX2* (*RNA on the X*) have been identified in a screen for sex-specific expression (Han et al., 1996; Amrein and Axel, 1997). The *roX* RNAs are implicated in dosage compensation, as they are only expressed in males, specifically coat the male X chromosome, co-localise to the same sites as MSL1, and can be detected by RT-PCR in immuno-precipitates of MSL1, MSL3, MLE, and MOF (Amrein and Axel, 1997; Meller et al., 1997; Franke and Baker, 1999; Meller et al., 2000; Akhtar et al., 2000; Smith et al., 2000; Buscaino et al., 2003). The known RNA helicase and dsRNA-binding abilities of MLE (Gibson and Thompson, 1994; Lee et al., 1997) make it the obvious candidate to bind *roX* RNA,

and indeed neither RNA is present if MLE carries a mutation in the helicase domain (Gu et al., 2000). Nevertheless, the chromo-barrel domains of recombinant MSL3 or MOF both bind RNA non-specifically *in vitro*, and a specific *in vivo* interaction between *roX2* and MOF seems likely, as the RNA is not co-immunoprecipitated when MOF is mutant in the chromo-barrel domain (Akhtar et al., 2000).

The importance of the RNAs in dosage compensation is clear, as a double *roX1* and *roX2* mutation results in a characteristic drop in male viability, and rescues the lethality caused by inappropriate compensation when dosage compensation is induced in females (as discussed below) (Franke and Baker, 1999; Meller and Rattner, 2002). Despite large sequence and length differences, the two RNAs appear to be functionally redundant, as flies with either gene mutated display no phenotype whatsoever (Meller et al., 1997; Franke and Baker, 1999; Meller and Rattner, 2002). Given that MSL1, MSL2, and MLE do not localise to the X chromosome in double *roX* mutants (Franke and Baker, 1999), the main role of the RNAs appears to be in the targeting of the MSL complex. This is elaborated on in Section 1.1.4.

The most well-defined role of the MSL complex is to acetylate the histones of the male X chromosome. Antibodies specific for the isoform of histone 4 acetylated at lysine 16 (H4K16Ac) bind at least three times higher to the male X than to the autosomes or female X chromosome (Turner et al., 1992). The acetylation is linked to dosage compensation, as the isoform co-localises with MSL1 and MLE, and the male X enrichment disappears with mutations in any of the MSLs (Bone et al., 1994; Hilfiker et al., 1997). Radio-labelled acetyl CoA is incorporated into free histones or nucleosomes when incubated with recombinant MOF, but not when the acetyl CoA binding site in MOF is mutated (Akhtar and Becker, 2000). Furthermore, a partially-purified MSL complex mediates the same specific histone acetylation, but not when MOF is mutated as before, implying that MOF is the only histone acetyl-transferase associated with the complex (Smith et al., 2000). Thus the ribonucleoprotein MSL complex specifically binds to and modifies the chromatin of the *Drosophila* male X chromosome.

Associated proteins are difficult to detect

Additional factors involved in any aspect of dosage compensation have a more general role in both sexes, thus cannot be identified by male-specific lethality. Association with dosage compensation can sometimes be inferred if hypomorphic alleles affect male more than female viability, but this method relies on the fortuitous generation of these alleles. Other factors can instead be identified through their effects on the male X chromosome morphology or gene expression. Physical interactions with, or proximity to, known MSL components also implies guilt by association.

The best characterised factor aside from the MSL proteins is the JIL-1 tandem kinase. Null alleles of *JIL-1* are recessive lethal, but hypomorphs have a lower male than female viability,

and a condensed chromatin structure most striking on the male X chromosome (Wang et al., 2001; Deng et al., 2005*a*). Indeed JIL-1 binds chromatin, and is enriched on the male X chromosome, where it co-localises and physically interacts with MSL proteins (Jin et al., 1999, 2000). The MSL complex appears to recruit JIL-1, as X-bound JIL-1 requires the MSL proteins, but not *vice versa*. JIL-1 phosphorylates both itself and histone 3 at serine 10 (H3S10), an epigenetic mark of active chromatin (Jin et al., 1999; Wang et al., 2001). Reduced levels of JIL-1 increase the amount of histone 3 di-methylated at lysine 9, and associated HP1, on the male X chromosome, signaling an increased heterochromatic state (Zhang et al., 2006). Although the effect on X-linked transcription levels remains to be determined, JIL-1 is necessary for dosage compensation of *white*-mediated eye colour (Lerach et al., 2005). Thus recruitment of JIL-1 by the MSL complex causes enrichment of a second chromatin modification, beyond H4K16 acetylation, on the male X chromosome, and is likely a vital part of the dosage compensation mechanism.

Similar lines of evidence also link the DNA supercoiling factor (SCF) with dosage compensation (Furuhashi et al., 2006). Like JIL-1, SCF binds chromatin and is enriched on the male X chromosome. SCF on the male X chromosome co-localises with, and is dependent upon, MSL proteins. Reduction of SCF by RNAi leaves MSL complex on the X chromosome, but abolishes the associated male-specific transcriptional increases, and reduces male viability. These results suggest that MSL complex also recruits SCF to aid in the up-regulation, although the consequent effect is less obvious than for the JIL-1 kinase.

Indirect associations with several modifiers of chromatin are also apparent. Hypomorphic alleles of the GAGA factor affect males more severely than females, but do not co-localise on the male X chromosome (Greenberg et al., 2004). Mutated components of the nucleosome remodeling factor (NURF) complex give the polytenised male X chromosome a decondensed appearance, and require MSL proteins to do so (Deuring et al., 2000; Badenhorst et al., 2002; Corona et al., 2002; Bai et al., 2007). Similarly, mutated or reduced amounts of heterochromatin-associated proteins SU(VAR)3-7 and HP1 de-condense all chromatin, most noticeably that of the male X chromosome, and are more lethal to males (Seum et al., 2002; Delattre et al., 2004; Liu et al., 2005). As discussed below (Section 1.1.5), a balance between repressive and stimulatory effectors of the male X chromosome may be necessary for correct levels of expression.

As a chromatin modifier, the MSL complex will necessarily interact with many other proteins. Mass spectrometry of MSL3 or MOF immuno-precipitates identifies several of these immediate contacts (Mendjan et al., 2006). In high-stringency conditions (such that MLE, JIL-1 and SCF are not retained), several components of nuclear pores and regulatory exosomes are robustly detected. Reduction of two nuclear pore components (Mtor and Nup153) removes MSL proteins from the male X chromosome, and abolishes dosage compensation. This provides

evidence for a structural aspect to dosage compensation; separable from chromatin regulation.

1.1.3 Sex-specific compensation

MSL2 is male-specific

Given that dosage compensation is a male-specific up-regulation of X-linked gene transcription, some molecular control must exist to activate the mechanism in males. Presence of a Y chromosome does not lead to dosage compensation, as males or females with an incorrect number of Y chromosomes are not abnormal (Belote and Lucchesi, 1980*b*; Seecof et al., 1969). Dosage compensation also acts independently from the *doublesex* and *transformer* genes, which are involved in sex determination (Belote and Lucchesi, 1980*b*; Fukunaga et al., 1975).

However, the MSL2 protein is absent in females, exhibits differential splicing between the sexes, and contains binding sites for the female-specific Sex-Lethal (SXL) protein in the 5′ and 3′ untranslated regions (UTR) of the *msl-2* transcript (Bashaw and Baker, 1995; Kelley et al., 1995; Zhou et al., 1995). SXL is an RNA-binding protein that sex-specifically splices its own transcript and that of *transformer*, ultimately controlling the sex determination pathway (Sosnowski et al., 1989; Bell et al., 1991). SXL also binds to and represses translation of the *msl-2* transcript in females, preventing MSL complex formation (Bashaw and Baker, 1997; Kelley et al., 1997). Sucrose gradient analysis of the assembly of the ribosome complex on *msl-2* RNA reveals that the presence of SXL inhibits association of the 40S ribosomal subunit, and therefore acts to block at least the initiation of translation (Gebauer et al., 2003). Ectopic expression of an *msl-2* gene lacking the SXL binding sites leads to MSL complex assembly in females, and a corresponding 80 % decrease in female viability (Kelley et al., 1995).

MSL2 induction leads to dosage compensation

MSL2 is not the only component expressed in a sex-specific manner. Amounts of both MSL1 and MSL3 are lower in females, *roX1* is not expressed after embryogenesis, and *roX2* transcripts are completely absent (Gorman et al., 1995; Amrein and Axel, 1997; Meller et al., 1997; Meller, 2003). Although the translation of MSL1 in females is also weakly repressed by SXL, the main reason for the lower protein levels is the lack of stabilisation by MSL2 (Chang and Kuroda, 1998). Induced MSL2 expression in females increases at least MSL1 levels, and depletion of MSL2 by RNAi in males leads to a decrease in both MSL1 and MSL3 levels (Kelley et al., 1995; Buscaino et al., 2003). Over-expression of both MSL1 and MSL2 in females results

in full dosage compensation, and corresponding 100 % female lethality (Chang and Kuroda, 1998).

Expression of the *roX* RNAs is also directly regulated by the MSL proteins. Levels of *roX1* RNA drop heavily with RNAi directed against MSL2, and no *roX* transcripts are detected when any of the MSLs are absent (Straub et al., 2005*b*; Amrein and Axel, 1997; Meller et al., 1997). MSL1 can be detected binding to translocations of the *roX* genes to the autosomes, and *roX2* DNA can be detected in immunoprecipitates of either MSL1 or MSL2 (Kelley et al., 1999; Smith et al., 2001). A male-specific DNaseI hypersensitive site (DHS) is present in each of the *roX* genes, and small regions spanning either DHS are sufficient to bind MSL1 at ectopic locations (Kageyama et al., 2001; Park et al., 2003). Short stretches of residues, including three that resemble binding sites for GAGA factor, are conserved between both DHS regions in several related *Drosophila* species, and mutations in each site reduce the MSL-binding ability (Park et al., 2003). Deletions of either DHS abolish transgenic *roX* RNA expression, and addition of the DHS to a *roX1* reporter gene provides male-specific up-regulation that is dependent on the MSL proteins (Bai et al., 2004). Accordingly, induction of MSL2 in females gives normal *roX* RNA levels (Amrein and Axel, 1997).

The ability of the MSLs to regulate *roX* expression may be distinct from the complex's normal dosage compensation function. In contrast to the known requirements for dosage compensation, *roX* regulation does not require presence of the *roX* RNAs, nor a functional RING finger in MSL2 (Bai et al., 2004; Rattner and Meller, 2004). Furthermore, *roX1* induction at the endogenous location does not appear to require MSL3, MLE, or MOF, nor even the DHS (Rattner and Meller, 2004). As the DHS can attract MSL complex, this surprising result may indicate an otherwise high local concentration of MSL proteins at the *roX1* location. In any case, the importance of MSL2 is clear. Given that MSL2 also stabilises the levels of the other MSL proteins, the key sex-specific control of dosage compensation is thus the regulation of MSL2.

1.1.4 Targeting the male X chromosome

Dosage compensation involves the co-ordinate regulation of hundreds of male genes on a specific chromosome, so some means of identifying the X chromosome must exist. Translocations of large fragments of the X to autosomes retain compensation of their corresponding genes, and the characteristic puffed chromatin morphology indicitive of hyper-transcription (Seecof et al., 1969; Tobler et al., 1971; Oh et al., 2004). Likewise, the activity levels of genes on autosome to X translocations remain not compensated (Lakhotia and Mukherjee, 1969; Roehrdanz and Lucchesi, 1977). This reinforces the natural assumption that gene-specific elements mark X-linked genes for compensation. However, transpositions of individual genes do not fully support this model. Transpositions of some genes from the X to an autosome retain compensation,

but autosomal genes inserted on the X often acquire a level of compensation (Krumm et al., 1985; Scholnick et al., 1983; Spradling and Rubin, 1983). This complicates the situation by further implying that the X chromosome environment is in some way special. Indeed, limiting the compensation to X-linked genes appears to involve several levels of control. The problem can be best addressed by identifying how the MSL proteins, the main effectors of dosage compensation, are targeted to the X chromosome.

The roX RNAs aid MSL complex targeting

Although deletion of both *roX* genes is generally not tolerated by males, a varying number of developmentally-delayed adults can survive, depending on the severity of the *roX* mutation (Meller and Rattner, 2002; Deng et al., 2005*b*). As the severity of *roX* deletion increases, the number of X chromosome sites bound by the MSLs decreases, and the number of autosomal sites increases, to the point where most MSL complex binds to the centromeric heterochromatin. MSL binding to *roX* DHS sites on autosomes is also abolished in the absence of endogenous *roX* genes (Park et al., 2003). Note that in both cases the targeting ability of *roX* RNA is unrelated to the chromosomal location of the *roX* gene, as correct MSL binding is restored upon expression of transgenic *roX* RNA from another chromosome (Park et al., 2003). Structurally, the same region of MSL2 that is necessary to incorporate *roX* RNAs is also needed to correctly locate the complex to the X chromosome (Li et al., 2008). Inclusion of the *roX* RNA in the complex therefore increases the binding specificity of the MSL proteins, to aid correct compensation of the X chromosome.

The roX genes as chromosome markers

As the genes for the *roX* RNAs are themselves on the X chromosome, and the regulation of these genes necessarily involves MSL binding (Section 1.1.3), they may serve as a molecular marker of the X chromosome. Once located to the X chromosome, the complex would then "spread" to the rest of the genes to be compensated. Indeed, the MSL complex is sometimes observed to spread up to 1 Mb into the chromatin surrounding an autosomal *roX* transgene, and can increase the transcription of a coupled reporter gene (Kelley et al., 1999; Meller et al., 2000; Henry et al., 2001). Spreading can even be observed from a multimerised nine-copy tandem repeat of the *roX1* DHS (Kageyama et al., 2001). While spreading may contribute to targeting, *roX* gene localisation on the X is not necessary, as an autosomal *roX* transgene can rescue the lethality of a double *roX* deletion, and correctly target the MSL complex to the X (Meller and Rattner, 2002; Park et al., 2002).

Much like the gene dosage problem that is being solved, the targeting of dosage compensation to the X appears to depend upon a delicate balance between *roX* RNAs and the MSL

proteins themselves. Over-expressed MLE or MOF bind to all chromosomes, indicating an inherent chromatin-binding ability, whereas over-expressed MSL1 and MSL2 together bind in excess to the X, giving a shortened, hyper-inflated chromosome (Richter et al., 1996; Gu et al., 2000; Oh et al., 2003). The affinity of MSL1 and MSL2 for the X chromosome is so great that the over-expressed pair can even partly rescue the lethality of a double *roX* mutant, restore the correct chromatin appearance, and return mis-localised MSL proteins to the X (Oh et al., 2003; Deng et al., 2005*b*). However, under normal circumstances *roX* RNAs appear to compete for a limited pool of MSL proteins, as the extent of spreading from a *roX* transgene is inversely proportional to the number of *roX* genes (Park et al., 2002). Spreading is only seen from a *roX* gene when transcribed, and is most extensive with high levels of MSLs or a reduced number of *roX* genes (Park et al., 2002; Oh et al., 2003; Demakova et al., 2003; Bai et al., 2004). Abnormally high spreading may thus reflect a locally high concentration of assembling complexes on the nascent RNA, and under a normal MSL / *roX* ratio the contribution of this mechanism to targeting is likely low.

A hierarchy of sites with affinity for the MSL complex

Although deletion of MSL1 or MSL2 removes MSL complex from the X under any circumstances, partial complexes lacking MSL3, MLE, or MOF bind a subset of sites in females with a modified MSL2 that allows complex formation (Palmer et al., 1994; Bashaw and Baker, 1995; Gu et al., 1998). These "high affinity" sites number about 60 to 70, differ slightly depending on which MSL component is lacking, and appear to overlap only moderately with the subset of sites bound with limiting *roX* RNA (Lyman et al., 1997; Demakova et al., 2003; Deng et al., 2005*b*; Dahlsveen et al., 2006). Strong MSL binding is also seen at the *roX* genes in these females, but as double *roX* mutations are lethal, and repeated searches for *roX*-like transcripts identify no new candidates, the other high affinity sites are unlikely to resemble the *roX* genes (Meller and Rattner, 2002; Deng et al., 2005*b*; Fujii and Amrein, 2002; Oh et al., 2003).

Indeed, the best characterised high affinity site (at cytological position 18D10[1]) produces no transcript, and has no GAGA-like sites (Oh et al., 2004). Similarities can nevertheless be seen to the MSL binding site in the *roX* genes, as the 18D10 site also spans a DHS that is more pronounced in males, and this DHS can attract MSL complex binding to an autosomal location, especially if multimerised. The MSL complex can occasionally spread into neighbouring chromatin from both the larger 18D10 fragment and the four-copy multimer of the DHS, and the spreading increases with reductions in *roX* genes, indicating that the high affinity sites also compete for the limiting MSL proteins. Several of the highest affinity sites (including that at 18D10) are detected in immuno-precipitates of MSL1, but the ability of each to attract MSL appears to differ (Dahlsveen et al., 2006). Systematic deletion analysis of the strongest sites

[1] recently correctly located at 18D11

shows the significance of general DHS elements containing degenerate sequences, including a GA nucleotide pair, for MSL attraction (Gilfillan et al., 2007).

It is tempting to conclude that both *roX* genes and the high affinity sites serve to mark the X chromosome, and that the MSL proteins once bound can then diffuse to the remaining X-linked chromatin, but the situation is yet more subtle than that. The large translocations of X segments to autosomes always retain their normal MSL binding pattern, even if they encompass no obvious high affinity nor *roX* site (Fagegaltier and Baker, 2004; Oh et al., 2004). The MSL complex is also unable to spread into surrounding autosomal chromatin, nor onto an autosomal fragment translocated to the X, even when the foreign chromatin is immediately adjacent to a *roX* gene (Fagegaltier and Baker, 2004). Furthermore, the number of X sites bound by the complex can also be controlled by limiting the *amount* of MSL complex (Demakova et al., 2003). Severe *msl-2* alleles allow binding to only four sites, increasingly functional alleles bind more sites until a wild-type pattern is reached, and over-expressed MSL1 and MSL2 bind additional autosomal and heterochromatic sites. Thus a hierarchy of sites exist, with differing affinities for the MSL complex. Each site appears to individually recruit the complex, with no obvious "spreading" linearly from high to low sites. Low amounts of complex, or partial complexes lacking MSL3, MLE, or MOF, are sufficient to bind the sites of highest affinity, but the full binding pattern is only achieved with correct levels of all MSL proteins.

Chromosomal studies of MSL targeting are inherently low resolution, with each fluorescent band representing many kilobases of chromatin-bound DNA. To accurately identify many high affinity sites, immuno-precipitates of MSL proteins in limited binding patterns can be passed over microarrays representing the X chromosome (ChIP on chip) (Alekseyenko et al., 2008; Straub et al., 2008). Many such sites possess a defined MSL recognition element (MRE) of 20–30 bp, encompassing a nearly exact eleven-copy repeat of GA di-nucleotides. High-throughput sequencing of MSL3 immuno-precipitates (ChIP-seq) bound in a wild-type male pattern confirms that between 150 and 300 of the sites with highest affinity for MSL complex encompass an MRE (Alekseyenko et al., 2008). The MREs in the high affinity sites lie predominantly in intronic or non-coding sequences near genes. However, although very strongly bound at MRE elements, MSL proteins also bind robustly over tens of kilobases of surrounding DNA in a high affinity site. An additional mechanism must be responsible for the final stage of MSL attraction to the X chromosome.

Targeting to active chromatin

As with the identification of high affinity sites, the nature of all bound sites can be resolved by microarray analysis of MSL immuno-precipitates from wild-type male tissues (Alekseyenko et al., 2006; Gilfillan et al., 2006; Legube et al., 2006). When at full occupancy, the MSL proteins bind over 700 sites. In contrast to the location of high affinity sites, the complete binding

pattern almost exclusively resides within genes rather than intergenic sequences, and within exons not introns, and is clearly enriched for coding sequences over un-translated regions. On binding profiles of all genes scaled by gene length, MSL proteins peak towards the 3′ end of target genes. About 90 % of the MSL target genes are expressed, and 88 % of the target sites clearly coincide with RNA polymerase II sites. It appears that the act of transcription can attract the MSL complex, as MSL binding is also seen at GAL4-induced transcription of random X chromosomal sequences (Sass et al., 2003). Furthermore, transcription of the X-linked *mof* gene at autosomal insertions attracts MSL complex even when different promoters are used, or an antisense transcript of the gene is produced instead (Kind and Akhtar, 2007). Thus the MSL complex is attracted to active chromatin.

The ability to identify active chromatin may be provided by an affinity of MSL3 for an associated chromatin epigenetic mark. Methylation of lysine 36 of histone 3 (H3K36) is generally associated with active chromatin (Saunders et al., 2006). Whilst di-methylation of H3K36 is enriched at active promoters, tri-methylation (H3K36me3) peaks towards the 3′ end of the CDS (Pokholok et al., 2005; Rao et al., 2005; Barski et al., 2007; Mikkelsen et al., 2007; Bell et al., 2008). This is reminiscent of the 3′ localisation of the *Drosophila* MSL proteins, which partially co-localise with H3K36me3 on polytene X chromosomes (Bell et al., 2008). At the high resolution of ChIP on chip, MSL3 binding co-localises with H3K36me3 on the male X chromosome, and binds autosomal genes surrounding a *roX2* translocation proportional to H3K36me3 enrichment (Larschan et al., 2007). Recombinant MSL3 binds nucleosomes containing H3K36me3, and binding is impaired by deletions in its chromodomain (Larschan et al., 2007; Buscaino et al., 2006; Sural et al., 2008). Depletion of the relevant histone methylase Set2 (alias KMT3) by RNAi reduces MSL binding to target genes, and impairs consequent H4K16 acetylation and dosage compensation (Larschan et al., 2007; Bell et al., 2008). Indeed, tri-methylation of H3K36 is a better predictor of MSL binding than gene transcription (Larschan et al., 2007).

Targeting of the MSL complex to the X chromosome is thus best described as a two-step process. Initial attraction of the complex to high affinity sites is likely a complex combination of sequence-specific MREs, nucleosome occupancy, and chromatin marks. Subsequent dispersal of the MSLs to all sites requires the recognition of active chromatin marks including H3K36me3 by the chromodomain of MSL3. This distinction in targeting steps is clearly seen in immunoprecipitates of MSL3, and MSL3 mutated in the chromodomain (Sural et al., 2008). Wild-type MSL3 binds around 1400 X-linked genes in 700 clusters, but MSL3 lacking the chromodomain binds on average only one gene per cluster. The direct effect on dosage compensation is a higher dependancy on MSL2 for compensation of genes close to a MRE, but more reliance on MSL3 (and MOF) for genes further away (Straub et al., 2008). Taken together with the demonstrated importance of transcription, it appears that the presence of active genes themselves leads to the final pattern of MSL occupancy.

Establishment of MSL binding

Although the nature of the MSL binding sites is mostly determined, the dynamics surrounding the initiation of MSL binding are relatively poorly understood. MSL proteins are expressed early in the developing embryo, and bound complex is detected in all somatic cells from the blastoderm stage onwards (about 2.5 hours after egg laying) (Rastelli et al., 1995; Franke et al., 1996). Maternal MSL1, MSL3, and MLE are provided, but complex cannot form until MSL2 expression after fertilisation. Bound complex requires prior *roX* expression, but differences between the two *roX* genes mean the very early complex may differ slightly (Meller, 2003). *roX1* is expressed at about 2 hours after egg laying; *roX2* at around 6 hours. Thus bound complex in the first few hours may differ slightly from that in later stages. The earliest available evidence of dosage compensation is microarray analysis of late (12–14 hour) embryo transcripts (Straub et al., 2005*a*).

The ability of the complex to disassociate with the X chromosome is also unclear. Bound sites on polytene chromosomes are mostly identical across developmental time and tissue type, but some tissue- and stage-specificity is also evident (Sass et al., 2003; Kotlikova et al., 2006). However, fluorescent signals of bound MSL2 do not recover in 20 minutes after photo-bleaching, suggesting a very stable association with the X chromosome (Straub et al., 2005*b*). Chromosome-wide ChIP on chip results confirm that the complex mostly binds similar genes in different tissues, with the differentially-bound genes reflecting tissue-specific transcripts (Alekseyenko et al., 2006; Legube et al., 2006). Furthermore, the genes that remain bound in larval salivary glands are mostly a subset of those bound in early (4–6 hour) embryos, implying that MSL binding is mostly established in embryogenesis, then maintained through development.

1.1.5 Potential mechanisms

Binding of the MSL complex does not necessarily lead to transcriptional up-regulation. Although undoubtedly the MSL complex and related proteins are responsible for the transcriptional increase, there are several observations that indicate MSL binding and up-regulation are separable events. The amount of bound MSL by a gene correlates to the transcriptional change when MSL is depleted, confirming the MSL proteins as the main agents (Gilfillan et al., 2006). However, the transcription of many un-bound genes are also affected by MSL depletion, and can score as 'compensated' in male to female transcript comparisons (Alekseyenko et al., 2006; Legube et al., 2006). The complex can also bind without up-regulation, as a minority of the high affinity sites that attract the complex to an autosomal location cannot also increase a corresponding reporter gene's activity (Alekseyenko et al., 2008). What, then, is the consequence of MSL complex binding? How is transcription increased?

MSL binding appears to indirectly affect transcription by modifying chromatin. The com-

plex does not increase polymerase initiation like a transcription factor would, but must rather facilitate transcription once begun. Although MSL binding correlates with transcription and active chromatin, the ChIP on chip results for MSL complex and RNA Polymerase II do not exactly overlap (Gilfillan et al., 2006). Neither do MSL proteins exactly co-localise with polymerase nor initiation factors on polytene chromosomes (Kotlikova et al., 2006; Legube et al., 2006). In fact the complex also binds several tRNA genes transcribed by RNA Polymerase III, and known to be compensated (Gilfillan et al., 2006; Birchler et al., 1982). There are no reports of MSL enrichment in promoter regions, no interactions with RNA Polymerase II or any general transcription factors, and no genetic or cytogenetic indication that any might be involved. Instead, MSL targets are predominantly 3′ coding sequences, and many of the proteins involved modify chromatin. As the acetylation of H4K16 by MOF also peaks towards the 3′ of coding sequences (Smith et al., 2001; Kind et al., 2008), it seems more likely that the complex facilitates increased transcription elongation through chromatin modification.

Positive effects on transcription counteract repressive actions. As MSL complex de-condenses chromatin and increases transcription, it necessarily antagonises factors that negatively affect gene activity. In some cases this effect can be specifically observed. Insertions of a *roX1* transgene marked by mini-*white* sometimes show position effect variegation of male eye colour that is affected by levels of heterochromatin-associated factors (Kelley and Kuroda, 2003). In these cases, female eyes are very pale or even white, implying specific repression of the mini-*white* gene. This repression is relieved in males or females expressing *msl2* in the eye, due to *roX1*-mediated attraction of MSL complex, and up-regulation of mini-*white*. More specifically, *in vitro* transcription can be repressed by assembly of nucleosomes on the template, but relieved by histone acetylation from recombinant MOF (Akhtar and Becker, 2000). Relief of repression is also evident at the *roX1* locus itself, which MSL complex positively regulates (Section1.1.3). Basal expression of *roX1* is constitutively suppressed in both sexes via the same DHS necessary for MSL up-regulation (Bai et al., 2004). The repression is dependent on the main component of the NURF complex, demonstrating direct antagonistic effects of the two chromatin-modifying complexes (Bai et al., 2007). If relief of repression is a common theme to MSL action, it may be that the apparent two-fold transcriptional increase is actually a product of a balance between repression and stimulation.

The MSL complex may also have a global effect. By affecting the chromosome as a whole, the expression of individual genes could be altered without direct binding by the complex. The male X chromosome tends to locate towards the periphery of the nucleus in interphase cells (Rastelli et al., 1995; Rastelli and Kuroda, 1998), and directly interacts with nuclear pore components (Mendjan et al., 2006). As nuclear pore mutants de-localise MSL proteins from the X chromosome, their effect on MSL targeting necessarily also impacts dosage compensation. However, the association with the pores may also help euchromatic regions of the male

X chromosome form sub-nuclear domains. Concentration of the genes to be compensated in a particular area would be an efficient use of MSL complex, JIL-1, SCF, and the corresponding chromatin modifications. The possibility of global regulation of the chromosome would explain the above discrepancy between MSL binding and gene regulation. Dosage compensation may thus be 'action at a distance'; not just chromosomal in effect, but in approach.

1.1.6 Summary

Dosage compensation is the equilisation of gene expression from unequal doses of genes. *Drosophila* males up-regulate transcription from their single X chromosome to match the level produced from XX females. Five *male-specific lethal* (MSL) genes are absolutely required in males, and encode the main agents of the up-regulation. At least these five proteins, together with either or both of two non-coding *roX* RNAs, form the MSL complex, which binds to and modifies the chromatin of the male X chromosome. Female-specific translational repression of a key MSL protein (MSL2), and subsequent loss of auto-regulatory and stabilising effects, limits the MSL complex to males.

Attraction of the MSL complex to the X chromosome is a complicated and finely-balanced process. A hierarchy of sites with differing affinities for the MSL complex compete for a limiting pool of MSL proteins. The *roX* loci are two of the most potent attractors, likely due to MSL regulation of *roX* expression, and complex assembly on nascent *roX* transcripts. A further 70–300 high affinity sites share a common 21 bp MSL recognition element (MRE). The differing affinities of each site are probably due to differing strengths of MRE consensus, combined with nucleosome occupancy and surrounding chromatin state. Attraction to the X chromosome may be a two-step process: first to the *roX* genes and (predominantly non-coding) high affinity sites, then to surrounding active genes. Indeed most MSL sites are within active genes. The establishment of chromosome binding and the dynamics of the association remain relatively unknown.

The mechanism by which gene dose is compensated is by modifying chromatin. The MSL complex and co-factors collectively alter X-linked euchromatin to create a de-condensed state that facilitates increased transcription. Rather than enhancing polymerase initiation at promoters, MSL proteins accumulate at $3'$ coding sequences, where they likely affect transcription elongation or RNA processing. The characterised activities of histone acetylation and phosphorylation act at least in some cases to relieve repression, demonstrating that final two-fold transcriptional changes may be a product of antagonistic effects. Such chromatin alteration could also have a long-range effect on transcription. By creating sub-nuclear domains of de-condensed chromatin, the modifying factors can be efficiently shared to globally modify the entire chromosome. Thus the effect on transcription is indirect: an average two-fold increase provided by a combination of local and global changes to chromatin structure.

1.2 Transvection

Homologous chromosomes of Dipteran species pair not just in meiosis but also through all stages of a mitotic cell cycle (McKee, 2004). Homologous pairing in somatic cells initiates during the maternal to zygotic transition of embryogenesis (Fung et al., 1998; Gemkow et al., 1998). The mechanism responsible for pairing remains unknown, but is likely a complex interaction of chromatin factors. A key contributing factor may be *Drosophila topoisomerase II*, as reduction of this enzyme by RNAi significantly reduces pairing in several cell lines (Williams et al., 2007).

The tight association of homologous chromosomes can affect gene expression levels, in a variety of ways known collectively as 'transvection'. *Drosophila* transvection effects were first observed in a study of *Ultrabithorax* (Lewis, 1954). Chromosomal re-arrangements with break points near *Ultrabithorax* affect the ability of certain alleles to complement others. In effect, such complementation only occurs if the alleles are homologously paired (Duncan, 2002; Ashburner et al., 2005).

Several transvection pairing effects appear to share a common mechanism. Gans (1953) first noticed that the recessive allele of *zeste* z^1 would produce yellow eyes only if flies carried two wild-type copies of *white*. This observation reflects a transvection effect on *white*, as both *white* alleles need to be homologously paired (Jack and Judd, 1979). The Zeste protein tends to form aggregates that suppress the eye enhancer of *white*, and the z^1 allele enhances this suppression (Bickel and Pirrotta, 1990; Qian et al., 1992). Suppression is probably mediated through local chromatin modifications, as Zeste binds the three *Polycomb* group proteins: *Psc*, *Scm*, and *E(z)* (Saurin et al., 2001). Pairing-dependent repression is not unique to *zeste* repression of *white*, and affects at least *cubitus-interruptus*, *Sex combs reduced*, and *brown* (Locke and Tartof, 1994; Pattatucci and Kaufman, 1991; Henikoff and Dreesen, 1989).

Pairing-dependent transcription enhancement, or relief of repression, similarly occurs at *decapentaplegic*, and *eyes absent* (Ashburner et al., 2005), but is most well-studied at *yellow*. As for *Ultrabithorax*, different alleles of *yellow* can complement each other, but only if homologously paired (Geyer et al., 1990). The *yellow* allele y^{82f29} lacks wing and body enhancers, but can be expressed in these tissues if the homologous allele (*e.g.* $y^{1\#8}$) supplies wild-type enhancer elements (Morris et al., 1998). Interestingly, the enhancers only work in *trans* when the *cis* promoter is inactive (Morris et al., 2004). The effect is also independent of the specific basal promoter, as wing and body enhancers can activate in *trans* a *yellow* transgene expressed from *white* or *eve* promoters (Lee and Wu, 2006).

Transvection is thus the synergistic action of enhancers or repressors between homologous copies of a gene (Duncan, 2002; Ashburner et al., 2005). The exact nature of gene regulation in each instance may vary considerably, and can include heterochromatic effects. Subtle interactions with *zeste* occur with at least *Ultrabithorax*, *decapentaplegic*, *eyes absent*, and *yellow*

(Duncan, 2002). To date, transvection events have mainly been observed by unusual interactions of different alleles, and no wide-scale survey of the phenomenon is available. How often transvection occurs, and how generally important it is, remain to be determined.

1.3 Basis of Experimental Design

1.3.1 Drosophila transformation and the ϕC31 system

The standard method to introduce recombinant DNA into the *Drosophila* germline relies on elements of the *Drosophila P* transposon (Spradling and Rubin, 1982; Rubin and Spradling, 1982). A plasmid containing the cassette of interest, flanked by the $5'$ and $3'$ termini of the *P* transposon, is injected into the early embryo. Co-injection with a 'helper' plasmid that transiently expresses *P* transposase, but is unable to transpose itself, mediates transfer of the cassette into a random genomic site. The level of gene expression from transgenic DNA is dependent on the site of insertion; a "position effect" of the surrounding chromatin (Spradling and Rubin, 1983). Such effects can be avoided by recombinase-mediated targeting to defined positions.

The large serine recombinase from the temperate *Streptomyces* phage ϕC31 mediates intramolecular recombination between defined sites both in *E. coli* and *in vitro* (Thorpe and Smith, 1998). The combination of the attachment site of the phage (*attP*) and the *Streptomyces ambofaciens* (*attB*) site, both imperfect inverted repeats, forms hybrid *attL* and *attR* sites that are not recognised by the recombinase (Thorpe et al., 2000). When supplied with helper recombinase, plasmids encoding *attB* can be targeted to *P* element-mediated insertions of *attP* in *Drosophila* (Groth et al., 2004). A collection of over 60 characterised *attP* "landing sites" has been generated by *mariner* transposase-mediated insertions, and the precise locations of each site determined (Bischof et al., 2007). With expression of ϕC31 recombinase from a genomic insertion, up to 55 % of fertile G_0 are transgenic, about five times more efficient than *P* element insertions (Bischof et al., 2007).

Although the ϕC31 recombinase requires as few as 50 or 51 bp of the *attP* and *attB* sites, respectively, for *in vitro* inter-plasmid transposition, and only 39 and 34 bp in *E. coli* cells, larger fragments of about 221 and 285 bp function well in mammalian cells and *Drosophila* (Thorpe et al., 2000; Groth et al., 2000, 2004).

1.3.2 Tetracycline-regulatable expression systems

The tetracycline repressor (TetR) of the *E. coli* tetracycline-resistance transposon Tn*10* binds as a homodimer to the tetracycline response element (TRE or tetO) only in the absense of tetracycline (Hillen et al., 1983; Hinrichs et al., 1994). When fused to the DNA binding and dimerisation domains from TetR, the VP16 transcriptional activation domain of the herpes simplex virus

can mediate transcription of genes downstream of a heptamer of tetO sequences in mammalian cells (Gossen and Bujard, 1992). The action of this tetracycline-controlled transcriptional activator (tTA) is also inhibited by tetracycline, or a chemical derivative, such as doxycycline. The mutation of four amino acids in the TetR region gives an 'reverse' activator (rtTA) with the exact opposite regulation: it only binds in the presence of tetracycline (Gossen et al., 1995).

With the use of appropriate regulatory sequences, both the tTA and rtTA function in *Drosophila* (Bello et al., 1998; Bieschke et al., 1998). Although modifications have been introduced to the rtTA that increase its induction capabilities and activity in *Drosophila* (Urlinger et al., 2000; Stebbins et al., 2001), similar modifications to tTA are not effective (Zhou et al., 2007).

1.3.3 DNA replication-related element (DRE)

The eight base-pair palindromic TATCGATA sequence is present in the promoters of genes for both *Drosophila* proliferating cell nuclear antigen (PCNA) and DNA polymerase α, and is thus termed DNA replication-related element (DRE) (Hirose et al., 1993). The consensus WATCGATW motif, with up to one nucleotide mismatch, is present in 26 % of a subset of EST sequences selected to likely represent regions around transcription start sites (Ohler et al., 2002).

Characterisation of a protein isolated by affinity purification with the DRE from *PCNA* revealed an 80 kDa trans-activating factor (DREF) capable of mediating transcription via associated DRE elements (Hirose et al., 1993, 1996). DREF, and some components of the nucleosome remodeling factor (NURF) complex (including the catalytic ATPase subunit ISWI), tightly associate with the TBP-related factor 2 (TRF2) to promote transcription (Hochheimer et al., 2002). The TRF2/DREF complex stimulates transcription of PCNA at a secondary site, independent of TBP/TAFIID-mediated transcription, and thus represents a core promoter selectivity factor. Induction of TRF2 causes up-regulation by microarray analysis of a subset of genes implicated mostly in DNA replication and cell proliferation.

Analysis of immunoprecipitates of MSL1 has implicated the DRE element in dosage compensation (Legube et al., 2006). The DRE motif is over-represented in the subset of cDNAs bound by MSL1, in comparison to those un-bound (28.2 versus 15.8 %, respectively), and is slightly enriched on the X chromosome. Thus the presence of a DRE element in a gene's promoter, and consequent induction of transcription from TRF2/DREF, may be an important factor in the selection of genes for dosage compensation.

Only the core 8 nucleotides of the DRE from *PCNA* are necessary to stimulate transcription of *lacZ* or the gene for chloramphenicol acetyl transferase, but the following two nucleotides (GA) are also required to compete in mobility shift assays with DREF (Yamaguchi et al., 1995). A double-stranded oligonucleotide representing three direct repeats of 18 bp (5′-CCTGCTATCGATAGATTC-3′) of the DRE from *PCNA* successfully competes with DREF

binding in a mobility shift assay (Seto et al., 2006).

1.3.4 The Glass Multimer Reporter

The *glass* gene encodes a multiple zinc-finger transcription factor that is required for correct photoreceptor development in the *Drosophila* eye (Moses et al., 1989). *glass* is predominantly expressed in the developing larval eye-antennal disc, in all cells posterior to the morphogenetic furrow (Moses and Rubin, 1991; Ellis et al., 1993). Some other tissues with photoreceptor cells also express *glass*, including the ocelli, small areas of the brain, and the embryonic Bolwig organ. Recombinant Glass protects a 27 bp fragment of the *Rh1* (alias *ninaE*) promoter from DNase digestion, and a pentamer of this fragment can drive expression in all cells where *glass* is expressed (Moses and Rubin, 1991; Ellis et al., 1993). An enhancer-promoter combination of the pentamer of *glass* binding sites and minimal TATA box-containing promoter of *hsp70* is known as the glass multimer reporter (GMR) (Hay et al., 1994). GMR also drives efficient expression of downstream coding sequences in *Drosophila* eyes and other *glass*-expressing tissues (Plautz et al., 1996; Hay et al., 1997).

1.3.5 A reporter gene system to detect changes in MSL complex activity

The *head involution defective* (*hid*) gene encodes a novel protein that functions to cause programmed cell death, or apoptosis (Grether et al., 1995). Expression of *hid* from the GMR enhancer-promoter construct (Section 1.3.4) causes a reduction in eye size (Grether et al., 1995). The effect is dependent on gene dose, as homozygotes have a smaller eye than heterozygotes. As the eye is a non-essential tissue, this dose-dependent phenotype is suitable for genetic screens to identify modifiers of apoptosis (Bergmann et al., 1998). Mutations that dominantly suppress the *GMR-hid* phenotype (have a larger eye) occur in factors whose wild-type function aids apoptosis; those that enhance the phenotype (have a smaller eye) represent apoptosis-inhibitors.

A similar approach was adopted to search for modifiers of dosage compensation. The male-specific DHS from *roX1* attracts MSL complex to autosomes, and a nine-copy multimer of the DHS attracts increased levels of complex that are sufficient for limited 'spreading' around the transgene insertion site (Kageyama et al., 2001). The same DHS mediates male-specific up-regulation of a down-stream *armadillo-lacZ* transgene at autosomal insertions (Henry et al., 2001). Both the monomer and nine-copy multimer of the *roX1* DHS were placed upstream of *GMR-hid* to create the plasmids RGH and R_9GH, respectively (Laverty, 2003). On preliminary analysis, male *Drosophila* expressing the modified constructs appear to exhibit a slightly smaller eye than females, likely due to attraction of the MSL complex. If the sex-specific phenotype is consistent, these lines would be suitable for a mutational screen for modifiers of dosage

compensation. Mutations in factors that help the MSL-mediated up-regulation of *roX1 DHS-GMR-hid* would suppress the male-specific size reduction; those in inhibitors of MSL complex would enhance the phenotype.

1.4 Aims

This study focused on the initiation of dosage compensation. The general aim was to investigate how the MSL complex is established on the X chromosome, and its subsequent effects on gene expression.

1.4.1 Determination of the relative order of key events in Drosophila dosage compensation

Due to the difficulty in observing the physiological initiation of such a developmentally early mechanism, many questions surrounding the establishment of compensation remain unanswered. Large-scale collection of tightly-staged embryos is difficult, not least due to the differing age of eggs at time of laying. The chromosomes of embryonic nuclei are also not suitable for detailed cytological analysis. Even the exact time of binding for each MSL component is controversial, perhaps due to differences in antibody affinities (McDowell et al., 1996). The system may be easier studied by artificially inducing compensation in females, and then observing the relative dynamics of initiation. The large, polytene chromosomes of third instar female larvae would be suitable for polytene observations.

For instance, as MSL1 and MSL2 seem to mediate the structural core of the complex, it is possible that these two proteins bind the X chromosome before any of the other proteins. It seems likely that MSL1 and MSL2 limit the activities of the other proteins to the male X chromosome. As the induced complex assembles in females, MSL1 and MSL2 might thus bind X chromosomal sites first. This can be observed both at the chromosomal level with fluorescent antibodies, and at specific genomic sites with chromatin immuno-precipitation.

Furthermore, the 'two-step' model of MSL association with the X chromosome is largely based on the relative affinities of each class of site, rather than the time of association. It remains to be tested that the high affinity sites are actually the first sites bound, rather than merely the sites with the highest affinity for the complex. If the complex indeed binds sites sequentially, such a process would be easily observed at the chromosomal level.

The importance of the $3'$ determinants in a compensated gene could also be clarified by the inducible system. It is possible that the $3'$ predominance of MSL proteins and histone modifications, and the demonstrated importance of transcription through $3'$ elements, could be artefacts of the direction of transcription. Alternatively, if the transcription unit forms a linked loop as occurs in yeast (O'Sullivan et al., 2004), apparent $3'$ importance may instead reflect a facilitation of re-initiation. With chromatin immunoprecipitation studies at points across a transcribed unit, the initial positions bound by the complex should be easily distinguished.

Specific Aims

- To construct an inducible dosage compensation system
- To determine the order of key initiation events upon induction

1.4.2 Identification of novel factors involved in Drosophila dosage compensation

Studies of factors that interact with the MSL complex provide valuable knowledge on the mechanism of dosage compensation. Each newly-identified component refines the model of action. The phosphorylation of H3S10 by JIL-1 is evidence that the chromatin is modified in more ways than just acetylation by MOF. Interactions with the NURF complex imply that MSL complex can act to relieve repression. Association with nuclear pore components indicate that spatial positioning of the chromatin within the nuclei may be important. It seems likely that such a complex system draws upon many factors that have less-specific roles than those of the MSL complex. However, the identification of factors associated with dosage compensation is not trivial. Whereas mutational screens for male-specific lethality easily identify the basic components, any additional factors with less-critical or sex-independent functions are more difficult to detect. To overcome this obstacle, a sensitive reporter gene system that measures levels of dosage compensation would be invaluable. With this approach, even subtle changes to MSL complex activity by interacting factors could be observed.

The reporter gene system that drives the apoptosis inducing gene *head involution defective* (*hid*) by the eye-specific Glass Multimer Reporter (GMR) enhancer-promoter was modified to reflect levels of *hid* compensation (Section 1.3.5). Placement of the MSL binding site at the DHS of *roX1* upstream of *GMR-hid* facilitated higher expression in males carrying the construct. Preliminary analysis appeared to indicate a sex-difference in eye sizes that would be appropriate for a mutational screen for factors interacting with dosage compensation. However, the system needed to be properly characterised to determine its suitability for such a mutagenesis screen.

Specific Aims

- To confirm the suitability of the *GMR-hid* system for a mutagenesis screen to identify new dosage compensation factors
- If needed, to modify the existing system, or construct an appropriate alternative, that would accurately monitor effects of MSL recruitment on gene transcription
- To identify additional factors involved in dosage compensation

1.4.3 Investigation of the deficiency in compensation of GMR-based expression

Whilst addressing the previous aims, it became apparent that GMR-based expression was to some degree inadequate to facilitate dosage compensation. This provided a unique opportunity. By investigating what was lacking in the GMR construct, insight could be gained as to what elements determine whether a gene becomes compensated.

The *GMR* expression system also allowed a functional test of a theory regarding the initiation of dosage compensation. MSL complex binds fewer genes in larval salivary glands than in early embryos, and the reduced number appear to be largely a subset of the earlier number (Legube et al., 2006). Taken together with the correlation to expressed genes, and the demonstrated stability of chromosomal association, the authors proposed that bound MSL complex may establish early, and then be maintained throughout development. The same conclusion was reached from observations that MSL complex can co-incide with RNA polymerase II on polytene chromosomes, but remain bound after polymerase dis-associates (Kotlikova et al., 2006). However, the relative importance of early expression for MSL binding was not established. Many tissue- and stage-specific genes (*e.g. white*) are clearly bound and compensated later in development, and thus competent to re-distribute MSL complex. Nevertheless, early expression of a gene may be an important determinant in the attraction of MSL complex. By providing the late, tissue-specific expression of *GMR-lacZ* with an additional early burst of expression, the relative importance of developmental stage on MSL attraction could be assessed.

Specific Aims

- To investigate what is necessary for GMR to be sufficient for dosage compensation
- To determine if an early burst of expression is sufficient for later compensation of *GMR-lacZ*

Chapter 2

Materials and methods

2.1 Bacterial methods

All plasmids were transformed in *Escherichia coli* DH5α, genotype: *supE*44 Δ*lac* U169 (ϕ80 *lacZ*ΔM15) *hsdR*17 *recA*1 *endA*1 *gyrA*96 *thi*-1 *relA*1. Glycerol stocks of all new strains bear the name of the transformed plasmid.

2.1.1 Growth and maintenance

Standard methods for handling and culturing *E. coli* were followed (Ausubel et al., 1994). Overnight plates were 1.5 % agar (w/v) LB medium, incubated at 37 °C for about 16 hours, then stored at 4 °C indefinitely. Overnight cultures were of LB medium, incubated at 37 °C, 200 rpm, for about 16 hours. Antibiotics were added to plates or broths at the recommended concentrations (Sambrook et al., 1989).

Luria-Bertani (LB) medium

1 % (w/v) bacto-tryptone, 0.5 % (w/v) bacto-yeast extract, 1 % (w/v) NaCl (Sambrook et al., 1989), autoclaved.

2.1.2 Creation of competent cells

Transformation-competent *E. coli* were made as described (Inoue et al., 1990), except that cultures were raised at 22 °C. A cell density with 0.6 absorbance at 600 nm was reached at about 30 hours.

2.1.3 Transformation

2.0 μL of a 20 μL ligation reaction (Section 2.3.9) was added to a gently-thawed 200 μL aliquot of *E.coli* DH5α competent cells (Section 2.1.2), mixed gently and incubated on ice for 30 minutes. The cells were heat-shocked for 30 seconds at 42 °C, and 800 μL of SOC (Sambrook et al., 1989) immediately added. The cells were left to recover for 1.5 hours at 37 °C, with agitation, and plated as 100 μL and 900 μL samples on overnight plates with LB medium and the appropriate antibiotic (Section 2.1.1).

2.2 Drosophila methods

All fly strains used in this study are listed in Appendix A.

Routine manipulations were carried out under Olympus NM-2 1x–4x or SZ61 zoom stereo dissecting microscopes, with halogen illumination from Olympus LG-PS2 or SZ2-LGB light

sources. Flies were immobilised with a continuous carbon dioxide stream through a porous stage. Fluorescent observations were made with an Olympus SZX12 stereo zoom microscope, U-LH100HG mercury UV-burner, and U-RFL-T controller. Green or yellow fluorescence was observed with the GFPA (Olympus) filter set, excitation 460–490 nm, emission 510–550 nm. Red fluorescence was observed with the RFP1 (Olympus) filter set, excitation 520–550 nm, emission 580+ nm.

2.2.1 Solutions

Ringer's solution

182 mM KCl, 46 mM NaCl, 3 mM $CaCl_2$, 10 mM Tirs-HCl, pH 7.2 (Ashburner, 1989), autoclaved.

2.2.2 Stock maintenance

To prepare culture media for *Drosophila*, agar (Neogen® acumedia® bacteriological), yeast (New Zealand food Industries Ltd Pinnacle active dried), and ground cornmeal (Kiwi Organic coarse polenta) or cornflour (Organic Living) were mixed in relative amounts as indicated below. Water was added slowly with stirring to form a paste, then a thick liquid. The mixture was boiled for at least two minutes to thicken the consistency, then allowed to cool slightly before addition of sugar (Simply pure cane white), moldex (methyl 4-hydroxybenzoate, SIGMA H3647), or molasses (Wholesome® Sweeteners organic). When needed, tetracycline (SIGMA T3383) was added to food that had further cooled below 60 °C, to a final concentration of 10 μg/mL. The food was dispensed into 100 mL glass bottles (Schott) or 30 mL plastic vials (LabServ LBS3560), and allowed to cool overnight, covered with Stockinette (Packaging House 1041648). Tight cotton bungs were added the next day, and the media used within seven to ten days, depending on ambient temperature and humidity.

Fly stocks were kept at 21 °C, with incident lab lighting, but experiments were generally conducted at 25 °C, in the dark. Humidity was not controlled.

Embryo collections were conducted on egg-laying medium, streaked with a freshly-prepared light yeast paste (yeast and water, to the consistency of motor oil).

Occasionally, 'Instant food' was prepared by mixing equal volumes of water with Carolina® Formula 4-24 flakes (about 7 mL to 1.5 g; 28 mL to 6 g).

Standard medium

Concentrations on volume before boiling: 1 %(w/v) agar, 4 %(w/v) yeast, 11 %(w/v) ground cornmeal, 13 %(w/v) sugar, 2 %(v/v) molasses, and 0.33 % moldex (pre-dissolved 8.9 % in

95 % ethanol).

High yeast medium

Concentrations on volume before boiling: 1.6 %(w/v) agar, 10 %(w/v) yeast, 10 %(w/v) sugar, and 0.33 % moldex (pre-dissolved 8.9 % in 95 % ethanol).

Egg-laying medium

Concentrations on volume before boiling: 2 %(w/v) agar, 6 %(w/v) cornflour, 4 %(v/v) molasses, and 0.33 % moldex (pre-dissolved 8.9 % in 95 % ethanol).

Soft medium

Concentrations on volume before boiling: 0.75 %(w/v) agar, 3 %(w/v) yeast, 11 %(w/v) cornflour, 13 % (w/v) sugar, 5 %(v/v) molasses, and 0.33 % moldex (pre-dissolved 8.9 % in 95 % ethanol).

2.2.3 Virgin collection

Cultures were cleared of adults, and incubated at 25 °C. Newly-emerged flies collected up to eight hours later were assumed to be virgin. Obviously-young flies were always treated as virgin.

2.2.4 Genetic crosses

Generally, five pair adult flies were set in a 30 mL vial (Labserv), or eight pair in a 100 mL bottle (Schott), on standard medium (Section 2.2.2) at 25 °C. Single males were crossed to three to four virgin females (Section 2.2.3), and single females received three to four males. Parents were removed at five to seven days, and offspring were collected morning and night at 11 to 13 days.

roX1 DHS-GMR-hid response to altered MSL levels

To compare the activity of *RGH* and R_9GH lines to *hsp83-msl2* (Figure 4.3a), w^*; ry^{506} Sb^1 $P\{ry^{+t7.2}=\Delta 2\text{-}3\}99B/TM6B$, Tb^1 (lab stock M1) males were mated to virgin *w;* $msl1^{L60}$*/CyO;* $P\{w^+=hsp83\text{-}msl2\}$ (lab stock M51) females, Sb^+, *Cy* male offspring selected, and mated to the *GMR-hid* transgenic line. Tubby pupae were separated from normal siblings. Both genotypes had one copy of the *GMR-hid* transgene, and Tb^+ flies also carried the *hsp83-msl2* transgene.

To compare the effects of both *hsp83-msl1* and *hsp83-msl2* on *RGH* and R_9GH lines (Figure 4.3b), *y w;* $msl2^1$ *cn;* $P\{w^+=hsp83\text{-}msl1\}2I$ $P\{w^+=hsp83\text{-}msl2\}6I$/+ (lab stock M48) males

were mated to virgin *y w; L^2/CyO,Cy pr cn^2 y^+* (lab stock #8) females, and *Cy, w^+* male offspring mated to virgin *w^*; ry^{506} Sb^1 P{$ry^{+t7.2}$=Δ2-3}99B/TM6B, Tb^1* (lab stock M1) females. *Cy, Tb, w^+* male offspring were mated to virgin *GMR-hid* females, and the resulting larvae separated to tubby and normal siblings. Both genotypes carried one copy of the *GMR-hid* transgene, and the *Tb^+* flies also carried one copy of the chromosome with *hsp83-msl1* and *hsp83-msl2*.

GMR-hid response to altered MSL levels

Both the *GMR-hid* (stock C60) and the *GMR-hid^{Ala5}* (stock C34) of Figure 4.5b were tested in the same manner.

To compare the reaction of *GMR-hid* to hsp83-*msl2*, two crosses were conducted to remove any contributory effect of the TM6 balancer chromosome. Males *w; $msl1^{L60}$/CyO; P{w+=hsp83-msl2}* (lab stock M51) by virgin *GMR-hid* females produced curly-winged offspring with one copy of the *GMR-hid*, and one copy of *hsp83-msl2* (and normal-winged *$msl1^{L60}$* heterozygotes, not assayed). In a separate cross, males *y^1 w^1; $msl1^{L60}$/CyO, pr cn^2 y^+* (lab stock #3) by virgin *GMR-hid* females produced curly-winged offspring with one copy of *GMR-hid* only (and normal-winged *$msl1^{L60}$* heterozygotes, not assayed).

For the reaction of *GMR-hid* to reduced MSL1, MSL2, and MLE, flies: *msl2 msl1 mle/CyO* (lab stock #18) males were crossed to virgin *GMR-hid* females. Curly-winged offspring had one copy of *GMR-hid* and wild-type MSL levels; normal-winged siblings also had one copy of *GMR-hid*, but only one functional copy of each of *msl1*, *msl2*, and *mle*.

GMR-lacZ response to hsp83-msl2

To observe the response of *arm-lacZ*, *GMR-lacZ*, *DRE-GAGA-GL*, or *GiL* at 2A to *hsp83-msl2* (Table 5.6, Figure 5.5), *w; $msl1^{L60}$/CyO; P{w^+=hsp83-msl2}* (lab stock M51) males were mated to virgin *w^{1118}; P{w^{+mC}=hs-hid}3,Dr^1/TM6C,cu^1 Sb^1* (Bloomington #9546) females. Resulting curly-winged, stubble-bristled males were mated to virgin homozygous *lacZ* females of the appropriate transgenic line, and *Cy^+* females collected. Stubble-bristled daughters carried one copy of the *lacZ* transgene; *Sb^+* sisters also had one copy of *hsp83-msl2*. These flies were weak and developmentally-delayed.

Creation of injection stocks C61–C65

To create stocks C61–C65 that contained both an *attP* landing site and a φC31 recombinase, *FM7; φC31* males and *attP; TM3* females were first created.

To create *FM7; φC31* males: *y^1 w^*;TM3,y^+ Ser^1/Sb^1* (lab stock #9) males were mated to virgin *w cv mof^1/FM7,y^{31d} sc^8 w^a B* (lab stock #17) females, and serate-winged, bar-eyed

males selected. These males were then mated to virgin *ϕC31* (lab stock J9) females, and serate-winged, bar-eyed virgin females selected. These females were mated to *ϕC31* (lab stock J9) males, and bar-eyed, normal-winged males selected.

To create *attP; TM3* females: y^1 w^*;*TM3*,y^+ Ser^1/Sb^1 (lab stock #9) males were crossed to *attP* flies of the corresponding landing site (Table A.1), and serate-winged males selected. These males were back-crossed to virgin *attP* females of the same stock, and serate-winged virgin females collected.

To create the stocks homozygous for both the landing site and the recombinase, the *FM7; ϕC31* males were crossed to the *attP; TM3* virgin females, and serate-winged progeny collected. The *Ser* siblings were mated to each other, and B^+; Ser^+ offspring collected. These flies were crossed as the founding parents of stocks C61–C65 (Table 2.1).

Table 2.1: Description of fly stocks C61–C65 for ϕC31 injections

Lab#	Landing site	attP stock	Genotype
C61	2A	zh-attP-2A	*y w pM{3xP3-RFP=attP}2A; pM{3xP3-RFP=ϕC31{3xP3-GFP=vas-ϕC31}}86F*
C62	6E	zh-attP-6E	*y w pM{3xP3-RFP=attP}6E; pM{3xP3-RFP=ϕC31{3xP3-GFP=vas-ϕC31}}86F*
C63	20C	zh-attP-20C	*y w pM{3xP3-RFP=attP}20C; pM{3xP3-RFP=ϕC31{3xP3-GFP=vas-ϕC31}}86F*
C64	3B	zh-attP-3B	*y w pM{3xP3-RFP=attP}3B; pM{3xP3-RFP=ϕC31{3xP3-GFP=vas-ϕC31}}86F*
C65	3Aa	zh-attP-3Aa	*y w pM{3xP3-RFP=attP}3Aa; pM{3xP3-RFP=ϕC31{3xP3-GFP=vas-ϕC31}}86F*

2.2.5 Transformation

Transformation of *Drosophila* with both *P* element and ϕC31 recombinase systems followed the same process; a modification of that by Spradling and Rubin (1982) to suit our laboratory conditions.

All *P*-element vectors were co-transformed with the Δ2-3 *P* transposase helper plasmid (Table A.3) into *y w* (lab stock #4) flies. Injection stocks for ϕC31 vectors were usually: C61 for insertions at 2A, C62 for insertions at 6E, C63 for insertions at 20C, and J5 for insertions at 86F (Tables 2.1, A.1). However, prior to the creation of stocks C61–C63 (Section 2.2.4), plasmids ALattB, GHattB, and GLattB were injected into crosses of J9 by zh-attP-2A, zh-attP-3Aa, zh-attP-3B, zh-attP-6E, or zh-attP-20C.

Plasmid preparation

High quality plasmid DNA was prepared with a Plasmid Maxi kit (Section 2.3.2). Alternatively, a large alkaline lysis plasmid preparation (Section 2.3.2) was purified by equilibrium centrifugation in an ethidium bromide-caesium chloride gradient (Section 2.3.3). Twenty or twenty-five micrograms of plasmid, with 5 μg of Δ2-3 *P* transposase helper plasmid if appropriate, was precipitated (Section 2.3.4), re-suspended in 50 μL injection buffer, filtered through a 0.45 μm Millex-HV (Millipore SLHVR04NL) filter, and stored at -20 °C. On each day of use the solution was centrifuged for 10 minutes prior to filling the needle, in an attempt to reduce clogging of the needle by particulate material.

Microinjection

About three days prior to injection, excess yeast was added to cultures of the appropriate stock or cross. On the injection day, 100 to 200 pairs of flies were tapped into polystyrene beakers (45 mm diameter, 90 mm deep), atop a 55 mm x 14 mm (Plastiques GOSSELIN BP53-02) petri dish of fresh egg-laying media streaked with a light yeast paste (Section 2.2.2), and incubated at room temperature under ambient light. Unlaid eggs were purged for one hour, and discarded, before 20 minute collection periods. Eggs were de-chorionated by rolling with fine forceps on double-sided tape, and aligned anterior left on a strip of double-sided tape attached to a microscope slide. About 10 to 20 embryos were aligned in a timed five minute dechorionation period. The slide was placed atop silica beads in a closed container, and incubated for an empirically-determined dehydration time of 3 to 10 minutes. Eggs were then immediately covered with Halocarbon 27 oil (SIGMA H8773), and injected as soon as possible.

Five to seven microlitres of plasmid preparation was loaded with a 20 μL microloader tip (eppendorf 5242 956.003) into either a Femptotip® II (eppendorf 5242 957.000) or a needle pulled from a quartz micropipette, and the needle attached to a micro-manipulator (Le-

ica/Leitz), and a Femptojet express transjector (eppendorf), supplied with pressurised nitrogen gas. A small but indeterminate amount of injection preparation was injected to the posterior end of the prepared embryos. Injections were conducted with a CK2 inverted compound microscope (Olympus), illuminated by a ULWCD0.30 light source. Damaged or improperly-aged embryos were destroyed with forceps under a dissecting microscope. The slides were incubated overnight at 18 °C in a chamber (QNA International MIC-101) under pressurised oxygen and high humidity. The chamber was placed at 21 °C on the following day, allowing larval collection in the morning of the third day. Hatched larvae were transferred on small squares of moist tissue paper (Kimberly-Clark Kimwipe) to vials of soft or standard media (Section 2.2.2) with furrows dug in the media, and raised at 25 °C.

Crosses to establish lines

For all injections, single G_0 (injected) flies were mated to *y w* (lab stock #4) flies (Section 2.2.4), and the progeny screened for the positive transformation marker of red or orange eyes. Single transformant first generation (G_1) flies were again mated to *y w* flies, and the progeny examined to determine the number of transgene insertions. A 1:1 ratio of coloured to white-eyed flies indicated a single insertion; a 3:1 ratio a double insertion. Double insertions could usually be separated by eye colour, and were never observed for ϕC31 transformations. Up to three lines were selected, with preference given to single insertions derived from male G_0 flies, then virgin female G_0 flies.

For *P*-element insertions, homozygous stocks were created at the same time as establishing genetic linkage of chromosomal insertion. Red-eyed G_2 males were crossed to virgin *y w; L^2/CyO,Cy pr cn^2 y^+* (lab stock #8) females, then red-eyed, curly-winged offspring were back-crossed to *y w* flies, and (independently) self-crossed. Insertions on the second chromosome would segregate from the curly balancer chromosome in the backcross, and homozygous stocks could thus be established from the Cy^+ progeny of the *w^+; Cy* self-cross. An independent crossing scheme with *y^1 w^*;TM3,y^+ Ser^1/Sb^1* (lab stock #9), or *w^{1118}; P{w^{+mC}=hs-hid}3,Dr^1/TM6C,cu^1 Sb^1* (Bloomington #9546) virgins was conducted to examine the third chromosome. Insertions on the X chromosome were obvious from a skewed male:female ratio, confirmed with a test cross of w^+ males to *y w* virgin females, and homozygous stocks established based on a darker eye colour of homozygous females. No insertions were observed on the Y or 4th chromosomes. Insertions that proved homozygous lethal were maintained over *FM7*, *CyO*, *TM3*, or *TM6* balancer chromosomes.

For ϕC31 insertions on the X chromosome, virgin transformant (w^+) G_2 females were crossed to *w cv mof^1/FM7,y^{31d} sc^8 w^a B* males (lab stock #17), and w^+ offspring self-crossed. Homozygous stocks were established from the normal-eyed progeny of the self cross. For ϕC31 insertions on the third chromosome, transformant G_2 were crossed to *w^{1118}; Pw^{+mC}=hs-*

hid3,Dr^{1}*/TM6C,cu*^{1} *Sb*^{1} (Bloomington #9546), and stubble-bristled offspring self-crossed. Homozygous stocks were established from the normal-bristled progeny of the self cross.

For all ϕC31 insertions, red fluorescent eyes were observed, as a confirmation of the *attP* landing site (marked with *3xP3-RFP*). Furthermore, green fluorescent eyes were selected against to remove the recombinase itself (marked with *3xP3-RFP*). This was most commonly done by selecting transformant G_1 that also did not fluoresce under blue light.

All transgenic lines generated in this study are listed in Appendix H.

PCR confirmation of transformation

To confirm a transformation event, genomic DNA was prepared (Section 2.3.2), and diluted 37.5-fold in water. Five microlitres of the diluted genomic DNA was used as template in a standard PCR reaction (Section 2.3.10) with primers to amplify a region of the inserted DNA, and examined by gel electrophoresis (Section 2.3.6). To confirm correct targeting of ϕC31 insertions, a second PCR experiment was conducted with the forward primer complementary to the inserted cassette, and the reverse primer complementary to the *attP* landing site. The particular reactions are listed in Table 2.2 (and reaction detail is in Table 2.3).

Injection Buffer

50 mM KCl, 0.1 mM sodium phosphate buffer, pH 6.8 (Sambrook et al., 1989), after Spradling and Rubin (1982).

Pulled injection needles

Needles were pulled from 1 mm quartz (SUTTER QF100-70-10) micropipettes, with a P-2000 (SUTTER INSTRUMENT) needle puller, set as follows: heat 700, filament 4, velocity 60, delay 150, and pull 166 to 169. The correct tip was created with a BV-10 K.T. Brown type (SUTTER INSTRUMENT) micro-pipette beveller, by grinding at an angle of 25 degrees on a dry 0.05 μm alumina abrasive plate. Needles were pulled and bevelled by Esther Belikoff[1].

2.2.6 Inverse PCR

To determine the position of insertion for *P*-element insertions, the inverse PCR protocol of Jay Rehm available from the Berkeley Drosophila Genome Project website (http://www.fruitfly.org/about/methods/inverse.pcr.html), as modified by Roger Hoskins, September 18, 2006, was followed. However, the flies were starved a few hours at 25 °C prior to freezing or homogenising

[1]Massey University, Palmerston North, New Zealand

Table 2.2: Reaction guide for PCR confirmations of ϕC31 *Drosophila* transformations

		PCR Reaction (Table 2.3)	
Plasmid	Location	Transformation	Targeting
ALattB	2A	H	I
	6E	H	I
	20C	H	I
GHattB	2A	A	B
	6E	A	B
	20C	A	B
GLattB	2A	A	B
	6E	A	B
	20C	A	B
CL12	2A	A	B
CL13	2A	A	B
	6E	C	B
CL15	2A	A	B
CL19	2A	A	B
CL21	2A	C	B
CL25	2A	A	B
GEG	2A	A	B
GiL	2A	C	B
GLSV40	2A	F	G
GL18D	2A	D	E
tetO-msl2-HA3-attB	20C	Q	E

in order to clear ingested yeast from the gut, and the flies were homogenised in a 1.7 mL micro-centrifuge tube (AXYGEN) with a micro-pestle (eppendorf 0030 120.973), in 200 μL Buffer A (0.1 M Tris-HCl, pH 7.5; 0.1 M EDTA, pH 8.0; 0.1 M NaCl; 0.5 % SDS). An additional 200 μL of Buffer A was added, and the flies ground further. Furthermore, the ligation was conducted with NEB T4 DNA ligase and a corresponding buffer, and the PCR step was as for a standard PCR (Section 2.3.10), but using 5 μL of the ligation reaction as template, 35 cycles instead of 30, and a 2 minute extension time at 72 °C. The PCR product was purified with a PCR purification kit (Section 2.3.3), and eluted in 50 μL.

Two nanograms of product per 100 bp amplicon length was used as the template for DNA sequencing (Secition 2.3.7). Vector sequences were trimmed from the resulting sequence, and the genomic portion compared to the *Drosophila* genome by BLAST analysis.

2.3 Molecular biology methods

All oligonucleotides used in this study were sourced from Proligo/SIGMA, and are listed in Appendix A. All plasmids used or created in this study are listed in Appendix A and new plasmids are illustrated in Appendix I. DNA solutions were stored in TE buffer (Section 2.3.1) at -20 °C indefinitely. Unless otherwise indicated, all enzymes, corresponding buffers and supplements were supplied by NEB. dNTPs and ATP were supplied by Roche. Centrifugation steps were carried out with a Heraeus® Biofuge *pico*® (Thermo Scientific), Heraeus® Fresco® 21 (Thermo Scientific), or Sorvall® RC-5C centrifuge.

2.3.1 Solutions

TE buffer

10 mM Tris-HCl, 1 mM EDTA, pH 8.0 (Sambrook et al., 1989).

dNTPs, 10 mM each

10 mM dATP, 10 mM dGTP, 10 mM dCTP, 10 mM dTTP.

2.3.2 DNA Isolation

Plasmid isolation by commercial kits

High-quality preparations of plasmid DNA for DNA sequencing were isolated from 5 mL overnight LB cultures (Section 2.1.1) with the QIAprep Miniprep (QIAGEN 27104), *AccuPrep*®

Plasmid Extraction (BIONEER K-3030-1), or AxyPrep™ Plasmid Miniprep (AXYGEN AP-MN-P-50) kits, according to the manufacturer's instructions.

Routine preparations of plasmid DNA sufficient for cloning purposes were isolated from 250 mL overnight LB cultures (inoculated 1:1000 from a 5 mL eight hour starter culture) with the Plasmid Midi (QIAGEN 12143) kit, mostly according to the manufacturer's instructions. However, DNA was precipitated with 1.0 volume isopropanol in 30 mL glass (COREX 8445) centrifuge tubes, and centrifuged at 8000 g, 4 °C, for 90 minutes.

Large amounts of high-quality plasmid DNA for *Drosophila* transformation were prepared with a Plasmid Maxi (QIAGEN 12162) kit from 250 mL overnight LB culture (inoculated 1:1000 from a 5 mL eight-hour starter culture), according to the manufacturer's instructions, and re-suspended in 500 μL of TE buffer (Section 2.3.1). DNA was precipitated in COREX tubes as for the Plasmid Midi kit.

Alkaline lysis plasmid isolation

Crude preparations of plasmid DNA for screening transformant clones were isolated from 5 mL overnight LB cultures (Section 2.1.1), and conducted mostly as described (Sambrook et al., 1989), except with 5 M potassium acetate (pH 4.8) as Solution III, and without the phenol-chloroform purification steps. DNA was re-suspended in 30 μL of TE buffer (Section 2.3.1).

Large-scale preparations of plasmid DNA, prior to purification by ultra-centrifugation on an ethidium bromide-caesium chloride gradient, were isolated from 250 mL overnight LB cultures (inoculated 1:1000 from a 5 mL eight hour starter culture) essentially according to Sambrook (1989). Briefly: cells were harvested by a 10 minute centrifugation at 4 °C, 6000 g, re-suspended in 8 mL of Solution I, supplemented with 1 mL of Lysozyme (SIGMA L-6876, 25 mg/mL in Solution I) and incubated 10 minutes at room temperature; mixed with 18 mL of Solution II and incubated 10 minutes on ice; then mixed with 13.5 mL Solution III (only 5 M potassium acetate, pH 4.8), and centrifuged for 10 minutes at 20 000 g, 4 °C. DNA was precipitated with 1.0 volume isopropanol for 10 minutes at room temperature, centrifuged for 10 minutes at 15 000 g, room temperature, washed with 8 mL ethanol (70 % v/v in water), dried for around one hour at 37 °C, and re-suspended in 4 mL TE buffer (Section 2.3.1).

Drosophila genomic DNA isolation

Drosophila genomic DNA was prepared according to the instructions for a "genomic prep" in the inverse PCR protocol (Section 2.2.6).

2.3.3 DNA purification

Routine purifications of DNA solutions with phenol:chloroform were conducted as described (Sambrook et al., 1989). If necessary, TE buffer (Section 2.3.1) was added to the starting volume to ensure at least 250 μL.

DNA was usually purified from enzymatic reactions with the QIAquick® PCR purification (QIAGEN 28104) kit, according to the manufacturer's instructions, and eluted in the supplied elution buffer (10 mM Tris-HCl, pH 8.5).

Linear fragments of DNA were purified from agarose gels with the QIAquick® gel extraction (QIAGEN 28704) kit, according to the manufacturer's instructions, and eluted in the supplied elution buffer (10 mM Tris-HCl, pH 8.5).

To prepare plasmid DNA for *Drosophila* transformation, crude alkaline lysis preparations (Section 2.3.2) were purified by equilibrium centrifugation on a caesium chloride-ethidium bromide gradient, largely as described (Sambrook et al., 1989), but with the following modifications: 1.05 g CsCl per mL of starting plasmid solution was used, the proteins were separated by centrifugation at 17 000 g for 10 minutes, the 6 mL PA Ultracrimp (Sorvall 03945) ultracentrifuge tube was filled with blank TE/CsCl-ethidium bromide instead of paraffin oil, and the ultracentrifugation was at 265 000 g for 16 to 18 hours with an OTD Combi ultra-centrifuge and TV-865 rotor (Sorvall). Ethidium bromide was extracted with TE-saturated iso-amyl-acohol (3-Methylbutanol, SIGMA 19392), and then the solution diluted with 2 volumes TE to prevent precipitation of CsCl. DNA was precipitated with 2 volumes ethanol (Section 2.3.4), and re-suspended in 500 μL TE buffer.

2.3.4 DNA precipitation

DNA was precipitated from solutions with ethanol or isopropanol, as described (Sambrook et al., 1989).

2.3.5 DNA quantification

Solutions of DNA were usually quantified with the NanoDrop® ND-1000 (NanoDrop Technologies) spectrophotometer, but the Ultraspec®3000 (Pharmacia Biotech) was also used. Occasionally, the concentration of electrophoresed DNA fragments was estimated by comparison with the brightness of ethidium bromide-stained fragments in the prepared λ DNA molecular weight marker (Section 2.3.6).

2.3.6 Gel electrophoresis

DNA electrophoresis was conducted in the Horizon® 58 (LIFE TECHNOLOGIES/GIBCO BRL), MINI-SUB® or MINI-SUB® CELL GT (BIO-RAD), or the MiniGel Electrophoresis System (Select BioProducts SBF100) horizontal electrophoresis apparatus.

Gels were set from heated solutions of TAE containing 0.5 to 1.7 % SeaKem® LE (Lonza 50004) agarose, and cast in the appropriate tray. Set gels were assembled in the electrophoresis tank, and covered with TAE buffer. 0.2 volumes of Gel loading dye was added to DNA to be electrophoresed, mixed, and the combined sample loaded to the prepared gel. An aliquot of 1Kb Plus ladder (Invitrogen 10787-018, diluted 10-fold in water, with a final 1 x concentration of Gel loading dye) or λ molecular weight markers was also loaded. Following electrophoresis, gels were stained in 0.5 μg/mL ethidium bromide for 15 to 30 minutes, and briefly de-stained in water. Gels were photographed with a Gel Doc (BIO-RAD) trans-illuminator, and recorded with the Quantity One® (BIO-RAD, version 4.4.0) software.

Gel loading dye, 6x

0.25 % bromophenol blue, 40 % sucrose (Sambrook et al., 1989).

TAE buffer, 50 x

2 M Tris, 0.1 M EDTA, 5.71 % Glacial acetic acid (Ausubel et al., 1994), autoclaved.

λ DNA molecular weight markers

250 μg bacteriophage λ DNA (NEB N3011) was digested overnight at 37 °C (Section 2.3.9) with 125 units each of *Hind* III and *Sst* II, 1 x NEBuffer 2, in a final volume of 600 μL. Digested DNA was diluted in water with a final concentration of 1 x Gel loading dye. Resulting linear fragments have the sizes (in kb): 20.3, 9.4, 4.4, 3.7, 2.8, 2.3, 2.0, 1.5, 1.1, 0.55, 0.21, 0.1.

2.3.7 DNA Sequencing

All DNA sequencing reactions and separations were performed by the Genome Service of the Allan Wilson Centre for Molecular Ecology and Evolution (Massey University, Private Bag 11 222 Palmerston North, New Zealand, http://www.allanwilsoncentre.ac.nz/), using an ABI3730 DNA analyser for Sanger sequencing. Samples were prepared as instructed. Primers were either as listed in Table A.2, or supplied by the Service.

2.3.8 Annealing of oligonucleotides

To prepare a double-stranded DNA fragment, 5 μg of each of the designed oligonucleotides (Table A.2) were mixed, in 50 mM NaAc and total volume 50 μL. The solution was boiled in a beaker over a bunsen burner for 5 minutes, then left to slowly cool for several hours. If the oligonucleotides were not designed with 5′ phosphate group modifications, 1.5 μL (15 units) of T4 polynucleotide kinase was added to the 50 μL solution, and DNA ligase buffer to 1 x concentration, then incubated for 1 hour at 37 °C, inactivated by heat, and purified by extraction with phenol/chloroform (Sambrook et al., 1989).

2.3.9 Enzymatic manipulations of DNA

Restriction endonuclease digestion

Preparative digestions of plasmid DNA for molecular cloning involved: 5 μg DNA, 2 μL each restriction enzyme, in 1 x appropriate buffer, with 1 x BSA if required, to a final volume of 50 μL water, for 3 hours at the appropriate temperature. A further three hours incubation at a second temperature was performed if two enzymes with different incubation temperatures were used. If possible, the reactions were inactivated by heat as directed.

Diagnostic digestions of transformed *E. coli* clones (Section 2.1.3) involved: 2 to 2.5 μL plasmid DNA, 0.3 μL each restriction enzyme, 1μg/mL RNase A (Roche 10 109 142 001), in 1 x appropriate buffer, with 1 x BSA if required, to a final volume of 10 μL, for 1 hour at the appropriate temperature. Diagnostic digestions of plasmids to confirm identity were similar, but with 200 to 500 ng of plasmid DNA, and without RNase.

Partial digestions of plasmid DNA to obtain molecules not fully digested were conducted with serial dilutions of the enzyme, as described (Ausubel et al., 1994).

Creation of blunt ends

Following restriction digestion and inactivation of the restriction enzyme(s), 1 unit per microgram plasmid DNA of DNA Polymerase I, Large (Klenow) Fragment (NEB M0210) was added, with dNTPs to a final concentration of 33 μM each. Reactions were incubated for 15 minutes at 25 °C, and inactivated by the addition of EDTA to 10 mM (if no further enzymatic manipulations were to be performed) and incubation for 10 minutes at 75 °C.

Alternatively, 1 unit per microgram of T4 DNA polymerase (NEB M0203) was added, with dNTPs to a final concentration of 100 μM each. Reactions were incubated for 15 minutes at 12 °C, then inactivated with EDTA to 10 μM (if appropriate), and incubation for 20 minutes at 75 °C.

DNA phosphatase treatment

Following a 50 μL restriction digestion reaction of 5 μg plasmid DNA, 1.5 μL (15 units) of Calf Intestinal Alkaline Phosphatase was added, and incubated at 37 °C for 1.5 hours.

Ligation

A typical molecular cloning reaction to ligate a DNA fragment to a digested vector involved: 50 ng prepared vector, an appropriate amount of prepared insert DNA for the desired molar ratio (typically 3:1 insert:vector), 1 μL (400 units) T4 DNA Ligase, in 1 x supplied buffer (or digestion buffer supplemented with ATP to 1mM), to a final volume of 20 μL, for 16 hours at 16 °C. Ligation solutions were stored at 4 °C until used for transformation, then stored indefinitely at -20 °C.

Ligation reactions to re-circularise a linear plasmid contained 15 ng of DNA in a 20 μL reaction.

Reactions to concatenate annealed oligonucleotides (3DRE4GAGA) were similar, but with 500 ng annealed oligonucleotides.

2.3.10 PCR

Unless otherwise indicated, a standard PCR reaction consisted of 0.2 mM each dNTP (Section 2.3.1), 0.2 μM each primer (Table A.2), 0.25 μL (1.25 U) Taq DNA Polymerase (Roche 11 418 432 001), and 1 x supplied buffer, in a total volume of 25 μL. The final $MgCl_2$ concentration was 1.5 mM. A master mix of dNTPs and primers was prepared first, template added, then a second master mix of polymerase and buffer dispensed immediately before cycling. Cycling conditions were an initial denaturation step of 5 minutes at 94 °C; 30 cycles of 30 seconds at 94 °C, 30 seconds at the annealing temperature, an extension time at 72 °C; and a final elongation step of 10 minutes at 72 °C. The annealing temperature was as recommended by MacVector (version 9.5) software, and the extension time was 45 seconds per kilobase amplicon length, rounded up to the nearest 30 second interval. Exact times used are listed in Table 2.3. A negative (no template) control was always performed. A 5 μL aliquot of the PCR product was examined by gel electrophoresis (Section 2.3.6). If required for further cloning purposes, the desired PCR product was purified with a PCR purification or gel extraction kit (Section 2.3.3).

Long-range PCR to amplify and add epitope tags to the *msl2* gene used 100 pg of primer as template, and 1.5 μL (7.5 U) of Pwo DNA Polymerase (Roche 11 644 947 001), with appropriate buffer, in a 50 μL reaction. The initial denaturation step was only 2 minutes at 94 °C, and the final elongation step 6 minutes at 72 °C.

PCR with genomic DNA as input (Section 2.3.2) used 5 μL of a 37.5-fold dilution of genomic DNA as template in a standard reaction.

A temperature gradient was used for amplification of the promoter region of *armadillo* and the *ninaE* second intron. Successful temperatures were 60, 63, and 66 °C for the *armadillo* fragment, and 50, 51, 53, or 55 °C for the *ninaE* intron.

Table 2.3: PCR Reaction Detail

Reaction	Primer		Annealing Temp(°C)	Extension Time(s)	Amplicon Length(bp)
	Forward	Reverse			
A	GMR-DsRedFseq2	h70p(A)rev	48.0	30	340
B	GMR-DsRedFseq2	attPTargetRev	53.2	30	691
C	lacZ3′Fwd	h70p(A)rev	52.8	30	577
D	lacZ3′Fwd	SV40RevPstI	55.6	90	1491
E	armlacZNotI	attPTargetRev	53.8	90	440
F	lacZ3′Fwd	SV40RevPstI	53.9	60	939
G	lacZ3′Fwd	attPTargetRev	54.4	60	1262
H	lacZ3′Fwd	Pry1	54.0	60	1059
I	armlacZseq	attPTargetRev	55.4	60	509
J	pXLsense4013/pr1	Nhe/Not/HA/M2/asen	60.0	36	600
K	pXLsense4013/pr1	HIS/M2/ASENSE/pr2	60.0	36	600
L	M2/HIS/SENSE/3	NotI/BirA/asense4	60.0	68	1100
M	pXLsense4013/pr1	NotI/BirA/asense4	60.0	102	1600
N	phic31For	phic31Rev	60.5	60	550
O	armDREGAGAfwd	armDREGAGArev	60–66[a]	30	350
P	ninaEin2F	ninaEin2R	50–55[a]	30	214
Q	pXLsense4013/pr1	AS5	54.5	30	748

[a]Successful temperatures from a gradient range

2.3.11 Molecular cloning

All vectors used or created in cloning steps are listed in Appendix A. The typical procedure for cloning an insert to a larger vector is described: The identity of all starting vectors was confirmed with diagnostic restriction digestions (Section 2.3.9), and analysis by gel electrophoresis (Section 2.3.6). Plasmid DNA to yield both vector and insert DNA was digested, and blunt ends created if need be (Section 2.3.9). The digested vector was de-phosphorylated (Section 2.3.9), and purified with a PCR purification kit (Section 2.3.3). The insert fragment was released from the vector backbone by separation with gel electrophoresis. The desired band was cut from the gel under ultra-violet illumination, and the DNA purified with a gel extraction kit (Section 2.3.3). Both vector and insert solutions were quantified (Section 2.3.5), and a ligation reaction (Section 2.3.9) set up with a 3:1 insert:vector molar (ng/kb) ratio. Insert fragments less than 200 bp in size were more successfully ligated at a 2:1 insert:vector ratio. An aliquot of the ligation solution was transformed into chemically-compentent *E. coli* cells (Section 2.1.3). Five millilitre overnight cultures (Section 2.1.1) of LB with the appropriate antibiotic were inoculated with single colonies picked from the transformation plates, and crude alkaline lysis preparations of plasmid (Section 2.3.2) obtained for each clone. At the same time, the single colonies were also streaked to a backup overnight plate of LB and antibiotic (Section 2.1.1). The successful cloning event and orientation of insert was determined by a diagnostic digestion with restriction endonuclease(s), and analysis by gel electrophoresis. Clones were also usually confirmed with DNA sequencing (Section 2.3.7). A larger overnight culture of the correct clone was then prepared by inoculation from the 5 mL culture or (more commonly) the backup plate. A 25 mL culture purified with the Plasmid Midi kit (Section 2.3.2) was sufficient to yield enough plasmid for subsequent cloning steps. A 250 mL culture was prepared if the new plasmid was to be injected to *Drosophila* (Section 2.2.5). In both cases, a 0.6 mL sample of the overnight culture was mixed with 0.4 mL of 50 % (v/v) glycerol, snap-frozen in liquid nitrogen, and stored at -80 °C. Note that stocks of pBBHC, pBBHN, and pFH10 are stored in 7.5 % glycerol. The identity of the larger plasmid preparation was again confirmed by restriction endonuclease digestion, gel electrophoresis, and (usually) DNA sequencing.

Cloning steps to destroy a restriction endonuclease site in a plasmid, or delete all sequences between two restriction endonuclease sites, involved digestion, creation of blunt ends if appropriate, and re-ligation at a low concentration to promote intra-molecular ligation events. The cloning step to remove the DNA between the two *Xba* I sites of pHF10 involved a ligation of 5 μg digested vector in 10 μL

Cloning steps to delete one of two restriction endonuclease sites in a plasmid involved partial digestion, creation of blunt ends, and a ligation at low DNA concentration (20 ng vector per 20 μL reaction) to promote re-circularisation of the vector (Section 2.3.9).

Cloning steps involving a linker oligonucleotide to mediate the annealing of non-compatable

DNA ends involved a three-way ligation with a molar ratio of 5:2:1 or 2:2:1 linker:insert:vector. The linker was first prepared by annealing two oligonucleotides (Section 2.3.8).

Cloning steps where the insert was derived from synthetic nucleotides involved annealing of the oligonucleotides (Section 2.3.8), digestion if necessary (Section 2.3.9), and a ligation reaction with a 1:2 molar ratio of vector to insert. For the insertion of the 3DRE4GAGA molecule in the creation of pCL24 and pCL19, the annealed double-stranded molecule was concatenated by ligation (Section 2.3.9) prior to digestion with *Xho* I, in order to avoid problems with digestion near DNA termini.

Cloning steps where the insert was derived from a purified PCR product (Section 2.3.10) involved a 50 μL digestion of the maximum available amount of PCR product, followed by purification with a PCR purification kit, prior to ligation.

2.4 Reverse Transcriptase (RT) -PCR

2.4.1 Precautions for RNA work

Glassware was baked for four hours at 160 °C in a dry oven. Homogenizers were soaked overnight in DEPC-H_2O, incubated at 37 °C for several hours, rinsed in DEPC-H_2O and autoclaved. Solutions were prepared with DEPC-H_2O. Plastic beakers were wiped with RNase *AWAY* (Molecular BioProducts 7000).

DEPC-H_2O

0.1 % (v/v) diethyl pyrocarbonate (SIGMA D5758) in water. The solution was prepared in baked glassware, left overnight, then autoclaved.

2.4.2 RT-PCR reactions

Fifty mixed-sex flies were starved four hours to remove ingested yeast, snap-frozen in liquid nitrogen, and stored at -80 °C. Total RNA was extracted with 50 μL TRIzol® Reagent (Invitrogen 15596-018), according to the manufacturer's instructions, and quantified with the NanoDrop spectrophotometer (Section 2.3.5). Thirty micrograms of RNA in 50 μL was treated with TURBO™DNase (Ambion 2238) according to the manufacturer's instructions, except with 3 x (3 μL) the recommended enzyme amount, and twice (60 min) the incubation time. The RNA was re-suspended in 15 μL DEPC-H_2O, and quantified again.

Complementary DNA (cDNA) was synthesised with Expand Reverse Transcriptase (Roche 11 785 826 001) according to the manufacturer's instructions. Each 20 μL reaction contained 1 μg of total RNA, 1 μL of RNaseOUT™(Invitrogen 1077-019), and either 80 pmol $p(dT)_{15}$

primer (Roche 10 814 270 001), 30 pmol lacZ493R (Table A.2) gene-specific primer, or 800 pmol $p(dN)_6$ (Roche 11 034 731 001) random primer to attempt accurate cDNA synthesis. Reactions lacking reverse transcriptase were also performed to ensure complete removal of genomic DNA. A standard PCR reaction (Section 2.3.10) was conducted on a template of 1.0 μL reverse transcription reaction, with primers GiLfor and GiLrev (Table A.2), for 35 cycles with an annealing temperature of 54 °C. The RNA quality was ascertained by separate PCR reactions with primers rp49 596F and rp49 720R (Table A.2), under the same conditions. Negative PCR controls for non-specific PCR products (no template), and positive controls (genomic DNA as template for GiL reactions, a known good RNA preparation for the rp49 reactions) were also performed. Five microlitre aliquots of PCR reaction were analysed on a 1.5 % agarose gel by electrophoresis (Section 2.3.6).

2.4.3 Band quantification

The intensities of the bands corresponding to RNA species were calculated using the 'Gels' function of ImageJ version 1.43o (NIH, Maryland; http://rsb.info.nih.gov/ij/), with reference to the advice of Luke Miller[2]. Areas of interest were drawn around the lanes, avoiding dust spots that would contribute to the pixel count (Appendix F). Bands of interest were sectioned from the generated intensity plots, and their areas calculated. The apparent sizes of the bands of interest were determined by comparison with a hand plot of the standard curve of the molecular marker. Areas of each band were divided by the length of each fragment to obtain an area proportional to molecular weight. These values were summed for each lane, and individual bands expressed as a proportion of the total.

2.5 Western blot analysis

For all test samples, male *tetO-msl2* flies were mated to *yp1-tTA* (lab stock M18) virgin females on high yeast medium, with or without tetracycline, in the dark at 25 °C (Section 2.2.2), and the progeny analysed. For the positive control, *hsp70-msl2* (lab stock S13) flies were raised on standard medium at 25 °C.

Adult females were collected at 14 days incubation. Samples to be heat-shocked were first placed in a 37 °C water bath for one hour, then left to recover (without food) for four hours at 25 °C. Non-heat shocked samples were also starved for four hours at 25 °C. All samples were then snap frozen in liquid nitrogen as aliquots of 10 females per 1.5 mL microfuge tube (eppendorf), and stored at -80 °C. Frozen flies were homogenised in 400 μL Laemmli

[2]Northeastern University, Massachusetts, http://www.lukemiller.org/journal/2007/08/quantifying-western-blots-without.html

buffer in 1 mL glass homogenisers (Kimble Kontes 885450-0020), incubated at 100 °C for five minutes, then sonicated for 10 minutes in an Elmasonic sonicating water bath (Elma S10H, 'sweep' setting) with ice. Samples were centrifuged at full speed for ten minutes at 4 °C, and the supernatant retained. Samples not immediately used were stored at -20 °C, and incubated at 100 °C for five minutes before loading on the gel.

Ten microlitres of sample was mixed with 2 μL 6x Gel loading dye (Section 2.3.6), and loaded alongside Benchmark™ Pre-stained Protein Ladder (Invitrogen 10748-010) markers to a poly-acrylamide gel in a Mini-PROTEAN 3 System (BIO-RAD 165-3302) electrophoresis cell, prepared according to the manufacturer's instructions. A separating gel of 10 % acrylamide, stacking gel of 4 % acrylamide, and the running buffer were all prepared according to Laemmli (1970). Samples were separated at 100 volts for about one hour. Proteins were transferred to Hybond-C nitrocellulose membranes (Amersham RPN303C) with the Mini Trans-Blot® Electrophoretic Transfer Cell (BIO-RAD 170-3930) according to the manufacturer's instructions. The transfer buffer was 25 mM Tris, 192 mM glycine, 20 % methanol, and the running conditions were 30 volts, approximately 150 mA, overnight.

Unless otherwise mentioned, all washing and antibody incubation steps were performed with gentle agitation, at room temperature. Membranes were washed ten minutes in TBS+Tween-20, then incubated one hour in blocking buffer. The membranes were then incubated between 3 hours and overnight with primary antibody in 10 mL blocking buffer, and washed three times in TBS+Tween-20. A 10 mL solution of blocking buffer containing the secondary horseradish peroxidase conjugate antibody was added, and the membranes incubated for one hour, followed by three more washes in PBS+Tween-20. Horseradish peroxidase was detected with 1.4 mL / membrane of Supersignal® West Dura Extended Duration Substrate (Pierce Thermo 37071), prepared according to the manufacturer's instructions, incubated for five minutes. Excess substrate was dripped off, and the membranes sealed between acetate sheets. Detection was between 30 seconds and three minutes with a Fujifilm LAS-1000 and Intelligent Dark Box II.

For the detection of the HA tag, 2 μL of rat monoclonal anti-HA (Roche 11 867 423 001) primary antibody was used, and incubation was at 4 °C. The secondary antibody was 3 μL of sheep anti-rat-horseradish peroxidase conjugate (Boehringer Mannheim 1348 758). To detect MSL2, 4 μL of either rabbit anti-MSL2 peptide (affinity-purified polyclonal against DMPTSC-PAETPPTSC peptide, from Genscript) or rabbit anti-MSL2 (Mitzi I. Kuroda[3]) primary antibodies, and 2 μL of donkey anti-rabbit IgG ECL-horseradish peroxidase linked (Amersham NA934) secondary antibody, were used.

[3]Harvard Medical School

Laemmli buffer

62.5 mM Tris-HCl pH6.8, 10 % glycerol, 2 % SDS, 5 % 2-mercaptoethanol, 1 x protease inhibitor (Roche 11 697 498 001), after Laemmli (1970). Stored at -20 °C, adding protease inhibitor and 2-mercaptoethanol only immediately prior to use.

TBS+Tween-20

10 mM Tris-HCl, 150 mM NaCl, 0.05 % (v/v) polyoxyethylene sorbitan monolaurate (Tween-20, Bio-Rad 170-6531).

Blocking buffer

5 % skim milk powder in TBS + Tween-20.

2.6 Female viability assay

Each cross was set in bottles of standard medium, with or without tetracycline (Section 2.2.2). Total males and females were counted in at least three replicate bottles of at least 200 flies, and the proportion females relative to total flies established (Appendix B). Note that for C48 crossed to *arm-tTA* with 10 μg/mL tetracycline, two of the three triplicates contained only around 100 flies. Statistical analyses were conducted with Minitab 15. The analysis of variance was conducted on lines C48 and C50 using a general linear model that treated 'line' as a random factor. Bonferonni, Sidak, and Dunnett's post-hoc simultaneous tests were conducted to compare individual 'treatments' to the situation without *arm-tTA* (as a 'control'). All three tests gave similar results.

2.7 Polytene chromosome immuno-fluorescence

Polytene chromosomes were immuno-stained according to a modified protocol of Zink and Paro (1989). Wandering third instar larvae were collected from stock bottles at five to seven days incubation at 25 °C, or eight to nine days at 18 °C, and washed in Ringer's solution (Section 2.2.1) on ice. Salivary glands were dissected out in PBT. Most solution was removed with Whatman 3MM paper, and the glands treated with Fix 1 for a maximum of 30 seconds. Fix 1 was replaced with Fix 2 for 2.5 to 3 minutes. Individual glands were carried in 2.5 μL of Fix 2 with a P20 autopipette to a siliconised coverslip (two per coverslip), and a microscope slide gently lowered on top. The slide was turned coverslip up, and struck sharply at a low angle with the eraser end of a pencil, moving the coverslip 2 to 3 mm. This was repeated if the nuclei of the glands mostly did not break open. Optionally, prior to the strike with eraser

end, the coverslip was tapped with the lead end of a pencil 3 to 4 times in each corner, and over the glands, then a zigzag pattern applied diagonally from lower left to upper right corner. This step is suggested by C. H. James et al. in protocol PROT01 of the Epigenome online protocols (http://www.epigenome-noe.net). In all cases, the coverslip was then covered with a lab tissue, and firmly squashed with my thumb.

Slides containing good chromosome spreads were frozen in liquid nitrogen, and the coverslips removed with a scalpel blade. After one hour dehydration in ethanol at -20 °C, the slides were washed twice in PBT for 30 minutes. Spreads were blocked with 100 μL blocking solution for one hour in a humid chamber, with a square of lab film applied to prevent evaporation. Slides were briefly drip-dried, then incubated in a similar manner with 50 μL primary antibody overnight at 4 °C. The spreads were washed three times with PBT, and treated with 100 μL secondary antibody in the dark at room temperature. Slides were washed a further three times with PBT, then mounted in 15 μL of VECTASHIELD mounting medium containing DAPI (Vector Laboratories H-1200), and the coverslip sealed with nail varnish.

All antibodies were diluted in blocking solution. For the MSL1 staining of line C47 in Figure 3.5, and all spreads of Figures 3.4 and 3.7, the primary antibody was a 1:100 dilution of rabbit anti-MSL1 (FLR3618, Mitzi I. Kuroda[4]), and the secondary antibody was a 1:1000 dilution of donkey anti-rabbit IgG AlexaFluor 594 fluorescent conjugate (Molecular Probes A-21207; excitation maximum 590 nm emission maximum 617 nm). For the MSL2 staining of M27xC47 in Figure 3.5, the primary antibody was a 1:50 dilution of rabbit anti-MSL2 (FLR3616, Mitzi I. Kuroda), and the secondary antibody was the same anti-rabbit IgG AlexaFluor 594 fluorescent conjugate. For the HA staining of M27 x C47 in Figure 3.5, the primary antibody was a 1:50 dilution of rat anti-HA fluorescein conjugate (Roche 11 988 506 001, excitation maximum 488 nm, emission maximum 515 nm), and the secondary antibody was a 1:600 dilution of rabbit anti-rat IgG FITC conjugate (SIGMA F-1763, excitation maximum 495 nm, emission maximum 525 nm).

Slides were viewed on a BX51 (Olympus) compound microscope, with U-LH100HGAPO mercury UV-burner, and U-RFL-T controller. DAPI stain was observed with the U-MWU2 (Olympus) filter set, excitation 330 to 385 nm, emission 420 nm. AlexaFluor 594 was observed with the U-MWIG2 (Olympus) filter set, excitation 520 to 550 nm, emission 580 nm. FITC was observed with the U-MWIBA2 (Olympus) filter set, excitation 460 to 490 nm, emission 510 to 550 nm. Monochromatic images from the individual filters were captured with the Magnafire camera and software (Optronics), and were artificially coloured, merged, enhanced, rotated and cropped in Adobe Photoshop 9.0.2. Individual monochromatic DAPI and fluorescent antibody images were aligned in Adobe Illustrator 12.0.1, and the colours inverted. Ticks were drawn along the antibody image at each site bound, then overlayed on the DAPI image, and printed.

[4]Harvard Medical School

Cytological bands were identified by hand on DAPI images by comparison with the drawings of Bridges, and the electron micrographs of Sorsa and Saura (Bridges, 1935, 1938; Sorsa and Saura, 1980*a*,*b*; Sorsa et al., 1983; Saura et al., 1993). The location of each fluorescent antibody band was inferred from the location of the corresponding tick mark. Comparisons with Demakova et al. (2003) were made on the assumption that an observed band at *e.g.* 12E was the same as that reported at *e.g.* 12E1–4.

PBT

0.2 % Triton X-100 in PBS (Sambrook et al., 1989).

Fix 1

10 % Triton X-100, 3.7 % formaldehyde in PBS (Sambrook et al., 1989). Warmed to disperse Triton X-100, then kept on ice. Made fresh hourly.

Fix 2

50 % glacial acetic acid, 3.7 % formaldehyde. Made fresh hourly.

Siliconised coverslips

New coverslips immersed in SIGMACOTE (SIGMA SL2), air-dried, and stored in absolute ethanol at room temperature.

Blocking solution

0.2 % BSA (New England Biolabs) in PBT

2.8 Eye size assays

Crosses to produce heterozygous flies were *y w* (lab stock #4) males by virgin females from the transgenic line. More detailed crossing schemes are described in Section 2.2.4. Each cross involved eight to twelve pairs of flies on standard food in 100 mL bottles (Schott) at 25 °C (Section 2.2.2). Adults were removed at four to six days, and the progeny assayed at 15 to 21 days, most commonly 18. Adults were assigned into the indicated eye categories by hand, prior to separation on basis of sex. Number of flies in each category are expressed as a percentage of the total population of that sex in that culture bottle. The 95 % confidence intervals were calculated with the STDEV and CONFIDENCE functions of Microsoft Excel 11.4.1. Representative eye images were captured as described (Section 2.11.3).

2.9 β-galactosidase enzymatic assays

Crosses to produce flies heterozygous for a transgene were *y w* (lab stock #4) males by virgin females of the transgenic line. For the pairing assay (Figure 5.3), the reciprocal crosses were males of one homozygous stock by virgin females of the other. For the crosses with *tetO-GMR-lacZ* (stock C100) in Figure 5.9, M27, EW140, or EW162 males were mated to virgin C100 females. Only Cy^+ offspring were assayed, as the EW140 stock had possible *s1-tTA/CyO* genotype flies. More detailed crosses are described in Section 2.2.4. Generally, eight to twelve pairs of flies were set on standard food in 100 mL bottles (Schott) at 25 °C (Section 2.2.2), in triplicate. If tetracycline (at 10 μg/mL) was added to any samples, all samples were foil-wrapped. Parents were removed at about four days (for twelve pair), or six days (for eight pair), and newly emerged flies collected over a 24 hour period around eleven days. Flies were aged three to five days at 25 °C before the assay.

2.9.1 β-galactosidase assay

The assay procedure was based on the procedure of Simon and Lis (1987). Briefly, nine females or twelve males (*arm-lacZ* lines) were hemisected to remove wings and abdomens, or 15 female or 20 male heads were colleceted (*GMR-lacZ* lines), and transferred to 1.5 mL (eppendorf) microfuge tubes on ice. Flies were homogenised in 1 mL phosphate assay buffer in 1 mL glass homogenisers (Kimble Kontes 885450-0020), and transferred back to the microfuge tubes. A one minute centrifugation separated the debris from the lysate, and the homogenate was transferred to new tubes. Triplicate aliquots of the homogenate were transferred to new tubes, 1 mM CPRG in phosphate assay buffer was added to make up 1 mL. The CPRG was added to each tube at 30 second intervals to facilitate spectrophotometer readings, and the tubes were then transferred to a 37 °C water bath. The absorbance at 574 nm was read for each sample in a quartz cuvette at 30 minutes and again at one hour incubation. A blank reading of phosphate buffer and CPRG was subtracted from each measurement. CPRG (Roche 10 884 308 001) was prepared from a 10 mM stock solution (in phosphate buffer), stored at -20 °C.

Phosphate assay buffer

50 mM Potassium-phosphate buffer, pH 7.5 (Ausubel et al., 1994), autoclaved. 1 mM $MgCl_2$ added after sterilisation, stored at 4 °C.

2.9.2 Total protein assay

Total protein concentration was measured with the Protein Assay Dye Reagent (Bio-Rad 500-0006), according to the supplied instructions. A standard curve of absorbance at 595 nm against

microgram protein was first prepared several times from 0 to 18 mg/mL bovine γ-gobulin (Bio-Rad 500-0005), and the average slope and Y intercept determined. Triplicate (15 μL if hemisected adults, 75 μL if heads) aliquots of the homogenate were taken, and 1 mL of 1x Dye Reagent added. After a 30 minute room temperature incubation, the absorbance at 595 nm was read and a blank subtracted. The protein concentration was determined from the standard curve values, and averaged to obtain a mean value for each homogenised sample.

2.9.3 Enzymatic activity calculations

An average sample absorbance at one hour was calculated from an assumed zero time absorbance of 0, and the 30 and 60 minute readings at 574 nm, using the TREND function of Microsoft Excel 11.4.1. The activity was then expressed as optical density per minute per milligram total protein, and averaged across the triplicate homogenate samples. Mean male values, female values, or ratios between the sexes were calculated from triplicate experimental populations. The standard errors (SEs) were calculated as standard deviation over the square root of the sample size (3).

2.9.4 Statistical analyses

Statistical analyses were conducted in Minitab 15 from raw triplicate values, or mean line ratios, as appropriate. Satterthwaite's approximation was applied for all *t*-tests, so as not to assume equal variances between samples. Three lines each of *DRE-GAGA-GL* and *GMR-lacZ* were compared at 2A with a two-sample, two-tailed Student's *t*-test on the line ratios, or the $\log_2$ values of the ratios. For *GiL*, $\log_2$ values of the raw triplicate ratios for the single line were compared in a one-sample, two-tailed Student's *t*-test to the mean of the three *GMR-lacZ* lines (0.333). The post-hoc analysis of variance on transvection effects for *arm-lacZ*, *GMR-lacZ*, and *DRE-GAGA-GL* was conducted with a general linear model of 'Construct' and 'Location', where 'Location' was treated as a random variable. Subsequent one-sample, two-tailed Student's *t*-tests were conducted with the data pooled for all *attP* locations, against a ($\log_2$) value of 1.0 (no transvection effect). The effects of *hsp83-msl2* on *DRE-GAGA-GL* and *GiL* lines were analysed with two-sample, two-tailed Student's *t*-tests against *GMR-lacZ*, using $\log_2$ transformations of the raw triplicate ratios as data.

2.9.5 Homogenate volume

Depending on the level of *lacZ* expression, different volumes of homogenate were used in the β-galactosidase assay to ensure the spectrophotometric absorbance readings were within the linear portion of the standard curve. The volume was accounted for in calculations. The

required volumes for each line are listed in Table 2.4. To determine the required volume, either a rough assay of un-aged, mixed-sex flies was conducted, or (more commonly) samples of the homogenate were tested towards the start of a full assay procedure. Both approaches involved 30 minute incubations only and were not repeated in triplicate.

2.10 β-galactosidase tissue staining

Flies of genotype *y w* were lab stock #4; *arm-lacZ*, V42; *arm-tTA*, M27; and *tetO-GMR-lacZ*, C100. Twenty pair of flies from #4 or V42, and crosses of #4 males by C100 virgin females, and M27 males by C100 virgin females, were set on standard food with or without tetracycline at 10 μg/mL in foil-wrapped 100 mL bottles (Schott) at 25 °C (Section 2.2.2). Parents were removed at about 4 days. Climbing third instar larvae were collected at 11 days, and washed in Ringer's solution (Section 2.2.1) on ice.

The staining procedure was adapted from Glaser et al. (1986), with only slight changes in the dissection, washing, and overnight incubation steps, and in the constitution of staining solution and glycerol concentration. Briefly, larvae were dissected in Ringer's solution in one cavity of a twin cavity (SUPERIOR MARIENFELD 13 201 02) microscope slide, Ringer's removed with Whatman 3MM paper, and fixed for 15 to 20 minutes. Two washes were performed with 1 x PBS (Sambrook et al., 1989), and tissue transferred to 150 μL staining solution in 96-well plates. The plate lid was sealed with lab film, and incubated overnight (21 to 23 hours) in the dark at 30 °C. Stained tissue was mounted in 90 % glycerol in PBS, and photographed (Section 2.11.3).

Fixitive

10 mM Sodium-phosphate buffer, pH 7.0 (Sambrook et al., 1989), 150 mM NaCl, 1 mM $MgCl_2$, and 1 % glutaraldehyde. A stock solution lacking glutaraldehyde was prepared and stored at -20 °C; glutaraldehyde was added immediately prior to use.

Staining solution

10 mM Sodium-phosphate buffer, pH 7.0 (Sambrook et al., 1989), 150 mM NaCl, 1 mM $MgCl_2$, 3.1 mM $K_3Fe(CN)_6$, 3.1 mM $K_4Fe(CN)_6 \cdot 3\,H_2O$, and 0.2 % 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal). A stock solution lacking the potassium salts and X-gal was prepared and stored at -20 °C; these were added immediately prior to use.

Table 2.4: Homogenate volume required for β-galactosidase assays.

Sample	Homogenate volume (μL)	
	1 copy[1]	2 copies
V40	50	
C66	50	25
C67	50	25
C68	50	25
C76	50	25
C78	50	25
C81	50	25
C84	100	50
C91	50	12.5
C95	25	12.5
C96	50	25
C98	50	25
C100	50	25
C109	50	50
C111	100	25
C112	25	12.5
C113	100	50
C66xC67		100
C84xC98		100
C76xC78		25
C96xC109		25
C66/hsp83-msl2	50[2]	
C76/hsp83-msl2	25[2]	
C84/hsp83-msl2	100[2]	
C96/hsp83-msl2	50[2]	
M27xC100	10	
M27xC100+tet	25	
EW140xC100	25	
EW162xC100	25	

[1]Number of copies of *lacZ* reporter construct
[2]Same volume for the control progeny

2.11 Drosophila photography

2.11.1 Embryonic preparation

To obtain staged embryos, eggs were collected over a one hour period from populations of the appropriate stock. One- to two-hundred pairs of flies were tapped into polystyrene beakers (45 mm diameter, 90 mm deep), atop a 55 mm x 14 mm (Plastiques GOSSELIN BP53-02) petri dish of fresh egg-laying media streaked with a light yeast paste (Section 2.2.2), and incubated at room temperature under ambient light. Unlaid eggs were purged for one hour, and discarded, before hourly collection periods. Collected embryos were incubated on the laying plates at room temperature in a chamber (QNA International MIC-101) under pressurised oxygen and high humidity for two to three hours. Eggs were then de-chorionated by rolling with forceps on double-sided tape, aligned dorsal up, anterior left, on double-sided tape in one cavity of a twin cavity (SUPERIOR MARIENFELD 13 201 02) microscope slide, and covered with Halocarbon 27 oil (SIGMA H8773) to a level flush with the slide surface. The slides were returned to the pressurised, humid, high-oxygen environment, and incubated at room temperature until the appropriate photography time.

2.11.2 Pupal dissection

Unstaged pupae were removed from the walls of emptied stock bottles with fine forceps. To obtain pupal imagos, the anterior section of pupal casing was first removed with a scalpel. Fine forceps were slid between the imago and the casing along the ventral side, and a slice with a 21 guage needle was made through the casing against the forcep arm. The casing was peeled away with a second pair of forceps, taking care not to disfigure the imago at the posterior end. Pairs of imagos were aligned dorsal up, anterior left, on a microscope slide for photography.

2.11.3 Photography

Most photographs of *Drosophila* adult eyes or heads, larvae, embryos, pupal imagos, and dissected larval tissues were taken with an Olympus DP-70 camera connected to the Olympus SZX12 stereo zoom microscope by a U-TVO 0.5 x adapter. Images were captured through a 108 x objective with either the Olympus DP Controller 2.1.1.183 or Soft Imaging System analySIS FiVE 5.0 software packages, and overlayed with a scale bar previously calibrated with a stage micrometer. Images of adult eyes from the *(roX1DHS)9-GMR-hid* lines were captured with the Magnafire camera and software (Optronics).

Embryos for the *GMR-GFP* expression test were photographed with transmitted white light at 1/200 second exposure, ISO 200, after an equal white balance with white card. Green fluorescent images were taken under the GFPA filter with a 3.5 second exposure, ISO 200, black

cutoff of 170.

Embryos for the *n1-tTA* and *s1-tTA* expression test were photographed with transmitted white light at 1/100 second, ISO 200, and a white balance of 1.0/1.0/0.9 with white card. YFP expression was detected with the GFPA filter at 3.5 second exposure, ISO 200, black cutoff of 160.

Pupal imagos and adult heads to test *GMR-GFP* expression were photographed under reflected white halogen light, through the 60 x objective, with variable exposures, and a white balance of 0.8/1.0/1.7 from white filter paper. Green fluorescent images of imagos were captured with the GFPA filter at 1/12 second exposure, ISO 200, and a black cutoff of 0. Fluorescent images of heads were captured at 2.0 second exposure, ISO 200, black cutoff of 32.

Larval tissue for detection of β-galactosidase was photographed with transmitted white light at 1/90 second exposure, ISO 200, and a white balance of 1.0/1.0/0.9 from white card.

Adult eyes for representative sizes or colours were photographed under reflected white halogen light, at an automatic exposure time, ISO 200, white balance with white card.

Images were oriented and cropped with Adobe Photoshop CS2 9.0.2. The colour balance of the photos from the *(roX1DHS)9-GMR-hid* lines was altered to attain an appearance resembling the other eye observations, as the camera and software provided no inbuilt balance. Individual photos of eyes from *18DGH* and *GMR-hid* lines were likewise adjusted to match the average appearance. The colour balance was either un-modified or co-ordinately modified for all larval tissue, and pupal images, and images of adult heads for *GMR-GFP* expression and comparison of position-dependent eye colours.

Chapter 3

Development of an inducible dosage compensation system in female Drosophila

The establishment of *Drosophila* dosage compensation is poorly understood. Early embryos are difficult to stage, and provide poor chromosomes for cytological analysis. It was thus proposed to study the initiation of dosage compensation after artificially inducing the process in females (Section 1.4.1). The large polytene chromosomes of third instar female larvae would be suitable for cytological observations. After initially constructing the induction system, its validity was determined against three conditions:

1. Absence of MSL complex prior to induction
2. High levels after induction
3. A consequent chain of events resembling those of the male system

3.1 A two-component system for MSL2 induction was created

In order to artificially induce dosage compensation in female *Drosophila*, a tetracycline-regulated system (Section 1.3.2) was designed that would allow induction of *msl2* expression, and consequent up-regulation of genes on the female X chromosome. As several lines were available that constitutively-expressed a tetracycline-controlled transcriptional activator (tTA), only the *tetO-msl2* component needed be created. A line expressing *tetO-msl2* was available (Scott et al., 2004), but was not optimal for immuno-precipitation studies. To allow efficient immuno-precipitation of MSL2, epitope tags were incorporated into a *tetO-msl2* cassette, and assembled in a vector suitable for *Drosophila* transformation. Additionally, flanking insulator sequences were added to prevent possible positional effects of repressive chromatin at insertion sites.

The *tetO-msl2* constructs were assembled in the HF10 *Drosophila* transformation vector, which already contained SCS and SCS′ insulator sequences (Section 2.3.11). The *Not* I site in pHF10 was first destroyed by digestion with *Xba* I, and re-ligation at a low concentration. The modified plasmid was re-named pCL04 (Appendix I). A 1.2 kb *Xho* I–*Hind* III fragment of plasmid W.T.P-2, containing seven repeats of the tetracycline operator (tetO), 5′ *P* minimal promoter, *hsp70* leader sequences and *hsp70* 3′ UTR, was excised. The cohesive ends were filled in with Klenow fragment to create blunt ends, and the fragment inserted in the *Eco* RI site of pCL04, treated likewise to have blunt ends, to create pSTPS′ (Appendix I). A clone with the tetO expression cassette in an inverse orientation was also isolated and retained as pS′TPS.

Three expression constructs were created to encode MSL2 with different C-terminal epitope tags: a single copy of the 9 amino acid (YPYDVPDYA) hemaglutanin (HA) tag, three copies of HA, or six histidine residues (HIS) and a biotinylation activation domain (BAD). The short (17 amino acid) BAD peptide is an effective target for biotinylation from the *E.coli* BirA enzyme

in mammalian cells and *Drosophila* (Viens et al., 2004; Mito et al., 2005). A *BirA* gene was included in the same transcript as the tagged *msl2*, following an internal ribosome entry site that facilitated high biotinylation efficiency in mammalian cells (Viens et al., 2004). Using the pXL1 plasmid as a template for PCR (Section 2.3.10), the epitope tags were added to the C-terminus of MSL2, immediately before the endogenous stop codon (Figure 3.1). pXL1 contains *msl2* cDNA from Kelley et al. (1995) that lacks UTR regions necessary for Sxl repression (Li, 2002).

The HA tag was incorporated in primer 5 ("Nhe/Not/HA/M2/asen"), followed by a new stop codon. A PCR reaction with primers 1("pXLsense4013/pr1") and 5 yielded a single product of the expected 600 bp size (not shown). The *Nar* I–*Not* I fragment of pXL1 was replaced with the *Nar* I–*Not* I fragment of the PCR product containing the modified *msl2* to create pM2HA (Appendix I). The plasmid was sequenced with the M2intseq, pXLsense4013/pr1, and T7 primers to obtain over 1 kb of sequence covering the 3′ end of *msl2*, the HA tag, and *msl2* 3′ UTR, which confirmed that no mutations had been introduced by PCR. The 2.37 kb *Eco* RI–*Not* I fragment of pM2HA was inserted in the *Eco* RI and *Not* I sites of the pSTPS′ expression vector to create pTM2HA (Appendix I).

To create the three repeats of HA, the NheHAXHANotTop and NheHAXHANotBot oligonucleotides (Table A.2) were annealed to create a short fragment containing two additional copies of HA followed by a new stop codon, and with single-strand overhangs to create cohesive *Nhe* I and *Not* I ends (Figure 3.1c). The short *Nhe* I to *Not* I section of pTM2HA containing a stop codon was replaced with the new fragment to create pTM2HA3 (Appendix I). Sequencing of the plasmid with the pXLsense4013/pr1 primer revealed problems in the second HA epitope tag. A single C, and a dinucleotide TG, were absent, which would translate to YP*TM*PDYA instead of YP*YDV*PDYA. This mutation was most likely due to errors in the synthesis of the oligonucleotides containing the tag, but as the result was essentially a frameshift mutation that left the two other HA epitopes and stop codon in frame and correct, the mutated clone of pTM2HA3 was not discarded. Thus in all experiments the 3 x HA epitope tag actually represents only two complete copies of the HA epitope.

To create the Histidine and Biotin epitope tag, parallel PCR reactions were conducted (Section 2.3.10) to amplify the 3′ end of *msl2* from pXL1 (with primers pXLsense4013/pr1 and HIS/M2/ASENSE/pr2), and the histidine epitope, BAD, and *BirA* cassette of pBBHC (with primers M2/HIS/SENSE/3 and NotI/BirA/asense/4). Both reactions yielded single products of the expected sizes (600 bp and 1 100 bp, respectively, not shown), and were co-purified from the agarose gel (Section 2.3.3) in a final 30 μL volume. By designing 15 base overhangs at the 5′ end of primers HIS/M2/ASENSE/pr2 and M2/HIS/SENSE/3 that were complementary to the other PCR amplicon, a second PCR with primers pXLsense4013/pr1 and NotI/BirA/asense/4, using 2 μL of the combined products from the first round reactions as template, amplified a

fused 1.6 kb fragment of the 3′ end of *msl2*, HIS, BAD, and *BirA* (Figure 3.1c, not shown). The *Nar* I–*Not* I fragment of pXL1 was replaced with the *Nar* I–*Not* I fragment of the 1.6 kb PCR product containing the modified *msl2* to create pM2HisB (Appendix I). The plasmid was sequenced with the M2intseq, pXLsense4013/pr1, and T7 primers to obtain over 1.5 kb of sequence including the 3′ end of *msl2*, HIS epitope tag, BAD, and *BirA* gene up to and including the *Not* I site, which confirmed that no PCR-induced errors were present. The 3.43 kb *Eco* RI–*Not* I fragment of pM2HisB was inserted in the *Eco* RI and *Not* I sites of the pSTPS′ expression vector to create pTM2HisB (Appendix I).

P element transformations of *y w Drosophila* were conducted (Section 2.2.5) with pTM2HA, pTM2HA3, and pTM2HisB to yield the transgenic lines listed in Appendix H. Two transgenic lines were isolated that carried the *tetO-msl2-HA* construct, both on chromosome 3. Three transgenic flies were isolated from the pTM2HA3 injection, however genetic analysis (Section 2.2.5) revealed one was carrying two insertions of the *tetO-msl2-HA3* on different chromosomes. By breeding to separate the double insertion, a fourth line was obtained. Of these four lines, three carry insertions on the X, and one on the second chromosome. From the pTM2HisB injection, three transgenic flies with one insertion, and three with two separable insertions, resulted in nine lines carrying *tetO-msl2-HisB*. The insertions were on chromosomes X, 2, or 3.

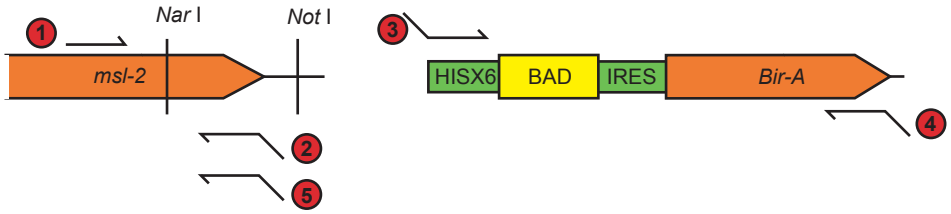
3.2 MSL2 was produced in females

3.2.1 MSL2 protein levels below detection limit

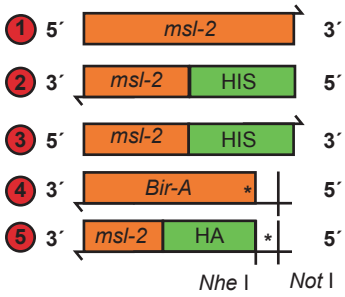
In order to confirm production of MSL2 in the transgenic lines, crude protein extracts were prepared from female samples of several TM2HA and TM2HA3 lines, and analysed by western blot (Section 2.5). To induce expression, the *tetO-msl2* lines were crossed to a line that expressed high levels of tTA in female fat cells from the *Yolk protein 1* (*yp1*) promoter (Heinrich and Scott, 2000), and raised in the absence of tetracycline.

Two different antibodies were first assessed for their ability to detect MSL2: an antibody raised against an MSL2 peptide (Genscript), and one raised against the full-length MSL2 (Kelley et al., 1995). Western blots were performed on the above samples, in comparison to extracts from *y w* females that should produce no MSL2, and from females of a *hsp70-msl2* line (lab stock #S13) that produces high levels of MSL2 (Li et al., 2008), heat shocked to induce expression. The peptide antibody detected a single band of the wrong size in all samples (not shown), including the negative control, and was thus unsuitable. However, the antibody against full-length MSL2 detected a single band of about 110 kDa only in the sample of heat-shocked *hsp70-msl2* (not shown). This blot was performed by Dr. Fang Li[1]. The size of this band is in

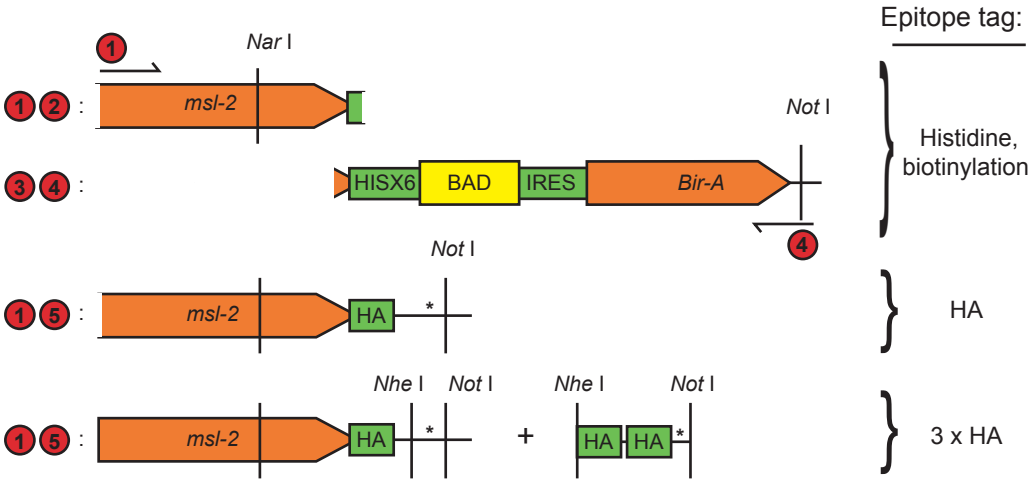
[1]Massey University, Palmerston North, New Zealand



(a) Primer positions



(b) Primer detail



(c) PCR products

Figure 3.1: Scheme to add epitope tags to *msl-2*. Epitope tags were appended to the C-terminus of MSL2 by PCR with the indicated primers, then replacing the *Nar*-I to *Not*-I fragment of the wild-type *msl-2* with the altered forms. The histidine epitope (HIS) and Biotinylation activation domain (BAD) were added with two rounds of PCR, first with primers 1+2 and 3+4, then with primers 1+4 on the combined products of the first round. One copy of the HA tag was introduced in primer 5. Two additional HA tags were later introduced in a linker that replaced the *Nhe*-I–*Not*-I fragment to create the HA3 construct. IRES, internal ribosome entry site; *, stop codon.

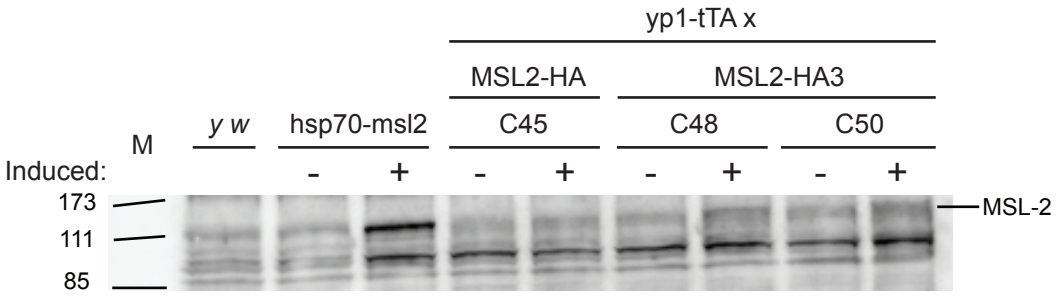


Figure 3.2: Un-detectable induction of MSL2 by tTA. Western blot of MSL2 expression in protein extracts from females of the indicated fly lines, with the full-length antibody (Kelley et al., 1995). *y w* (lab stock #4) females should not express MSL2. hsp70-*msl2* flies produce high amounts of full-length MSL2 with (+) heat shock. One line of tetO-MSL2-HA (C45), and two of tetO-msl2-HA3 (C48, C50) are tested with or without repression by tetracycline. Absence of tetracycline, or heat treatment are the induced (+) states. M is Invitrogen Benchmark™ protein ladder, sizes in kDa.

agreement with the 115 to 130 kDa reported for MSL2 (Bashaw and Baker, 1995; Kelley et al., 1995; Zhou et al., 1995).

Western blots were thus performed with the antibody raised against full-length MSL2, on female samples from the TM2HA and TM2HA3 lines, crossed to *yp1-tTA* and raised without tetracycline (Figure 3.2). These samples were compared to extracts from *y w* females, and heat-shocked *hsp70-msl2* females, as negative and positive controls for MSL2 expression, respectively. Un-induced samples are from *hsp70-msl2* without heat-shock, or from *tetO-msl2* lines crossed to *yp1-tTA* but raised with tetracycline. Although a strong band at the expected size of MSL2 was detected in heat-shocked *hsp70-msl2*, very low levels of MSL2 were detected in any of the *tetO-msl2* lines. A secondary band at 90 to 100 kDa is likely to be a protein known to be non-specifically detected by this antibody (Kelley et al., 1995). Two further bands were evident at about 55 to 70 kDa, that are not shown with the original antibody report (Kelley et al., 1995), but similar bands have been noted in detections of MSL2 with another antibody (Bashaw and Baker, 1995, 1997). The age of the antibody used may also have been a factor in the poor detection of MSL2. Given the background of non-specific bands, it was difficult to determine if any MSL2 was expressed from the induction lines. If any was expressed, it was below the detection limit of this technique.

3.2.2 MSL2 expressed sufficient to kill females

As ectopic expression of *hsp83-msl2* in females kills about 80 % of the flies (Kelley et al., 1995), it was reasoned that expression of the *tetO-msl2* might be detected genetically as a decreased female viability. The *tetO-msl2* lines could be crossed to a line that constitutively expresses tTA from an *armadillo* (*arm*) promoter, *arm-tTA* (Scott et al., 2004), and the female survival

compared to that of the *tetO-msl2* stock alone. By raising flies with tetracycline in the diet, the expression of *msl2* could also be specifically inhibited.

However, because it appeared that the inclusion of tetracycline in the food had a detrimental effect on fly survival, development time, and health, the effect of different concentrations of tetracycline on the *y w* injection line was first established. By raising triplicate populations of *y w* on instant food with tetracycline concentrations of 0, 1, 3, 10, 30, and 100 μg/mL (Section 2.2.2), and counting the number of offspring, it was observed that high tetracycline concentrations did indeed lower the number of offspring (not shown). Whereas 1 μg/mL tetracycline had no effect, 3 μg/mL produced 71 % of the population recovered from 0 μg/mL, and 10μg/mL produced only 45 %. Both 30 and 100 μg/mL resulted in about 35 % the expected population sizes. A similar toxicity effect was observed for concentrations of doxycycline $\geq$ 10 μg/mL on *y w* or *act5C-rtTA* (Table A.1) fly strains (not shown). To ensure large enough samples for statistical analysis, it was decided to attempt the inhibition of the tTA with levels of tetracycline no higher than 10 μg/mL. This value is consistent with the concentration previously used to inhibit tTA expressed from the *armadillo* promoter (Scott et al., 2004).

To confirm an equal survival of each sex without induction of *msl2*, populations of several *tetO-msl2* lines were raised in triplicate, along with *y w*, and the mean female viability established as a proportion of total flies (Figure 3.3, series without *arm-tTA*). Curiously, the two HA- and two histidine-tagged lines, and one of the four lines with three HA tags, showed a decreased male viability (not shown). The possible presence of uncontrolled factors that affected the normal male:female ratio rendered these lines unsuitable for further analysis.

To test for induction of MSL2, two of the remaining lines (C48, and C50) with three HA tags were crossed to *arm-tTA*, and raised with or without tetracycline (Section 2.6). The number of females as a proportion of total flies was established for crosses on 0, 1, and 10 μg/mL of tetracycline (Appendix B, Figure 3.3). A slight decrease in female viability with induction by *arm-tTA* (but without tetracycline) indicated that levels of MSL2 were being induced sufficient to kill females. An analysis of variances was conducted for both lines C48 and C50. Post-hoc comparisons between flies with *arm-tTA* (without tetracyline), and those without showed that the effect, although slight, was very highly statistically significant (T = -5.041, p = 0.0001). The effect was specific to production of MSL2, as when the crosses were raised in the presence of 10 μg/mL tetracycline, no significant female death was observed ($T = 1.904$, $p = 0.1641$ with Dunnett simultaneous test). Note that the estimated values for C48 x *arm-tTA* with 10 μg/mL of tetracycline are less accurate due to low total fly numbers resulting from this cross. An intermediate level of tetracycline (1 μg/mL) gave variable results, depending on the genomic position of transgenic insertion.

Overall, the levels of MSL2 expressed from *tetO-msl2* lines were likely too low to be detected by western blot, but were sufficient to mediate a decrease in female viability consistent

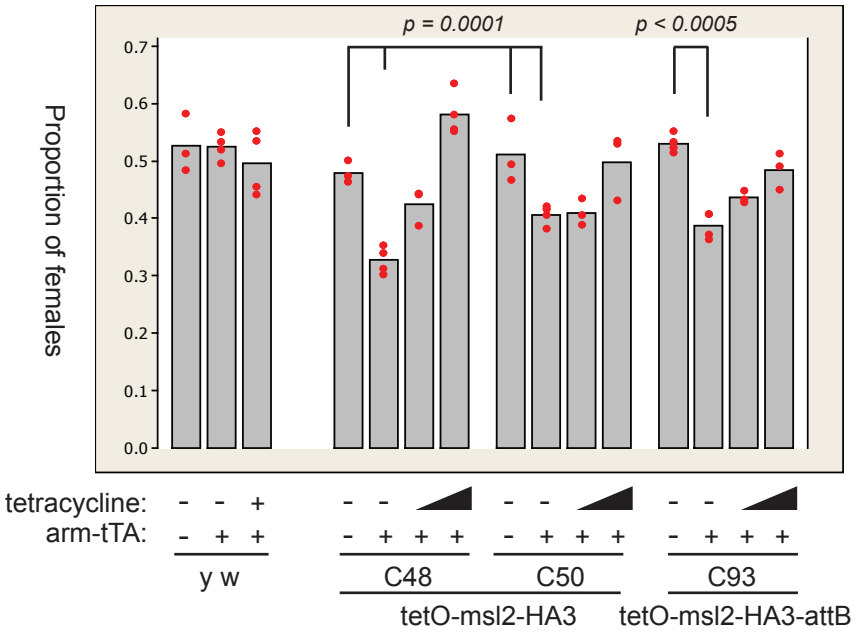


Figure 3.3: Induction of MSL2 sufficient to lower female viability. The proportion females per total flies are shown for populations of flies carrying the indicated constructs, and with (+) or without (-) one copy of *arm-tTA*. *arm-tTA* was inhibited with tetracycline at 1 or 10 μg/mL. The responses from the injection line (*y w*) are provided as a negative control. Individual values from at least three independent experiments are plotted as red dots, with bars representing the means. *p* values are from post-hoc comparisons following an analysis of variance on combined C48 and C50 lines, and from a Student's *t*-test for C93.

with aberrant dosage compensation.

3.3 Transgenic MSL2 localised to the female X chromosome

To confirm that the induced MSL2 resulted in MSL complex assembly in females, the large polytene chromosomes of the larval salivary glands were probed with fluorescent antibodies against MSL proteins (Section 2.7). The absence of MSL binding on the female X chromosome without induction of *tetO* was first established by examining female larvae of several *tetO-msl2* lines (Figure 3.4). Unfortunately, bound MSL1 was detected at a number of X chromosome sites in all lines tested, and to most of the X chromosome for the histidine-tagged line C51. Immunofluorescence with antibodies against MSL2 or the HA epitope exactly co-localised with that against MSL1, but no binding was ever observed with an antibody against the HIS epitope (not shown). Even presuming a male larvae was accidentally examined for the C51 spread, the amount of binding in all other lines represented a significant obstacle for

the planned experiments. One *tetO-msl2-HA3* line (C47) was selected for further analysis, and the chromosome spreads examined in greater detail to identify the position of all MSL1 bands (Section 2.7). In comparison to the minimal set of MSL "high affinity" sites (Section 1.1.4) as listed in Table 1 of Demakova *et al.* (2003), the two data sets appear nearly identical, differing mostly in the relative intensity of individual bands (not shown). The most obvious high affinity sites not bound with *tetO-msl2-HA3* were at cytological positions 5D6–8 and 12E1–4, but in contrast *tetO-msl2-HA3* induced weak MSL sites at several positions not listed as high affinity sites, most notably 14A.

The surprising observation that un-induced *tetO-msl2* could still produce MSL binding prompted a re-evaluation of the line that expressed the un-insulated *tetO-msl2* (Table A.1, line M32). It was previously reported that the MSL binding to the X chromosome mediated by *arm-tTA* was completely abolished with repression by tetracycline (Scott et al., 2004). However, examination of the MSL binding pattern in female larvae of the *tetO-msl2* line also revealed several strong bands of bound MSL complex on the X chromosome, in a pattern similar to that observed in lines carrying the epitope-tagged, insulated constructs (not shown).

As MSL complex was produced without induction of *tetO-msl2* by the tetracycline system, or indeed without even the tetracycline trans-activator, there must exist a low level of 'leaky' *msl2* expression from the transgenic 5′ *P* basal promoter. This leakiness has also been noted with tetO-mediated expression of luciferase (Yokoyama et al., 2007). Alternatively, if the SCS and SCS′ insulators were inadequate to prevent the action of surrounding genomic elements, *msl2* expression would be possible in situations where the transgene lay near to enhancer sequences. As each of the six lines analysed represented a different genomic insertion of *tetO-msl2*, this possibility seemed unlikely. Whatever the cause, the low amounts of MSL2 produced were sufficient to form MSL complex capable of binding the X chromosome. This complicated any study of the induction of dosage compensation, as to some degree there appeared to be always MSL complex in these females.

As discussed above, female lethality was not affected by the *tetO-msl2* alone, but was clearly increased with additional induction by the tetracycline system. The consequent increase in MSL complex with *tetO-msl2* induction might also have been visible with chromosome immunofluorescence. To investigate this possibility, chromosomes were prepared from the larval salivary glands of female flies carrying one copy of *tetO-msl2-HA3* (from line C47) and one copy of *arm-tTA*, and the amount of bound MSL complex compared to the background binding pattern (Figure 3.5). Several more sites of bound MSL complex are observed in the presence of the tTA activator, and the relative intensity of the background sites appears increased. The increased binding is specific to production of MSL2, as inhibition with tetracycline results in a pattern resembling the background level of binding. As the number of sites bound in the induced situation is significantly less than the many hundreds of sites the MSL complex binds on

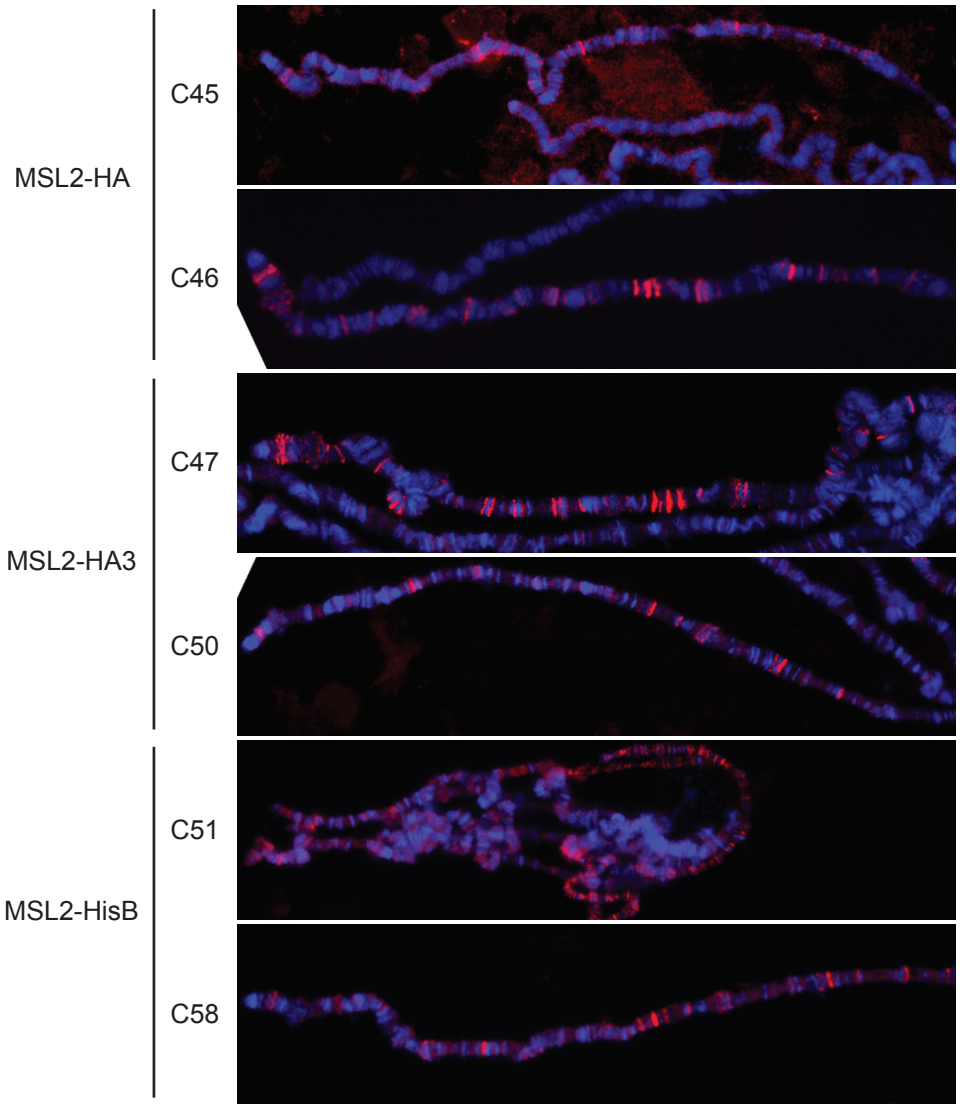


Figure 3.4: Background level of MSL binding in transgenic females. Representative combined DNA (DAPI, blue) and MSL1 (red) fluorescent images of X chromosomes from female larval salivary gland nuclei of the indicated transgenic lines. C46 and C58 carry two copies of the transgene; all other lines have only one copy.

a wild-type male X chromosome, the levels of transgenic MSL2 produced are likely lower than in wild-type males.

3.4 Repressive chromatin effects reduced un-induced chromosome binding

Any study of the initiation of dosage compensation with the induction of *msl2* in females necessarily requires that there be no compensation prior to induction. Although the amounts of MSL protein produced from *tetO-msl2* may have been low in the un-induced state, they were sufficient to effect MSL complex binding to the X chromosome, and thus probably resulted in downstream effects on gene expression. Therefore, a lower background expression was desirable. In mammalian cells, the level of background expression from tet-regulated cassettes can vary depending on the site of insertion (Wong et al., 2005; Brough et al., 2007). If the *tetO-msl2* construct was also inserted in a repressive chromatin environment, the background level of *msl2* expression from the minimal promoter might be lowered or even abolished. To target the construct to a precise location, the ϕC31 system (described in Section 1.3.1) was used.

3.4.1 The 20C environment negatively affected gene expression

In the course of evaluating the activity of *arm-lacZ* at several X chromosome positions (Section 4.5.3), it was observed that insertions of the transgene at the 20C location resulted in a pale and variegated eye colour from the mini-*white* transformation marker, whereas insertions at other locations had strong mini-*white* expression (Figure 3.6, 'arm-*lacZ*'). As heterochromatin-mediated repression of mini-*white* is known to cause a variegated eye colour (Hazelrigg et al., 1984; Pirrotta et al., 1985), and as the 20C insertion location is nearly centromeric, the repression was most likely caused by the variable spreading of heterochromatin. However, the effect was not observed with insertions of *GMR-lacZ* or *GMR-hid* at exactly the same location. The lack of response in these plasmids may be due to differences in plasmid construction that alter the position of the mini-*white* gene. The differing construction of these vectors (Appendix I) causes the mini-*white* gene of the GMR plasmids to be located about the middle of the integrated DNA in a ϕC31-mediated insertion, whereas the mini-*white* gene of *arm-lacZ* positions to the extreme right hand side (centromere-proximal) of the integrated DNA. A variable degree of spreading of the facultative centromeric heterochromatin would be consistent with varying repression on the centromeric end of the inserted cassette.

This explanation is reinforced by the observation that insertion of *GMR-lacZ* or *GMR-hid* at exactly the same location did not result in a pale eye colour (Figure 3.6, 'GMR-*lacZ*'). The differing position of the *attB* sequence in the GMR plasmids as compared to the *arm-lacZ*

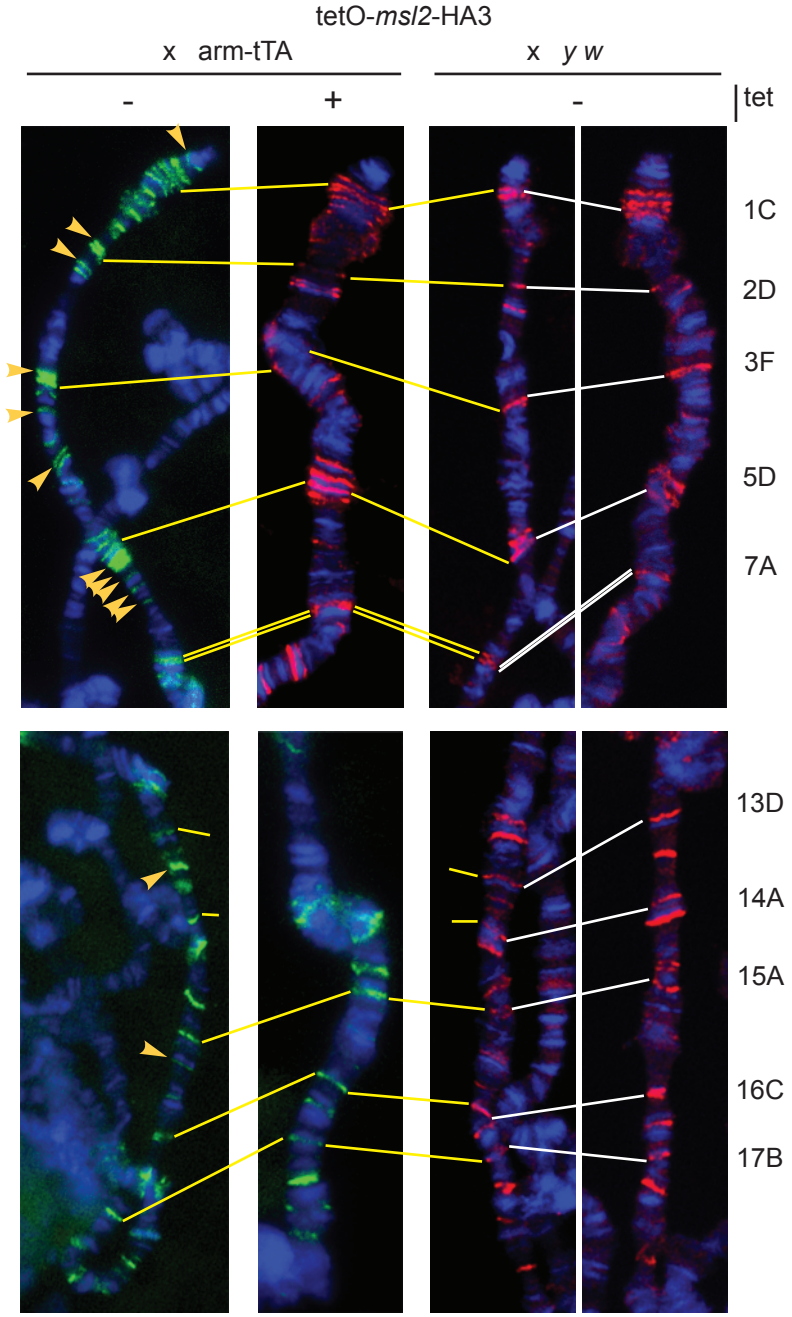


Figure 3.5: Modest increase in MSL binding to X chromosome with *msl2* induction. Combined DNA (DAPI, blue), HA (green), and MSL (red) fluorescent images of chromosome X tip (upper) and mid section (lower) from female larval salivary gland nuclei carrying one copy of the tetO-*msl2*-HA3 construct, and with or without one copy of the tetracycline trans-activator (arm-tTA). Flies were raised with (+) or without (-) the tetO inhibitor tetracycline (tet). Extra bands upon induction are marked with yellow arrow-heads. Bound MSL2 was detected for the top image of tetO-*msl2*-HA3 x arm-tTA (+ tet), all other MSL detection images were of MSL1. Antibodies against MSL1, MSL2 and HA exactly co-localise (not shown).

plasmid (Appendix I) causes the mini-*white* gene of the GMR plasmids to be located about the middle of the integrated DNA in a ϕC31-mediated insertion, whereas the mini-*white* gene of *arm-lacZ* positions to the extreme right hand side (centromere-proximal) of the integrated DNA. A variable degree of spreading of the facultative centromeric heterochromatin would be consistent with varying repression on the centromeric end of the inserted cassette.

3.4.2 Less MSL complex bound the X chromosome when MSL2 was expressed from 20C

To allow targeting of the *tetO-msl2* construct to specific location with the ϕC31 system, the 3.0 kb *Eco* RV-*Not* I fragment of pTM2HA3 containing the *tetO-msl2-HA* cassette was inserted into the *Hind* III (blunted) and *Not* I sites of the *Drosophila* transformation vector pUASTattB (Section 2.3.11), replacing the UAS and *hsp70* promoter sequences. The resulting plasmid, pTM2HA3attB (Appendix I), used the 3′ UTR and poly(A) sequences of SV40, contained an *attB* site for insertion, and lacked insulator elements that might otherwise impair the action of the heterochromatin. This design also places the *msl2* gene at the extreme right (centromeric) end of the integrated DNA. Sequencing of the construct with the AS5 primer confirmed the construction (Section 3.1).

ϕC31 transformations of *Drosophila* were conducted (Section 2.2.5) to target the *tetO-msl2-HA3* construct to two different *attP* sites that were hypothesized to reduce expression. One *attP* site was located at the 20C location, one was located in heterochromatin of the 3L chromosome arm (Bischof et al., 2007; Venken et al., 2006). Embryos from stock C63 (Table 2.1) and from a cross between J9 and #9746 (Table A.1) were injected to obtain insertions at 20C and at the 3L heterochromatin, respectively. Ten lines were obtained from 33 fertile G_0 of the C63 transformations, and three were bred to homozygousity. As PCR analysis confirmed that all three lines represented correctly-targeted insertions, only one was retained as stock C93 (Appendix H). Despite injecting 337 J9 x #9746 embryos, and screening 43 fertile G_0 flies, no transgenic flies were identified. This may be due to the relatively inaccessible target site, or to repressive effects on the mini-*white* transformation marker gene, or both.

The insertion of *tetO-msl2-HA3* at 20C also resulted in a paler, more variegated eye colour than the previous X chromosome insertions (Figure 3.6, compare '20C' to the 'other X' line C48). As the mini-*white* gene is located further to the left (centromere distal) than the *msl2* gene in pTM2HA3, it seems likely that the heterochromatin would also repress *msl2* expression. Following induction with *arm-tTA*, sufficient MSL2 protein was produced to reduce female viability (Figure 3.3). A Student's *t*-test shows the effect to be very highly significant ($T = 9.97$, $p < 0.0005$).

Chromosomes from the salivary glands of female larvae of both C48 and C93 lines were

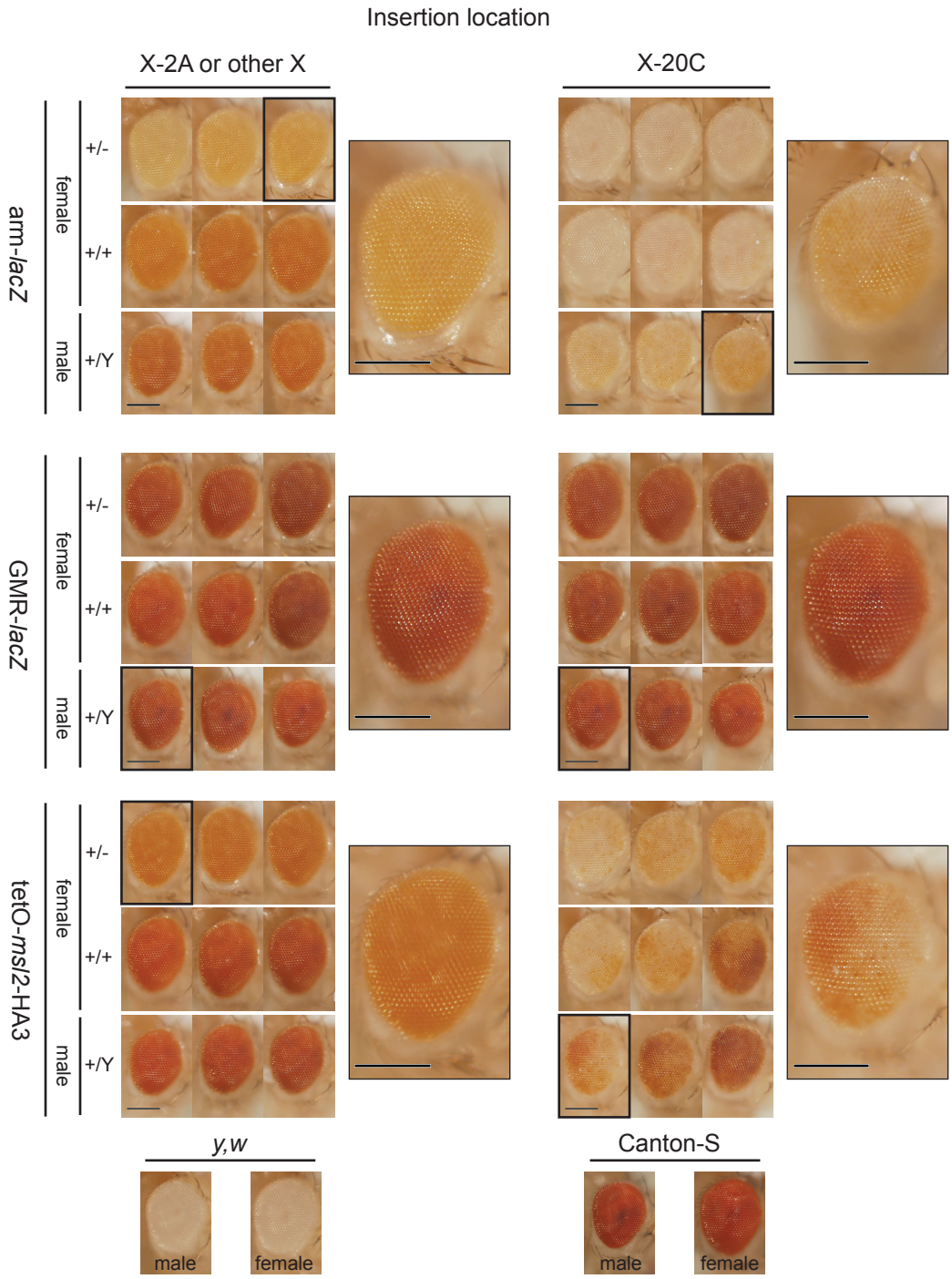


Figure 3.6: Position-dependent repression on the mini-*white* marker gene. Representative eye phenotypes of the arm-*lacZ*, GMR-*lacZ*, and tetO-*msl2*-HA3 constructs at the putatively repressive 20C location, and at another X site (arm-*lacZ* and GMR-*lacZ* at 2A). Males carry one insertion on their single X chromosome (XY), females have insertions on either one X chromosome (+/-), or both (+/+). Triplicate photos illustrate the range of phenotypes, and the enlarged eyes are indicated with black boxes. White eyes of *y w* and wild-type eyes of Canton S provided for comparison. All photos taken under identical conditions. Bar = 0.2 mm.

prepared and analysed for bound MSL1 (Section 2.7). Over the cytological regions 1 to 15, less bright bands (20–30) of bound MSL complex were observed when *msl2* was expressed from 20C (Figure 3.7a, "Restrictive") than when the insulated version was expressed from elsewhere on the X chromosome ("Permissive", >40 bands). On detailed analysis (Figure 3.7c), faint MSL1 binding could still be observed to many of the same locations, but the relative intensity of most bands was much reduced. Placement of the *tetO-msl2* construct at the 20C location had thus reduced, but not abolished, the level of MSL complex produced.

3.5 Evaluation of the inducible dosage compensation system

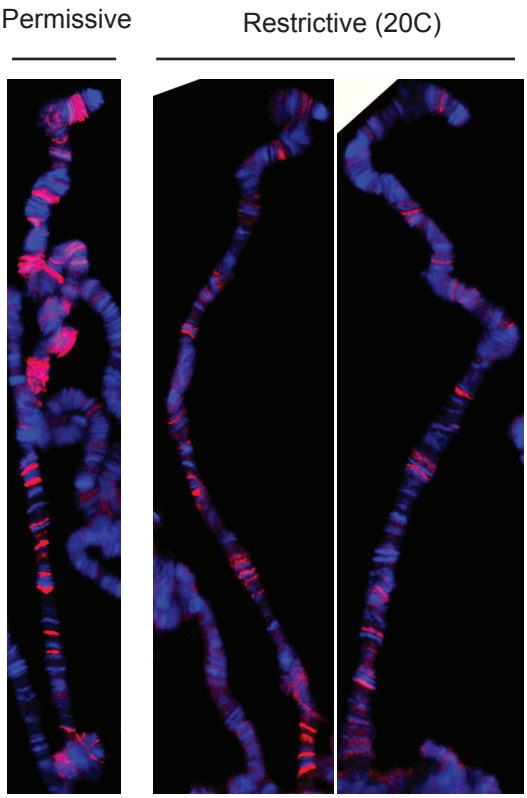
Desirable characteristics of the induced female dosage compensation system were:

1. Absence of MSL complex prior to induction
2. High levels after induction
3. A consequent chain of events resembling those of the male system

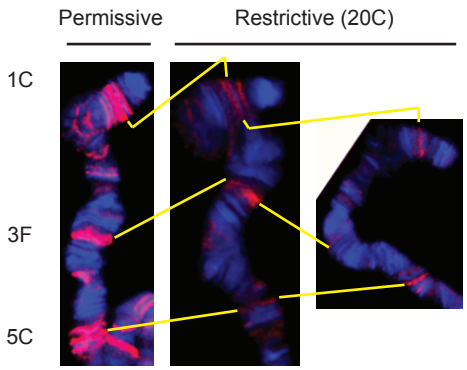
The tetracycline-repressible expression of *msl2* did not satisfy all these conditions.

In the absence of tetracycline, MSL2 was produced but the protein levels were lower than expected. A drop in female viability confirmed that MSL2 was produced (Section 3.2). Consequent effects mimicked male dosage compensation, as evidenced by the the localisation of MSL components to the female X chromosomes in salivary gland preparations (Section 3.3). However, the level of MSL2 produced with induction by *yp1-tTA* or *arm-tTA* was less than desired. Western blots could not detect MSL2, and the MSL binding pattern on the female X chromosome was less complete than that of the wild-type male X chromosome. Although dosage compensation is primarily regulated through the repression of *msl2* translation by Sxl, a weak female repression of *msl1* means that the amounts of *msl1* are also a limiting factor (Section 1.1.3). Perhaps simultaneous induction of *msl1* with the tetracycline system would increase the levels of MSL complex formed. Co-expression of *hsp83-msl1* with *hsp83-msl2* kills 100 % of females at the late larval stage (Chang and Kuroda, 1998). Although these dying females would not be suitable for polytene chromosome preparation from the third instar larval salivary glands, preparations of earlier tissues should be useful for gene expression or other binding studies (see below).

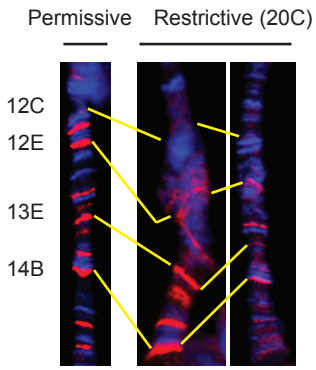
Absence of MSL complex prior to induction is absolutely necessary to study the initiation of dosage compensation. Levels of transgenic MSL2 produced in females without the presence of a trans-activator were not detectable by western blot, and had no effect on female viability.



(a) Whole X chromosome



(b) Telomeric



(c) Mid-section

Figure 3.7: Reduced X chromosome binding of MSL complex with expression of *msl2* from repressive 20C environment. Chromosomes were isolated from larval salivary glands of females expressing tetO-*msl2*-HA3 inserted in a permissive (line C48) or restrictive (20C, line C93) X chromatin environment. Combined DNA (DAPI, blue) and MSL1 (red) fluorescent images of X chromosomes are shown (a), with cytological regions 1–5 (b) and 12–14 (c) enlarged below.

However, sufficient MSL2 was produced to bind the X chromosome in salivary gland preparations. The limited pattern of sites bound resembled that of bound male complex in *msl* mutations (Demakova et al., 2003), and thus the sites likely represent the previously characterised high affinity sites (Section 1.1.4). The low level of MSL2 was likely a result of basal transcription from the $5'$ *P* TATA sequences downstream of the *tetO* regulator in the transgenic cassette. Similar 'leaky' transcription occurs sometimes with the same *tetO* in mammalian cells, and indeed in *Drosophila* (Furth et al., 1994; Kistner et al., 1996; Yokoyama et al., 2007). Perhaps this basal expression somehow caused the drop in male viability in five of the combined eight transgenic lines when compared to *y w*. Expression of the *tetO-msl2* cassette from a repressive chromatin location reduced, but did not abolish the amount of MSL complex bound to the X chromosome, and produced equal male and female viability when not induced(Section 3.4).

Significant work is still required to implement the system in any study of early dosage compensation events. Repression of basal transcription could be provided with co-expression a tetracycline-controlled transcriptional silencer (tTS). A tTS is based on a fusion of a species-specific repressor domain to tetR that also binds tetO sequences in the absence of tetracycline, but mediates trans-repression of a downstream gene (Deuschle et al., 1995; Belli et al., 1998; Freundlieb et al., 1999). A three-component (*e.g.* rtTA, tTS and tetO) system provides tight control in yeast and mammalian cells (Belli et al., 1998; Forster et al., 1999; Freundlieb et al., 1999). Transcriptional silencers based on *Knirps*, *Giant*, or *dCtBP* function well in an analogous *Drosophila* system (Ryu et al., 2001). Alternatively, the leaky aspect of the tetO could be improved. Modifications to the spacing of the tetO elements and distance of the elements upstream of the TATA box reduce basal transcription in mammalian cells (Agha-Mohammadi et al., 2004; Pluta et al., 2005). These modifications could similarly be employed in *Drosophila* systems.

Finally, if the co-expression of *tetO-msl1* did not sufficiently elevate levels of MSL complex produced, an alternative basal promoter, or indeed alternative regulatory system, might be required. Simple *hsp* elements (*e.g.* *hsp70-msl2*) do not provide the tight control of the tetracycline system, as the necessary heat-shock would also cause an unwanted down-stream stress response. The GAL4 / UAS system is more direct, but does not allow temporal induction. Combination with a temperature-sensitive version of the Gal4-inhibitor Gal80 (Zeidler et al., 2004) could achieve this aspect. This approach would consist of: *arm-GAL4*, *arm-GAL80^{ts}*, and *UAS-msl2*. A shift from the lower permissive temperature to a higher restrictive temperature would abolish the inhibition and allow Gal4 activation. Although any alternative system may elevate the levels of MSL2, it may be that the female death caused by high levels of MSL complex and aberrant dosage compensation is problematic for certain experiments, as would be the case for polytene analysis. Regardless of its form, any new system should be similarly

analysed on the above induction characteristics before proceeding with initiation studies.

3.6 Future directions

Immuno-fluorescence detection of MSL2 on the polytene chromosomes is very sensitive, and can detect binding when no apparent protein is visible by western blot (Kelley et al., 1997). The low level of MSL2 produced from the 'leaky' *tetO-msl2* transgenes was also detectable only by polytene chromosome analysis, and was undetectable by both western blot or a drop in female viability (Sections 3.2, 3.3). While this may complicate any similar chromosomal studies, the protein levels may be suitable for experiments less sensitive to a low background.

Binding to individual targets could be detected by PCR following chromatin immuno-precipitation (ChIP) against MSL components or the consequent modified histone isoforms, as is possible in the male system (Smith et al., 2001; Straub et al., 2005*b*; Kind and Akhtar, 2007). Certainly many known targets appear un-bound by background protein levels at the resolution of polytene chromosome analysis. With induction of the female system, the order of binding of individual MSL components could be investigated, or the position on the gene first bound (*i.e.* 5′ to 3′) sought. Chromosome-wide binding profiles could be similarly analysed by micro-array or "next-generation" sequencing of the immuno-precipitates. The hypothesis that a hierarchy of sites are bound in order as the complex assembles could be tested.

Gene expression changes on known MSL targets at the time of complex formation could be measured through quantitative RT-PCR. It may be possible to observe a transcription increase following MSL binding, or even initiation of transcription (*e.g.* at *roX1*). Chromosome-wide transcription profiles could be generated by microarray analysis, as was the case for male binding profiles (Hamada et al., 2005; Deng and Meller, 2006; Gupta et al., 2006). Improvements in quantification of RNA-seq results may prove to be a superior technique (Sackton and Clark, 2009).

Regardless of the experiment performed, the induction parameters of the system will first need to be carefully established. Post-induction steps of transcription, translation, modification, and transport to the nucleus will all delay the time of MSL complex formation. A tetracycline-derivative such as doxycycline may give more desirable induction profiles (Gossen et al., 1995; Horn and Wimmer, 2003). As initially planned, tight control over induction may be best studied in a cell culture system. Suitable systems would be the Kc *Drosophila* cell line, which is derived from female embryonic tissue (Hoshijima et al., 1991; Lynch and Maniatis, 1996), or a culture of female tissues, *e.g.* salivary glands. The plasmids generated in this study should be equally suitable for transient transfection in cell culture.

Chapter 4

Development of the GMR-hid reporter gene system to search for novel factors involved in Drosophila dosage compensation

The line of enquiry presented here extends that of earlier studies (Laverty, 2003). A reporter gene had been developed to visualise the efficiency of dosage compensation in a phenotypic readout, specifically eye size (Section 1.3.5). The subsequent goal was to use this system as the basis for a mutational screen to identify novel components of the dosage compensation apparatus (Section 1.4.2). The existing system was first extensively characterised to ensure that it reliably reported dosage compensation, and was suitable for the proposed screen.

4.1 The roX1DHS-GMR-hid system displayed an unreliable sex-specific difference

The reporter constructs RGH and R_9GH aim to attract MSL complex to the *roX1* DHS site, or a 9 copy multimer of the site, respectively, and observe the corresponding effect on GMR-*hid* expression (Section 1.3.5, Figure 4.1). From previous *Drosophila* transformations, one line carrying RGH and one carrying R_9GH were available (Laverty, 2003). These were designated stocks C4 and C5, respectively (Appendix H).

4.1.1 Male-specific phenotype less significant than predicted

The eyes of flies carrying the insertion of RGH were smaller than those of wild-type flies, indicating that the *GMR-hid* cassette was expressed. If the MSL complex was attracted to the cassette by the *roX1* DHS, then males should have had smaller eyes than females, due to elevation of *GMR-hid* expression. However, male eyes were only very slightly smaller than female eyes, indicating an inability to attract MSL complex, or report changes to *GMR-hid* expression. Representative individuals are shown as 'C4' in Figure 4.2a.

Preliminary observations of the line C5 carrying R_9GH suggested that males had smaller eyes than females (Laverty, 2003). To examine the line more extensively, C5 adults were back-crossed (Section 2.2.4) to the *y w* injection stock (Table A.1), and larger populations were raised. Regardless of whether the mothers or fathers of the cross were C5, a large range in the eye size of the offspring was evident (not shown). Although most males had smaller eyes than most females, male eye sizes ranged between ablated to nearly half the size of a wild-type eye, and female eyes were between 20 % of wild-type to nearly full-sized. Representative individuals are shown as 'C5' in Figure 4.2b. The significant overlap in the phenotypes would complicate the use of the line in a large genetic screen for modifiers of dosage compensation, as changes in dose could not quickly be identified. Furthermore, the large degree of variation in eye size may indicate a poor response to the expression level of *GMR-hid*. Therefore, the existing lines were unsuitable for the proposed genetic screen. Several options were explored in an effort to improve the system.

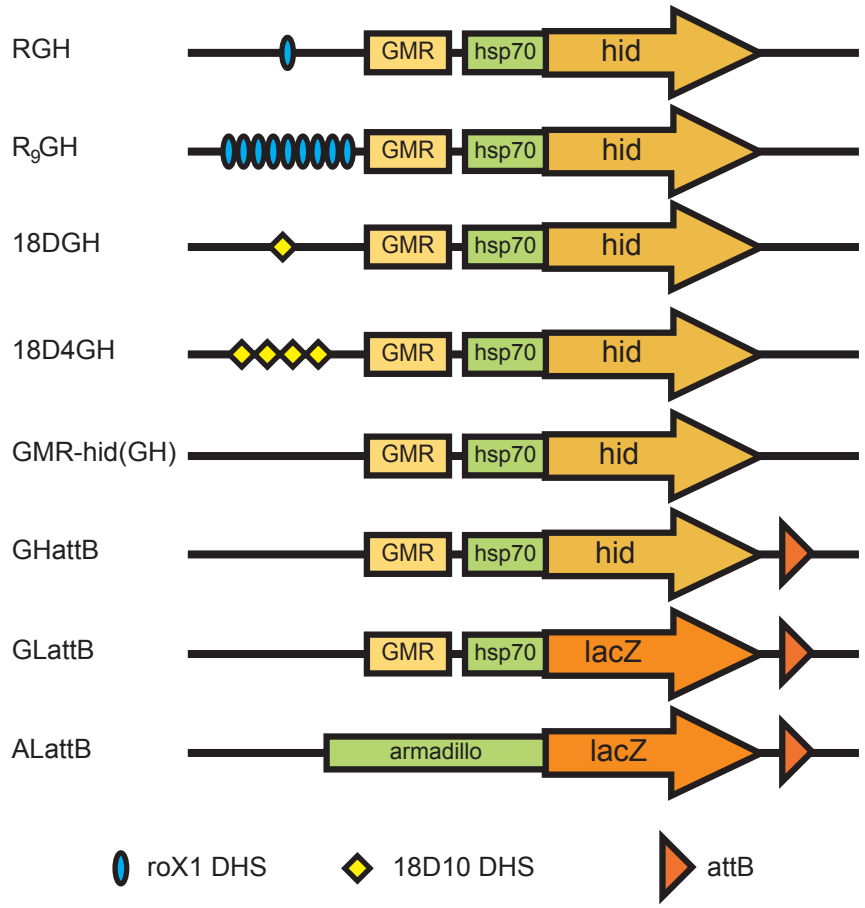


Figure 4.1: Schematic diagram of constructs *GMR-hid* and related constructs. A depiction of RGH and R9GH (from Laverty, 2003), with the modified constructs below. GLattB and ALattB are constructs *GMR-lacZ* and *arm-lacZ*, respectively, of Figure 5.1. The first five constructs (without an *attP* site) also contain *P* termini surrounding the expression cassette (not pictured).

4.1.2 New transformant locations did not improve the system

In an effort to improve the system through use of differing chromatin environments, new insertions of both RGH and R_9GH were generated. *P* element transformations of *y w Drosophila* were conducted (Section 2.2.5) with the plasmids RGH and R9GH, to yield the transgenic lines listed in Appendix H. Four transgenic flies carrying single insertions of RGH were identified, with insertions on either the second or third chromosome. Four flies carrying a single insertion of R_9GH, and one carrying a separable double insertion, resulted in six lines. Of these six, insertion 3 (stock C18) was at an X chromosome location; the others were on chromosomes two or three.

Flies with representative eye sizes were photographed for RGH (Figure 4.2a) and R_9GH (Figure 4.2b) lines. The phenotype of R_9GH line C17 was similar to line C18, and of line C21 to line C19. Due to the small eye size in line C22, equal to that observed in all homozygotes, this line likely represents a double insertion on the same chromosome. Accordingly, flies homozygous for this insertion were very rarely observed, perhaps due to increased expression of *hid* in some other neuronal cells where GMR also acts (Section 1.3.4).

All lines carrying *GMR-hid* preceded by one copy of the *roX1* DHS exhibited smaller eyes than wild-type flies, but regardless of the site of insertion no increased sex-specific difference in eye size was observed. Therefore, differing chromatin locations did not improve the ability of RGH to report on dosage compensation. In contrast, the five different phenotypes generated with insertions of R_9GH must have resulted from the nine copy multimer of the *roX1* DHS. The R_9GH female eyes were generally paler and larger than those of RGH. Male R_9GH eyes were significantly darker but not much smaller than the corresponding female eyes. This is consistent with the observation that the *roX1* DHS confers female-specific repression on a transgene, that is alleviated with the MSL complex produced by *hsp83-msl2* (Bai et al., 2004). Female-specific silencing of a mini-*white* marker gene when the transgene is inserted into the heterochromatic fourth chromosome is also relieved by MSL complex induced with *hsp83-msl2* (Kelley and Kuroda, 2003). Thus it was likely that MSL complex was attracted to the nine copy multimer of the *roX1* DHS, and was capable of mediating up-regulation of the associated mini-*white* transgene. Unfortunately, any effect of MSL on *GMR-hid* expression was slight. No line exhibited a male eye size small enough to allow rapid distinction from the female eye.

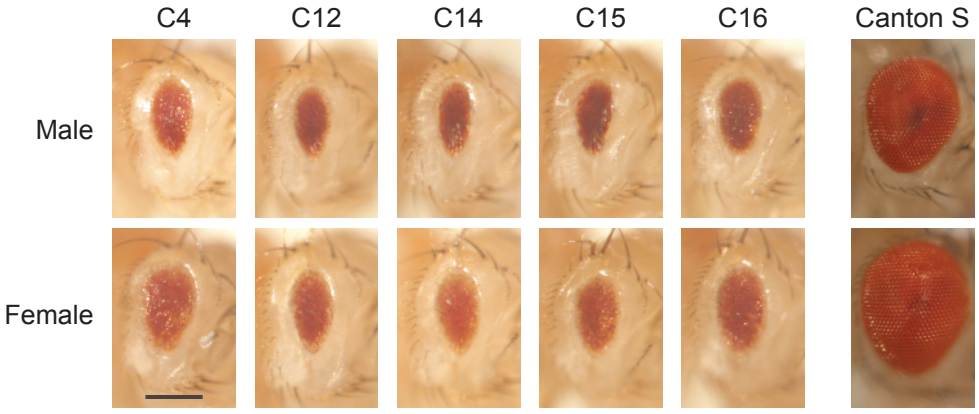
4.1.3 The roX1DHS-GMR-hid system was dose responsive

If the male expression of *GMR-hid* was indeed doubled by the attracted MSL complex in the R_9GH lines, then the system did not report a corresponding two-fold decrease in eye size. Two doses of the reporter gene cassette were present in the flies homozygous for the insert, and all such flies displayed nearly ablated eyes (not shown). However, the complicating effects

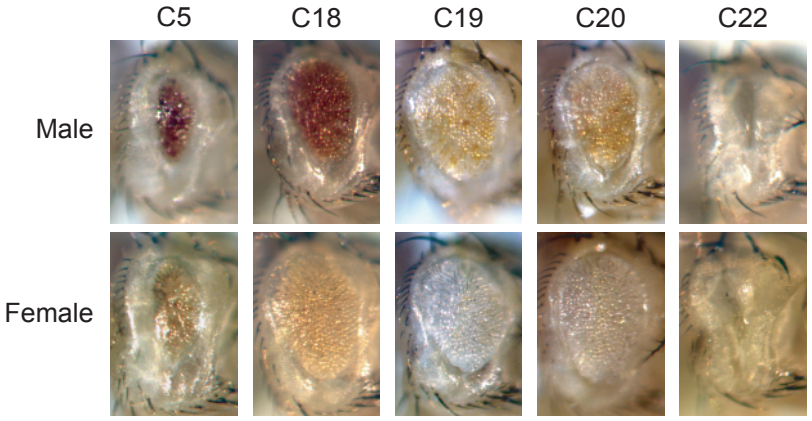
of transvection (Section 1.2) meant the observed eye phenotype for homozygotes may reflect levels of *GMR-hid* expression above a normal two-dose level. To obtain an accurate two-dose phenotype free from transvection, flies homozygous for one insertion of R_9GH were crossed to flies homozygous for another, and eyes of the offspring photographed (Figure 4.2c). The offspring carried one copy of R_9GH at each of the two different locations, and thus could not share transcription factors between chromatids. As the true two-dose phenotype was also a nearly ablated eye, it was concluded that the (*roX1* DHS)9-GMR-*hid* system could accurately respond to a two-fold increase in *GMR-hid* expression. The failure of males with one copy of the reporter gene to show an eye size comparable to that produced from two doses thus reflects an increase in transcription of somewhat less than two-fold. That is, dosage compensation was inefficient. This is most surprising for the insertion on the X chromosome, where a degree of endogenous compensation might also have been acquired.

4.2 The roX1DHS-GMR-hid system unreliably reported MSL complex activity

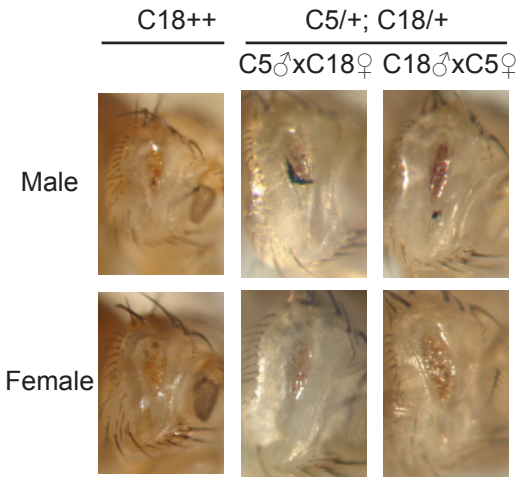
If the slight male-specific increase in *GMR-hid* expression in *R_9GH* lines was due to the action of the MSL complex, then the increase did not reflect efficient dosage compensation. This might be due to insufficient levels of MSL complex recruited to the autosomal transgene. To test if the male-specific increase was less than two-fold because of low MSL protein levels, several *R_9GH* lines were crossed to lines carrying *hsp83-msl2*, or both *hsp83-msl1* and *hsp83-msl2* (Section 2.2.4). Expression of *msl2* in females induces MSL complex formation in females and reduces viability, an effect enhanced by the co-expression of *msl1* (Section 1.1.3). Although not measured, it was expected that the transgenes may also alter MSL levels in males. Flies carrying one copy of a *GMR-hid* construct were compared to siblings that also carried one copy of the over-expressed transgene(s) (Figure 4.3). Surprisingly, increased MSL levels corresponded to a larger eye, and thus a decrease in *GMR-hid* expression, in most lines, including a line carrying a single copy of the *roX1* DHS. In lines where *GMR-hid* expression was normally quite low (*e.g.* *R_9GH* lines C19, C20, C21), no difference in eye size was observed. In all lines, the increased MSL levels did result in a darker eye colour (not shown), implying increased expression of the mini-*white* transformation marker gene. Thus MSL complex was indeed recruited to the autosomal insertions of *R_9GH*, and was responsible for increased expression of mini-*white*. However, as the level of *GMR-hid* expression did not correlate to amounts of MSL complex, the reporter gene system did not accurately provide information on the effects of the complex.



(a) *roX1DHS-GMR-hid* lines

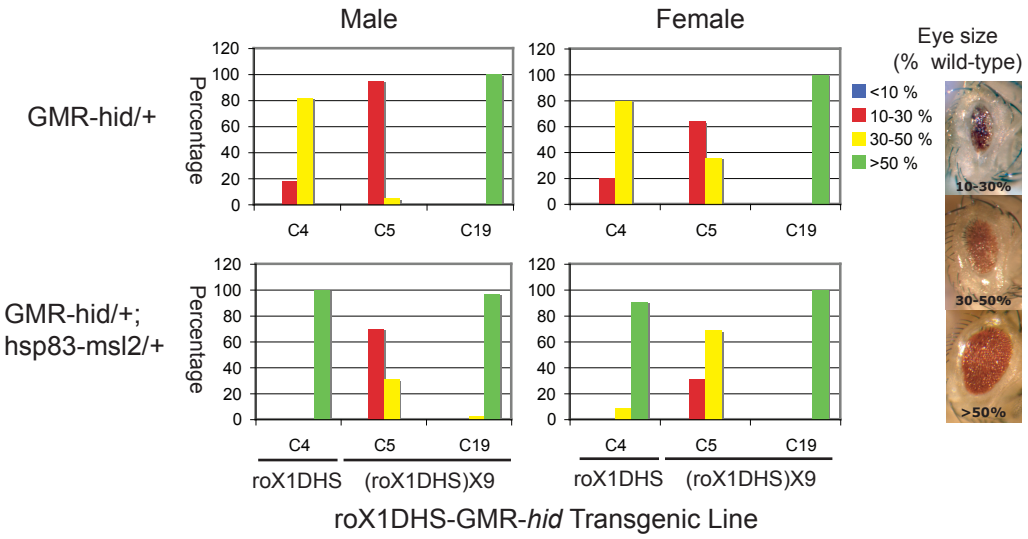


(b) *(roX1DHS)9-GMR-hid* lines

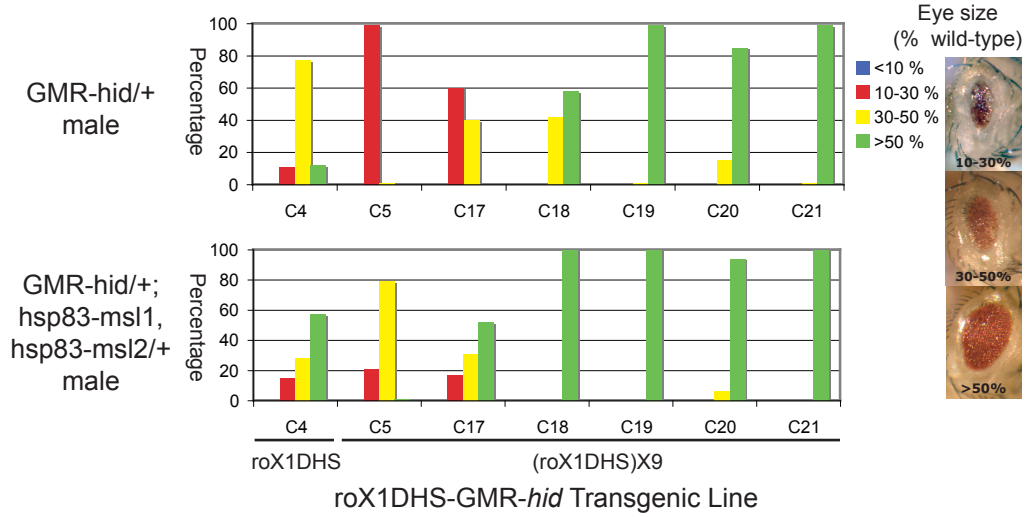


(c) Two doses of construct at different locations

Figure 4.2: Subtle sex-specific phenotype of *roX1DHS*-GMR-*hid*. Photos of representative eyes from indicated lines (*e.g.* C4) of transgenics carrying one copy of *roX1DHS-GMR-hid* (a), or *(roX1DHS)9-GMR-hid* (b). Wild-type (Canton S) eyes provided for comparison. Two doses of either construct give an ablated eye, even without transvection effects (c). Bar = 0.2 mm.



(a) Response to *hsp83-msl2*



(b) Response to *hsp83-msl1* and *hsp83-msl2*

Figure 4.3: Increased eye size of *roX1DHS-GMR-hid* with over-expression of *msl* genes. Crosses were performed to create flies carrying one copy of the reporter construct from each of several lines (*e.g.* C4), and either normal or over-expressed *msl* genes. Flies were sorted into eye size categories as indicated, based on percentage of a wild-type eye, and the percentage of total flies of that sex plotted. (a) Response to *hsp83-msl2* only. (b) Response of males to *hsp83-msl1* and *hsp83-msl2*. Females of this genotype die (Chang and Kuroda, 1998).

4.3 A replacement MSL binding site did not improve the reporting ability

As the regulation of the *roX* genes by MSL complex may represent a transcriptional enhancement activity distinct from dosage compensation (Section 1.1.3), the MSL complex attracted by the *roX1* DHS may not have been sufficient to confer dosage compensation on the associated *GMR-hid* transgene. Thus the *roX1* DHS was replaced with both the DHS from the high affinity site at 18D10 (Section 1.1.4), and a four copy multimer of this DHS known to recruit MSL complex more effectively (Oh et al., 2004). A 3.9 kb *Eco* RI–*Eco* RI fragment of the plasmid "BS SK hid cDNA" containing the *hid* cDNA was inserted into the *Eco* RI site of pGMR-1 upstream of GMR to create pGMRhid (Section 2.3.11, Appendix I). A 510 bp *Bam* HI–*Not* I fragment of "pCaSpeR3 18D monomer" containing the 18D10 DHS, treated with Klenow fragment to fill in the cohesive *Not* I end, was inserted into the *Bgl* II and *Stu* I sites of pGMRhid to create p18DGH (Figure 4.1). A 2040 bp fragment of "pCaSpeR3 18DL 4mer" containing the 4 copies of the 18D10 DHS was likewise inserted into pGMRhid to create p18D4GH (Figure 4.1).

P element transformations of *y w Drosophila* (Section 2.2.5) were conducted with both p18DGH and p18D4GH, to yield the transgenic lines listed in Appendix H. Seven transgenic lines were obtained from seven independent insertions of p18DGH, on chromosomes 2 and 3. None of the lines exhibited a smaller male eye (Figure 4.4). Two separate transgenic lines carrying insertions of p18D4GH on the third chromosome likewise displayed similar eye sizes in both sexes. Because the 18D10 DHS weakly attracts MSL complex to its autosomal site of insertion, and the multimer strongly so (Oh et al., 2004), the failure to observe a male-specific increase in *GMR-hid* expression must have been due to a lack of conferred up-regulation. Either the attracted complex was not capable to mediate compensation of the transgene, or the *GMR-hid* cassette itself was not amenable to male hyper-transcription by the MSL complex.

4.4 GMR-hid insufficient to report dosage compensation

To test the ability of X-linked *GMR-hid* to respond to dosage compensation, *P* element transformations of *y w Drosophila* were conducted (Section 2.2.5) with the *GMR-hid* plasmid pGMR-hid (Figure 4.1). Two transgenic flies with single insertions of *GMR-hid*, one fly with two separable insertions, and one with at least three insertions, were obtained. Further injections by Esther Belikoff[1] produced seven transgenic flies, including two carrying separable double insertions, and one with three separable insertions of *GMR-hid*. With controlled breeding and

[1]Massey University, Palmerston North, New Zealand

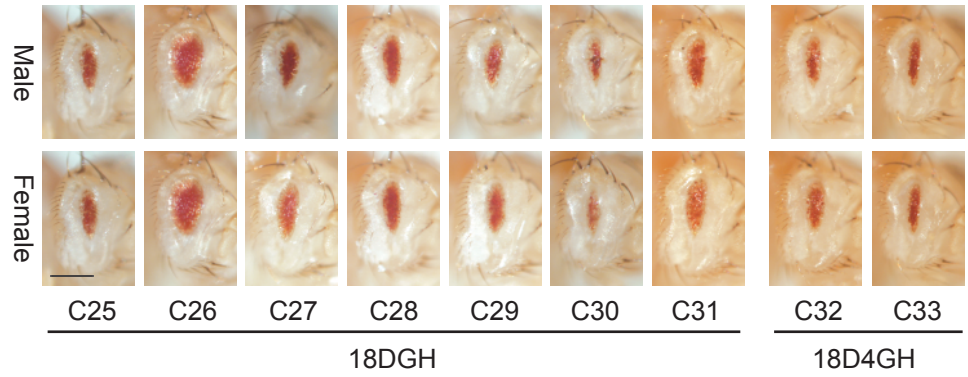


Figure 4.4: No increase in sex-specific phenotype of GMR-*hid* with use of 18D MSL binding site. Representative photos of indicated transgenic lines (*e.g.* C25) carrying GMR-*hid* preceded by either the DHS site of 18D10 (18DGH) or a four copy multimer of the same site (18D4GH). Bar = 0.2 mm.

genetic linkage analysis (Section 2.2.5), a combined total of 18 lines were obtained, with insertions on chromosomes 2, 3, and X. Ten lines were retained, including four with insertions on the X chromosome (Appendix H).

4.4.1 X insertions of GMR-hid did not acquire full compensation

Eyes from representative flies of each of the ten lines carrying insertions of *GMR-hid* were photographed (Figure 4.5a). Eye sizes of the lines carrying autosomal insertions varied depending on the position of insertion, reflecting the effects of the surrounding chromatin on gene expression. Although photographic conditions were not controlled for, low levels of *GMR-hid* expression correlated with low mini-*white* expression. In all six lines with autosomal insertions, no sex-specific difference in either *GMR-hid* or mini-*white* expression is observed, indicating that both genes are expressed equally in each sex. However, in three of the four insertions on the X chromosome, male eyes were very slightly smaller than female eyes, and the fourth line (C60) noticeably so. The C60 male eye was still larger than the eye of the two-dose homozygous female (not shown). In all lines, the male-specific increase in *hid* expression may have resulted from the effects of the MSL complex already present on the X chromosome.

4.4.2 GMR-hid on the X did not respond to altered MSL levels

To confirm that the sex-specific difference in eye size of *GMR-hid* fly lines was due to the action of MSL complex, crosses were performed to observe the response of the transgene to altered levels of MSL proteins (Section 2.2.4). As the *GMR-hid* line C60 showed the greatest difference in eye size, the response of this line to both increased amounts of MSL2, and a reduction in MSL1, MSL2, and MLE, was observed, in comparison to wild-type levels of MSLs

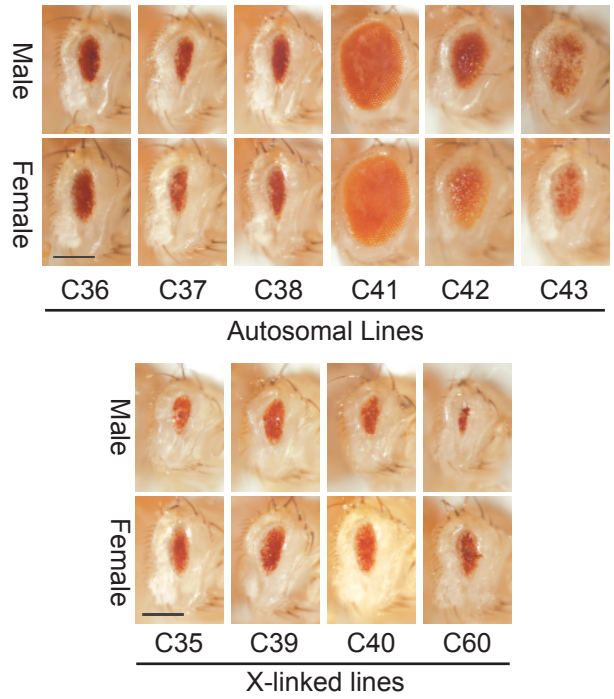
(Appendix C, Figure 4.5b). No histograms of phenotypes are provided, as the overall variation in phenotypes was slight. If the eye size accurately reported the level of MSL-mediated dosage compensation, then a reduction in MSL1, MSL2, and MLE would be expected to decrease the amount of MSL complex available, and may lead to a male eye more resembling the female eye. This was particularly likely given that the eye size seemed sensitive to position-dependent changes in *GMR-hid* expression (Figure 4.5a). Likewise, over-expression of *msl2* from the constitutive heat-shock promoter should induce MSL complex formation in females (Section 1.1.3), causing a smaller eye. However, neither situation significantly altered the phenotypes of *GMR-hid*. A reduction of MSL proteins had no effect whatsoever. An increase in *msl2* did not affect females, and if anything slightly increased the size of the male eye. Thus the male-specific increase in *GMR-hid* expression of line C60 appeared to be not due to MSL complex-mediated dosage compensation, and may have instead resulted from the effect of sex-specific enhancer elements at or near the insertion site.

4.4.3 A similar X-linked GMR-hid was also unaffected by MSL levels

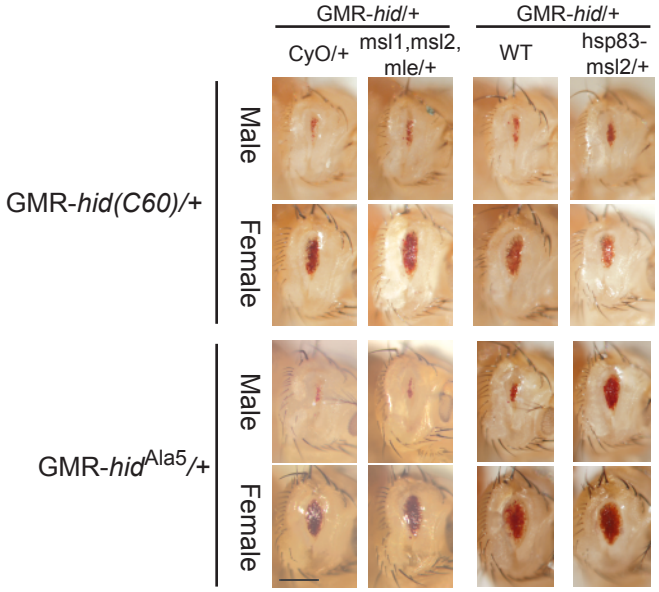
An X-linked insertion of a related transgene, *GMR-hid*^{Ala5} (Bergmann et al., 1998), also resulted in eyes that were significantly smaller in males than in females (Andreas Bergmann, personal communication). The *Ala5* allele is a mutation of the five MAPK phosphorylation sites of *hid* that prevents the post-translational repression of endogenous Hid normally required to suppress programmed cell death (Bergmann et al., 1998). Expression of *hid*^{Ala5} by the *GMR-hsp70 TATA* thus generally results in a smaller eye than similar expression of *hid*$^{+}$ (Bergmann et al., 1998), but both should be equally susceptible to MSL-mediated dosage compensation.

To confirm that the male-specific increase in *GMR-hid*^{Ala5} expression from the line with an X-linked insertion was due to MSL-mediated dosage compensation, the line was also crossed (Section 2.2.4) to lines with altered levels of MSL proteins, and the effects compared to wild-type levels of MSLs (Appendix C, Figure 4.5b). As with line C60 of *GMR-hid*, a reduction in MSL1, MSL2, and MLE did not affect the male-specific level of *GMR-hid*^{Ala5} expression. Similarly, over-expression of *msl2* did not effect a smaller eye in either sex. In fact, about 5 % of the *GMR-hid*/+*; hsp83-msl2*/+ males (14 of 306) displayed an increased eye size (as pictured). As the sex-specific differences were not substantially affected by MSL levels, the smaller male eyes in both *GMR-hid* line C60 and the X-linked line of *GMR-hid*^{Ala5} thus most likely resulted from the effects of sex-specific enhancer elements nearby the region of insertion.

Because X-linked insertions of *GMR-hid* did not generally give a substantially greater male-specific expression, and the insertions that did were not susceptible to MSL levels, it was concluded that *GMR-hid* itself was incapable of reporting dosage compensation.



(a) Phenotypes of *GMR-hid* lines



(b) *GMR-hid* phenotype in response to altered MSL levels.

Figure 4.5: *GMR-hid* alone a poor reporter of dosage compensation. (a) Representative eye phenotypes of indiated lines carrying GMR-*hid* at both autosomal and X chromosomal locations. (b) Phenotypes of the sex-specific GMR-*hid* line C60 and a line carrying a related construct (*GMR-hid*^{Ala5}, a gift from A. Bergmann) in response to either a reduction of MSL1, MSL2, and MLE, or an over-expression of *msl2*. Each cross also gave an internal genetic control with normal MSL levels (CyO/+ and WT, respectively). Bar = 0.2 mm.

4.5 GMR-based expression only weakly affected by endogenous dosage compensation

Because the MSL complex generally binds to the 3′ coding sequence of X-linked genes (Section 1.1.4), it was likely that the *hid* gene itself was not susceptible to dosage compensation, perhaps due to a lack of appropriate 3′ sequence or chromatin features. However, it remained to be verified that GMR-mediated expression was capable of driving genes that could be dosage compensated. To this end, the expression of *lacZ* from a GMR expression cassette inserted on the X chromosome was examined.

The approach was as follows:

1. Establish a targeted transformation system to allow accurate comparisons of cassettes.
2. Optimise the β-galactosidase activity assay to measure *lacZ* expression from GMR.
3. Confirm ability of transgenes to acquire dosage compensation at defined X-linked sites, with constitutively-expressed *lacZ* from the *armadillo* promoter (*arm-lacZ*).
4. Confirm that *GMR-hid* cannot acquire compensation at the same sites.
5. Determine the response of *GMR-lacZ* at the same sites.

4.5.1 A targeted system was developed

To accurately compare *GMR-lacZ* and *GMR-hid*, it was desirable to eliminate the variable position effects inherent with *P* element transformation. The ϕC31 recombinase-mediated transformation system (Section 1.3.1) allows precise insertion of transgenic cassettes at pre-determined 'landing' sites. Furthermore, as each insertion event of a plasmid at a landing site should be genetically identical to every other, only one transgenic line of each need be retained and examined, significantly reducing the workload.

Three X-linked landing sites were selected

In order to utilise the system, an *attB* site needed to be introduced into the recombinant reporter gene cassettes. Five lines of *Drosophila* carrying *mariner*-mediated insertions of *attP* landing sites (at 2A, 3A, 3B, 6E, and 20C, marked with *3xP3-RFP*) on the X chromosome were obtained from Johannes Bischof [2], along with a line that expressed the ϕC31 recombinase from a genomic insertion. Four of the lines (not zh-3Aa) were known to support targeting and subsequent gene expression (Bischof et al., 2007). In comparison to the microarray data from ChIP

[2] Institute of Molecular Biology, University of Zurich

against MSL components (Gilfillan et al., 2006), sites at 2A, 6E, and 3B were within 20 kb of significant bound MSL in late embryos (not shown). To target a particular *attP* landing site, the corresponding line was crossed to the recombinase line, and the offspring injected with the *attB*-encoding plasmid (Section 2.2.5). The five *attP* sites were screened for targeting ability with the *arm-lacZ* cassette (described below), and the relevant data summarized in Table 4.1.

ϕC31 transformations of *Drosophila* were first conducted (Section 2.2.5) to target the *arm-lacZ* cassette to the X-linked *attP* site at cytological position 2A. Eleven independent insertions were detected, six of which were bred to homozygous lines. All six insertions were single insertions genetically linked to the X chromosome (Section 2.2.5), and displayed co-segregation of the mini-*white* transformation marker and the *3xP3-RFP* landing site marker (not shown). Furthermore, all six gave bands of the expected sizes upon PCR both within the *arm-lacZ* cassette, and across the cassette / landing site junction (Figure 4.6a, not shown). Lastly, all six gave an identical band upon inverse PCR (Section 2.2.6) with the primers mar3f and mar3r designed to detect the 3′ *mariner* element of the 2A *attP* landing site (Figure 4.6a, primer pair '3', not shown). DNA sequencing (Section 2.3.7) of two reactions from the inverse PCR with the mar3s primer produced identical sequence starting in the 3′ *mariner* element and running into genomic sequence at 2A (Figure 4.6b). As all six, and likely all 11, insertions appeared identical, it was concluded that the ϕC31 targeting strategy was robust, at least to 2A. One line carrying *arm-lacZ* at 2A was retained (as C66).

Additional ϕC31 transformations of *Drosophila* were conducted to target the *arm-lacZ* construct to the X-linked *attP* sites at cytological positions 3A, 3B, 6E, and 20C.

Six transgenic flies were isolated from the 3A transformations, and were bred to create homozygous lines, however none proved useful. Four of the six co-segregated with *3xP3-RFP*, and gave bands of expected sizes with both PCR reactions as for insertions at 2A, and thus were likely inserted correctly. Sequencing of bands produced by inverse PCR as above confirmed that the landing site was at 3A (Appendix E). However, all four of these insertions at 3A were male lethal, and thus useless for evaluating dosage compensation. The two remaining 'insertions' of *arm-lacZ* segregated independently from *3xP3-RFP*, were not genetically linked to the X chromosome, were male-viable, and did not produce bands of the expected sizes from the PCR tests for insertions. The likelihood of two events of back-mutation of *white* in one transformation experiment is not high. These two lines may thus have represented ectopic insertions of *arm-lacZ* at locations other than the *attP* landing site at 3A.

Despite repeated transformation attempts to target *arm-lacZ* to the *attB* at 3B, and the screening of 24 fertile G_0, no transformant flies were identified. A BLAST analysis of the genomic sequence of insertion of the 3B landing site (provided by Johannes Bischof) against the *Drosophila* genome reveals nearly 40 kb of gene-devoid sequence surrounding the landing site (not shown). The chromatin state of such a gene-poor region may increase the difficulty of

ϕC31-mediated recombination at the location.

Two insertion events at 6E were detected, both of which co-segregated with *3xP3-RFP*, were linked to the X chromosome, and produced bands of the expected size with PCR and inverse PCR reactions. Sequencing of the inverse PCR products confirmed the landing site was at 6E. One line was retained (as C67).

Insertions at 20C were very difficult to detect. Expression of the mini-*white* transformation marker from this location was so low that females with one copy of the transgenic cassette were not distinct from the pale pink eye phenotype often produced from the *3xP3-RFP* transformation marker (Figure 3.6). Thus despite screening 45 fertile G_0 flies (mostly by Esther Belikoff[3]), only two w^+ males were found. One resulting line was examined, and exhibited co-segregation of mini-*white* and *3xP3-RFP*. Sequencing of a band produced by inverse PCR of the 3′ *mariner* end confirmed the landing site was at 20C. This line was retained (as C68).

The targeted system was more efficient than P-element transformation

Over the course of the studies presented in this report, *Drosophila* were transformed by both *P* element and ϕC31 methods many times. Each session of microinjection of a particular line with a particular plasmid could be described as an independent 'event', with an associated frequency of transformation. These frequencies were calculated as proportion transgenic flies from total fertile 'G_0' (injected) flies (Appendix E). The frequencies were compared between transformation methods, and between landing sites for the ϕC31 method (Figure 4.7).

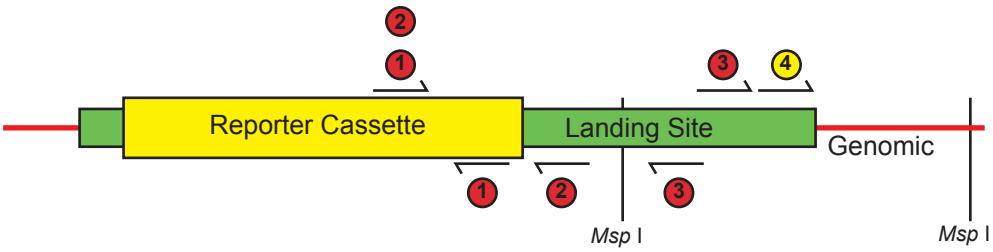
In 36 experiments with useful data, ϕC31 transformation frequencies ranged between 0 and 0.5, similar to 17 available *P* element experiments (Figure 4.7a). After excluding frequencies of zero, small sample sizes, and low confidence estimates, a more valid comparison was established from a limited data set (12 ϕC31 experiments, 5 *P* element experiments, Figure 4.7b). The median transformation efficiency for ϕC31 experiments was 0.23, nearly twice the frequency of 0.13 for *P* element transformations. For the ϕC31 transformations, insufficient data was available to establish whether particular landing sites supported better targeting (Figure 4.7c), especially when the estimations with low confidence were excluded (Figure 4.7d). Thus the ϕC31 method was demonstrably more efficient for *Drosophila* transformation than the *P* element approach. Not only did the targeted technique reduce the required number of lines, but also the effort needed to generate those lines.

[3]Massey University, Palmerston North, New Zealand

Table 4.1: Summary of ϕC31 transformations with *arm-lacZ*

		Transgenic Lines		Co-segregate with		Band with PCR		Genomic Position	
attP	Fertile G_0	Created	Analysed	X chromosome	RFP	Within Cassette	Across boundary	Cytological	Nucleotide[a]
2A	37	11	6	Yes	Yes	Yes	Yes	2A	1308929-
3A	48	6	4	-[b]	Yes	Yes	Yes	3A	2386670+
			2	No	No	No	No	-	-
3B	24	0	-	-	-	-	-	-	-
6E	17	2	2	Yes	Yes	Yes	Yes	6E	6796832+[c]
20C	45	2	1	Yes	Yes	-	-	20C	22023051-

[a]First nucleotide after ϕC31 landing site, BDGP genome version 4.3. +/- is relative to genome.
[b]Male lethal
[c]Repetitive sequence; possibly before 6796525



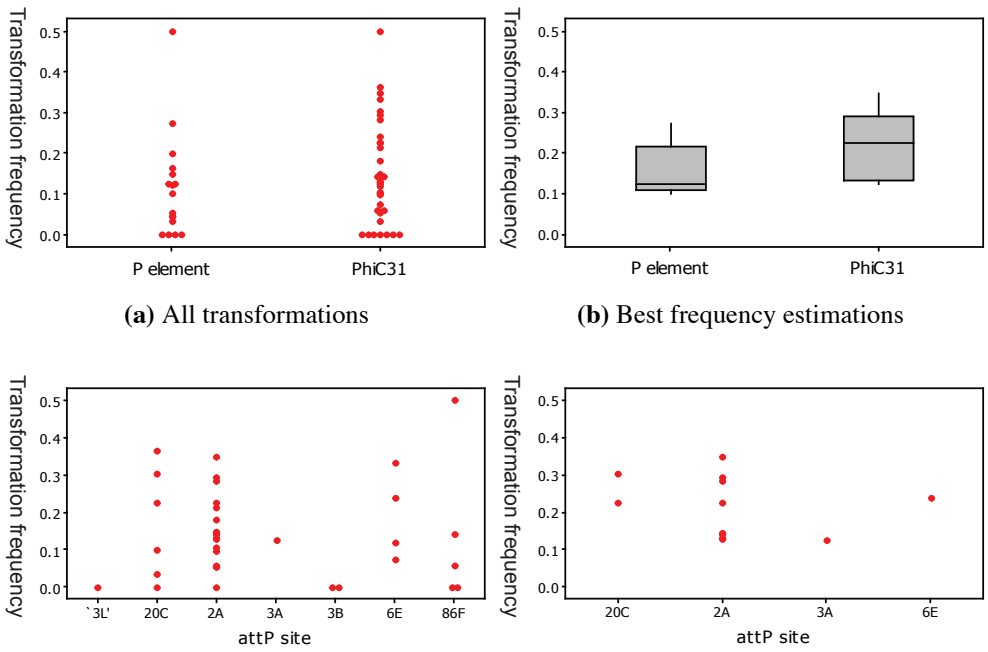
(a) Landing site detail

From 4:

```
         10         20         30         40         50
AWWWRWSGMTARTATTKACGTTTGCGAGACATCTATATGTTCGAACCGAC
TWWWYWSCKATYATAAMTGCAAACGCTCTGTAGATATACAAGCTTGGCTG

         60         70         80         90        100
ATTCCCTACTTGTACWMCTGGTATATCTCCGCCTGTCTAAGCCTTAAGTC
TAAGGGATGAACATGWKGACCATATAGAGGCGGACAGATTCGGAATTCAG

        110        120        130        140        150
AGCTCATTTCCTAAATGCTCACTAATCATCCAAACTACCGGGGTGGACAC
TCGAGTAAAGGATTTACGAGTGATTAGTAGGTTTGATGGCCCCACCTGTG

        160        170
GTACGCGGGTGCTTACGACCGTCAGTCGA
CATGCGCCCACGAATGCTGGCAGTCAGCT
                          3
```

(b) Example sequence from primer 4

Figure 4.6: Design of experiments to confirm transformation. (a) Schematic of the integrated reporter cassette in a landing site. Transformations were confirmed with PCR within the cassette (primer pair '1'), and targeting was confirmed with PCR across the cassette / landing site boundary (primer pair '2'). Genomic location was confirmed by digestion with *e.g. Msp* I, ligation to circularise, inverse PCR using primer pair '3', and sequencing from primer '4'. (b) Example sequence from an insertion of *arm-lacZ* at 20C. Green is landing site; red is genomic. The *Msp* I site (CCGG) is in bold. The position of primer 3 is indicated. Genomic sequences from all four sites are listed in Appendix G.



(a) All transformations

(b) Best frequency estimations

(c) ϕC31 transformations by landing site

(d) Best frequency ϕC31 estimations

Figure 4.7: Transformation frequencies. Transformation frequencies were calculated from number of transgenic lines per fertile G_0 (Appendix G), for all *P*-element and ϕC31 experiments. (a) Individual value plot of all experiments. (b) Box plot of limited experiments. (c) Individual values of ϕC31 sites per landing site (d) Individual values of limited ϕC31 data set; insufficient points for a boxplot. Limited sites were those with $np \geq 5$, where n is sample size (number of G_0), and p is frequency of transformation, thus eliminating poor frequency estimators of low sample size or low proportions.

4.5.2 Modified β-galactosidase assay to measure activity from the GMR-hsp70 promoter

In contrast to expression from the armadillo promoter, expression from the *GMR-hsp70* promoter limited *lacZ* expression to eyes and other neuronal cells (Section 1.3.4). Therefore, the protocol to assay β-galactosidase activity needed to be modified to detect expression in adult heads.

β-galactosidase activity from *armadillo* was measured by the absorbance at 574 nm of product resulting from catalysis of CPRG (Section 2.9). A one milli-litre homogenisation of nine hemi-sected females, or 12 hemi-sected males (being smaller), provided enough homogenate to allow triplicate samples of (usually 25 or 50 μL) homogenate, each with sufficient transgenic β-galactosidase to give spectrophotometric readings in the linear range. Triplicate homogenate samples of 15 μL were subsequently measured for total protein mass, using a Bradford assay (Bradford, 1976) standardised to bovine γ-globulin. Enzymatic activity can be reported as specific to total protein, or per wet weight of fly tissue.

Due to the smaller mass of fly heads in comparison to hemisected adults, the number of flies and homogenate volumes were adjusted to allow readings in the linear range. Fifteen female or twenty male heads were homogenised in 1 mL, and triplicate volumes of between 12.5 and 100 μL were assayed, depending on the activity of the transgene. The total protein was established from triplicate 75 μL samples, and the globulin standard assay adjusted accordingly. The small mass of fly heads prohibited accurate measurements of wet weight with a fine balance (Metler Toledo Classic Plus AB135-3/FACT), precise to 10 μg, but in my experience could not be relied on for accuracy below 100 μg. However, the calculation of activity relative to wet weight relies on an estimation of volumes that vary with homogenisation efficiency, and is thus inferior to specific activity.

The specificity of the modified assay was confirmed. Endogenous *Drosophila* β-galactosidase is expressed in limited adult tissues, including the ventriculus and the reproductive tract (Glaser et al., 1986; Schnetzer and Tyler, 1996). Lower levels are also present in the testes, antennae, brain, and a few specialised cells throughout the thorax, but are difficult to detect at the pH used to assay *E.coli* β-galactosidase. The lack of endogenous contribution in heads was confirmed by an assay with the *y w* injection stock (lab stock #4). Likewise, to confirm no background contribution from the increased concentration of *white*-deposited eye pigment, wild-type Canton S (lab stock #6) flies were assayed. Neither line gave any significant absorbance at 574 nm (not shown). Thus the modified assay was specific to transgenic β-galactosidase.

4.5.3 Acquired compensation detected at defined X chromosome sites

Prior to the comparison of *GMR-lacZ* and *GMR-hid*, it was first established that insertions at each of the *attP* sites on the X chromosome could acquire dosage compensation, by inserting an *arm-lacZ* cassette to each location. Insertions of *arm-lacZ* on the X chromosome generally acquire a significant degree of male-specific enhancement (Fitzsimons et al., 1999; Henry et al., 2001; Weake and Scott, 2007). An approximately 300 bp *Eco* RI–*Eco* RI fragment of pTAattB containing the *attB* site was inserted into the *Eco* RI site of pCaSpeR-arm-βgal to create the plasmid ALattB (Section 2.3.11, Figure 4.1). Sequencing across the insertion site with the primer armlacZseq confirmed the insertion and orientation of the *attB* sequence (not shown).

ϕC31 transformations of *Drosophila* were conducted (Section 2.2.5) to target the *arm-lacZ* cassette to the X-linked *attP* sites at cytological position 2A, 3A, 3B, 6E, and 20C. As discussed above, only insertions at 2A, 6E, and 20C proved useful. The β-galactosidase activity of *arm-lacZ* at each of these sites was measured in hemisected adults (Section 2.9, Table 4.2). To establish the degree of male-specific up-regulation acquired when inserted on the X chromosome, a ratio of male:female activity was established from flies carrying one copy of *arm-lacZ* (Figure 4.8a). A significant degree of male-specific enhancement was detected at each of the three X-linked locations. This response was also observed with a line carrying a *P* element-generated X-linked insertion, previously reported as being able to acquire compensation (Weake and Scott, 2007). To determine if the increase was sufficient to compensate for the male dose deficiency, separate measurements were made to obtain the ratio of activity from single copy (hemizygous) males to two copy (homozygote) females (Figure 4.8b). At all three locations, the male-specific increase was to a level sufficient to compensate for the decreased dose; the insertions had acquired dosage compensation. This agrees with the reported response of the *P* element insertion of *arm-lacZ* (Weake and Scott, 2007), reproduced here for comparison. As all three locations were capable of sustaining dosage compensation of an introduced transgene, all subsequent insertions should thus be equally susceptible to a similar up-regulation.

Table 4.2: β-galactosidase activity in hemisected flies carrying *arm-lacZ* at X-linked *attP*sites

			Heterozygous (1 dose)			Homozygous (2 doses)		
			Activity[a]		**Ratio**[b]	Activity		**Ratio**
Construct	Location	Stock	Male	Female	**Male/Female**	Male[c]	Female	**Male/Female**
arm-lacZ	*P*(10D8)	V40	0.50±0.04	0.26±0.01	**1.92**±0.21			
arm-lacZ	2A	C66	0.56±0.02	0.37±0.03	**1.53**±0.18	0.53±0.01	0.53±0.02	**1.01**±0.06
	6E	C67	0.50±0.04	0.35±0.02	**1.46**±0.08	0.43±0.02	0.50±0.02	**0.85**±0.02
	20C	C68	0.47±0.05	0.26±0.03	**1.86**±0.07	0.42±0.01	0.44±0.02	**0.95**±0.07

[a] Mean OD/min/mg protein, ± 1 SE, n=3
[b] Mean male/female activity, ± 1 SE, n=3
[c] X-linked males hemizygous; only 1 dose of transgenic cassette

4.5.4 GMR-hid at compensation-permissive sites acquired very little compensation

The behaviour of *GMR-hid* at each of the useful *attP* sites was then determined. An approximately 300 bp *Eco* RI–*Eco* RI fragment of pTAattB containing the *attB* site was isolated, the cohesive ends made blunt with Klenow fragment, and inserted into the *Sma* I and *Eco* RV sites of pBS II KS (Stratagene) to create pBSattB (Section 2.3.11, Appendix I). Sequencing with the M13fwd primer confirmed the insertion and orientation of the *attB* site (not shown). The *Xho* I site of pBSattB was removed by digestion with *Xho* I, the creation of blunt ends with Klenow fragment, and re-ligation at a low concentration. The resulting plasmid, pBSattB-X (Appendix I), was verified by sequencing with both M13fwd and M13rev primers (not shown). A 5.8 kb *Hin* dIII–*Hin* dIII fragment of pGMR-1, containing the GMR-hsp70TATA-hsp70 p(A) expression cassette and mini-*white* gene, was inserted into the *Hin* dIII site of pBSattB-X to create pBSw$^+$GMRattB (Appendix I). Finally, a 3.9 kb *Eco* RI–*Eco* RI fragment of pBS SK hid cDNA containing the *hid* cDNA was inserted into the *Eco* RI site of pBSw$^+$GMRattB to create pGHattB (Figure 4.1).

ϕC31 transformations of *Drosophila* were conducted (Section 2.2.5) to target the *GMR-hid* cassette to each of the landing sites at 2A, 6E, and 20C. Two to four insertion events at each locus were characterised. All insertions co-segregated with *3xP3-RFP* and were genetically linked to the X chromosome. All insertions gave bands of the expected sizes from PCR both within the transgenic cassette and across the cassette / landing site boundary (not shown).

A large scale observation of the eye size was conducted for several insertions of *GMR-hid* at each position on the X chromosome (Section 2.8, Appendix D). Bottles of at least 200 flies were counted (except for insertion 4 at 20C, with about 100 flies / bottle). As separate insertion events gave similar phenotypes, representative lines were chosen, and the average male eye compared to the average female eye (Figure 4.8c). Although the eye sizes varied, all lines displayed only a slight male-specifc reduction in eye size, consistent with X-linked insertions of *GMR-hid* by *P* element (Section 4.4.1). In no case was the reduced eye size equivalent to that of two-dose females, which all displayed an ablated eye (not shown). Thus at all three locations where *arm-lacZ* had acquired a significant degree of dosage compensation, *GMR-hid* did not.

4.5.5 GMR-lacZ at compensation-permissive sites acquired very little compensation

Finally, the response of X-linked *GMR-lacZ* to the endogenous compensation machinery was established. To directly compare *GMR-lacZ* to *GMR-hid*, a targeting cassette was designed that would differ from that of *GMR-hid* only in the sequence of the gene itself (Figure 4.1). A 3.5 kb *Sma* I–*Xba* I fragment of pCaSpeR-arm-βgal containing the *lacZ* gene was isolated, the

cohesive *Xba* I end made blunt with T4 polymerase, and the fragment inserted into the *Sma* I site of pBC KS+ (Stratagene) to create pBClacZ (Section 2.3.11, Appendix I). The BClacZ plasmid was digested with *Eco* RI, the cohesive ends made blunt with T4 polymerase, then the DNA digested with *Bam* HI. Due to the difficulty of separating the 3.5 kb *lacZ* fragment from the 3.4 kb pBC backbone, the mixture of fragments was cloned into the *Bgl* II and *Stu* I sites of pBSw^{+}GMRattB. Clones that were resistant to ampicillin but sensitive to chloramphenicol represented an insertion of the 3.5 kb *Bam* HI–*Eco* RI *lacZ* fragment, which created pGLattB (Figure 4.1).

ϕC31 transformations of *Drosophila* were conducted (Section 2.2.5) to target the *GMR-lacZ* cassette to each of the landing sites at 2A, 6E, and 20C. Two to four lines of insertion events at each site were characterised. All insertions co-segregated with *3xP3-RFP*, and were genetically linked to the X chromosome. All gave bands of the expected sizes in PCR tests for insertion and targeting (not shown). Several lines of each insertion were retained (Appendix H).

The activity from *GMR-lacZ* was then measured (Section 2.9, Table 4.3). This activity was measured with the modified β-galactosidase assay designed for adult heads (Section 4.5.2). The activity of *arm-lacZ* at 2A was also measured in adult heads for comparison. Due to the surprising nature of the *GMR-lacZ* results, the activity in several of the lines resulting from independent insertions (but genetically identical) was established. In contrast to the full compensation of *arm-lacZ*, males carrying one dose of *GMR-lacZ* had only slightly more β-galactosidase than one-dose females (Figure 4.8d). Furthermore, the slight increase was in no case sufficient to account for the difference in dose, being equal to about half the level in females carrying two copies (Figure 4.8e). Clearly, the *lacZ* gene can be compensated at these locations. Therefore, it was concluded that GMR-mediated expression of a transgene is insufficient to allow acquisition of full dosage compensation. Any attempts to screen for additional factors that influence dosage compensation that were based on a GMR expression cassette would thus be doomed to failure.

Although the unusual behavior of GMR-based constructs rendered the promoter unsuitable for the proposed reporter gene, this observation was in itself interesting. As the GMR-based constructs failed to acquire dosage compensation, they may have lacked something that compensated genes normally possess. This is explored in Chapter 5. Furthermore, a direct comparison of *arm-lacZ* and *GMR-lacZ* reveals that differing promoters can affect recruitment of dosage compensation. This conclusion is the principal finding of this investigation, and is discussed further below (Section 6.1).

Table 4.3: β-galactosidase activity in heads of flies carrying *arm-lacZ* or *GMR-lacZ* at X-linked *attP* sites

			Heterozygous (1 dose)			Homozygous (2 doses)		
			Activity[a]		**Ratio**[b]	Activity		**Ratio**
Construct	Location	Stock	Male	Female	**Male/Female**	Male[c]	Female	**Male/Female**
arm-lacZ	2A	C66	0.35±0.03	0.20±0.03	**1.84**±0.14	0.31±0.02	0.32±0.01	**0.95**±0.04
GMR-lacZ	2A	C76	2.35±0.07	1.85±0.04	**1.28**±0.06	1.28±0.05	2.52±0.11	**0.51**±0.02
		11GL3	2.23±0.01	1.72±0.03	**1.30**±0.03			
		11GL4	2.36±0.08	1.98±0.17	**1.21**±0.07			
	6E	C78	1.81±0.11	1.60±0.14	**1.13**±0.04	0.95±0.02	1.79±0.08	**0.53**±0.01
		13GL6	2.22±0.04	1.94±0.08	**1.15**±0.06			
	20C	C81	2.11±0.02	1.72±0.04	**1.23**±0.03	1.13±0.06	2.14±0.11	**0.53**±0.03

[a]Mean OD/min/mg protein, ± 1 SE, n=3
[b]Mean male/female activity, ± 1 SE, n=3
[c]X-linked males hemizygous; only 1 dose of transgenic cassette

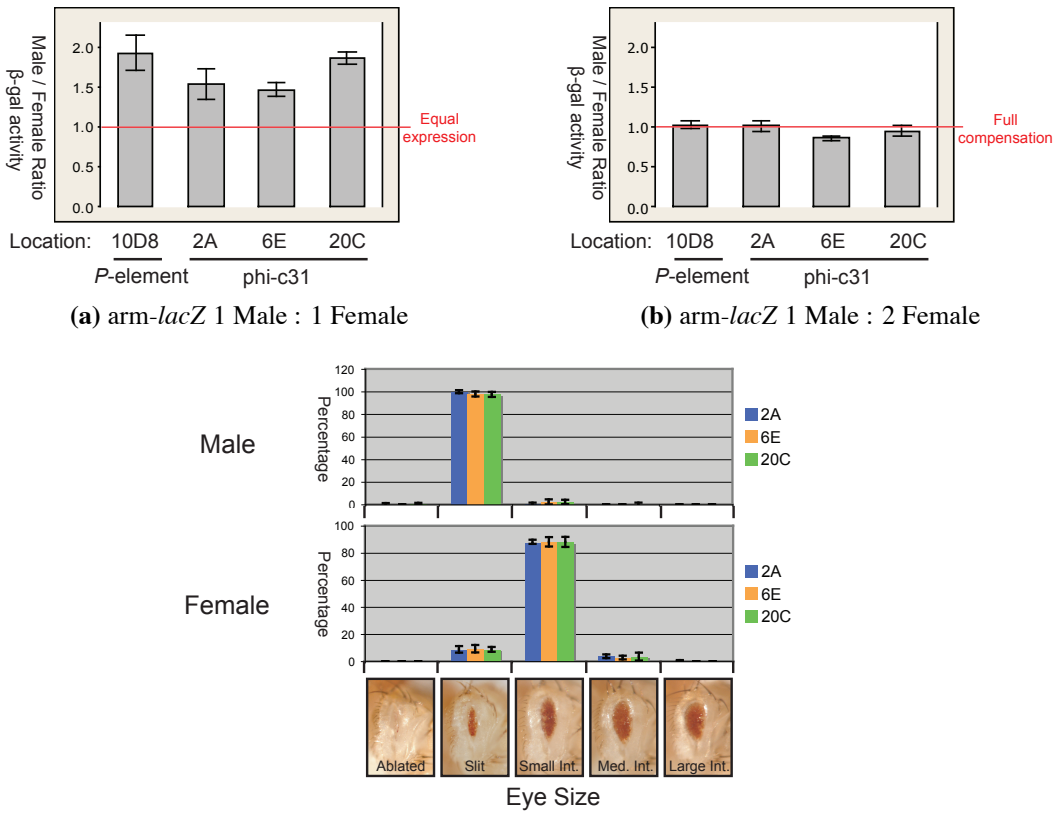
4.6 Advantages of ϕC31 targeted integration for studying dosage compensation

To accurately compare different reporter constructs and observe small changes in *lacZ* expression, it was necessary to adopt a targeted transformation system. Through repeated targeting to a defined landing site, any number of constructs could be accurately compared without position effects. By eliminating the need to observe responses at many locations, a targeted system requires creation of only one transgenic line per construct, significantly reducing the workload. Theoretically, any line lost could also be re-generated at a later stage.

Several systems were considered. Recombination-mediated cassette exchange allows the integration of a cassette flanked by heterologous *FRT* sites when supplied with FLP recombinase (Schlake and Bode, 1994), and works in *Drosophila* (Horn and Handler, 2005). However, at the time of investigation a consensus system was not fully developed. An analogous system with Cre recombinase and *loxP* sites was more established in *Drosophila*, but no lines with landing sites on the X chromosome were available (Oberstein et al., 2005). Cassette exchange with a ϕC31 recombinase requires a cassette to be flanked by inverted *attB* sites, and can mediate insertions in both orientations (Bateman et al., 2006). The ϕC31-mediated system described in the introduction (Section 1.3.1) allows targeted insertion of a whole plasmid to pre-constructed genomic *attP* sites, several of which were available on the X chromosome (Bischof et al., 2007). Furthermore, as it relies on only one recombination event instead of two as in the cassette exchange systems, transformation by ϕC31 is especially efficient. With only one recombination event, insertions can also only occur in one orientation, thus eliminating subtle effects on transgene expression. The only advantage of cassette-exchange systems is that no extra plasmid DNA is inserted. Excess bacterial DNA can repress transcription from ϕC31-inserted transgenes in mammalian systems (Chen et al., 2003; Ehrhardt et al., 2005). However, any effect of the surrounding plasmid DNA in the anticipated ϕC31 transformations would be equal across all compared constructs.

The system was successfully adopted to our laboratory conditions, and only minor abnormalities detected (Section 4.5.1). Plasmids were modified to carry an full-length *attB* site (a gift from M. Calos[4]), and inserted into the *zh-attP* sites of several X-linked lines, and one autosomal line (gifts from J. Bischof[5]). One attempt was also made to target an *attP* in the heterochromatin of chromosome 2L (Venken et al., 2006). Initial transformations were conducted with embryos heterozygous for both the *attP* landing site and a *vasa-ϕC31* transgene, but most later experiments were conducted with flies crossed to be homozygous for each element (Section 2.2.4). In a comparison of all *Drosophila* transformations conducted, the frequency of a transformation

[4]Stanford University School of Medicine
[5]University of Zurich



(a) arm-*lacZ* 1 Male : 1 Female

(b) arm-*lacZ* 1 Male : 2 Female

(c) GMR-*hid* (1 copy)

(d) GMR-*lacZ* 1 Male : 1 Female

(e) GMR-*lacZ* 1 Male : 2 Female

Figure 4.8: Acquired compensation of arm-*lacZ*, GMR-*hid*, and GMR-*lacZ* at defined X chromosomal sites. β-galactosidase (β-gal) activity from arm-*lacZ* in hemisected adults (a, b) or from GMR-*lacZ* in adult heads (d, e) was measured at defined X chromosome locations (2A, 6E, 20C). Ratios from one copy male over one copy female (a, d), and over two copy female (b, e) are plotted, as a mean of three experiments ± 1 SE. (c): GMR-*hid* expression from one copy at 2A, 6E, and 20C (lines C71, C72, 17GH3). Flies were sorted according to eye size, and percentage of total flies per category determined. Means of three experiments were plotted, with 95 % confidence intervals.

event was higher for the ϕC31 than the *P* element approach.

Although reducing the number of lines required saved much time and effort, confirmation of a successful targeting event by PCR necessitated more effort than with *P*. This was necessary, as several false positive lines expressed the *white* transformation marker gene, but had incorrectly-targeted insertions of the transgenic DNA. Of the 22 characterised transgenic 'insertions' at 2A, two of the four with pCL21 were ectopic integrations: one on the X chromosome and one on the third. Ectopic insertions can occur at pseudo-*attP* sites of about 25–45 % identity in mammalian cells and *Drosophila* (Thyagarajan et al., 2001; Olivares et al., 2002; Groth et al., 2004; Bischof et al., 2007). It may be co-incidence that both ectopic events occurred in transformations with the same plasmid. Two of the six insertions of *arm-lacZ* at 3A were *white*$^+$, but not targeted correctly. The integrated DNA was autosomal, but could not be detected by PCR within the plasmid. These lines may have resulted from partial integration of the transgenic plasmid.

Selection of a particular *attP* for transgenic comparison is important, due to the random nature of the *mariner* insertions that created the landing sites. No insertions were detected in *zh-attP-3B*, nor at the heterochromatic *attP* of 3L. In addition to the two mis-integrated insertions with *zh-attP-3Aa*, the remaining four were correctly targeted but male lethal. This landing site is in the large second intron of some transcripts of the DNA methylase *trol*, most mutations of which are recessive-lethal (Tweedie et al., 2009). Although both sexes of the landing site are viable, perhaps the extra DNA upon targeted insertion was not tolerated in males. A break-down of transformation rates by landing site did not reveal clear patterns in differing efficiency, mostly due to the small data set. Expression of *arm-lacZ* and *tetO-msl2* from 20C was low, and *white*-deposited pigment difficult to detect especially in females, implying transformation rates may actually be higher than calculated. However, expression of *GMR-hid* and *GMR-lacZ* was no lower at 20C than at other X-linked sites. Consistent levels of expression of all genes were observed at 2A, 6E, and 86F.

Although adequate, the system could be improved in several areas. Expression from the *white* transformation marker in *tetO-msl2*, *arm-lacZ*, *GMR-hid*, *GMR-lacZ*, and indeed any pCaSpeR-derived vector, obscured the *3xP3-RFP* landing site marker and the *3xP3-GFP* marker for the *ϕC31* insertion. Screening for co-segregation with the landing site and loss of the recombinase was thus conducted with the small ocelli on the top of *Drosophila* heads, and was time-consuming. Expression of *white* is especially sensitive to position effect variegation (Hazelrigg et al., 1984; Pirrotta et al., 1985), and thus positive transformant events may go undetected. When directly compared to *white*, EGFP is a superior transformation marker (Handler and Harrell, 1999; Horn et al., 2000), but would conflict with the *3xP3-EGFP* marker of the recombinase. Alternatively, as the lines with landing sites are also mutant for *yellow*, this gene would be an obvious choice of transformation marker.

A further improvement would be the removal of the step to confirm targeting by PCR. A

carefully-designed cassette-exchange system allows genetic detection of correct target events with the loss of a marker gene in the excised cassette (Horn and Handler, 2005). A solution for the whole-plasmid ϕC31 system is the "split-*white*" design where the exons of the *white* gene are split between target and landing site, and transcription occurs only in correctly-targeted insertions (Bischof et al., 2007), although this approach suffers from the conflict with the *3xP3* markers, as discussed above. The split need not be exonic. Perhaps the *3xP3* and *RFP* could be split by a small *attP* site, such that the same *3xP3* promoter could drive transcription of an appropriate gene on the plasmid. Loss of *RFP* expression and gain of another fluoresence in the eye would indicate a correct transformant. In this case the transgenic insertion of *vasa-ϕC31*, itself inserted in a *3xP3-RFP*-marked *attP* site, would need to be replaced. Careful choice of fluorescent proteins would be necessary to ensure no conflict between the three marker genes. Regardless of design, any particular *attP* site of the existing library would need to be first modified to accommodate the new targeting technique, as was envisioned in the design of the *zh-attP* landing sites (Bischof et al., 2007). For a suitably large experiment, this effort would be worthwhile.

Finally, the source of helper transposase could be optimised. Expression of *vasa-ϕC31* from 86F appears to mediate more transformation than when expressed from 102D (Bischof et al., 2007), but a systematic screen of different landing sites may identify a better position. Additionally, a stronger embryonic promoter than *vasa* may yet be developed, which would likely further increase the transformation efficiency.

4.7 Critical issues for measuring β-galactosidase activity

The *E. coli lacZ* gene is a useful reporter gene well established in *Drosophila* (Lis et al., 1983; Wakimoto et al., 1986; Hiromi and Gehring, 1987). By minimising input from endogenous *Drosophila* β-galactosidase, an enzymatic assay on hemi-sected adults (lacking wings and abdomens) specifically measures transgenic β-galactosidase activity, and is appropriate for expression of *lacZ* from constitutive promoters such as *armadillo* (Simon and Lis, 1987; Fitzsimons et al., 1999). However, the *GMR-hsp70* enhancer-promoter construct restricts *lacZ* expression to eye and neuronal tissues (Section 1.3.4). The β-galactosidase enzymatic assay from Simon and Lis (1987) was therefore adapted to measure activity from *GMR-lacZ* in *Drosophila* heads (Section 2.9).

The adapted procedure was sufficient for the required experiments, but could still be improved. To measure male and female activity of one transgenic line in three independent experiments requires only 63 hemi-sected flies, but 105 heads. This significantly increased the time for sample preparation. Although minimised by keeping samples on ice, increased protein degradation prior to homogenisation in phosphate buffer was also likely an issue. Numbers

could not be reduced, as for lines with low β-galactosidase activity, the limit of linearity was approached even with 100 μL of homogenate. By using a spectrophotometric plate reader, the absorbance of smaller volumes could be measured, reducing the number of flies required. Three measurements of male:female (or genotype x:y) activity, with absorbance readings in triplicates, requires 18 samples, and a further 18 for the Bradford assay. Two lines were usually measured at one time, and would use 72 wells of a 96-well plate. The automation would also significantly reduce the labour required.

A high degree of variation in activities was observed between lines that should have been genetically identical, in some cases up to a two-fold difference. The source of this variation remains unknown, but may result from slight differences in population crowding, health or age of the flies, temperature or humidity of lab conditions, or even the exact pH of the assay buffer. Crowding was controlled by setting similar numbers of parent stocks for similar lengths of time, and flies were always assayed at 3 to 4 days age to obtain consistent activity levels. Stocks were raised at 25 °C, but ambient lab temperature during assay preparation varied, as did humidity at all stages. The consistency, quality, and precise amount of culture media also varied. With increased technique proficiency and speed, the variation decreased, but was not eliminated. The activity appeared to decrease proportional to the length of time a transgenic stock had been established, but this remains to be investigated. Most experimental set-ups included an internal genetic control by normalising male to female, or *e.g* Sb^{+} to *Sb*, which should eliminate this variation. However, the balancer chromosome used in this particular cross likely carried many allelic differences that may have affected gene expression. Direct comparisons of absolute activity between stock bottles or lines without first normalising to a reference gene proved problematic (Section 5.3). For such experiments, this variation should be eliminated.

4.8 Summary and future directions

The *roX1 DHS-GMR-hid* system was designed to report small changes in the activity of the MSL complex-mediated dosage compensation as changes in eye size (Section 1.3.5). None of the many lines with any of the variations on this system provided such an effect.

The DHS from *roX1* can attract MSL complex capable of increasing transcription of an *arm-lacZ* transgene in males (Henry et al., 2001), but it had no significant effect on the level of *GMR-hid* (Section 4.1). A nine-copy multimer of the DHS attracts more complex to transgenic insertion sites (Kageyama et al., 2001), and was thus also tested with *GMR-hid*. Despite clearly attracting MSL complex capable of relieving positional repression on the *white* transgenic marker gene in males, no significant up-regulation of *GMR-hid* was observed, and the strength of phenotype did not correlate to MSL levels (Section 4.2). The DHS from 18D10 can similarly attract MSL complex, and a four-copy multimer strongly so (Oh et al., 2004). As a

"high affinity site", it more likely attracts complex sufficient for dosage compensation (Section 1.1.4), but it also had no effect on the sex-specific expression of *GMR-hid* (Section 4.3).

The lack of male-hypertranscription was not a problem of the MSL attractor, but the reporter gene system itself. Lines carrying *GMR-hid* cassettes displayed a range of eye sizes, and were capable of reporting two-fold differences in gene dose. Indeed, the *GMR-hid* system is sensitive enough to identify modifiers of Hid in an analogous mutational screen (Bergmann et al., 1998). However, no line of *GMR-hid* on the X chromosome acquired significant dosage compensation, and those that appeared to did not respond appropriately to levels of MSL complex (Section 4.4). These lines most likely exhibited some location-specific response to sex-specific enhancers, but may have possessed some genetic background that affected Glass or Hid activity. The non-responsiveness of *GMR-hid* was absolutely confirmed by comparison to *arm-lacZ* at defined locations by the ϕC31 recombinase targeting system (Section 4.5). At three distinct X chromosome locations *arm-lacZ* acquired dosage compensation, yet neither *GMR-hid* nor *GMR-lacZ* responded. GMR-mediated expression appeared un-responsive to dosage compensation, and was thus unsuitable for the proposed reporter gene system.

Further development is required to obtain a working system. It is possible that modifications to the GMR cassette may lead to better acquisition of dosage compensation, although the DRE and GAGA motifs effective for *GMR-lacZ* did not work for *GMR-hid* (Section 5.7). Instead, the use of a different eye-specific promoter to drive *hid* expression is suggested. A three-copy multimer of the Pax-6 transcription factor binding site (3xP3) drives strong expression in *Drosophila* eyes (Horn et al., 2000), but is less tissue-specific than GMR, and thus *3xP3-hid* may prove lethal. The new system should be as sensitive to small transcription changes, and as easy to screen for differences, as *GMR-hid*. With use of the targeted ϕC31 transformations to X chromosomal sites known to mediate dosage compensation, the MSL attractor (*e.g. roX1 DHS*) could be omitted. Furthermore, the endogenous X chromosome environment is more likely than autosomal transgene insertions to represent normal dosage compensation. Finally, a male-specific response of the new system should be tested to ensure it is likely due to dosage compensation. The response should not occur when inserted on an autosome, should respond to *hsp83-msl2* in females, and may even be sensitive to MSL levels in males. PCR after ChIP against the acetylated H4K16 would likely detect the transgene in males only. A mutational screen for novel components or modifiers of dosage compensation could then be conducted.

Chapter 5

The deficiency in compensation of GMR-based expression

The observation that a transgene expressed from a GMR-mediated promoter was not able to acquire dosage compensation provided a unique opportunity. It was now possible to identify reasons why GMR-based expression could not acquire adequate dosage compensation, and thus infer knowledge about the basic requirements for *Drosophila* dosage compensation. In related experiments, the relative importance of early expression for dosage compensation was then investigated, as a functional test of predictions on the establishment of dosage compensation (Section 1.4.3).

5.1 Identified differences between arm- and GMR-based vectors

In order to identify likely reasons for the lack of compensation, the *GMR-lacZ* cassette was compared to the *arm-lacZ* cassette (Figure 5.1a, top two constructs). Because *arm-lacZ* responded to the endogenous compensation at the same X-linked sites where *GMR-lacZ* did not (Section 4.5), a comparison of the constructs should identify differences that could explain the differing responses.

Both pALattB and pGLattB (Appendix I) are based on pBR322, and carry the mini-*white* *Drosophila* transformation marker gene. The expression cassette and mini-*white* gene of pALattB are flanked by *P* termini for *P* element transformation, but pGLattB also has part of the 5′ *P* terminus as an artifact of construction. pGLattB contains the f1 phage origin of replication, which pALattB lacks, but this is trivial. Other minor differences include multiple cloning sequences, remnants of *lacZ* selection sequences, and the position of the *attB* (before the pALattB expression cassette, after that in pGLattB).

More significant differences are in the transgenic cassettes. Both cassettes contain an identical fragment of the *lacZ* gene, and thus only differ in the regulatory sequences surrounding *lacZ*. The 3′ UTR and p(A) site from SV40 are used for *arm-lacZ*, whereas *GMR-lacZ* uses *hsp70*-derived sequences. The promoter and 5′UTR regions also differ. The GMR is a 5 copy multimer of the glass response element (Section 1.3.4), upstream of a basal TATA box-containing promoter from *hsp70* that begins transcription immediately before the ORF, to create a transcript of only one exon. In contrast, the endogenous *armadillo* promoter fragment used in *arm-lacZ* (Vincent et al., 1994) contains all intergenic sequence, at least two alternative first exons (Riggleman et al., 1989), and the start of the second exon, until just prior the ATG codon where translation normally begins (Figure 5.1b).

Furthermore, several interesting elements are present in the *armadillo* promoter. *armadillo* contains no consensus TATA box (Riggleman et al., 1989), but instead a software search revealed a cluster of four DRE-like elements (Figure 5.1b), a motif associated with TBP-independent

transcription, and implicated in dosage compensation (Section 1.3.3). Finally, several GAGA (or TCTC) sequences are present in the *armadillo* promoter (Figure 5.1b), which might be able to bind the GAGA factor (van Steensel et al., 2003). The GAGA factor is also genetically linked to dosage compensation, and may play a subtle role in targeting the MSL complex (Greenberg et al., 2004). Although these GAGA sites resemble the core of the MRE consensus sequence (Alekseyenko et al., 2008), the nearest bound MRE site is in the 3′ coding sequence of *armadillo* (A. Alekseyenko, personal communication). It should also be noted that as transcription from the 'E16' exon is favoured over that from the 'E9' (Riggleman et al., 1989), the cluster of DRE-like and most of the GAGA-like elements of *armadillo* are intronic. Four GAGA sites are upstream of the E16 exon, and 12 are in what is normally the first intron. Similar searches of the GMR promoter region show no more DRE- nor GAGA-like elements than would be expected by chance (not shown).

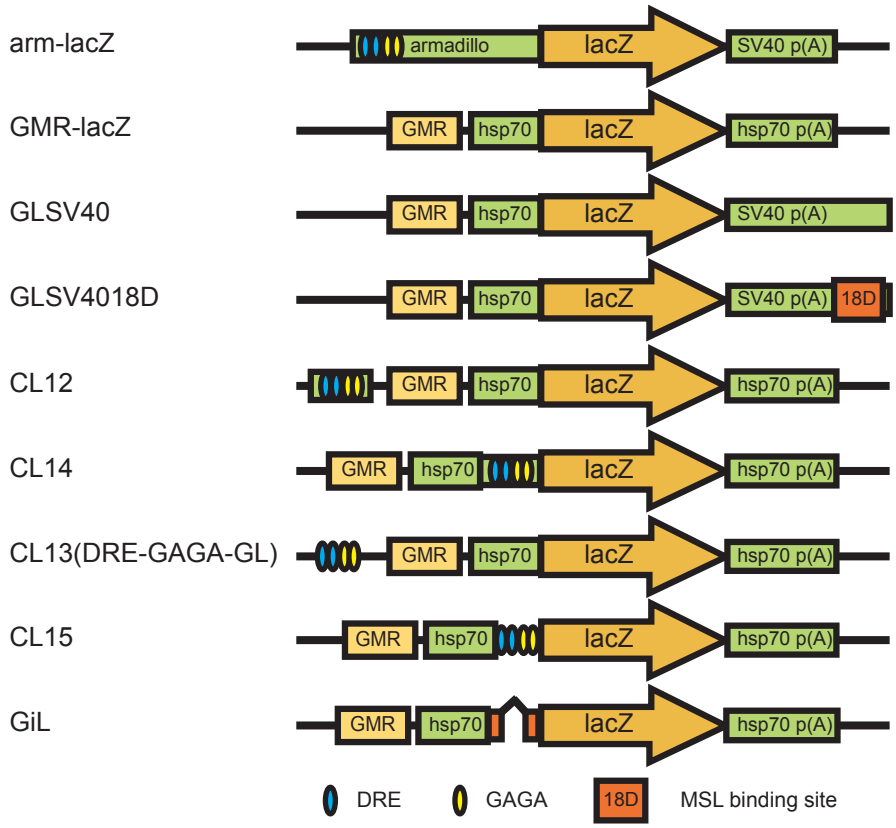
5.2 Promoter modifications increased the male:female expression ratio of GMR-lacZ

To avoid potentially deleterious changes to the *armadillo* promoter, a systematic reverse genetics approach was adopted to alter *GMR-lacZ* with regards to each point of difference from *arm-lacZ*. Any difference that led to an increase in the male:female ratio of expression from an X-linked location above that observed for *GMR-lacZ* should thus be involved in the acquisition of dosage compensation. Each construct was made (Section 2.3.11) as described below, then injected into *Drosophila* appropriate for ϕC31-mediated targeting to the *attP* at 2A, as confirmed by PCR analysis of genomic DNA (Section 2.2.5). The 2A landing site was chosen because of consistently high expression of transgenes, no observed false-positive transformation events, and high transformation efficency (Section 4.5). By targeting all constructs to 2A, the position effect variable, inherent in *P*-element transformations, was removed, and thus only one line of each was assayed for β-galactosidase activity (Section 2.9). The degree of male hyper-transcription acquired was established by comparing the male:female ratio of β-galactosidase activity in flies carrying one copy of each construct (Table 5.1, 'heterozygous' measurements; Figure 5.2a). Separate measurements were made for the ratio of β-galactosidase activity in single copy males to two copy females, to determine if the male-specific increase was sufficient to compensate for the gene dose deficiency (Table 5.1, 'homozygous' measurements, Figure 5.2b). The corresponding responses of *arm-lacZ* and *GMR-lacZ* at 2A (from Table 4.3) are reproduced for comparison.

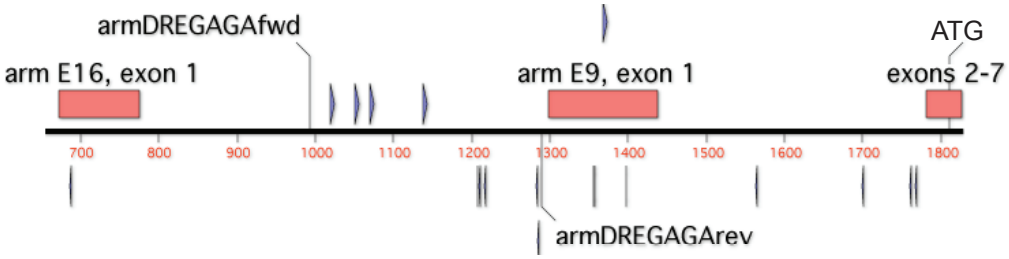
5.2.1 SV40 3′ UTR altered expression level but not male:female ratio

To test the effect of sequences downstream of *lacZ* on dosage compensation, the SV40 3′ UTR and p(A) site was inserted in place of the *hsp70* sequences of the *GMR-lacZ* cassette. The *Hin* dIII site of pGLattB at 2240 was deleted to create pGLattB-H (Appendix I). pGLattB-H was digested with *Eco* RI, treated with Klenow fragment to make blunt ends, then digested with *Hin* dIII. The *lacZ* and 3′ UTR sequences were replaced with a 4.3 kb *Sma* I–*Hin* dIII fragment of pCaSpeR-arm-βgal containing the *lacZ* gene and SV40 sequences to create pGLSV40attB.

Expression of *GMR-lacZ* from 2A was markedly increased with the use of the SV40 sequences in both males and females (Figure 5.2, '*GLSV40*'), suggesting that these 3′ sequences were superior to those of *hsp70* for mediating *lacZ* expression. The male-specific increase ('Heterozygous', '1 Male : 1 Female') appeared to be slightly higher than that of *GMR-lacZ*, but this increase was not sufficient to compensate for gene dose ('Homozygous', '1 Male : 2 Female'). Thus the SV40 sequences were not sufficient for *GMR-lacZ* to acquire full compensation when inserted on the X chromosome. This suggested that the important differences between



(a) Constructs modifying GMR-*lacZ*



(b) The *armadillo* promoter

Figure 5.1: Schematic diagrams of *lacZ* constructs. (a) A comparison between *arm-lacZ*, and *GMR-lacZ*, with the modified constructs below. All sites also contain an *attB* site (not pictured). (b) Detail from the endogenous *armadillo* promoter, with alternative first exons depicted as boxes. E9 and E16 refer to the respective transcripts produced. Sequences identical to, or one mis-match from, the DRE consensus motif WATCGATW (Ohler et al., 2002) are shown as arrowheads above the sequence line, and GAGA or TCTC motifs are shown below the line. Searches performed with MacVector 10.0. The positions of the primers (armDREGAGAfwd and armDREGAGArev) for amplification of a promoter fragment entailing the DRE and GAGA motifs are indicated. Sequence numbers as in Genbank accession number X54468.

arm-lacZ and *GMR-lacZ* were in the 5′ regulatory regions.

5.2.2 Local MSL complex was not limiting

Before examining the 5′ regulatory sequences, it was first investigated whether the amount of MSL complex at the landing site was a limiting factor. Perhaps local concentrations of MSL complex were sufficient to mediate up-regulation of some genes and not others. A single copy of the 510 bp MSL high-affinity site from 18D10 (Section 1.1.4) inserted in the 3′ UTR attracts MSL complex (by immunofluorescent polytene studies), and up-regulates β-galactosidase activity of *arm-lacZ-SV40* at autosomal locations (Anja Schiemann[1], personal communication).

To increase the local concentration of MSL complex around *GMR-hid* at 2A, the same strategy was employed. This places the 18D10 and its corresponding MRE close to the consensus location for MREs (3′ of CDS), but avoids interrupting the coding sequence of *lacZ*. pGLattB-H was digested with *Eco* RI and *Hin* dIII, and treated with Klenow fragment to create blunt ends. The *lacZ* and *hsp70* 3′ UTR sequences were replaced with a 4.85 kb *Sma* I–*Sma* I fragment of "pBS2N17mer HF12-1x12" , containing *lacZ* and the SV40 sequence with 18D10 fragment, to create pGL18DSV40attB (Appendix I).

The effect of the additional 18D sequence should be observed by comparison to the *GMR-lacZ-SV40* cassette (Figure 5.2, *GLSV4018D* to *GLSV40*), to account for the differing 3′ sequences from *GMR-lacZ* alone. The addition of the 18D high affinity site appeared to further increase β-galactosidase activity in both sexes. However, the male-specific up-regualtion was not altered, and the dosage compensation was still incomplete.

5.2.3 Promoter elements increased the male:female ratio

The importance of the 5′ DRE-like and GAGA motifs, and how to test these elements in *GMR-lacZ*, was then examined. Although the DRE promoter elements in the genes for PCNA and DNA Polymerase α are immediately 5′ of the transcription start point (tsp), there is no preferential location in 1941 EST sequences aligned at the tsp (Ohler et al., 2002). In the analysed region -60 to +30, DRE sites occurred with similar frequency between -60 and -20, were slightly less likely to occur in -20 to -1, and were even less frequently observed after the tsp. The DRE from the PCNA gene is able to complement a *PCNA-CAT* construct with mutated DRE equally well at positions between -92 to -1399 (Hirose et al., 1993), indicating that the position relative to the tsp is not critical. However, the cluster of four DRE-like motifs in the *armadillo* promoter are usually in the large first intron (when transcribed with the E16 exon), and this location may have some importance for dosage compensation. Likewise, consensus GAGA motifs are predominantly bound only in introns throughout transcribed sequences (van Steensel et al., 2003).

[1]Massey University, Palmerston North, New Zealand

But well-characterised clusters of GAGA motifs are usually 5′ of the tsp, mostly within 200 bp (Granok et al., 1995). Thus the important locations for both DRE and GAGA elements was likely upstream of the tsp, although the relative distance was less critical. Alternatively, an intronic location in the transcribed region might be suitable for the proposed experiments.

A 349 bp fragment of the *armadillo* promoter containing four DRE-like motifs and 5 GAGA elements (Figure 5.1b) was amplified from pCaSpeR-arm-βgal (200 ng) by PCR with the armDREGAGAfwd and armDREGAGArev primers, which also added flanking *Xho* I and *Eco* RI sites (Section 2.3.10). The fragment was inserted into the *Xho* I site of pGLattB, upstream of both the tsp (at -209) and GMR enhancer, in the same direction as found in *armadillo*, to create pCL12 (Figure 5.1a). To test the importance of the motifs in a position similar to that in *armadillo*, they were also placed downstream of the tsp (at +146), in the 5′ UTR. The *Eco* RI site of pGLattB at 10 972 was destroyed to created pGLattB-RI (Appendix I). The *armadillo* promoter fragment was then inserted in the remaining *Eco* RI site of pGLattB-RI, in the same orientation, to create CL14 (Figure 5.1a). Unfortunately, as several ATG sequences were present in the inserted sequence, it was possible that translation would be incorrectly initiated and a non-functional protein produced.

To directly test the influence of the DRE-like and GAGA motifs, and to prevent erroneous translation initiation, four oligonucleotides (Table A.2, "3DRE4GAGA ... ") were designed to create a 132 bp fragment of three DRE sites and four GAGA motifs, without any ATG trinucleotides. Although not tested, the design also allowed future removal of either DRE elements or GAGA elements. A pair of *Nhe* I sites flanked the DRE sequences, and a pair of *Bgl* II sites flanked the GAGA sequences. Furthermore, the entire sequence was flanked with *Xho* I sites, and possessed overhangs that created cohesive *Nhe* I (5′) and *Eco* RI (3′) sites, to facilitate cloning. Thus the restriction sites and elements in 5′ to 3′ order: *Nhe* I cohesive end, *Xho* I, *Nhe* I, 3 DRE motifs, *Nhe* I, *Bgl* II, 4 GAGA motifs, *Bgl* II, *Xho* I, *Eco* RI cohesive end. The three directly repeated DRE sites closely mimicked that (3DREP) of Seto et al. (2006), bound by DREF in a mobility shift assay (Section 1.3.3). The four alternating TCTC and GAGA motifs were highly similar to an oligonucleotide that was protected from DNAseI digestion by GAF, and bound in a mobility shift assay (Katsani et al., 1999). The 3DRE4GAGA oligonucleotides were inserted in the same *Xho* I and *Eco* RI sites of pGLattB and pGLattB-RI, to create plasmids CL13 and CL15, respectively (Figure 5.1a).

Sequencing of plasmids CL12, 13, 14, and 15 with the primers white3f and lacZnewrev confirmed the correct insert orientation (not shown), although reverse orientation clones were in all cases also isolated. Note that due to difficulties in the synthesis of long oligonucleotides, pCL13 contained an unexpected extra G between the first two DRE motifs that should not significantly alter the DREF binding, and pCL15 lacks a T in the first *Bgl* II site that would prevent the subsequent excision of GAGA sequences. Although not tested, a correct pCL15

clone was later isolated. Furthermore, minor differences in the *arm* promoter from the reported sequence (Riggleman et al., 1989), but consistent with the genomic sequence, were detected in both pCL12 and pCL14. Finally, all clones contained a single nucleotide G to C mutation that changes amino acid #39 of the *Adh-lacZ* fusion from E to D; likely to be also present in pCaSpeR-arm-βgal, and thus unlikely to destroy the activity of *lacZ*.

The β-galactosidase activities and male:female ratios were then determined for all four cassettes expressed from 2A (Table 5.1, Figure 5.2). Unfortunately, the exonic insertion of the *armadillo* fragment (*CL14*) did not result in functional protein, presumably due to the presence of extraneous ATG codons upstream of the correct translational start site. The specific DRE and GAGA oligonucleotides in this position (*CL15*) avoided this problem and produced functional β-galactosidase. In comparison to *GMR-lacZ*, no difference in the level of β-galactosidase activity was observed. Neither did this modification affect the male-specific increase nor the dosage compensation. Similarly, the insertion of the *armadillo* fragment upstream of the GMR and tsp (*CL12*) did not substantially alter β-galactosidase activity, male-specific increase in expression, nor dosage compensation.

However, the specific DRE and GAGA motifs in the *GMR-lacZ* promoter (*CL13*) had a noticeable effect on the expression of the cassette. This modification substantially reduced β-galactosidase activity in both sexes, but increased the male-specific increase in expression. Surprisingly, the increase was not sufficient to correspondingly alter the compensation of dose deficiency, an observation explored below. As the difference in male up-regulation from *GMR-lacZ* was slight, the measurements were repeated in two further (and presumably genetically identical) transgenic lines of *CL13*. Some variation in the absolute β-galactosidase assays from day to day is inevitable with this method (discussed in Section 4.7), but the male:female ratios are more consistent. The triplicate readings of *CL13* (and *GMR-lacZ*) effectively controlled this slight variation, giving an average male:female (1:1) ratio of 1.70 for *CL13*, and 1.26 for *GMR-lacZ*. A Student's t-test of the $\log_2$ values of these ratios showed the effect to be tending towards statistical significance ($T = 4.17$, d.f. $= 2$, $p = 0.053$). A similar result was obtained if the logarithm transformation was not applied ($T = 3.73$, d.f. $= 2$, $p = 0.065$), but this is less valid for a comparison of ratios. Thus the DRE and GAGA motifs in the promoter of *GMR-lacZ* decreased the absolute expression in both sexes, but increased the male-specific up-regulation acquired when inserted at 2A. This construct (*CL13*) is hereafter referred to as *DRE-GAGA-GL*.

5.2.4 Addition of an intron affected the male:female ratio

As the cluster of DRE-like and GAGA motifs in *armadillo* are usually in the large first intron, a similar arrangement was attempted in *GMR-lacZ*. An intron was inserted after the transcription start point in *GMR-lacZ*, with the intention of then inserting both the fragment of the *armadillo* promoter, and the synthetic oligonucleotide with consensus sites. This approach would also

avoid the aberrant translation initiation of *lacZ*. To avoid likely involvement with dosage compensation machinery, an intron was chosen from autosomal, tissue-specific genes. The small second intron of *neither inactivation nor afterpotential E* (*ninaE*) also contained an *Mfe* I site useful for cloning purposes. The splice acceptor and donor sites were in good agreement with the consensus *Drosophila* sequences (Weir and Rice, 2004). Curiously, both 'Spliceport' and 'Genie' algorithms (Dogan et al., 2007; Reese et al., 1997) predicted the $3'$ splice acceptor to be a strong splice donor site, similar to the "re-splicing" site of *Ubx* (Hatton et al., 1998), but this should not affect the intended use. Finally, potential branch sites were all downstream of the *Mfe* I site (which is 62 nt from the $3'$ acceptor site), so not likely to be disrupted by subsequent cloning steps.

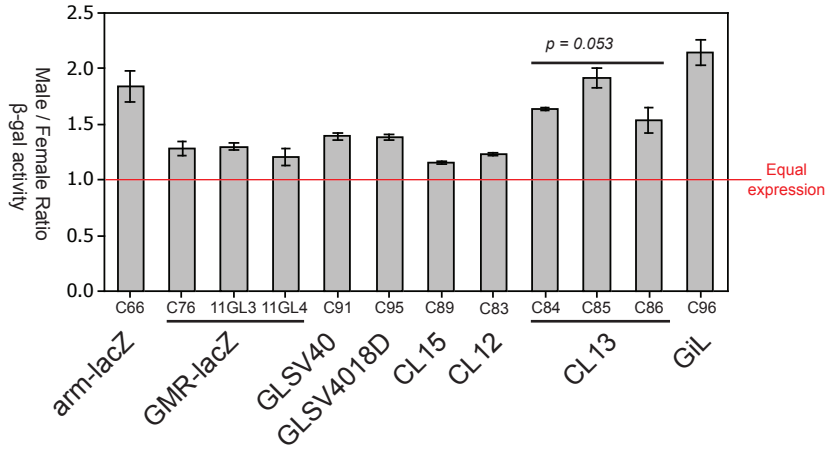
A 214 bp fragment encompassing the *ninaE* second intron was amplified from Canton S genomic DNA by PCR with the primers ninaEin2F and ninaEin2R, designed to append *Eco* RI sites, and destroy the ATG just after the splice acceptor site (Section 2.3.10). The amplified fragment was inserted into the *Eco* RI site of pGLattB-RI to create pGiLattB (Appendix I). Sequencing with the primers white3f, lacZ493R, and hsp70TATAfwd revealed what appeared to be an allelic difference from the expected *Drosophila* genome sequence: an equal length but different sequence between the $5'$ donor site and the *Mfe* I site, almost identical to *ninaE* from the rucuca strain (Shapiro et al., 2007). Further digestion analysis and sequencing with the white3r primer identified an unexpected extra *Mfe* I site at 6259 of pGiLattB, which was removed to create pGiL-Mfe (Appendix I, Figure 5.1a, 'GiL'). This construction was confirmed by sequencing with the primers hsp70TATAfwd and lacZ493R (not shown). The injected version of the construct was pGiL-Mfe.

Addition of the intron had an unexpected effect on the response of *GMR-lacZ* at 2A (Figure 5.2). The expression of *GMR-lacZ* was not disrupted by the presence of the *ninaE* intron ('*GiL*'), and the level of activity was not greatly different from that without the intron. However, the male-specific up-regualtion was greatly increased; a near doubling in comparison to the single copy female. As with *DRE-GAGA-GL*, the increased up-regulation was not sufficient to compensate for the dose deficiency (discussed below). A Student's t-test of the $\log_2$ transformations of the male:female (1:1) ratio readings against the average *GMR-lacZ* response (1.26, or 0.333) revealed the effect to be highly statistically significant ($T = 9.46$, $p = 0.011$), although this analysis suffers from a lack of repetition. This surprising observations implied that the very presence of the intron was sufficient to increase the ability of *GMR-lacZ* to acquire male-specific up-regulation.

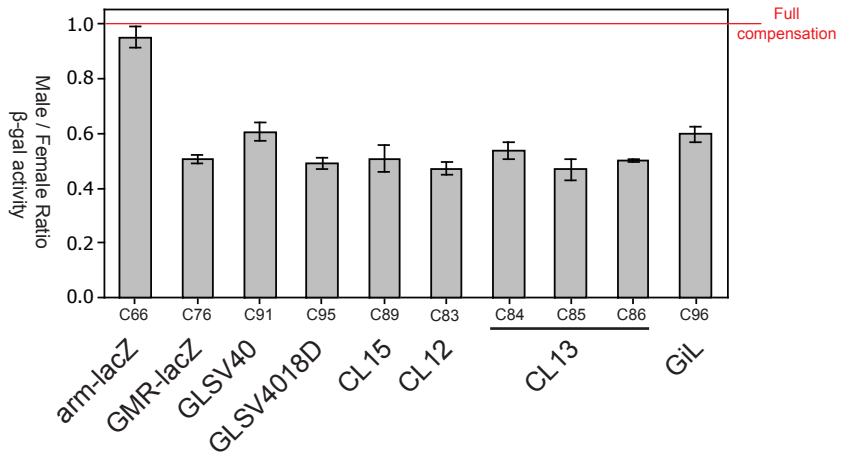
Table 5.1: Altered β-galactosidase activity in flies carrying modified GMR-lacZ constructs at 2A

		Heterozygous (1 dose)			Homozygous (2 doses)		
		Activity[a]		**Ratio**[b]	Activity		**Ratio**
Construct	Stock	Male	Female	**Male/Female**	Male[c]	Female	**Male/Female**
arm-lacZ	C66	0.35±0.03	0.20±0.03	**1.84**±0.14	0.31±0.02	0.32±0.01	**0.95**±0.04
GMR-lacZ	C76	2.35±0.07	1.85±0.04	**1.28**±0.06	1.28±0.05	2.52±0.11	**0.51**±0.02
	11GL3	2.23±0.01	1.72±0.03	**1.30**±0.03			
	11GL4	2.36±0.08	1.98±0.17	**1.21**±0.07			
GLSV40	C91	4.07±0.07	2.93±0.07	**1.39**±0.04	4.57±0.69	7.45±0.73	**0.61**±0.03
GLSV4018D	C95	5.38±0.12	3.88±0.09	**1.39**±0.03	4.59±0.37	9.31±0.54	**0.49**±0.02
CL14	C87	0.01±	0.01±				
CL15	C89	2.19±0.12	1.90±0.11	**1.15**±0.01	2.25±0.12	4.44±0.19	**0.51**±0.05
CL12	C83	2.08±0.02	1.68±0.01	**1.24**±0.01	2.02±0.11	4.29±0.23	**0.47**±0.02
CL13(DRE-GAGA-GL)	C84	0.35±0.00	0.21±0.00	**1.64**±0.01	0.29±0.01	0.54±0.03	**0.54**±0.03
	C85	0.30±0.03	0.16±0.01	**1.92**±0.09	0.26±0.01	0.55±0.04	**0.47**±0.03
	C86	0.64±0.03	0.42±0.03	**1.53**±0.11	0.68±0.03	1.35±0.06	**0.50**±0.00
GiL	C96	2.39±0.02	1.12±0.07	**2.15**±0.12	2.51±0.08	4.20±0.17	**0.60**±0.03

[a]Mean OD/min/mg protein, ± 1 SE, n=3
[b]Mean male/female activity, ± 1 SE, n=3
[c]X-linked males hemizygous; only 1 dose of transgenic cassette



(a) 1 Male : 1 Female



(b) 1 Male : 2 Female

Figure 5.2: Modifications to GMR-*lacZ* altered the X chromosome expression ratio. Male to female ratios of β-galactosidase (β-gal) activity in adult fly heads for lines carrying the indicated constructs (as per Figure 5.1a) at X-linked insertion point 2A. Ratios from (single copy) males over single copy females (a) and over double copy females (b) are shown. Means of three experiments were plotted, $\pm$ 1 SE. The p value is from a Student's t-test of the three *CL13* against the three *GMR-lacZ* values.

5.3 Females with two DRE-GAGA-GL doses were affected by transvection

There appeared to be a disparity between the measurements of male up-regulation and dosage compensation. For both *DRE-GAGA-GL* and *GiL* cassettes a significant degree of male-specific up-regulation had been observed, but this increase was not sufficient to compensate for the dose deficiency. This was most surprising for *GiL*, where the two-fold higher expression in males would be expected to fully compensate for the XY male to XX female deficiency. As both experiments involved genetically identical males, the problem lay with females. If female activity levels in the homozygote were not two-fold higher than in the heterozygote, then a different amount of compensation would have been necessary to fully compensate for the dose deficiency. This 'transvection' effect can result from the close association of transcriptional regulators, caused by the pairing of homologous chromosomes in *Drosophila* (Section 1.2).

To remove the effects of transvection, measurements were made of females carrying two doses at different locations. Firstly, ϕC31-mediated transformations of *Drosophila* were conducted to target the *DRE-GAGA-GL* and *GiL* cassettes to the X-linked *attP* at 6E (Section 2.2.5). Ratios of β-galactosidase activity were measured in heads of flies carrying each of the constructs (Section 2.9, Table 5.2). When expressed from 6E, *DRE-GAGA-GL* mediated similar β-galactosidase activities, male-specific up-regulation, and dosage compensation as was observed at 2A (Table 5.1). In contrast, the β-galactosidase activity of *GiL* at 6E was much lower than that at 2A. Somewhat surprisingly, the male-specific up-regulation at 6E is also much lower than that at 2A.

Reciprocal crosses were performed between lines carrying homozygous insertions at 2A and at 6E for each of the cassettes *GMR-lacZ*, *DRE-GAGA-GL*, and *GiL*, and the β-galactosidase activity measured in the offspring (Section 2.2.4, Table 5.3). However, comparisons between all cases were difficult. By normalising male expression to female expression, male:female ratios effectively control for observed line, day to day, and even replicate variation in absolute β-galactosidase activity. Comparisons of activity level per sex of each construct at each location did not benefit from this control, and thus can vary substantially. For example, three assays of the expression level from males with one dose of *GMR-lacZ* at 6E give the estimates 0.95, 1.81, and 2.22 OD/min/mg (Table 4.3). Nevertheless, the best estimates of absolute activity at each location were compared to the paired situation (Figure 5.3). Unfortunately, the activity from paired female transgenes differed between 2A and 6E for all constructs. This made it difficult to determine if the two-copy value differed when un-paired.

As an alternative approach, the established female activities of all *arm-lacZ*, *GMR-lacZ*, and modified *GMR-lacZ* lines were re-analysed. With all X-linked insertions, single copy males were measured simultaneously with the female measurements, providing an internal control.

As with male:female estimates of dosage compensation, normalising the female activities to single copy male levels controlled for inter-experiment variability. Thus the increase in female activity when gaining a second copy could be established for each line (Table 5.4). Activities from two-copies of *arm-lacZ* were in most cases just less than twice the single-copy level. Thus a male-specific increase of between 1.5- and 2.0-fold is sufficient to fully compensate. However, two-copy activities from all GMR cassettes were considerably more than twice the single-copy level. An analysis of variance was conducted on the $\log_2$(ratio) values for *arm-lacZ*, *GMR-lacZ*, and *DRE-GAGA-GL*. No significant difference was observed for the choice of *attP* location ($F = 1.01$, d.f. $= 2, 6$, $p = 0.418$), but the effect of construct choice was highly significant ($F = 15.46$, d.f. $= 2, 6$, $p = 0.004$). Assuming then that the transvection effect was equal at each *attP* location, Student's t-tests were conducted to determine if the pooled data at all sites were significantly different from a ratio of 2.0 ($\log_2$ of 1.0). The four estimates for *arm-lacZ* were not significantly different from the expected homozygous increase ($T = -1.99$, $p = 0.141$), confirming a lack of transvection effects for this construct. The three estimates for *GMR-lacZ* tended towards statistical significance ($T = 3.31$, $p = 0.080$). More measurements would be needed to confirm the transvection effect for *GMR-lacZ*. However, the four estimates of *DRE-GAGA-GL* were significantly greater than the expected increase ($T = 4.87$, $p = 0.017$), confirming a transvection effect with this construct.

Taken together, these observations suggested that expression from the GMR-based promoters was subject to positive transvection effects when homozygous. Although the two modified constructs increased male expression (as compared to single copy females), the increase was not sufficient to compensate against a far higher level of homozygous female expression; equating in most cases to only half the two-copy female levels.

Table 5.2: β-galactosidase activity in flies carrying constructs *DRE-GAGA-GL* or *GiL* at 6E *attP* site

			Heterozygous (1 dose)			Homozygous (2 doses)		
			Activity[a]		Ratio[b]	Activity		Ratio
Construct	Location	Stock	Male	Female	Male/Female	Male[c]	**Female**	Male/Female
DRE-GAGA-GL	6E	C98	0.46±0.02	0.33±0.01	1.38±0.05	0.73±0.01	**1.39**±0.05	0.52±0.02
GiL	6E	C109	0.61±0.01	0.48±0.02	1.28±0.09	0.60±0.02	**1.06**±0.04	0.56±0.01

[a]Mean OD/min/mg protein, ± 1 SE, n=3
[b]Mean male/female activity, ± 1 SE, n=3
[c]X-linked males hemizygous; only 1 dose of transgenic cassette

Table 5.3: β-galactosidase activity in flies with *lacZ* insertions at mixed sites

Cross			Male (1 dose)		Female (2 doses)		Ratio[b]
Father x Mother		Construct	Location	Activity[a]	Location	**Activity**	Male/Female
C76 x C78		GMR-lacZ	6E	1.61±0.03	2A/6E	**3.71**±0.11	0.44±0.01
C78 x C76			2A	2.19±0.10	2A/6E	**3.40**±0.10	0.65±0.03
C84 x C98		DRE-GAGA-GL	6E	0.46±0.04	2A/6E	**0.83**±0.02	0.56±0.06
C98 x C84			2A	0.56±0.03	2A/6E	**0.85**±0.03	0.66±0.03
C96 x C109		GiL	6E	0.96±0.03	2A/6E	**2.32**±0.07	0.42±0.02
C109 x C96			2A	2.50±0.00	2A/6E	**2.55**±0.10	0.98±0.04

[a]Mean OD/min/mg protein, ± 1 SE, n=3
[b]Mean male/female activity, ± 1 SE, n=3

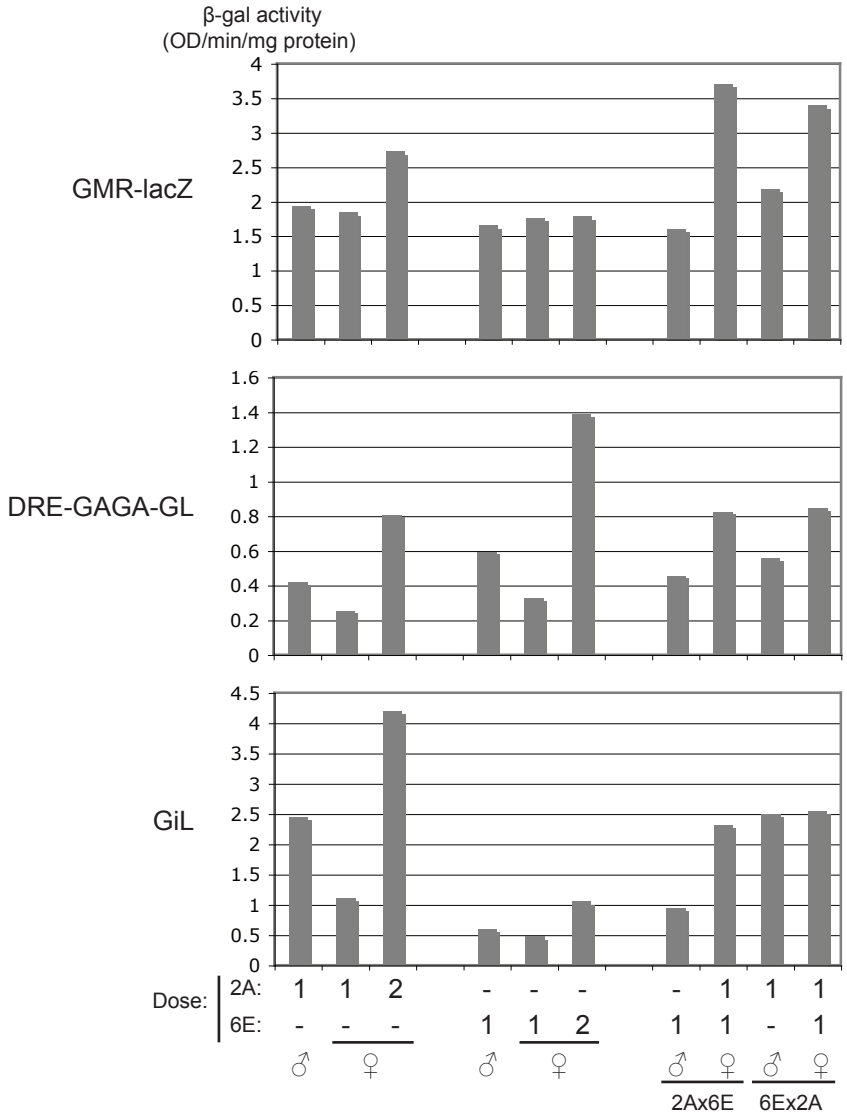


Figure 5.3: Effect of two separate transgene doses on female expression levels. Best estimates of β-galactosidase (β-gal) activity are plotted from measurements in heads of flies carrying one or two doses of the indicated constructs. In cases where several lines were measured, the estimates are averages of those lines. Otherwise means from individual lines were plotted. Flies with transgenic insertions at 2A or 6E were compared to flies resulting from crosses between 2A males and 6E females (giving 6E male progeny), or 6E males by 2A females (giving 2A male progeny).

Table 5.4: Two copy β-galactosidase activities differed from expected two-fold increases

Construct	Location	Line	Female Activity[a] One Dose	Female Activity[a] Two Doses	**Homozygous Increase**
arm-lacZ	2A	C66	0.55±0.04	1.05±0.04	**1.91**
		C66[b]	0.67±0.09	1.00±0.07	**1.49**
	6E	C67[b]	0.69±0.04	1.17±0.03	**1.70**
	20C	C68[b]	0.54±0.02	1.07±0.08	**1.98**
GMR-lacZ	2A	C76	0.79±0.04	1.97±0.06	**2.50**
	6E	C78	0.88±0.03	1.89±0.05	**2.14**
	20C	C81	0.81±0.02	1.90±0.09	**2.33**
DRE-GAGA-GL	2A	C84	0.61±0.01	1.86±0.11	**3.06**
		C85	0.52±0.02	2.16±0.17	**4.12**
		C86	0.66±0.05	2.00±0.20	**3.03**
	6E	C98	0.73±0.03	1.91±0.07	**2.63**
GiL	2A	C96	0.47±0.03	1.68±0.09	**3.58**
	6E	C109	0.79±0.05	1.79±0.03	**2.27**
GLSV40	2A	C91	0.72±0.02	1.65±0.09	**2.30**
GLSV4018D	2A	C95	0.72±0.01	2.04±0.08	**2.83**
CL12	2A	C83	0.81±0.01	2.13±0.11	**2.63**
CL15	2A	C89	0.87±0.01	2.00±0.20	**2.30**

[a]OD/min/mg protein, normalised to single dose male activity, ±1 SE, n=3
[b]Measured in hemisected adults; all others, adult heads

5.4 Increased GMR-lacZ male:female ratios observed specifically on the X chromosome

To confirm that the increased male:female ratios were specifically observed on the X chromosome, ϕC31-mediated transformations of *Drosophila* were conducted to target constructs *GMR-lacZ*, *DRE-GAGA-GL*, and *GiL* to an *attP* at 86F, on the 3rd chromosome (Section 2.2.5). The male:female ratio of β-galactosidase activity was measured for each construct at 86F (Section 2.9, Table 5.5, Figure 5.4). As this location is autosomal, both heterozygote and homozygote cases would be expected to give a ratio of 1.0 if equally expressed. Indeed, neither modification to *GMR-lacZ* exhibited an altered male:female ratio when expressed from the autosomal location.

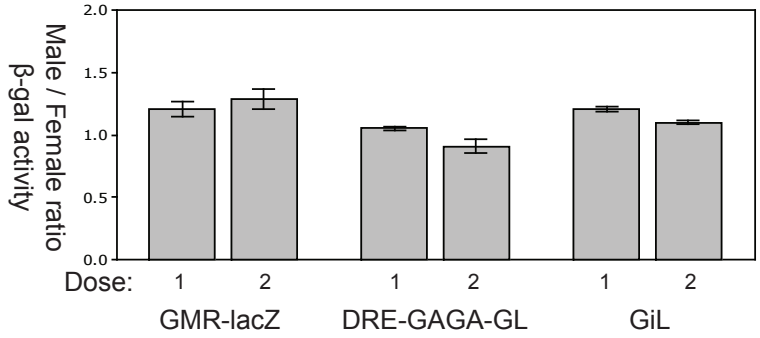


Figure 5.4: Modifications to GMR-*lacZ* did not alter the autosomal expression ratio. Male and female β-galactosidase (β-gal) activity was measured in adult fly heads for lines with each of the indicated constructs at insertion point 86F of chromosome 3. Ratios from single dose (heterozygote) and double dose (homozygote) are plotted as the mean of three experiments, ± 1 SE.

Table 5.5: Unaltered β-galactosidase activity in flies carrying modified *GMR-lacZ* constructs at 86F

		Heterozygous (1 dose)			Homozygous (2 doses)		
		Activity[a]		**Ratio**[b]	Activity		**Ratio**
Construct	Stock	Male	Female	**Male/Female**	Male	Female	**Male/Female**
GMR-lacZ	C112	2.46±0.13	2.05±0.03	**1.20**±0.06	6.28±0.36	4.89±0.05	**1.28**±0.08
DRE-GAGA-GL	C113	0.29±0.01	0.28±0.01	**1.05**±0.01	0.90±0.03	1.00±0.03	**0.91**±0.05
GiL	C111	0.53±0.00	0.44±0.01	**1.21**±0.02	2.41±0.08	2.19±0.10	**1.10**±0.01

[a]Mean OD/min/mg protein, ± 1 SE, n=3
[b]Mean male/female activity, ± 1 SE, n=3

5.5 DRE-GAGA-GMR-lacZ responded to altered MSL levels

To further test that the elevated male:female ratios of *GMR-lacZ* activitiy with DRE/GAGA sequences and intronic sequences were specifically due to acquisition of dosage compensation, MSL complex was induced in females. Expression of *msl2* from the constitutive *hsp83* promoter in females induces MSL complex formation and lowers viability (Section 1.1.3). Crosses were performed to observe the response of a single copy of *arm-lacZ*, *GMR-lacZ*, *DRE-GAGA-GL*, and *GiL* at 2A to *hsp83-msl2* (Section 2.2.4). The fold increase in female β-galactosidase activity with induction of *msl2* was established for each construct (Section 2.9, Table 5.6, Figure 5.5). As expected, *msl2* expression in females increased the β-galactosidase activity from *arm-lacZ*, but not from *GMR-lacZ*. The activity from *DRE-GAGA-GL* also increased, indicating a response to the dosage compensation machinery. A Student's t-test against the values for *GMR-lacZ* showed the increase to be statistically significant ($T = 6.46$, d.f. $= 2$, $p = 0.023$). Surprisingly, *hsp83-msl2* did not affect the female activity from *GiL* ($T = -2.85$, d.f. $= 2$, $p = 0.065$), implying that the increased male:female (1:1) ratio was not due to the action of MSL complex.

Table 5.6: Effect on female β-galactosidase activity with over-expressed *msl2*

lacZ source		Activity[a]		**Ratio**[b]
Construct	Stock	*lacZ/+; hsp83-msl2*	*lacZ/+; Sb*	***Sb⁺/Sb***
arm-lacZ	C66	0.28±0.02	0.22±0.00	**1.27**±0.10
GMR-lacZ	C76	2.52±0.14	2.69±0.08	**0.94**±0.04
DRE-GAGA-GL	C84	0.65±0.03	0.51±0.03	**1.29**±0.01
GiL	C96	1.02±0.04	1.29±0.07	**0.79**±0.03

[a]Mean OD/min/mg protein, ± 1 SE, n=3
[b]Mean Sb^{+}/Sb activity, ± 1 SE, n=3

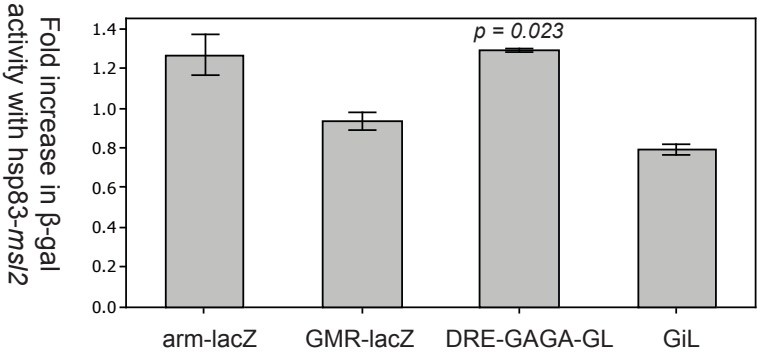


Figure 5.5: *DRE-GAGA-GL* responded to altered MSL levels. β-galactosidase (β-gal) activity was measured in adult female heads of lines carrying the indicated constructs, with and without hsp83-*msl-2*. The fold increase in females that express MSL2 relative to females that do not was plotted as a mean of three experiments, ± 1 SE. The p value is from a Student's t-test against *GMR-lacZ*

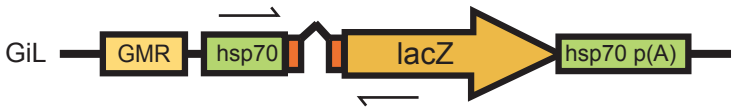
5.6 Inefficient processing of ninaE intron

The differing response of *GiL* from *GMR-lacZ* must have been due to the presence of the intron. To investigate the splicing status of the *ninaE* intron, cDNA was prepared from flies carrying *GiL* at 2A, and PCR conducted with primers flanking the intron (Section 2.4, Figure 5.6). In three different preparations of cDNA, the spliced product was much harder to detect than the un-spliced product. The intensities of the bands corresponding to RNA species were determined by software (Section 2.4.3, Appendix F). In the three preparations (Figure 5.6, lanes 2,4,6), the correctly-spliced product accounted for 14, 6, or 13 % of the PCR product, respectively. The abundance of the spliced product relative to the un-spliced product was 18, 49, or 16 %, respectively, with the large variation due to the abundance of the unknown smaller species in the preparation from the GiL-specific primer. These results indicated that a mixture of spliced and un-spliced products were present, and that the removal of the *ninaE* intron by splicing was inefficient.

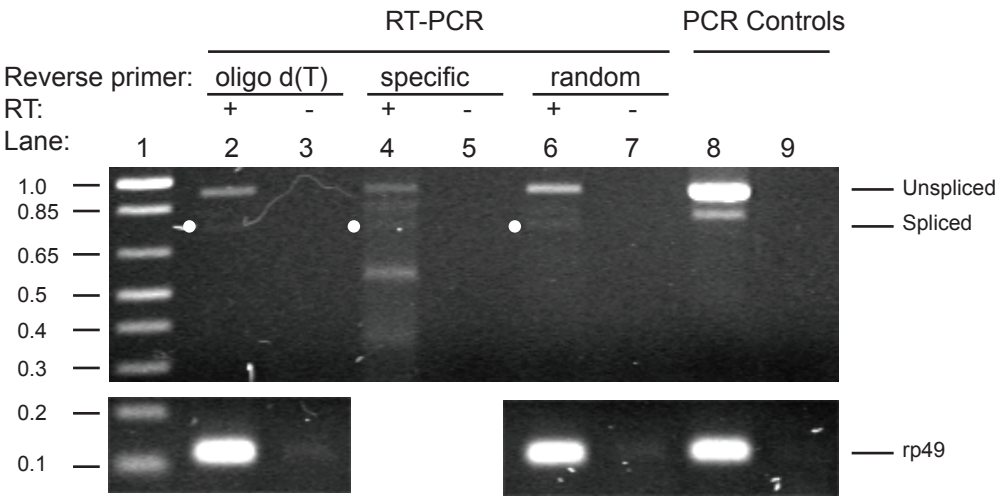
5.7 DRE and GAGA sites insufficient to alter GMR-hid expression pattern

To determine whether the DRE and GAGA motifs could likewise increase the acquisition of dosage compensation by a different gene, the clustered motifs were introduced into *GMR-hid*, and tested under identical conditions. The annealed 3DRE4GAGA oligonucleotides were inserted into the *Xho* I site of pBSw$^+$GMRattB, to create pCL24 (Section 2.3.11, Appendix I). Sequencing with primers armlacZseq and Corey_Seq_roX confirmed the insertion and orientation, but revealed two deleted nucleotides: shortening a spacer between two DRE motifs, and destroying an *Nhe* I site (not shown). A 3.9 kb *Eco* RI–*Eco* RI fragment of pBS SK hid cDNA containing the *hid* cDNA was inserted into the *Eco* RI site of pCL24 to create pCL25 (*DRE-GAGA-GMR-hid*) (Appendix I). The construction was confirmed by sequencing with the primers armlacZseq, hsp70TATAfwd, and h70p(A)rev (not shown). ϕC31-mediated transformations of *Drosophila* were conducted to target the construct to the *attP* at 2A, and were confirmed by PCR (Section 2.2.5, not shown).

A large-scale observation of eye sizes was conducted to determine the level of *DRE-GAGA-GMR-hid* expression in each sex, and compared to the level from *GMR-hid* (Section 2.8, Figure 5.7). As was the case with *GMR-lacZ*, the addition of DRE and GAGA motifs upstream of the promoter dropped the level of *GMR-hid* expression, giving a larger eye. However, the sequences had no noticeable effect on the slight sex-specific difference in *GMR-hid* expression.



(a) Schematic of *ninaE* intron 2 PCR amplicon



(b) PCR products

Figure 5.6: Inefficient splicing of *ninaE* intron. RNA was isolated from a line carrying the GiL construct, and reverse-transcribed with oligo d(T), GiL-specific, or p(dN)6 (random) reverse primers. Tandem reactions were conducted with (+) and without (-) reverse transcriptase (RT). PCR was conducted on all samples with primers spanning the *ninaE* intron (a) to test for splicing, and within the *rp49* gene to confirm efficient production of cDNA. Five microlitres of PCR products were electrophoresed on a 1.5 % agarose gel (b). Products from PCR over the intron are in the upper panel; within *rp49*, the lower panel. Reverse transcription with the GiL-specific primer would create no transcripts of *rp49*. Negative controls for PCR (lane 9) contained no template, and templates for positive controls (lane 8) were genomic DNA (upper panel) or a cDNA preparation prepared previously and known to amplify well (lower panel). Invitrogen 1Kb Plus ladder was loaded in lane 1; marker sizes in kilobases indicated along the left of the gel. The expected product sizes are 928 bp unspliced intron, 713 bp spliced intron, 125 bp rp49. Correctly-spliced products indicated with white dots.

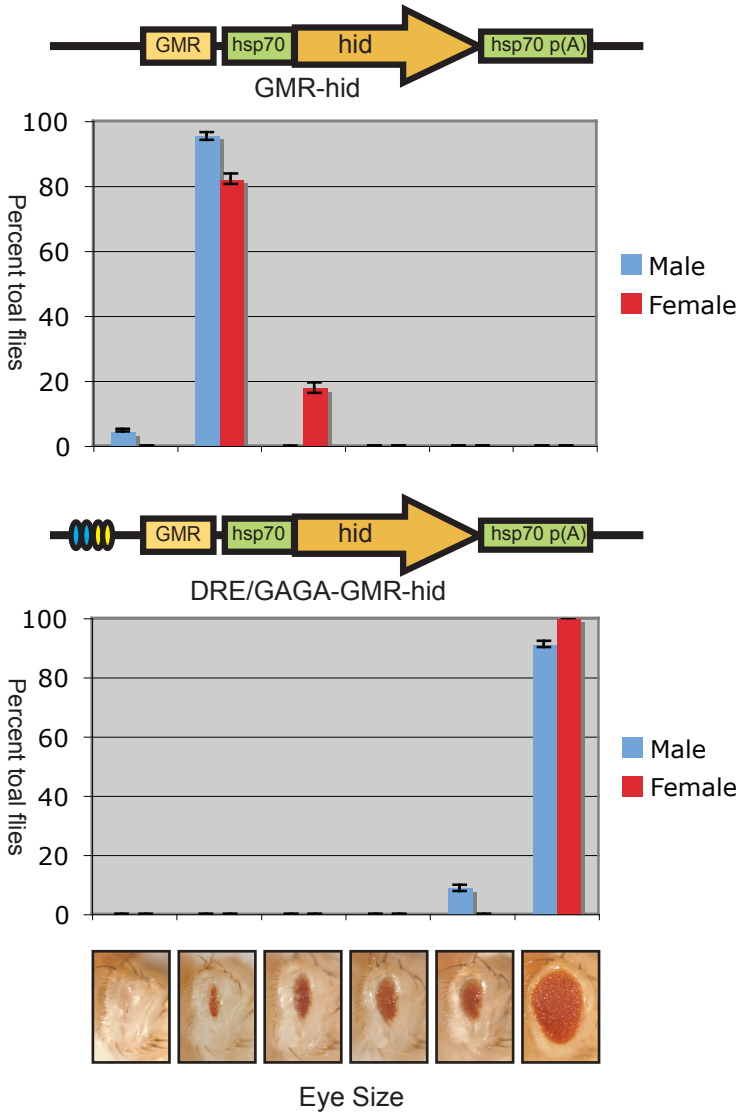


Figure 5.7: No enhancement of GMR-*hid* sex-specific difference with addition of DRE and GAGA sites. Male and female flies of each line (lines C70 and C107) were sorted into eye size categories, and the percentage in each category of total flies of that sex calculated. Means of three independent experiments plotted, with 95 % confidence interval.

5.8 Effect of DRE and GAGA elements on GMR-GFP difficult to detect

The time of expression may be important for dosage compensation. MSL complex binds the male embryonic X chromosome first at 2.5 hours after egg laying, and dosage compensation is at least present by late (12–14 hour) embryogenesis (Section 1.1.4). Genes commonly-expressed across cell types and developmental time ("housekeeping" genes) are enriched in bound MSL targets from ChIP on chip studies (Alekseyenko et al., 2006; Gilfillan et al., 2006). Perhaps GMR-mediated expression escaped dosage compensation due to its relatively late time of expression (Section 1.3.4). The DRE and GAGA motifs were designed to attract DREF and GAGA factor, both significant modifiers of transcription, and thus may have altered the basic transcription profile of the *GMR-hsp70* promoter construct. To determine if DRE and GAGA motifs also have an effect on the time or distribution of GMR-mediated expression, their effect on *GMR-GFP* was observed.

A 0.7 kb *Bam* HI–*Hin* cIII fragment of pBC-EGFP, containing the enhanced green fluorescent protein (EGFP) cDNA from Clontech, was inserted into the *Bgl* II and *Stu* I sites of pBSw^{+}GMRattB to create pGEG (*GMR-EGFP*) (Section 2.3.11, Appendix I). Sequencing with the primers hsp70TATAfwd and h70p(A)rev confirmed the construction (not shown). The annealed 3DRE4GAGA oligonucleotides were inserted into the *Xho* I site of pGEG to create pCL19 (*DRE/GAGA-GMR-GFP*). Sequencing from the armlacZseq primer confirmed the insertion and correct orientation (not shown). ϕC31-mediated transformations of *Drosophila* were conducted with pGEG and pCL19 to target the cassettes to the *attP* at 2A, and were confirmed by PCR (Section 2.2.5, not shown). The absence of the ϕC31 recombinase (also carrying the *3xP3-GFP* marker gene) was confirmed for all lines by PCR with the primers phic31For and phic31Rev (not shown).

The GFP expression was observed in lines carrying both constructs, in comparison to *y w* (lab stock #4, Table A.1) for a negative control, and *arm-GFP* (Bloomington #8556) or *pUB-GFP* (lab stock V62) lines for a positive GFP expression control (Section 2.11, Figure 5.8). Anatomical designations were made with reference to Hartenstein (1993). It was difficult to observe *GMR-GFP* expression above background levels of auto-fluorescence in embryos, with much non-specific and gut fluorescence observed. In some later (22 hr) embryos and early larvae of both test lines, spots of expression in the central nervous system were faintly observed (not shown). Although still faint, observations at later stages were more successful. Fluorescence was detected in the developing pupal eye and the adult ocelli for both *GMR-GFP* and *DRE/GAGA-GMR-GFP*. The pigment in the adult eye obscured the relatively weak fluorescence. As exposure and image manipulation were controlled, the images could be directly compared (by eye). All observations indicated that DRE and GAGA motifs lowered the ex-

pression level of *GMR-GFP*, as had been seen for both *GMR-lacZ* and *GMR-hid*. However, the difficulty of detection, especially in embryos, prevented any conclusions as to an earlier time of expression.

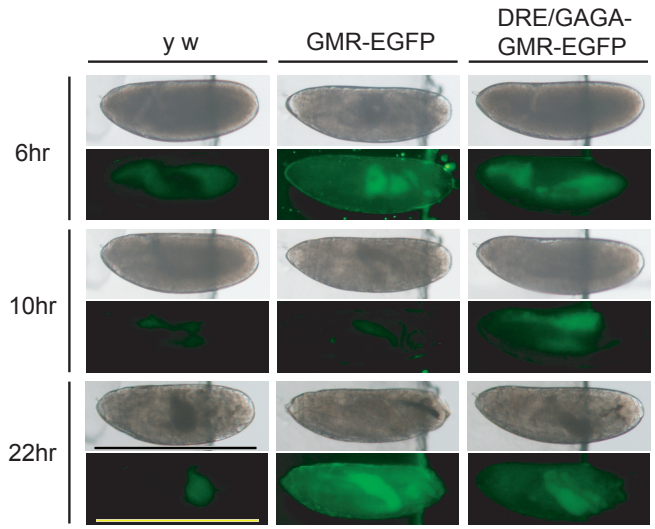
5.9 An earlier burst of expression insufficient for compensation of GMR-lacZ

In order to test the relative importance of early expression on dosage compensation of a gene, the *GMR-lacZ* cassette was placed under control of the tetracycline-regulated system (Section 1.3.2), and provided with additional bursts of activation at an earlier developmental point than that of GMR. If early expression was an important determinant in the attraction of MSL complex, the early expression should increase the dosage compensation of *GMR-lacZ* in adult heads. The genes *nullo* and *serendipity* α encode structural proteins of the early blastoderm (Ibnsouda et al., 1993; Postner and Wieschaus, 1994). Both genes are specifically expressed only at the blastoderm stage (Vincent et al., 1986; Rose and Wieschaus, 1992; Ibnsouda et al., 1995). Expression of *tTA* from promoters of both genes (dubbed *n1* and *s1*, respectively) is sufficient to kill flies expressing *tetO-hid*, but only in the absence of tetracycline (Horn and Wimmer, 2003). Lines expressing *n1-tTA* and *s1-tTA* were obtained from Ernst Wimmer [2].

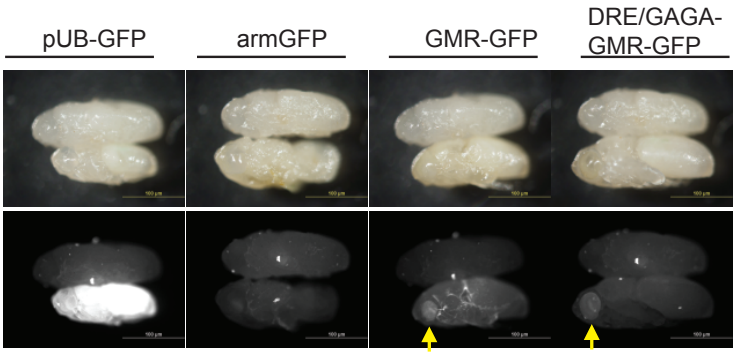
A 0.3 kb *Xho* I–*Sac* I fragment of pW.T.P-2 (Bello et al., 1998) containing the *tetO* operator was inserted in the *Xho* I site of pGLattB, with the aid of the "SacI_XhoI ..." linker oligonucleotides, to create pCL21 (*tetO-GMR-lacZ*) (Section 2.3.11, Appendix I). Sequencing with the primers armlacZseq and corey_seq_roX confirmed the insertion and orientation (not shown). ϕC31 transformations of *Drosophila* were conducted to target the construct to the *attP* at 2A, and were confirmed by PCR (Section 2.2.5, not shown). This was the only occasion where targeting to 2A gave false positives. Two of the four independent insertion events were ectopic. Genetic linkage analysis revealed one ectopic insertion was at another X chromosome location, and the other on the third chomosome, indeed the same chromatid as carried the ϕC31 recombinase (not shown). The X-linked ectopic insertion gave a band of the expected size following PCR within the cassette, but not across the cassette / landing site boundary (not shown). The other two insertion events were correctly targeted, and retained.

The male:female (1:1) ratio of β-galactosidase activity from *tetO-GMR-lacZ* at 2A with induction from *n1-tTA* or *s1-tTA* was measured in adult heads (Table 5.7), and compared to the known ratios of *arm-lacZ* and *GMR-lacZ* at 2A (Figure 5.9a). A separate cross of *tetO-GMR-lacZ* and *y w* showed that the *tetO* sequence had no effect on the acquired dosage compensation of *GMR-lacZ*. The system also allowed a separation of the temporal and spatial features of

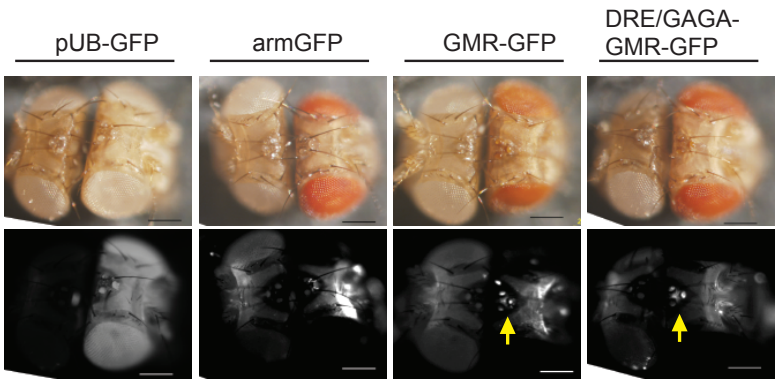
[2]Universität Bayreuth, Germany



(a) Embryos



(b) Pupal imagos



(c) Heads

Figure 5.8: DRE and GAGA sites were insufficient to alter GMR-GFP expression. Indicated lines photographed under white light (top of pair), or blue light for GFP expression (bottom, as grayscale in b,c). *pUB-GFP* and *arm-GFP* lines provided as positive controls for GFP expression. (a) embryos at different time points after egg laying, bar = 0.5 mm. (b) Pupal imagos of each line below a *y w* imago, arrow = developing eye, bar = 0.1 mm. (c) adult heads of each line (right) beside a *y w* head, arrow = ocelli, bar = 0.2 mm.

armadillo expression from the specific promoter elements, as *arm-tTA* should give a sustained early and general level of activation through the tetracycline system. Neither *n1-tTA* nor *s1-tTA* altered the male:female ratio. Likewise, *arm-tTA* did not affect the activity ratio, although the sustained additional expression increased the absolute β-galactosidase activity in heads of either sex (Figure 5.9b). Inhibition with tetracycline showed that the increase was specific to *arm-tTA*.

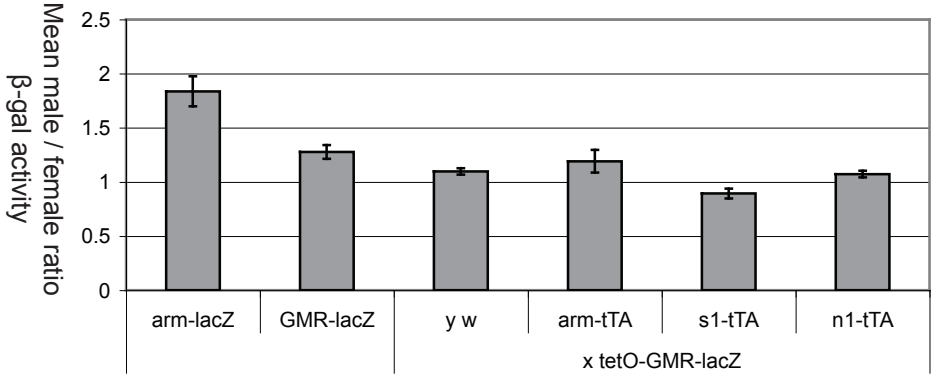
To confirm that the *n1-tTA* and *s1-tTA* lines (lab stock EW162 and EW140, respectively) expressed active *tTA* drivers, they were crossed to *tetO-YFP* (lab stock EW89), and the yellow fluorescence observed in early embryos, in comparison to that of the *arm-tTA* driver (Section 2.11, Figure 5.10). Both drivers produced fluorescence above background, and the *s1* promoter was particularly strong.

To confirm that *tetO-GMR-lacZ* was capable of expressing *lacZ* in a manner separate from GMR through the tetracycline system, the expression pattern in response to *arm-tTA* was examined in third-instar larval tissues with histochemical staining (Section 2.10, Figure 5.11). *tetO-GMR-lacZ* responded to *arm-tTA* in the brain and other tissues, but not in imaginal discs. Inhibition with tetracycline showed that the response was specific to the action of *arm-tTA*. Staining from *arm-lacZ* was more pronounced than from *arm-tTA* x *tetO-GMR-lacZ*, due to two copies of the transgene, but it was also more wide-spread. As the fragment of the *armadillo* promoter used in *arm-lacZ* and *arm-tTA* was the same (Scott et al., 2004), it seemed likely that the *GMR-hsp70* promoter was specifically repressed in all imaginal discs. If this is the case, then in the eye-antennal imaginal disc there is no repression posterior to the morphogenetic furrow, or in the Bolwig's nerve. Nevertheless, as *tetO-GMR-lacZ* could clearly respond to additional transcription induction from the tetracycline system, and as all *tTA* drivers were active in embryos, it was concluded that an earlier burst of expression was not sufficient to increase the dosage compensation of *GMR-lacZ* in adult heads.

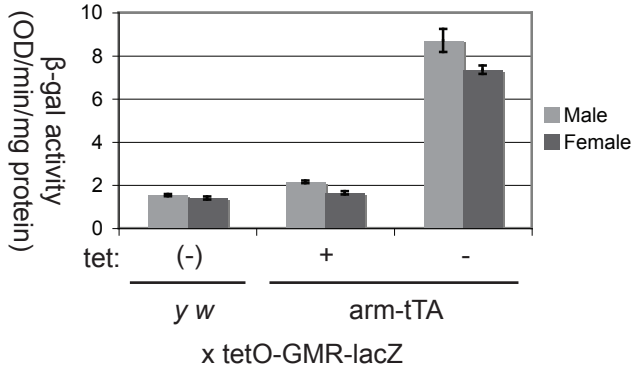
Table 5.7: β-galactosidase activity from a single copy of *tetO-GMR-lacZ*, with or without an additional burst of early expression

Cross			Activity[a]		**Ratio[b]**
lacZ	x Driver	Tetracycline	Male	Female	**Male/Female**
tetO-GMR-lacZ	x (*y w*)	-	1.53±0.05	1.40±0.08	**1.10**±0.03
	x arm-tTA	+	2.15±0.06	1.64±0.07	**1.31**±0.03
	x arm-tTA	-	8.70±0.53	7.35±0.20	**1.19**±0.10
	x s1-tTA	-	1.70±0.04	1.91±0.06	**0.89**±0.04
	x n1-tTA	-	1.74±0.03	1.62±0.03	**1.07**±0.03

[a]Mean OD/min/mg protein, ± 1 SE, n=3
[b]Mean male/female activity, ± 1 SE, n=3



(a) Male:female ratios of β-gal activity from *tetO-GMR-lacZ* with and without earlier expression burst



(b) β-gal expression of *tetO-GMR-lacZ* with and without repression by tetracycline

Figure 5.9: Unchanged compensation of *tetO-GMR-lacZ* with additional burst of early expression. (a) Male:female ratios of β-galactosidase (β-gal) activity, measured in heads of flies carrying one copy of *tetO-GMR-lacZ*, and one copy of the indicated tTA tetracycline driver. *y w* flies have no driver. Means of three independent experiments plotted, ± 1 SE. Ratios for one copy of *arm-lacZ* (line C66) and *GMR-lacZ* (C76) provided as positive and negative dosage compensation controls, respectively. All *lacZ* constructs were at 2A. Flies were raised without tetracycline. (b) β-galactosidase activity in heads of lines with indicated constructs, with (+) and without (-) repression by tetracycline (tet). Means of three experiments were plotted, ± 1 SE.

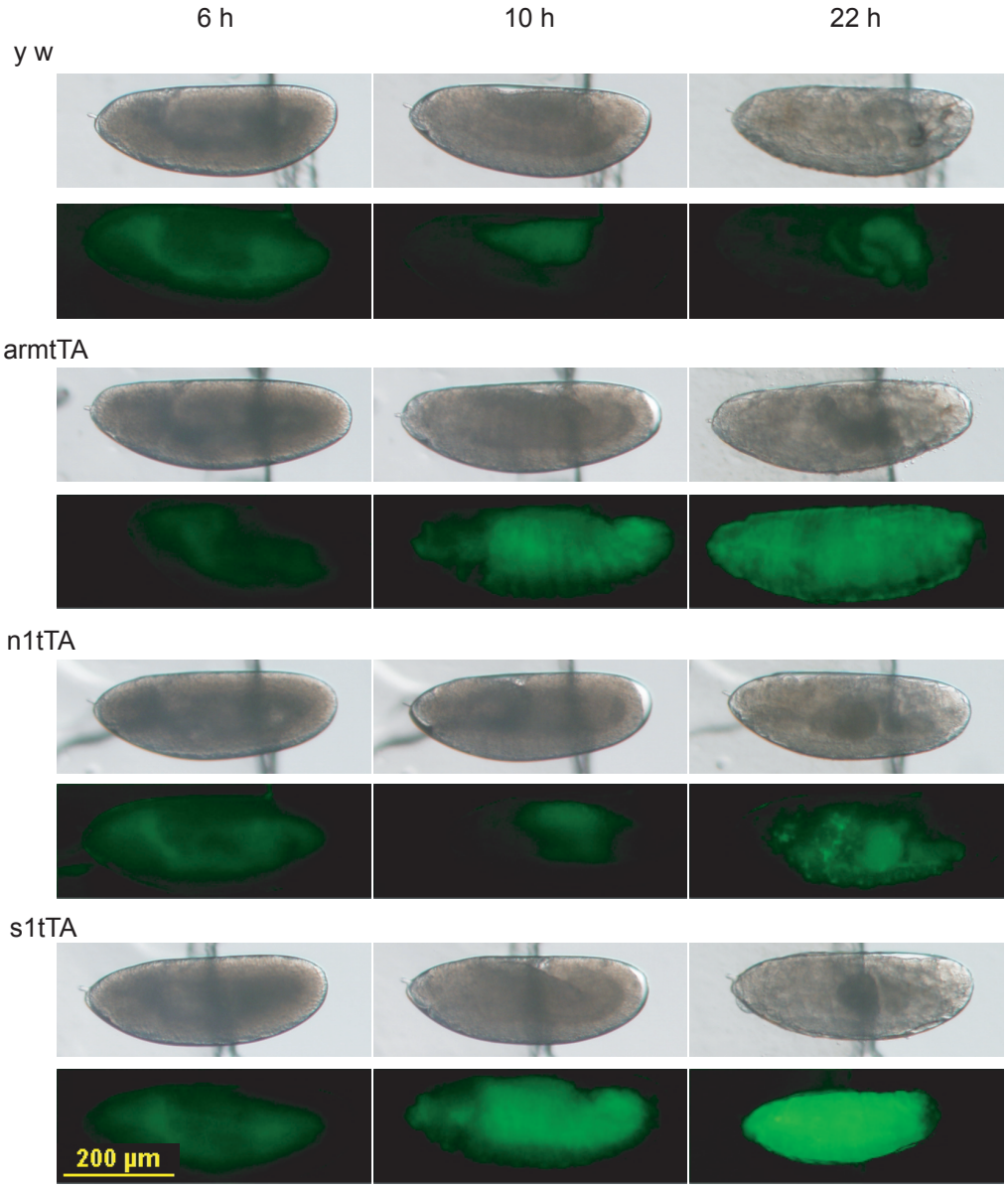


Figure 5.10: Activity measurement of tetracycline transactivator drivers. Males with each transactivator driver, or *y w*, were crossed to *tetO-YFP* females. White light (top of each pair) and fluorescent images (bottom) of the offspring at different time points after egg laying were recorded. YFP expression observed through a green filter set.

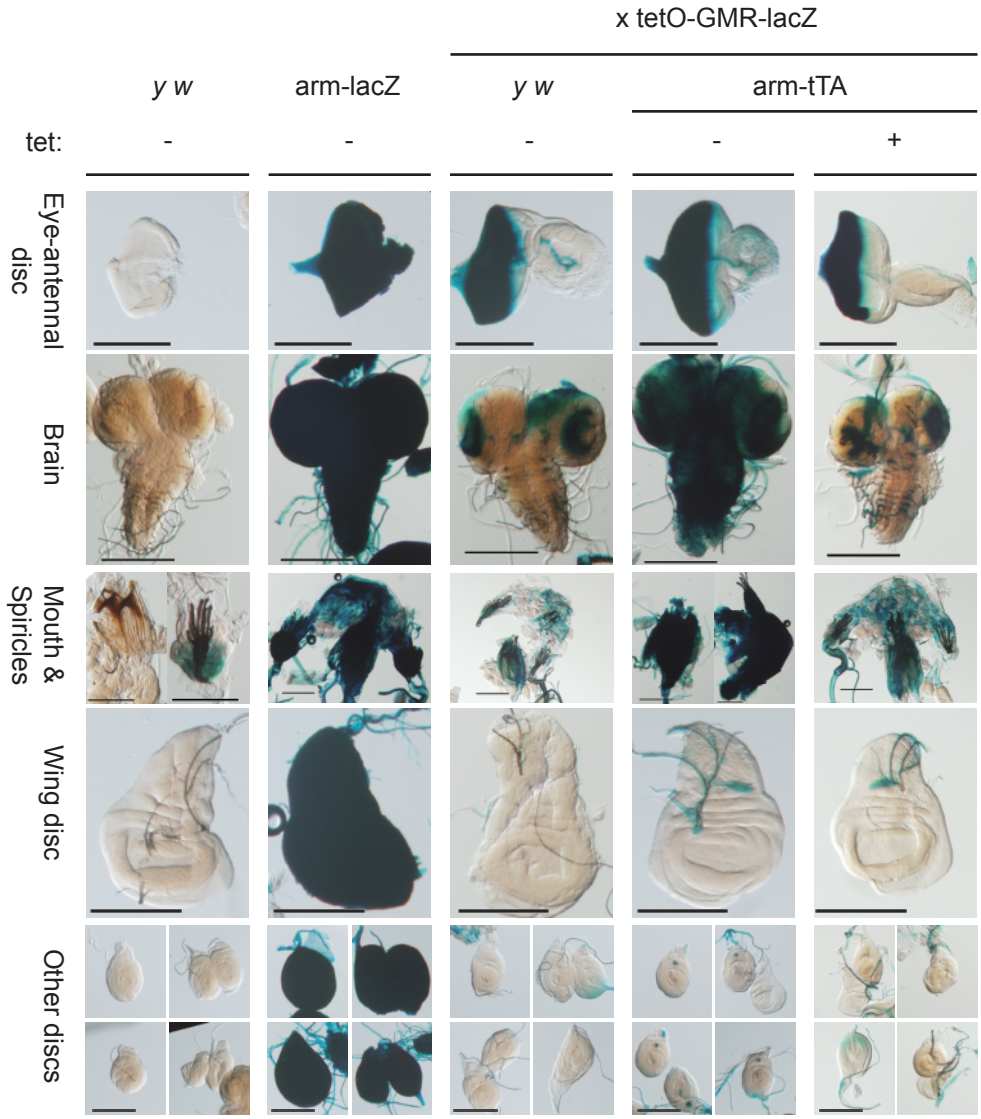


Figure 5.11: *tetO-GMR-lacZ* responded to tTA in brain and other tissues but not imaginal discs. Third instar larvae of the indicated strains, raised with (+) or without (-) tetracycline (tet) in the diet, were dissected and stained with X-gal for *lacZ* expression. Additional induction of *tetO-GMR-lacZ* occurs with *arm-tTA*, in the absence of tetracycline. *y w* and *arm-lacZ* provided as negative and positive controls for *lacZ* expression, respectively. All staining and photographic conditions were equal across all sets. Bar = 0.2 mm.

Chapter 6

General discussion

6.1 Importance of the promoter in dosage compensation of a X-linked transgene

The markedly different responses of *arm-lacZ* and *GMR-lacZ* transgenes when inserted on the X chromosome is a novel observation (Section 4.5). Transgenes often acquire a degree of compensation (usually less than two-fold) when inserted on the X chromosome (Scholnick et al., 1983; Spradling and Rubin, 1983; Krumm et al., 1985). Accordingly, *P*-element insertions of *arm-lacZ* are expressed equally in either sex if autosomally-located, but are compensated when inserted on the X chromosome (Fitzsimons et al., 1999). ϕC31-mediated insertions of *arm-lacZ* at three defined autosomal *attP* sites (2A, 6E, 20C) established that each location could support dosage compensation of a transgene. At all three sites where *arm-lacZ* acquired compensation, *GMR-lacZ* responded only weakly. At the same sites, the response of *GMR-hid* was similarly weak. A corresponding *arm-hid* construct would likely kill flies due to constitutive expression of the pro-apoptotic gene, and was not tested.

The differing response at X-linked locations was due to the promoter. By comparing each construct at defined and identical *attP* landing sites, position effect was eliminated. The ϕC31 system employed can only support insertions in one orientation, removing corresponding subtle differences on gene expression. Detailed comparisons and subsequent tests of regions in the two constructs could not identify differences other than the *GMR-hsp70* promoter that may have caused the differing response (Section 5.2).

This observation is difficult to correlate with reports claiming the act of transcription is sufficient to lead to dosage compensation of X-linked genes. A transgenic *UAS-hsp70TATA* promoter construct ectopically transcribes random genomic fragments at insertions sites when induced by GAL4, and recruits MSL complex to 7 of 10 such sites on the X chromosome (Sass et al., 2003). However, the inability to recruit MSL complex to all sites may indicate that sequence-specific or position-specific elements are also involved. MSL binding to autosomal insertions of the X-linked *mof* gene also occurs only when transcribed, and is maintained if the promoter is replaced by that from *tubulin*, or the GAL4 UAS (Kind and Akhtar, 2007). As recruitment occurs even with anti-sense transcription of *mof*, the authors conclude that "the type of promoter and direction of transcription are not [important]". However, Kind and Akhtar also note that "transcription alone is unlikely to be the sole targeting signal", as binding does not occur when *mof* is transcribed but lacks the 3′ terminus. As *mof* is encoded on the X chromosome and normally dosage compensated, it seems likely that the combination of transcription and the 3′ elements attract MSL complex to autosomal sites. By directly testing two promoters at identical sites, the observations of *arm-lacZ* and *GMR-lacZ* refute the claim that promoter type is unimportant.

If different promoters can affect the compensation status of a gene, then it seems likely

that endogenous genes on the X chromosome also escape compensation due to inappropriate promoters. The only genome-wide data on compensation levels is a measurement from microarrays of an average two-fold decrease in expression levels with knock-down of MSL components (Hamada et al., 2005). The authors noted that around 19 % of expressed genes changed less than 1.1 fold after knock-down, but were not confident of this estimate due to current technological constraints. I expect that similar experiments with RNA-seq will confirm that a significant minority of expressed X-linked genes will escape compensation. However, it is too early to speculate why these genes fail to be compensated. The observations with GMR indicate that inappropriate promoter type may be a common theme. It will have to be determined experimentally for each gene whether "inappropriate promoter" refers to an altered transcription or chromatin state that is not amenable to the compensation machinery, or merely to the absence of sequence elements that can help attract that machinery.

6.2 Effect of promoter elements on dosage compensation

Two modifications to the GMR promoter significantly increased the ability of the *GMR-lacZ* transgene to acquire dosage compensation at an X-linked location (Section 5.2). The modified constructs were compared at a single location capable of supporting dosage compensation of the *arm-lacZ* transgene, but not sufficient for up-regulation of *GMR-lacZ*. By cloning all modifications in the same manner as *GMR-lacZ* itself, base vector differences were eliminated. The differing 3′ sequences of *arm-lacZ* were not sufficient to mediate compensation. Thus the changes to the level of compensation must have been caused by changes to the promoter region. These observations re-inforce the assertion that promoter type can influence compensation, and provide further insights as to the mechanism of attraction to the X chromosome.

The presence of the second intron from *ninaE* dramatically increased the single dose male: female ratio of *GMR-lacZ* when expressed from 2A. Although the absolute expression levels appear to indicate a decreased female expression relative to *GMR-lacZ*, these measurements are difficult to compare between different assays. Furthermore, a large decrease in activity was also observed at the autosomal 86F location. Thus it seems most likely that the intron decreased *GMR-lacZ* transcription in all cases, with the response at 2A being a male-specific increase of this lowered value. The effect was specifically observed at the X chromosomal location, as would be expected for acquisition of endogenous dosage copensation. However, the ratio did not increase in females expressing *hsp83-msl2*, arguing against an effect of the MSL complex. A stretch of seven thymidines in the intron suggested possible female repression by Sxl, but the sequence does not otherwise resemble consensus Sxl binding sites (Samuels et al., 1994; Singh et al., 1995; Crowder et al., 1999), and would not explain the apparent general repression and male-specific activation. The presence of nearby sex-specific enhancer elements unrelated

to dosage compensation is also unlikely, as several independent insertions of *GMR-lacZ* and other constructs failed to exhibit sex-specific expression. Given that the female MSL complex induced by *hsp83-msl2* does not fully resemble the male situation (Section 1.1.3), perhaps the addition of the intron was sufficient to attract only male complex.

In what way might the intron attract more MSL complex? A software search did not identify any regions of the intron similar to the DRE, MRE or even GAGA consensus sequences (with MacVector 10.0, not shown). Perhaps the presence of the intron itself was enough to enhance dosage compensation. The addition of an intron to a previously mono-cistronic transcript should have lead to association with many RNA processing elements not normally involved in the transcription of *GMR-lacZ*. Components of both nuclear pores and regulatory exosomes co-immunoprecipitate with MSL proteins, and at least two pore proteins appear directly involved in dosage compensation (Mendjan et al., 2006). Transcription induction in yeast often leads to association with nuclear pores (Akhtar and Gasser, 2007), and in at least two cases the association is required for the correct level of activation (Brickner and Walter, 2004; Taddei et al., 2006). It is proposed that the MSL-nucleoporin association may similarly compartmentalise X-linked genes within the nucleus to further aid transcriptional up-regulation (Mendjan et al., 2006; Mendjan and Akhtar, 2007). Consistent with this, correct splicing, or perhaps splicing itself, appears necessary for sufficient levels of *roX2* transcripts (Park et al., 2005). As the MSL complex mediates *roX* expression and stabilises *roX* transcripts (Section 1.1.3), this may be further evidence that splicing can affect dosage compensation. However, the *ninaE* intron was poorly removed by splicing. If the intron does indeed lead to transcriptional up-regulation, then it may be that increased association with RNA processing machinery, rather than functional splicing, is sufficient for up-regulation.

This speculation needs to be verified. The direct results of the intron addition could be clarified. An increased association with nuclear pore, exosome, or other processing components could be detected with ChIP. Any change in sub-nuclear localisation might be visible with a combination of fluorescent in-situ hybridisation and confocal imaging. More generally, the impact of sub-nuclear localisation on transcription in *Drosophila* needs to be established. Likewise, the ability of the MSL complex to influence nuclear position remains theoretical. If dosage compensation were induced (as above), or perhaps targeted on demand to a marked site, any consequent change in nuclear position should be visible with real-time, live cell polytene imaging (Yao et al., 2006, 2008).

A modest increase in compensation of *GMR-lacZ* at 2A also resulted from the inclusion of multiple copies of transcription factor response elements found in the *armadillo* promoter. Although the presence of ATG tri-nucleotides complicated the use of a section of the *arm* promoter, a direct test of the DRE and GAGA elements in the $5'$ UTR of the transcript significantly increased the single dose male:female ratio. The increase was specific to the X chromosome,

and was also observed at 6E or in females carrying *hsp83-msl2*, indicating a likely effect of the MSL complex. The altered promoter reduced *lacZ* and *GFP* expression, but did not appear to re-distribute the transgene expression pattern spatially or temporally. The sequences alone were not sufficient to attract dosage compensation in all cases, as they had no effect on *GMR-hid*. Nonetheless, in a strict comparison with *GMR-lacZ* at 2A, the DRE and GAGA sequences clearly aided the acquisition of dosage compensation.

The observations and tests of DRE and GAGA sequences could be extended. The elements may be sufficient to increase compensation of other transgenes. A quantitative assay for activity of GFP in *Drosophila* eyes is available (Sarkar et al., 2006), and could be applied to the established lines with *GMR-GFP*. Detection of temporal or spatial changes in expression could be further refined. If fluorescent studies continued to be inconclusive, tissue staining of *lacZ* should suffice. The sufficiency of either GAGA or DRE elements could be tested, perhaps in shorter strings or in different locations relative to the transcription start point. As the response elements were designed to attract DREF and GAF, the presence of these and related transcription factors could be confirmed with mobility shift assays or PCR following ChIP. Genetic interactions of DREF or GAF with the MSL complex could also be investigated.

Measurements of dosage compensation are necessarily complex and often controversial. Cell culture systems and females with induced MSL complex may not accurately represent the male situation. Ectopic up-regulation at autosomal transgenes takes no account of features specific to the X-chromosome environment. The system presented here permits the study of normal male dosage compensation in an endogenous X chromosome environment. Further weak or collabarative contributions to compensation could be detected with the same system. For example, the GMR constructs contain the TATA and tsp of the autosomal *hsp70* gene. The effect of a basal promoter from an endogenous X-linked (and dosage compensated) gene could be tested. Constructs could be designed to establish the effect of other local X-linked sequences, such as the *white* transgenic marker gene. Cassette exchange targeting systems would eliminate such extraneous sequences. In a complementary approach, deletion analysis of compensation-capable promoters (*e.g. arm*) or genes could be undertaken at the same site. The relative effects of important sequences in the 3′ coding sequences as compared to surrounding positions could be established. Different introns, intron position and length, alternatively-spliced transcripts, and strength of splice sites could all be analysed. Finally, as the attraction of MSL complex is a separable event from dosage compensation (Section 1.1.5), the presence, necessity, and contribution of the complex could be analysed following ChIP of MSL components or the MOF-mediated acetylation of H4K16.

6.3 Early expression insufficient for dosage compensation of a tissue-specific transgene

As discussed, dosage compensation in embryos may differ from that in later stages (Section 1.1.4). Early complex may differ in compostion of *roX* RNAs, and may preferentially bind genes that will remain constitutively expressed throughout development. Two recent studies have independently proposed that MSL binding is mainly established early, then maintained throughout development (Legube et al., 2006; Kotlikova et al., 2006). The finding that early expression of *GMR-lacZ* is not sufficient for compensation in adult heads does not lend support to this theory (Section 5.9).

It was established that *GMR-lacZ* at 2A did not acquire dosage compensation, but could be modified to do so. This cassette was altered to allow additional regulation through the tetracycline system, and was shown to respond to this control in tissues where GMR alone cannot drive expression. Early induction was provided by tTA activators driven by promoters of genes known to be expressed in the early embryo. It was demonstrated that tTAs expressed from these early promoters could activate tetO-mediated expression in embryos. However, the additional embryonic expression of *GMR-lacZ* provided by such activators did not increase the level of acquired dosage compensation measured in adult heads. This was observed both for a short burst of expression around the time of MSL complex assembly, and for an expression level sustained throughout development.

Early gene expression thus does not appear to be a powerful determinant for attraction of MSL complex. Undoubtedly the early expression profile will affect the deposition of MSL complex as it first assembles. Bound complex is clearly quite stably associated with the X chromosome, and is not recycled within 20 minutes, at least at the resolution offered by FRAP analysis of cultured cells (Straub et al., 2005*b*). However, the complex must also be capable of re-distribution as genes become active at later stages, and these changes though slight at chromosome resolution correlate with transcriptional changes when analysed by ChIP on chip (Sass et al., 2003; Kotlikova et al., 2006; Alekseyenko et al., 2006). The relevant issue might instead be viewed as one of timing. Inhibition of transcription in cultured cells with 20 hours of α-amanitin treatment reduces bound MSL complex on X-linked genes, demonstrating that MSL complex can disassociate from specific genes as transcription ceases (Kind and Akhtar, 2007). Indeed, Kind and Akhtar propose that "stable binding throughout development could be the rule with some differentially transcribed genes being the exception". The data presented here support this hypothesis, by demonstrating that a tissue-specific gene can avoid compensation, even if expressed from an early stage. Early expression itself may not be sufficient to determine the compensation status of a gene.

An unexpected and complicating observation was the lack of *tetO-GMR-lacZ* response to

arm-tTA induction in larval imaginal discs. As *arm-lacZ* was clearly expressed in these tissues, and no reports are known of tissue-specific repression of tetR binding or VP16 activation, it seems likely that the *GMR-hsp70* promoter was specifically repressed, and that the *arm-tTA* activator was not able to overcome this repression. A similar repression pattern occurs with another Glass reporter ('38-1'), a pentamer of larger (38 bp) fragments spanning the Glass binding site (Ellis et al., 1993). 38-1 cannot respond to constitutively-expressed *hsp70-glass* in imaginal tissue, except after the morphogenetic furrow in the eye-antennal disc, where the repressor is presumably absent. As GMR (five copies of a shortened 29 bp glass binding site) is not susceptible to this repression, it was concluded that the binding site for the putative repressor was removed when shortening the fragment (Ellis et al., 1993). The observation that *tetO-GMR-lacZ* was susceptible to the same repression characteristics now suggests that the repressor binding site was not removed, merely weakened to allow dominance of glass over the repressor.

It seems unlikely that this putative repression would have invalidated the test of expression time on dosage compensation. MSL complex predominantly binds 3′ coding sequences, and would likely be unperturbed by the transient presence of the repressor in the promoter. Furthermore, β-galactosidase activity in adult heads would also have been provided by expression in the brain, where no repressive action was evident. A related experiment could be conducted without GMR (*tetO-TATA-lacZ*), to test the effect of expression time on dosage compensation. However, this approach could only address expression time, not the relative importance of time on a tissue-specific gene.

The extent and duration of the putative repression remain to be determined. Given that *arm-tTA* appeared sufficient to mediate *tetO-GMR-lacZ* expression when repression was absent, the nature of the repressor could be established by analysing embryos carrying both these transgenes. A detailed study of embryonic and larval stages with the above histochemical staining *lacZ* assay may thus be informative.

6.4 On the combinatorial nature of Drosophila dosage compensation

Drosophila dosage compensation is a complex biological system. Some of the observations presented here are good evidence of this complexity. The differing levels of dosage compensation acquired by *arm-lacZ*, *GMR-lacZ*, and related constructs were most likely due to differences in the affinity of each for the MSL complex. As discussed, attraction of the complex to the X chromosome is a complicated two-step procedure involving cis-encoded factors, chromatin modifications and density, and degenerate sequence motifs (Section 1.1.4). By measuring dosage compensation in an endogenous X chromosome environment, all these factors will affect the

local level of MSL complex and subsequent transcriptional up-regulation of transgenes. Thus the degree of up-regulation acquired can be viewed as the result of a combination of different attraction factors.

The limited response of *GiL* may be due to a unique combination of attraction factors. The male-specific up-regulation of *GiL* was far greater at 2A than at 6E. Certainly the concentration of complex differs across the X chromosome, as visible with polytene preparations or in MSL immuno-precipitates. It may be that the local concentration of endogenous MSL complex also differs slightly between these two *attP* locations. Perhaps then the response of *GiL* at 2A was a combined result of local complex, and the ability of the transgene itself to recruit complex. This location-specific factor should be carefully considered in the choice of locations at which to study dosage compensation.

At any particular location, different elements of a gene unit must also combine to attract MSL complex. The effect of the promoter has been demonstrated, but the transcribed portion of a gene can also affect attraction of MSL complex. A transgenic promoter designed to transcribe genomic sequences adjacent to its insertion site can recruit MSL complex to the transcribed sequences, but only when the promoter is induced through the GAL4 system (Sass et al., 2003). This attraction occurs at most X-linked insertions, but rarely when inserted on the autosome. These observations imply that transcription itself, when on the X chromosome, is sufficient to attract local MSL complex to the transcribed region. Alternatively, sequence-specific features of the transcribed sequence may be important for complex recruitment. Consistent with this, *tublin-mof* transgenes recruit MSL complex to autosomal locations, but *tubulin-pho* transgenes do not (Kind and Akhtar, 2007). Transcription indeed facilitates binding, but a sequence-specific element of the transcript must also be important.

The apparent promoter/gene contradiction is mostly resolved if the MSL-attracting sequences function as a combination of degenerate sequence motifs. Sites of high affinity for the MSL complex are usually co-incident with an MRE recognition element in non-coding DNA (Section 1.1.4), highly similar to the GAGA motifs employed in this study. However, transcription of a *mof* transgene only recruited MSL complex if a $3'$ portion of the transcript was present, demonstrating that sequence motifs can also reside within a transcribed unit (Kind and Akhtar, 2007). Consistent with an additive attraction model is the increased attraction observed when known MSL attraction motifs are multimerised. Direct repeats of the DHSs from *roX1* or the 18D10 site, or the MRE-containing portions of other high-affinity sites, attract MSL complex to autosomal locations more robustly than single copies (Kageyama et al., 2001; Oh et al., 2004; Gilfillan et al., 2007). It seems likely that the multimerisation of DRE and GAGA elements contributed to their ability to attract MSL complex.

The response of any one reporter gene at any one site is thus likely to depend upon a combination of degenerate motifs in the gene or promoter, nucleosome occupancy or chromatin den-

sity such as the motifs might be exposed, and location relative to local concentrations of bound MSL complex. The splicing status of the transcript may also contribute to MSL-mediated mechanisms of up-regulation. Any experiment aimed to test the ability of any one factor on MSL recruitment must therefore be carefully designed to control or measure all others.

6.5 Dosage compensation and transvection

Measurements of dosage compensation were also complicated by transvection effects (Section 5.3). An apparent difference in two-copy β-galactosidase activity from paired transgenes, as compared to non-paired loci, indicated that pairing-dependent transcription enhancement may occur. However, as expression from 6E appeared to differ from that at 2A, the effect was difficult to separate from a position effect. By normalising female activities to male activities in data from previous dosage compensation experiments, the position variable was eliminated, but the transvection effect remained. Analysis was only possible on the X chromosome, as both heterozygote and homozygote female measurements could be normalised to the same one copy male levels. Autosomal homozygous readings were made with two copies of male transgenes. Nevertheless, several observations of *arm-lacZ*, *GMR-lacZ*, and *DRE-GAGA-GL* transgene activities on the X chromosome were available. A positive transvection effect was repeatedly observed for *GMR-lacZ*, and *DRE-GAGA-GL*. This is the first observation of transvection effects on the *GMR-hsp70* promoter.

The molecular nature of transvection at GMR remains unknown. It seems likely that Glass-mediated activation is enhanced by the close association of the paired homologues. Alternatively, it may be that the affinity of the enhancer for Glass, or a co-factor, is enhanced. A third possibility is a change to the interactions between Glass and putative repressive factors. Glass may overcome the repression easier if two copies of GMR are paired. Characterisation of these interactions would be of interest to the study of transvection. But the effect at GMR may be of interest more generally. Most observations of transvection involve differences in gene expression from the pairing of two different alleles (Section 1.2). However, as single and double copies of the GMR constructs were directly compared, the observations are not just indicative of allelic differences, but of a more general phenomenon.

A more general transvection effect would have significant implications on dosage compensation. It seems likely that shared enhancers or repressors would contribute to the normal level of gene expression from many endogenous genes, including those on the paired female X chromosomes. The equalisation of transcription between the sexes would thus depend on a combination of effects. Male hyper-transcription would in some cases need to be more or less than two-fold to equal the paired female level. Indeed, reduction of MSL2 by RNAi dropped X-linked transcripts by different amounts, averaging 0.6 to 0.8 of the wild-type expression level

(Hamada et al., 2005). Similarly, different X-linked transcripts in males decrease differing amounts if MSL proteins are lacking, from a third of wild-type levels to no significant decrease (Chiang and Kurnit, 2003). And in analogous experiments to those reported here, male hyper-transcription of a *white* transgene can vary between 1.5- and 10-fold, depending on promoter strength (Qian and Pirrotta, 1995). The concept of interplay between transvection and the MSL complex is not far removed from that of action through relief of repression (Section 1.1.5). It is already widely accepted that MSL up-regulation may counter repressive actions to achieve the apparent two-fold difference. Similarly, the individual repressive or stimulatory transvection effects on female genes may affect the level to which male transcription needs to be equated.

If transvection can indeed affect dosage compensation, care must be taken to measure the relevant biological process. If transvection effects are likely present, as with *GMR-lacZ*, then they must be removed if the process to be studied is male hyper-transcription. That is why the conclusions above are drawn only from the heterozygous (1:1) observations; homozygous measurements were confounded by transvection enhancement of the two-copy female transgene. The alterations to the *GMR-hsp70* promoter increased the male-specific up-regulation, but were not sufficient to compensate for the dose difference. Subsequent measurements of dosage compensation should also carefully distinguish between these two effects.

Chapter 7

Conclusions

Many aspects of *Drosophila* dosage compensation remain unknown. The many factors that co-operate or antagonise MSL binding or up-regulation highlight the complexity of the system. It seems likely that additional factors that influence the process remain undiscovered. A genetic system to detect these factors has advantages over biochemical approaches, particularly in the detection of subtle interactions. In light of the attempts with *GMR-hid*, such a system must accurately and consistently report small changes in the level of male hyper-transcription. Recombinase-based targeting to locate the reporter at an X-linked location that supports good levels of dosage compensation would be a clear advantage.

The details of targeting dosage compensation to the X chromosome are becoming more clear. If the two-step model of MSL attraction continues to be supported, and the nature of the sequence-specific determinants can be determined, then the general picture will be apparent. However, some issues are still poorly understood, such as the involvement of *roX* RNAs, and their dual function as MSL complex components and chromosome markers. An inducible dosage compensation system would provide unique insights concerning the establishment of X chromosome binding. The observations with *tetO-msl2* indicate that such a system must be tightly controlled. The advent of ChIP on chip and ChIP-seq technologies has largely eclipsed cytogenetic observations of polytene chromosomes. If combined with an inducible system, the resolution offered by these new technologies could provide significant observations on the dynamics of MSL association.

Concerning the establishment of dosage compensation on the male X chromosome, I sought to provide functional support for the proposed importance of early expression. The compensation-deficient *GMR-lacZ* transgene was provided with additional early transcription through the tetracycline system. A burst of embryonic expression was insufficient to establish compensation of the later tissue-specific transcription. Similarly, constitutive expression of the transcript did not facilitate male-specific up-regulation of the transgene. These observations do not exclude the possibility that early or constitutive expression may affect dosage compensation. Instead, these factors appear relatively less important than the sequence-specific factors of the gene unit.

A surprising observation was the apparent repression on the GMR reporter. The *tetO-GMR-lacZ* transgene did not respond to constitutively-expressed tTA in most larval imaginal tissue, excluding the differentiating eye-antennal disc. This repression pattern appeared identical to that of an earlier reporter that used larger Glass binding sites. It seems that shortening the Glass binding site did not exclude the putative repressor from binding, but merely reduce the affinity of the site for the repressor. Thus at the GMR, Glass can overcome the repression, but when Glass is absent, the repressor may affect the interaction of upstream enhancers, such as the activation through the tetracycline system. The extent of the repression, and indeed existence of the repressor, remain to be determined.

I have presented several observations on the targeting of dosage compensation to the X chromosome. Male expression from the *GMR-hsp70* enhancer-promoter was not doubled in repeated attempts to attract MSL complex to the reporter. Neither did the cassette acquire compensation from the local machinery when inserted on the X chromosome. A comparison of *GMR-lacZ* and *arm-lacZ* at identical X-linked sites revealed that the regulatory sequences can be important for determining the compensation status of the transcript. A *GMR-lacZ* cassette with additional upstream DRE and GAGA motifs acquired more male hyper-transcription at the X-linked locations, strengthening the argument that promoter type can affect compensation status. Yet, other reports highlight the importance of transcribed sequences in attraction of dosage compensation, particularly $3'$ coding sequences. The importance of the promoter need not conflict with these reports, if the attraction is presumed to act through a combinatorial mechanism. Accordingly, I propose that the different cassettes recruited MSL complex to different degrees, which varied depending on the strength of various degenerate sequences. However, presence of MSL complex was not determined. Rather, the experiments revealed differences in male-specific up-regulation. In this manner, all mechanisms that co-operate to hyper-transcribe the male X chromosome should have been equally reported. It is thus also possible that MSL recruitment was affected indirectly through associated factors.

I have argued that transvection can affect dosage compensation. When modified *GMR-lacZ* cassettes attracted more male hyper-transcription, the corresponding increase in dosage compensation was less pronounced. This apparent paradox is resolved if female levels increased by more than two-fold when paired. Retrospective analyses of all X-linked insertions of *GMR-lacZ* found this to be the case. If transvection can more generally affect paired female transcription levels, then dosage compensation will at least in some cases be attained with male hyper-transcription of more or less than two-fold. It is my contention then that at the individual gene level, we need study male hyper-transcription only, not equalisation between the sexes. The final compensation for dose deficiency is merely the product of several different biological mechanisms, chief among them MSL-mediated up-regulation.

Bibliography

Agha-Mohammadi, S., M. O'Malley, A. Etemad, Z. Wang, X. Xiao, and M. T. Lotze. 2004. Second-generation tetracycline-regulatable promoter: repositioned tet operator elements optimize transactivator synergy while shorter minimal promoter offers tight basal leakiness. J Gene Med **6**:817–828.

Akhtar, A., and P. B. Becker. 2000. Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in Drosophila. Molecular Cell **5**:367–75.

Akhtar, A., and P. B. Becker. 2001. The histone H4 acetyltransferase MOF uses a C2HC zinc finger for substrate recognition. EMBO Rep **2**:113–118.

Akhtar, A., and S. M. Gasser. 2007. The nuclear envelope and transcriptional control. Nat Rev Genet **8**:507–17.

Akhtar, A., D. Zink, and P. B. Becker. 2000. Chromodomains are protein-RNA interaction modules. Nature **407**:405–9.

Alekseyenko, A. A., E. Larschan, W. R. Lai, P. J. Park, and M. I. Kuroda. 2006. High-resolution ChIP-chip analysis reveals that the Drosophila MSL complex selectively identifies active genes on the male X chromosome. Genes Dev **20**:848–57.

Alekseyenko, A. A., S. Peng, E. Larschan, A. A. Gorchakov, O.-K. Lee, P. Kharchenko, S. D. McGrath, C. I. Wang, E. R. Mardis, P. J. Park, and M. I. Kuroda. 2008. A sequence motif within chromatin entry sites directs MSL establishment on the Drosophila X chromosome. Cell **134**:599–609.

Amrein, H., and R. Axel. 1997. Genes expressed in neurons of adult male Drosophila. Cell **88**:459–69.

Aronson, J., G. Rudkin, and J. Schultz. 1954. A comparison of giant X chromosomes in male and female Drosophila-melanogaster by cytophotometry in the ultraviolet. Journal of Histochemistry & Cytochemistry **2**:458–459.

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

Ashburner, M., K. G. Golic, and S. Hawley, 2005. Drosophila: A laboratory handbook, Chapter 29 . Cold Spring Harbor Laboratory Press, 2nd edition.

Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl, editors. 1994. Current Protocols in Molecular Biology. John Wiley and Sons.

Badenhorst, P., M. Voas, I. Rebay, and C. Wu. 2002. Biological functions of the ISWI chromatin remodeling complex NURF. Genes and Development **16**:3186–3198.

Bai, X., A. A. Alekseyenko, and M. I. Kuroda. 2004. Sequence-specific targeting of MSL complex regulates transcription of the roX RNA genes. Embo J **23**:2853–61.

Bai, X., E. Larschan, S. Y. Kwon, P. Badenhorst, and M. I. Kuroda. 2007. Regional control of chromatin organization by noncoding roX RNAs and the NURF remodeling complex in Drosophila melanogaster. Genetics **176**:1491–9.

Barski, A., S. Cuddapah, K. Cui, T.-Y. Roh, D. E. Schones, Z. Wang, G. Wei, I. Chepelev, and K. Zhao. 2007. High-resolution profiling of histone methylations in the human genome. Cell **129**:823–37.

Bashaw, G. J., and B. S. Baker. 1995. The msl-2 dosage compensation gene of Drosophila encodes a putative DNA-binding protein whose expression is sex specifically regulated by Sex-lethal. Development **121**:3245–58.

Bashaw, G. J., and B. S. Baker. 1997. The regulation of the Drosophila msl-2 gene reveals a function for Sex-lethal in translational control. Cell **89**:789–98.

Bateman, J. R., A. M. Lee, and C. T. Wu. 2006. Site-specific transformation of Drosophila via phi C31 integrase-mediated cassette exchange. Genetics **173**:769–777.

Bell, L. R., J. I. Horabin, P. Schedl, and T. W. Cline. 1991. Positive autoregulation of sex-lethal by alternative splicing maintains the female determined state in Drosophila. Cell **65**:229–239.

Bell, O., T. Conrad, J. Kind, C. Wirbelauer, A. Akhtar, and D. Schübeler. 2008. Transcription-coupled methylation of histone H3 at lysine 36 regulates dosage compensation by enhancing recruitment of the MSL complex in Drosophila melanogaster. Mol Cell Biol **28**:3401–9.

Belli, G., E. Gari, L. Piedrafita, M. Aldea, and E. Herrero. 1998. An activator/repressor dual system allows tight tetracycline-regulated gene expression in budding yeast. Nucleic Acids Res **26**:942–947.

Bello, B., D. Resendez-Perez, and W. J. Gehring. 1998. Spatial and temporal targeting of gene expression in Drosophila by means of a tetracycline-dependent transactivator system. Development **125**:2193–2202.

Belote, J. M., and J. C. Lucchesi. 1980*a*. Control of X chromosome transcription by the maleless gene in Drosophila. Nature **285**:573–5.

Belote, J. M., and J. C. Lucchesi. 1980*b*. Male-specific lethal mutations of Drosophila melanogaster. Genetics **96**:165–86.

Bergmann, A., J. Agapite, K. McCall, and H. Steller. 1998. The Drosophila gene hid is a direct molecular target of Ras-dependent survival signaling. Cell **95**:331–341.

Bickel, S., and V. Pirrotta. 1990. Self-association of the Drosophila zeste protein is responsible for transvection effects. EMBO J **9**:2959–67.

Bieschke, E. T., J. C. Wheeler, and J. Tower. 1998. Doxycycline-induced transgene expression during Drosophila development and aging. Mol Gen Genet **258**:571–579.

Birchler, J. A., U. Bhadra, M. P. Bhadra, and D. L. Auger. 2001. Dosage-dependent gene regulation in multicellular eukaryotes: implications for dosage compensation, aneuploid syndromes, and quantitative traits. Developmental Biology **234**:275–88.

Birchler, J. A., R. K. Owenby, and K. B. Jacobson. 1982. Dosage compensation of serine-4 transfer RNA in Drosophila melanogaster. Genetics **102**:525–37.

Birchler, J. A., M. Pal-Bhadra, and U. Bhadra. 2003. Dosage dependent gene regulation and the compensation of the X chromosome in Drosophila males. Genetica **117**:179–90.

Bischof, J., R. K. Maeda, M. Hediger, F. Karch, and K. Basler. 2007. An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases. Proc Natl Acad Sci U S A **104**:3312–3317.

Bone, J. R., J. Lavender, R. Richman, M. J. Palmer, B. M. Turner, and M. I. Kuroda. 1994. Acetylated histone H4 on the male X chromosome is associated with dosage compensation in Drosophila. Genes and Development **8**:96–104.

Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem **72**:248–54.

Breen, T. R., and J. C. Lucchesi. 1986. Analysis of the dosage compensation of a specific transcript in Drosophila melanogaster. Genetics **112**:483–91.

Brickner, J. H., and P. Walter. 2004. Gene recruitment of the activated INO1 locus to the nuclear membrane. PLoS Biol **2**:e342.

Bridges, C. B. 1935. Salivary chromosome maps: With a Key to the Banding of the Chromosomes of Drosophila Melanogaster. J Hered **26**:60–64.

Bridges, C. B. 1938. A revised map of the salivary gland X-chromosome of Drosophila melanogaster. J Hered **29**:11–13.

Brough, R., A. M. Papanastasiou, and A. C. G. Porter. 2007. Stringent and reproducible tetracycline-regulated transgene expression by site-specific insertion at chromosomal loci with pre-characterised induction characteristics. BMC Mol Biol **8**:30.

Buscaino, A., T. Kocher, J. H. Kind, H. Holz, M. Taipale, K. Wagner, M. Wilm, and A. Akhtar. 2003. MOF-regulated acetylation of MSL-3 in the Drosophila dosage compensation complex. Molecular Cell **11**:1265–77.

Buscaino, A., G. Legube, and A. Akhtar. 2006. X-chromosome targeting and dosage compensation are mediated by distinct domains in MSL-3. EMBO Rep **7**:531–8.

Chang, K. A., and M. I. Kuroda. 1998. Modulation of MSL1 abundance in female Drosophila contributes to the sex specificity of dosage compensation. Genetics **150**:699–709.

Chen, Z.-Y., C.-Y. He, A. Ehrhardt, and M. A. Kay. 2003. Minicircle DNA vectors devoid of bacterial DNA result in persistent and high-level transgene expression in vivo. Mol Ther **8**:495–500.

Chiang, P. W., and D. M. Kurnit. 2003. Study of dosage compensation in Drosophila. Genetics **165**:1167–81.

Copps, K., R. Richman, L. M. Lyman, K. A. Chang, J. Rampersad-Ammons, and M. I. Kuroda. 1998. Complex formation by the Drosophila MSL proteins: role of the MSL2 RING finger in protein complex assembly. EMBO Journal **17**:5409–17.

Corona, D. F., C. R. Clapier, P. B. Becker, and J. W. Tamkun. 2002. Modulation of ISWI function by site-specific histone acetylation. EMBO Rep **3**:242–7.

Crowder, S. M., R. Kanaar, D. C. Rio, and T. Alber. 1999. Absence of interdomain contacts in the crystal structure of the RNA recognition motifs of Sex-lethal. Proceedings of the National Academy of Sciences of the United States of America **96**:4892–7.

Dahlsveen, I. K., G. D. Gilfillan, V. I. Shelest, R. Lamm, and P. B. Becker. 2006. Targeting determinants of dosage compensation in Drosophila. PLoS Genet **2**:e5.

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

Demakova, O. V., I. V. Kotlikova, P. R. Gordadze, A. A. Alekseyenko, M. I. Kuroda, and I. F. Zhimulev. 2003. The MSL complex levels are critical for its correct targeting to the chromosomes in Drosophila melanogaster. Chromosoma **112**:103–15.

Deng, H., W. Zhang, X. Bao, J. N. Martin, J. Girton, J. Johansen, and K. M. Johansen. 2005*a*. The JIL-1 kinase regulates the structure of Drosophila polytene chromosomes. Chromosoma **114**:173–82.

Deng, X., and V. H. Meller. 2006. roX RNAs are required for increased expression of X-linked genes in Drosophila melanogaster males. Genetics **174**:1859–66.

Deng, X., B. P. Rattner, S. Souter, and V. H. Meller. 2005*b*. The severity of roX1 mutations is predicted by MSL localization on the X chromosome. Mech Dev **122**:1094–105.

Deuring, R., L. Fanti, J. A. Armstrong, M. Sarte, O. Papoulas, M. Prestel, G. Daubresse, M. Verardo, S. L. Moseley, M. Berloco, T. Tsukiyama, C. Wu, S. Pimpinelli, and J. W. Tamkun. 2000. The ISWI chromatin-remodeling protein is required for gene expression and the maintenance of higher order chromatin structure in vivo. Mol Cell **5**:355–65.

Deuschle, U., W. K. Meyer, and H. J. Thiesen. 1995. Tetracycline-reversible silencing of eukaryotic promoters. Mol Cell Biol **15**:1907–1914.

Dogan, R. I., L. Getoor, W. J. Wilbur, and S. M. Mount. 2007. SplicePort–an interactive splice-site analysis tool. Nucleic Acids Res **35**:W285–91.

Duncan, I. W. 2002. Transvection effects in Drosophila. Annu Rev Genet **36**:521–56.

Ehrhardt, A., H. Xu, Z. Huang, J. A. Engler, and M. A. Kay. 2005. A direct comparison of two nonviral gene therapy vectors for somatic integration: In vivo evaluation of the bacteriophage integrase phi c31 and the Sleeping Beauty transposase. Molecular Therapy **11**:695–706.

Ellis, M. C., E. M. O'Neill, and G. M. Rubin. 1993. Expression of Drosophila glass protein and evidence for negative regulation of its activity in non-neuronal cells by another DNA-binding protein. Development **119**:855–65.

Fagegaltier, D., and B. S. Baker. 2004. X chromosome sites autonomously recruit the dosage compensation complex in Drosophila males. PLoS Biol **2**:e341.

Fang, S., K. L. Lorick, J. P. Jensen, and A. M. Weissman. 2003. RING finger ubiquitin protein ligases: implications for tumorigenesis, metastasis and for molecular targets in cancer. Semin Cancer Biol **13**:5–14.

Fitzsimons, H. L., R. A. Henry, and M. J. Scott. 1999. Development of an insulated reporter system to search for cis-acting DNA sequences required for dosage compensation in Drosophila. Genetica **105**:215–26.

Forster, K., V. Helbl, T. Lederer, S. Urlinger, N. Wittenburg, and W. Hillen. 1999. Tetracycline-inducible expression systems with reduced basal activity in mammalian cells. Nucleic Acids Res **27**:708–710.

Franke, A., and B. S. Baker. 1999. The rox1 and rox2 RNAs are essential components of the compensasome, which mediates dosage compensation in Drosophila. Molecular Cell **4**:117–22.

Franke, A., A. Dernburg, G. J. Bashaw, and B. S. Baker. 1996. Evidence that MSL-mediated dosage compensation in Drosophila begins at blastoderm. Development **122**:2751–60.

Freundlieb, S., C. Schirra-Muller, and H. Bujard. 1999. A tetracycline controlled activation/repression system with increased potential for gene transfer into mammalian cells. J Gene Med **1**:4–12.

Fujii, S., and H. Amrein. 2002. Genes expressed in the Drosophila head reveal a role for fat cells in sex-specific physiology. EMBO J **21**:5353–63.

Fukunaga, A., A. Tanaka, and K. Oishi. 1975. Maleless, a recessive autosomal mutant of Drosophila melanogaster that specifically kills male zygotes. Genetics **81**:135–141.

Fung, J. C., W. F. Marshall, A. Dernburg, D. A. Agard, and J. W. Sedat. 1998. Homologous chromosome pairing in Drosophila melanogaster proceeds through multiple independent initiations. J Cell Biol **141**:5–20.

Furth, P. A., L. St Onge, H. Böger, P. Gruss, M. Gossen, A. Kistner, H. Bujard, and L. Hennighausen. 1994. Temporal control of gene expression in transgenic mice by a tetracycline-responsive promoter. Proc Natl Acad Sci U S A **91**:9302–6.

Furuhashi, H., M. Nakajima, and S. Hirose. 2006. DNA supercoiling factor contributes to dosage compensation in Drosophila. Development **133**:4475–83.

Ganguly, R., N. Ganguly, and J. E. Manning. 1985. Isolation and characterization of the glucose-6-phosphate dehydrogenase gene of Drosophila melanogaster. Gene **35**:91–101.

Gans, M. 1953. Etude genetique et physiologique du mutant z de Drosophila melanogaster. Bull. Biol. Fr. Belge (suppl.) **38**:1–90.

Gebauer, F., M. Grskovic, and M. W. Hentze. 2003. Drosophila sex-lethal inhibits the stable association of the 40S ribosomal subunit with msl-2 mRNA. Molecular Cell **11**:1397–404.

Gemkow, M. J., P. J. Verveer, and D. J. Arndt-Jovin. 1998. Homologous association of the Bithorax-Complex during embryogenesis: consequences for transvection in Drosophila melanogaster. Development **125**:4541–52.

Geyer, P. K., M. M. Green, and V. G. Corces. 1990. Tissue-specific transcriptional enhancers may act in trans on the gene located in the homologous chromosome: the molecular basis of transvection in Drosophila. EMBO J **9**:2247–56.

Gibson, T. J., and J. D. Thompson. 1994. Detection of dsRNA-binding domains in RNA helicase A and Drosophila maleless: implications for monomeric RNA helicases. Nucleic Acids Res **22**:2552–2556.

Gilfillan, G. D., C. Konig, I. K. Dahlsveen, N. Prakoura, T. Straub, R. Lamm, T. Fauth, and P. B. Becker. 2007. Cumulative contributions of weak DNA determinants to targeting the Drosophila dosage compensation complex. Nucleic Acids Res **35**:3561–72.

Gilfillan, G. D., T. Straub, E. de Wit, F. Greil, R. Lamm, B. van Steensel, and P. B. Becker. 2006. Chromosome-wide gene-specific targeting of the Drosophila dosage compensation complex. Genes Dev **20**:858–70.

Glaser, R. L., M. F. Wolfner, and J. T. Lis. 1986. Spatial and temporal pattern of hsp26 expression during normal development. EMBO J **5**:747–754.

Gorman, M., A. Franke, and B. S. Baker. 1995. Molecular characterization of the male-specific lethal-3 gene and investigations of the regulation of dosage compensation in Drosophila. Development **121**:463–75.

Gorman, M., M. I. Kuroda, and B. S. Baker. 1993. Regulation of the sex-specific binding of the maleless dosage compensation protein to the male X chromosome in Drosophila. Cell **72**:39–49.

Gossen, M., and H. Bujard. 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc Natl Acad Sci U S A **89**:5547–5551.

Gossen, M., S. Freundlieb, G. Bender, G. Muller, W. Hillen, and H. Bujard. 1995. Transcriptional activation by tetracyclines in mammalian cells. Science **268**:1766–1769.

Granok, H., B. A. Leibovitch, C. D. Shaffer, and S. C. Elgin. 1995. Chromatin. Ga-ga over GAGA factor. Curr Biol **5**:238–41.

Greenberg, A. J., J. L. Yanowitz, and P. Schedl. 2004. The Drosophila GAGA factor is required for dosage compensation in males and for the formation of the male-specific-lethal complex chromatin entry site at 12DE. Genetics **166**:279–89.

Grether, M. E., J. M. Abrams, J. Agapite, K. White, and H. Steller. 1995. The head involution defective gene of Drosophila melanogaster functions in programmed cell death. Genes Dev **9**:1694–708.

Groth, A. C., M. Fish, R. Nusse, and M. P. Calos. 2004. Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31. Genetics **166**:1775–82.

Groth, A. C., E. C. Olivares, B. Thyagarajan, and M. P. Calos. 2000. A phage integrase directs efficient site-specific integration in human cells. Proceedings Of The National Academy Of Sciences Of The United States Of America **97**:5995–6000.

Gu, W., P. Szauter, and J. C. Lucchesi. 1998. Targeting of MOF, a putative histone acetyl transferase, to the X chromosome of Drosophila melanogaster. Developmental Genetics **22**:56–64.

Gu, W., X. Wei, A. Pannuti, and J. C. Lucchesi. 2000. Targeting the chromatin-remodeling MSL complex of Drosophila to its sites of action on the X chromosome requires both acetyl transferase and ATPase activities. EMBO Journal **19**:5202–11.

Gupta, V., M. Parisi, D. Sturgill, R. Nuttall, M. Doctolero, O. K. Dudko, J. D. Malley, P. S. Eastman, and B. Oliver. 2006. Global analysis of X-chromosome dosage compensation. J Biol **5**:3.

Hamada, F. N., P. J. Park, P. R. Gordadze, and M. I. Kuroda. 2005. Global regulation of X chromosomal genes by the MSL complex in Drosophila melanogaster. Genes Dev **19**:2289–94.

Han, P. L., V. Meller, and R. L. Davis. 1996. The Drosophila brain revisited by enhancer detection. J Neurobiol **31**:88–102.

Handler, A. M., and R. A. Harrell, 2nd. 1999. Germline transformation of Drosophila melanogaster with the piggyBac transposon vector. Insect Mol Biol **8**:449–57.

Hartenstein, V. 1993. Atlas of Drosophila Development. Cold Spring Harbor Laboratory Press, New York.

Hatton, A. R., V. Subramaniam, and A. J. Lopez. 1998. Generation of alternative Ultrabithorax isoforms and stepwise removal of a large intron by resplicing at exon-exon junctions. Mol Cell **2**:787–796.

Hay, B. A., R. Maile, and G. M. Rubin. 1997. P element insertion-dependent gene activation in the Drosophila eye. Proc Natl Acad Sci U S A **94**:5195–200.

Hay, B. A., T. Wolff, and G. M. Rubin. 1994. Expression of baculovirus P35 prevents cell death in Drosophila. Development **120**:2121–9.

Hazelrigg, T., R. Levis, and G. M. Rubin. 1984. Transformation of white locus DNA in drosophila: dosage compensation, zeste interaction, and position effects. Cell **36**:469–81.

Heinrich, J. C., and M. J. Scott. 2000. A repressible female-specific lethal genetic system for making transgenic insect strains suitable for a sterile-release program. Proc Natl Acad Sci U S A **97**:8229–8232.

Henikoff, S., and T. D. Dreesen. 1989. Trans-inactivation of the Drosophila brown gene: evidence for transcriptional repression and somatic pairing dependence. Proc Natl Acad Sci U S A **86**:6704–8.

Henry, R. A., B. Tews, X. Li, and M. J. Scott. 2001. Recruitment of the male-specific lethal (MSL) dosage compensation complex to an autosomally integrated roX chromatin entry site correlates with an increased expression of an adjacent reporter gene in male Drosophila. Journal of Biological Chemistry **276**:31953–8.

Hilfiker, A., D. Hilfiker-Kleiner, A. Pannuti, and J. C. Lucchesi. 1997. mof, a putative acetyl transferase gene related to the Tip60 and MOZ human genes and to the SAS genes of yeast, is required for dosage compensation in Drosophila. EMBO Journal **16**:2054–60.

Hillen, W., C. Gatz, L. Altschmied, K. Schollmeier, and I. Meier. 1983. Control of expression of the Tn10-encoded tetracycline resistance genes. Equilibrium and kinetic investigation of the regulatory reactions. J Mol Biol **169**:707–21.

Hinrichs, W., C. Kisker, M. Düvel, A. Müller, K. Tovar, W. Hillen, and W. Saenger. 1994. Structure of the Tet repressor-tetracycline complex and regulation of antibiotic resistance. Science **264**:418–20.

Hiromi, Y., and W. J. Gehring. 1987. Regulation and function of the Drosophila segmentation gene fushi tarazu. Cell **50**:963–74.

Hirose, F., M. Yamaguchi, H. Handa, Y. Inomata, and A. Matsukage. 1993. Novel 8-base pair sequence (Drosophila DNA replication-related element) and specific binding factor involved in the expression of Drosophila genes for DNA polymerase alpha and proliferating cell nuclear antigen. J Biol Chem **268**:2092–2099.

Hirose, F., M. Yamaguchi, K. Kuroda, A. Omori, T. Hachiya, M. Ikeda, Y. Nishimoto, and A. Matsukage. 1996. Isolation and characterization of cDNA for DREF, a promoter-activating factor for Drosophila DNA replication-related genes. J Biol Chem **271**:3930–3937.

Hochheimer, A., S. Zhou, S. Zheng, M. C. Holmes, and R. Tjian. 2002. TRF2 associates with DREF and directs promoter-selective gene expression in Drosophila. Nature **420**:439–445.

Horn, C., and A. M. Handler. 2005. Site-specific genomic targeting in Drosophila. Proceedings Of The National Academy Of Sciences Of The United States Of America **102**:12483–12488.

Horn, C., B. Jaunich, and E. A. Wimmer. 2000. Highly sensitive, fluorescent transformation marker for Drosophila transgenesis. Dev Genes Evol **210**:623–629.

Horn, C., and E. A. Wimmer. 2003. A transgene-based, embryo-specific lethality system for insect pest management. Nat Biotechnol **21**:64–70.

Hoshijima, K., K. Inoue, I. Higuchi, H. Sakamoto, and Y. Shimura. 1991. Control of doublesex alternative splicing by transformer and transformer-2 in Drosophila. Science **252**:833–6.

Ibnsouda, S., F. Schweisguth, G. de Billy, and A. Vincent. 1993. Relationship between expression of serendipity alpha and cellularisation of the Drosophila embryo as revealed by interspecific transformation. Development **119**:471–83.

Ibnsouda, S., F. Schweisguth, D. Jullien, C. Kücherer, J. A. Lepesant, and A. Vincent. 1995. Evolutionarily conserved positive and negative cis-acting elements control the blastoderm-specific expression of the Drosophila serendipity alpha cellularisation gene. Mech Dev **49**:71–82.

Inoue, H., H. Nojima, and H. Okayama. 1990. High efficiency transformation of Escherichia coli with plasmids. Gene **96**:23–28.

Jack, J. W., and B. H. Judd. 1979. Allelic pairing and gene regulation: A model for the zeste-white interaction in Drosophila melanogaster. Proc Natl Acad Sci U S A **76**:1368–1372.

Jin, Y., Y. Wang, J. Johansen, and K. M. Johansen. 2000. JIL-1, a chromosomal kinase implicated in regulation of chromatin structure, associates with the male specific lethal (MSL) dosage compensation complex. Journal of Cell Biology **149**:1005–10.

Jin, Y., Y. Wang, D. L. Walker, H. Dong, C. Conley, J. Johansen, and K. M. Johansen. 1999. JIL-1: a novel chromosomal tandem kinase implicated in transcriptional regulation in Drosophila. Molecular Cell **4**:129–135.

Kageyama, Y., G. Mengus, G. Gilfillan, H. G. Kennedy, C. Stuckenholz, R. L. Kelley, P. B. Becker, and M. I. Kuroda. 2001. Association and spreading of the Drosophila dosage compensation complex from a discrete roX1 chromatin entry site. EMBO Journal **20**:2236–45.

Kaiser, K., M. Furia, and D. M. Glover. 1986. Dosage compensation at the sgs4 locus of Drosophila melanogaster. Journal of Molecular Biology **187**:529–36.

Katsani, K. R., M. A. Hajibagheri, and C. P. Verrijzer. 1999. Co-operative DNA binding by GAGA transcription factor requires the conserved BTB/POZ domain and reorganizes promoter topology. EMBO J **18**:698–708.

Kelley, R. L., and M. I. Kuroda. 2003. The Drosophila roX1 RNA gene can overcome silent chromatin by recruiting the male-specific lethal dosage compensation complex. Genetics **164**:565–74.

Kelley, R. L., V. H. Meller, P. R. Gordadze, G. Roman, R. L. Davis, and M. I. Kuroda. 1999. Epigenetic spreading of the Drosophila dosage compensation complex from roX RNA genes into flanking chromatin. Cell **98**:513–22.

Kelley, R. L., I. Solovyeva, L. M. Lyman, R. Richman, V. Solovyev, and M. I. Kuroda. 1995. Expression of msl-2 causes assembly of dosage compensation regulators on the X chromosomes and female lethality in Drosophila. Cell **81**:867–77.

Kelley, R. L., J. Wang, L. Bell, and M. I. Kuroda. 1997. Sex lethal controls dosage compensation in Drosophila by a non-splicing mechanism. Nature **387**:195–9.

Kind, J., and A. Akhtar. 2007. Cotranscriptional recruitment of the dosage compensation complex to X-linked target genes. Genes Dev **21**:2030–40.

Kind, J., J. M. Vaquerizas, P. Gebhardt, M. Gentzel, N. M. Luscombe, P. Bertone, and A. Akhtar. 2008. Genome-wide analysis reveals MOF as a key regulator of dosage compensation and gene expression in Drosophila. Cell **133**:813–28.

Kistner, A., M. Gossen, F. Zimmermann, J. Jerecic, C. Ullmer, H. Lübbert, and H. Bujard. 1996. Doxycycline-mediated quantitative and tissue-specific control of gene expression in transgenic mice. Proc Natl Acad Sci U S A **93**:10933–8.

Korge, G. 1975. Chromosome puff activity and protein synthesis in larval salivary glands of Drosophila melanogaster. Proc Natl Acad Sci U S A **72**:4550–4554.

Kotlikova, I. V., O. V. Demakova, V. F. Semeshin, V. V. Shloma, L. V. Boldyreva, M. I. Kuroda, and I. F. Zhimulev. 2006. The Drosophila dosage compensation complex binds to polytene chromosomes independently of developmental changes in transcription. Genetics **172**:963–74.

Krumm, A., G. E. Roth, and G. Korge. 1985. Transformation of salivary gland secretion protein gene Sgs-4 in Drosophila: stage- and tissue-specific regulation, dosage compensation, and position effect. Proceedings of the National Academy of Sciences of the United States of America **82**:5055–9.

Kuroda, M. I., M. J. Kernan, R. Kreber, B. Ganetzky, and B. S. Baker. 1991. The maleless protein associates with the X chromosome to regulate dosage compensation in Drosophila. Cell **66**:935–47.

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227**:680–685.

Lakhotia, S. C., and A. S. Mukherjee. 1969. Chromosomal basis of dosage compensation in Drosophila. I. Cellular autonomy of hyperactivity of the male X-chromosome in salivary glands and sex differentiation. Genetical Research **14**:137–50.

Larschan, E., A. A. Alekseyenko, A. A. Gortchakov, S. Peng, B. Li, P. Yang, J. L. Workman, P. J. Park, and M. I. Kuroda. 2007. MSL complex is attracted to genes marked by H3K36 trimethylation using a sequence-independent mechanism. Mol Cell **28**:121–33.

Laverty, C., 2003. Development of a reporter gene construct to identify further components of the Drosophila melanogaster dosage compensation complex. Bachelor of Science (Honours) dissertation, Massey University, Palmerston North, New Zealand.

Lee, A. M., and C.-T. Wu. 2006. Enhancer-promoter communication at the yellow gene of Drosophila melanogaster: diverse promoters participate in and regulate trans interactions. Genetics **174**:1867–80.

Lee, C. G., K. A. Chang, M. I. Kuroda, and J. Hurwitz. 1997. The NTPase/helicase activities of Drosophila maleless, an essential factor in dosage compensation. EMBO Journal **16**:2671–81.

Legube, G., S. K. McWeeney, M. J. Lercher, and A. Akhtar. 2006. X-chromosome-wide profiling of MSL-1 distribution and dosage compensation in Drosophila. Genes Dev **20**:871–83.

Lerach, S., W. Zhang, H. Deng, X. Bao, J. Girton, J. Johansen, and K. M. Johansen. 2005. JIL-1 kinase, a member of the male-specific lethal (MSL) complex, is necessary for proper dosage compensation of eye pigmentation in Drosophila. Genesis **43**:213–5.

Lewis, E. 1954. The theory and application of a new method of detecting chromosomal rearrangements in drosophila-melanogaster. American Naturalist **88**:225–239.

Li, F., D. A. Parry, and M. J. Scott. 2005. The amino-terminal region of Drosophila MSL1 contains basic, glycine-rich, and leucine zipper-like motifs that promote X chromosome binding, self-association, and MSL2 binding, respectively. Mol Cell Biol **25**:8913–24.

Li, F., A. H. Schiemann, and M. J. Scott. 2008. Incorporation of the noncoding roX RNAs alters the chromatin-binding specificity of the Drosophila MSL1/MSL2 complex. Mol Cell Biol **28**:1252–64.

Li, X., 2002. Progress towards the development of a genetically modified strain of the Australia Sheep Blowfly *Lucilia cuprina* suitable for a sterile release program. Ph.D. thesis, Massey University, Palmerston North, New Zealand.

Lis, J. T., J. A. Simon, and C. A. Sutton. 1983. New heat shock puffs and beta-galactosidase activity resulting from transformation of Drosophila with an hsp70-lacZ hybrid gene. Cell **35**:403–10.

Liu, L.-P., J.-Q. Ni, Y.-D. Shi, E. J. Oakeley, and F.-L. Sun. 2005. Sex-specific role of Drosophila melanogaster HP1 in regulating chromatin structure and gene transcription. Nat Genet **37**:1361–6.

Locke, J., and K. D. Tartof. 1994. Molecular analysis of cubitus interruptus (ci) mutations suggests an explanation for the unusual ci position effects. Mol Gen Genet **243**:234–43.

Lucchesi, J., T. Skripsky, and F. Tax. 1982. A new male-specific lethal mutation in Drosophila melanogaster. Genetics **100**:s42.

Lyman, L. M., K. Copps, L. Rastelli, R. L. Kelley, and M. I. Kuroda. 1997. Drosophila male-specific lethal-2 protein: structure/function analysis and dependence on MSL-1 for chromosome association. Genetics **147**:1743–53.

Lynch, K. W., and T. Maniatis. 1996. Assembly of specific SR protein complexes on distinct regulatory elements of the Drosophila doublesex splicing enhancer. Genes Dev **10**:2089–101.

Lyon, M. F. 1961. Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature **190**:372–373.

McDowell, K. A., A. Hilfiker, and J. C. Lucchesi. 1996. Dosage compensation in Drosophila: the X chromosome binding of MSL-1 and MSL-2 in female embryos is prevented by the early expression of the Sxl gene. Mechanisms of Development **57**:113–9.

McKee, B. D. 2004. Homologous pairing and chromosome dynamics in meiosis and mitosis. Biochim Biophys Acta **1677**:165–80.

Meller, V. H. 2003. Initiation of dosage compensation in Drosophila embryos depends on expression of the roX RNAs. Mech Dev **120**:759–67.

Meller, V. H., P. R. Gordadze, Y. Park, X. Chu, C. Stuckenholz, R. L. Kelley, and M. I. Kuroda. 2000. Ordered assembly of roX RNAs into MSL complexes on the dosage-compensated X chromosome in Drosophila. Current Biology **10**:136–43.

Meller, V. H., and B. P. Rattner. 2002. The roX genes encode redundant male-specific lethal transcripts required for targeting of the MSL complex. Embo J **21**:1084–91.

Meller, V. H., K. H. Wu, G. Roman, M. I. Kuroda, and R. L. Davis. 1997. roX1 RNA paints the X chromosome of male Drosophila and is regulated by the dosage compensation system. Cell **88**:445–57.

Mendjan, S., and A. Akhtar. 2007. The right dose for every sex. Chromosoma **116**:95–106.

Mendjan, S., M. Taipale, J. Kind, H. Holz, P. Gebhardt, M. Schelder, M. Vermeulen, A. Buscaino, K. Duncan, J. Mueller, M. Wilm, H. G. Stunnenberg, H. Saumweber, and A. Akhtar. 2006. Nuclear pore components are involved in the transcriptional regulation of dosage compensation in Drosophila. Mol Cell **21**:811–23.

Meyer, B. J., and L. P. Casson. 1986. Caenorhabditis elegans compensates for the difference in X chromosome dosage between the sexes by regulating transcript levels. Cell **47**:871–881.

Mikkelsen, T. S., M. Ku, D. B. Jaffe, B. Issac, E. Lieberman, G. Giannoukos, P. Alvarez, W. Brockman, T.-K. Kim, R. P. Koche, W. Lee, E. Mendenhall, A. O'Donovan, A. Presser, C. Russ, X. Xie, A. Meissner, M. Wernig, R. Jaenisch, C. Nusbaum, E. S. Lander, and B. E. Bernstein. 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature **448**:553–60.

Mito, Y., J. G. Henikoff, and S. Henikoff. 2005. Genome-scale profiling of histone H3.3 replacement patterns. Nat Genet **37**:1090–7.

Morales, V., T. Straub, M. F. Neumann, G. Mengus, A. Akhtar, and P. B. Becker. 2004. Functional integration of the histone acetyltransferase MOF into the dosage compensation complex. Embo J **23**:2258–68.

Morris, J. R., J. L. Chen, P. K. Geyer, and C. T. Wu. 1998. Two modes of transvection: enhancer action in trans and bypass of a chromatin insulator in cis. Proc Natl Acad Sci U S A **95**:10740–5.

Morris, J. R., D. A. Petrov, A. M. Lee, and C.-T. Wu. 2004. Enhancer choice in cis and in trans in Drosophila melanogaster: role of the promoter. Genetics **167**:1739–47.

Moses, K., M. C. Ellis, and G. M. Rubin. 1989. The glass gene encodes a zinc-finger protein required by Drosophila photoreceptor cells. Nature **340**:531–6.

Moses, K., and G. M. Rubin. 1991. Glass encodes a site-specific DNA-binding protein that is regulated in response to positional signals in the developing Drosophila eye. Genes Dev **5**:583–93.

Mukherjee, A. S., and W. Beermann. 1965. Synthesis of ribonucleic acid by the X-chromosomes of Drosophila melanogaster and the problem of dosage compensation. Nature **207**:785–6.

Muller, H., B. League, and C. Offermann. 1931. Effects of dosage changes of sex-linked genes, and the compensatory effects of other gene differences between male and female. Anat. Rec. Suppl. **51**.

Muller, H. J. 1932. Further studies on the nature and causes of gene mutations. Proceedings of the International Congress on Genetics, 6th **1**:215–255.

Oberstein, A., A. Pare, L. Kaplan, and S. Small. 2005. Site-specific transgenesis by Cre-mediated recombination in Drosophila. Nature Methods **2**:583–585.

Offermann, C. 1936. Branched chromosomes as symmetrical duplications. Journal of Genetics **32**:103–116.

Oh, H., J. R. Bone, and M. I. Kuroda. 2004. Multiple classes of MSL binding sites target dosage compensation to the X chromosome of Drosophila. Curr Biol **14**:481–7.

Oh, H., Y. Park, and M. I. Kuroda. 2003. Local spreading of MSL complexes from roX genes on the Drosophila X chromosome. Genes and Development **17**:1334–9.

Ohler, U., G.-c. Liao, H. Niemann, and G. M. Rubin. 2002. Computational analysis of core promoters in the Drosophila genome. Genome Biol **3**:RESEARCH0087.

Olivares, E. C., R. P. Hollis, T. W. Chalberg, L. Meuse, M. A. Kay, and M. P. Calos. 2002. Site-specific genomic integration produces therapeutic Factor IX levels in mice. Nat Biotechnol **20**:1124–8.

O'Sullivan, J. M., S. M. Tan-Wong, A. Morillon, B. Lee, J. Coles, J. Mellor, and N. J. Proudfoot. 2004. Gene loops juxtapose promoters and terminators in yeast. Nat Genet **36**:1014–8.

Palmer, M. J., V. A. Mergner, R. Richman, J. E. Manning, M. I. Kuroda, and J. C. Lucchesi. 1993. The male-specific lethal-one (msl-1) gene of Drosophila melanogaster encodes a novel protein that associates with the X chromosome in males. Genetics **134**:545–57.

Palmer, M. J., R. Richman, L. Richter, and M. I. Kuroda. 1994. Sex-specific regulation of the male-specific lethal-1 dosage compensation gene in Drosophila. Genes and Development **8**:698–706.

Park, Y., R. L. Kelley, H. Oh, M. I. Kuroda, and V. H. Meller. 2002. Extent of chromatin spreading determined by roX RNA recruitment of MSL proteins. Science **298**:1620–1623.

Park, Y., G. Mengus, X. Bai, Y. Kageyama, V. H. Meller, P. B. Becker, and M. I. Kuroda. 2003. Sequence-specific targeting of Drosophila roX genes by the MSL dosage compensation complex. Molecular Cell **11**:977–86.

Park, Y., H. Oh, V. H. Meller, and M. I. Kuroda. 2005. Variable Splicing of Non-Coding roX2 RNAs Influences Targeting of MSL Dosage Compensation Complexes in Drosophila. RNA Biol **2**.

Pattatucci, A. M., and T. C. Kaufman. 1991. The homeotic gene Sex combs reduced of Drosophila melanogaster is differentially regulated in the embryonic and imaginal stages of development. Genetics **129**:443–61.

Pirrotta, V., H. Steller, and M. P. Bozzetti. 1985. Multiple upstream regulatory elements control the expression of the Drosophila white gene. EMBO Journal **4**:3501–8.

Plautz, J. D., R. N. Day, G. M. Dailey, S. B. Welsh, J. C. Hall, S. Halpain, and S. A. Kay. 1996. Green fluorescent protein and its derivatives as versatile markers for gene expression in living Drosophila melanogaster, plant and mammalian cells. Gene **173**:83–7.

Pluta, K., M. J. Luce, L. Bao, S. Agha-Mohammadi, and J. Reiser. 2005. Tight control of transgene expression by lentivirus vectors containing second-generation tetracycline-responsive promoters. J Gene Med **7**:803–817.

Pokholok, D., C. Harbison, S. Levine, M. Cole, N. Hannett, T. Lee, G. Bell, K. Walker, P. Rolfe, E. Herbolsheimer, J. Zeitlinger, F. Lewitter, D. Gifford, and R. Young. 2005. Genome-wide map of nucleosorne acetylation and methylation in yeast. Cell **122**:517–527.

Postner, M. A., and E. F. Wieschaus. 1994. The nullo protein is a component of the actin-myosin network that mediates cellularization in Drosophila melanogaster embryos. J Cell Sci **107 (Pt 7)**:1863–73.

Qian, S., and V. Pirrotta. 1995. Dosage compensation of the Drosophila white gene requires both the X chromosome environment and multiple intragenic elements. Genetics **139**:733–44.

Qian, S., B. Varjavand, and V. Pirrotta. 1992. Molecular analysis of the zeste-white interaction reveals a promoter-proximal element essential for distant enhancer-promoter communication. Genetics **131**:79–90.

Rao, B., Y. Shibata, B. Strahl, and J. Lieb. 2005. Dimethylation of histone H3 at lysine 36 demarcates regulatory and nonregulatory chromatin genome-wide. Molecular and Cellular Biology **25**:9447–9459.

Rastelli, L., and M. I. Kuroda. 1998. An analysis of maleless and histone H4 acetylation in Drosophila melanogaster spermatogenesis. Mechanisms of Development **71**:107–17.

Rastelli, L., R. Richman, and M. I. Kuroda. 1995. The dosage compensation regulators MLE, MSL-1 and MSL-2 are interdependent since early embryogenesis in Drosophila. Mechanisms of Development **53**:223–33.

Rattner, B. P., and V. H. Meller. 2004. Drosophila male-specific lethal 2 protein controls sex-specific expression of the roX genes. Genetics **166**:1825–32.

Reese, M. G., F. H. Eeckman, D. Kulp, and D. Haussler. 1997. Improved splice site detection in Genie. J Comput Biol **4**:311–23.

Richter, L., J. R. Bone, and M. I. Kuroda. 1996. RNA-dependent association of the Drosophila maleless protein with the male X chromosome. Genes to Cells **1**:325–36.

Riggleman, B., E. Wieschaus, and P. Schedl. 1989. Molecular analysis of the armadillo locus: uniformly distributed transcripts and a protein with novel internal repeats are associated with a Drosophila segment polarity gene. Genes Dev **3**:96–113.

Rio, D. C., and G. M. Rubin. 1985. Transformation of cultured Drosophila melanogaster cells with a dominant selectable marker. Mol Cell Biol **5**:1833–8.

Roehrdanz, R. L., and J. C. Lucchesi. 1977. Transcriptional controls and dosage compensation in Drosophila melanogaster. Nature **269**:243–5.

Rose, L. S., and E. Wieschaus. 1992. The Drosophila cellularization gene nullo produces a blastoderm-specific transcript whose levels respond to the nucleocytoplasmic ratio. Genes Dev **6**:1255–68.

Rubin, G. M., and A. C. Spradling. 1982. Genetic transformation of Drosophila with transposable element vectors. Science **218**:348–53.

Rudkin, G., 1964. The Nucleohistones, Chapter the proteins of polytene chromosomes, pages 184–192 . Holden-Day, San Francisco.

Ryu, J. R., L. K. Olson, and D. N. Arnosti. 2001. Cell-type specificity of short-range transcriptional repressors. Proc Natl Acad Sci U S A **98**:12960–5.

Sackton, T. B., and A. G. Clark. 2009. Comparative profiling of the transcriptional response to infection in two species of Drosophila by short-read cDNA sequencing. BMC Genomics **10**:259.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory Press.

Samuels, M. E., D. Bopp, R. A. Colvin, R. F. Roscigno, M. A. Garcia-Blanco, and P. Schedl. 1994. RNA binding by Sxl proteins in vitro and in vivo. Mol Cell Biol **14**:4975–90.

Sarkar, A., A. Atapattu, E. J. Belikoff, J. C. Heinrich, X. Li, C. Horn, E. A. Wimmer, and M. J. Scott. 2006. Insulated piggyBac vectors for insect transgenesis. BMC Biotechnol **6**:27.

Sass, G. L., A. Pannuti, and J. C. Lucchesi. 2003. Male-specific lethal complex of Drosophila targets activated regions of the X chromosome for chromatin remodeling. Proceedings of the National Academy of Sciences of the United States of America **100**:8287–91.

Saunders, A., L. J. Core, and J. T. Lis. 2006. Breaking barriers to transcription elongation. Nat Rev Mol Cell Biol **7**:557–67.

Saura, A. O., T. I. Heino, and V. Sorsa. 1993. Electron-microscopic analysis of the banding-pattern in the salivary-gland chromosomes of drosophila-melanogaster - divisions 11 through 20 of X. Hereditas **119**:123–141.

Saurin, A. J., Z. Shao, H. Erdjument-Bromage, P. Tempst, and R. E. Kingston. 2001. A Drosophila Polycomb group complex includes Zeste and dTAFII proteins. Nature **412**:655–60.

Schlake, T., and J. Bode. 1994. Use of mutated FLP recognition target (FRT) sites for the exchange of expression cassettes at defined chromosomal loci. Biochemistry **33**:12746–51.

Schnetzer, J. W., and M. S. Tyler. 1996. Endogenous beta-galactosidase activity in the larval, pupal, and adult stages of the fruit fly, Drosophila melanogaster, indicates need for caution in lacZ fusion-gene studies. Biol Bull **190**:173–87.

Scholnick, S. B., B. A. Morgan, and J. Hirsh. 1983. The cloned dopa decarboxylase gene is developmentally regulated when reintegrated into the Drosophila genome. Cell **34**:37–45.

Scott, M. J., J. C. Heinrich, and X. Li. 2004. Progress towards the development of a transgenic strain of the Australian sheep blowfly (Lucilia cuprina) suitable for a male-only sterile release program. Insect Biochem Mol Biol **34**:185–192.

Scott, M. J., L. L. Pan, S. B. Cleland, A. L. Knox, and J. Heinrich. 2000. MSL1 plays a central role in assembly of the MSL complex, essential for dosage compensation in Drosophila. EMBO Journal **19**:144–55.

Seecof, R. L., W. D. Kaplan, and D. G. Futch. 1969. Dosage compensation for enzyme activities in Drosophila melanogaster. Proceedings of the National Academy of Sciences of the United States of America **62**:528–35.

Seto, H., Y. Hayashi, E. Kwon, O. Taguchi, and M. Yamaguchi. 2006. Antagonistic regulation of the Drosophila PCNA gene promoter by DREF and Cut. Genes Cells **11**:499–512.

Seum, C., D. Pauli, M. Delattre, Y. Jaquet, A. Spierer, and P. Spierer. 2002. Isolation of Su(var)3-7 mutations by homologous recombination in Drosophila melanogaster. Genetics **161**:1125–36.

Shapiro, J. A., W. Huang, C. Zhang, M. J. Hubisz, J. Lu, D. A. Turissini, S. Fang, H.-Y. Wang, R. R. Hudson, R. Nielsen, Z. Chen, and C.-I. Wu. 2007. Adaptive genic evolution in the Drosophila genomes. Proc Natl Acad Sci U S A **104**:2271–6.

Simon, J. A., and J. T. Lis. 1987. A germline transformation analysis reveals flexibility in the organization of heat shock consensus elements. Nucleic Acids Res **15**:2971–2988.

Singh, R., J. Valcárcel, and M. R. Green. 1995. Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. Science **268**:1173–6.

Smith, E. R., C. D. Allis, and J. C. Lucchesi. 2001. Linking global histone acetylation to the transcription enhancement of X-chromosomal genes in Drosophila males. Journal of Biological Chemistry **276**:31483–6.

Smith, E. R., A. Pannuti, W. Gu, A. Steurnagel, R. G. Cook, C. D. Allis, and J. C. Lucchesi. 2000. The drosophila MSL complex acetylates histone H4 at lysine 16, a chromatin modification linked to dosage compensation. Molecular and Cellular Biology **20**:312–8.

Sorsa, V., and A. O. Saura. 1980*a*. Electron-microscopic analysis of the banding-pattern in the salivary-gland chromosomes of drosophila-melanogaster - division 1 and 2 of X. Hereditas **92**:73–83.

Sorsa, V., and A. O. Saura. 1980*b*. Electron-microscopic analysis of the banding-pattern in the salivary-gland chromosomes of drosophila-melanogaster - divisions 3, 4 and 5 of X. Hereditas **92**:341–351.

Sorsa, V., A. O. Saura, and T. I. Heino. 1983. Electron-microscopic analysis of the banding-pattern in the salivary-gland chromosomes of drosophila-melanogaster - division-6 through division-10 of X. Hereditas **98**:181–200.

Sosnowski, B. A., J. M. Belote, and M. McKeown. 1989. Sex-specific alternative splicing of RNA from the transformer gene results from sequence-dependent splice site blockage. Cell **58**:449–459.

Spradling, A. C., and G. M. Rubin. 1982. Transposition of cloned P elements into Drosophila germ line chromosomes. Science **218**:341–347.

Spradling, A. C., and G. M. Rubin. 1983. The effect of chromosomal position on the expression of the Drosophila xanthine dehydrogenase gene. Cell **34**:47–57.

Stebbins, M. J., S. Urlinger, G. Byrne, B. Bello, W. Hillen, and J. C. Yin. 2001. Tetracycline-inducible systems for Drosophila. Proc Natl Acad Sci U S A **98**:10775–10780.

Straub, T., G. D. Gilfillan, V. K. Maier, and P. B. Becker. 2005*a*. The Drosophila MSL complex activates the transcription of target genes. Genes Dev **19**:2284–8.

Straub, T., C. Grimaud, G. D. Gilfillan, A. Mitterweger, and P. B. Becker. 2008. The chromosomal high-affinity binding sites for the Drosophila dosage compensation complex. PLoS Genet **4**:e1000302.

Straub, T., M. F. Neumann, M. Prestel, E. Kremmer, C. Kaether, C. Haass, and P. B. Becker. 2005*b*. Stable chromosomal association of MSL2 defines a dosage-compensated nuclear compartment. Chromosoma **114**:352–64.

Sural, T. H., S. Peng, B. Li, J. L. Workman, P. J. Park, and M. I. Kuroda. 2008. The MSL3 chromodomain directs a key targeting step for dosage compensation of the Drosophila melanogaster X chromosome. Nat Struct Mol Biol **15**:1318–25.

Taddei, A., G. Van Houwe, F. Hediger, V. Kalck, F. Cubizolles, H. Schober, and S. M. Gasser. 2006. Nuclear pore association confers optimal expression levels for an inducible yeast gene. Nature **441**:774–8.

Thorpe, H. M., and M. C. M. Smith. 1998. In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. Proceedings Of The National Academy Of Sciences Of The United States Of America **95**:5505–5510.

Thorpe, H. M., S. E. Wilson, and M. C. M. Smith. 2000. Control of directionality in the site-specific recombination system of the Streptomyces phage phi C31. Molecular Microbiology **38**:232–241.

Thyagarajan, B., E. C. Olivares, R. P. Hollis, D. S. Ginsburg, and M. P. Calos. 2001. Site-specific genomic integration in mammalian cells mediated by phage phiC31 integrase. Mol Cell Biol **21**:3926–34.

Tobler, J., J. T. Bowman, and J. R. Simmons. 1971. Gene modulation in Drosophila: dosage compensation and relocated v + genes. Biochemical Genetics **5**:111–7.

Turner, B. M., A. J. Birley, and J. Lavender. 1992. Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in Drosophila polytene nuclei. Cell **69**:375–84.

Tweedie, S., M. Ashburner, K. Falls, P. Leyland, P. McQuilton, S. Marygold, G. Millburn, D. Osumi-Sutherland, A. Schroeder, R. Seal, H. Zhang, and FlyBase Consortium. 2009. FlyBase: enhancing Drosophila Gene Ontology annotations. Nucleic Acids Res **37**:D555–9.

Urlinger, S., U. Baron, M. Thellmann, M. T. Hasan, H. Bujard, and W. Hillen. 2000. Exploring the sequence space for tetracycline-dependent transcriptional activators: novel mutations yield expanded range and sensitivity. Proc Natl Acad Sci U S A **97**:7963–7968.

van Steensel, B., J. Delrow, and H. J. Bussemaker. 2003. Genomewide analysis of Drosophila GAGA factor target genes reveals context-dependent DNA binding. Proc Natl Acad Sci U S A **100**:2580–5.

Venken, K. J. T., Y. He, R. A. Hoskins, and H. J. Bellen. 2006. P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in D. melanogaster. Science **314**:1747–51.

Viens, A., U. Mechold, H. Lehrmann, A. Harel-Bellan, and V. Ogryzko. 2004. Use of protein biotinylation in vivo for chromatin immunoprecipitation. Anal Biochem **325**:68–76.

Vincent, A., H. V. Colot, and M. Rosbash. 1986. Blastoderm-specific and read-through transcription of the sry alpha gene transformed into the Drosophila genome. Dev Biol **118**:480–7.

Vincent, J. P., C. H. Girdham, and P. H. O'Farrell. 1994. A cell-autonomous, ubiquitous marker for the analysis of Drosophila genetic mosaics. Dev Biol **164**:328–331.

Wakimoto, B. T., L. J. Kalfayan, and A. C. Spradling. 1986. Developmentally regulated expression of Drosophila chorion genes introduced at diverse chromosomal positions. J Mol Biol **187**:33–45.

Wang, Y., W. Zhang, Y. Jin, J. Johansen, and K. M. Johansen. 2001. The JIL-1 tandem kinase mediates histone H3 phosphorylation and is required for maintenance of chromatin structure in Drosophila. Cell **105**:433–443.

Weake, V. M., and M. J. Scott. 2007. The non-dosage compensated Lsp1alpha gene of Drosophila melanogaster escapes acetylation by MOF in larval fat body nuclei, but is flanked by two dosage compensated genes. BMC Mol Biol **8**:35.

Weir, M., and M. Rice. 2004. Ordered partitioning reveals extended splice-site consensus information. Genome Res **14**:67–78.

Whitney, J., and J. Lucchesi. 1972. Ontogenic expression of fumarase activity in Drosophila-melanogaster. Insect Biochemistry **2**:367–&.

Williams, B. R., J. R. Bateman, N. D. Novikov, and C.-T. Wu. 2007. Disruption of topoisomerase II perturbs pairing in drosophila cell culture. Genetics **177**:31–46.

Wong, E. T., J. L. Kolman, Y.-C. Li, L. D. Mesner, W. Hillen, C. Berens, and G. M. Wahl. 2005. Reproducible doxycycline-inducible transgene expression at specific loci generated by Cre-recombinase mediated cassette exchange. Nucleic Acids Res **33**:e147.

Yamaguchi, M., Y. Hayashi, Y. Nishimoto, F. Hirose, and A. Matsukage. 1995. A nucleotide sequence essential for the function of DRE, a common promoter element for Drosophila DNa replication-related genes. J Biol Chem **270**:15808–15814.

Yao, J., K. M. Munson, W. W. Webb, and J. T. Lis. 2006. Dynamics of heat shock factor association with native gene loci in living cells. Nature **442**:1050–3.

Yao, J., K. L. Zobeck, J. T. Lis, and W. W. Webb. 2008. Imaging transcription dynamics at endogenous genes in living Drosophila tissues. Methods **45**:233–41.

Yokoyama, R., A. Pannuti, H. Ling, E. R. Smith, and J. C. Lucchesi. 2007. A plasmid model system shows that Drosophila dosage compensation depends on the global acetylation of histone H4 at lysine 16 and is not affected by depletion of common transcription elongation chromatin marks. Mol Cell Biol .

Zeidler, M. P., C. Tan, Y. Bellaiche, S. Cherry, S. Hader, U. Gayko, and N. Perrimon. 2004. Temperature-sensitive control of protein activity by conditionally splicing inteins. Nat Biotechnol **22**:871–876.

Zhang, W., H. Deng, X. Bao, S. Lerach, J. Girton, J. Johansen, and K. M. Johansen. 2006. The JIL-1 histone H3S10 kinase regulates dimethyl H3K9 modifications and heterochromatic spreading in Drosophila. Development **133**:229–35.

Zhou, S., Y. Yang, M. J. Scott, A. Pannuti, K. C. Fehr, A. Eisen, E. V. Koonin, D. L. Fouts, R. Wrightsman, and J. E. Manning. 1995. Male-specific lethal 2, a dosage compensation gene of Drosophila, undergoes sex-specific regulation and encodes a protein with a RING finger and a metallothionein-like cysteine cluster. EMBO Journal **14**:2884–95.

Zhou, X., J. Symons, R. Hoppes, C. Krueger, C. Berens, W. Hillen, B. Berkhout, and A. T. Das. 2007. Improved single-chain transactivators of the Tet-On gene expression system. BMC Biotechnol **7**:6.

Zink, B., and R. Paro. 1989. In vivo binding pattern of a trans-regulator of homoeotic genes in Drosophila melanogaster. Nature **337**:468–471.

Appendices

Appendix A

Biological material

All fly strains used in this study are listed in Table A.1. All oligonucleotides are in Table A.2. The plasmids used or created are in Table A.3.

Table A.1: Fly strains used in this study

Lab Stock	*Genotype*	Transgene Chromosome	Comments	Reference
#3	$y^1\ w^1$; $msl1^{L60}$/CyO, pr $cn^2\ y^+$			Richard Kelley[a]
#4	*y w*		Injection stock	John Lucchesi[b]
#6	*(Wild-type)*		Canton S strain	Bloomington[c] #1
#8	*y w;* L^2/CyO, Cy pr $cn^2\ y^+$		2nd chromosome balancer	John Lucchesi
#9	y^1 w*;TM3,y^+ Ser^1/Sb^1		3rd chromosome balancer	Bloomington #1614
#17	*w cv* mof^1/FM7, $y^{31d}\ sc^8\ w^a$ *B*		1st chromosome balancer	John Lucchesi
#18	*msl2 msl1 mle/CyO*			Belote, 1983
#8556	w*; P{$w^{+mW.hs}$=arm-GFP.P}57	2		Bloomington #8556
#9546	w^{1118}; P{w^{+mC}=hs-hid}3, Dr^1/TM6C, $cu^1\ Sb^1$		3rd chromosome balancer	Bloomington #9546
#9746	$y^1 w^{1118}$; PBac{y+=attP-9A}VK00029	3	attP heterochromatic	Bloomington #9746
C34	*w* P{w^+=GMR-hid^{Ala5}}	X(18F)		Bergmann et al., 1998
EW89	*CyO/Sp; PBac{DsRed=TRE-eyfp}M6.III*	3	Some heterozygous over TM2	Horn and Wimmer, 2000
EW140	*pBac{EYFP=HS4-s1-tTA-HS4}M3.II; MKRS/TM2*	2	Some heterozygous over CyO	Horn and Wimmer, 2003
EW162	*pBac{EYFP=HS4-n1-tTA-HS4}M4.II; MKRS/TM2*	2		Horn and Wimmer, 2003
F1	*y w* P{w^+=GMR-hid^{Ala5}}13	X		Li et al., 2005
J5	*y w pM?{hs-flp}; pM{3xP3-RFP=attP}86F; pM{3xP3-RFP=*ϕ*C31{3xP3-GFP=vas-*ϕ*C31}}102F*	X, 3(86F), 4(102F)		Bischof et al., 2007
J9	*y w pM?{hs-flp}; pM{3xP3-RFP=*ϕ*C31{3xP3-GFP=vas-*ϕ*C31}}86F*	X, 3(86F)		Bischof et al., 2007
M1	w*; $ry^{506}\ Sb^1$ P{$ry^{+t7.2}$=Δ2-3}99B/TM6B, Tb^1			Bloomington #1798
M18	*w;* P{w^+=yp1-tTA}	3		Heinrich and Scott, 2000
M27	*y w;* P{w^+=arm-tTA}19	Not X		Scott et al., 2004
M32	*y w;* P{w^+=tetO-msl2}5	2		Scott et al., 2004

Lab Stock	*Genotype*	Transgene Chromosome	Comments	Reference
M35	*w^{1118} P{w^+=actin5C-rtTA}901*	X		Bieschke et al., 1998
M48	*y w; $msl2^1$ cn; P{w^+=hsp83-msl1}2I P{w^+=hsp83-msl2}6I/+*	3	Females lack transgenes	Chang and Kuroda, 1998
M51	*w; $msl1^{L60}$/CyO; P{w^+=hsp83-msl2}*	3		Kelley et al., 1995*b*
S13	*y w P{w^+=hsp70-msl2}*	X		Li et al., 2008
V40	*y w P{w^+=arm-lacZ}E3*	X(10D8)		Weake and Scott, 2007
V42	*y w; P{w^+=arm-lacZ}F1*	3(79A4)		Weake and Scott, 2007
V62	*y w; P{pUB-EGFP}B1*	3		A. Handler[d]
zh-attP-2A	*y w P{3xP3-RFP=attP}2A*	X	a.k.a. zh-attP-11	Bischof et al., 2007
zh-attP-3Aa	*y w P{3xP3-RFP=attP}3Aa*	X	a.k.a. zh-attP-56	Bischof et al., 2007
zh-attP-3B	*y w P{3xP3-RFP=attP}3B*	X	a.k.a. zh-attP-46	Bischof et al., 2007
zh-attP-6E	*y w P{3xP3-RFP=attP}6E*	X	a.k.a. zh-attP-13	Bischof et al., 2007
zh-attP-20C	*y w P{3xP3-RFP=attP}20C*	X	a.k.a. zh-attP-17	Bischof et al., 2007

[a]Baylor College of Medicine, Houston

[b]Emory University, Atlanta

[c]Bloomington Drosophila Stock Center, Indiana University

[d]U.S. Department of Agriculture, Gainesville, Florida

Table A.2: Oligonucleotides used in this study

Name	Sequence (5′ to 3′)	Modifications	Tm(°C)	Comments
3DRE4GAGATop5′	AATTGCTCGAGCTAGCTATCGATAGATTCCC-TGCTATCGATAGATTCCCTGCTATCGA	5′ P	79	These four oligonucleotides bind to form artificial DRE/GAGA sites, and cloning sites.
3DRE4GAGATop3′	TAGATTCGCTAGCAGATCTCTCTCGTTCATT-GAGAGAGCAAAGGCCTCTCTCGTTCATTGAG-AGAGATCTCGAG	5′ P	85	
3DRE4GAGABottom5′	AATTCTCGAGATCTCTCTCAATGAACGAGAG-AGGCCTTTGCTCTCTCAATGAACGAGA	5′ P	81	
3DRE4GAGABottom3′	GAGATCTGCTAGCGAATCTATCGATAGCAGG-GAATCTATCGATAGCAGGGAAATCTATCGAT-AGCTAGCTCGAGC	5′ P	83	
armDREGAGAfwd	CCGGAATTCTCGAGTGGAATGTAAACAATGC-CACAGAC		74	Amplifies 317 bp of armadillo promoter.
armDREGAGArev	CCGGAATTCTCGAGTAAACGGAACAGAATCA-CAGATGC		73	
armlacZseq	CAGAGAAGGAGGCAAACAGC			
armlacZNotI	AATAAAGCAATAGCATCACA			
AS5	CTCTGTAGGTAGTTTGTCC			
attPTargetRev	CGGAATCGCACCGAATG		54	Binds between attP and marIR(L) in attP landing site.
corey_seq_roX	AGCGCCTCTATTTATACTCCG		50	Sequences upstream of hsp70 TATA.
GiLfor	GAGCGCCGGAGTATAAATAG		60	Amplifies across ninaE intron in pGLattB. 928 bp unspliced.
GiLrev	ATTGACCGTAATGGGATAGG		58	

Name	Sequence (5′ to 3′)	Modifications	Tm(°C)	Comments
GMRDsRedFseq2	ACTGATAAGAATGTTTCGATCG			
h70p(A)rev	GTCCATCGCAATAAAATGAGCC		55	Binds towards the 3′ end of hsp70 3′UTR.
hsp70TATAfwd	CCAAGTAAATCAACTGCAACTACTG		52	Binds in the hsp70 minimal promoter sequence.
lacZ3′Fwd	GCCGCTATTTCTCTGTTCTCGC		57	Binds at nt 3340 in lacZ ORF.
lacZnewRev	ATCTGCCAGTTTGAGGGGAC		62	Binds at nt 1310 in lacZ ORF.
M2intseq	CAACAACTTTGCTGTCTCCTAC		49	Sequences downstream of nt 807 in msl2 ORF.
pXLsense4013/pr1	GGACGTTCCAGAGCCGTTGACTCAG		64	Amplifies 3′ end of msl2.
HIS/M2/ASENSE/pr2	GTGGTGATGGTGATGCAAGTCATCCGAGCCCG		76	Amplifies 3′ end of msl2, without stop codon, and adding overlapping M2/HIS/SENSE/3 sequences.
lacZ493R	TGTCCTGGCCGTAACCG			
M2/HIS/SENSE/3	GGCTCGGATGACTTGCATCACCATCACCACC-ATCG		78	Amplifies HIS tags, BAD, and BirA from pBBHC, adding overlapping HIS/M2/ASENSE/pr2 sequences.
NotI/BirA/asense/4	CATCGCGGCCGCCTATTTTTCTGCACTACG		73	Amplifies HIS tags, BAD, and BirA from pBBHC, adding a NotI site.
Nhe/Not/HA/M2/asen	CATCGCGGCCGCTTAGCTAGCGGCGTAATCG-GGCACATCGTAGGGGTACAAGTCATCCGAGC-CCGAC		92	Amplifies 3′ end of msl2, without stop codon, and adding an HA tag, NheI and NotI sites.
NheHAXHANotTop	CTAGCTACCCCTACGATGTGCCCGATTACGC-CTCGAGCTACCCCTACGATGTGCCCGATTAC-GCCTAAGC		88	Forms NheI/NotI linker containing 2x HA sites, separated by a XhoI site.
NheHAXHANotBot	GGCCGCTTAGGCGTAATCGGGCACATCGTAG-GGGTAGCTCGAGGCGTAATCGGGCACATCGT-AGGGGTAG		90	
mar3f	TGAGAGTAAACAATTATGACGCT		48	For iPCR of mariner.

Name	Sequence (5′ to 3′)	Modifications	Tm(°C)	Comments
mar3r	CGACTGACGGTCGTAAG		54	
mar3s	AAGGTTGACACTTCACAAGG		58	Sequences from the mariner 3′ terminus into the genome
ninaEin2F	CCGGAATTCGGCTGGAGCAGGTGAGT		69	Amplifies second intron of ninaE, adding EcoRI sites.
ninaEin2R	CCGGAATTCCGGCAGATACCTTTTAGATTAAG		66	
pCL04seq_minus	TCTGAGTGTCGTATTGAGTCTG		47	Sequences from 3′P.
phic31For	GTTCAGGTTCGTCGGGCATTTC		59	Amplifies 550 bp of phic31.
phic31Rev	GGGAAGGTGTTTGTGCGTCTTG		58	
Pry1	CCTTAGCATGTCCGTGGGGTTTGAAT			
rp49_596F	CGGTTACGGATCGAACA		52	Amplifies 125 bp of rp49
rp49_720R	CGATCTCGCCGCAGTAAA		56	
SacI_XhoItop	CTTTAAACGGTTGGCG	5′ P	48	Forms SacI/XhoI linker with SacI_XhoIbot.
SacI_XhoIbot	TCGACGCCAACCGTTTAAAGAGCT	5′ P	62	
SV40RevPstI	TTTCTGCAGGATCCAGACATGATAAG			
white3f	TGATCAGATAAGTTCAATGATATCCAG			Binds in white 3′
white3r	CCAGGTTCTTTCGATTACCTC			Binds in white 3′

Table A.3: Plasmids used or created in this study

Plasmid	Size(bp)	Resistance	Base Vector	Insert(s), Comments	Reference
UChs$\pi\Delta$2,3 (pΔ2-3)	7300	Ampicillin		Expresses *P* transposase	(Rio and Rubin, 1985)
HF10	~10220	Ampicillin			(Fitzsimons et al., 1999)
W.T.P-2	8848	Ampicillin			(Bello et al., 1998)
XL1	5306	Ampicillin			Xueli Li[c]
BBHC	~9000	Ampicillin		Source of BAD and BirA	(Viens et al., 2004)
GMR-1	8851	Ampicillin			(Hay et al., 1994)
BS SK hid cDNA	~6890	Ampicillin		Source of hid cDNA. A.K.A. cDNA 5A1B.	(Grether et al., 1995)
CaSpeR-3-18D-monomer	~8300	Ampicillin		Source of 18D monomer.	(Oh et al., 2004)
CaSpeR-3-18D-L4mer	~10000	Ampicillin		Source of 18D multimer.	(Oh et al., 2004)
TAattB	4222	Ampicillin, Kanamycin		Source of attB.	(Groth et al., 2000)
UASTattB	8489	Ampicillin			(Bischof et al., 2007)
CaSpeR-arm-βgal	13874	Ampicillin			(Vincent et al., 1994)
BS2N17mer HF12-1x12	9649	Ampicillin		Source of SV40 p(A) and embedded 18D10.	A. Schiemann[c]
BC-EGFP	4096	Chloramphenicol		Source of EGFP.	C. Concha[c]
RGH	13020	Ampicillin		*roX1 DHS-GMR-hid* in a pCaSpeR derivative.	Laverty 2003
R9GH	14975	Ampicillin		*(roX1 DHS)9-GMR-hid* in a pCaSpeR derivative.	Laverty 2003
CL04	10183	Ampicillin	HF10	None; have destroyed the (only) XbaI site.	This Study
STPS′	11298	Ampicillin	CL05	XhoI*[a]-HindIII* tetO, 5′ P TATA, and hsp70 p(A) from pW.T.P-2 into EcoRI*.	This Study
M2HA	5289	Ampicillin	XL1	NarI-NotI msl-2, HA tag from PCR "1,5" amplicon (from pXL1) into NarI-NotI.	This Study

Plasmid	Size(bp)	Resistance	Base Vector	Insert(s), Comments	Reference
M2HisB	6359	Ampicillin	XL1	NarI-NotI msl-2, HIS tag, BAD, and BirA from PCR "1,4" amplicon (from pBBHC) into NarI-NotI.	This Study
TM2HA	13644	Ampicillin	STPS′	EcoRI-NotI msl-2, HA tag from pM2HA into EcoRI-NotI.	This Study
TM2HA3	13701	Ampicillin	TM2HA	NheI-NotI 2xHA tag from annealed "NheHAXHANotI" oligonucleotides into NheI-NotI.	This Study
TM2HisB	14714	Ampicillin	STPS′	EcoRI-NotI msl-2, HIS tag, BAD, and BirA from pM2HisB into EcoRI-NotI.	This Study
TM2HA3attB	11081	Ampicillin	UASTattB	EcoRV-NotI tetO, msl-2 and HA tags from pTM2HA3 into HindIII*-NotI.	This Study
GMRhid	12763	Ampicillin	GMR-1	EcoRI-EcoRI hid cDNA from pBS SK hid cDNA into EcoRI. A.K.A. pGH	This Study
18DGH	13328	Ampicillin	GMRhid	BamHI-NotI* 18D10 monomer from pCaSpeR3 18D10-L monomer into BglII-StuI.	This Study
18D4GH	14966	Ampicillin	GMRhid	BamHI-NotI* 18D10 4-mer from pCaSpeR3 18D10-L 4mer into BglII-StuI.	This Study
BSattB	3255	Ampicillin	Bluescript II KS+	EcoRI*-EcoRI* attB from pTAattB into EcoRV-SmaI.	This Study
BSattB-X	3258	Ampicillin	BSattB	None; have destroyed the (only) XhoI site.	This Study
BClacZ	6856	Chloramphenicol	BC KS+	SmaI-XbaI* lacZ from pCaSpeR-arm-βgal into SmaI.	This Study
BSw+GMRattB	9072	Ampicillin	BSattB-X	HindIII-HindIII mini-white, GMR, hsp70 TATA and p(A) from pGMR-1 into HindIII.	This Study
ALattB	14183	Ampicillin	CaSpeR-arm-βgal	EcoRI-EcoRI attB site from pTAattB into EcoRI.	This Study
GHattB	12980	Ampicillin	BSw+GMRattB	EcoRI-EcoRI hid cDNA from pBS SK hid cDNA into EcoRI.	This Study
GLattB	12526	Ampicillin	BSw+GMRattB	BamHI-EcoRI* lacZ from pBClacZ into BglII-StuI.	This Study

Plasmid	Size(bp)	Resistance	Base Vector	Insert(s), Comments	Reference
GLattB-H	12530	Ampicillin	GLattB	None; have destroyed the HindIII site at 2240 of pGLattB.	This Study
GLattB-RI	12530	Ampicillin	GLattB	None; have destroyed the EcoRI site at 10972 of pGLattB.	This Study
GLSV40attB	12848	Ampicillin	GLattB-H	SmaI-HindIII lacZ, SV40 p(A) from pCaSpeR-arm-βgal into EcoRI*-HindIII.	This Study
CL12	12853	Ampicillin	GLattB	XhoI-XhoI armadillo promoter PCR amplicon (from pCaSpeR-arm-βgal) into XhoI.	This Study
CL13	12647	Ampicillin	GLattB	XhoI-XhoI DRE/GAGA annealed oligonucleotides into XhoI.	This Study
CL14	12867	Ampicillin	GLattB-RI	EcoRI-EcoRI armadillo promoter PCR amplicon (from pCaSpeR-arm-βgal) into EcoRI. β-Galactosidase inactive.	This Study
CL15	12662	Ampicillin	GLattB-RI	(MfeI)-EcoRI DRE/GAGA annealed oligonucleotides into EcoRI. Clone has an incorrect BglII site.	This Study
GiLattB	12745	Ampicillin	GLattB-RI	EcoRI-EcoRI ninaE intron 2 PCR amplicon (from Canton-S genomic DNA) into EcoRI. Two MfeI sites (at 6259 and 7628).	This Study
GLSV4018DattB	13400	Ampicillin	GLattB-H	SmaI-SmaI lacZ, SV40 p(A) with 18D10 monomer from pBS2N17mer HF12-1x12 into EcoRI*-HindIII.	This Study
GEG	9794	Ampicillin	BSw+GMRattB	BamHI-HincII* EGFP from pBC-EGFP into BglII-StuI. Originally called “pBSw+GMR ZsGreen attB” or “GZattB” when thought was cloning ZsGreen.	This Study
GiL-MfeI	12752	Ampicillin	GiLattB	None; have destroyed the MfeI site at 6259 of pGiLattB.	This Study
CL19	9915	Ampicillin	GEG	XhoI-XhoI DRE/GAGA annealed oligonucleotides into XhoI.	This Study

Plasmid	Size(bp)	Resistance	Base Vector	Insert(s), Comments	Reference
CL21	12846	Ampicillin	GLattB	XhoI-SacI tetO from pW.T.P-2, and SacI/XhoI linkers into XhoI. Linkers 3′ to tetO. 3′ XhoI site not re-made. A.K.A. ptetOGL.	This Study
CL24	9196	Ampicillin	BSw+GMRattB	XhoI-XhoI DRE/GAGA annealed oligonucleotides into XhoI. Clone V(injected) has an incorrect NheI site; E an incorrect BglII site.	This Study
CL25	13105	Ampicillin	CL24	EcoRI-EcoRI hid cDNA from pBS SK hid cDNA into EcoRI.	This Study

[a]An asterix (*) denotes that a cleavage site was made blunt with Klenow fragment

[b]Brackets () around a restriction site indicate that it was not re-constituted upon ligation

[c]Massey University, Palmerston North, New Zealand

Appendix B

Raw data for effect of msl2 expression on female viability

Raw data is provided for the comparisons of female viability (Figure 3.3). The proportion (p) of females per total fly count was used for the graph and statistical analyses.

Number of flies						
Female	Male	Total	p(female)	Line	'Treatment'	Replicate
233	249	482	0.483	*y w*	x self	1
222	159	381	0.583	*y w*	x self	2
131	125	256	0.512	*y w*	x self	3
50	60	110	0.455	*y w*	x M27 + tet 10 μg/mL	1
48	39	87	0.552	*y w*	x M27 + tet 10 μg/mL	2
46	40	86	0.535	*y w*	x M27 + tet 10 μg/mL	3
41	52	93	0.441	*y w*	x M27 + tet 10 μg/mL	4
123	108	231	0.532	*y w*	x M27 - tet	1
155	158	313	0.495	*y w*	x M27 - tet	2
176	144	320	0.550	*y w*	x M27 - tet	3
181	168	349	0.519	*y w*	x M27 - tet	4
170	170	340	0.500	C48	x self	1
197	219	416	0.474	C48	x self	2
120	139	259	0.463	C48	x self	3
48	39	87	0.552	C48	x M27 + tet 10 μg/mL	1
61	44	105	0.581	C48	x M27 + tet 10 μg/mL	2
180	104	284	0.634	C48	x M27 + tet 10 μg/mL	3
20	16	36	0.556	C48	x M27 + tet 10 μg/mL	4

Number of flies						
Female	Male	Total	p(female)	Line	'Treatment'	Replicate
374	472	846	0.442	C48	x M27 + tet 1 μg/mL	1
238	378	616	0.386	C48	x M27 + tet 1 μg/mL	2
300	379	679	0.442	C48	x M27 + tet 1 μg/mL	3
66	121	187	0.353	C48	x M27 - tet	1
44	97	141	0.312	C48	x M27 - tet	2
39	90	129	0.302	C48	x M27 - tet	3
36	70	106	0.340	C48	x M27 - tet	4
152	156	308	0.494	C50	x self	1
210	156	366	0.574	C50	x self	2
86	98	184	0.467	C50	x self	3
262	347	609	0.430	C50	x M27 + tet 10 μg/mL	1
416	371	787	0.529	C50	x M27 + tet 10 μg/mL	2
446	388	834	0.535	C50	x M27 + tet 10 μg/mL	3
323	475	798	0.405	C50	x M27 + tet 1 μg/mL	1
214	337	551	0.388	C50	x M27 + tet 1 μg/mL	2
364	473	837	0.435	C50	x M27 + tet 1 μg/mL	3
78	126	204	0.382	C50	x M27 - tet	1
78	110	188	0.415	C50	x M27 - tet	2
100	138	238	0.420	C50	x M27 - tet	3
98	144	242	0.405	C50	x M27 - tet	4
133	126	259	0.514	C93	x self	1
123	108	231	0.532	C93	x self	2
138	112	250	0.552	C93	x self	3
139	127	266	0.523	C93	x self	4
343	420	763	0.450	C93	x M27 + tet 10 μg/mL	1
381	363	744	0.512	C93	x M27 + tet 10 μg/mL	2
363	377	740	0.491	C93	x M27 + tet 10 μg/mL	3
308	404	712	0.433	C93	x M27 + tet 1 μg/mL	1
318	427	745	0.427	C93	x M27 + tet 1 μg/mL	2
317	390	707	0.448	C93	x M27 + tet 1 μg/mL	3
105	153	258	0.407	C93	x M27 - tet	1
91	154	245	0.371	C93	x M27 - tet	2
101	147	248	0.407	C93	x M27 - tet	3
79	139	218	0.362	C93	x M27 - tet	4

Appendix C

Raw data for effect of altered MSL levels on GMR-hid lines C60 and GMR-hid^{Ala5}

The following table lists absolute fly numbers for crosses to compare response of select X-linked lines to altered levels of MSL proteins (Figure 4.5b).

Table C.1: *GMR-hid* eye sizes with altered MSL levels

				Eye size					
Line	Genotype	Sex	Rep	Ablated	Slit	Sm Int	Med Int	L Int	Large
C60	*CyO*/+	M	1	23	-	-	-	-	-
			2	21	-	-	-	-	-
		F	1	-	-	92	-	-	-
			2	-	-	89	-	-	-
	msl2 msl1 mle/+	M	1	34	-	-	-	-	-
			2	38	-	-	-	-	-
		F	1	-	-	87	-	-	-
			2	-	-	102	-	-	-
	CyO/+	M	1	31	2	-	-	-	-
		F	1	-	-	83	-	-	-
	CyO/+; *hsp83-msl2*/+	M	1	-	17	-	-	-	-
		F	1	-	-	6	-	-	-
Ala5	*CyO*/+	M	1	34	-	-	-	-	-
			2	61	-	-	-	-	-
			3	59	-	-	-	-	-
		F	1	-	-	41	-	-	-
			2	-	-	87	-	-	-
			3	-	-	95	-	-	-
	msl2 msl1 mle/+	M	1	32	-	-	-	-	-
			2	45	-	-	-	-	-
			3	49	-	-	-	-	-
		F	1	-	-	70	-	-	-
			2	-	-	77	-	-	-
			3	-	-	91	-	-	-
	CyO/+	M	1	-	111	-	-	-	-
			2	-	100	-	-	-	-
		F	1	-	-	157	-	-	-
			2	-	-	140	-	-	-
	CyO/+; *hsp83-msl2*/+	M	1	-	142	-	13	-	-
			2	-	150	-	1	-	-
		F	1	-	-	-	35	-	-
			2	-	-	-	19	-	1

Appendix D

Raw data for GMR-hid-attB at X-linked attP sites

Table D.1: Flies per category eye size

				Eye Size					
Location	Line	Sex	Replicate	Ablated	Slit	Sm Int	Med Int	L Int	Total
2A	1(C70)	Male	1	0	77	1	0	0	78
			2	2	103	0	0	0	105
			3	1	79	2	0	0	82
		Female	1	0	12	123	3	0	138
			2	0	10	157	7	0	174
			3	0	15	136	2	0	153
	2(C71)	Male	1	0	92	2	0	0	94
			2	0	86	1	0	0	87
			3	2	100	1	0	0	103
		Female	1	0	13	164	7	0	184
			2	0	12	136	7	2	157
			3	0	6	145	9	0	160
	3^{1}	Male	1	0	126	0	0	0	126
			2	0	120	0	0	0	120
			3	0	151	1	0	0	152
		Female	1	0	3	164	7	0	174
			2	0	21	178	3	0	202
			3	0	43	244	19	0	306
6E	$1(C72)^{1}$	Male	1	0	90	3	0	0	93

				Eye Size					
Location	Line	Sex	Replicate	Ablated	Slit	Sm Int	Med Int	L Int	Total
			2	0	103	0	0	0	103
			3	0	111	4	0	0	115
		Female	1	0	16	115	5	0	136
			2	0	13	134	2	0	149
			3	0	11	142	5	0	158
	2(C73)	Male	1	0	104	3	0	0	107
			2	0	85	2	0	0	87
			3	0	83	1	0	0	84
		Female	1	0	8	138	7	0	153
			2	0	17	131	4	0	152
			3	0	14	126	4	0	144
20C	2(C74)	Male	1	0	82	1	2	0	85
			2	0	54	1	0	0	55
			3	1	65	3	0	0	69
		Female	1	0	19	134	20	0	173
			2	0	18	141	11	0	170
			3	0	24	138	17	0	179
	3[1]	Male	1	0	109	3	0	0	112
			2	0	130	2	0	0	132
			3	2	165	4	0	0	171
		Female	1	0	11	117	2	0	130
			2	0	18	135	8	0	161
			3	0	13	182	7	0	202
	4(C75)	Male	1	0	57	2	0	0	59
			2	0	44	2	0	0	46
			3	0	42	1	0	0	43
		Female	1	0	12	60	4	0	76
			2	0	8	58	3	1	70
			3	0	8	45	1	0	54

[1] Used for Figure 4.8c

Appendix E

Genomic sequences from inverse PCR of attP sites

The following sequences are the genomic portion of sequencing reads, immediately following *mariner* inverted repeat sequences of the landing site.

```
Sequence Range: 80 to 158

 80        90       100
  WATATATTTCTTCGCGGTAAC
  WTATATAAAGAAGCGCCATTG

         110       120       130       140       150
  CTCTGCCCATTTYCSGATCARATAAGTTCAATGATATCCAGTGCAGTAAA
  GAGACGGGTAAARGSCTAGTYTATTCAAGTTACTATAGGTCACGTCATTT

  AAAAAAAA
  TTTTTTTT
```

Figure E.1: zh-attP-2A

```
Sequence Range: 72 to 141

         80        90       100
  TATATCTCCGCCTGTCTAAGCCTTAAGTC
  ATATAGAGGCGGACAGATTCGGAATTCAG

         110       120       130       140
  AGCTCATTTCCTAAATGCTCACTAATCATCCAAACTACCGG
  TCGAGTAAAGGATTTACGAGTGATTAGTAGGTTTGATGGCC
```

Figure E.2: zh-attP-20C

```
Sequence Range: 77 to 571

   80        90       100
  KATATTGTTGCTTCGGATTTTCAC
  MTATAACAACGAAGCCTAAAAGTG

         110       120       130       140       150
  ATTTTATTTTTTYCCGAKTTTCAYYKYMKCTYYWACGGACCAAAAGACTT
  TAAAATAAAAAARGGCTMAAAGTRRMRKMGARRWTGCCTGGTTTTCTGAA

         160       170       180       190       200
  GGGATATTTATTTAGTGCGCATTTGTTGTTTYKYTGTTACRATTGTTGCT
  CCCTATAAATAAATCACGCGTAAACAACAAARMRACAATGYTAACAACGA

         210       220       230       240       250
  AACTTTTATGACCCATGTTCGCCGCATTTTCSTGGCTCCTAACATACGGT
  TTGAAAATACTGGGTACAAGCGGCGTAAAAGSACCGAGGATTGTATGCCA

         260       270       280       290       300
  TCGTTCTTATTTYGTATTTCAATGTTTTTATGGCCTCTTGTTTTGACATT
  AGCAAGAATAAARCATAAAGTTACAAAAATACCGGAGAACAAAACTGTAA

         310       320       330       340       350
  CMTGCMTATTTTGTAATAATCTTAAACTGAKTAATAACCCCTTGTGCCTT
  GKACGKATAAAACATTATTAGAATTTGACTMATTATTGGGGAACACGGAA

         360       370       380       390       400
  CYCTTGGTTATRGCATTTTATGGWACCACGMGWGTCACGGCRCYYCTTWT
  GRGAACCAATAYCGTAAAATACCWTGGTGCKCWCAGTGCCGYGRRGAAWA

         410       420       430       440       450
  CCATCASAATTTTATATGGYGATYTKTWTGGCTAAAGAMATACCACAATG
  GGTAGTSTTAAAATATACCRCTARAMAWACCGATTTCTKTATGGTGTTAC

         460       470       480       490       500
  TGTWKYGTTTTCAYTAWATCTATTTTATTTGACAGCCCMATAATGGGAAA
  ACAWMRCAAAAGTRATWTAGATAAAATAAACTGTCGGGKTATTACCCTTT

         510       520       530       540       550
  TGCTTATACACACATATGTACATACATCTTGWGTCTWCAATTTGCTTACA
  ACGAATATGTGTGTATACATGTATGTAGAACWCAGAWGTTAAACGAATGT

         560       570
  TGTGTTCTTCWTCTATTTTCA
  ACACAAGAAGWAGATAAAAGT
```

Figure E.3: zh-attP-3Aa

```
Sequence Range: 77 to 620

   80        90       100
  KAATCGTTGGTCACCTCCTGGAAT
  MTTAGCAACCAGTGGAGGACCTTA

         110       120       130       140       150
  TCCTGTTCSCYTGGTAATTTRAGCCGAAAATCTCTAAATTTTGCTTACAA
  AGGACAAGSGRACCATTAAAYTCGGCTTTTAGAGATTTAAAACGAATGTT

         160       170       180       190       200
  AATGAGCATTTTAGTGGCAAATTTTGGGATGTTTKGTACCAGGCCATCAT
  TTACTCGTAAAATCACCGTTTAAAACCCTACAAAMCATGGTCCGGTAGTA

         210       220       230       240       250
  ACTCAGTAGCGGGTGTCTACTGAAAACCAACTAATCGTTGGTCGCCTCCT
  TGAGTCATCGCCCACAGATGACTTTTGGTTGATTAGCAACCAGCGGAGGA

         260       270       280       290       300
  GGAATTCCTGTTCGCCTGGTAATTTAAGCCGAAAATCTCTAAATTTTGCT
  CCTTAAGGACAAGCGGACCATTAAATTCGGCTTTTAGAGATTTAAAACGA

         310       320       330       340       350
  TACAAAATGAGCATTTTAGTGGCAAATTTTGGATTTTTGGTACCAGGCCA
  ATGTTTTACTCGTAAAATCACCGTTTAAAACCTAAAAACCATGGTCCGGT

         360       370       380       390       400
  TCATACTCAGTAGCGGGTGTCTACTGAAAACCAACTAATCGTTGGTCGCC
  AGTATGAGTCATCGCCCACAGATGACTTTTGGTTGATTAGCAACCAGCGG

         410       420       430       440       450
  TCCTGGAATTCCTGTTCGCCTGGTAATTTATGCCGAAAATCTGTTAATTT
  AGGACCTTAAGGACAAGCGGACCATTAAATACGGCTTTTAGACAATTAAA

         460       470       480       490       500
  TGCTTACAAAATGAGCTTTTTTGTGGCAAATTTTGGGATGTTTGGTACCA
  ACGAATGTTTTACTCGAAAAAACACCGTTTAAAACCCTACAAACCATGGT

         510       520       530       540       550
  GGCCATCATACTCAGTAGCGGGTGTCTACTGAAAACCAACTAATCGTTGG
  CCGGTAGTATGAGTCATCGCCCACAGATGACTTTTGGTTGATTAGCAACC

         560       570       580       590       600
  TCGCCTCCTGGAATTCCTGTTCGCCTGGTAATTAAAACCGAATACTCACA
  AGCGGAGGACCTTAAGGACAAGCGGACCATTAATTTTGGCTTATGAGTGT

         610       620
  CAATGTGCCTTAAAAACAAT
  GTTACACGGAATTTTTGTTA
```

Figure E.4: zh-attP-6E

Appendix F

Quantification of band intensitiies

Images are provided (Figure F.1) to indicate the ares of interest for the quantification of the band intensities of the gel in Fgure 5.6. The calculations following from those areas are in Table F.1.

Table F.1: Calculations of band intensities for RT-PCR gel

Lane	Band	Area[a]	Length(kb)	Area/kb	Molar freq.	Freq. of 3 vs. 1
2	1	560.1	0.977	573.33	0.796	0.177
	2	38.8	0.851	45.57	0.063	
	3	78.9	0.776	101.67	**0.141**	
4	1	206.3	0.977	211.17	0.131	0.490
	2	122.0	0.851	143.39	0.089	
	3	80.3	0.776	103.50	**0.064**	
	4	336.7	0.575	585.53	0.364	
	5	215.3	0.380	566.67	0.352	
6	1	813.1	0.977	832.28	0.802	0.163
	2	59.5	0.851	69.90	0.067	
	3	105.0	0.776	135.34	**0.130**	

[a] Area under intensity plots corresponding to band peaks

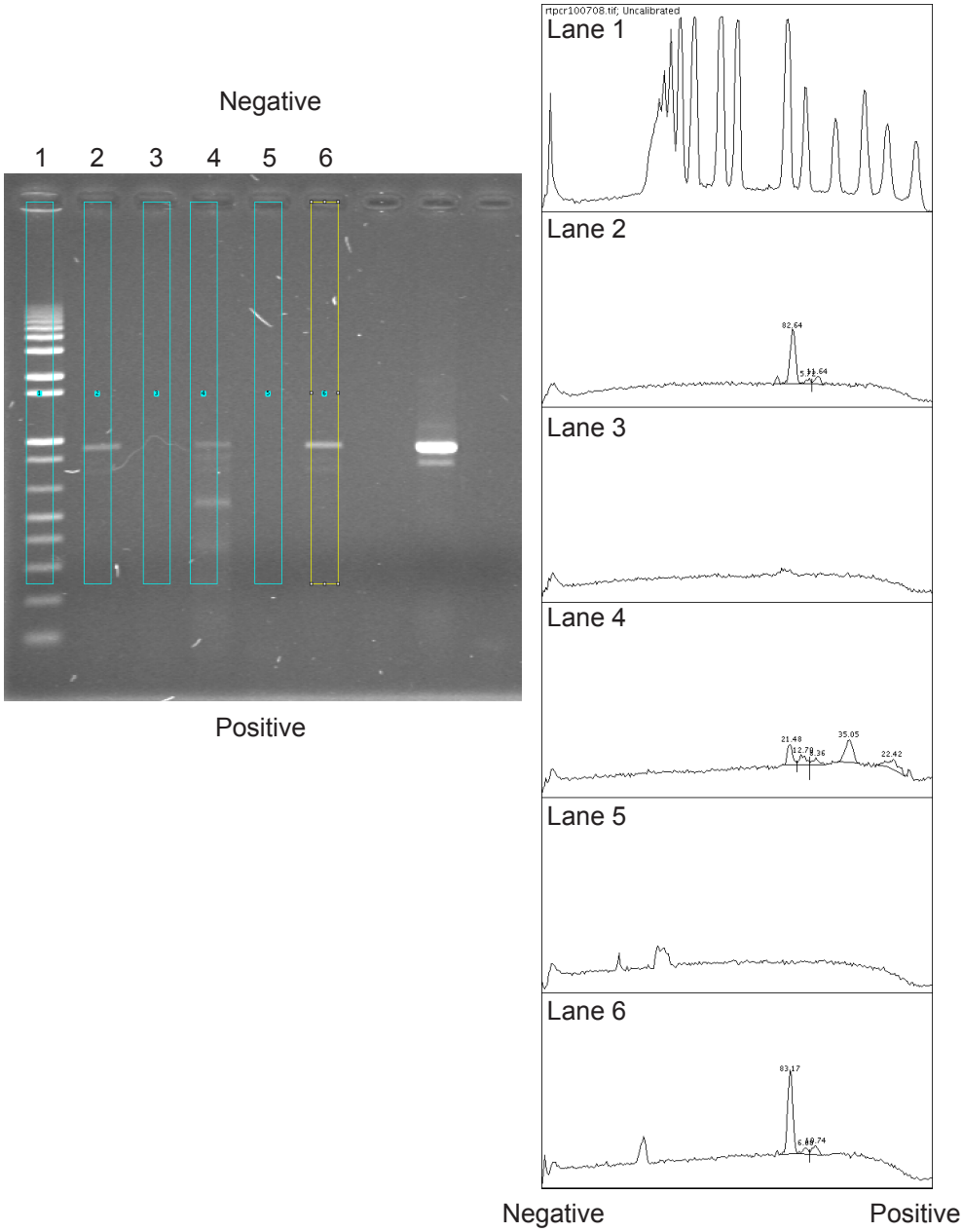


Figure F.1: Intensity plots of the RT-PCR gel. The indicated areas of each lane (left panel) were used to generate the cross-sectional intensity plots (right panel). Areas were selected to avoid dust spots in the regions of interest. The areas inside the indicated peaks were used to quantify the relative intensities of the bands in lanes 2, 4, and 6.

Appendix G

Drosophila Transformation data

The relevant data are presented from micro-injection of *Drosophila* embryos, grouped by transformation type. Number of surviving larvae and adults are listed, along with the total fertile adults and frequency of transgenesis (number of transgenic lines per fertile G_0 flies). Gaps indicate no records available. Some experiments are not discussed in this report.

Table G.1: ϕC31 recombinase transformations

					G_0 Adults			
Date	Plasmid	attP	Embryos	Larvae	Total	Fertile	Transgenic	Freq.[1]
13 Nov 06	ALattB	2A	290	54	22[2]	14	3	0.21
14 Nov 06	ALattB	6E	138	46	18[2]	17	2	0.12
15 Nov 06	ALattB	20C	219	15	6	5	0	0.00
16 Nov 06	ALattB	3B	142	23	12	11	0	0.00
17 Nov 06	ALattB	3A	281	88	59	48	6	0.13
4 Nov 06	ALattB	20C	203	21	15	10	1	0.10
5 Nov 06	ALattB	2A	145	73	48	23	8	0.34
26 Feb 07	GHattB	2A	194	13	5	2	0	0.00
28 Feb 07	GHattB	6E	223	44	16	9	3	0.33
12 Mar 07	GLattB	6E	207	55	40	25	6	0.24
14 Mar 07	GLattB	2A	172	28	14	11	2	0.18
	GHattB	2A	193	52	24	17	5	0.29
15 Mar 07	GHattB	20C	177	35	18	11	4	0.36
16 Mar 03	GLattB	20C	207	54	29	22	5	0.23
28 Mar 03	GHattB	2A	160	48	35			
	GLattB	2A	173	51	44	39	5	0.13
01 Aug 07	ALattB	20C	323	53	39	30	1	0.03

					G_0 Adults			
Date	Plasmid	attP	Embryos	Larvae	Total	Fertile	Transgenic	Freq.[1]
02 Aug 07	ALattB	3B	74	5	1	0	0	0.00
03 Aug 07	GLSV40	2A	297	94	59	38	5	0.13
06 Aug 07	CL13	2A	284	131	76	54	3	0.06
07 Aug 07	CL14	2A	217	69	29	17	1	0.06
09 Aug 07	CL12	2A	312	168	118	98	14	0.14
13 Aug 07	CL15	2A	250	110	74	55	8	0.15
14 Aug07	ALattB	3B	201	63	25	13	0	0.00
19 Dec 07	TM2HA3attB	20C	264	95	40	33	10	0.30
14 Jan 08	TM2HA3attB	'3L'	337	67	48	43	0	0.00
30 Jan 08	GiL	2A	170	115	75	62	14	0.22
13 Feb 08	CL13	6E	271	67	45	40	3	0.08
22 Feb 08	CL19	2A	164	62	25	19	2	0.10
26 Feb 08	CL25	2A	172	48	22	19	1	0.05
27 Feb 08	CL21	2A	173	83	51	41	4	0.09
03 Mar 08	CL25	2A	85	40	32	27	4	0.14
Aug 08[3]	GiL	86F		4	2	2	1	0.50
Sep 08[3]	GiL	6E	62	46	30		>4	
15 Sep 08[3]	GiL	86F	53	22	14	13	0	0.00
23 Sep 08[3]	CL13	86F	29	19	9	8	0	0.00
	GLattB	86F	32	20	14			
23 Oct 08[3]	CL13	86F	46	25	24	21	3	0.14
28 Oct 08[3]	GLattB	86F	36	25	20	17	1	0.06
26 Apr 09[3]	pgdTATA[4]	2A	89	56	50	46	13	0.28
08 May 09[3,4]	CL13	20C	99	46	31	28	>3	

[1]Frequency of transgenic flies per fertile G_0

[2]More larvae survived but were accidentally killed or lost

[3]Injected by Esther Belikoff

[4]Not included in this report

Table G.2: *P* element transformations

				G_0 Adults			
Date	Plasmid	Embryos	Larvae	Total	Fertile	Transgenic	Freq.[1]
Jul 04	R9GH			55	50	5	0.10
Jan 05	RGH			23	20	4	0.20
Apr 05	tetO-msl2-HisB			47	33	9	0.27
	tetO-msl2-HA			9	4	2	0.50
Jun 05	tetO-msl2-HisB	312	120	73	56	3	0.05
22 Jun 05	18D4GH	145	65	20	16	2	0.13
23 Jun 05	RcGH[2]	224	102	36	30	0	0.00
24 Jun 05	18DGH	193	98	53	43	7	0.16
Jun 05	A2[2]	303	128	59	44	2	0.05
28 Jun 05	A1[2]	200	99	57	48	0	0.00
	A4[2]	269	107	58	56	7	0.13
Sep 05	RcGH[2]	321	185	75	63	2	0.03
22 Sep 05	A1[2]	64	48	15	11	0	0.00
18 Oct 05	A1[2]	211	75	8	5	0	0.00
9 Nov 05	GMR-hid	308	93	32	23	1	0.04
14 Nov 05	GMR-hid	251	95	24	20	3	0.15
14 Nov 05[3]	GMR-hid			57	57	7	0.12

[1] Frequency of transgenic flies per fertile G_0
[2] Not included in this report
[3] Injected by Esther Belikoff

Appendix H

Transgenic Stocks

Relevant transgenic stocks are listed below, with the internal lab reference. All *P* element transposon-based insertions were injected in *y w* (lab stock #4) flies (Table 2.1), and the transformation marker was w^+. Thus the genotype can be inferred from the indicated chromosome position in the table below, *e.g.* strain C46 has the genotype: *y w; P{*w^+*=tetO-msl2-HA}2*. Insertions with the ϕC31 recombinase also used the w^+ transformation marker, and their location is at the *attP* site of the indicated injection strain or cross (as in Table A.1). As the ϕC31 recombinase was removed by selective breeding against the *3xP3-GFP* marker (Section 2.2.5), the genotype for *e.g.* strain C76 would be: *y w P{3xP3-RFP=*ϕ*C31{*w^+*=GMR-lacZ}}2A*.

Table H.1: Transgenic *Drosophila* stocks created in this study

Plasmid	Alias	Line	Location	Base Strain	Lab#
tetO-msl2-HA	HA	1	3	#4	C45
		2	3	#4	C46
tetO-msl2-HA3	HA3	1	X	#4	C47
		2	X	#4	C48
		3	X	#4	C49
		4	2	#4	C50
tetO-msl2-HisB	HisB	1	3	#4	C51
		2	3	#4	
		3	X	#4	C53
		4	X	#4	
		5	3	#4	
		6	2	#4	C56
		7	X	#4	C57
		8	3	#4	C58

Plasmid	Alias	Line	Location	Base Strain	Lab#
		9	2	#4	
tetO-msl2-HA3-attB	63HA3	1	X(20C)	C63	C93
RGH		1^{1}	2(32A)	#4	C4
		2	3	#4	C12
		3	3	#4	C14
		5	2	#4	C15
		6	3	#4	C16
R9GH		1^{1}	2	#4	C5
		3	X(12C)	#4	C17
		4	3	#4	C18
		5	2	#4	C19
		8	3	#4	C20
		9	3	#4	C21
		13	2	#4	C22
18DGH	18D	1	3(64B)	#4	C25
		2	2(42B)	#4	C26
		3	3	#4	C27
		4	2(36F)	#4	C28
		5	3	#4	C29
		6	3(62D)	#4	C30
		7	2(22A)	#4	C31
18D4GH	18D4	1	3(85C)	#4	C32
		2	3(95C)	#4	C33
GMR-hid	GH	1	X	#4	C35
		2	3	#4	C36
		3	2	#4	C37
		4	3	#4	C38
		5	X(19E)	#4	C39
		7	X	#4	C60
		8	X(18A)	#4	C40
		9	2	#4	C41
		10	3	#4	C42
		11	3	#4	C43
ALattB	zh11	2	X(2A)	J9 x zh-attP-2A	C66
	zh13	1	X(6E)	J9 x zh-attP-6E	C67
	zh17	1	X(20C)	J9 x zh-attP-20C	C68

Plasmid	Alias	Line	Location	Base Strain	Lab#
GHattB	11GH	1	X(2A)	J9 x zh-attP-2A	C70
		2	X(2A)	J9 x zh-attP-2A	C71
		3	X(2A)	J9 x zh-attP-2A	
	13GH	1	X(6E)	J9 x zh-attP-6E	C72
		2	X(6E)	J9 x zh-attP-6E	C73
	17GH	2	X(20C)	J9 x zh-attP-20C	C74
		3	X(20C)	J9 x zh-attP-20C	
		4	X(20C)	J9 x zh-attP-20C	C75
GLattB	11GL	1	X(2A)	J9 x zh-attP-2A	C76
		3	X(2A)	J9 x zh-attP-2A	
		4	X(2A)	J9 x zh-attP-2A	
	13GL	4	X(6E)	J9 x zh-attP-6E	C78
		6	X(6E)	J9 x zh-attP-6E	C79
	17GL	4	X(20C)	J9 x zh-attP-20C	C81
	-	6	3(86F)	J5	C112
GLSV40attB	SV	3	X(2A)	C61	C91
GLSV4018DattB	GL18D	(1)	X(2A)	C61	C95
CL12		3	X(2A)	C61	C83
DRE-GAGA-GL	CL13	1	X(2A)	C61	C84
		2	X(2A)	C61	C85
		3	X(2A)	C61	C86
	13-13	2	X(6E)	J9 x zh-attP-6E	C98
	-	19	3(86F)	J5	C113
CL14		1	X(2A)	C61	C87
CL15		2	X(2A)	C61	C89
GiL		1	X(2A)	C61	C96
		10	X(6E)	C63	C109
		b	3(86F)	J5	C111
CL25		2	X(2A)	C61	C107
GMR-EGFP	GEG	1	X(2A)	C61	C103
CL19		1	X(2A)	C61	C105
tetO-GMR-lacZ	CL21	1	X(2A)	C61	C100

[1]Created in (Laverty, 2003).

Appendix I

Plasmid maps

Relevant plasmid maps follow, including those for all plasmids made in this report. For descriptions, see Table A.3. Restriction endonuclease sites are displayed for: *Afl* II, *Bam* HI, *Bgl* II, *Eco* RI, *Eco* RV, *Hin* dIII, *Kpn* I, *Mfe* I, *Nhe* I, *Not* I, *Psp* OMI, *Sac* I, *Sma* I, *Spe* I, *Stu* I, *Xba* I, and *Xho* I. Some plasmids were rotated to space out restriction sites, else nucleotide 1 is at the top. Maps were generated with MacVector 10.0 software. Note that an extra *Eco* RV restriction site was detected in pGMR-1 beyond that pictured, which probably remains in all derived plasmids. The extra site maps to about 4700 of pGMR-1, between the *Afl* II and *Xho* I sites.

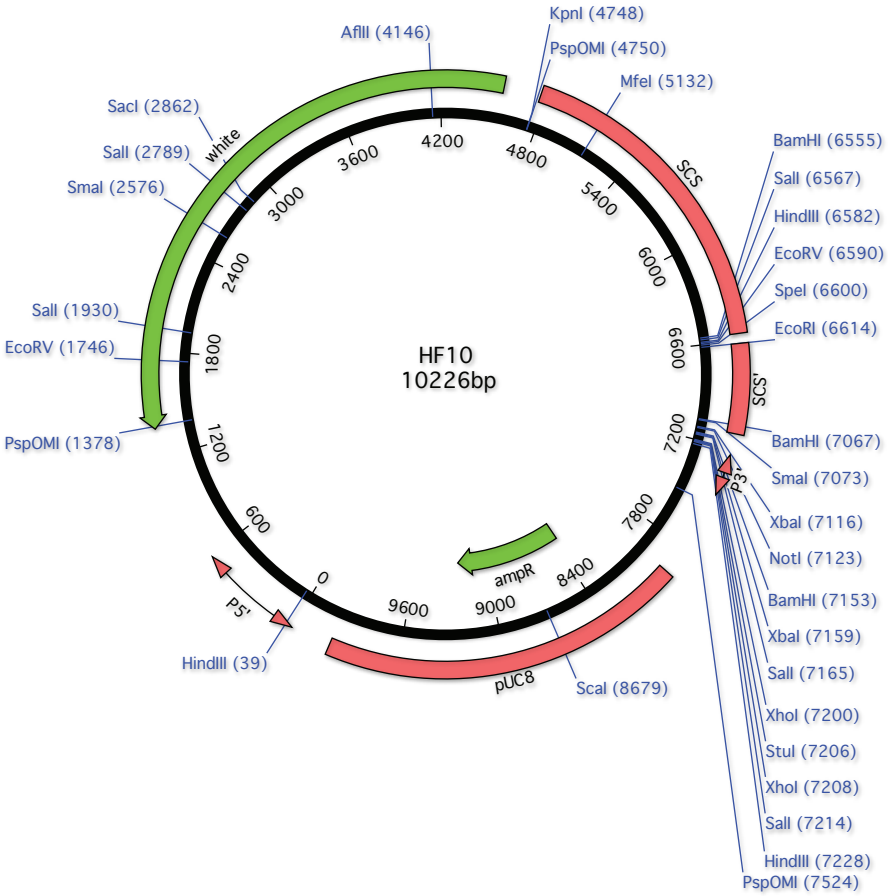


Figure I.1: pHF10

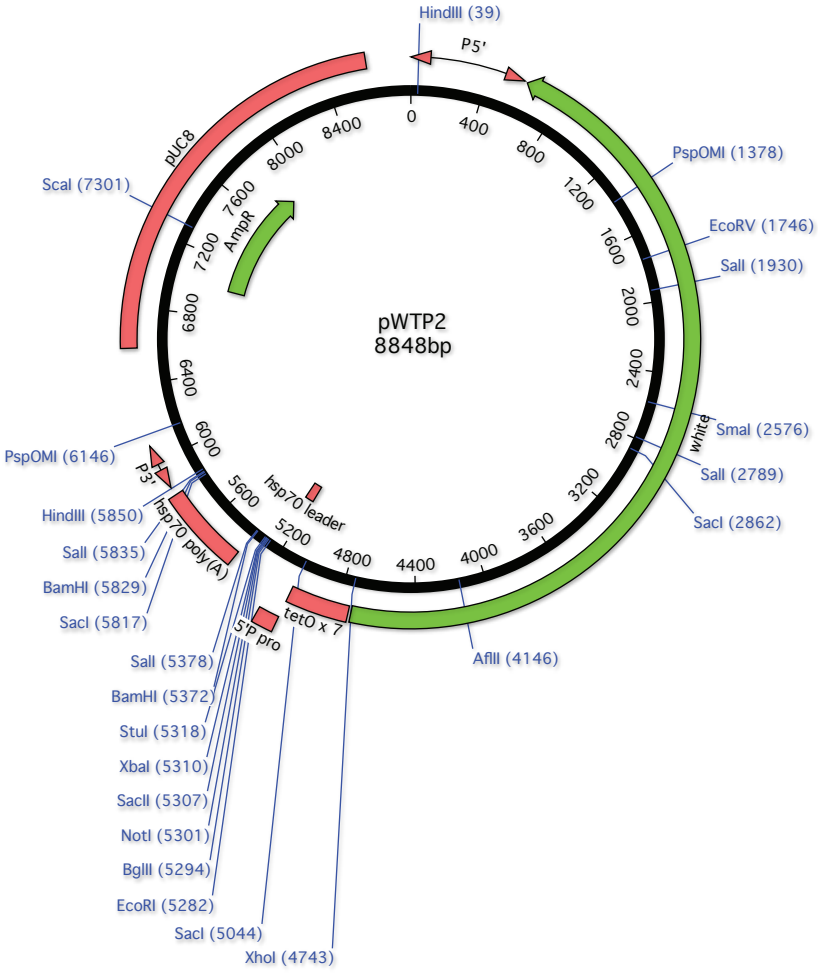


Figure I.2: pW.T.P.-2

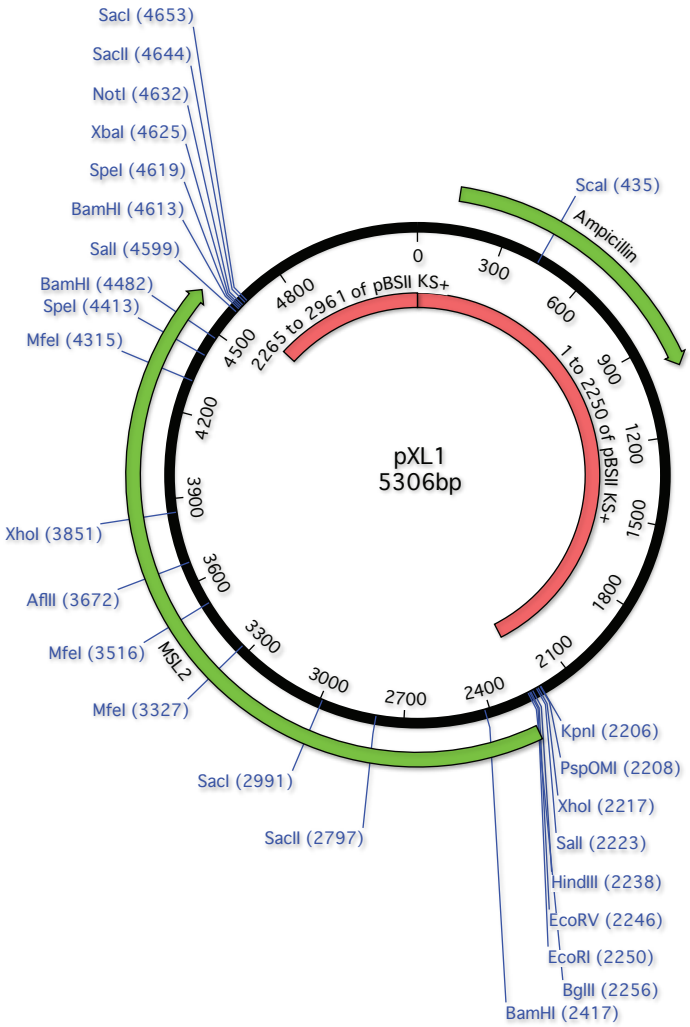


Figure I.3: pXL1

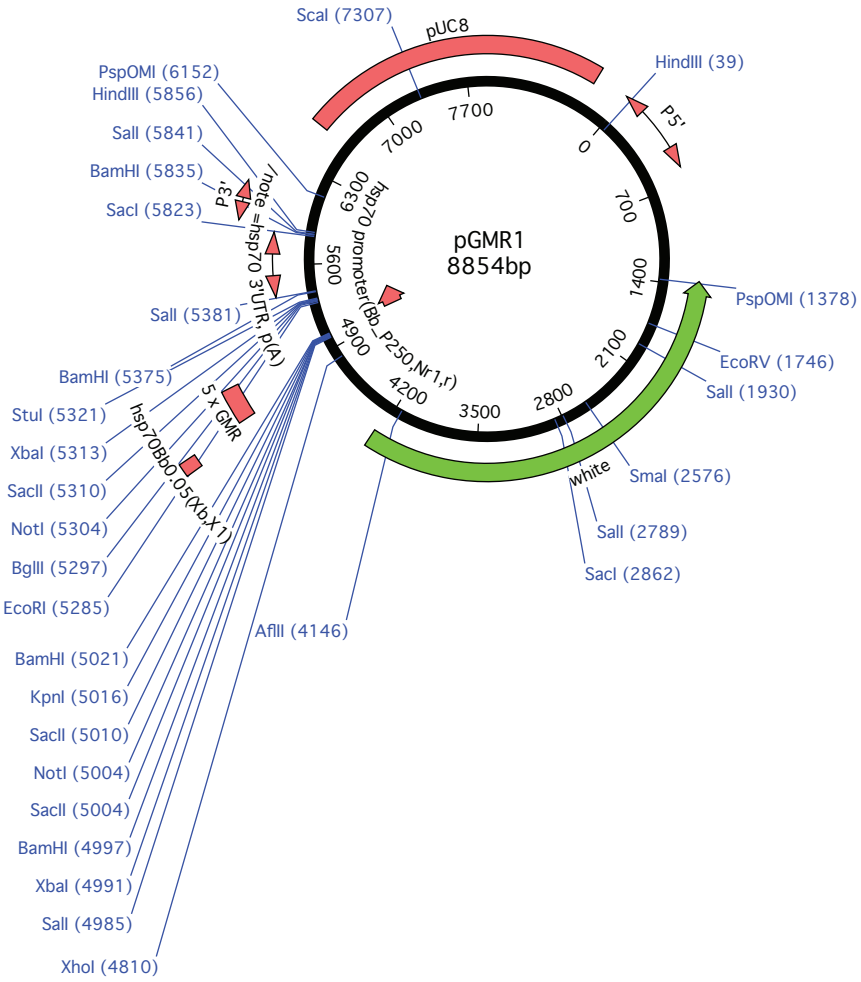


Figure I.4: pGMR-1

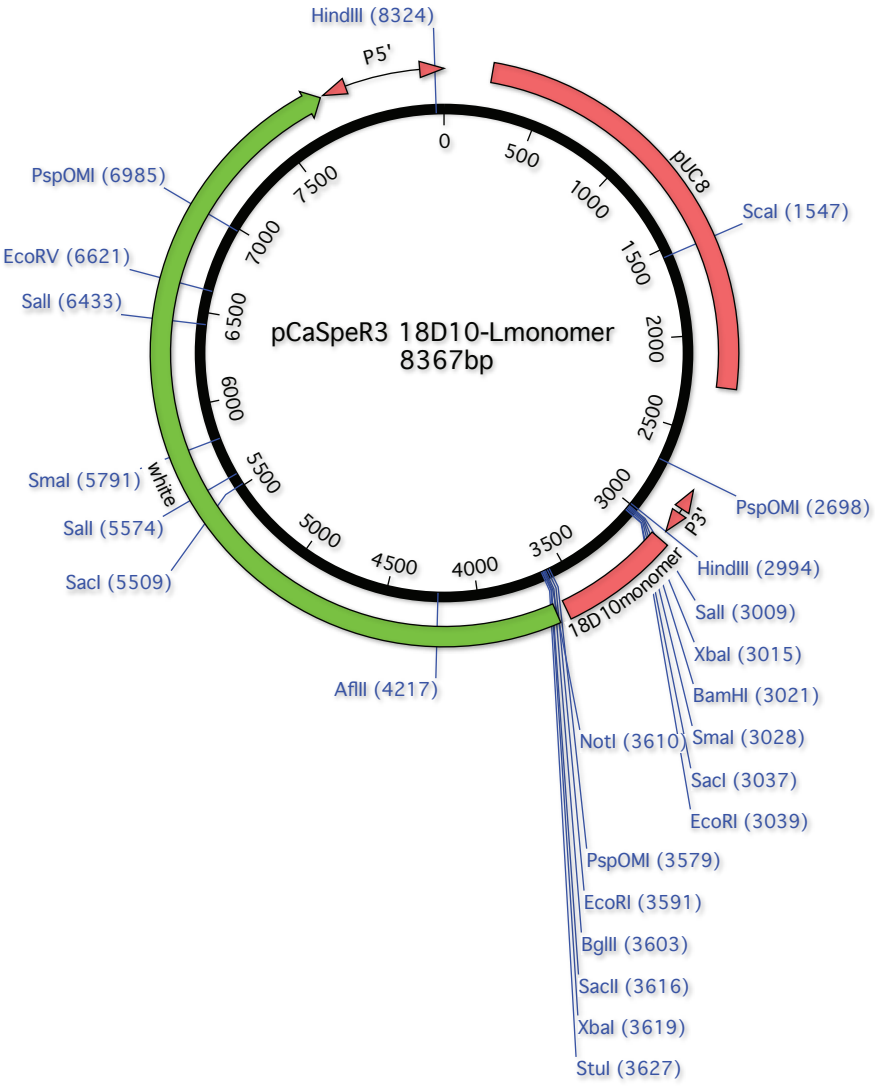


Figure I.5: pCaSpeR-3-18D-monomer

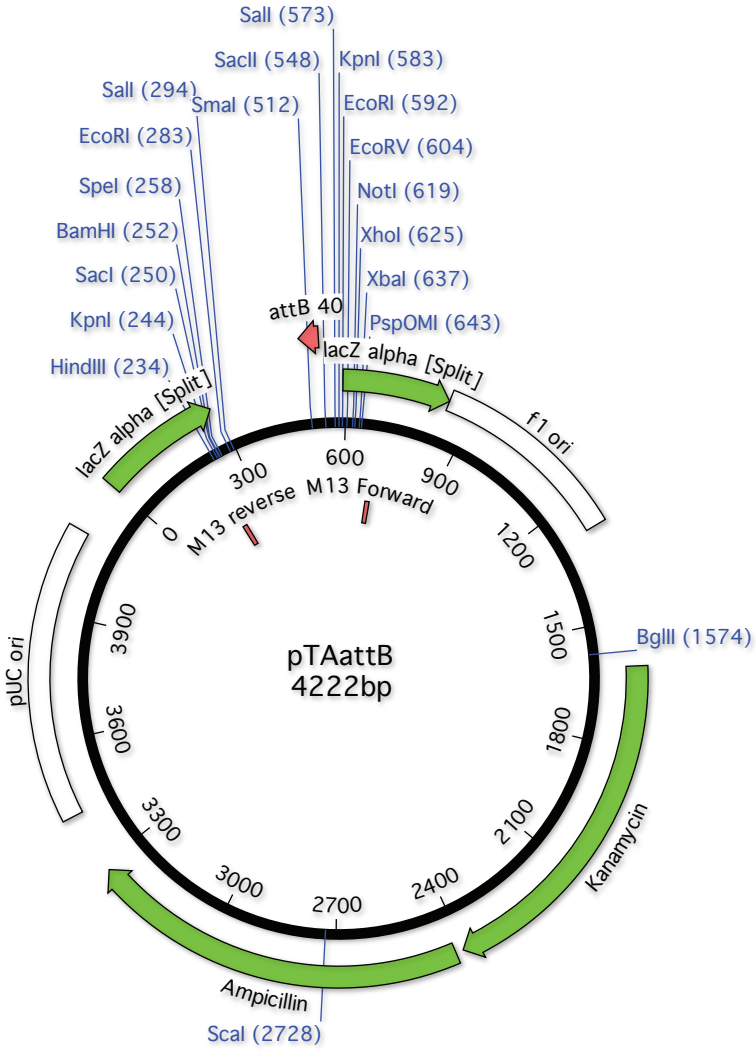


Figure I.6: pTAattB

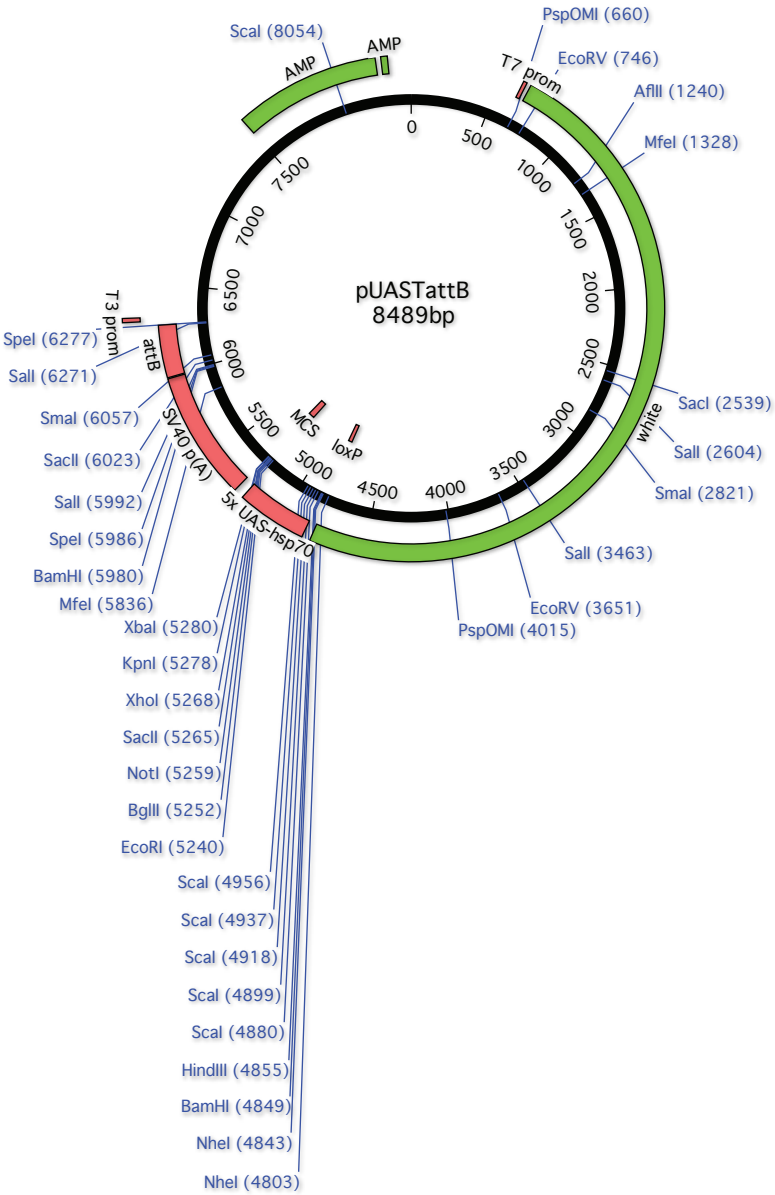


Figure I.7: pUASTattB

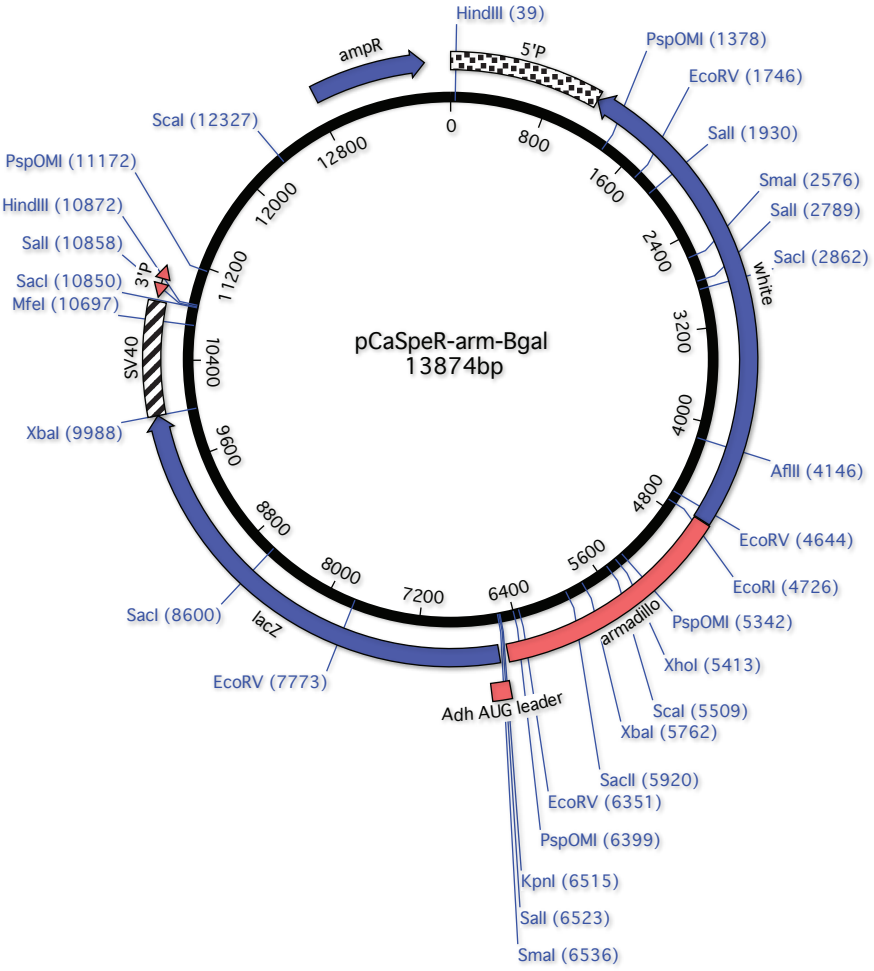


Figure I.8: pCaSpeR-arm-βgal

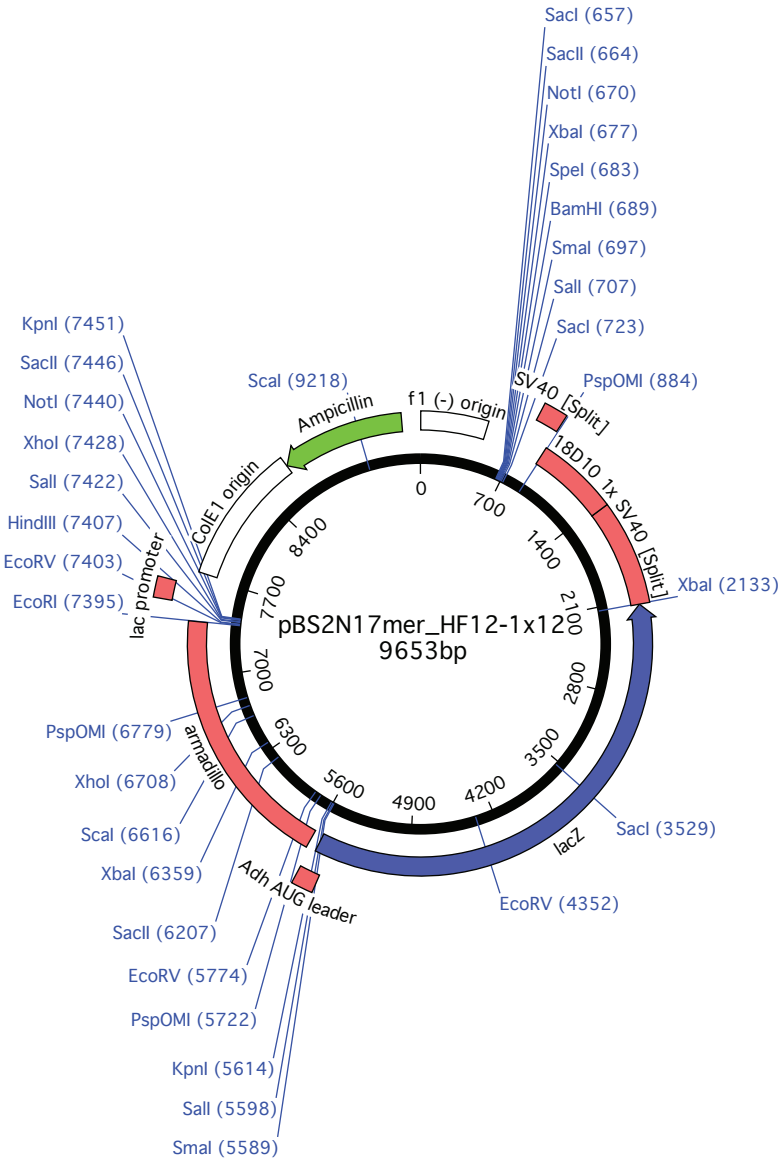


Figure I.9: pBS2N17mer HF12-1x12

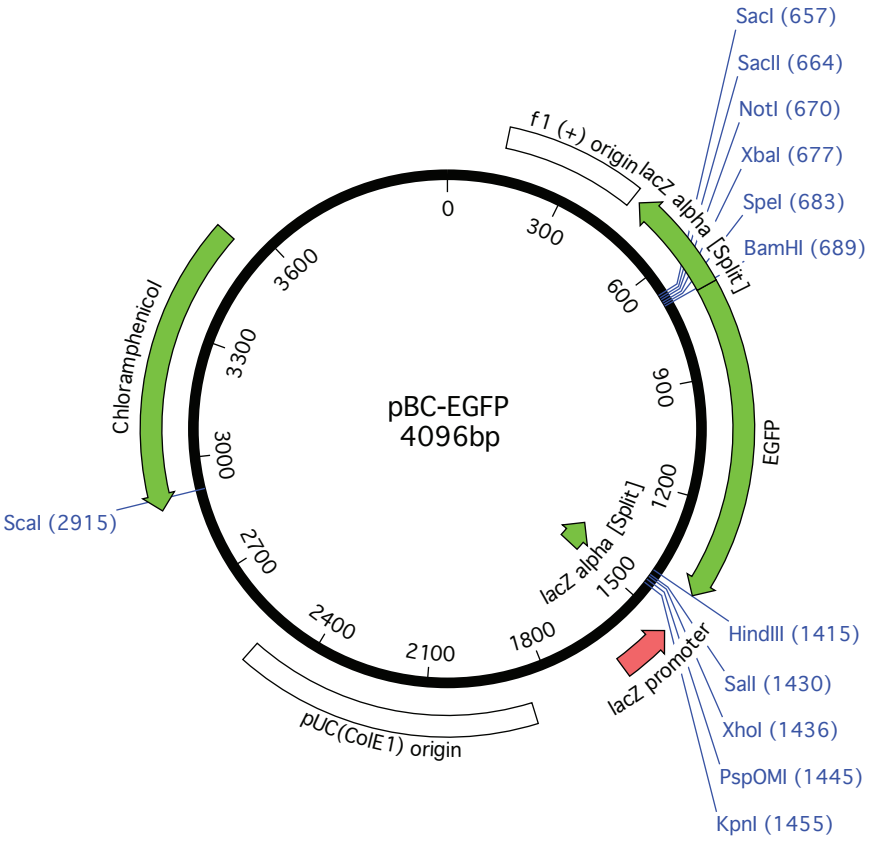


Figure I.10: pBC-EGFP

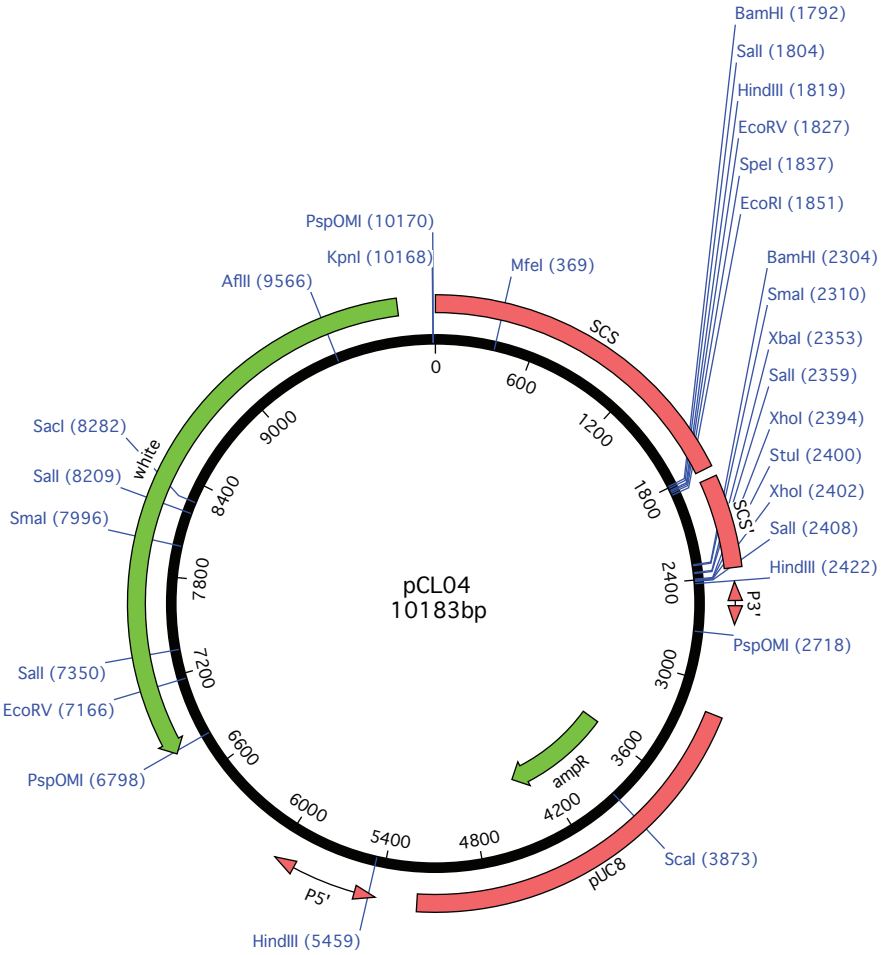


Figure I.11: pCL04

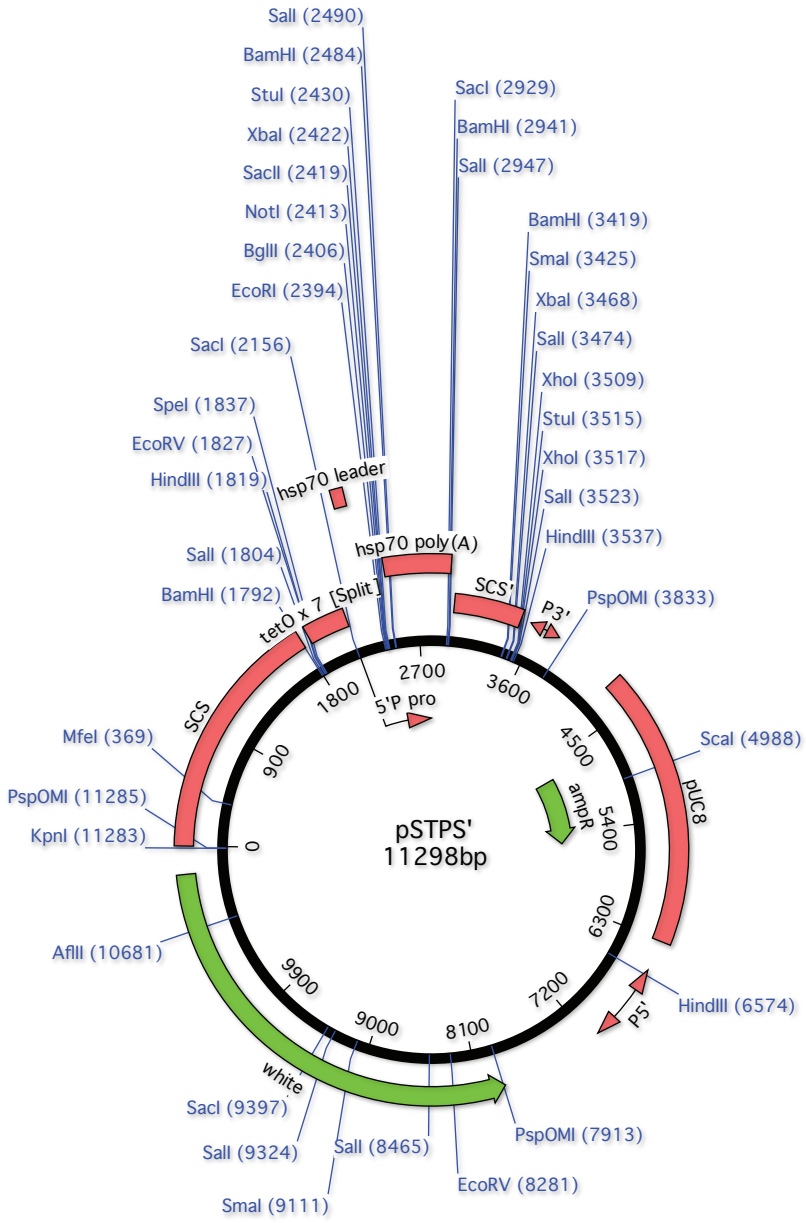


Figure I.12: pSTPS′

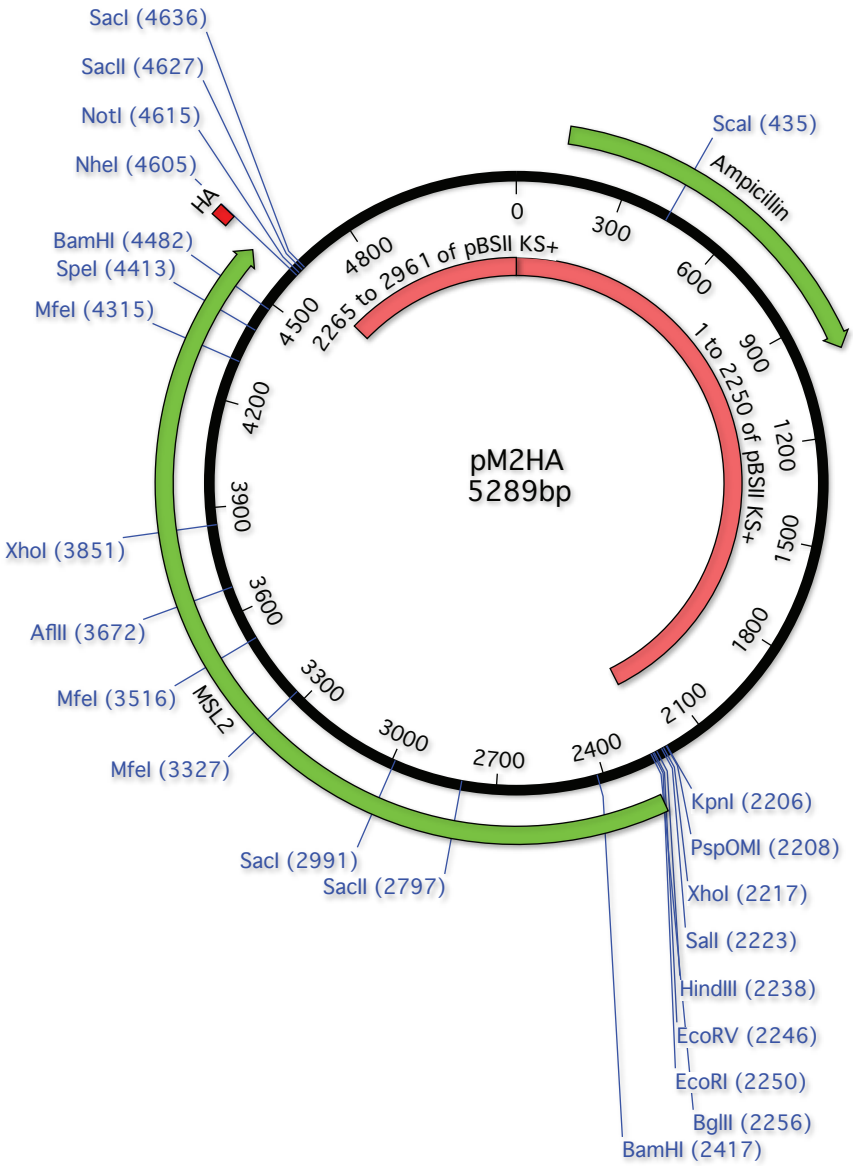


Figure I.13: pM2HA

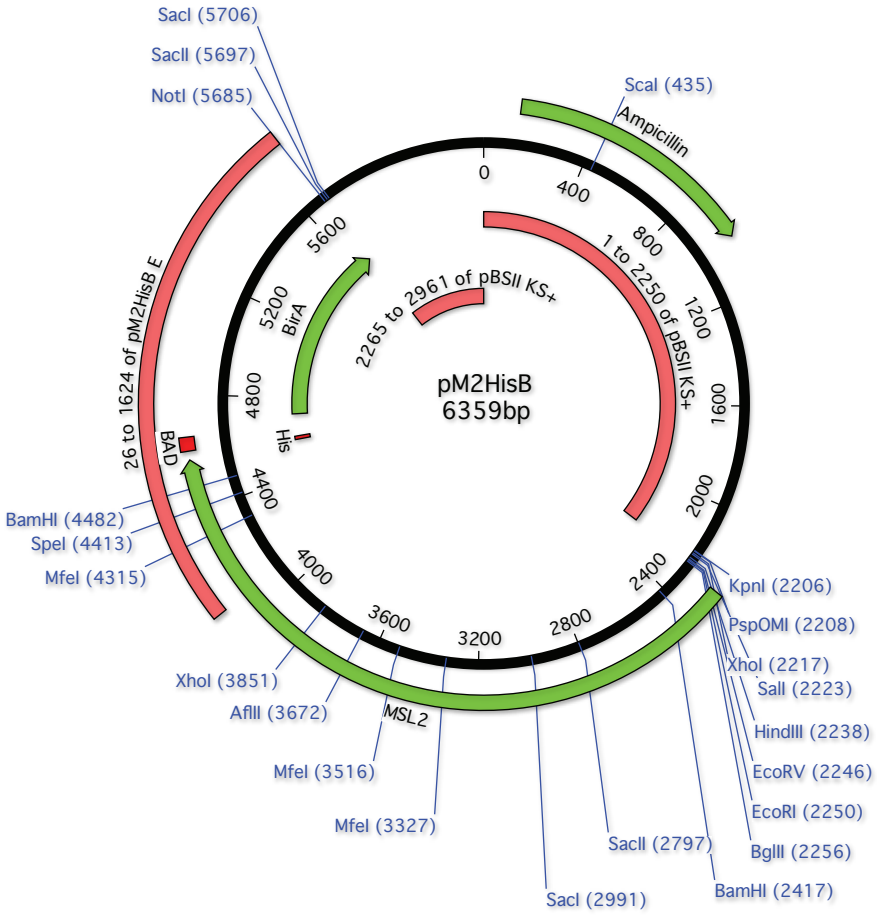


Figure I.14: pM2HisB

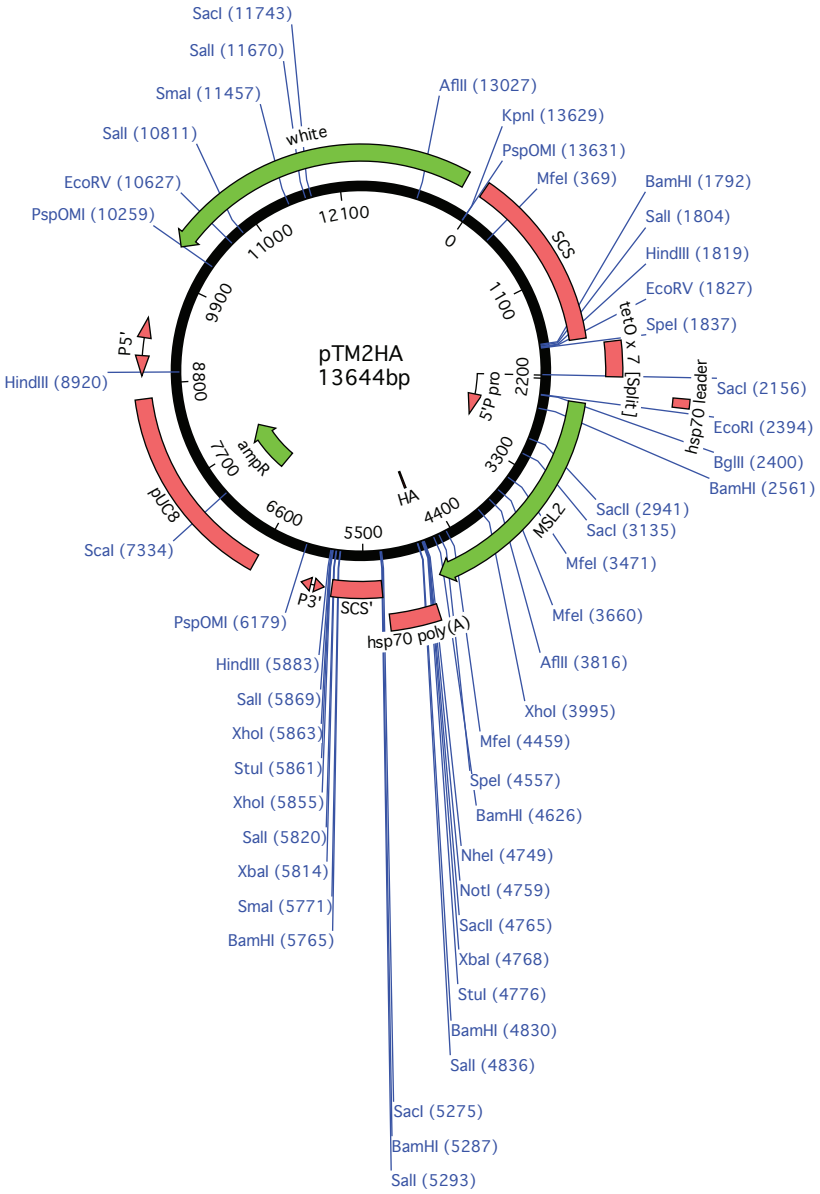


Figure I.15: pTM2HA

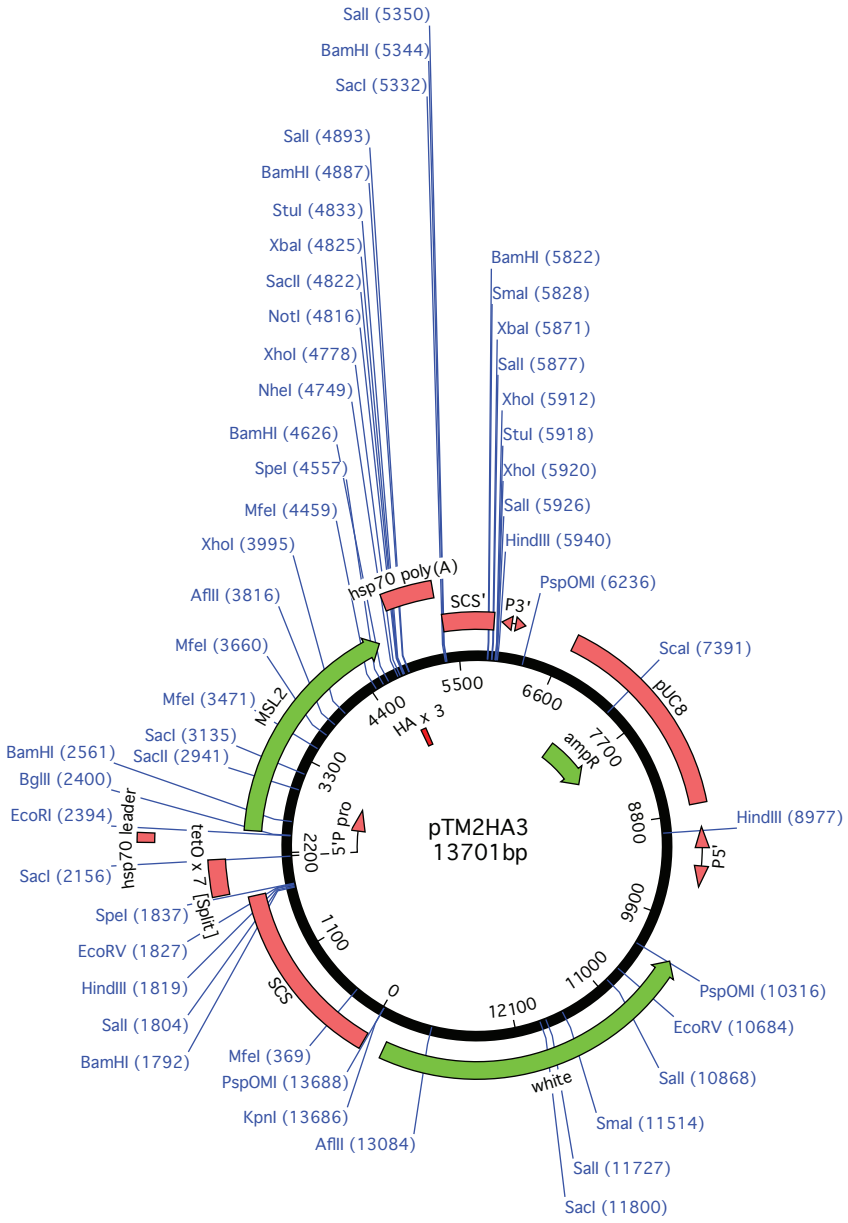


Figure I.16: pTM2HA3

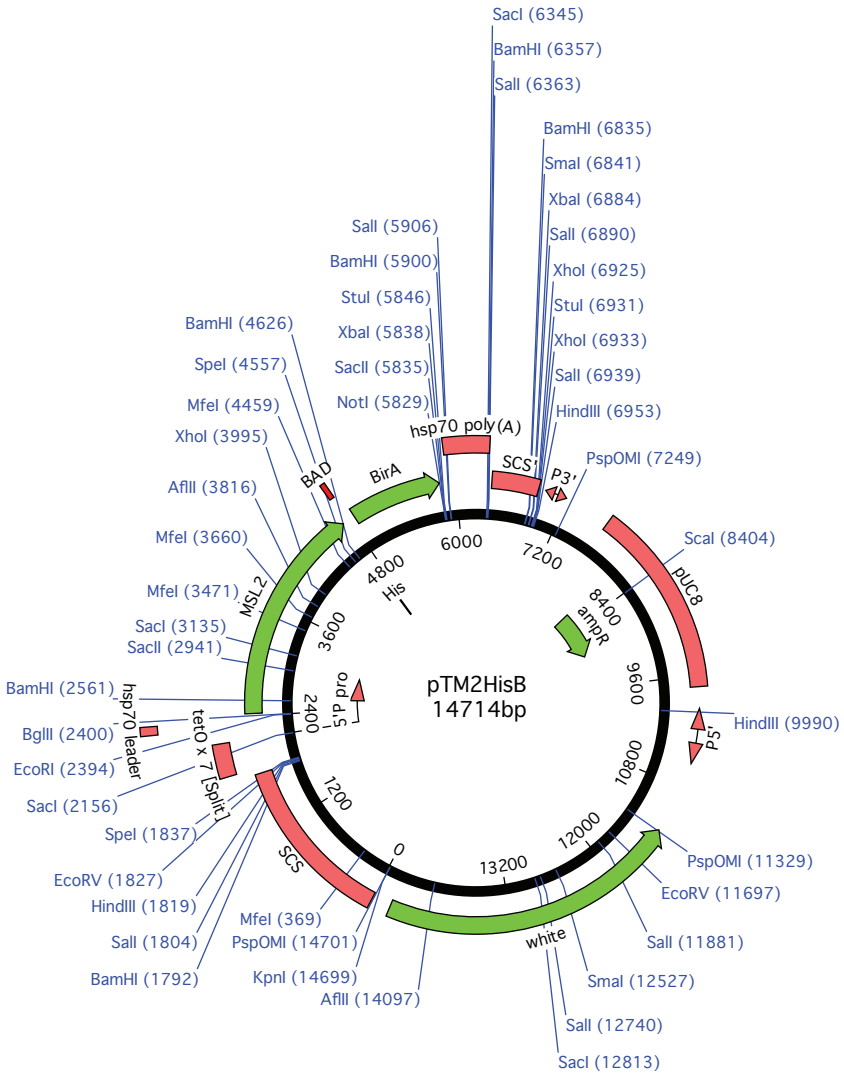


Figure I.17: pTM2HisB

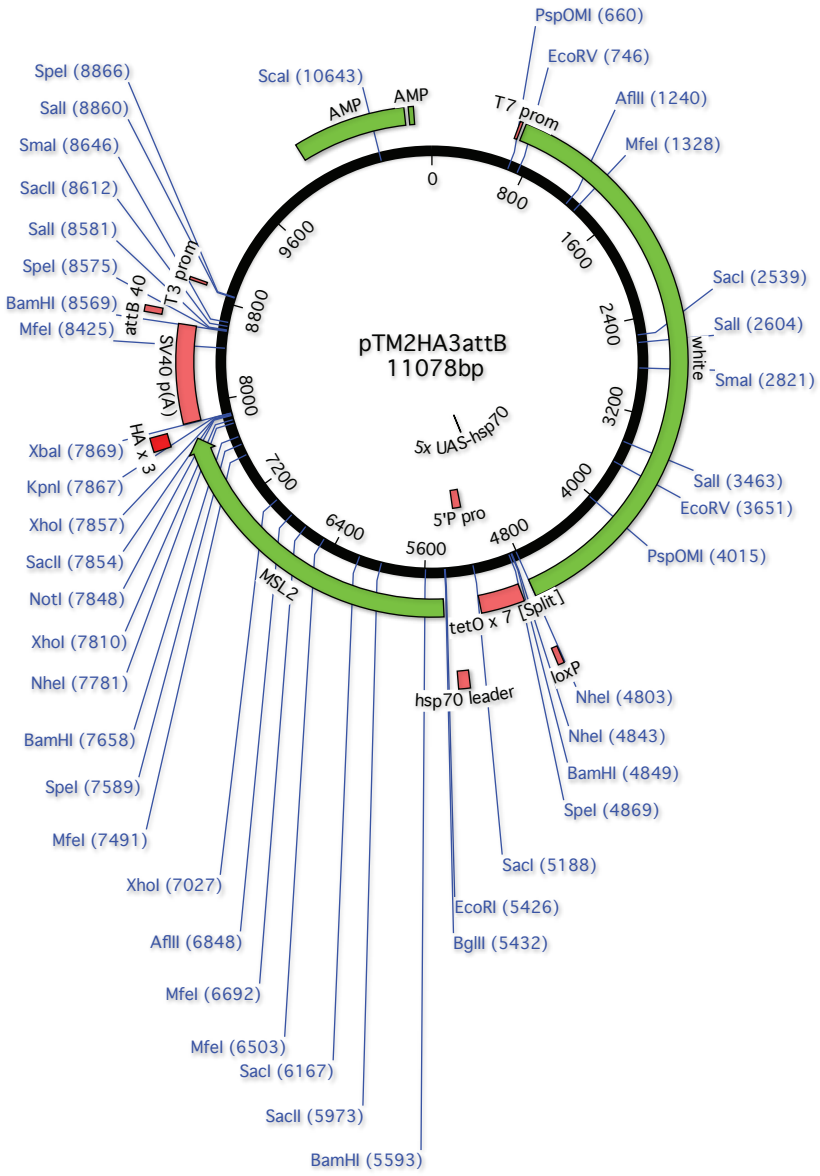


Figure I.18: pTM2HA3attB

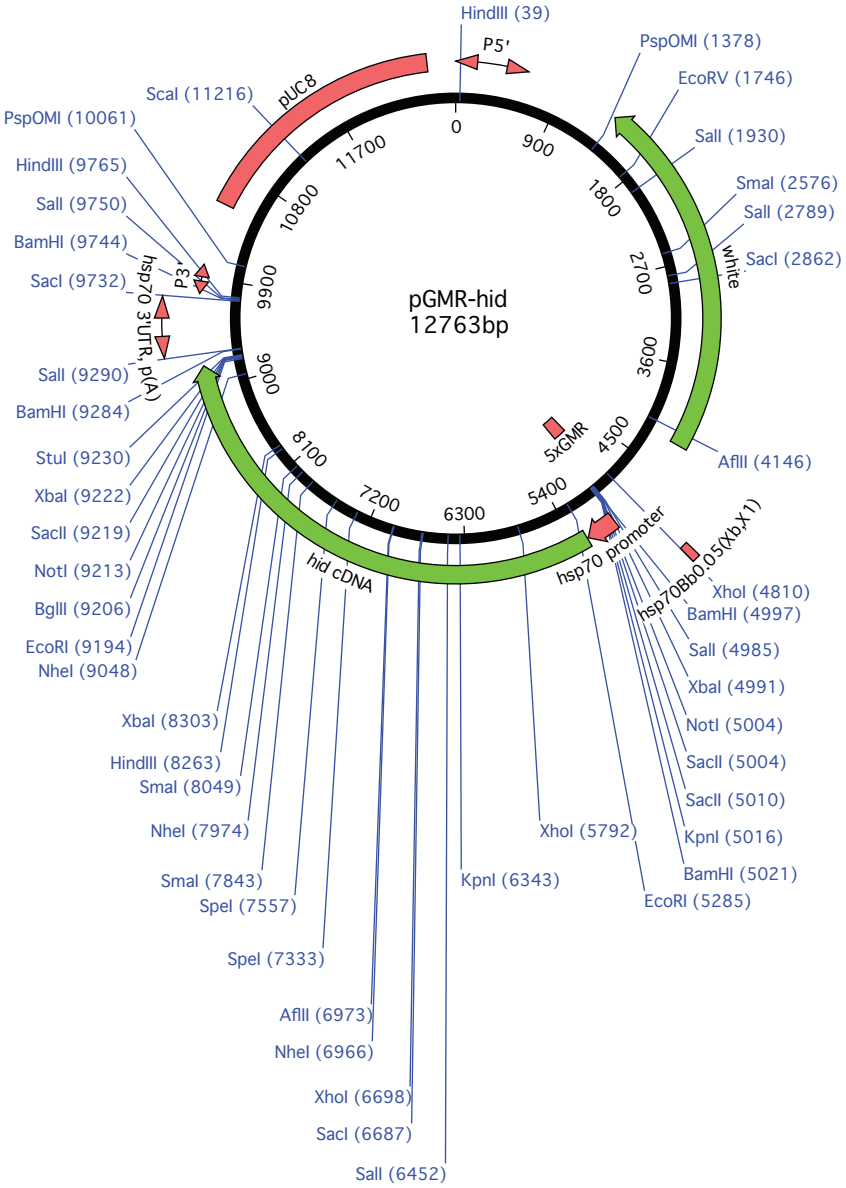


Figure I.19: pGMRhid

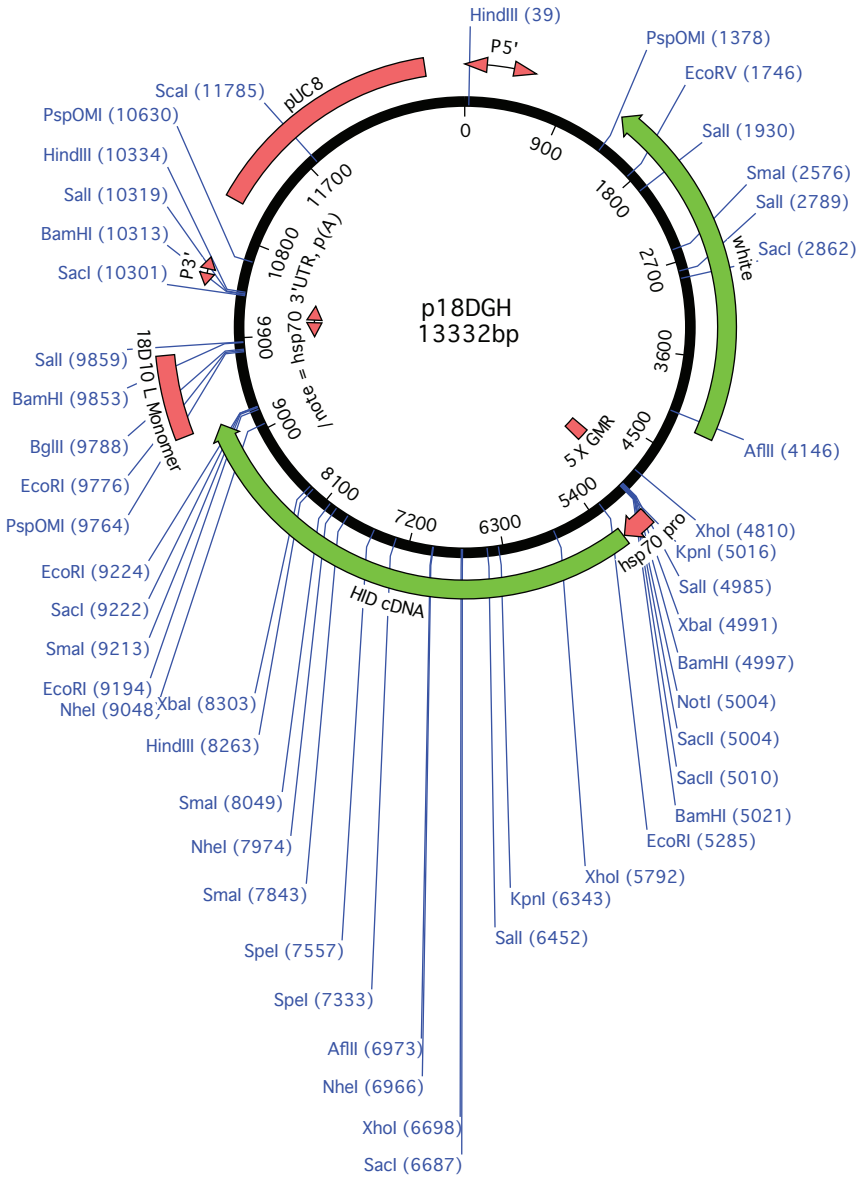


Figure I.20: p18DGH

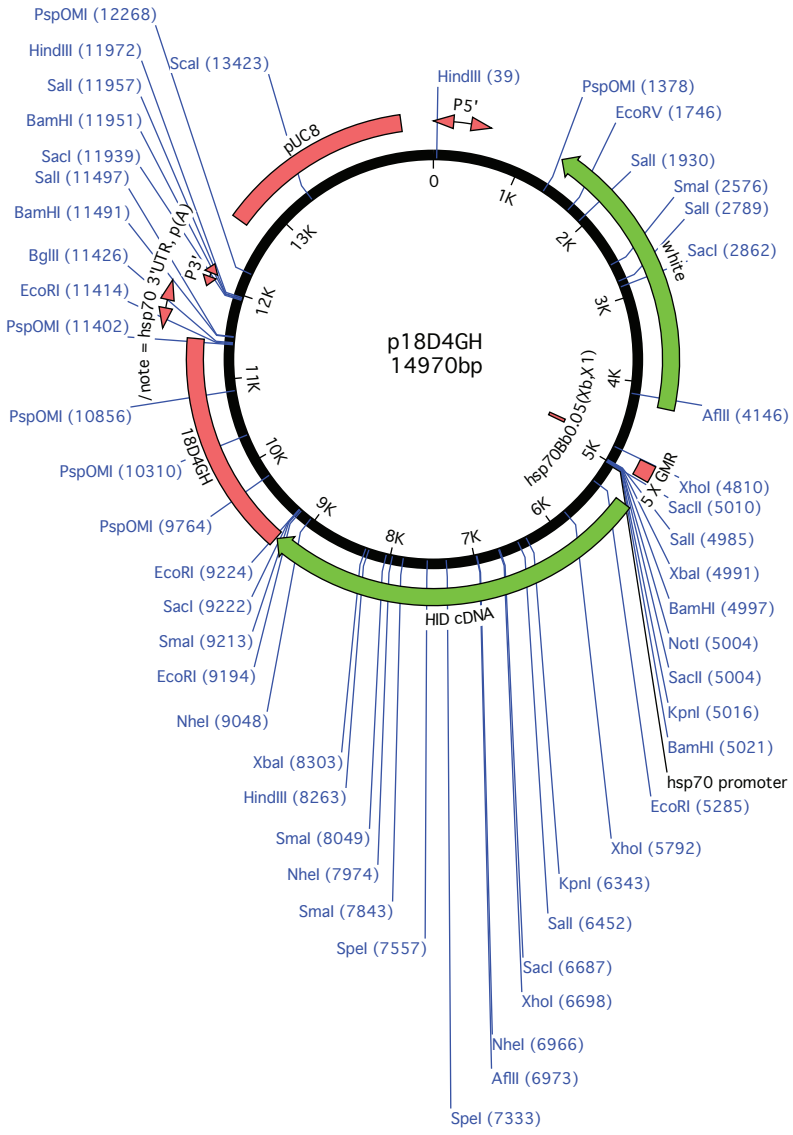


Figure I.21: p18D4GH

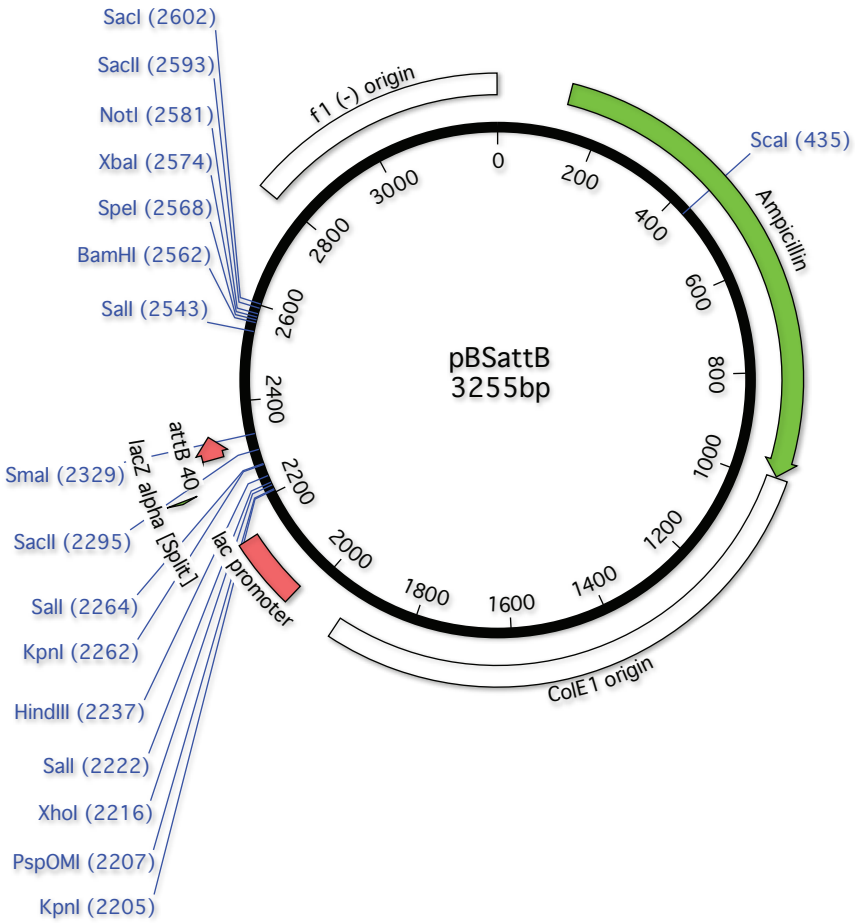


Figure I.22: pBSattB

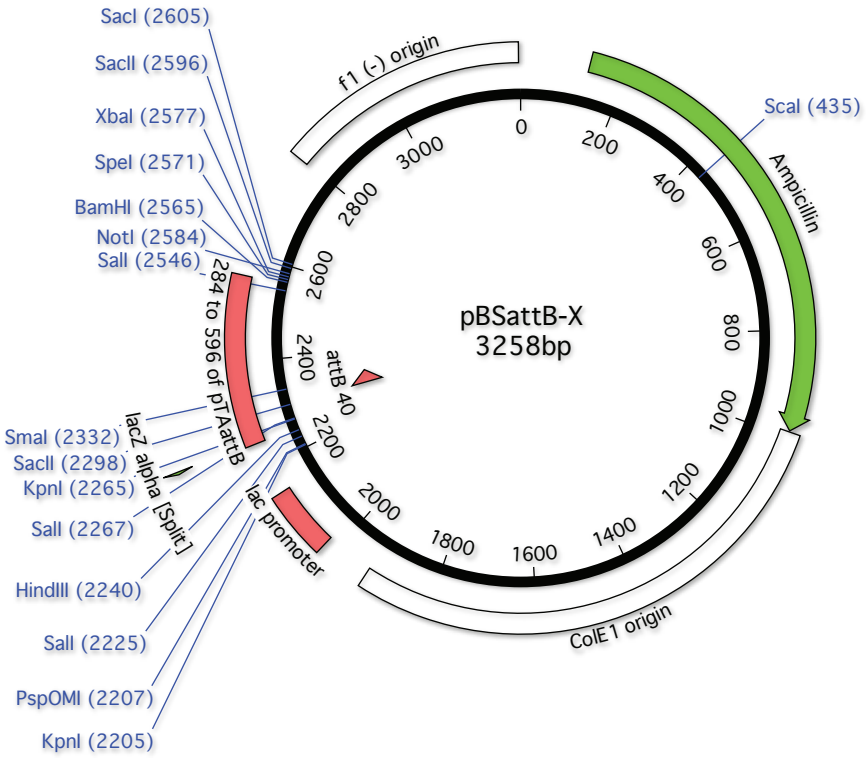


Figure I.23: pBSattB-X

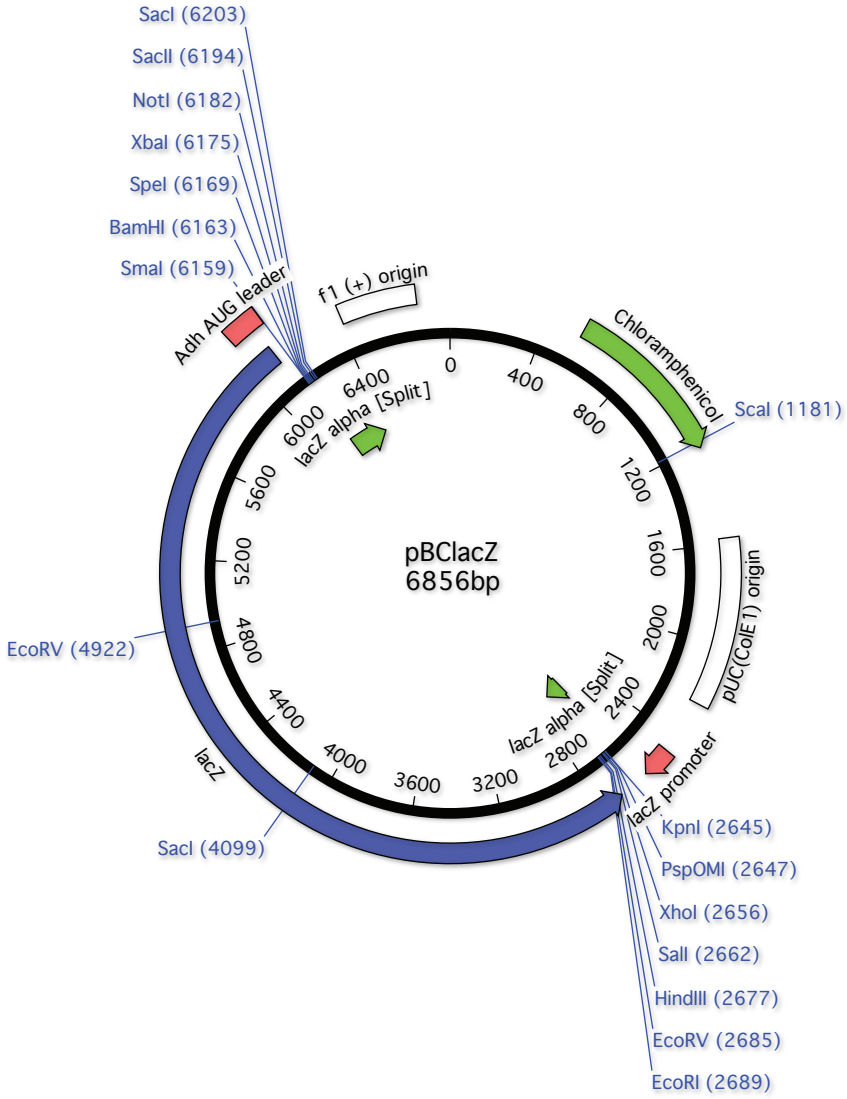


Figure I.24: pBClacZ

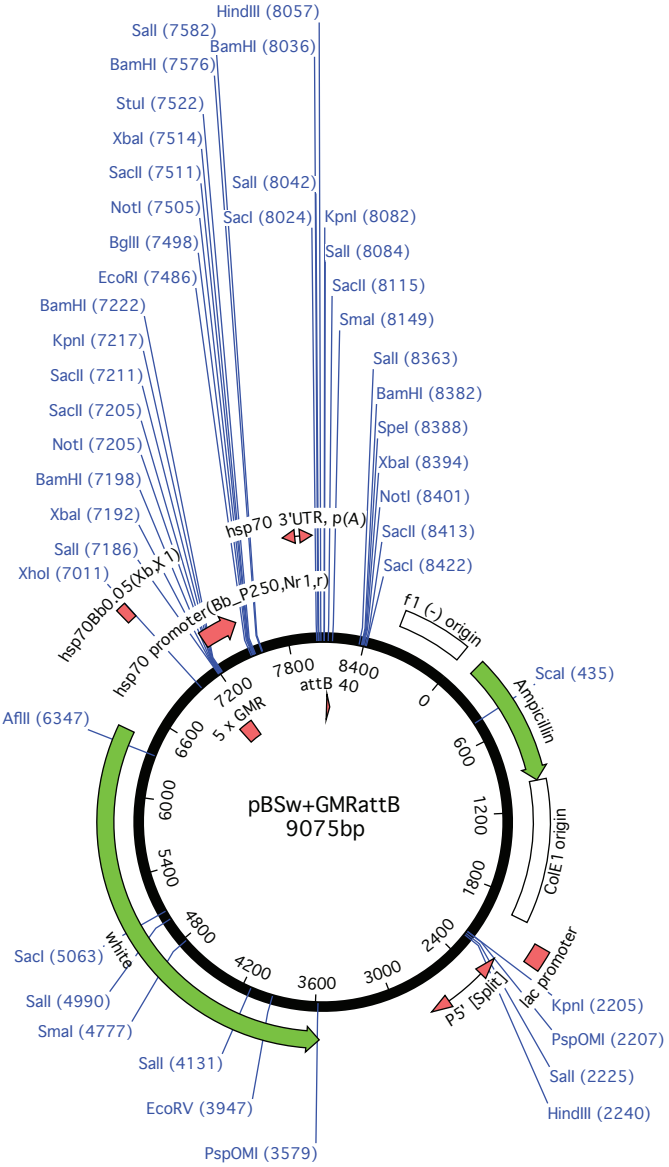


Figure I.25: pBSw+GMRattB

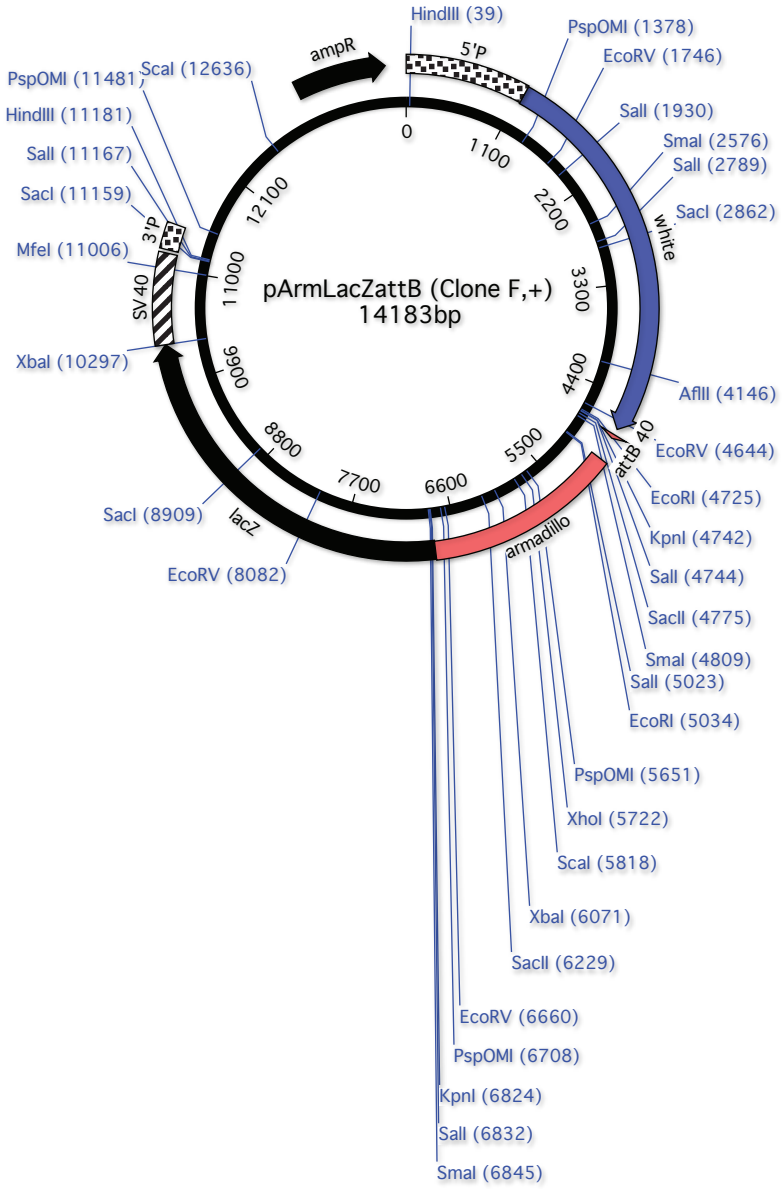


Figure I.26: pALattB

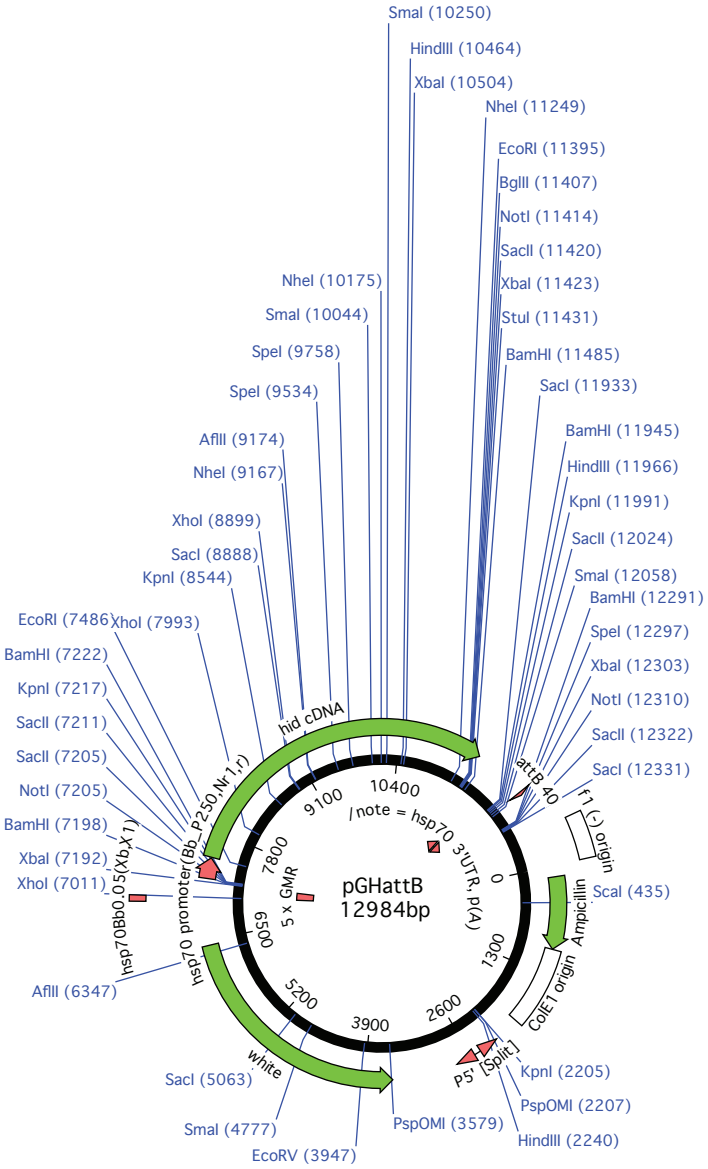


Figure I.27: pGHattB. *Sal* I sites removed for clarity.

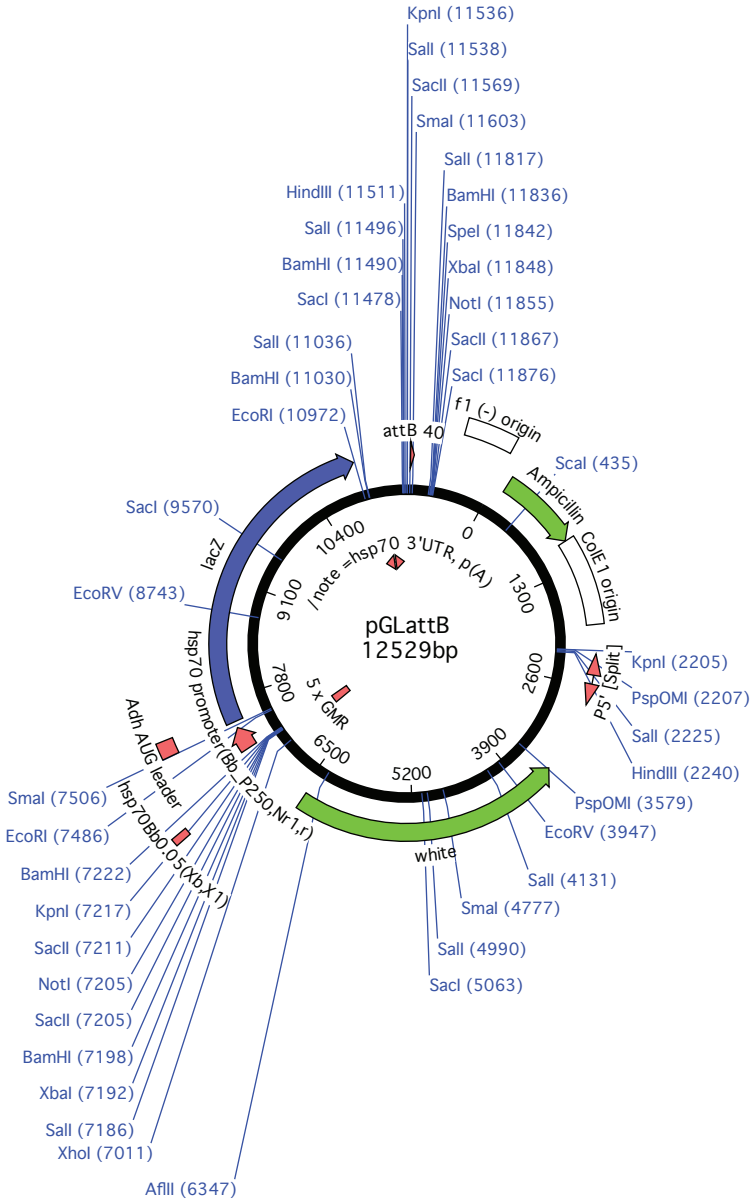


Figure I.28: pGLattB

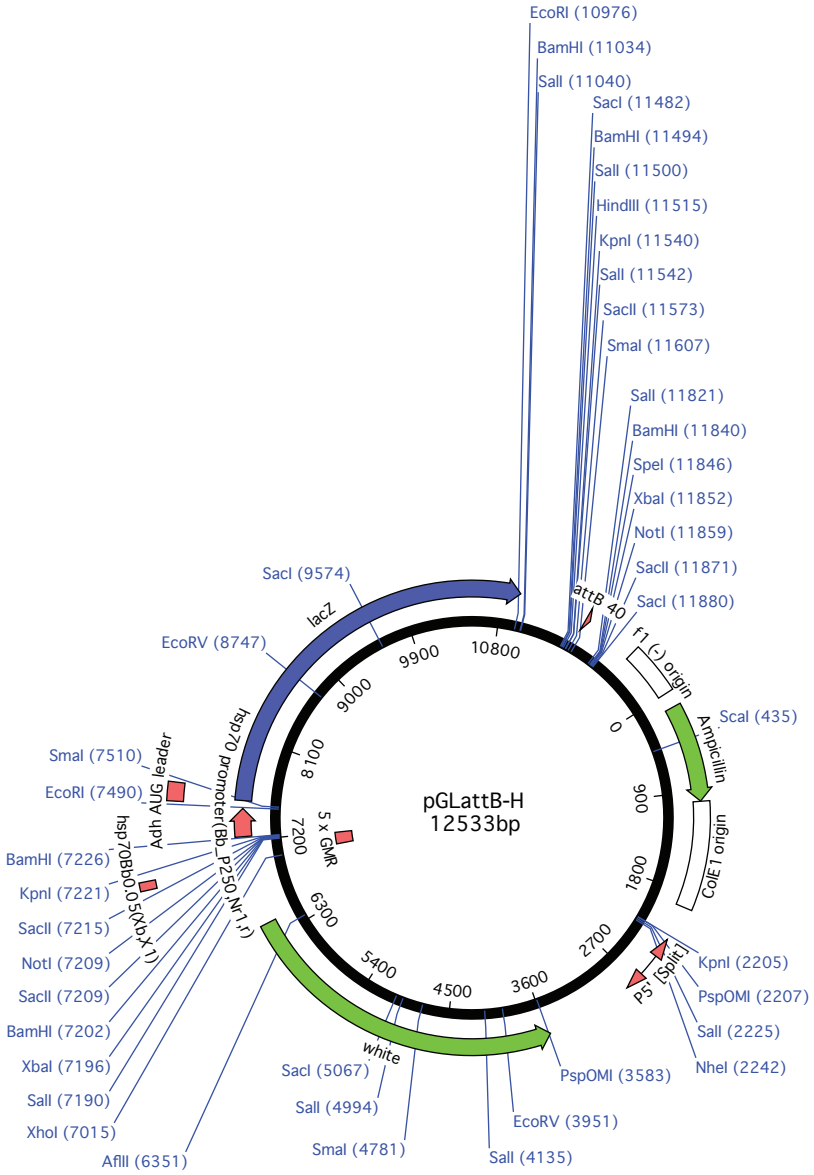


Figure I.29: pGLattB-H

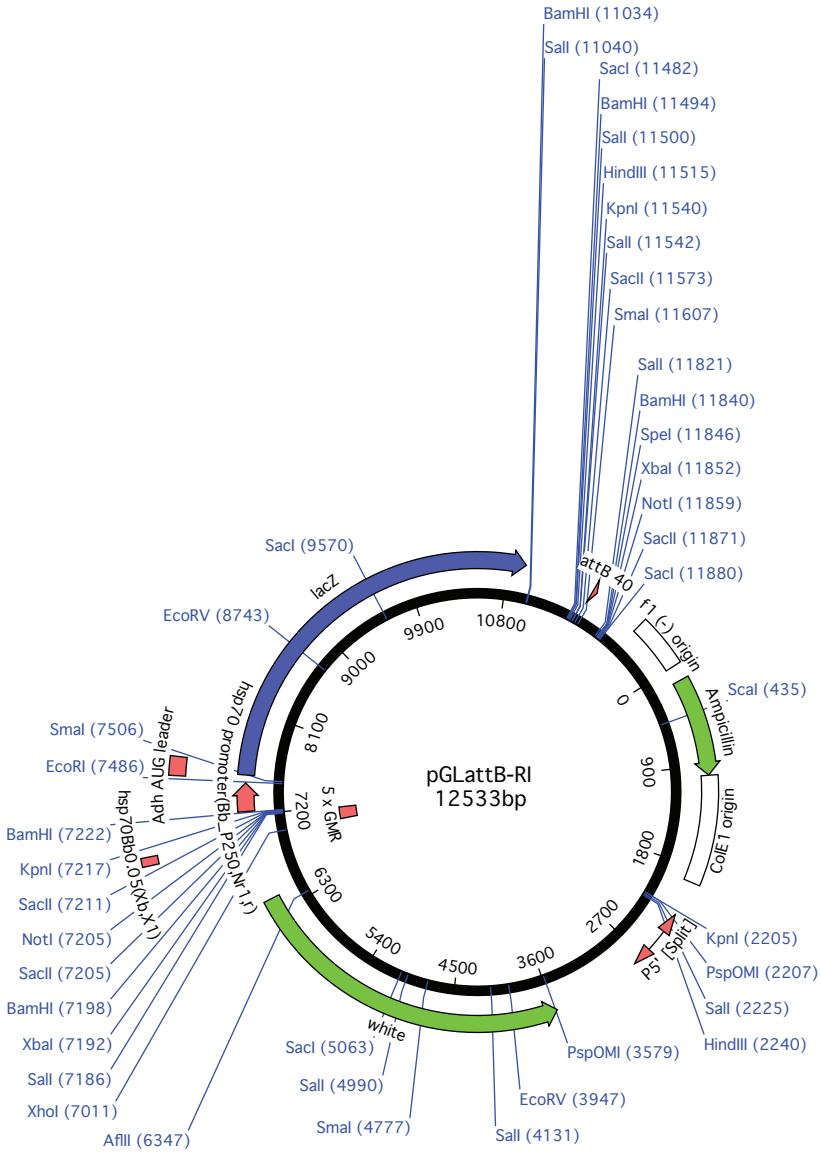


Figure I.30: pGLattB-RI

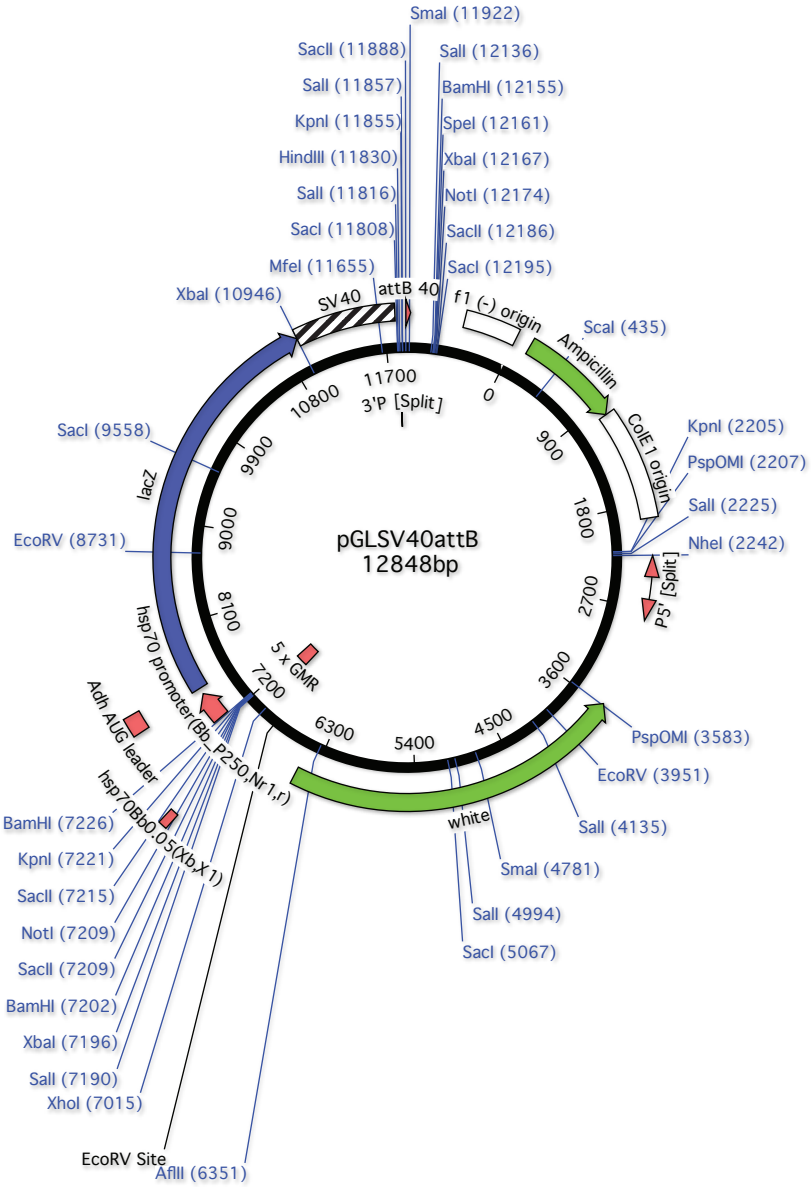


Figure I.31: pGLSV40

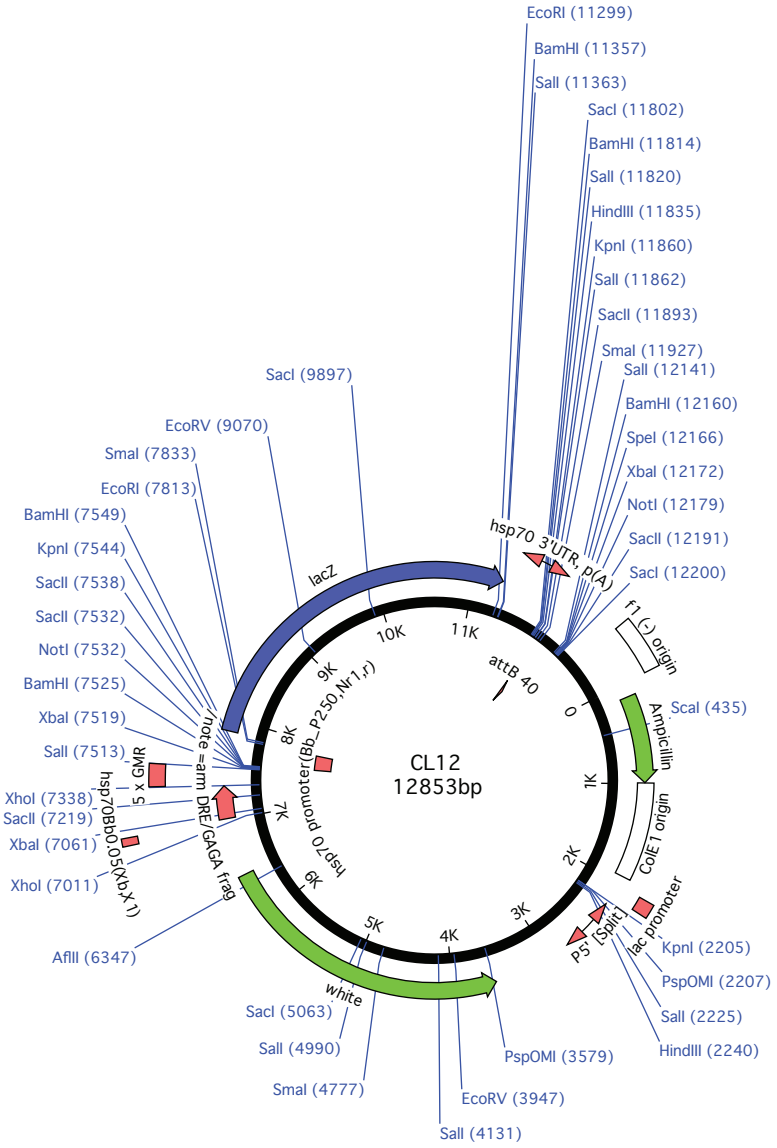


Figure I.32: pCL12

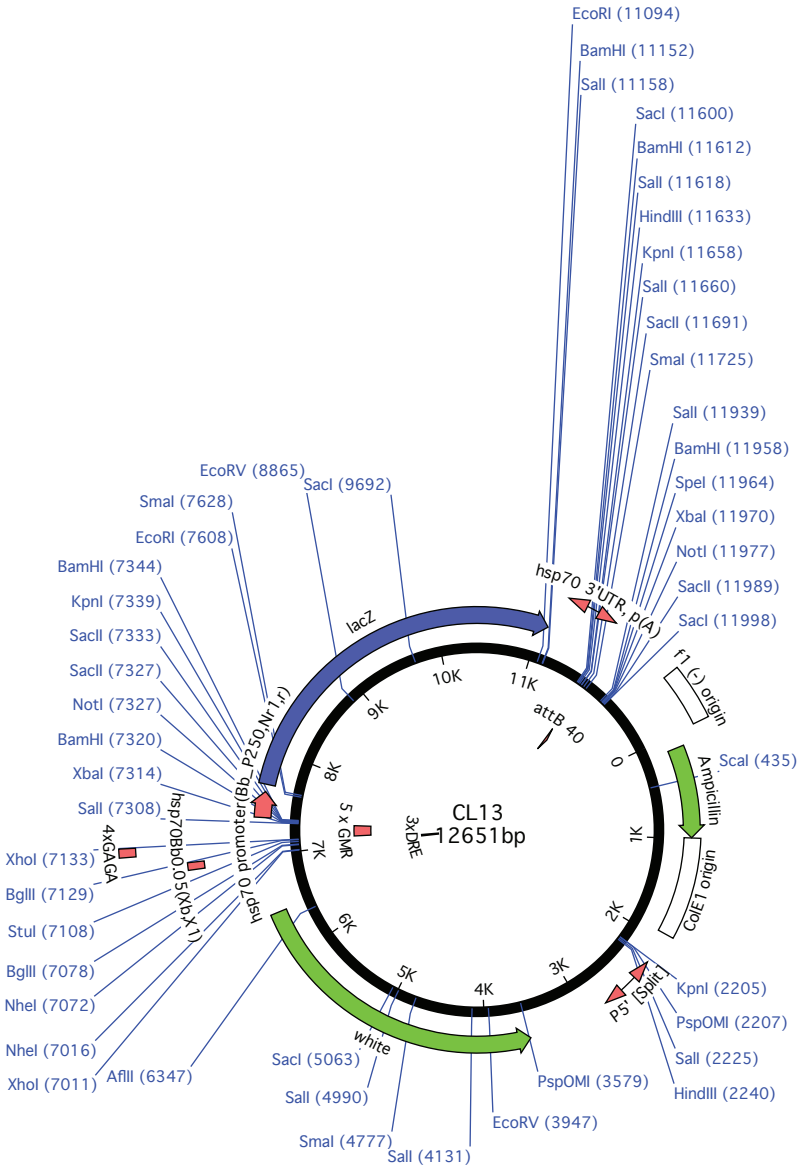


Figure I.33: pCL13

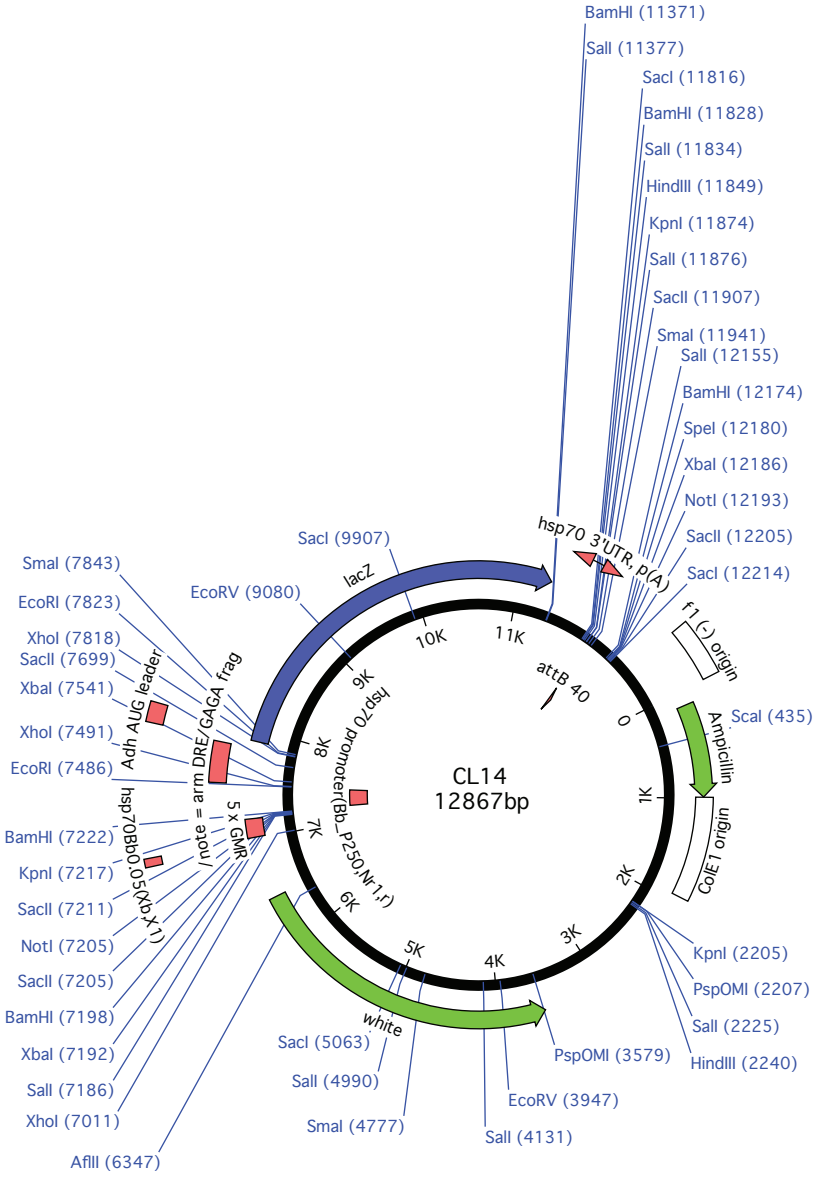


Figure I.34: pCL14

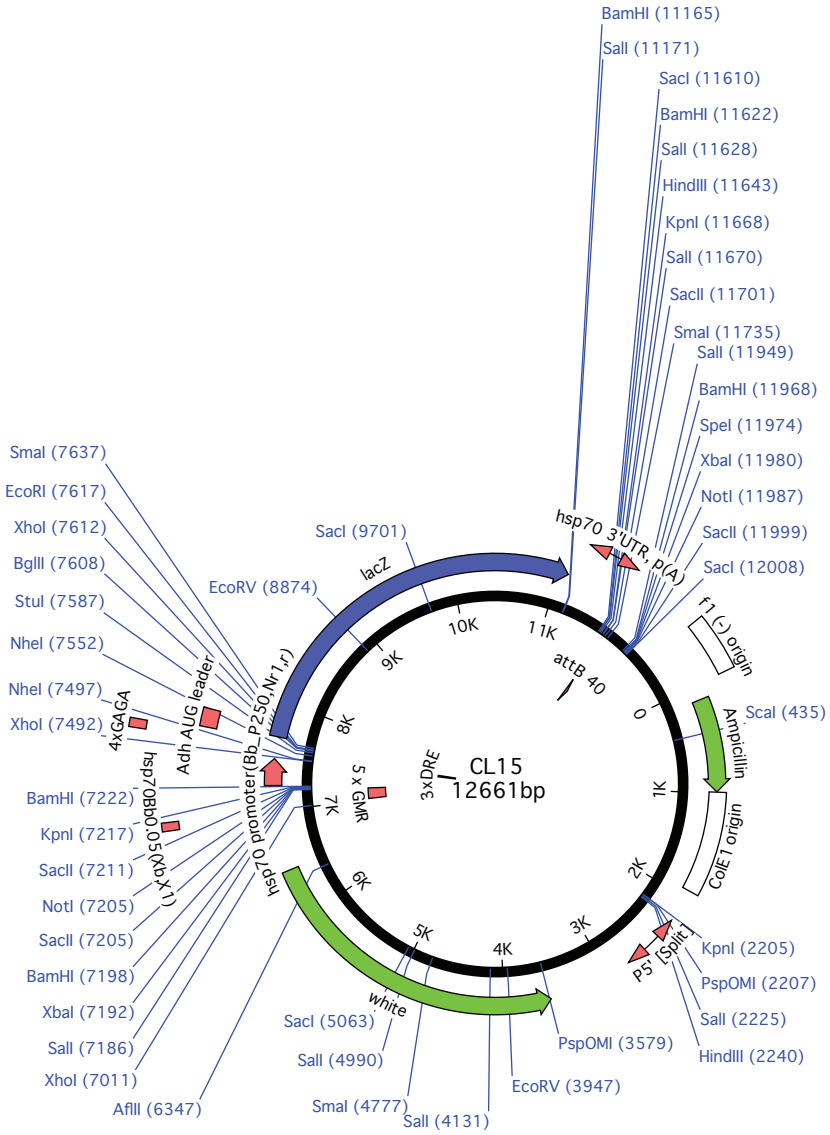


Figure I.35: pCL15

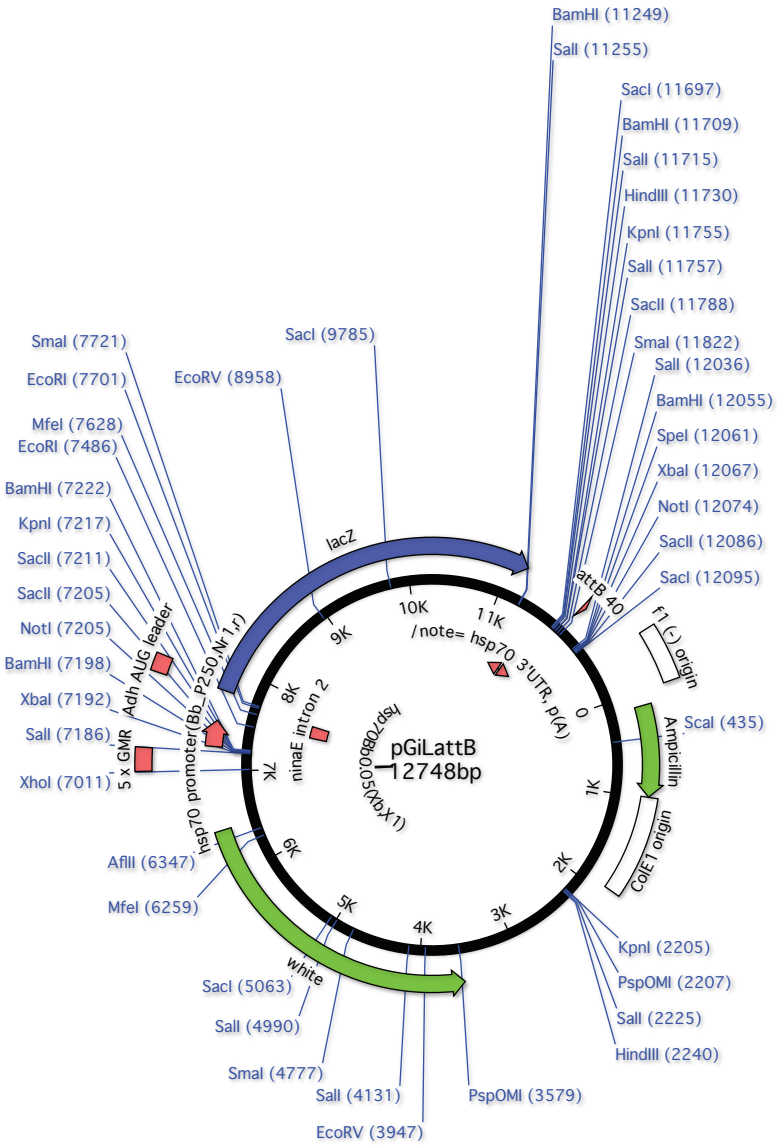


Figure I.36: pGiLattB

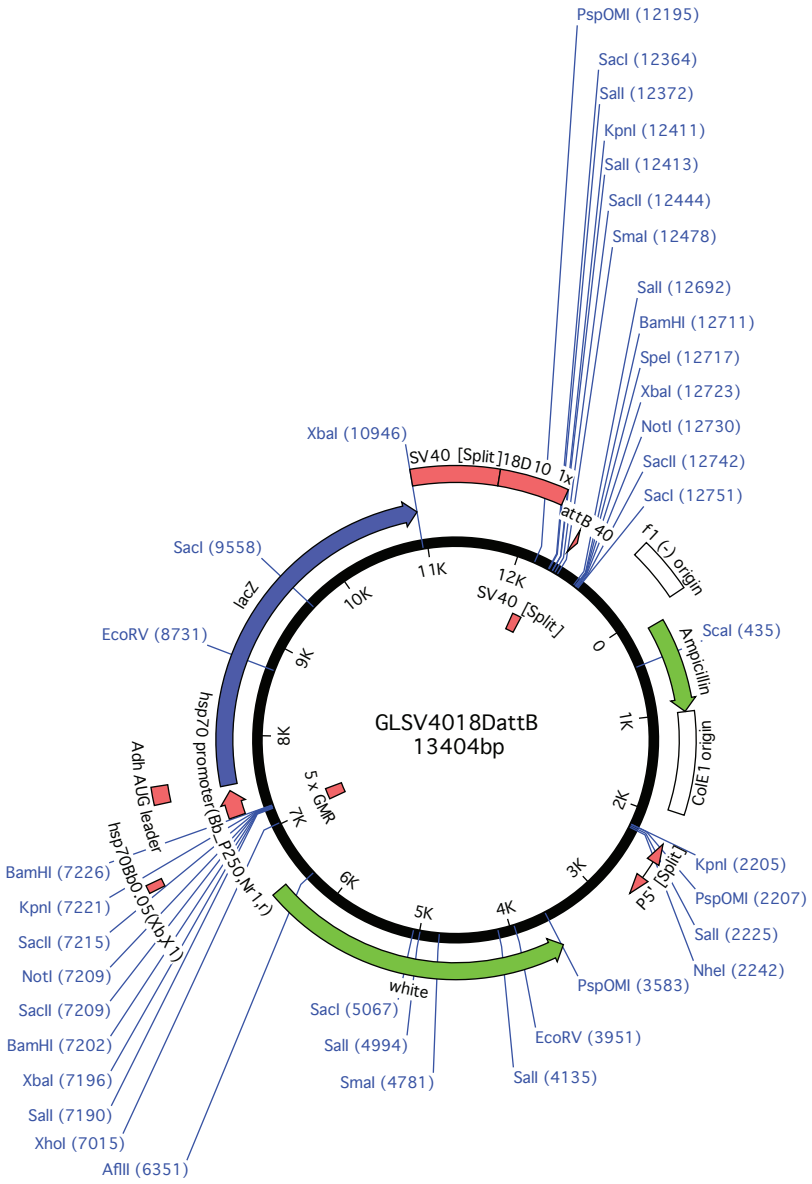


Figure I.37: pGLSV4018DattB

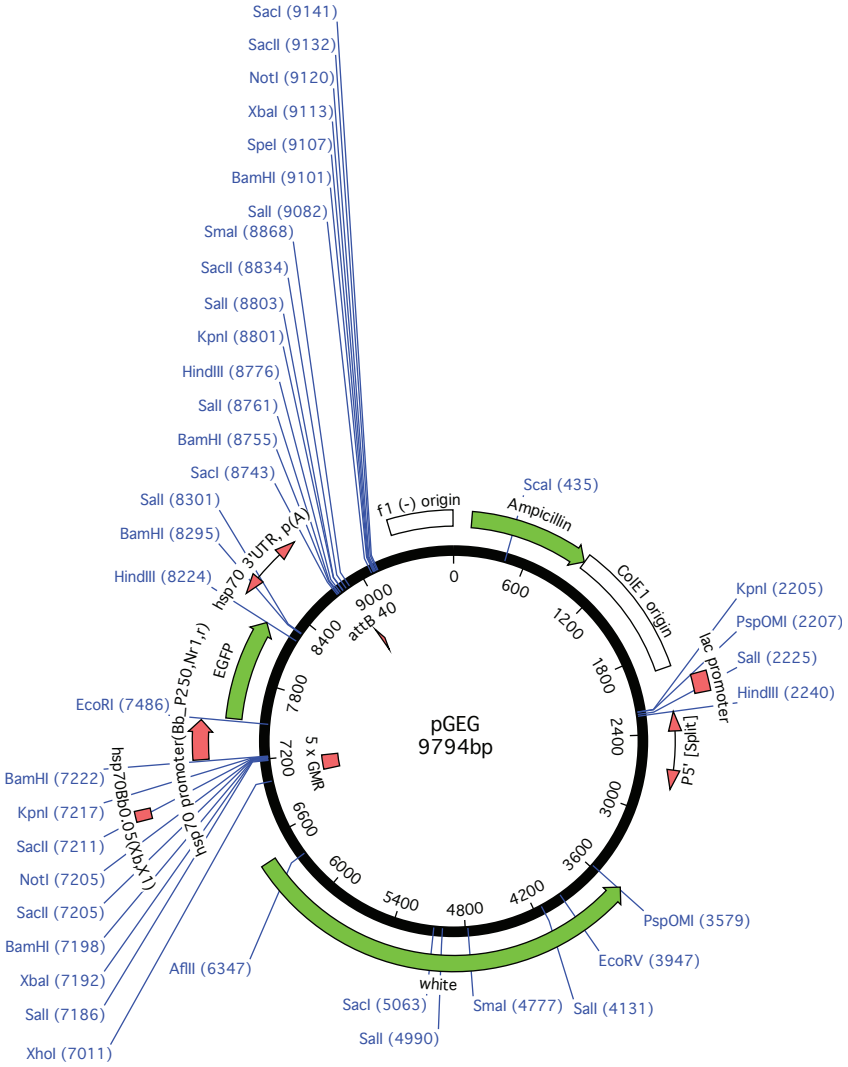


Figure I.38: pGEG

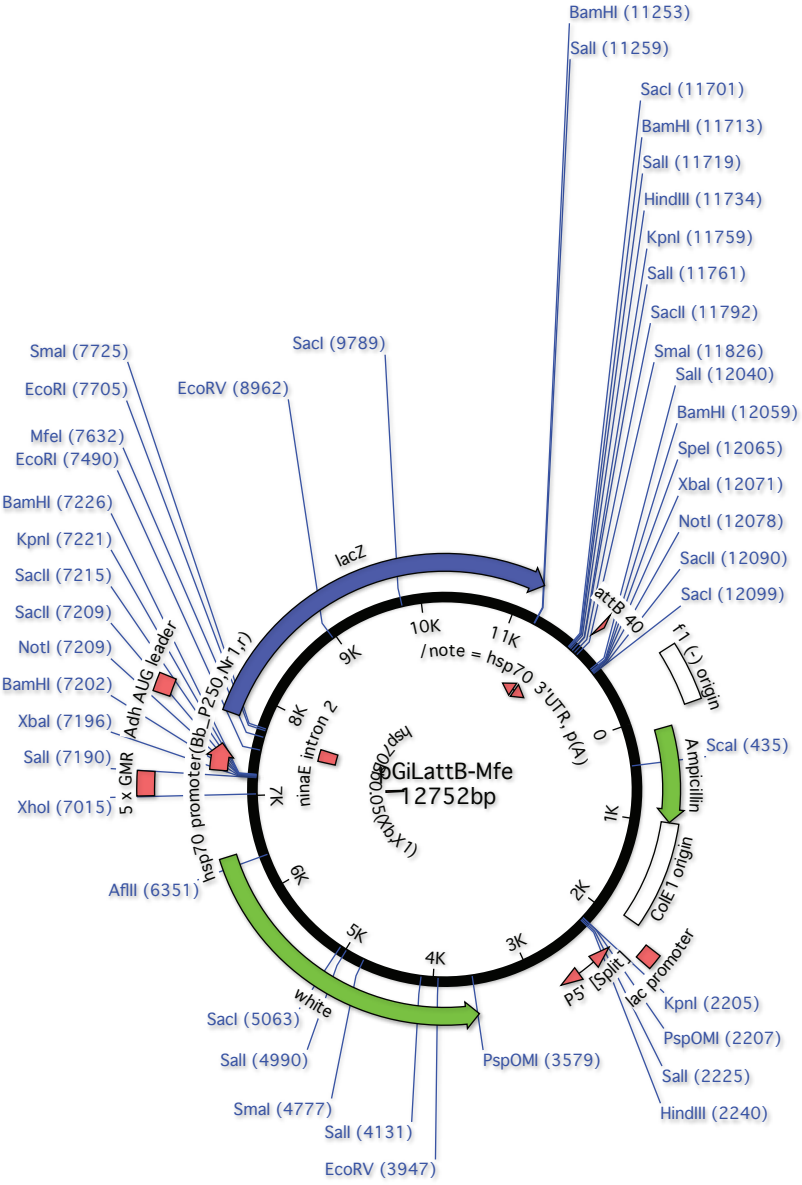


Figure I.39: pGiLattB-MfeI

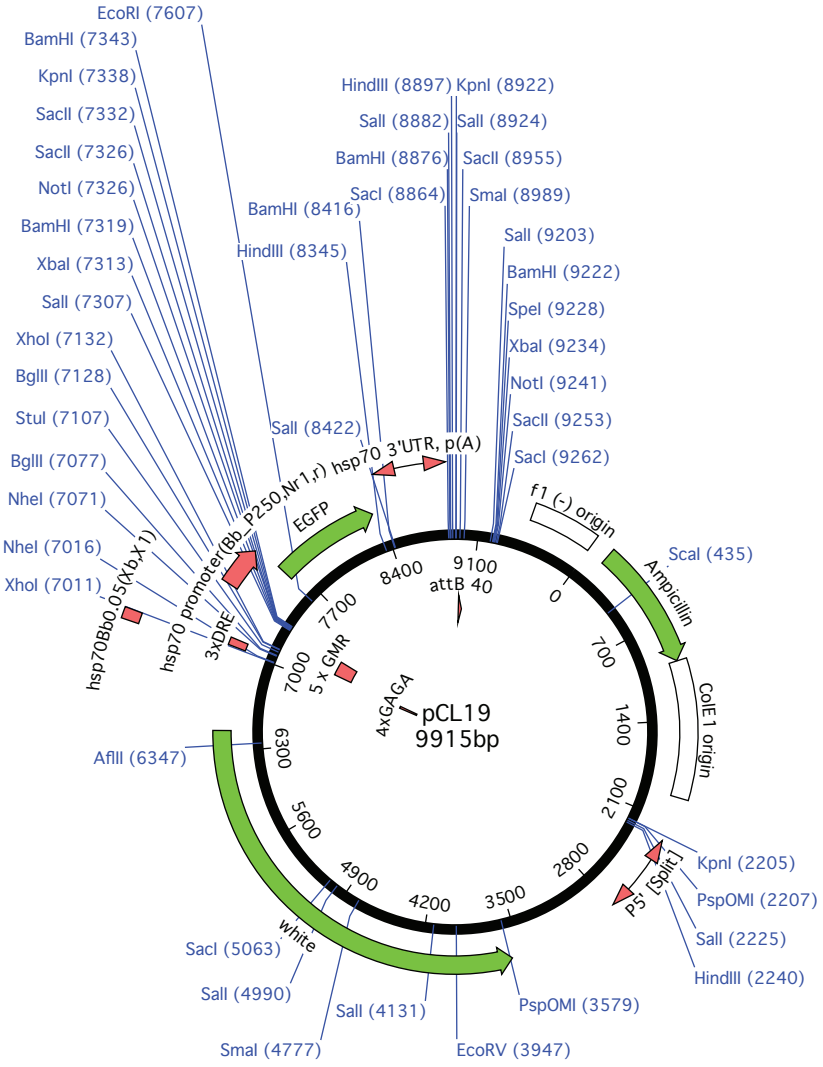


Figure I.40: pCL19

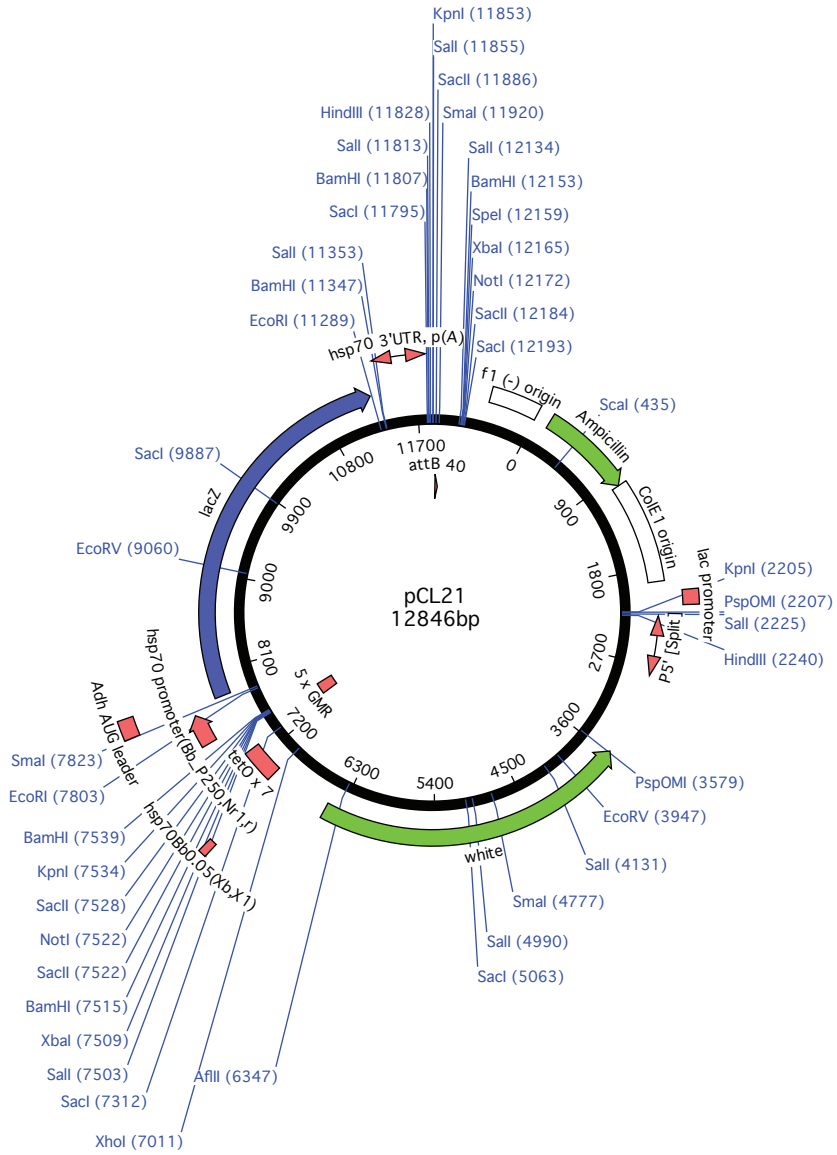


Figure I.41: pCL21

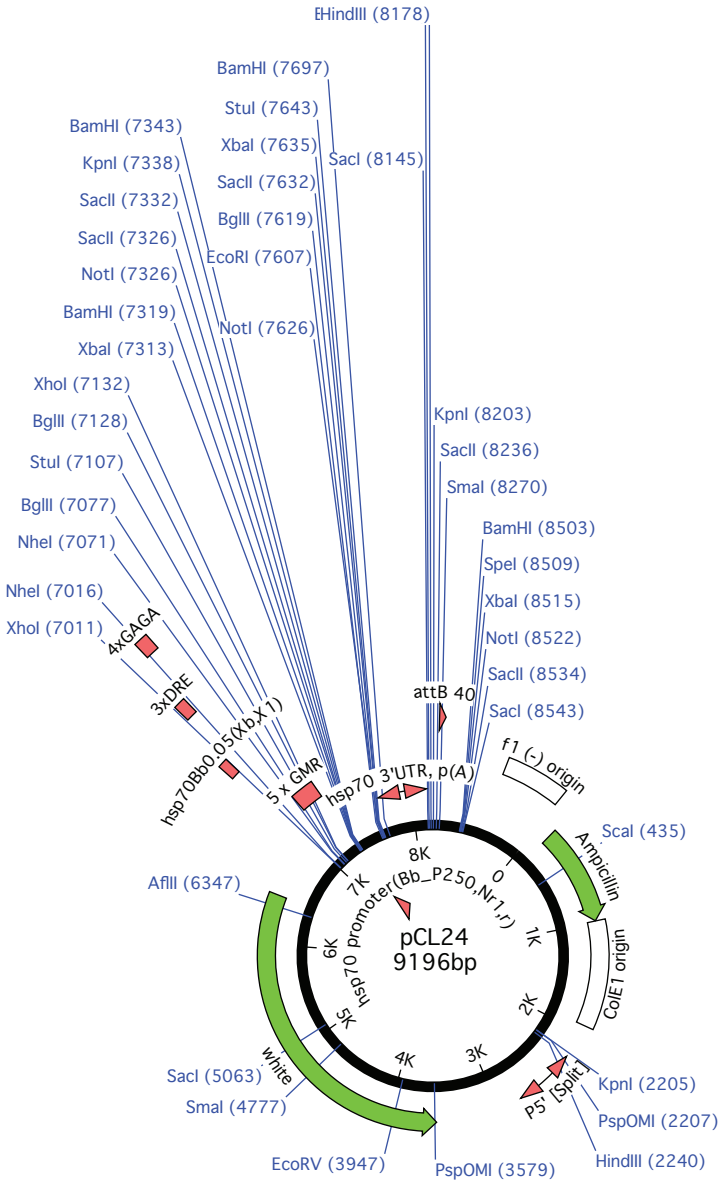


Figure I.42: pCL24. *Sal* I sites removed for clarity.

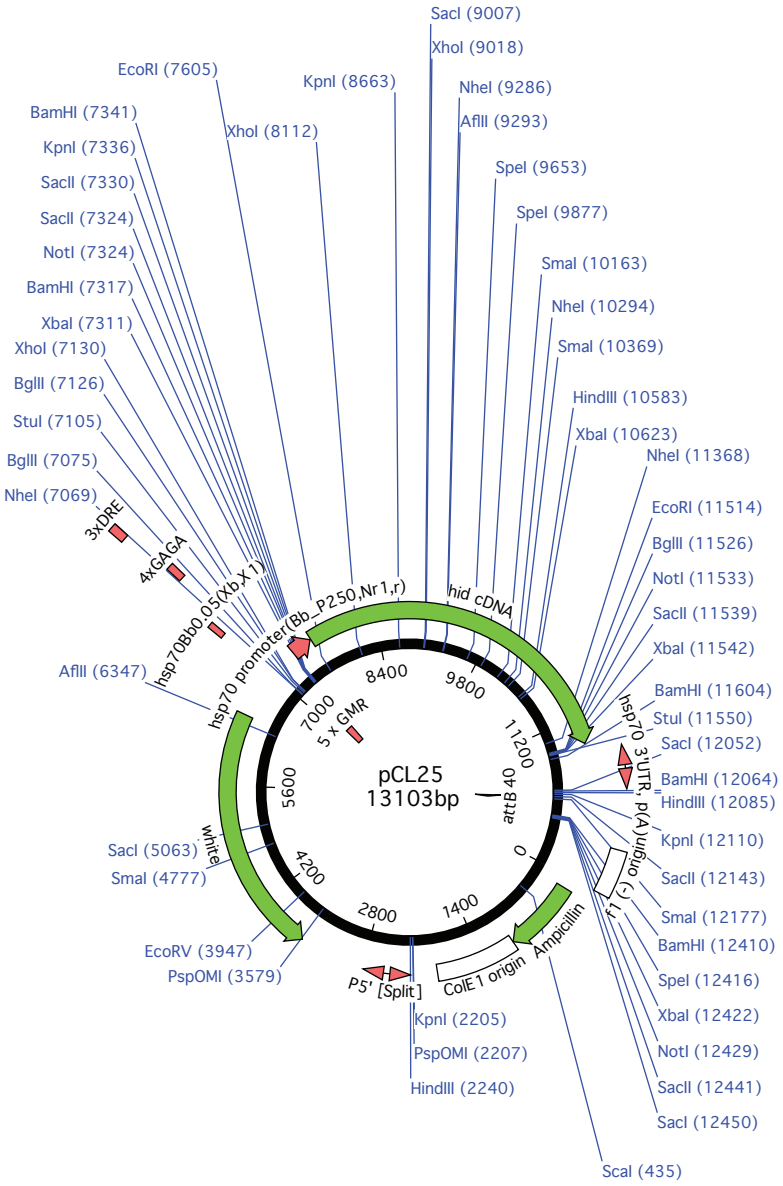


Figure I.43: pCL25. *Sal* I sites removed for clarity.