DIGITALCOMMONS
@WAYNESTATE

Wayne State University

Wayne State University Dissertations

1-1-2011

Rox1 function in dosage compensation: structural / functional analysis of a non-coding rna

Ying Kong
Wayne State University,

Follow this and additional works at: http://digitalcommons.wayne.edu/oa_dissertations

 Part of the Biology Commons, Genetics Commons, and the Molecular Biology Commons

Recommended Citation

Kong, Ying, "Rox1 function in dosage compensation: structural / functional analysis of a non-coding rna" (2011). *Wayne State University Dissertations.* Paper 415.

This Open Access Dissertation is brought to you for free and open access by DigitalCommons@WayneState. It has been accepted for inclusion in Wayne State University Dissertations by an authorized administrator of DigitalCommons@WayneState.

roX1 FUNCTION IN DOSAGE COMPENSATION: STRUCTURAL / FUNCTIONAL ANALYSIS OF A NON-CODING RNA

by

YING KONG

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2012

MAJOR: BIOLOGICAL SCIENCES

Approved by:

_____

Advisor Date

_____

_____

_____

ACKNOWLEDGEMENTS

At the outset I express my most sincere gratitude to my advisor Dr. Victoria H. Meller. She taught me not only how to do my research, but also helped me to develop scientific abilities which will benefit my future career. I am really grateful for her guidance and support in my work and study for all these years.

I thank my other committee members, Dr. Lori Pile, Dr. Miriam Greenberg and Dr. Russell Finley for their patience and advice through my graduate work.

I would also like to thank the past and current members of the Meller laboratory for their assistance and friendship. They create a nice environment to work in that we share our thoughts, experiences, and humors together.

I would like to thank all my friends at Wayne State University. They give me a lot of helps, supports and wonderful moments for my life here which makes me not feel lonely even far away from my hometown.

At last I thank the most important people in my life, my mother Suhua Pan and my father Xiangyue Kong, for their constant, unconditional love, support and understanding during all these years, which become the source of my power to pursue my life and dreams.

TABLE OF CONTENTS

Acknowledgements……..……..……..ii

List of Tables……..…v

List of Figures……..……vi

CHAPTER 1 – INTRODUCTION ……..…..……..….1

CHAPTER 2 – FUNCTION OF *roX* RNAS IN REGULATION OF AUTOSOMAL GENES IN DROSOPHILA

Introduction……..……..…25

Materials and Methods……..……30

Results..……..…40

Discussion……..……..…51

CHAPTER 3 – NONCODING *roX1* RNA 5' SEQUENCES CONTRIBUTE TO X CHROMOSOME LOCALIZATION AND LOCAL SPREADING OF THE MSL COMPLEX

Introduction……..……..…55

Materials and Methods……..……63

Results..……..…75

Discussion……..…100

CHAPTER 4 – FUTURE DIRECTIONS ……..……112

Appendix A. Minimum information about microarray analysis of *roX* in females… ……..……..117

Appendix B. Investigation of a *roX1* transgene insertion reveals female specific lethality……..……..……127

Appendix C. JIL-1 enrichment on the X chromosome is blocked in *roX1 roX2* males……..…… 140

References……..……..……..…..143

Abstract……..……...162

Autobiographical Statement……..……164

LIST OF TABLES

Table 2-1 Primers used for creation and confirmation of *roX1* transgenes, and for confirmation of integration……..…..…….……..……32

Table 2-2 Primers used for Q-RT-PCR analysis of gene expression…..…….36

Table 2-3 Q-RT-PCR validation of microarray analysis……..…….…..43

Table 3-1 Primers used for creation of *roX1* transgenes……..……..…65

Table 3-2 Primers used for comfirmation of *roX1* transgenes and verification of integration of *roX1* transgenes and recombinant chromosomes….. 66

Table 3-3 Primers used to analyze gene expression by Q-RT-PCR……73

Table 3-4 Survival of *roX1 roX2* males that are rescued by $roX1^{R1}$, $roX1^{R2}$ and $roX1^{R3}$ transgenes inserted at 68A4 on third chromosome……..…..78

Table 3-5 Survival of *roX1 roX2* males that are rescued by $roX1^{R1}$, $roX1^{R2}$ and $roX1^{R3}$ transgenes inserted on the X chromosme……...93

Table 3-6 Rescue of *roX1 roX2* $roX1^{R1-3}$ males by autosomal *roX1* transgenes …….99

Table A-1 Rescue of 18D3 lethal insertions by duplications of 18D3……137

Table A-2 Rescue of female-specific lethal insertions by Dp(1;3) DC363….. 138

Table A-3 Rescue of female-specific lethal insertions by Dp(1;3) DC364……139

LIST OF FIGURES

Figure 1-1 Organisms use divergent strategies to compensate sex chromosome gene dosage……...…….2

Figure 1-2 MSL proteins and *roX* RNA form a complex that binds to the *Drosophila* X chromosome……...5

Figure 1-3 *roX1* and *Xist* have distinct regions necessary for gene function…..12

Figure 1-4 Strategies for X chromosome recognition in mammals, flies and worms……....18

Figure 2-1 The crossing scheme for generating mutant and control females with identical genetic background for microarray study……..….34

Figure 2-2 Expression of the fourth chromosome is reduced in *roX1 roX2* males but not in females…….……..……..41

Figure 2-3 Genes situated in proximal heterochromatin require *roX* RNA for full expression in males but not in females……....44

Figure 2-4 A full length *roX1* transgene does not fully rescue autosomal gene expression……..……46

Figure 2-5 Expression of 4th-linked genes is unchanged in females expressing the *roX1*^{c20} transgene…….……...……..……49

Figure 3-1 The crossing scheme for recombining X-linked *roX1*^{R1}, *roX1*^{R2} and *roX1*^{R3} transgenes onto a *roX1 roX2* chromosome ……69

Figure 3-2 Schematic structure of the *roX1* gene and *roX1*^{R1}, *roX1*^{R2} and *roX1*^{R3} transgenes……..……..……76

Figure 3-3 Autosomal *roX1* transgenes partially restore MSL complex localization to the X chromosome…….……..……80

Figure 3-4 MLE immunostaining of polytene chromosomes from male larvae..82

Figure 3-5 *roX1* transcript levels in *roX1 roX2* male larvae carrying *roX1* transgenes…….……..…84

Figure 3-6 Reverse transcription followed by PCR reveals multiple *roX1* transcripts from the *roX1*$^{R1-3}$ transgenes…….……..…86

Figure 3-7 The autosomal integration site of 68A4 enlarged to show MSL binding and spreading from *roX1*^{R1}, and binding but no spreading from *roX1*^{FULL}, *roX1*^{R2} and *roX1*^{R3}……..……88

Figure 3-8 Rescue of X-linked gene expression by *roX1*$^{R1-3}$ transgenes..……89

Figure 3-9 *Beadex* responds to activity of *roX1*$^{R1-3}$ transgenes……..…91

Figure 3-10 Survival of *roX1*^{SMC17A} *roX2 roX1*$^{R1-3}$ males after manipulation of MSL protein levels……..……..……95

Figure 3-11 X-linked *roX1* transgenes partially restore MSL complex localization to the X chromosome……..……97

Figure 3-12 Models of *roX1*^{R1} directed spreading of MSL complex from its transcription site……..……..…..…….105

Figure 3-13 Mobility of the MSL complex depends on *roX1* and MLE…….……...109

Figure 3-14 Schematic illustration of functional *roX1* domains…..…..…..……111

Figure A-1 Annotation of the 18D3 insertion site and the MSL High Affinity Site (HAS) located at this region……..……..…….129

Figure A-2 Structure of viable and lethal integrations at 18D3…….……..130

Figure A-3 Schematic of insertions before and after mobilization by transposase……..……..…….134

Figure A-4 Summary of insertions that show different viabilities……..……135

Figure A-5 JIL-1 is mislocalized in *roX1 roX2* males…….……...…….141

CHAPTER 1

INTRODUCTION

This chapter has been published as: How to get extra performance from a chromosome: recognition and modification of the X chromosome in male *Drosophila melanogaster*. Kong Y. and Meller V.H. (2007) The Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics. Part I Genetics. Section 1.3 Epigenetics. It is reproduced here with updates that reflect recently published material.

Differentiated sex chromosomes cause genetic imbalance

Many organisms have a single X chromosome in males and two in females. The X chromosome is gene-rich and carries genes required in both sexes. By contrast, the Y is often gene-poor, and may carry genes only required in males. The resulting imbalance in the dosage of X-linked genes is fatal if not addressed early in life. Equalization of expression between the sexes is an essential feature of differentiation in flies, mammals, and the worm *Caenorhabditis elegans*. Although the problem is common, the strategy used to solve it in each of these organisms is distinct (Figure 1-1). To equalize expression between *C. elegans* males (XO; one X chromosome but no Y chromosome) and XX hermaphrodites, hermaphrodites reduce transcription from both X chromosomes by 50%. Mammalian females silence most genes on a

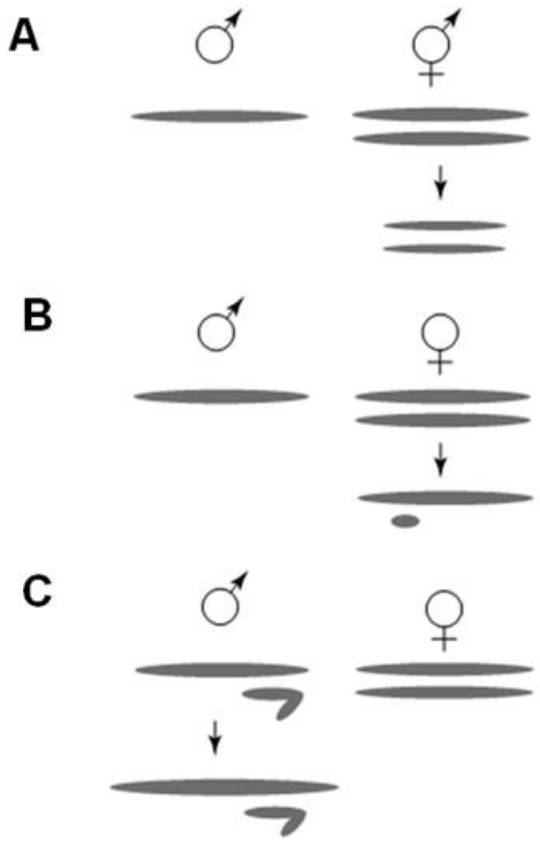

Figure 1-1. Organisms use divergent strategies to compensate sex chromosome gene dosage. (A) *C. elegans* hermaphrodites (right) have 2 X chromosomes, whereas males have a single X chromosome and no Y chromosome (left). Association of the repressive DCC complex reduces expression of both hermaphrodite X chromosomes by about 50%. **(B)** Mammalian females (right) randomly silence a single X chromosome. The remaining active X chromosome is transcriptionally equivalent to the single X chromosome of males. **(C)** *Drosophila* males increase expression from their X chromosome by modulation of chromatin structure (left). Female X chromosomes remain unchanged.

single X chromosome. The remaining X chromosome is transcriptionally equivalent to the single X chromosome of males. By contrast, *Drosophila* males increase expression from their single X chromosome about two fold. Although these methods for equalizing expression are overtly very different, each organism regulates the X chromosome through modulation of chromatin architecture.

These animals must accurately and selectively modulate a single chromosome or a pair of chromosomes in one sex. Interestingly, both mammals and flies use large noncoding RNAs to direct chromatin-modifying proteins that regulate expression. The large *Xist* (*X inactive specific transcript*) is transcribed from and directs silencing to the inactive X chromosome of mammalian females (reviewed by (Plath et al., 2002). As silencing of both X chromosomes would be lethal, this process is restricted to chromatin *in cis* to a single *Xist* allele. *Drosophila* males have two large, noncoding transcripts, *roX1* (*RNA on the X1*), and *roX2* (*RNA on the X2*), that are necessary for localization of a complex of proteins and *roX* RNA to the male X chromosome (Meller and Rattner, 2002). In both mammals and flies, this process likely involves two steps: recruitment of a protein complex, followed by modulation of gene expression. This review will focus on advances in understanding the process of recognition and modulation in flies. It will center on the role of the *roX* transcripts in recognition and modification of the X chromosome.

RNA and protein coat the X chromosome of *Drosophila* males

Many of the genes necessary for dosage compensation in flies were identified through male-specific lethal (*msl*) mutations. These genes, *maleless* (*mle*), the *male-specific lethals 1*, *2*, and *3* (*msl1*, *2*, and *3*), and *males absent on first* (*mof*), are collectively known as the *male-specific lethals* (reviewed by (Mendjan and Akhtar, 2007). Mutations in these genes cause developmental delay and lethality in males, but are not essential in females. The genes that encode the *roX* RNAs are X-linked and functionally redundant for dosage compensation. Both properties make them unlikely to be identified by conventional mutagenesis and phenotypic analysis. Accordingly, the *roX* genes were discovered serendipitously (Amrein and Axel, 1997; Meller et al., 1997). Immunolocalization of MSL proteins or *in situ* hybridization to *roX* on polytene preparations reveals finely banded enrichment along the X chromosome (Figure 1-2). The MSL proteins and *roX* RNA coimmunoprecipitate, demonstrating that they form a complex (Meller et al., 2000; Smith et al., 2000). Removal of individual members of the complex disrupts its localization and can reduce the stability of remaining molecules. This is particularly dramatic for the *roX* RNAs, which are unstable upon elimination of any MSL protein (Amrein and Axel, 1997; Meller et al., 1997). Mutation of *mle*, *msl3*, or *mof* reduces X-chromosome binding by remaining members of the complex, but a subset of sites able to bind the remaining proteins is retained on the X chromosome. The most prominent of these sites are the *roX* genes themselves (reviewed by (Kelley, 2004). MSL1 and MSL2 have a more central role in regulation and assembly of the MSL complex

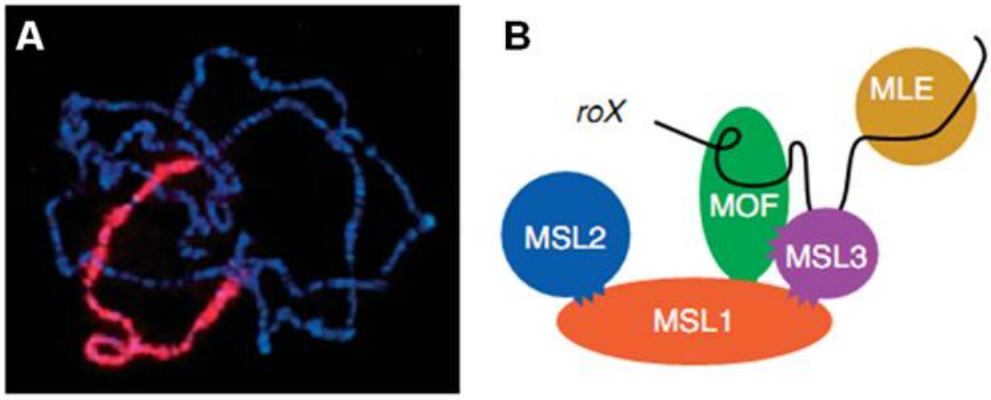

Figure 1-2. MSL proteins and *roX* RNA form a complex that binds to the *Drosophila* X chromosome. (A) Immunodetection of MSL2 on a polytene chromosome preparation from a male larva. The X chromosome binds MSL2, detected with Texas Red. DNA appears blue. **(B)** Molecular interactions between MSL proteins and RNA. Interactions between proteins are denoted by teeth. Potential interactions are modeled between a single *roX* transcript (black line) and proteins reported to have RNA-binding activity. Protein/protein and protein/RNA interactions are reported by Akhtar *et al*. (2000); Buscaino *et al*. (2003); Copps *et al*. (1998); Li *et al*. (2005); Morales *et al*. (2004); Scott *et al*. (2000).

as elimination of either of these proteins prevents all chromatin binding by remaining complex members (Lyman et al., 1997). In contrast, simultaneous elimination of both *roX* RNAs shifts the MSL proteins from the X chromosome to ectopic autosomal sites and results in reduced X-linked gene expression (Deng and Meller, 2006; Deng et al., 2005; Meller and Rattner, 2002). Recognition of the X chromosome is thus a property of the intact MSL complex, and is not attributable solely to a single participating molecule.

Proteins associated with the MSL complex modify chromatin

Increased expression of the male X chromosome is believed to result from changes in chromatin architecture induced by MSL complex. One member of the MSL complex, MOF, is an acetyltransferase specific for lysine 16 of histone H4 (Akhtar and Becker, 2000; Hilfiker et al., 1997; Smith et al., 2000). This modification is generally associated with active chromatin, and is highly enriched on the male X chromosome of flies (Turner et al., 1992). Acetylation of H4K16 by MOF increases transcription *in vitro* and *in vivo* (Akhtar and Becker, 2000). Effector proteins that mediate transcriptional change bind to some modified histones, but none specific for H4K16ac has been found. A study demonstrated that acetylation of H4K16 inhibits the formation of highly compact chromatin by disrupting charge-based internucleosomal interactions (Shogren-Knaak et al., 2006). This structural effect partially decondenses chromatin, thereby increasing the accessibility of the DNA template. In humans, H4K16ac is found ubiquitously on all chromosomes except for the inactive X chromosome (Jeppesen and

Turner, 1993). In flies, H4K16ac is found at the promoter region on both X-linked and autosomal genes in both sexes, but only in males displays a bimodal distribution to promoters and the 3' ends of X-linked genes (Gelbart et al., 2009; Kind et al., 2008). These findings suggest that H4K16ac has a general role in *Drosophila* gene expression. However, the role of MOF, and even whether MOF is responsible for 5' H4K16ac, remains controversial (Gelbart et al., 2009; Kind et al., 2008). On the male X chromosome, H4K16ac colocalizes with the MSL complex in the body of actively transcribed genes (Bone et al., 1994; Gelbart et al., 2009; Smith et al., 2001; Turner et al., 1992). However, the distribution of H4K16ac is much broader than MSL complex (Gelbart et al., 2009). It is supposed that broad H4K16ac distribution is due to transient association of the MSL complex, or chromosome looping that allows the complex to modify distant histones.

A second modification linked to increased expression, phosphorylation of H3 on serine 10 (H3S10p), is also enriched on the male X chromosome (Jin et al., 1999; Mahadevan et al., 2004). H3S10p, produced by the *aurora* kinase, is abundant on mitotic chromosomes (Adams et al., 2001; Giet and Glover, 2001). In interphase cells, H3S10p is directed by the JIL-1 kinase (Wang et al., 2001). Proper dosage compensation of the X-linked *white* gene requires JIL-1 function (Lerach et al., 2005). In addition to compensation of the male X chromosome, JIL-1 has a general role in maintenance of chromatin structure and limits the spread of heterochromatin into euchromatic regions (Zhang et al., 2006). Accordingly, JIL-1 is an essential gene required in both sexes (Wang et al., 2001).

The male X chromosome is therefore marked with at least two histone modifications that are associated with elevated transcription. Recently MSL2 was illustrated as an E3 ubiquitin ligase for histone H2B (Wu et al., 2011). This ubiquitylation then crosstalks with histone H3 K4 and K79 methylation. It is likely that the primary function of the MSL complex is to direct and control these modifications.

The MSL complex increases transcription by a general method

X chromosome compensation affects hundreds of genes with different expression levels and profiles. It must therefore be superimposed on genes with distinct regulatory strategies. Interestingly, chromatin immuno and affinity precipitation of DNA bound by the MSL complex detects modest levels of these proteins in promoter regions, but higher levels within the body of most actively transcribed X-linked genes in males, with a bias towards 3' ends (Alekseyenko et al., 2006; Gilfillan et al., 2006; Legube et al., 2006). H4K16ac has also been found to be high in the body of X-linked genes in males (Gelbart et al., 2009; Kind et al., 2008; Smith et al., 2001). The reduced compaction of chromatin enriched with H4K16ac may increase the speed or processivity of RNA polymerase within the body of genes. Enhanced expression is thus likely to result from facilitation of transcriptional elongation, rather than increased initiation (Henikoff and Meneely, 1993). This hypothesis is verified by a recent genome-wide gene run-on study from isolated nuclei to examine the effect of the MSL complex on RNA Polymerase II (Larschan et al., 2011). However, the possibility

that modifications at the 3' end of transcription units enhances reinitiation by recycled RNA polymerase has not been eliminated (Dieci and Sentenac, 2003).

A second theory for how the MSL complex enhances expression stems from a study of proteins copurifying with tagged MOF and MSL3 (Mendjan et al., 2006). This study found no classical transcriptional factors but identified exosome components, interband binding proteins and nuclear pore components in association with MSL proteins. Knock down of the nuclear pore proteins Mtor and Nup153 disrupted the location of MSL proteins and compensation of some X-linked genes, suggesting that interaction with the nuclear pore is important for localization of the MSL complex (Mendjan et al., 2006). An association with the nuclear pore might facilitate transcriptional elongation by affecting RNA processing and transportation. Alternatively, the nuclear pore might establish a transcriptionally active compartment, or a region of facilitated chromatin remodeling within the nucleus (Casolari et al., 2005; Feuerbach et al., 2002). Tethering transcription units to nuclear pores facilitates expression in yeast, supporting the involvement of this structure in activation (Cabal et al., 2006; Taddei et al., 2006). Recruitment of silenced genes into a repressive nuclear compartment has been proposed as the mechanism of X-chromosome inactivation in mammals. The inactive X chromosome (Xi) occupies a region adjacent to the nucleolus during replication (Zhang et al., 2007). This may ensure epigenetic maintenance of the silent state through replication. X-linked gene inactivation is accompanied by movement of individual genes from the outer fringe of the domain occupied by the Xi to a more interior position from which

RNA polymerase is excluded (Chaumeil et al., 2006). Thus X inactivation and its perpetuation may rely on recruitment of genes and the X chromosome into specific nuclear compartments. Recent studies suggest that dosage compensation in flies may involve repositioning of genes in a transcriptionally active domain (Grimaud and Becker, 2009). Dosage compensated genes on the X chromosome were closer together in interphase nuclei from males than in those from females. This was dependent on MSL proteins, indicating that these proteins promote a male-specific conformation of the X chromosome. The idea of a dosage compensated nuclear compartment is appealing as it provides a mechanism through which compensation can be superimposed on genes with different regulatory strategies on one chromosome.

Large RNAs that control X chromosomes: powerful but mysterious molecules

Regulatory RNAs that coat the X chromosome play a key role in dosage compensation in mammals and *Drosophila*. In spite of the central role of *roX* transcripts in fly dosage compensation, how they interact with the MSL proteins, and how this changes the properties of the MSL complex, remains speculative. A comparison with *Xist* may prove valuable. *Xist* is well studied and shares unusual properties with the *roX* RNAs. Both RNAs coat dosage-compensated X chromosomes, direct protein complexes to chromatin, and are able to recruit chromatin-modifying activities *in cis* to the site of RNA synthesis.

Xist is transcribed from the *Xic* (*X inactivation center*) and is essential for initiation and propagation of X-chromosome inactivation in mammals (reviewed by (Plath et al., 2002). *Xist* is selectively expressed from one X chromosome and spreads *in cis* from the site of synthesis to coat most of the inactive X chromosome (Xi). *Xist* recruits Polycomb proteins that introduce repressive histone modifications (Plath et al., 2004; Silva et al., 2003). Several days after the initiation of X inactivation, the inactivation becomes largely independent of *Xist*. Additional changes in chromatin, such as enrichment for variant histones and methylation of CpG islands, characterize the differentiated Xi (reviewed by (Lucchesi et al., 2005). Distinct sequences within *Xist* are responsible for localization to the X chromosome and for silencing (Figure 1-3 A, (Wutz et al., 2002)). Several widely separated *Xist* sequences act cooperatively to direct X localization, and a repeated element that folds into short stem loops mediates silencing as well as localization (Wutz et al., 2002). This repeat is also necessary for relocalization of silenced genes inside the domain occupied by the Xi (Chaumeil et al., 2006).

roX1 also has multiple regions necessary for function (Figure 1-3 B). In spite of their redundancy, *roX1* and *roX2* share little sequence similarity that can be used to identify potentially important sequences (Amrein and Axel, 1997; Park et al., 2003). Several highly conserved "*roX* boxes" (short GUUNUACG sequences) are present at the 3' end of both *roX* RNAs (Franke and Baker, 1999; Park et al., 2008). A stem loop at the 3' end of *roX1* is formed in part by *roX* box sequences. These conserved 3' "*roX* boxes" are able to form alternative stem

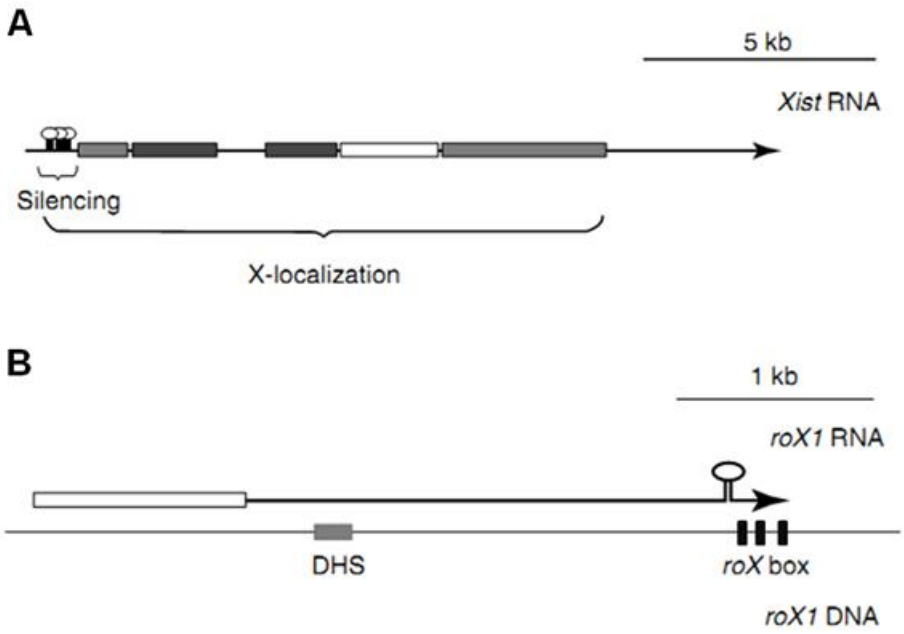

Figure 1-3. *roX1* and *Xist* have distinct regions necessary for gene function. **(A)** The *Xist* transcript has a series of 15 short stem loops near the 5' end that are necessary for silencing. Distributed elements that contribute to X localization are shown as gray and white boxes. The strongest of these are darkest. Figure is based on (Wutz, 2003; Wutz *et al*., 2002). **(B)** Functional and conserved regions of *roX1* RNA (top) and DNA (bottom) are represented. One kb at the 5' end of *roX1* (open box on left) is necessary for wild-type localization of the MSL complex. Between this and the 3' stem loop (right) there is no identified element necessary for RNA function. The 200 bp *roX1* DNase hypersensitive site (DHS) is shown as a gray box on the *roX1* DNA. This sequence attracts the MSL complex. The "*roX* boxes" (black) are at the right. This is based on (Kageyama *et al*., 2001; Park *et al*., 2003; Stuckenholz *et al*., 2003; Park et al., 2008).

loops (Kelley et al., 2008; Park et al., 2008). Transgenes deleted for the stem loop or with disrupted pairing of the stem have low rescue of *roX1 roX2* males in spite of substantial recruitment of MSL proteins to the X chromosome (Stuckenholz et al., 2003). The stem loop is therefore expected to influence chromatin modification or gene activation by the MSL complex, rather than X chromosome targeting. This is supported by studies revealing that the histone acetyltransferase activity of the MSL complex requires elements in the 3' end of *roX1* (Kelley et al., 2008; Park et al., 2008; Stuckenholz et al., 2003). A weakly conserved 200 bp sequence within each *roX* gene strongly attracts the MSL proteins and forms a male-specific DNaseI hypersensitive site (Kageyama et al., 2001; Park et al., 2003). The *roX1* DNaseI hypersensitive site (DHS) acts as an enhancer of *roX1* transcription in males and a repressor in females (Bai et al., 2004). However, internal deletions of *roX1* lacking the DHS are still regulated in a sex-specific manner, and *roX1* alleles and transgenes lacking this sequence retain full activity (Deng et al., 2005; Rattner and Meller, 2004; Stuckenholz et al., 2003). Although the role of the DHS remains speculative, all evidence points to its function as DNA, rather than RNA.

Deletion of 1 kb at the 5’ end of *roX1* also destroys activity. When large portions of this region are removed by internal deletion, *roX1* activity is reduced commensurate with the amount deleted (Deng et al., 2005). Small (<300 bp) deletions scanning the 5’ end have failed to identify discrete elements, suggesting redundancy (Stuckenholz et al., 2003). Males that carry a *roX1* allele with a large part of the 5’ end missing display ectopic MSL binding and reduced

coverage of the X chromosome, suggesting that this region is necessary for recognition of the X chromosome (Deng et al., 2005). An internal deletion of 2.4 kb that retains 0.8 kb of 5' end and 0.6 kb of the 3' end, including the stem loop, supports full male survival (Deng et al., 2005). However, simultaneous expression of separate 3' and 5' fragments of *roX1* does not rescue either MSL localization or male survival (Meller and Rattner, 2002; Stuckenholz et al., 2003). Taken together, these observations suggest that *roX* activity requires simultaneous interaction with different molecules. An attractive model is that *roX1*, like *Xist*, has distinct domains necessary for X chromosome localization and gene activation.

The major *roX2* splice form is 600 bp, but a multitude of alternative *roX2* splice forms with decreased activity has been found (Park et al., 2005). *roX2* molecules with different levels of activity may modulate the activity of the MSL complex, thus fine-tuning the level of X chromosome activation.

MSL proteins have RNA-binding activity

MLE is an ATP-dependent RNA/DNA helicase with higher activity on RNA substrates (Lee et al., 1997). The ATPase activity is necessary for transcriptional enhancement by the MSL complex (Morra et al., 2008). The helicase activity of MLE is essential for normal localization of the MSL complex on the X chromosome, and for movement of the *roX* RNAs from their sites of synthesis, suggesting that MLE may integrate *roX* into the mature MSL complex (Gu et al., 2000; Meller et al., 2000; Morra et al., 2008). MLE itself does not interact with

other MSL proteins and can only be coimmunoprecipitated under nonstringent conditions using antibodies that pull down other MSL proteins (Smith et al., 2000). MLE can be released from polytene chromosomes by RNase A digestion, suggesting that it associates with the MSL complex through an RNA (Figure 2B; (Richter et al., 1996)). The stability of *roX1* is particularly dependent on MLE, supporting the idea of a direct interaction between these molecules (Meller, 2003).

Both MSL3 and MOF have RNA-binding activity *in vitro* and their localization on the X chromosome is destabilized by RNase digestion (Akhtar et al., 2000; Buscaino et al., 2003). Both proteins have variant chromo domains that have been implicated in RNA binding. Whereas the canonical chromo domains of Heterochromatin protein 1 (HP1) and Polycomb (Pc) contain aromatic residues in a recognition site that binds methylated histones, the variant structures found in MOF and MSL3, named *chromo barrel domains*, may have different functions (Bannister et al., 2001; Lachner et al., 2001). The MOF chromo barrel domain lacks particular residues that recognize methylated peptides (Nielsen et al., 2005). This region contributes to MOF's ability to bind RNA *in vitro* (Akhtar et al., 2000). The chromo barrel domain of MSL3 has also been implicated in RNA binding, but retains the aromatic residues necessary for methyl group binding (Nielsen et al., 2005). A previous study showed that mutation of the MSL3 chromo barrel domain prevents increased transcription of X-linked genes, but does not affect localization of the complex to the X chromosome (Buscaino et al., 2006). However, a later study showed that the MSL3 chromo barrel domain binds

histone H3 Lysine 36 trimethylation (H3K36me3), and this is essential for MSL spreading into transcribed regions of the X chromosome (Sural et al., 2008). The failure of Buscaino *et al*. to see any effect of mutation of the MSL3 chromo barrel domain on binding of the complex to the X chromosome may due to the low resolution of polytene chromosome immunostaining. The decreased expression of X-linked genes in the MSL3 chromo barrel domain mutant is likely the result of failure of the MSL complex to spread onto active genes.

The H4 acetyltransferase activity of MOF is greatly increased by association with MSL1 and MSL3, suggesting a mechanism for limiting MOF activity until it assembles with a regulatory complex (Morales et al., 2004). MSL1, 2, 3, and MOF continue to associate in the absence of *roX*, but only low levels of H4K16ac are detected at sites bound by these proteins (Deng and Meller, 2006). This may reflect a reduced MOF activity in the absence of *roX* RNA. Taken together, these studies suggest that multiple RNA/protein and protein/protein contacts within the MSL complex are necessary for precise regulation of the activity of the MSL complex.

Do *roX* and MLE recruit a preexisting chromatin-binding complex?

The discovery that the yeast NuA4 transcriptional regulator contains subunits similar to MSL3 and MOF, and characterization of a mammalian complex containing MSL homologs suggests that the association of these proteins is ancient (Eisen et al., 2001; Marin and Baker, 2000; Smith et al., 2005; Taipale et al., 2005). Human MOF (hMOF) participates in multiple protein

assemblies and is required for normal function of human ATM (ataxia-telangiectasia-mutated) protein in DNA repair (Gupta et al., 2005; Taipale et al., 2005).

roX RNAs have only been identified in closely related *Drosophilids*, but helicases with similarity to MLE have been identified from yeast to mammals (Park et al., 2003; Sanjuan and Marin, 2001). MLE homologs have not yet been isolated in complexes of MSL-like proteins outside of flies. MLE has a peripheral association with the fly MSL complex, and is presumably tethered by RNA. It thus seems plausible that the addition of MLE to the MSL complex depends on the presence of *roX*. The importance of *roX* in correct targeting of the MSL complex suggests that the addition of noncoding RNA was a major step in recruitment for the purpose of X chromosome compensation.

Recognition of X chromosomes

A mechanism that targets changes in expression to a single chromosome is a fundamental requirement of dosage compensation. Two distinct strategies for accomplishing this have been described. The chromosome may be controlled by a *cis*-acting element. The *Xic* is a strong *cis*-acting element capable of directing silencing to the entire X chromosome (Figure 1-4 A). The *Xic* can also silence autosomal chromatin if inserted on an autosome (Lee and Jaenisch, 1997; Lee et al., 1996; Wutz and Jaenisch, 2000). *Xist* RNA produced from the *Xic* does not work *in trans*, thus protecting one X chromosome from inactivation. An alternative mechanism for distinguishing a chromosome is finely dispersed

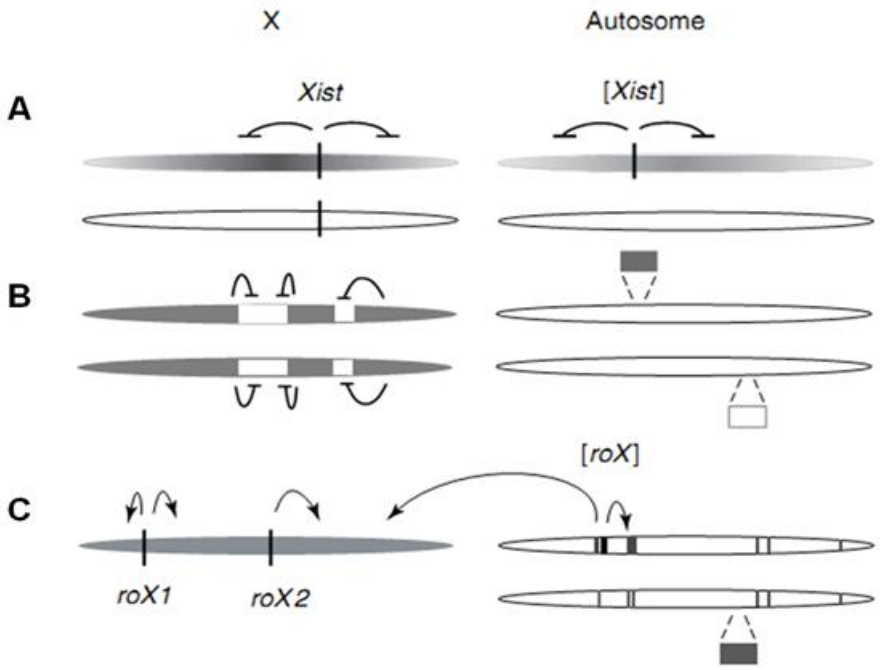

Figure 1-4. Strategies for X chromosome recognition in mammals, flies and worms. (A) One of two *X inactivation centers* present in females produces *Xist* RNA (top left). The chromosome carrying this allele becomes the inactive X (shaded). A transgene carrying *Xist* can silence autosomal chromatin *in cis* (shaded, right). **(B)** The *C. elegans* X chromosome is distinguished by sequence elements (gray shading). The distribution of these elements is uneven, leaving large gaps (white). The repressive DCC spreads into these gaps from flanking regions. Segments separated from the rest of the X chromosome attract the DCC if they have X-recognition sequences (autosomal insertion, top) but remain uncompensated if they lack these elements (autosomal insertion, bottom). **(C)** The *Drosophila* X chromosome is finely marked by sequences that attract the MSL complex (gray). Translocated X chromosome fragments are recognized accurately (autosomal insertion, bottom). Weak and scattered MSL-binding sites on the autosomes do not attract the MSL complex in normal males (gray lines, right). *roX1* and *roX2* (vertical black lines) produce *roX* RNA and are *cis*-acting elements that enhance recognition of the X chromosome. A *roX* transgene (top right) enables MSL binding to closely linked autosomal sites. The *roX* transgene also produces transcript that acts *in trans* to compensate an X chromosome.

sequence elements. Two short sequences that participate in recognition of the *C. elegans* X chromosome have been identified (McDonel et al., 2006). Interestingly, these are not exclusive to the X but are found near one another on the X chromosome. This suggests that cooperativity between multiple DNA-binding molecules underlies recognition X chromatin in worms. However, large regions of the *C. elegans* X chromosome fail to bind the repressive dosage compensation complex (DCC) when separated from the X chromosome, but are coated by it when on the X chromosome (Figure 1-4 B) (Csankovszki et al., 2004). This indicates that the ability of the DCC to spread *in cis* is necessary for a complete coverage of the *C. elegans* X chromosome.

Translocated segments of the *Drosophila* X chromosome are faithfully recognized by the MSL complex, indicating the presence of finely distributed sequences marking this chromosome (Figure 1-4 C, (Fagegaltier and Baker, 2004; Oh et al., 2004)). In addition, autosomal *roX* transgenes can fully rescue male viability, indicating that *roX* RNA can act *in trans* to its site of synthesis. But under some conditions, autosomal *roX* insertions also direct MSL binding to chromatin flanking the insertion site (Kelley et al., 1999; Park et al., 2002). Regional spreading of chromatin modification from the *roX* genes can also be observed on the X chromosome (Bai et al., 2007; Oh et al., 2003). It therefore appears that recognition of the *Drosophila* X chromosome involves strong, *cis*-acting elements as well as sequences identifying the X chromosome.

Subdivision of DNA clones that recruit the MSL complex and a functional assay for MSL recruitment have identified short sequences that contribute to

MSL binding (Dahlsveen et al., 2006; Gilfillan et al., 2007; Oh et al., 2004). These sequences are divergent and display a wide range in affinity for the MSL proteins. More recently, ChIP-chip and ChIP-seq studies identified over one hundred MSL High Affinity Sites (HAS), also called Chromatin Entry Sites (CES) on the X chromosome (Alekseyenko et al., 2008; Straub et al., 2008). These are regions that retain residual MSL binding in an *msl3* mutant background. A GA-rich MSL recognition element (MRE) has been identified within the CES (Alekseyenko et al., 2008). However, the MRE consensus is only two-fold enriched on the X chromosome, compared to the autosomes (Alekseyenko et al., 2008; Straub et al., 2008). Many MREs on the X chromosome do not reside within a CES. Therefore, the MRE is not the only feature that forms a CES or distinguishes X and autosomal chromatin. An attractive hypothesis is that a dense distribution of MREs and weak recruitment sites act cooperatively to mark the X chromosome (Dahlsveen et al., 2006; Demakova et al., 2003; Fagegaltier and Baker, 2004). Local elevation of the MSL complex by strong sites will enable weaker ones to be bound. The DHS within the *roX1* and *roX2* genes are extraordinarily strong MSL recruitment sites (Kageyama et al., 2001). Both *roX* genes can induce binding of the MSL complex in autosomal chromatin flanking transgene insertions, likely due to enhancement of weak autosomal binding sites (Dahlsveen et al., 2006; Kelley et al., 1999). Although not absolutely essential for compensation, the situation of the *roX* DHS on the X chromosome will enhance recognition of the X. Additional features, such as features linked to actively expressed genes might be also important to distinguish functional MREs on X chromosome.

X chromosome binding of the MSL proteins is disrupted in *roX1 roX2* males, but these proteins continue to colocalize at ectopic autosomal sites. The *roX* transcripts are therefore not essential for chromatin binding, but ensure high selectivity of the intact MSL complex for the X chromosome. Assembly with *roX* might enhance the ability of the MSL complex to recognize sequence elements on the X chromosome. Alternatively, a change in the complex following *roX* incorporation could promote cooperative binding at closely situated sites. This would favor the X chromosome, proposed to have dense mix of strong and weak sites, over the autosomes, which have more scattered sites capable of recruiting MSL proteins (Demakova et al., 2003).

The distribution of the MSL complex in the body of active genes suggests that localization is largely established by transcriptional activity (Alekseyenko et al., 2006; Gilfillan et al., 2006; Legube et al., 2006). This could occur by association with the transcription machinery or interaction with nascent transcripts. Alternatively, the MSL complex could be targeted to modified histones in the wake of a transcribing polymerase. The MSL3 chromodomain binds histone H3 Lysine 36 trimethylation (H3K36me3), which marks active genes, supporting this mechanism (Sural et al., 2008). These methods would identify transcribed regions, but are unable to distinguish between X-linked and autosomal genes. The current model for establish MSL binding is a two step model: first, sequence-dependent MSL recognition of specific chromatin entry sites on the X chromosome occurs, followed by sequence-independent spreading to active genes. Closely linked genes will be transcribed in proximity to

one another, and thus may be influenced by their neighbors. Linked mammalian genes often associate at “transcription factories” (Osborne et al., 2004). Although there is no evidence for analogous transcription factories in *Drosophila*, identification of nuclear pore proteins in association with the MSL complex suggests recruitment to a particular region of the nucleus. It is possible that some elements marking the X chromosome direct transcribed genes to regions where MSL loading can occur, rather than interacting directly with the MSL complex itself. For example, a male-specific conformation of the X chromosome that clusters the CES may facilitate distribution of the MSL complex along the X chromosome (Grimaud and Becker, 2009).

Concluding remarks

Dosage compensation in *Drosophila* is a remarkable model for the study of epigenetic regulation. It is also rich in common principles of chromatin-based transcription control. Regulation of the male X chromosome involves histone modifications that are proposed to act by increasing the speed, processivity, or reinitiation rate of RNA polymerase. Similar mechanisms will be relevant for the regulation of all eukaryotic genes. Modification of the *Drosophila* X chromosome is directed by cues including *cis*-acting DNA elements, transcriptional activity and possibly recruitment to regions where MSL loading is promoted or transcription is facilitated. Together these produce highly selective recognition and modulation of a single chromosome. Understanding how noncoding RNAs such as *roX*

coordinate this process will enhance our understanding of epigenetic processes in all eukaryotes.

Project outline

Non-coding RNAs are postulated to be involved in gene regulation in many different ways. One of the important ways is the involvement of *roX* RNA in the chromosome-wide gene regulation that occurs during dosage compensation in *Drosophila*. In this dissertation, I present my work on the structural and functional analysis of *roX1* RNA. In Chapter 1, I introduce background information about dosage compensation, the MSL complex, recognition of X chromosome and how the MSL binding pattern is established.

In Chapter 2, I examine whether *roX* RNA contributes to the expression of autosomal genes in *Drosophila* females. I show that *roX1* has no detectable affect on gene expression in females, although present in female embryos. I also analyze candidate regions suspected of involvement in regulating autosomal genes.

In Chapter 3, I dissect the function of the 5' end of *roX1*, a region critical for localization of the MSL complex. I describe the generation and analysis of *roX1* transgenes containing different sections of the 5' end. I show that multiple redundant elements contributing to X chromosome targeting are present throughout the 5' end of *roX1.* I also show that the extreme 5' *roX1* region has a unique function that promotes MSL complex spreading from sites of *roX1* transcription. Paradoxically, a transgene containing this region is unable to recruit

MLE to the X chromosome. I then show that all of the 5' *roX1* transgenes display different stability, but all can partially restore X-linked gene expression in *roX* mutant. *roX1* has multiple transcription start sites that produce several isoforms. My work suggests that the choice of transcription start site will have a striking effect on *roX1* function.

Data from these studies provide initial insight into the mechanism of a chromosome-wide regulation by a large non-coding RNA. Our results generate some interesting questions. Chapter 4 is a summary of my study and perspectives for future study that are based on my findings.

CHAPTER 2

FUNCTION OF *roX* RNAS IN REGULATION OF AUTOSOMAL GENES IN DROSOPHILA

Part of the work described in this chapter has been published in 'Deng X., Koya S.K., Kong Y., and Meller V.H. (2009) Coordinated regulation of heterochromatic genes in *Drosophila melanogaster* males. Genetics. 182(2): 481-491.'

INTRODUCTION

Many organisms have two X chromosomes in females but only one X chromosome and a Y chromosome in males. The X chromosome is gene-rich and majorly euchromatic. By contrast, the Y chromosome is often gene-poor, and largely heterochromatic (Charlesworth, 1991). Highly differentiated sex chromosomes result in imbalance of the dosage of X-linked genes that is fatal if not addressed early in life. Dosage compensation is the process which equalizes expression of X-linked genes between males and females. Although strategies for equalizing expression differ between species, there is a unifying theme which involves a coordinated regulation of the whole X chromosome through modulation of chromatin architecture. This is mediated by selective recruitment of chromatin modifying proteins (Lucchesi et al., 2005). In *Drosophila*, the male specific lethal (MSL) complex is the chromatin modifying complex which

selectively binds the male X chromosome and increases expression from the male X chromosome about two fold to achieve the dosage compensation.

The MSL complex is composed of both proteins and noncoding RNAs. Two noncoding RNAs, *roX1* and *roX2* (RNA on the X 1 and -2), are essential components of MSL complex (Meller and Rattner, 2002). *roX* RNAs assemble with five MSL proteins, the Male Specific Lethals 1, -2, and -3 (MSL1, MSL2, and MSL3), Maleless (MLE) and Males Absent On First (MOF), to form the complex. The MSL complex binds numerous sites on the X chromosome and acetylates histone H4 on lysine 16 (H4K16ac), which ultimately increases expression of the whole X chromosome (Akhtar and Becker, 2000). *roX1* and *roX2* RNAs are functionally redundant. Deletion of a single *roX* gene shows no phenotype, but simultaneous mutation of both *roX* genes is lethal in males (Meller and Rattner, 2002). A microarray study of *roX1 roX2* male larvae revealed a global decrease of X-linked gene expression, which confirmed that *roX* RNAs are necessary for full expression of dosage compensated genes in males (Deng and Meller, 2006).

The highly specific binding of the MSL complex to the X chromosome suggested the idea that the MSL components contribute solely to X-linked gene expression. However, expression of virtually all genes on the small fourth chromosome also decreased in the absence of both *roX* transcripts (Figure 2-2A, B, (Deng et al., 2009)). This finding suggested that there are other functions of MSL components beyond dosage compensation. In agreement with this, some of the MSL proteins act as general transcriptional regulators outside of the complex. Both MLE and MOF are found at autosomal sites of active transcription in males

and females (Kind et al., 2008; Kotlikova et al., 2006). But their role as general transcriptional regulators is not essential, as mutations in *mle* and *mof* are not lethal in females.

The fourth chromosome of *Drosophila* is a tiny chromosome, containing only around 100 genes. It is enriched for heterochromatin, consisting of interspersed heterochromatic and euchromatic regions (Locke et al., 1999; Riddle and Elgin, 2006). Many fourth-linked genes contain these heterochromatic regions. Genes within or near heterochromatin have specialized regulatory features that enable their expression in this repressive environment (Yasuhara and Wakimoto, 2006). It is possible that the heterochromatic feature of the fourth chromosome contributes to its regulation by *roX* RNAs. Studies in our lab confirmed that *roX* RNAs are required for full expression of heterochromatic genes in male larvae, including those on the fourth chromosome and in heterochromatic regions of the second and third chromosomes (Figure 2-3A, (Deng et al., 2009)).

Investigation of the genetic basis of heterochromatic regulation revealed that MSL1, MSL3 and MLE, but not MSL2, are also necessary for normal expression of heterochromatic genes in males (Deng et al., 2009). MSL2 is the only member of the MSL complex that is only expressed in males, thus limiting MSL complex assembly to males (Rastelli et al., 1995). Regulation of heterochromatic genes is therefore not mediated by the intact MSL complex. All of the MSL proteins necessary for regulation of heterochromatic genes are present in females (Rastelli et al., 1995). *roX1* is also present in female embryos

(Meller, 2003). Since females have *roX1*, and the MSL proteins necessary for regulation of heterochromatic genes, I asked if *roX* RNA also regulates heterochromatic gene expression in females. To do this, I compared genome-wide expression in *roX1 roX2* females with that of control females. I discovered that *roX* RNAs do not affect expression of the female fourth chromosome and heterochromatic genes, revealing that regulation of these genes by *roX* RNAs is male-limited.

What limits *roX*-dependent regulation of heterochromatic genes to males is still mysterious. Our preliminary data suggested a candidate *roX1* region that might contribute to regulation of heterochromatic genes. I tested this by comparing gene expression in *roX1 roX2* males carrying transgenes with and without this region. I found that this region has no influence on the expression of heterochromatic genes. Sexlethal (SXL) is a multifunctional RNA-binding protein that controls sexual differentiation in flies. The SXL protein is only expressed in females and determines the sexual development by modulating the expression of a set of downstream genes (reviewed in (Salz and Erickson, 2010)). It is expected that all sex-limited processes will be directly or indirectly regulated by SXL. SXL inhibits expression of MSL2 in females (Bashaw and Baker, 1997; Kelley et al., 1997; Penalva and Sanchez, 2003). MSL2 is, in turn, essential for *roX* expression and stability in older animals, but not in embryos (Rattner and Meller, 2004). *roX1* has two potential SXL binding sites in a small intron (Figure 2-5A, (Meller et al., 1997)). The function, if any, of these sites is unknown, but it is possible that they bind SXL in females, limiting the ability of *roX1* to regulate

heterochromatin. I examined the possibility that SXL binding to *roX1* inhibits its function in female heterchromatin by comparing gene expression in females carrying transgenes lacking the SXL binding site to wild type females. I demonstrated that deletion of the SXL binding site did not enable *roX1* to regulate heterochromatic genes in females.

MATERIALS AND METHODS

Drosophila strains

Flies were maintained in vials containing standard cornmeal agar media at 25°C. The *roX1*^{SMC17A} and *roX1*^{ex33A} mutations have been previously described (Deng et al., 2005). Df(1)52 is a lethal deletion of *roX2* and essential flanking genes. Elimination of *roX2* is accomplished by complementing Df(1)52 with [w^{+}4Δ4.3], a cosmid insertion carrying essential deleted genes but lacking *roX2* (Meller and Rattner, 2002). For convenience this combination is referred to as *roX2*.

Generation of a full length *roX1* expression construct

To create a full length *roX1* transgene, the 5' end of a slightly truncated 4.1 kb *roX1* genomic fragment (*roX1* bp 42-4165, numbering from (Amrein and Axel, 1997) is used throughout) was replaced by a fragment containing the entire 5' end of *roX1* (starting from *roX1* bp -78). The missing 5' fragment was generated by PCR amplification of genomic DNA from wild type flies using oligonucleotides *roX1* F-78 and *roX1* R1262 (Table 2-1), and subcloned into pCR4-TOPO (Invitrogen). A BamHI site is present in *roX1* F-78 and a MluI site is present in *roX1* R1262. A 1.3 kb *roX1* BamHI-MluI fragment from the pCRT-TOPO clone was used to replace a 1.2 kb BamHI-MluI fragment from the existing *roX1* clone. The full length *roX1* was released by EcoRI and ligated into a modified pUASTB. pUASTB is a transformation vector carrying an *attB* site and a mini-*white* gene

(Groth et al., 2004). pUASTB was modified by introducing the strong and constitutive *hsp83* promoter. The *hsp83* promoter was amplified from the pCasper HS83T3 vector (Meller et al., 2000) by PCR (primers in Table 2-1). A KpnI site was introduced at the 5' end and BamHI, EcoRI and SacII sites were introduced at the 3' end. A KpnI-SacII fragment containing the *hsp83* promoter was subcloned in the pUASTB vector. The unique BamHI and EcoRI sites were used to clone *roX1* adjacent to the promoter.

Generation of transgenic flies

The full length *roX1* transgene was integrated at 16C1 on the X chromosome and at 68A4 on the third chromosome using the site-specific ΦC31 integrase system (Groth et al., 2004). Stocks with an insertion of an *attP* landing site P[y^{+}Cary*attP*] at 68A4 were obtained from the Bloomington *Drosophila* stock center. Mobilization of this P element was accomplished by introducing a transposase source (TMS, Sb P[$ry^{+t7.2}\Delta$2-3]99B) and selecting for insertions on the X chromosome (P. Frolov, unpublished). The exact location of the 16C1 X-linked insertion was determined by inverse PCR and sequencing (Inverse PCR Protocol available from Berkeley Drosophila Genome Project; http://www.fruitfly.org/about/methods/inverse.pcr.html). My construct was injected into *Drosophila* embryos derived from mating female virgins carrying P[y^{+}Cary*attP*] with males having a ΦC31 integrase transgene on the third chromosome. Adult flies derived from injected embryos were mated to *yw*, and transgenics expressing the mini-*white* marker identified in their offspring. Transgenics were made homozygous,

Table 2-1. Primers used for creation and confirmation of *roX1* transgenes, and for confirmation of integration.

Primers	Sequences (5' – 3')
roX1 F-78	AGGATCCGAATTCGCGGCGTTACCGGCTCG
roX1 R1262	GTACGCGTCTTCTCGAAACGCAAGTGGCGA
hsp83 Link F	ATGGTACCGATGATCCTTAACGGGGAACTTG
hsp83 Link R	AGCCGCGGGAATTCGGATCCTCGACGGTATCGATAAGCTAG
hsp83 662F	TTCGGACCACTTAGACGAATTT
roX1 F8	TCAGTGTTCAGCACCTCGTC
roX1 R8	TTTTGGGCACTTGGTGAAG
y1456F	GGTCCACCGTTATATACGAAACA
BPR08wh	TAGCTCCTGATCCTCTTG
attB_210R	TGACCGTCGAGAACCCGCTGACG

and integration of the transgenes was verified by single fly PCR using primers in the P[y^+Cary*attP*] target site (y1456F) and pUASTB transgene (BPR08wh) (Table 2-1).

Generation of *roX1 roX2* and control females with identical genetic background

The genotype of mutant females used for the Microarray gene expression study is $roX1^{SMC17A}$ *roX2*. $roX1^{SMC17A}$ is a severe *roX1* allele that has a complex rearrangement with bp 204–2362 replaced by a LacZ gene, the SV40 poly adenylation signal and 3' P-element end (Deng et al., 2005). All subsequent studies in this chapter use this allele. A previous study to compare gene expression in *roX1 roX2*, $roX1^+$ *roX2* and *roX1* $roX2^+$ females produced ambiguous results, probably because of differences in genetic background between strains. To eliminate this effect, $roX1^{SMC17A}$ *roX2* mutant and control females that have identical genetic background were generated. Control females contain the same $roX1^{SMC17A}$ *roX2* X chromosome, but carry a [w^+Hsp*83-roX1*$^+$] transgene. The [w^+Hsp*83-roX1*$^+$] transgene is a 4.1 kb *roX1* genomic fragment (*roX1* bp 42-4165) driven by the hsp83 promoter. It rescues both male survival and X localization of the MSL proteins (V. Meller, data unpublished). The genetic crossing scheme to generate control females is shown in Figure 2-1.

Total RNA isolation

Total RNA was extracted from three groups of fifty third instar larvae of

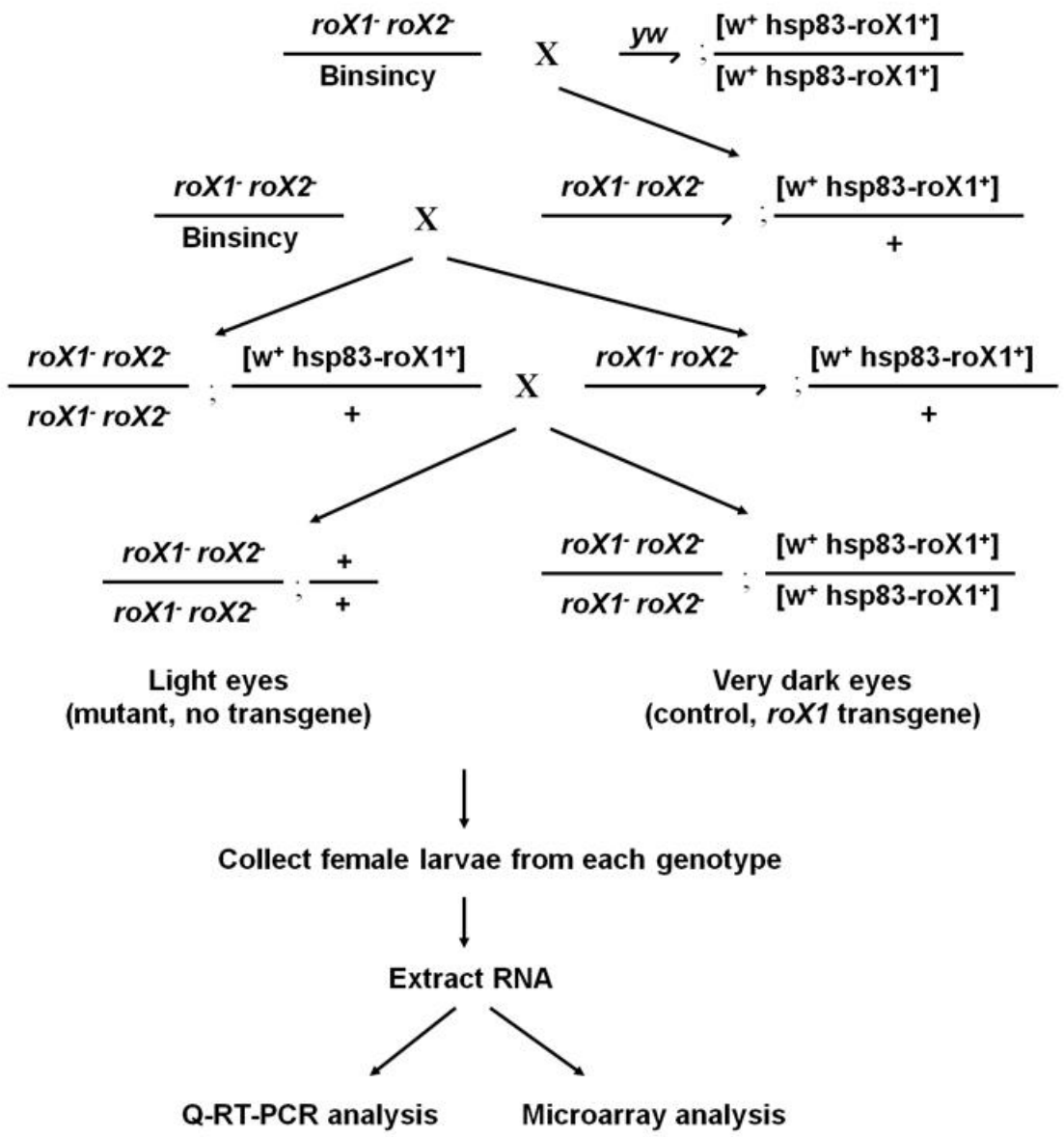

Figure 2-1. The crossing scheme for generating mutant and control females with identical genetic background for microarray study. "+" represents a wild type chromosome, X means cross.

each genotype using TRIzol (Invitrogen). Briefly, fifty third instar larvae were homogenized in 1ml of TRIzol by Tissue Tearor (Model 985-370, Biospec Products, Inc.). After centrifugation at 12000 rpm for 12 minutes, the supernatant was transferred into a new tube, mixed with 0.2ml chloroform and centrifuged again at 12000 rpm for 15 minutes. The aqueous phase was transferred into a new tube. Total RNA was precipitated by adding 0.5ml of isopropanol to the aqueous phase. The RNA pellet was washed with 75% ethanol, air dried, and dissolved in 50μl RNase-free water. Total RNA was then purified using the RNeasy Mini Kit following instructions (Qiagen).

Quantitative RT–PCR

1μg of total RNA was reverse transcribed at 42 °C for 1 hour using random hexamers and ImProm-II reverse transcriptase (Promega). The reverse transcription reaction consisted of 1 μg RNA template, 0.5 μg random hexamers (primers), 3 mM $MgCl_2$, 0.5 mM dNTP mix, 0.5 μl Recombinant RNasin (RNase inhibitor), 1 μl ImProm-II reverse transcriptase (Promega) and 4 μl 5XImProm-II reaction buffer in a total volume of 20 μl. Quantitative PCR was performed as previously described (Deng et al., 2005) using an Mx3000P Realtime PCR system (Stratagene). Briefly, two PCR reactions for each template were performed in parallel. Each PCR reaction consisted of 5 μl cDNA template (1:20 dilution), 0.3 μM of each primer, and 12.5 μl 2XSYBR Green PCR Master Mix (Applied Biosystem) in a total volume of 25 μl. Q-RT–PCR data were analyzed by the efficiency corrected comparative quantification method (Pfaffl, 2001). Ct

Table 2-2. Primers used for Q-RT-PCR analysis of gene expression*.

Genes	Primers	Sequences (5' - 3')	Primer work Con. (nM)	Efficiency (%)
Dmn	*Dmn* F	GACAAGTTGAGCCGCCTTAC	300	98.5
	Dmn R	CTTGGTGCTTAGATGACGCA	300	
Ytr	*Ytr* F	ATTTTGGACCAGCACCACTC	300	90.6
	Ytr R	CAAAATCCCTGCAATTTCGT	300	
LanB1	*LanB1* F	TCAACGAGCACCTGATTCAC	300	94.5
	LanB1 R	GCAAATGGATGTTTCCCAAT	300	
Xbp1	*Xbp1* F	GGGAGAGCAACTTTGACGAG	300	97.1
	Xbp1 R	GCCGGCCAAACTTAACAATA	300	
Atp-a	*Atp-a* F	ACCCACACTGCTACACTCCC	300	106.3
	Atp-a R	TCCTGGTTGCTCTTGTTGTG	300	
Dip-B	*Dip-B* F	AGGATCACGCCAGAAGACTG	300	92.8
	Dip-B R	AGTCACTGGGACGGAGAATG	300	
CKII	*CKII* F	CCTGGTTCTGTGGACTTCGT	300	98.4
	CKII R	GTAGTCCTCATCCACCTCGC	300	
skpA	*skpA* F	CTAAAAGTCGACCAGGGCAC	300	90.4
	skpA R	CCAGATAGTTCGCTGCCAAT	300	
CG1702	*CG1702* F	GACATCTTTGCAGCCTGTGA	300	92.7
	CG1702 R	GCCCTGATCTTGGGGTACTT	300	
Ppv	*Ppv* F	TTGACCACCCATGAACTCAA	300	94.2
	Ppv R	GTGTTTGCTATGCTTGGGGT	300	
Arc70	*Arc70* F	ATCGTACAACAACGAGCCCT	300	86.4
	Arc70	CAGCGTGAAAGAAACGTCAA	300	
cals	*cals* F	AGTTTGTCAGCCCTCACCTT	500	89.2
	cals R	CTCCTATGCATTGCGACAGA	500	
Ephrin	*Ephrin* F	TTGCAATTCTTGGCATTCAC	300	95.2
	Ephrin R	CATAGAGGTCGCGGTGATTT	300	
plexA	*plexA* F	AAAGCAGCGATTGGCTTTTA	500	86
	plexA R	GGCGCAGCTCTTATTCTGAC	500	
plexB	*plexB* F	AACGGAACCACAAAAGATCG	300	98.8
	plexB R	ATGTTACCGAGCGAACCAAC	300	
Rad23	*Rad23* F	GCGGATAACGAAGACTTGGA	300	99
	Rad23 R	TAGCCGTTCTATTGCGTCCT	300	
Crk	*Crk* F	AACATTAATGGGCAATGGGA	300	92.6
	Crk R	CATCGACAAATTCAACGTGC	300	
unc-13	*unc-13* F	GCGTTGGACGACTTAGCTTC	300	99.9
	unc-13 R	CATGTCTCCAAGTTCTCGCA	300	
Ank	*Ank* F	TGCAGAGTTTGGCACTCATC	300	100.1
	Ank R	TCGCCATCTTTTTCAATTCC	300	
Mav	*Mav* F	GATAAAATCGACGAGGCCAA	300	104.4
	Mav R	TTTTCCTAGATCCTGGCCCT	300	

*Information from X. Deng dissertation.

values for three biological replicates (each containing two technical replicates) per genotype were averaged into one Ct value per gene (performed by Mx3000P software, Stratagene). Statistical significance was assessed by performing an unpaired two-tail t test. A total of 19 genes were selected from three different gene groups (Autosomal, fourth chromosomal, and X chromosomal). The selected genes were expressed at moderate levels, displayed uniform absorbance in arrays of the same genotype, and reflected the average change in expression for their gene group in *roX1 roX2* males. The autosomal gene *Dmn* was selected as a reliable transcript for expression normalization (data not shown). The primers used in this study are presented in Table 2-2.

Gene expression microarrays

Total RNA was prepared from three groups of 50 third instar larvae of both genotypes. Each RNA preparation served as a template for probe synthesis. A total of six probes were made: three from female *roX1 roX2* larvae, and three from female *roX1 roX2* larvae carrying the [w^+Hsp*83*-*roX1*$^+$] transgene. Probes were synthesized at the ATGC core facility following manufacturer's instructions (www.Affymetrix.com). Hybridization of Affymetrix Drosophila Genome 2.0 chips (Santa Clara, CA) was also performed at the core facility. Background corrected intensity values were quantile normalized (Irizarry et al., 2003). In brief, all probe intensities from mutant and control arrays were assembled into a single ranking. Probes from individual chips were assigned the value of the corresponding quantile, thus preserving the rank order within a chip and standardizing intensity

distribution across all chips. Signal intensities were summarized into one expression value per sample and probe set using the robust multi-array average (RMA) algorithm. The Affymetrix MAS5.0 Present/Absent calls were used to filter out probe sets not present in at least two out of three replicates of each genotype.

Genes and probe sets (Berkeley Drosophila Genome Project annotation release 5.8) were sorted to enrich for heterochromatic genes on the basis of the boundaries between heterochromatic and euchromatic regions (Hoskins et al., 2007; Smith et al., 2007). The coordinates of these boundaries are: 2R;1-1285689, 2L;22000975-23011544, 3R;1-378656, 3L;22955576-24543557, X;22030326-22422827. The coordinates for heterochromatin that is not contiguous with assembled arm sequences are 2LHet;1-368872, 2RHet;1-3288761, 3LHet;1-2555491, 3RHet;1-2517507, XHet;1-204112, YHet; 1-347038. Only probe sets assigned to a chromosome were used. Genes and probe sets assigned to heterochromatic regions were obtained from FlyBase GBrowse. The corresponding gene and probe set information was obtained from the Affymetrix Drosophila_2 annotation file (Drosophila_2.na25) released on March 17, 2008 (Liu et al., 2007). Detailed microarray information is present in Appendix A.

Statistical methods and descriptions

The $\log_2$ fold change of each gene was computed as the $\log_2$ mean RMA expression of mutant samples minus the $\log_2$ mean RMA expression of control samples. The significance of differences between groups was assessed by the Wilcoxon test. Analyses were performed in the R software environment (www.r-

project.org) using Bioconductor (www.bioconductor.org) (Gautier et al., 2004). The raw data can be downloaded from the Gene Expression Omnibus (http:// www.ncbi.nlm.nih.gov/geo,GSE12076).

RESULTS

Females do not require *roX* for full expression of genes in heterochromatic environments

Simultaneous mutation of *roX1* and *roX2* is lethal for male adults, but males mutated for only one *roX* gene are normal with full survival. A microarray study conducted to compare genome wide expression in *roX1 roX2* males (null for *roX* function) and *roX2* (control) males primarily showed a global decrease of expression from the entire fourth chromosome, as well as X chromosome (Figure 2-2, A and B, (Deng et al., 2009; Deng and Meller, 2006)). The *roX* RNAs are redundant for their effect on fourth-linked genes, as well as for dosage compensation (Deng et al., 2009). The data that suggested a *roX*-dependent fourth chromosome regulation is based on a study of male flies. However, the idea that autosomal genes are regulated the same way in males and females is more plausible. *roX1* is abundant in early development stages of both sexes, and thus might contribute to the expression of fourth-linked genes in both sexes (Meller, 2003). Studies by our lab showed that MSL1, MSL3 and MLE are also necessary for normal expression of fourth-linked genes in males. All MSL proteins, except MSL2, which is not necessary for regulation of fourth-linked genes, are also present in females (Deng et al., 2009; Rastelli et al., 1995). To investigate the effect of loss of *roX* transcripts on expression of fourth-linked genes in females, microarrays were performed to compare the genome wide expression between *roX1 roX2* and *roX1 roX2;* [w^+Hsp*83-roX1*$^+$] female larvae.

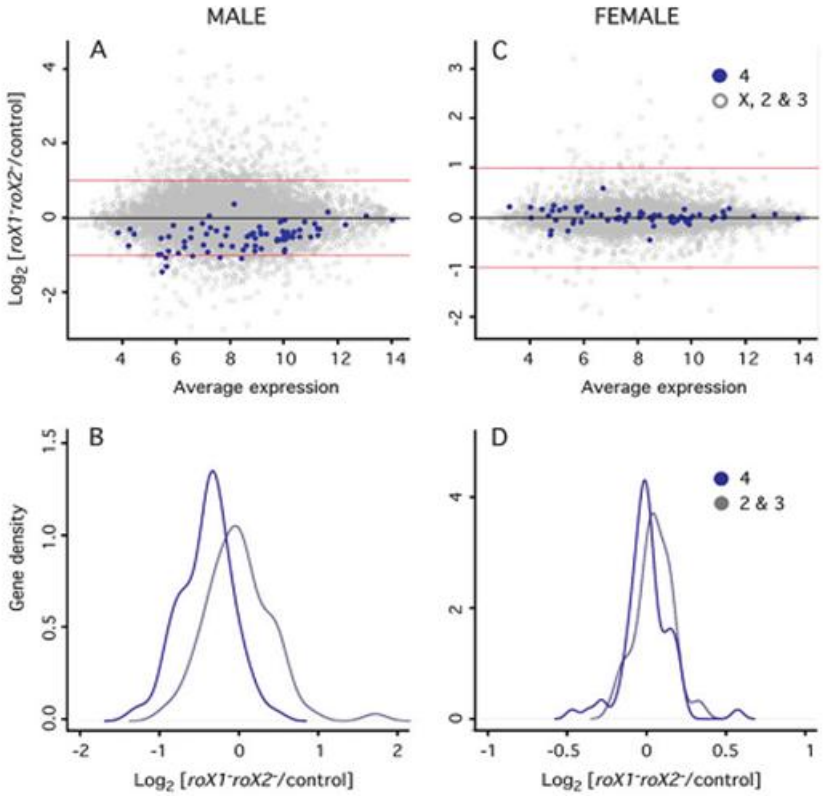

Figure 2-2. Expression of the fourth chromosome is reduced in *roX1 roX2* males but not in females. (A) In *roX1*^{SMC17A} *roX2* males the expression of fourth-linked genes (blue) decreases in comparison with the rest of the genome (gray). Points represent the $\log_2$ of the ratio of gene expression in *roX1*^{SMC17A} *roX2* males to control males (*roX2*) plotted against expression level ($\log_2$ absorbance). There are 9880 non-fourth-linked genes and 74 fourth-linked genes plotted. **(B)** The density distribution of $\log_2$ expression (mutant/control) for fourth-linked genes (blue) and second and third chromosome genes (gray) in males. The distribution of fourth-linked genes differs significantly from the remaining autosomal genes (adjusted P-value $6.6X10^{-16}$; Wil-coxon test). **(C)** In *roX1*^{SMC17A} *roX2* females the expression of fourth-linked genes (blue) is unchanged. The rest of the genome is shown in gray. Data is presented as the $\log_2$ of the ratio of gene expression in *roX1*^{SMC17A} *roX2* females to control females (*roX1*^{SMC17A} *roX2*; [w^+*Hsp83-roX1*$^+$]) plotted against expression level ($\log_2$ absorbance). Genes contributing to this analysis are 8433 non-fourth-linked and 69 fourth-linked genes. **(D)** The density distribution of $\log_2$ expression (mutant/control) for fourth-linked genes (blue) and second and third chromosome genes (gray) in female larvae. The distribution of fourth-linked genes is not significantly different from that of the second and third chromosomes (adjusted P-value 0.92). Only genes called as present in at least two out of three replicates were included. Microarray data was analyzed by S. K. Koya and A. Tarca.

This strategy enabled me to compare expression in *roX1 roX2* females with control females of virtually identical genetic background but carrying a *roX1* transgene, which rescues male survival and expression of both X-linked and fourth-linked genes of *roX1 roX2* males (Meller, unpublished, (Deng et al., 2009)). This minimizes inconsistency due to differences in genetic background. Two different plots were used to show this data. In figure 2-2 C, every gene is presented as a dot with expression level (X axis) plotted against $\log_2$ ratio of expression change (mutant : control) on the Y axis. The range of expression ratios is narrower in females than males (compare Figure 2-2, A and C). This is consistent with the fact that females are not developmentally disrupted by elimination of both *roX* RNAs. The expression of fourth-linked genes did not change in *roX1 roX2* females as compared to control females that express *roX1* from the [w^+Hsp*83*-*roX1*$^+$] transgene (Figure 2-2 C). A plot of the $\log_2$ of the expression ratio (mutant : control) of euchromatic genes on the second and third chromosomes has a distribution centered near zero. In contrast, the distribution of fourth-linked genes is shifted left in males (Figure 2-2 B), but is not significantly different from that of the second and third chromosomes in females (Figure 2-2 D). Therefore, the presence of the constitutively expressed *roX1* transgene did not affect expression of fourth-linked genes in *roX1 roX2* females. Quantitative RT-PCR of cDNAs from the same templates confirmed the microarray results (Table 2-3). I conclude that *roX* RNA are only required for full expression of fourth-linked genes in males.

The fourth chromosome is enriched for heterochromatin (reviewed in

Table 2-3. Q-RT-PCR validation of microarray analysis

Genes			Microarray Average		Fold Change	
Name	Chip_ID	Position	Mutant	Control	ΔArray (p-value)	ΔQPCR (std. dev.)
Xbp1	1635355	57C3-4	10.45	10.38	1.01 (0.41)	1.10 (0.13)
ytr	1623384	60A3	10.24	10.28	1.00 (0.68)	1.08 (0.22)
bigmax	1628490	97F1	9.42	9.32	1.01 (0.71)	1.00 (0.32)
skpA	1625801	1B14	12.04	12.04	1.00 (0.98)	0.98 (0.23)
Sgs4	1626630	3C10	14.06	13.98	1.01 (0.35)	1.09 (0.20)
plexB	1623432	102A1	7.92	7.85	1.01 (0.55)	1.05 (0.14)
Rad23	1625068	102B3	11.05	11.03	1.00 (0.85)	1.11 (0.15)
Crk	1634217	102A8	11.39	11.38	1.00 (0.95)	1.15 (0.16)
Ank	1635711	102A3	9.94	9.96	1.00 (0.91)	1.09 (0.16)
cals	1628842	102F8	10.15	10.16	1.00 (0.91)	1.10 (0.27)

To validate the microarray data of female *roX1*^{SMC17A} *roX2* flies, expression of selected genes was examined by Q-RT-PCR. cDNA templates used for Q-RT-PCR were generated from the same RNA samples used for microarray analysis. For each genotype, three independent RNA samples were used. Two PCR reactions per template were performed in parallel using an Mx3000P Realtime PCR system (Stratagene). *Dmn* was used as the normalizing gene. The microarray gene expression data was background corrected, normalized and summarized into a one expression value per sample and probeset using the RMA (robust multi-array average) algorithm. The microarray average is the average of three samples per genotype. The p-value of microarray analysis is generated by t-test. P-value > 0.2 means no significant changes.

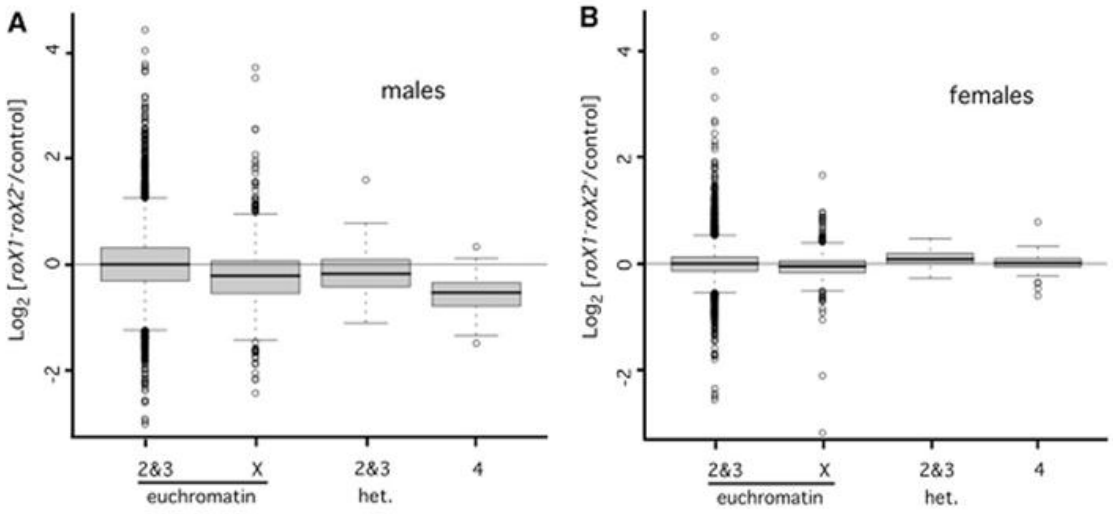

Figure 2-3. Genes situated in proximal heterochromatin require *roX* RNA for full expression in males but not in females. (A) Genes in proximal heterochromatin have reduced expression in *roX1*^{SMC17A} *roX2* male larvae. Box plots were generated using the $\log_2$ expression ratios (mutant/control) presented in Figure 2-2 A. The mean expression of genes in proximal heterochromatin on the second and third chromosomes decreases by 0.17 in *roX1*^{SMC17A} *roX2* males (adjusted P-value of 0.003). The mean expression of X-linked genes decreases by 0.24, and expression of fourth-linked genes decreases by 0.58. Changes of the X and fourth chromosome have an adjusted P-value of < $6.6X10^{-16}$. Only genes called as present in at least 2 out of 3 arrays contributed to this analysis (8347 in second and third euchromatin; 1533 in X euchromatin, 73 in second and third heterochromatin, and 74 on the fourth chromosome). **(B)** fourth-linked and heterochromatic genes do not require *roX* RNA for full expression in females. Box plots were generated using the $\log_2$ expression ratios (mutant/control) presented in Figure 2-2 C. The mean change in expression of X-linked genes in *roX1*^{SMC17A} *roX2* females is -0.04. Second and third chromosome heterochromatic genes and fourth-linked genes have a slight average increase (0.06 and 0.01, respectively) that is not statistically significant. Only genes called as present in at least 2 out of 3 arrays contributed to this analysis (7097 in second and third euchromatin, 1336 in X euchromatin, 57 in second and third heterochromatin, and 69 on the fourth chromosome). Microarray data was analyzed by S. K. Koya and A. Tarca.

(Riddle and Elgin, 2006)). Genes located in heterochromatic regions are presumed to have specialized regulatory features that enable their expression from repressive heterochromatic environment (Yasuhara and Wakimoto, 2006). It is possible that the involvement of *roX* RNAs in a general pathway to antagonize heterochromatin accounts for fourth-linked regulation. Reanalysis of microarrays showed that expression of heterochromatin-enriched genes on the second and third chromosomes also require *roX* RNA (Figure 2-3 A, (Deng et al., 2009)). Consistent with fourth-linked genes, expression of heterochromatic genes on the second and third chromosomes remained unchanged in *roX1 roX2* females (Figure 2-3 B). This suggests that there is a male-limited regulatory system for heterochromatic gene expression which involves *roX* RNAs.

A full length *roX1* transgene does not fully rescue expression of autosomal genes in males

roX RNAs are redundant for full expression of fourth-linked and heterochromatic genes in males. To confirm that reduced expression in *roX1 roX2* males is due to loss of the *roX* transcripts, the expression of fourth-linked genes was examined in $roX1^{SMC17A}$ *roX2* males carrying an autosomal *roX1* transgene, [w^{+}*Hsp83-roX1*$^{+}$], which fully rescues both male survival and X localization of the MSL proteins of *roX1 roX2* males (Meller, unpublished). Although this transgene rescues X-linked gene expression fully, only partial restoration of fourth-linked gene expression was observed (PhD thesis, X. Deng). As shown in figure 2-4, the expression level of X-linked genes from $roX1^{SMC17A}$

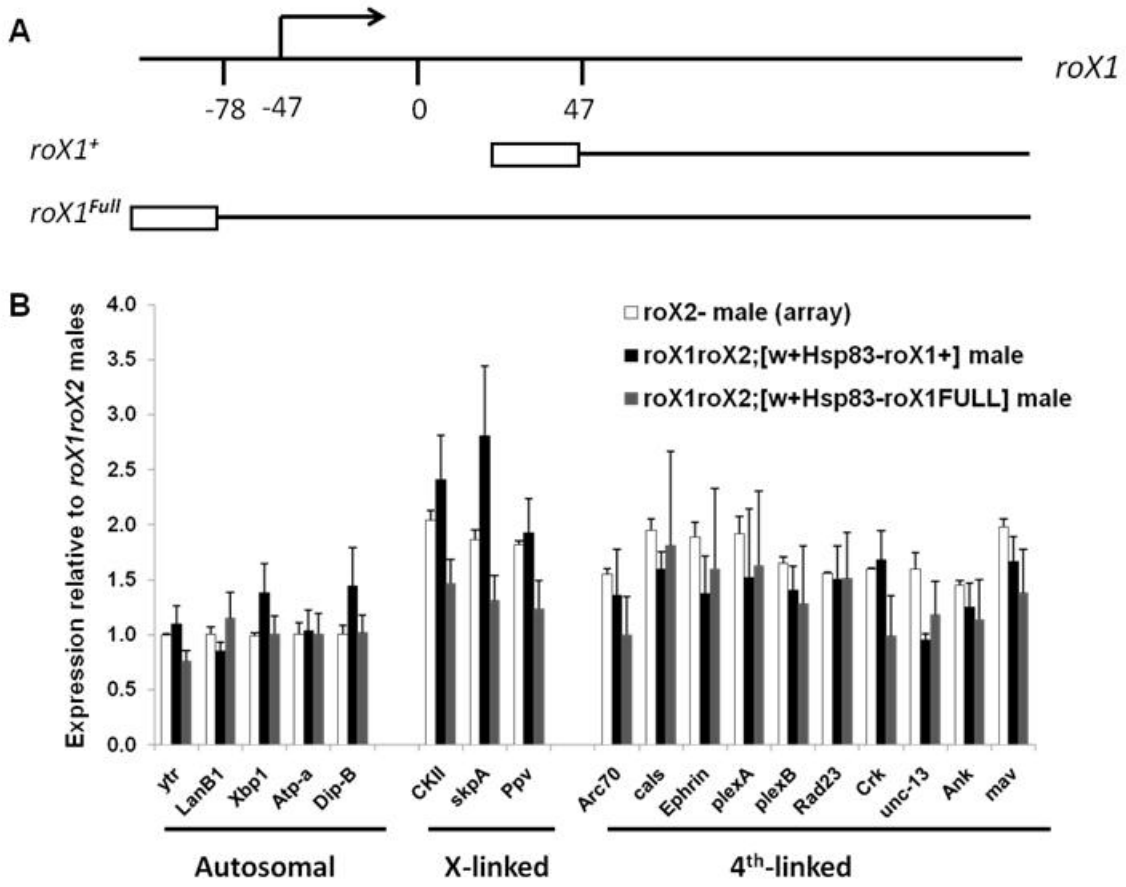

Figure 2-4. A full length *roX1* transgene does not fully rescue autosomal gene expression. (A) 5' ends of *roX1* transcript and transgenes. The top line indicates the *roX1* genomic region, numbering from (Amrein and Axel, 1997). The bent arrow indicates a transcription start site. Two additional transcription start sites are 306 and 675 bp downstream. $roX1^{+}$ and $roX1^{FULL}$ (short for [w^{+}*hsp83-roX1*$^{+}$] and [w^{+}*hsp83-roX1*^{Full}]) are illustrated below. The box represents *hsp83* promoter. **(B)** Expression of fourth-linked genes is partially restored by *roX1* transgenes. Five autosomal, three X-linked and 10 fourth-linked genes were analyzed by Q-RT-PCR in $roX1^{SMC17A}$ *roX2* (control), $roX1^{SMC17A}$ *roX2*; [w^{+}*hsp83-roX1*$^{+}$] (black) and $roX1^{SMC17A}$ *roX2*; [w^{+}*hsp83-roX1*^{Full}] (gray) male larvae. Expression from microarray analysis of $roX1^{+}$ $roX2^{-}$ male larvae (white) is shown for comparison. Error bars represent standard deviations from three independent measurements. Q-RT-PCR was normalized to the autosomal gene *Dmn*. Expression in $roX1^{SMC17A}$ *roX2* males was set to 1.

roX2 males carrying this transgene is almost the same as *roX2* single mutant. However, the expression of fourth-linked genes is lower than that in the *roX2* single mutant, but higher than that in the $roX1^{SMC17A}$ *roX2* double mutant. The $roX1^{+}$ transgene contains almost all of *roX1*, only lacking 100 bp at the extreme 5' end (Figure 2-4 A). Analysis of *roX1* ESTs (Expressed Sequence Tag, Flybase) showed that this 100 bp sequence is present in only a minor fraction of *roX1* transcripts. This suggested that the 5' end of *roX1* may have a separate function in the regulation of fourth chromosome or heterochromatic genes. To investigate this possibility, a full length *roX1* transgene including the additional 100 bp 5' sequence, [w^{+}*Hsp83-roX1*^{FULL}], was created (Figure 2-4 A). $roX1^{SMC17A}$ *roX2* flies carrying this transgene were generated, and the expression of a panel of euchromatic autosomal, fourth-linked and X-linked genes were measured by quantitative RT-PCR. When compared to the expression of these genes in *roX1 roX2* males carrying [w^{+}*Hsp83-roX1*$^{+}$], no significant difference in fourth-linked gene expression could be detected (Figure 2-4 B). This suggests that the extra 5' sequence is dispensable for regulation of autosomal genes. The incomplete rescue by *roX1* transgenes could be due to other aspects of the intrinsic *roX1* gene. One possibility is that the regulation is achieved by cooperation between different *roX1* transcripts. Interestingly, [w^{+}*Hsp83-roX1*^{FULL}] rescue of X-linked gene expression is also lower than that of [w^{+}*Hsp83-roX1*$^{+}$] (Figure 2-4 B). This is consistent with the fact that the majority of *roX1* transcripts do not contain this extra 5' sequences but still fully rescue X-linked genes expression.

Investigation of candidate *roX1* sequence that may limit *roX1* regulation of autosomal genes in females

The *roX* RNAs only affect expression of heterochromatic genes in males. However, *roX1* is present in both sexes in the early stages of development (Meller, 2003). All MSL proteins, with the exception of MSL2, are maternally provided to early embryos of both sexes (Rastelli et al., 1995). MSL2 is only produced in the male zygote, limiting MSL complex formation to males (Rastelli et al., 1995). While all other MSL proteins are required for heterochromatic regulation, MSL2 is not necessary for full expression of heterochromatic genes (Deng et al., 2009). What limits *roX*-dependent modulation of heterochromatin to males is still mysterious. Sexlethal (SXL) is a candidate factor that might block *roX* function in females. SXL is a multifunctional RNA-binding protein that controls sexual differentiation in flies and is expected to regulate all sex-limited processes directly or indirectly. The SXL protein is only expressed in females and determines sexual development by modulating the expression of a set of downstream genes including MSL2 (reviewed in (Salz and Erickson, 2010)). *roX1* has two potential SXL binding sites in a small intron (Figure 2-5 A, (Meller et al., 1997)). The function, if any, of these sites is unknown. It is possible that they bind SXL in females, limiting the ability of *rox1* to regulate heterochromatin. It is anticipated that if SXL blocks up-regulation of heterochromatic genes in females by binding to *roX1*, removal of SXL-binding sites from *roX1* will result in an elevated expression of heterochromatic genes in females. To test this possibility, females with a *roX1*^{c20} transgene that lacks SXL binding sites were used. The

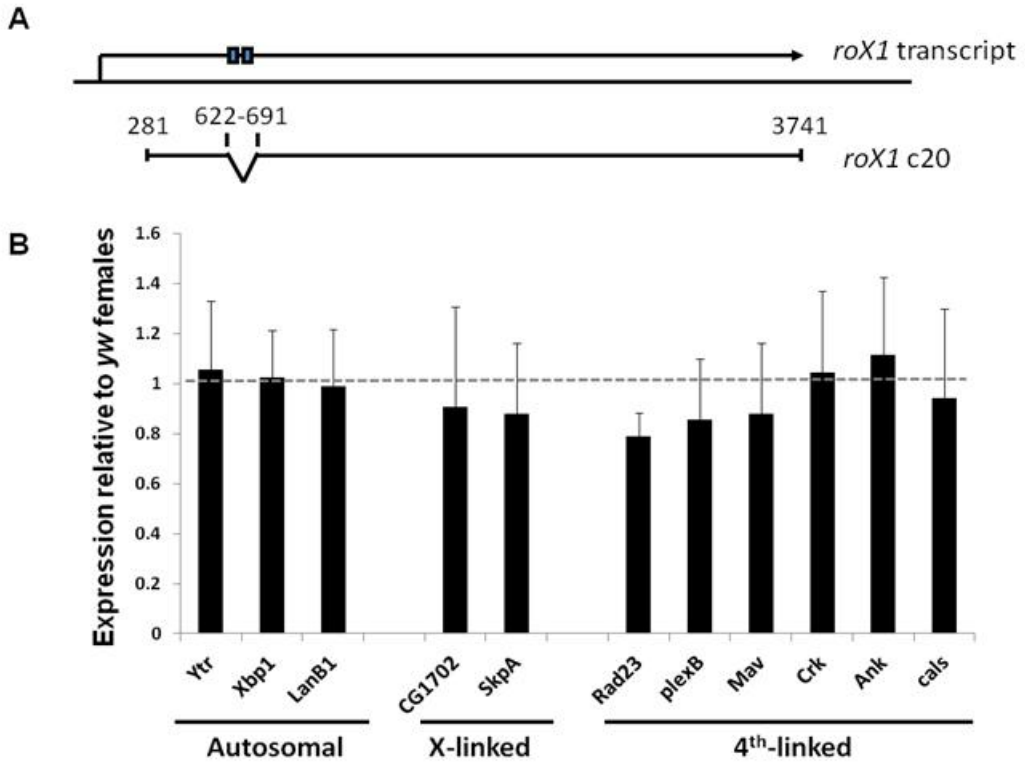

Figure 2-5. Expression of fourth–linked genes is unchanged in females expressing the *roX1*^{c20} transgene. (A) The *roX1*^{c20} transgene lacks SXL binding sites and initiates at one of the major transcription start sites. The middle line indicates the *roX1* genomic region, numbering from (Amrein and Axel, 1997). The arrow above represents the primiry *roX1* transcript. Boxes represent SXL binding sites. Below is the structure of the *roX1*^{c20} cDNA. *roX1*^{c20} lacks SXL binding sites, present in a small intron, which has been spliced out. **(B)** Quantitative reverse transcription PCR (Q-RT-PCR) of fourth-linked genes in females. Two autosomal, three X-linked and six fourth-linked genes were analyzed in female larvae carrying a *hsp83*-deiven *roX1*^{c20} transgene and in control females. Values are the ratio of gene expression of transgenic females to control (set to 1). Error bars represent standard deviations of at least three independent measurements. Q-RT-PCR was normalized to the autosomal gene *Dmn*. The full transgenic fly genotype is *yw;* [*w*$^{+}$hsp83-*roX1*^{c20}].

roX1^{c20} transgene was made from a spliced cDNA that lacks the intron containing SXL binding sites. It also initiates from a downstream transcription start site (Figure 2-6A, (Stuckenholz et al., 2003)). We choose several fourth-linked genes to test because of their heterchromatic features, and because these genes appear very sensitive to loss of *roX* RNA in males. The expression of fourth-linked genes in female larvae was assessed by Q-RT-PCR. As shown in Figure 2-5 B, the expression of genes on fourth chromosome was unchanged. I conclude that the SXL binding sites in *roX1* do not limit *roX1* function in females.

DISCUSSION

roX RNAs participate in two distinct biological processes in flies: X chromosome dosage compensation and normal heterochromatin function. Both of these coordinately regulate large regions of the genome. Although all molecules necessary for heterochromatin regulation are present in females, *roX* RNA has no affect on heterochromatin in females. Therefore, *roX*-dependent modulation of heterochromatin is limited to males, as dosage compensation.

Heterochromatin is present in both sexes, and autosomal heterochromatic regions are present in two copies in both sexes. The male-specific involvement of *roX* in heterochromatin function is an unexpected feature of heterochromatic regulation. However, mutations in heterochromatin proteins do show sex-biased phenotypes. Depletion of a major component of heterochromatin (HP1) causes higher lethality and more gene misregulation in males (Liu et al., 2005). The same study also identified differences in HP1 distribution in males and females. These differences suggested that heterochromatin itself is different in males and females. The male specific involvement of *roX* in heterochromatin function may reflect the differences of heterochromatin content in the male and female karyotypes.

Although *roX*-dependent regulation of heterochromatin is limited to males, we could not exclude the possibility that there is a female-specific noncoding RNA, rather than *roX* RNA, involved in the regulation of heterochromatin in females. However, if a female-specific RNA is involved in regulation of hetero-

chromatin, this represents a system different from the *roX*-dependent one that we have described. Q-RT-PCR of several fourth-linked genes in *mle* mutants showed that their expression was reduced in males, but unaffected in females (X. Deng, dissertation). Therefore, at least one protein component of the *roX*-dependent heterochromatin regulation system is necessary for normal expression of heterochromatic genes in males, but not in females. This suggested that the components of the *roX*-dependent heterochromatin regulation system do not assemble with a female-specific RNA instead of *roX* RNA in females to modulate heterochromatin.

Studies in our lab showed that heterochromatic genes display similar expression in males and females (Deng et al., 2009). The fact that they are differentially regulated in males and females, in spite of similar expression level, raises the question of why this difference exists. One possibility is the Y chromosome. The *Drosophila* Y chromosome represents 12% of the male genome and is almost entirely heterochromatic. It absorbs a large portion of the proteins that assemble into heterochromatin. Therefore, it affects other heterochromatic regions in male genome. Loss of the Y chromosome frees these proteins to enable them to bind elsewhere, promoting heterochromatin formation and enhancing Position Effect Variegation (PEV) throughout the genome (Weiler and Wakimoto, 1995). Therefore, loss of the Y chromosome silences transgenes in proximal heterochromatin and on the fourth chromosome. Our studies show that loss of *roX* RNAs increases expression from these transgenes (Deng et al., 2009). *roX* and the Y chromosome thus exert opposing effects on

heterochromatic expression.

Therefore, the role of the newly discovered *roX*-dependent system in sex-specific regulation of heterochromatin may accommodate the different heterochromatin environment in males, which carry a large, heterochromatic Y chromosome. Dosage compensation evolved in animals with differentiated X and Y chromosomes to balance expression of X-linked genes. During the evolution of sex chromosome pairs, the Y chromosome irreversibly loses coding sequences and accumulates repetitive sequences (Rice, 1996). This process promotes the formation of heterochromatin on the Y chromosome. The *roX*-dependent system that we have identified may have evolved to resolve a problem raised by the presence of a highly heterochromatic Y chromosome in the male nucleus. The Y chromosome changes the chromatin environment in the nucleus, as demonstrated by its influence on PEV. It is possible that the *roX*-dependent process arose to accommodate this difference in chromatin environment. Both processes of *roX*-dependent dosage compensation and modulation of heterochromatin occur only in males. They may have evolved to accommodate two different problems resulting from sex chromosome differentiation.

The molecular basis of *roX* regulation of heterochromatic genes is currently unclear. MSL1, MSL3 and MLE, but not MSL2, the only factor limited to males, are necessary for full expression of heterochromatic genes in males (Deng et al., 2009). *roX1* is present in females in the early stages of development (Meller, 2003). In contrast, *roX2* expression is limited to males. MSL1, MSL3 and MLE are also stably present in early embryos and in females (Rastelli et al.,

1995). It is possible that *roX* RNA, MSL1, MSL3 and MLE associate to form a complex that plays a transient role in the initial formation of heterochromatin in early embryos. However, what prevents this from occurring in females is still unclear. My study suggests that SXL binding to *roX1* is not the mechanism that prevents *roX1* activity in females. However, there may be other sex-limited factors that block regulation of heterochromatin in females, or enable it in males.

roX1 has several regions important for dosage compensation. The 5' region of *roX1* is important for MSL localization and a 3' stem-loop is essential for the chromatin modification by MSL (Deng et al., 2005; Park et al., 2008; Stuckenholz et al., 2003) It is possible that there are *roX1* regions that participate in regulation of heterochromatic genes, but not in dosage compensation. One candidate *roX1* region suggested by our preliminary data was the extreme 5' end, a region that is absent from most of the *roX1* transcripts produced by the fly (S. K. Koya, unpublished). As my full length *roX1* transgene including this region was not more active in rescue of heterochromatic genes compared to a transgene without this region but otherwise identical, I conclude that this region is dispensable for regulation of heterochromatic genes. The incomplete rescue by the full length *roX1* transgene could be due to some aspects of the intrinsic *roX1* gene. There is mixture of different *roX1* transcripts that are regulated by choice of transcription start sites, alternative splicing and transcript stability in the wild type flies. One possibility is that the regulation of heterochromatin is achieved by cooperation between different *roX1* transcripts.

CHAPTER 3

NONCODING *roX1* RNA 5' SEQUENCES CONTRIBUTE TO X CHROMOSOME LOCALIZATION AND LOCAL SPREADING OF THE MSL COMPLEX

This chapter is organized as manuscript in preparation (Kong Y. et al., in preparation)

INTRODUCTION

Transcriptional regulation in eukaryotes involves a large number of non-coding RNAs of various sizes. Previous studies show that they play critical roles in gene regulation through multiple mechanisms, including DNA methylation, chromatin modification, gene silencing, and dosage compensation (reviewed in (Mattick, 2003; Mattick and Makunin, 2005)). It is suggested that non-coding RNAs played a crucial evolutionary role by increasing the complexity and regulation of gene expression (reviewed in (Mattick, 2003; Prasanth and Spector, 2007)). Several long non-coding RNAs are misregulated in various diseases (Gupta et al., 2010; Prasanth and Spector, 2007; Taft et al., 2010). While an increasing number of long non-coding RNAs that have important functions has been discovered, the mechanisms by which they act are poorly understood (reviewed in (Mercer et al., 2009)). *roX1* is a long non-coding RNA involved in the epigenetic regulation of the X chromosome during dosage compensation in *Drosophila*. Functional dissection of *roX1* will contribute to our understanding of

other long non-coding RNA involved epigenetic processes, including those involved in human diseases.

Dosage compensation is a chromatin-mediated process that maintains a constant ratio of X-linked and autosomal gene expression in males and females (reviewed in (Lucchesi et al., 2005)). In *Drosophila*, dosage compensation occurs by a two-fold up-regulation of transcription from the single male X chromosome to match the transcription from the two female X chromosomes (reviewed in (Kong and Meller, 2007)). Dosage compensation in *Drosophila* is mediated by the Male-Specific Lethal (MSL) complex. The MSL complex specifically binds the male X chromosome and directs acetylation of histone H4 at lysine 16 (H4K16ac) (Bone et al., 1994; Smith et al., 2000). MSL localization along the male X is discontinuous. The MSL complex binds preferentially to actively transcribed genes, and is especially enriched within 3' transcribed regions (Alekseyenko et al., 2006; Gilfillan et al., 2006). Little MSL complex binds silent genes or intergenic regions of the X.

The MSL complex is composed of five Male-Specific Lethal (MSL) proteins and two non-coding *roX* (RNA on the X) RNAs. The five MSL proteins are MSL1 (Male-Specific Lethal 1), MSL2 (Male-Specific Lethal 2), MSL3 (Male-Specific Lethal 3), MLE (Maleless) and MOF (Males Absent on First) (reviewed in (Kong and Meller, 2007)). MSL1 is proposed to serve as the scaffold for assembly of the MSL complex (Li et al., 2005). MSL2 is a RING finger protein that regulates assembly of the complex, has DNA binding activity, and is only expressed in males (Fauth et al., 2010; Kelley et al., 1995). Recently it was found

to be an E3 ubiquitin ligase that modifies histone H2B on lysine 34 (Wu et al., 2011). The role of this activity in dosage compensation remains unknown. MSL3 contains a chromodomain that recognizes the cotranscriptional histone H3 trimethylation mark H3K36me3 (Larschan et al., 2007). The MSL3 chromodomain is required for the enrichment of the MSL complex within transcribed genes, and explains why the complex is specifically localized within the bodies of active genes (Sural et al., 2008). MOF is a histone acetyltransferase which acetylates histone H4 at lysine 16 (Hilfiker et al., 1997). MLE is an ATP-dependent DexH RNA/DNA helicase (Lee et al., 1997). The ATPase activity is necessary for transcriptional enhancement by the MSL complex (Morra et al., 2008). The helicase activity of MLE is essential for normal localization of the MSL complex on the X chromosome, and for movement of the *roX* RNAs from their sites of synthesis (Gu et al., 2000; Meller et al., 2000; Morra et al., 2008). MLE is suggested to associate with the MSL complex through an RNA (Richter et al., 1996; Smith et al., 2000). However, a recent study suggests that MLE is also capable of direct interaction with MSL2 (Morra et al., 2011). The stability of *roX1* is particularly dependent on MLE, supporting the idea of a direct interaction between these molecules (Meller, 2003). X-localization of remaining complex members is reduced or absent in males missing any one of the MSL proteins, or in males mutated for both *roX* genes. The mechanisms that underlie targeting of the MSL complex to the male X chromosome are poorly understood.

A two-step model is proposed for MSL complex targeting to X chromosome (reviewed in (Gelbart and Kuroda, 2009). In this model, the MSL

complex specifically recognizes and binds to a number of X-linked sites called chromatin entry sites (CES), also known as High Affinity Sites (HAS). These are regions that retain residual MSL binding in an *msl3* mutant background (Alekseyenko et al., 2008; Straub et al., 2008). ChIP-chip and ChIP-seq studies have identified over one hundred CESs on the X chromosome (Alekseyenko et al., 2008; Straub et al., 2008). Binding to CES sites is postulated to depend on interaction of the MSL complex with DNA sequences. After binding to the CES, the MSL complex is thought to spread to nearby, transcribed genes in a sequence-independent manner. This local spreading is partly mediated by the recognition of the H3K36me3 mark by the MSL3 chromodomain (Larschan et al., 2007; Sural et al., 2008). Other features, such as lower affinity binding sites that act cooperatively with the CES to recruit MSL complex and interphase chromosome architecture, which may bring X-linked sequences close together, could also contribute to MSL localization on X (Gilfillan et al., 2007; Grimaud and Becker, 2009; Kotlikova et al., 2006).

roX RNAs are also required for the localization of the MSL complex to the male X chromosome (Deng et al., 2005; Meller and Rattner, 2002). Upon elimination of both *roX* transcripts, mislocalization of the MSL proteins to heterochromatic regions and autosomal sites is observed (Deng et al., 2005). This is accompanied by a global decrease of X-linked gene expression (Deng and Meller, 2006). *roX* RNAs are male-preferential (*roX1*) or male-specific (*roX2*) non-coding RNAs that are transcribed from the X chromosome and assemble with the MSL proteins to coat the X chromosome (Amrein and Axel, 1997; Meller

et al., 1997). *roX1* and *roX2* are functionally redundant, in spite of a lack of extensive sequence similarity (Meller and Rattner, 2002). Assembly of *roX* RNAs into the MSL complex stabilizes the RNA, since they are rapidly degraded in the absence of MSL proteins (Meller et al., 2000).

The *roX* RNAs appear to direct MSL complex localization by two genetically distinct mechanisms. *roX* RNA is normally transcribed from the X chromosome, but autosomal *roX* transgenes enable the MSL complex to bind the X chromosome and rescue *roX1 roX2* males (Meller and Rattner, 2002). The *roX* RNAs can thus act when they originate in *trans* to the chromosome that they modify. The *roX* RNAs can also direct the spreading of the MSL complex into chromatin flanking autosomal sites of *roX* RNA synthesis in flies carrying a *roX* transgene (Kelley et al., 1999). Autosomal spreading is discontinuous, which is similar to the spreading pattern on the X, suggesting that this spreading occurs by a mechanism similar to spreading on the X chromosome (Kelley et al., 1999). This was confirmed by ChIP studies that detected enrichment of MSL3 and H4K16ac in genes flanking an autosomal *roX1* insertion (Park et al., 2010). The mechanism of spreading of MSL complex from a site of *roX* transcription into flanking chromatin is unknown, but one model to explain local spreading proposes that it is controlled by MSL complex assembly on nascent *roX* transcripts. If assembly of the complex is complete before the *roX* transcript is released from the DNA template, the MSL complex will then bind chromatin immediately, localizing in the vicinity of the *roX* genes (Oh et al., 2003; Park et al., 2002). However, if MSL protein is limiting, or *roX* transcription is fast, formation of

the intact complex will occur in the nucleoplasm after release of *roX* RNA from the chromatin template. The complex will then be able to travel to sites far from *roX* transcription. While unproven, this idea is supported by studies in which the level of MSL proteins or *roX* RNA has been manipulated. A high ratio of MSL proteins to *roX* transcripts promotes local spreading from sites of *roX* transcription, but when *roX* transcripts are more abundant, local spreading is less likely to occur (Kelley et al., 2008; Oh et al., 2003; Park et al., 2002).

A similar, RNA-dependent process occurs in mammalian females during X inactivation. The *X inactive specific transcript* (*Xist*) RNA is a large, non-coding transcript that is central to dosage compensation in mammals. *Xist* is transcribed from the *Xic* (*X inactivation center*) and coats the entire inactive X chromosome (Jaenisch et al., 1998). *Xist* sequences required for silencing and localization have been identified. A 5' repeat element (repeat A) that can form a two stem-loop structure is essential for X chromosome silencing, while redundant, spatially separated sequences throughout *Xist* act co-operatively to promote *Xist* localization to the X (Wutz, 2003).

We postulated that *roX1* might be similarly organized into regions with distinct function. The major *roX1* transcript is 3.7 kb (Amrein and Axel, 1997; Meller et al., 1997). It contains redundant 3' '*roX* boxes', which may form alternative secondary structures and are necessary for the histone acetyltransferase activity of the MSL complex (Kelley et al., 2008; Park et al., 2008). The '*roX* box' is the only highly conserved element between *roX1* and *roX2*, and between *roX* genes in different species (Park et al., 2008). A small 3'

stem loop incorporating one of these *roX* boxes was the first functional element identified in *roX1* (Stuckenholz et al., 2003). The 5' end of *roX1* also contains sequences that are necessary for function. Deletion of the entire 5' end destroys *roX1* activity, but small (<300 bp) deletions scanning across the 5' end do not detectably influence the activity of *roX* transgenes (Stuckenholz et al., 2003). However, loss of large amounts of 5' sequences has a particularly strong effect on MSL localization (Deng et al., 2005). *roX1* transcripts thus appear to contain distinct domains for X localization and chromatin modification, as does *Xist*. And, like *Xist*, the internal *roX1* sequence between the 5' and 3' ends is not essential for function. We hypothesize that, like *Xist*, multiple redundant elements that contribute to X chromosome targeting might be distributed throughout the 1.5kb sequence at the 5' end of the *roX1* transcript.

In this study, *roX1* transgenes containing different sections of the 5' end were created and analyzed to investigate the function of this region. This study demonstrated that multiple redundant elements contributing to X chromosome targeting are present throughout the 5' end of *roX1*. I also demonstrated that the extreme 5' end of *roX1* has a unique function that promotes MSL complex spreading in *cis* from sites of *roX* production. Intriguingly, *roX1* has multiple transcription start sites that are developmentally regulated. Spreading *in cis* is promoted by a portion of *roX1* that is present in few naturally occurring *roX1* transcripts, suggesting that production of *roX* RNA capable of directing spreading in *cis* is tightly controlled by the cell. It also suggests that *roX1* activities are regulated by choice of transcription start site. My studies reveal that while all 5'

regions tested can recruit MSL1 to the X chromosome, one region is deficient in recruitment of MLE. This suggests that different 5' segments make different molecular contacts within the MSL complex. Finally, I demonstrate that while all of my transgenes are able to partially restore the viability of *roX1 roX2* males when transcribed from an autosome, one of them is specifically deficient in male rescue when transcribed from X-linked sites. I anticipate that future studies of this unexpected finding may explain how the *roX* genes direct chromatin modification *in cis* to sites of transcription.

Materials and Methods

Drosophila strains

Flies were maintained in vials containing standard cornmeal agar media at 25°C. The *roX1*^{mb710}, *roX1*^{ex40A} and *roX1*^{SMC17A} mutations have been previously described (Deng et al., 2005; Meller et al., 1997). Df(1)52 is a lethal deletion of *roX2* and essential flanking genes. Elimination of *roX2* is accomplished by complementing Df(1)52 with [w^+4Δ4.3], a cosmid insertion carrying essential deleted genes but lacking *roX2* (Meller and Rattner, 2002). For convenience this combination is referred to as *roX2*. *y w*; P[y^+Cary*attP*]2, which contains P[y^+Cary*attP*] at 68A4, was obtained from the Bloomington *Drosophila* stock center (BDSC stock# 8622). The *roX1* transgene, [w^+*hsp83-roX1*$^+$], is a hsp83-driven *roX1* genomic fragment (*roX1* bp 42-4165, numbering from Amrein and Axel, 1997). The *roX1* transgene, [w^+*hsp83-roX1*^{FULL}], is described in chapter 2. The mutations in *msl1*^{L60}, *msl3*^{2} and *mle*^{1} have been previously described (Chang and Kuroda, 1998; Lindsley and Zimm, 1992). The *msl2* transgene, [w^+hsp83-M2]6I, has also been described previously (Kelley et al., 1995).

Generation of *roX1* deleted transgenes for site-specific integration

Three *roX1* transgenes containing different 5' fragments and the essential 3' region, [w^+*hsp83-roX1*^{R1}], [w^+*hsp83-roX1*^{R2}], [w^+*hsp83-roX1*^{R3}], were generated. These are named as [w^+*hsp83-roX1*$^{R1-3}$]. For convenience these are also referred to as *roX1*$^{R1-3}$. The *roX1*^{R2} transgene was created by PCR

amplification of genomic sequence from the *roX1*^{ex40A} mutant using oligonucleotides *roX1* 190F and *roX1* R4150 (Deng et al., 2005). *roX1*^{ex40A} mutation has an internal deletion of *roX1* bp 809-3150. The *roX1*^{R1} and *roX1*^{R3} transgenes were generated by joining *roX1* 5' fragments (bp -85-310 and 790-1490) to the 3' sequence present in *roX1*^{ex40A} (bp 3150-4150). In brief, the *roX1* 5' fragments and 3' sequence were amplified separately by PCR from wild type genomic template using oligonucleotides *roX1* -85F and *roX1* 310R (*roX1*^{R1} 5' fragments), *roX1* 790F and *roX1* 1490 R (*roX1*^{R3} 5' fragments), *roX1* 3150F and *roX1* R4150 (3' fragments). A XhoI site is introduced by the reverse primer of 5' fragment and the forward primer of 3' fragment. The 5' and 3' *roX1* fragments were then cut with XhoI and ligated together, and then subcloned into pCR4-TOPO vector (Invitrogen). After amplification of the plasmid, the *roX1*^{R1}, *roX1*^{R2} and *roX1*^{R3} fragments were released from pCR4-TOPO by EcoRI and ligated into a modified pUASTB. pUASTB is a transformation vector carrying an *attB* site and a mini-*white* gene (Groth et al., 2004). pUASTB was modified by introducing the strong and constitutive *hsp83* promoter. The *hsp83* promoter was amplified from the pCasper HS83T3 vector (Meller et al., 2000) by PCR using oligonucleotides *hsp83* Link F and *hsp83* Link R. A KpnI site was introduced at the forward primer *hsp83* Link F and BamHI, EcoRI and SacII sites were introduced at the reverse primer *hsp83* Link R. A KpnI-SacII fragment containing the *hsp83* promoter was subcloned in the pUASTB. The unique BamHI and EcoRI sites were used to clone *roX1* adjacent to the promoter. Primers used to create *roX1* transgenes are listed in Table 3-1. To confirm correct cloning, the primers listed in Table 3-2 were

Table 3-1. Primers used for creation of *roX1* transgenes.

Primers	Sequences (5' – 3')
roX1 -85F	ATGTGCAATGCATGTATAACAGAAA
roX1 310R	ATCACTCGAGGGCAGGCCCTGGTAACTA
roX1 190F	AAGACATGGGCGTAGTTTCATATAC
roX1 790F	CTGATAGGGTGACCTAACGCAACAG
roX1 1490R	TACACTCGAGTCGATGCGTCGTTTATTG
roX1 3150F	TAGTCTCGAGCCAACCCCACATCAGGCC
roX1 R4150	ATCGAACTGTATATAATAATGGCATCAG
hsp83 Link F	ATGGTACCGATGATCCTTAACGGGGAACTTG
hsp83 Link R	AGCCGCGGGAATTCGGATCCTCGACGGTATCGATAAGCTAG

Table 3-2. Primers used for comfirmation of *roX1* transgenes and verification of integration of *roX1* transgenes and recombinant chromosomes.

Genotype or integration site	Primer pairs	Seqences	Product length (bp)
[$roX1^{R1}$] [a]	*hsp83* 662F	TTCGGACCACTTAGACGAATTT	600
	roX1 BPR02	GAGGGTACCGGGACAGCTCGTATATGA	
[$roX1^{R2}$] [a]	*hsp83* 662F	TTCGGACCACTTAGACGAATTT	600
	roX1 557R	GTTTTTCTATTGTCCGGACTCG	
[$roX1^{R3}$] [a]	*hsp83* 662F	TTCGGACCACTTAGACGAATTT	600
	roX1 R5	ATATGGGGCTCATCCACTCC	
$roX1^{SMC17A}$ [b]	7BXbal	GCTCTAGAATTCCAGTGATCGATCGGTAATAGTAAA	350
	lacZ 5' R	CCAGTCACGACGTTGTAA	
68A4 / 16C1 [c]	y1456F	GGTCCACCGTTATATACGAAACA	1750
	BPR08wh	TAGCTCCTGATCCTCTTG	
2A3 [d]	*roX1* BPR25	GAGGTCTAGAGCTGGCAAACGACCTGAGCAATACT	700
	ZH2A-3'	GTTACAAACAAGAGCCCAGCC	
19C4 [e]	attL.For	GGGCGTGCCCTTGAGTTCTCTC	700
	19C4.Rev	GACCATCCAACTTCCAACATCG	

a To confirm correct structure of $roX1^{R1-3}$ transgene, indicated primers that only produce a fragment of the correct size when correct transgene is subcloned after the *hsp83* promoter, were used to amplify sequences from the *hsp83* promoter and linked *roX1* transgene.
b To confirm the presence of $roX1^{SMC17A}$ mutation, indicated primers were used to amplify sequences from *roX1* genomic region and inserted P element.
c To confirm the integration of $roX1^{R1-3}$ transgenes at 68A4 and 16C1, indicated primers were used to amplify sequences from *yellow* gene and linked *white* gene. *yellow* and *white* genes are close together after correct integration.
d To confirm the integration of $roX1^{R1-3}$ transgenes at 2A3, indicated primers were used to amplify sequences from $roX1^{R1-3}$ transgene and the genomic region proximal to the integration site.
e To confirm the integration of $roX1^{R1-3}$ transgenes at 19C4, indicated primers were used to amplify sequences from attL formed by integration and the genomic region proximal to the integration site.

used.

Generation of transgenic flies

The *roX1*$^{R1-3}$ transgenes were integrated at 2A3, 16C1, 18D3 and 19C4 on the X chromosome and at 68A4 on the third chromosome using the site-specific ΦC31 integrase system (Groth et al., 2004). Stocks with an insertion of an *attP* landing site in the P[*y*$^{+}$Cary*attP*] P element at 68A4 were obtained from the Bloomington *Drosophila* stock center. Mobilization of this P element was accomplished by introducing a transposase source (TMS, Sb P[ry$^{+t7.2}$Δ2-3]99B) and selecting for insertions on the X chromosome (P. Frolov, unpublished). The exact location of the 16C1 and 18D3 insertions was determined by inverse PCR and sequencing (Inverse PCR Protocol available from Berkeley Drosophila Genome Project; http://www.fruitfly.org/about/methods/inverse.pcr.html). All of my constructs were purified using Qiagen Plasmid Midi Kit (Qiagen). To generate transgenic flies, I injected my constructs into *Drosophila* embryos derived from mating female virgins carrying P[*y*$^{+}$Cary*attP*] at 68A4, 16C1 and 18D3 with males carrying a ΦC31 integrase transgene on the 3rd chromosome. Adult flies derived from injected embryos were mated to *yw*, and transgenics expressing the mini-*white* marker were identified in their offspring. Transgenics were made homozygous, and integration of the transgenes was verified by single fly PCR using primers in the pUASTB transgene (BPR08wh) and P[*y*$^{+}$Cary*attP*] target site (y1456F) (Table 3-2). The *roX1* transgenes inserted at 2A3 on the X chromosome were generated by injecting my constructs into M[vas-int.Dm]ZH-

102D, M[3xP3-RFP.attP]ZH-2A fly embryos by Rainbow Transgenic Flies, Inc (3251 Corte Malpaso, Suite 506, Camarillo, CA 93012 USA). The *roX1* transgenes inserted at 19C4 on the X chromosome were generated by injecting my constructs into embryos derived from mating female virgins carrying P[y^+Cary*attP*] at 19C4 with males having a ΦC31 integrase transgene at 2A3, performed by BestGene, Inc (2918 Rustic Bridge, Chino Hills, CA 91709 USA). The identification of integrants and verification was performed as described above using primers in the pUASTB transgene (*roX1* BPR25) and flanking genomic region (ZH2A-3' for site 2A3, 19C4.Rev for site 19C4).

Phenotypic testing of *roX1* transgenes and rescue efficiency

All transgenes were analyzed in a *roX1 roX2* double mutant background. To analyze autosomal $roX1^{R1\text{-}3}$ transgenes, $roX1^{SMC17A}$ *Df(1)52;* [w^+4Δ4.3] female virgins were mated with *yw*/Y; [w^+*hsp83-*$roX1^{R1\text{-}3}$] males to produce sons. Survival of these sons was calculated based on the recovery of females from the same cross. Survival of females ($roX1^{SMC17A}$ *Df(1)52*/*yw;* [w^+4Δ4.3]/+; [w^+*hsp83-*$roX1^{R1\text{-}3}$]/+) is set at 100%. To analyze X-linked $roX1^{R1\text{-}3}$ transgenes, each was recombined onto a $roX1^{SMC17A}$ *Df(1)52* chromosome to produce $roX1^{SMC17A}$ *Df(1)52* [w^+*hsp83-*$roX1^{R1\text{-}3}$] chromosomes. These were maintained over a Binsincy balancer (Figure 3-1). The survival of males carrying $roX1^{SMC17A}$ *Df(1)52* [w^+*hsp83-*$roX1^{R1\text{-}3}$] X chromosomes was determined by mating females to males carrying an autosomal insertion of [w^+4Δ4.3]. Male survival was calculated using all male and female non-Binsincy offspring from the cross. These offspring carry

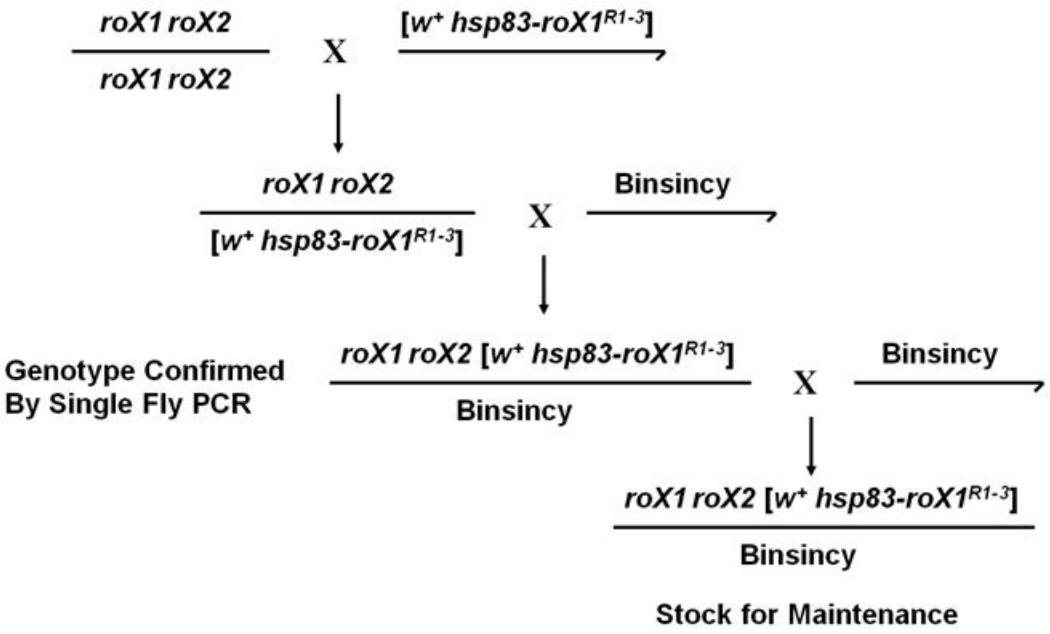

Figure 3-1. The crossing scheme for recombining X-linked *roX1*^{R1}, *roX1*^{R2} and *roX1*^{R3} transgenes onto a *roX1 roX2* chromosome. The full genotype of the *roX1 roX2* chromosome is *roX1*^{SMC17A} *Df(1)52*. *Df(1)52* is complemented with a cosmid which contains essential genes deleted by Df(1)52 but lacks *roX2,* [w^{+}4Δ4.3]. For simplicity, this combination is referred to as *roX2.* X represents mating.

the *roX1*^{SMC17A} *Df(1)52* [*w*$^{+}$*hsp83-roX1*$^{R1-3}$] X chromosome.

Rescue of *roX1*^{SMC17A} *Df(1)52* [*w*$^{+}$*hsp83-roX1*$^{R1-3}$] males with *msl* mutations and *roX1* transgenes

To determine if *roX1*^{SMC17A} *Df(1)52* [w^{+}*hsp83-roX1*$^{R1-3}$] males could be rescued by reduction of MSL protein levels or overexpression of MSL proteins and *roX1*, the mutations *msl1*^{L60}, *msl3*^{2} and *mle*^{1} (Chang and Kuroda, 1998; Lindsley and Zimm, 1992), a third chromosome carrying both the [*w*$^{+}$*hsp83*-M1] and [*w*$^{+}$*hsp83*-M2] transgenes (for convenience this chromosome is denoted as [M1][M2], (Chang and Kuroda, 1998)) and autosomal *roX1* transgenes *roX1*^{Full}, *roX1*$^{R1-3}$ were used. *roX1*^{Full} is described in chapter 2. *roX1*^{SMC17A} *Df(1)52* [*w*$^{+}$*hsp83-roX1*$^{R1-3}$]/Binsincy; [*w*$^{+}$4Δ4.3] female virgins were mated with *w*/Y; *msl1*/[*w*$^{+}$], *w*/Y; *msl3*/[*w*$^{+}$], *w*/Y; *mle*/[*w*$^{+}$], *w*/Y; [M1][M2]/+, *yw*/Y; [*w*$^{+}$*hsp83-roX1*$^{R1-3}$] and *yw*/Y; [*w*$^{+}$*hsp83-roX1*^{Full}] males to produce *roX1*^{SMC17A} *Df(1)52* [w^{+}*hsp83-roX1*$^{R1-3}$]; [*w*$^{+}$4Δ4.3] /+ sons with reduced level of MSL protein or elevated MSL protein or *roX1*. Male survival was determined as described above.

Extraction of genomic DNA from a single fly

A single fly is homogenized in an Eppendorf tube with a pestle in 50 µl of squash buffer (10mM Tris-HCl pH8.2, 1mM EDTA, 25mM NaCl, 0.2mg/ml proteinase K). The homogenate is incubated at 42°C for 1.5 hours, followed by 95°C for 10 minutes. 1.5 µl of this preparation is used as template for PCR amplification.

Polytene chromosome immunostaining

MSL1, MSL2 and MLE immunohistochemical detection on polytene chromosomes was performed as previously described (Kelley et al., 1999). In brief, polytene chromosome spreads were prepared from salivary glands of third instar male larvae. Immunostaining was performed with MSL1, MSL2 or MLE antibody (raised in rabbit) and detected with Texas Red conjugated α-rabbit antibody. Slides were briefly dipped in Hoechst (0.1µg/mL) to visualize DNA before mounting in 80% glycerol.

Photography

Observation and photography of immuno-histological images was performed with a Zeiss Axioscope 2 fluorescent compound microscope fitted with a Q-Imaging Retiga 2000R digital camera.

Generation and analysis of *Beadex* flies

To analyze the level of dosage compensation, flies carrying a dose sensitive *Beadex* allele (Bx^{r49k}) on X the chromosome were used. *roX2*Δ is a viable *roX2* deletion mutation generated by FLP/FRT recombination (V. Meller, unpublished). The $roX1^{mb710}$ *roX2*Δ Dp(1;Y) Bx^{r49k} X chromosome was made by recombination and maintained over a Binsincy balancer. $roX1^{mb710}$ *roX2*Δ Dp(1;Y) Bx^{r49k}/ Binsincy females were mated to *yw*; [w^{+}*hsp83-roX1*$^{R1-3}$] males to generate $roX1^{mb710}$ *roX2*Δ Dp(1;Y)Bx^{r49k}/Y; [w^{+}*hsp83-roX1*$^{R1-3}$]/+ sons. Wings were removed from flies, mounted in 4:5 lactic acid : ethanol and photographed. Wings

were measured and analyzed as previously described (Menon and Meller, 2009). Briefly, the length of wing margin lost from anterior and posterior edges and vein L3 length between the L2 junction and the edge of the wing were measured using ImageJ software (http://rsb.info.nih.gov/ij/). The wing notching is expressed as a ratio of the length of marginal lost to the length of the L3 vein to normalize for variation in wing size. The significance of differences in notching was determined by performing an ANOVA test.

Total RNA isolation

Total RNA was extracted from three groups of fifty third instar larvae of each genotype using TRIzol (Invitrogen). Briefly, fifty third instar male larvae isolated based on gonadal morphology were homogenized in 1ml of TRIzol by Tissue Tearor (Model 985-370, Biospec Products, Inc.). After centrifugation at 12000 rpm for 12 minutes, the supernatant was transferred into a new tube, mixed with 0.2ml chloroform and centrifuged again at 12000 rpm for 15 minutes. The aqueous phase was transferred into a new tube. Total RNA was precipitated by adding 0.5ml of isopropanol to the aqueous phase. The RNA pellet was washed with 75% ethanol, air dried, and dissolved in 50 µl RNase-free water. Total RNA was then purified using the RNeasy Mini Kit following instructions (Qiagen).

Quantitative-RT-PCR

Expression of *roX1* and the X-linked genes *Dlmo*, *SkpA*, and *Ck-II* was

Table 3-3. Primers used to analyze gene expression by Q-RT-PCR.

Genes	Primers	Sequences (5' - 3')	Primer work Con. (nM)	Efficiency (%)
Dmn	*Dmn* F	GACAAGTTGAGCCGCCTTAC	300	98.5
	Dmn R	CTTGGTGCTTAGATGACGCA	300	
CKII	*CKII* F	CCTGGTTCTGTGGACTTCGT	300	98.4
	CKII R	GTAGTCCTCATCCACCTCGC	300	
skpA	*skpA* F	CTAAAAGTCGACCAGGGCAC	300	90.4
	skpA R	CCAGATAGTTCGCTGCCAAT	300	
Dlmo	*Dlmo* F	TGAGATTGTTTGGCAACACG	500	95.3
	Dlmo R	ACGCATCACCATCTCGAAG	500	
roX1	*roX1* F3377	TTTTGTCCCACCCGAATAAC	300	107.3
	roX1 R3448	CCTTTTAATGCGTTTTCCGA	300	

measured by Quantitative-RT-PCR (Q-RT-PCR) as previously described (Deng et al., 2005). Briefly, one microgram of total RNA was reverse transcribed at 42 °C for 1 hour using random hexamers and ImProm-II reverse transcriptase (Promega). The reverse transcription reaction consisted of 1 µg RNA template, 0.5 µg random hexamers (primers), 3 mM $MgCl_2$, 0.5 mM dNTP mix, 0.5 µl Recombinant RNasin (RNase inhibitor), 1 µl ImProm-II reverse transcriptase (Promega) and 4 µl 5XImProm-II reaction buffer in a total volume of 20 µl. The quantitative PCR was performed using an Mx3000P Realtime PCR system (Stratagene). Two PCR reactions for each template were performed in parallel. Each PCR reaction consisted of 5 µl cDNA template (1:20 dilution), 0.3 µM of each primer, and 12.5 µl 2XSYBR Green PCR Master Mix (Applied Biosystem) in a total volume of 25 µl. Q-RT–PCR data were analyzed by the efficiency corrected comparative quantification method (Pfaffl, 2001). Ct values for three biological replicates (each containing two technical replicates) per genotype were averaged into one Ct value per gene (performed by Mx3000P software, Stratagene). The selected X-linked genes are all down regulated in *roX1 roX2* mutant (Deng and Meller, 2006). The autosomal gene *Dmn* was selected as a reliable transcript for expression normalization (data not shown). The primers used in this study are presented in Table 3-3. The significance of differences attributable to different *roX1* transgenes was determined by performing an unpaired two-tail t-test.

RESULTS

Deleted *roX1* transgenes partially rescue *roX1 roX2* males.

To investigate the function of 5' *roX1* sequence, three transgenes, [w^{+}*hsp83-roX1*^{R1}], [w^{+}*hsp83-roX1*^{R2}] and [w^{+}*hsp83-roX1*^{R3}] (also referred as *roX1*$^{R1-3}$ for convenience), containing fragments of the *roX1* 5' end were created. Design of these *roX1* transgenes was based on previous studies of *roX1* mutants and transgenes (Figure 3-2, (Deng et al., 2005) and data unpublished). These studies indicated that activity was present in at least three regions in the 5' end of *roX1*. The first region, present in *roX1*^{R1}, reflects the portion of the 5' end retained in the *roX1*^{ex7B} mutant, which retains partial *roX* activity. The second region, present in *roX1*^{R2}, when combined with the region present in *roX1*^{R1} and the 3' functional sequence, supports full male viability. This is reflected by the survival of *roX1*^{ex40A} *roX2* males (Figure 3-2). The third region, present in *roX1*^{R3}, can be deleted from *roX1* without any phenotype (Deng et al., 2005). But the activity of this region could not be excluded. *roX1* transcripts from a major transcription start site contain this region but lack region 1 and most of region 2. Possible redundancy of region 3 with the other two regions is therefore likely. These 5' fragments (*roX1*^{R1}, bp -85–310; *roX1*^{R2}, bp 190–805; *roX1*^{R3}, bp 790–1490 of *roX1* (numbering from (Amrein and Axel, 1997)) were fused to the 3' fragment (bp 3150–4150) that contains the essential stem loop, *roX* boxes and *roX1* termination signals (Figure 3-2). The sufficiency of this 3' fragment is implied by the survival of *roX1*^{ex40A} *roX2* males (Figure 3-2). In order to drive uniform

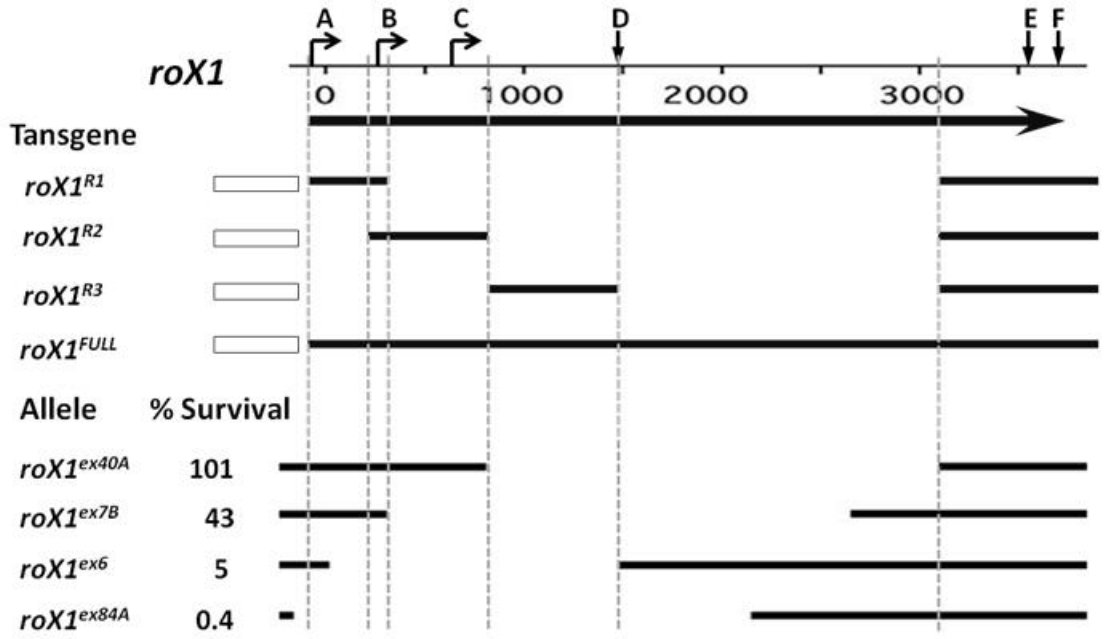

Figure 3-2. Schematic structure of the *roX1* gene and *roX1*^{R1}, *roX1*^{R2} and *roX1*^{R3} transgenes. Alignment of *roX1*^{R1}, *roX1*^{R2} and *roX1*^{R3} sequences shows the 5' fragments that were fused to the essential 3' end. Break points of selected excision mutations that contributed to the design of these *roX1* transgenes are shown below. The top line indicates the *roX1* genomic region, numbering from (Amrein and Axel, 1997). The arrows above are: A. The start sites of a minor class of *roX1* ESTs (expressed sequence tags) at -47 bp. B. A popular *roX1* transcription start site at 259 bp. C. The transcription start site of a major *roX1* transcript at 628 bp. D. the site of the *roX1*^{mb710} P-element insertion. E and F. the site of polyadenylation signals. Below are the *roX1* transgenes generated for this study (top 4) and selected excision mutations (bottom 4) that are illustrated with respect to the transcribed region of *roX1* (arrow). The *hsp83* promoter is shown as white box in front of each *roX1* transgene. The survival of selected *roX1* excision mutations is from (Deng et al, 2005)

expression, the strong, constitutive *hsp83* promoter was used. Transgenic flies carrying each construct were generated. To avoid differences due to position effect, all transgenes were integrated at the same site at 68A4 on the third chromosome using the site-specific ΦC31 integrase system (Groth et al., 2004). To enable comparison between X-linked and autosomal insertions, the $roX1^{R1\text{-}3}$ transgenes were also inserted at 2A3, 16C1, 18D3 and 19C4 on the X chromosome by the same site-specific system (see Materials and Method). For a control, a full length *roX1* transgene, $roX1^{FULL}$ (bp -78–4150 of *roX1*), were also generated and integrated at 68A4 and 16C1 by the ΦC31 integrase system (Figure 3-2).

Rescue of *roX1 roX2* male lethality was used to measure the biological activity of *roX* transgenes. Males carrying $roX1^{R1\text{-}3}$ transgenes at the autosomal site 68A4 were mated to $roX1^{SMC17A}$ *roX2* virgin females. The only source of *roX* RNA in the sons is from the $roX1^{R1\text{-}3}$ transgene. The full length *roX1* transgene, $roX1^{FULL}$, driven by *hsp83* and integrated at the 68A4 site, supports over 90% male survival (Table 3-4). As shown in Table 3-4, male viability is partially rescued by each of the three deleted *roX1* transgenes. $roX1^{R2}$ supports highest male survival, 54%, and $roX1^{R1}$ supports lowest male survival, 33%. Rescue of $roX1^{mb710}$ *roX2* chromosome is similar. Low rescue rate was expected, as each *roX1* transgene is deleted for most of the 5' end of *roX1*. However, all *roX1* transgenes are still capable of partially rescuing *roX1 roX2* males. This confirms previous studies suggesting that the 5' end of *roX1* contains multiple functional elements with at least partial redundancy. All subsequent studies of transgene

Table 3-4. Survival of *roX1 roX2* males rescued by *roX1*^{R1}, *roX1*^{R2} and *roX1*^{R3} transgenes integrated at 68A4 on third chromosome.

X chromosome	% male survival with autosomal transgene				
	No TG	[*roX1*^{R1}]	[*roX1*^{R2}]	[*roX1*^{R3}]	[*roX1*^{FULL}]
roX1^{SMC17A} *roX2*	7(978)	33(1722)	54(1744)	43(1725)	96(643)
roX1^{mb710} *roX2*	5(1680)[a]	44(352)	51(416)	49(313)	NA

roX1 roX2 females were mated to *yw* males homozygous for a *roX1* transgene at 68A4. The male and female adults were counted. Female viability is set as 100%. Male survival was expressed as the male to female ratio of offspring. The total number of female adults counted in each case is in parentheses.
No TG, no transgene.
NA, not available.
[a] Meller and Rattner, 2002.

effect utilize the *roX1*^{SMC17A} *roX2* chromosome.

Restoration of MSL complex localization to the X chromosome by *roX1*$^{R1-3}$ transgenes.

roX1 roX2 males display reduced MSL localization to the X chromosome and ectopic MSL binding to the chromocenter and autosomal sites (Deng et al., 2005; Meller and Rattner, 2002). To assess the role of 5' *roX1* sequences in MSL localization, we prepared polytene chromosome spreads from male *roX1*^{SMC17A} *roX2*/Y*; roX1*$^{R1-3}$/+ larvae and detected MSL1 localization by immunostaining. *roX1*^{FULL}, also integrated at the 68A4 site, is used as a control. *roX1*$^{R1-3}$ partially restore MSL recruitment to the *roX1*^{SMC17A} *roX2* X chromosome, but to different extents (Figure 3-3). All three autosomal *roX1* transgenes support high levels of MSL complex binding to the X chromosome, but less than wild type or *roX1*^{SMC17A} *roX2*/Y*; roX1*^{FULL}/+ males (Figure 3-3 A, C). Some ectopic MSL binding is still observed, especially at the chromocenter of males carrying *roX1*$^{R1-3}$. By contrast, *roX1*^{FULL} supports wild type level of MSL complex binding to the X chromosome and no ectopic MSL binding except the *roX1* integration site (compare panels A, C, D, E and F of Figure 3-3). Among all *roX1*$^{R1-3}$–rescued *roX1*^{SMC17A} *roX2* males, the least ectopic MSL binding was observed in males carrying *roX1*^{R1}. Males carrying *roX1*^{R2} retained the most prominent ectopic binding of MSL at the chromocenter. These results suggested that *roX1* 5' end contains redundant elements that act co-operatively to direct MSL localization to the X chromosme. Comparing this with the ability of *roX1*$^{R1-3}$ to rescue male survival reveals that the

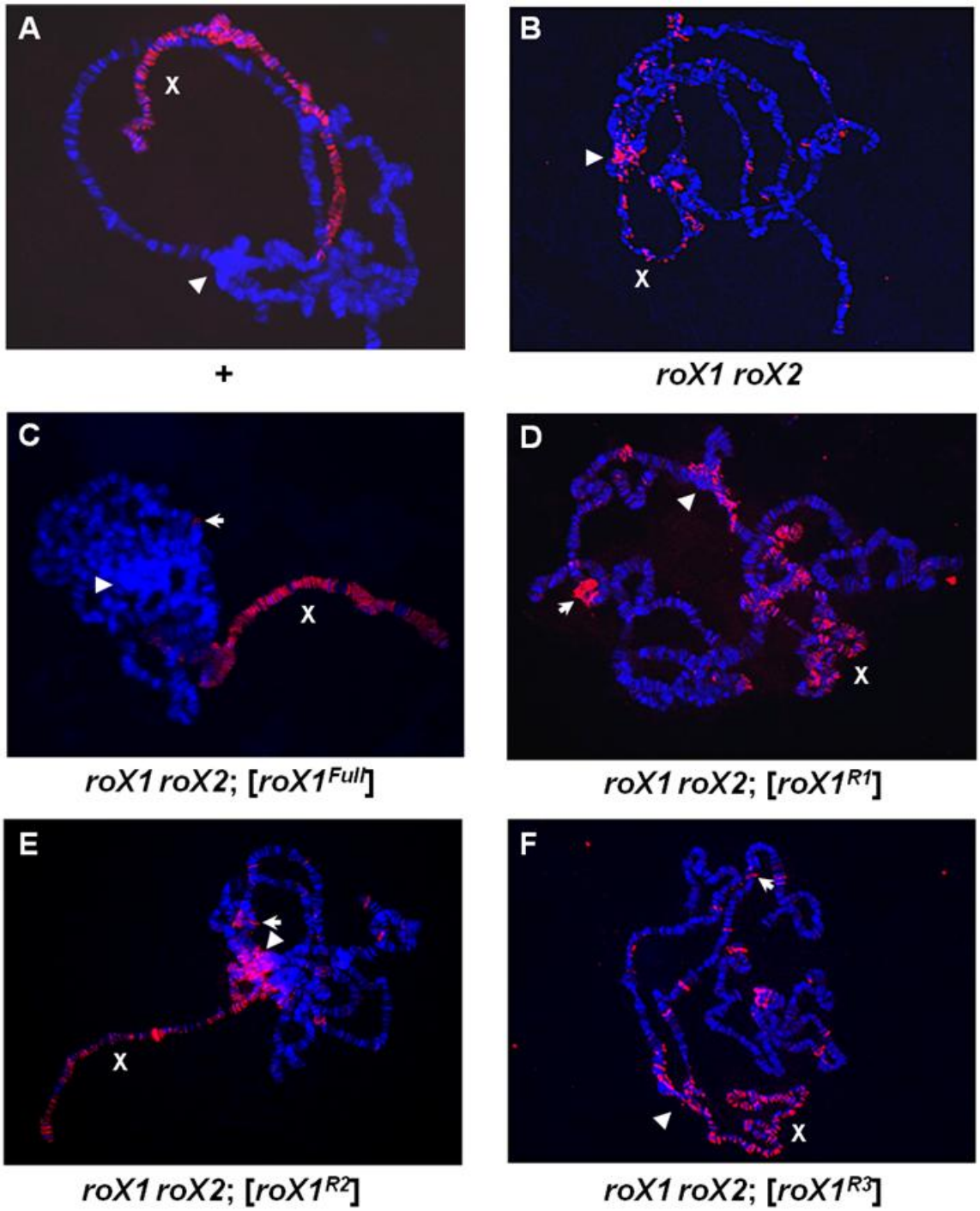

Figure 3-3. Autosomal *roX1* transgenes partially restore MSL complex localization to the X chromosome. MSL1 immunostaining of polytene chromosomes from male larvae of **(A)** wild type, **(B)** *roX1*^{SMC17A} *roX2*, **(C)** *roX1*^{SMC17A} *roX2*; [*roX1*^{Full}], **(D)** *roX1*^{SMC17A} *roX2*; [*roX1*^{R1}], **(E)** *roX1*^{SMC17A} *roX2*; [*roX1*^{R2}], **(F)** *roX1*^{SMC17A} *roX2*; [*roX1*^{R3}]. *roX1* transgenes are integrated at 68A4 on third chromosome. The chromocenter is indicated by the triangle. The integration site is indicated by the arrow. X marks the X chromosome. MSL1 is detected with Texas Red (red). DNA is detected with DAPI (blue).

ability of each transgene to promote X-localization does not agree precisely with its ability to rescue males. *roX1*^{R2} support the highest male survival in spite of strong ectopic binding of MSL at the chromocenter. However, although it directs good MSL binding on X, *roX1*^{R1} support the lowest male survival.

MLE is an RNA/DNA helicase which may not associate with other MSL proteins directly (Lee et al., 1997; Smith et al., 2000). It is possible that MLE associates with the MSL complex by interaction with *roX* RNAs (Meller, 2003; Richter et al., 1996). To determine whether *roX1* transcripts from autosomal *roX1*$^{R1-3}$ transgenes are able to recruit MLE to the X chromosome, I performed MLE immunostaining on polytene chromosome spreads from male *roX1* ^{SMC17A} *roX2; roX1*$^{R1-3}$ larvae. For *roX1*^{R2} and *roX1*^{R3} rescued males, the binding pattern of MLE is indistinguishable from that of MSL1 (Figure 3-4 C and D). This suggests that RNA transcribed from the *roX1*^{R2} and *roX1*^{R3} transgenes enables MLE to associate with the MSL proteins and localize to the X chromosome. However, I could not see any restoration of MLE localization to the X chromosome by *roX1*^{R1}. Therefore, the *roX1*^{R1} RNA appears unable to interact with MLE, or to mediate its integration into MSL complex. This could partially explain the observations that although *roX1*^{R1} restores MSL1 localization to the X as well as *roX1*^{R2} and *roX1*^{R3}, the survival of males rescued by *roX1*^{R1} is lower. This also suggests that although there is some redundancy among *roX1*$^{R1-3}$, these segments of *roX1* may also contain distinct functions.

Accumulation of *roX1*$^{R1-3}$ is determined by the 5' sequence.

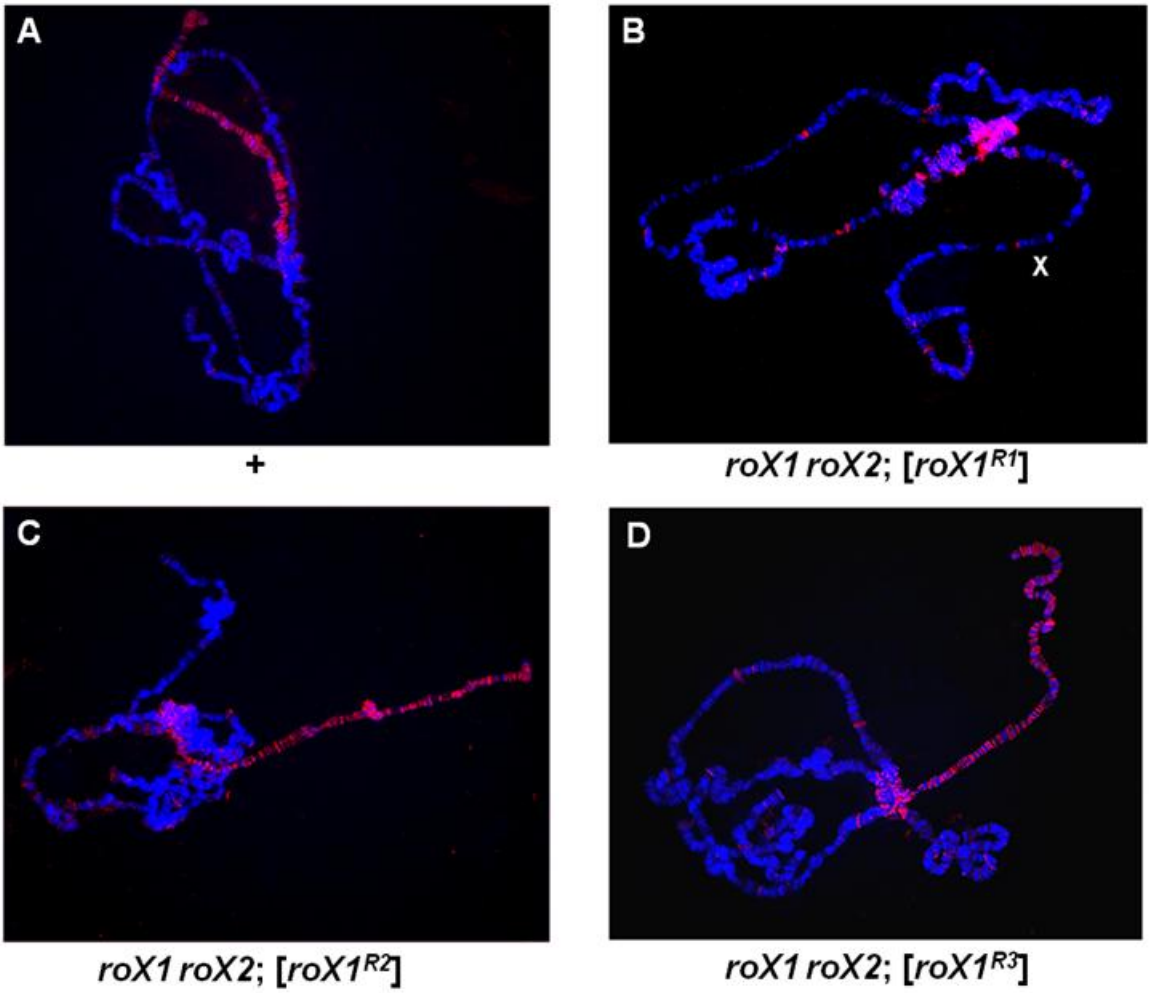

Figure 3-4. MLE immunostaining of polytene chromosomes from male larvae. (A) wild type, **(B)** *roX1*^{SMC17A} *roX2*; [*roX1*^{R1}], **(C)** *roX1*^{SMC17A} *roX2*; [*roX1*^{R2}], **(D)** *roX1*^{SMC17A} *roX2*; [*roX1*^{R3}]. *roX1* transgenes are integrated at 68A4 on the third chromosome. X marks the X chromosome. MLE is detected with Texas Red (red). DNA is detected with DAPI (blue).

Quantitative reverse transcriptional real-time PCR (Q-RT-PCR) was used to determine the accumulation of transcripts from all transgenes. As expected, *roX1*^{SMC17A} *roX2* males showed no detectable *roX1* RNA accumulation, which confirmed that the transgene was the sole source of *roX1* transcript in the males we tested (Figure 3-5). As *roX1*$^{R1-3}$ transgenes are driven by the same promoter and inserted at the same site, similar levels of transcription are anticipated. However, the levels of *roX1* transcript detected by Q-RT-PCR are different for each transgene. All *roX1*$^{R1-3}$ transgenes produce much lower level of *roX1* RNA than wild type (*yw*) males when integrated at the autosomal site 68A4. By contrast, the full length transgene, *roX1*^{FULL}, which is integrated at the same site, accumulated about 3.5-fold more *roX1* RNA than wild type males (Figure 3-5). *roX1*^{R2} showed 7-fold higher *roX1* RNA level than *roX1*^{R1}, and transcript from *roX1*^{R3} was virtually undetectable (Figure 3-5, black bars). We then compared accumulation of *roX1* transcripts produced from *roX1*$^{R1-3}$ transgenes inserted at 68A4 (autosomal), 16C1 (X-linked) and 2A3 (X-linked). The trend of transcript accumulation was essentially the same for *roX1*$^{R1-3}$ (*roX1*^{R2} highest, *roX1*^{R3} lowest), regardless of the site of integration (Figure 3-5, black, gray and brown bars). The high accumulation of stable transcript from *roX1*^{FULL} confirms that transcription driven by the *hsp83* promoter is high. This suggests that the dramatic difference in accumulation of transcripts from *roX1*$^{R1-3}$ reflects the stability of the RNA produced. *roX1* RNA stability therefore depends on specific sequences of *roX1* transcript. Deletion of large amounts of *roX1* sequence in *roX1*$^{R1-3}$ transgenes decreases the transcript stability dramatically.

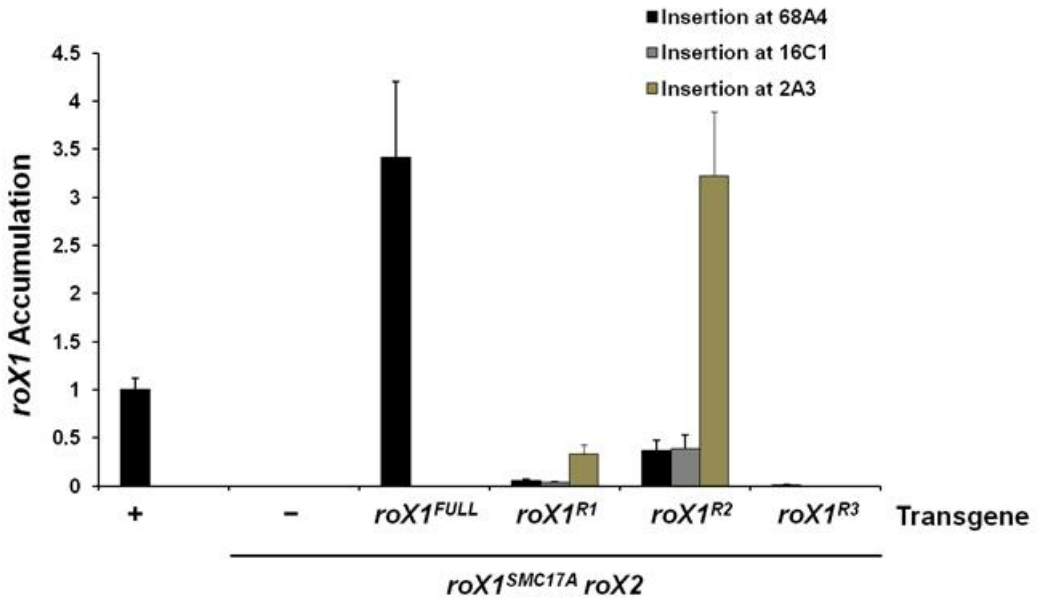

Figure 3-5. *roX1* transcript levels in *roX1 roX2* male larvae carrying *roX1* transgenes. *roX1* expression in the wild type control (+) was set to 1. The Y axis is the level of *roX1* accumulation in transgenic males divided by that in the wild type control. Q-RT-PCR was normalized to the autosomal gene *Dmn*. Error bars represent standard deviations of three independent biological replicates.

Although the *roX1* transcripts produced from *roX1*^{R3} were undetectable by Q-RT-PCR, reverse-transcription followed by a regular PCR for 40 cycles confirmed the presence of the anticipated gene product. This is consistent with a previous study, which found that even undetectable levels of mutated *roX1* transcript can support dosage compensation and surprisingly high levels of male survival (Deng et al., 2005). PCR, followed by sequencing, revealed alternative splicing within the 5' region of all *roX1*$^{R1-3}$ transgenes (Figure 3-6). Since the region amplified by Q-RT-PCR is located in the 3' region of *roX1* transgenes and is present in all alternative splice forms, the accumulation of *roX1* transcripts detected by Q-RT-PCR reflects all of the *roX1* transcripts present (primers for Q-RT-PCR are presented in Table 3-3).

roX1^{R1} directs MSL spreading _in cis_, while _roX1_^{R2} and _roX1_^{R3} do not.

The different levels of transcript accumulation enabled us to test the role of *roX* abundance in spreading. The prevailing model of cotranscriptional assembly of MSL proteins onto nascent *roX* transcripts predicts that lower *roX* transcript accumulation promotes local spreading from sites of *roX* production, but when *roX* transcripts are more abundant, local spreading is less likely to occur (Kelley et al., 2008; Oh et al., 2003; Park et al., 2002). We therefore examined polytene preparations for evidence of MSL localization at 68A4, the autosomal insertion site of our *roX1*$^{R1-3}$ transgenes. We observed that MSL binds to, but does not spread from, the *roX1*^{R2} and *roX1*^{R3} insertions. In contrast, strong MSL recruitment, and discontinuous spreading into flanking chromatin, was seen at

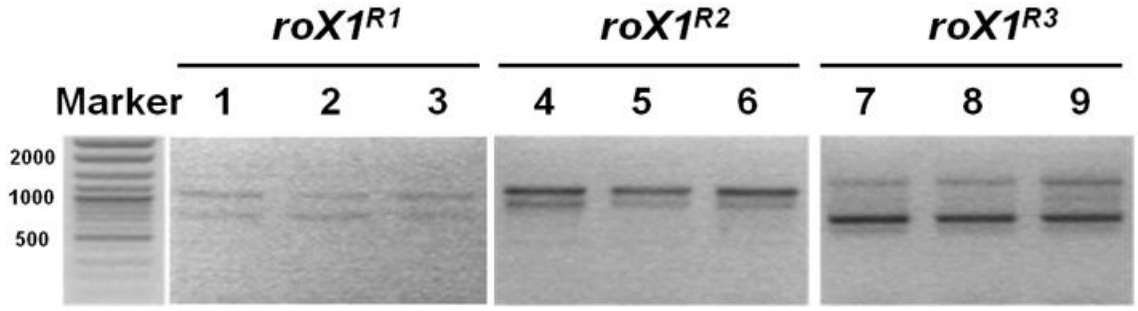

Figure 3-6. Reverse transcription followed by PCR reveals multiple *roX1* transcripts from the *roX1*$^{R1-3}$ transgenes. The template was the same cDNA used for Q-RT-PCR to detect the *roX1* transcript level. The cDNAs were reverse transcribed from three independent RNA samples extracted from 50 third instar larvae of each genotype. Lane 1-3, *roX1*^{SMC17A} *roX2*/Y; [*roX1*^{R1}]/+. Lane 4-6, *roX1*^{SMC17A} *roX2*/Y; [*roX1*^{R2}]/+. Lane 7-9, *roX1*^{SMC17A} *roX2*/Y; [*roX1*^{R3}]/+. The marker is 100bp DNA ladder. The size of selected band is labeled on left of the band. The highest band of each lane corresponds to the unspliced full length transcript from *roX1*$^{R1-3}$.

the *roX1*^{R1} insertion site (Figure 3-3 arrow head, Figure 3-7). This is contrary to our expectation that *roX1*^{R3}, with lowest RNA accumulation, would spread most readily from a transgene integration site. We conclude that transcript accumulation is not the sole determinant of spreading in *cis*, and that the *roX1*^{R1} transgene either contains specific sequences that promote *in cis* spreading of the MSL complex, or lacks sequences that block spreading *in cis*.

Transcript accumulation does not determine *roX1*$^{R1-3}$ activity.

roX1 roX2 mutant males show a global decrease in X-linked gene expression (Deng and Meller, 2006). To test whether *roX1*$^{R1-3}$ transgenes are able to rescue X-linked gene expression in *roX1 roX2* males, individual genes were measured by Q-RT-PCR in *roX1*^{SMC17A} *roX2* males carrying each transgene. As expected, X-linked gene expression was partially restored by *roX1*$^{R1-3}$ transgenes (Figure 3-8). Introduction of a copy of a *roX1*$^{R1-3}$ transgene increases expression of all X-linked, dosage compensated genes that were tested, as compared to *roX1*^{SMC17A} *roX2* males. While the increase was variable and often small (between 6 and 54%), in most cases the change was significant ($p<0.01$, Figure 3-8).

Beadex mutations are dose-sensitive, gain of function alleles of *Dlmo* (Shoresh et al., 1998). The X-linked *Dlmo* gene is dosage compensated by the MSL complex, making *Beadex* mutations sensitive reporters for MSL complex activity (Menon and Meller, 2009). Flies carrying *Beadex* mutations display notching of wing margins, and the notching becomes more severe with increased

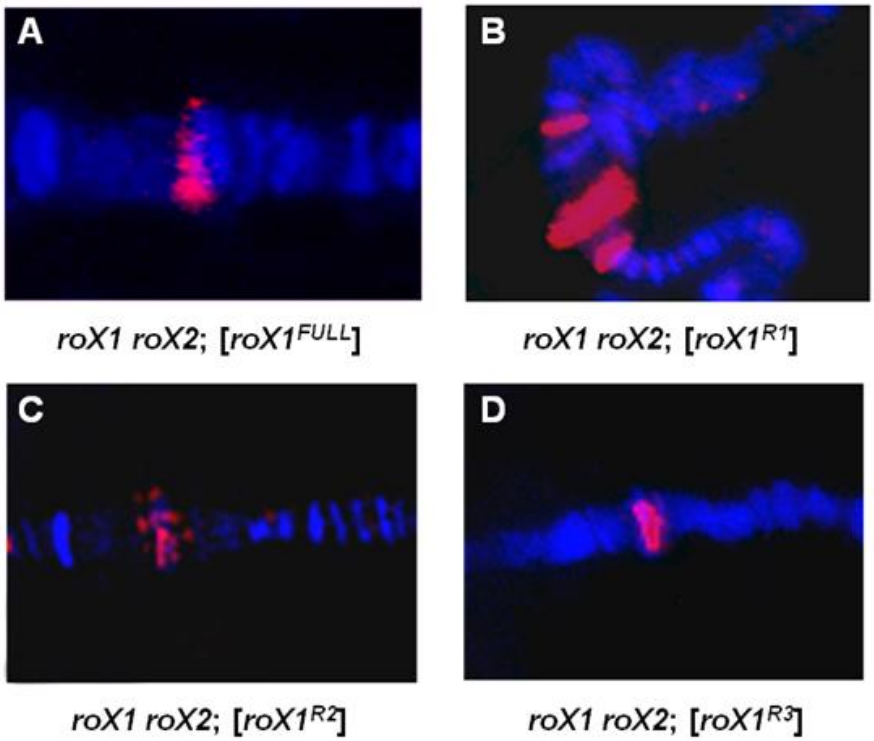

Figure 3-7. The autosomal integration site of 68A4 enlarged to show MSL binding and spreading from *roX1*^{R1}, and binding but no spreading from *roX1*^{FULL}, *roX1*^{R2} and *roX1*^{R3}. MSL1 immunostaining of polytene chromosomes from male larvae of **(A)** *roX1*^{SMC17A} *roX2*; [*roX1*^{Full}], **(B)** *roX1*^{SMC17A} *roX2*; [*roX1*^{R1}], **(C)** *roX1*^{SMC17A} *roX2*; [*roX1*^{R2}], **(D)** *roX1*^{SMC17A} *roX2*; [*roX1*^{R3}]. MSL1 is detected with Texas Red (red). DNA is detected with DAPI (blue).

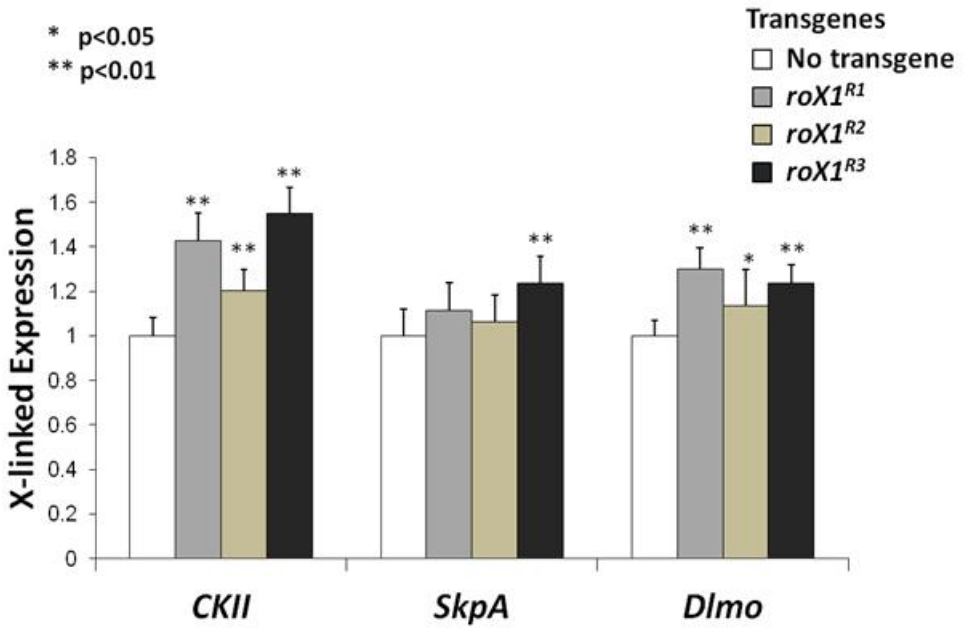

Figure 3-8. Rescue of X-linked gene expression by *roX1*$^{R1-3}$ transgenes. Three X-linked dosage compensated genes were analyzed by Q-RT-PCR in *roX1*^{SMC17A} *roX2* male larvae carrying *roX1*$^{R1-3}$ transgene at 68A4. Expression in *roX1*^{SMC17A} *roX2* males was set to 1. Values are the ratio of gene expression in *roX1*^{SMC17A} *roX2* males carrying a *roX1* transgene to that in *roX1*^{SMC17A} *roX2* males without a transgene. Q-RT-PCR was normalized to the autosomal gene *Dmn*. Error bars represent standard deviations of three independent biological replicate. P-values are determined by performing a two-sample unpaired t-test between each genotype and control.

expression of *Dlmo*. *Beadex* reporter analysis confirmed Q-RT-PCR findings (Figure 3-9). There is no notching of wing margins in females carrying one *Beadex* allele (Bx^{r49k}) and one wild type allele. $roX1^{mb710}$ *roX2* Bx^{r49k} males show a low level of notching of wing margins (Menon and Meller, 2009), because of the lower expression of *Dlmo* due to the failure of dosage compensation. $roX1^{mb710}$ *roX2*Δ Bx^{r49k} is a more severe allele that is lethal because of failure of dosage compensation. No adult male is recovered from this allele. Survival is partially rescued by introducing one copy of a $roX1^{R1-3}$ transgene. These males show high levels of notching of wing margins, suggesting that X chromosome dosage compensation is improved. The level of notching is consistent with the expression level of X-linked genes detected by Q-RT-PCR. $roX1^{R2}$ showed least notching, as well as least expression of *Dlmo* as detected by Q-RT-PCR, while $roX1^{R1}$ and $roX1^{R3}$ were quite similar. However, there is no correlation between the level of transcript from the $roX1^{R1-3}$ transgenes and X-linked gene expression levels. Transcript accumulation does not predict the ability of $roX1^{R1-3}$ to rescue *roX1 roX2* male viability, or to direct MSL spreading (Figure 3-5). Taken together, these studies revealed that *roX1* transcript levels alone do not determine activity.

$roX1^{R1}$ has a unique activity.

The ability of the extreme 5' end of *roX1* to direct MSL spreading *in cis* from its autosomal transcription site (Figure 3-7**)** suggests that it might have a role in marking the X chromosome. Both *roX* genes are located on the X

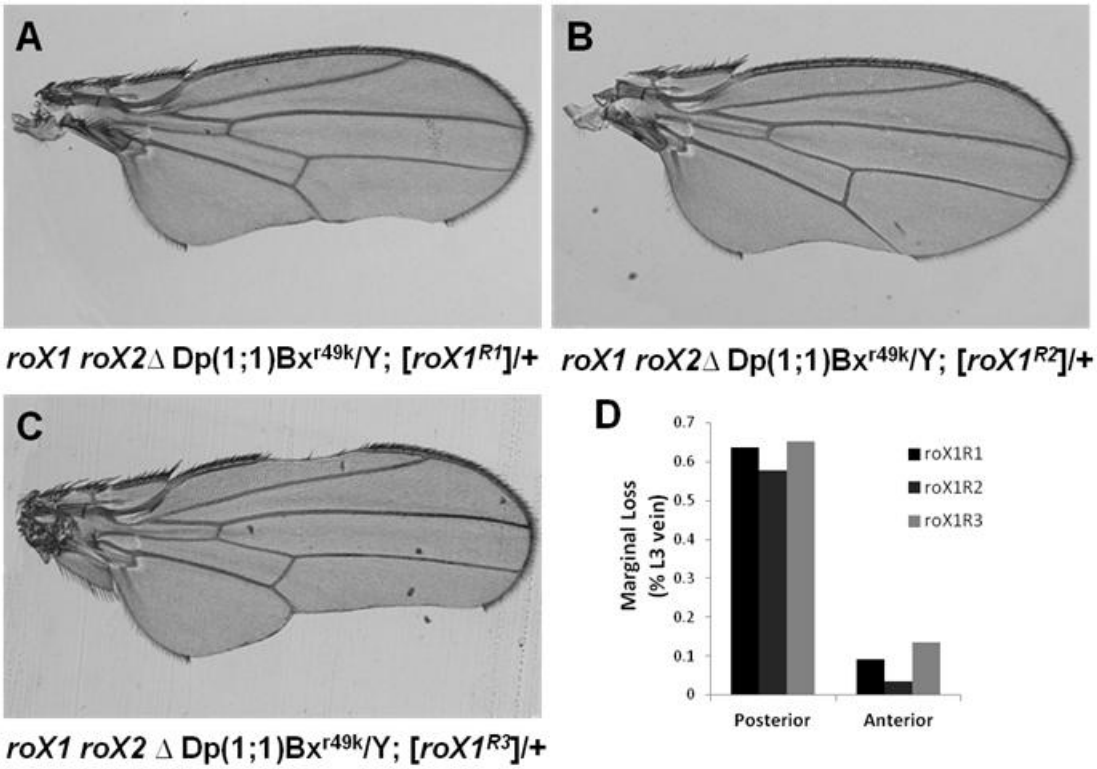

Figure 3-9. *Beadex* responds to activity of *roX1*$^{R1-3}$ transgenes. Wing from **(A)** *roX1*^{mb710} *roX2*Δ Dp(1;1)*Bx*^{r49k}/Y; [*roX1*^{R1}]/+ male, **(B)** *roX1*^{mb710} *roX2*Δ Dp(1;1)*Bx*^{r49k}/Y; [*roX1*^{R2}]/+ male, **(C)** *roX1*^{mb710} *roX2*Δ Dp(1;1)*Bx*^{r49k}/Y; [*roX1*^{R3}]/+ male. **(D)** The amount of wing margin lost is represented as the percentage of L3 vein length. Twenty wings from each group were measured.

chromosome, and it has been suggested that this helps target dosage compensation to the X chromosome. However, full length autosomal *roX* transgenes can fully rescue survival of *roX1 roX2* males, calling into question the role of spreading in *cis*. To examine this, $roX1^{R1\text{-}3}$ transgenes were integrated at four more sites, 2A3, 16C1, 18D3 and 19C4, on the X chromosome. The 18D3 landing site, in an intron of the *Tao-1* gene, is viable but integration of any of the *roX1* transgenes into this site is lethal. Integrations at the other three sites are viable. $roX1^{R1\text{-}3}$ transgenes integrated at 2A3, 16C1, and 19C4 were recombined onto the $roX1^{SMC17A}$ *roX2* chromosome. $roX1^{SMC17A}$ *roX2* $roX1^{R1\text{-}3}$/Binsincy females were mated to *yw* males to produce $roX1^{SMC17A}$ *roX2* $roX1^{R1\text{-}3}$ sons. The only source of *roX* RNA in the sons is the X-linked $roX1^{R1\text{-}3}$ transgene. $roX1^{R2}$ and $roX1^{R3}$ transgenes integrated at 19C4 rescued $roX1^{SMC17A}$ *roX2* males somewhat less than the autosomal integration, while $roX1^{R1}$ from this site showed no rescue of $roX1^{SMC17A}$ *roX2* males (Table 3-5**)**. 19C4 is a heterochromatic region. The expression of any transgenes inserted in this region could be suppressed by the heterochromatic environment. We expect that the lower rescue is due to low expression of $roX1^{R1\text{-}3}$ transgenes. This is supported by the fact that the expression of the *white* marker gene at this site was variegating and hard to detect. When integrated at 16C1, $roX1^{R2}$ and $roX1^{R3}$ displayed similar rescue to that observed for integrations at the autosomal 68A4 site. Unexpectedly, $roX1^{R1}$ showed no rescue of $roX1^{SMC17A}$ *roX2* males from 16C1 (Table 3-4 and 3-5**)**. This was puzzling, as the accumulation of $roX1^{R1}$ is identical when it is integrated at 16C1 and 68A4 (Figure 3-5). When integrated at

Table 3-5. Survival of *roX1 roX2* males that are rescued by *roX1*^{R1}, *roX1*^{R2} and *roX1*^{R3} transgenes inserted on the X chromosme.

Transgene Integration Site	% male survival with X-liked transgenes [*roX1*^{R1}]	[*roX1*^{R2}]	[*roX1*^{R3}]	[*roX1*^{FULL}]
2A3	16(1010)	70(1517)	60(1205)	NA
16C1	9(5150)	52(1865)	33(2546)	99(570)
19C4	8(679)	29(316)	31(304)	NA

To rescue *roX1 roX2* males by X-linked *roX1* transgene, *roX1*$^{R1-3}$ transgenes inserted at the indicated sites were recombined to *roX1*^{SMC17A} Df(1)52 X chromosome. *roX1* Df(1)52 [*roX1*$^{R1-3}$]/Binsincy females were mated to *yw* males homozygous for [$w^{+}4\Delta4.3$]. The non-Binsincy male and female adults were counted. The female viability is set at 100%. Male viability was expressed as the ratio of males to females recovered. The total number of female adults counted in each case is in parentheses.
NA: not available.

another X-linked site, 2A3, *roX1*^{R2} and *roX1*^{R3} displayed higher male rescue than all other sites. *roX1*^{R1} also showed higher male rescue from this site than from 16C1 and 19C4 (Table 3-5**)**. This is consistent with a higher transcript accumulation of *roX1*^{R1} and *roX1*^{R2} transcribed from this site than from 16C1 and 68A4 (transcript from *roX1*^{R3} is still undetectable, Figure 3-5 brown bars). However, the rescue of *roX1*^{SMC17A} *roX2* males by *roX1*^{R1} integrated at 2A3 was only about half of that by *roX1*^{R1} integrated on an autosome. Because the *roX1*^{R1} transgene has a unique ability to direct spreading of the MSL complex *in cis*, we hypothesized that the failure of rescue was due to abnormal local accumulation of MSL complex adjacent to the site of *roX1*^{R1} integration. We reasoned that this prevents spreading to dosage compensate the rest of the X chromosome.

To resolve this question, I asked whether genetic manipulations that reduce MSL spreading increase male survival when *rox1*^{R1} is X-linked. Reduced levels of the MSL proteins MSL1 or MSL2 inhibit MSL spreading in *cis* from a *roX* transcription site (Oh et al., 2003). To reduce the dose of MSL1, I introduced a loss of function copy of *msl1* (*msl1*^{L60}) by mating *roX1*^{SMC17A} *roX2 roX1*^{R1} 16C1 females to males carrying this mutation. Reduction of MSL1 did not rescue males with X-linked *roX1*^{R1} at 16C1 (Figure 3-10). Indeed it dramatically decreased *roX1*^{SMC17A} *roX2 roX1*^{R1} male survival. In contrast, reduction in the level of MSL3 and MLE had no discernable effect on male survival (Figure 3-10). Similar results were obtained with *roX1*^{R2} and *roX1*^{R3} transgenes integrated at the same site (Figure 3-10). This is consistent with previous studies demonstrating that the level of MSL1, but not MSL3 or MLE, is limiting for dosage compensation.

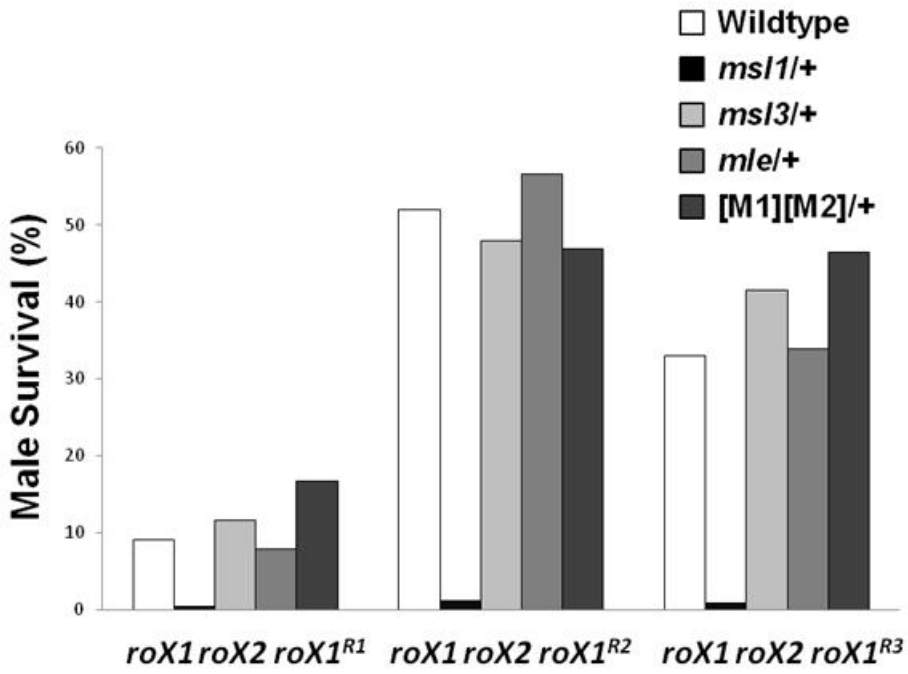

Figure 3-10. Survival of *roX1*^{SMC17A} *roX2 roX1*$^{R1-3}$ males after manipulation of MSL protein levels. *roX1*$^{R1-3}$ transgenes were located at 16C1 on the X chromosome. The survival of males is based on emergence of females of the same genotype, with the exception that females carry one wild type X chromosome and one *roX1 roX2 roX1*$^{R1-3}$ chromosome (female viability is set to 100%, see Materials and Methods for a description of crosses). The X chromosome genotype of each male is depicted on the X axis. The labels on the top right is the genotype of autosome in each case. "+" means wild type. The *msl* mutants used are *msl1*^{L60}, *msl3*^{2} and *mle*^{1}. [M1][M2] chromosome carries *hsp83*-driven transgenes that overexpress *msl1* and *msl2*. Full genotype is *roX1*^{SMC17A} Df(1)52 *roX1*$^{R1-3}$; [w^{+}4Δ4.3].

Simultaneous overexpression of MSL1 and MSL2 from transgenes increased the survival of *roX1*^{SMC17A} *roX2 roX1*^{R1} males by almost two fold, but *roX1*^{SMC17A} *roX2 roX1*^{R2} and *roX1*^{SMC17A} *roX2 roX1*^{R3} males did not benefit from MSL overexpression (Figure 3-10). Overexpression of MSL1 and MSL2 promotes speading in *cis* from *roX* genes and transgenes (Oh et al., 2003; Park et al., 2002). My results indicate that uneven localization of the MSL complex along the X chromosome is not the reason why *roX1*^{R1} can not rescue from 16C1 on the X chromosome. My findings instead suggest that the failure of *roX1*^{R1} to support dosage compensation when transcribed from the X chromosome is the reason for low male survival.

To investigate whether the failure of *roX1*^{R1} to support dosage compensation is a result of failure of proper localization of the MSL complex on the X chromosome, we prepared polytene chromosome spreads from male *roX1*^{SMC17A} *roX2 roX1*$^{R1-3}$ larvae and detected MSL1 localization by immuno-staining. As shown in figure 3-11, when integrated at 16C1, *roX1*^{R1} failed to restore MSL complex localization to the X chromosome, while *roX1*^{R2} and *roX1*^{R3} partially restored MSL complex localization to the X chromosome. Therefore, *roX1*^{R1} behaves differently in directing the localization of the MSL complex. It is able to direct MSL localization *in trans* from an autosomal site, but not from this X-linked site.

I then asked whether *roX1*^{R1} is toxic *in cis* to the X chromosome. I reasoned that there might be a very dose-sensitive gene close to the 16C1 integration site. The local accumulation of MSL complex in *roX1*^{SMC17A} *roX2*

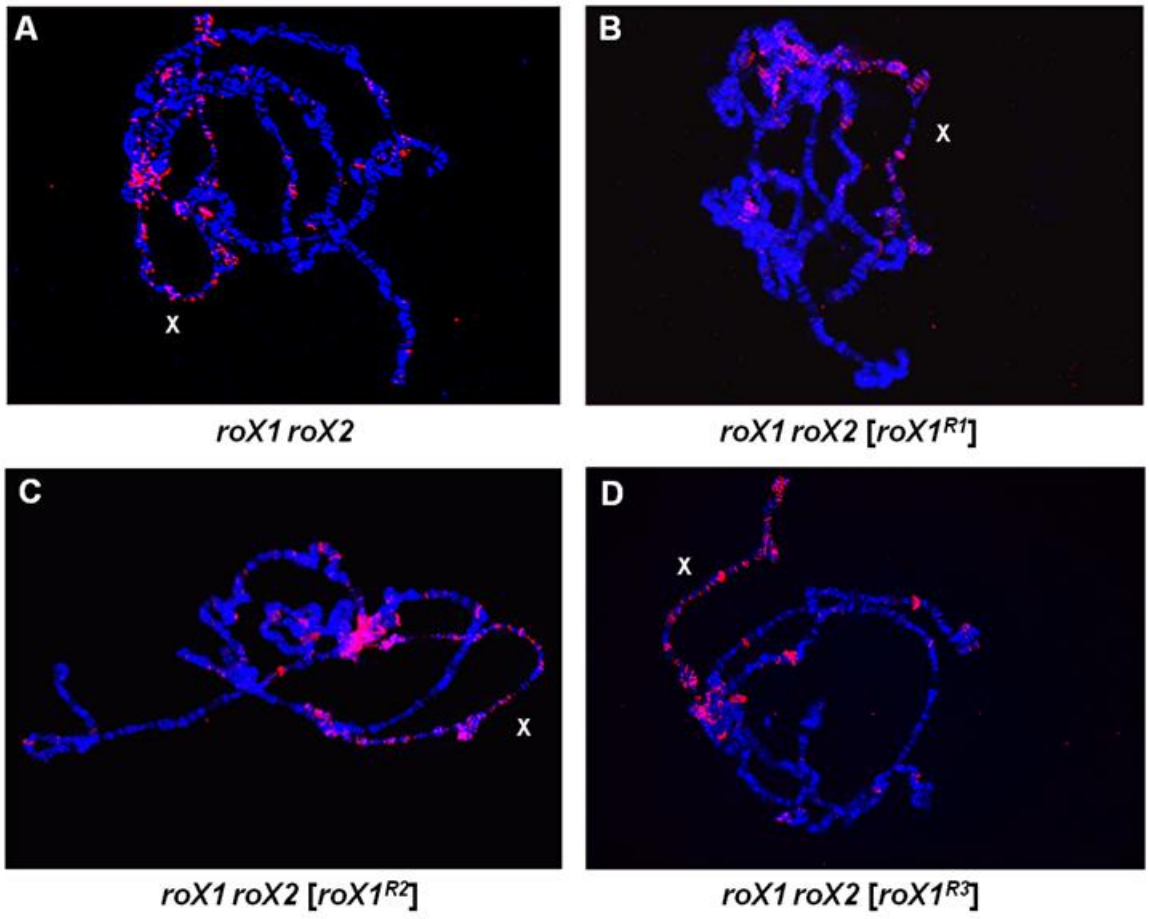

Figure 3-11. X-linked *roX1* transgenes partially restore MSL complex localization to the X chromosome. MSL1 immunostaining of polytene chromosomes from male larvae of **(A)** *roX1*^{SMC17A} *roX2*, **(B)** *roX1*^{SMC17A} *roX2* [*roX1*^{R1}], **(C)** *roX1*^{SMC17A} *roX2* [*roX1*^{R2}], **(D)** *roX1*^{SMC17A} *roX2* [*roX1*^{R3}]. *roX1* transgenes are integrated at 16C1 on X chromosome. X marks the X chromosome. MSL1 is detected with Texas Red (red). DNA is detected with DAPI (blue).

roX1^{R1} males might disrupt expression of this gene, leading to lethality. If this is the case, the *roX1*^{R1} transgene is likely to be toxic in a fly with wild type *roX1* supplied as well. A full length *roX1* transgene driven by *hsp83* and integrated at 68A4, *roX1*^{FULL}, was introduced to test this idea. The autosomal *roX1*^{FULL} almost fully rescues *roX1*^{SMC17A} *roX2 roX1*^{R1} males (Table 3-6). Transcript produced from *roX1*^{Full} is stable and accumulates to high levels compared to *roX1*^{R1} (Figure 3-5). This transcript could overcome the effect of *roX1*^{R1} transcript by competing for binding to MSL proteins. We next tested whether the much less stable *roX1* transcripts produced from *roX1*^{R2} and *roX1*^{R3} could rescue *roX1*^{SMC17A} *roX2 roX1*^{R1} males. Introducing a single copy of *roX1*^{R2} or *roX1*^{R3} integrated at 68A4 rescued *roX1*^{SMC17A} *roX2 roX1*^{R1} males to a similar extent as achieved by *roX1*^{R2} and *roX1*^{R3} when there is no *roX1*^{R1} transgene on the X chromosome (Table 3-6). Remarkably, *roX1*^{R1} at 68A4 also rescues *roX1*^{SMC17A} *roX2 roX1*^{R1} males. The extent of rescue is identical to that when there is no *roX1*^{R1} transgene at 16C1. Taken together, these results indicate that *roX1*^{R1} integrated at 16C1 is not toxic. Rather, it appears that the RNA produced from this site is uniquely unable to participate in some vital aspect of dosage compensation. When performing the same rescue of *roX1*^{SMC17A} *roX2 roX1*^{R2} and *roX1*^{SMC17A} *roX2 roX1*^{R3} by autosomal *roX1* transgenes, the effect of two transgenes together is not additive (Table 3-6). This suggests that different *roX1* transcripts do not work together to complement deficiencies in the function of each transcript.

Table 3-6. Rescue of *roX1 roX2 roX1^{R1-3}* males by autosomal *roX1* transgenes.

X chromosome	% male survival with autosomal transgenes				
	No TG	[*roX1^{R1}*]	[*roX1^{R2}*]	[*roX1^{R3}*]	[*roX1^{FULL}*]
roX1^{SMC17A} roX2	7(978)	33(1722)	54(1744)	43(1725)	96(643)
roX1^{SMC17A} roX2 roX1^{R1}	9(5150)	32(274)	40(252)	52(249)	83(563)
roX1^{SMC17A} roX2 roX1^{R2}	52(1865)	47(175)	59(170)	67(166)	97(274)
roX1^{SMC17A} roX2 roX1^{R3}	33(2546)	49(185)	61(186)	60(174)	115(234)

roX1^{SMC17A} roX2 roX1^{R1-3} X chromosomes carry the *roX1* transgene at 16C1. To rescue *roX1 roX2 roX1^{R1-3}* males by autosomal *roX1* transgenes, *roX1 roX2 roX1^{R1-3}*/Binsicy females were mated to *yw* males homozygous for the indicated *roX1* transgene at 68A4. Adults, all of which carry one copy of the autosomal *roX1* transgene, were counted. The female viability is set to 100%. Male viability was expressed as the percentage of males to females emerging from each cross. The total number of female adults counted in each case is in parentheses.
No TG: no autosomal transgene.

DISCUSSION

Both *roX* and *Xist* genes direct binding of chromatin complexes that coat the X chromosome to achieve dosage compensation. Both genes are able to recruit chromatin complexes in *cis* to the site of RNA production, but for neither molecule do we understand how this occurs at the molecular level. Our studies now reveal that a portion of *roX1* that is only present in a minor fraction of naturally occurring transcripts contributes to the MSL complex spreading *in cis*. When *roX1* transgenes are transcribed from an autosomal site, the MSL complex spreads discontinuously from the $roX1^{R1}$ transcription site to flanking regions. In contrast, it binds to but does not spread from the $roX1^{R2}$ and $roX1^{R3}$ transcription site (Figure 3-7). MSL spreading from the $roX1^{R1}$ transcription site shows a discontinuous pattern, as does spreading from a full length *roX1* autosomal transgene (Kelley et al., 1999; Meller et al., 2000). This observation argues that all the spreading occurs by a similar mechanism. However, the region of spreading from a wild type *roX1* transgene can cover an extensive region, while spreading from $roX1^{R1}$ is only into two adjoining, strongly-labeled bands.

It has been proposed that the MSL complex assembles co-transcriptionally on nascent *roX* transcripts, and spreads from these sites to adjacent, transcribed genes. The concept of spreading comes from the observed properties of heterochromatin. Heterochromatin can spread in *cis* from an initiation site along a chromosome for a long distance (Schuettengruber et al., 2007; Talbert and Henikoff, 2006; Wutz, 2003). This is believed to occur by linear propagation along

chromatin. Propagation involves silencing histone modifications (i.e., H3K9me) that serve to recruit proteins that make silencing histone modifications (i.e., Clr4), leading to spreading from sites of initiation. The mechanism of MSL complex spreading from a *roX* nucleation site to flanking regions is still unknown. It was found that a high ratio of MSL proteins to *roX* transcripts promoted local MSL spreading from sites of *roX* transcription (Kelley et al., 2008; Oh et al., 2003; Park et al., 2002). A model that local MSL spreading is controlled by the efficiency of MSL complex assembly onto nascent *roX* transcripts before their release was proposed. According to this model, there is an inverse relationship between the speed of MSL complex assembly and the ability of the MSL complex to spread from the site of *roX* transcription. Slower transcription or more MSL proteins produces more active MSL complex in the vicinity of the *roX* genes. This promotes binding to nearby chromatin. Decreased abundance of *roX* transcript also promotes assembly of intact MSL complexes prior to release of the *roX* transcript from the chromatin template. However, the $roX1^{R1}$ transgene has intermediate transcript accumulation compared to $roX1^{R2}$ and $roX1^{R3}$ (Figure 3-5). Therefore, the transcript abundance of $roX1^{R1}$, $roX1^{R2}$ and $roX1^{R3}$ does not predict their ability to direct MSL spreading *in cis*. We propose that the ability to direct spreading *in cis* is intrinsic to RNA or DNA sequences that are contained within the $roX1^{R1}$ transgene, but missing from $roX1^{R2}$ and $roX1^{R3}$.

All $roX1^{R1-3}$ transgenes are driven by the same promoter, inserted at the same site, and thus expected to be transcribed at the same level. However, each of the $roX1^{R1-3}$ transgenes displays a different level of transcript accumulation.

The *hsp83* promoter drives strong transcription of *roX1* transgenes. However, the accumulation of the transcripts from *roX1*$^{R1-3}$ is much lower than wild type. Transcript from *roX1*^{FULL} transgene, driven by the same promoter and inserted at the same site, showed much higher RNA accumulation than wild type. The high accumulation of the transcript from *roX1*^{FULL} confirms that transcription driven by the *hsp83* promoter is strong. This suggests that the dramatic difference in transcript accumulation reflects the stability of the RNA produced. *roX1*^{R3} is quite unstable compared to *roX1*^{R2} and *roX1*^{R1}, and *roX1*^{R2} is more stable than *roX1*^{R1}, but accumulation remains ten times less than *roX1*^{FULL} when expressed from the same site. *roX1*$^{R1-3}$ transgenes have deletion of large *roX1* sequences, which decreases the transcript stability dramatically. This indicates that *roX1* RNA stability needs cooperation of several *roX1* sequences, consistent with the idea that *roX1* contacts several proteins following integration into the MSL complex. However, we can not exclude the possibility that the transcription elongation rate of each *roX1* transgene is different, and this contributes to different transcript accumulation from each transgene. An *in vivo* transcription run-on assay could be used to investigate the transcription elongation rate from each *roX1* transgene. Binding of the MSL complex to the *roX1* transgene could interfere with the transcription elongation. However, the accumulation of the MSL complex at the integration site of *roX1* transgenes does not correlate with the accumulation of the transcript from corresponding *roX1* transgenes. Moreover, the *roX*^{FULL} transgene contains all 5' sequences, transcription of *roX*^{FULL} would presumably be limited by any 5' sequences that slow elongation. Since accumulation of

roX^{FULL} is highest, elongation alone does not explain the transcript accumulation levels from these transgenes. My data is most consistent with the stability of each transcript being primarily responsible for RNA accumulation.

Males carrying a *roX1* allele missing a large part of the 5' sequence display ectopic MSL binding to heterochromatic region and some autosomal sites and reduced binding on the X chromosome (Deng et al., 2005). All *roX1*$^{R1-3}$ transgenes with different fragments of the 5' end partially rescue male survival and MSL localization to the X chromosome (Table 3-4, Figure 3-3), supporting the idea that this region contains multiple elements necessary for targeting of the MSL complex to the X chromosome. The incomplete rescue of male survival and MSL localization by each *roX1*$^{R1-3}$ transgene indicates that several of these elements act co-operatively for MSL localization. As the *roX1* 5' end lacks obvious conserved sequences, this redundancy may operate at the level of RNA structure.

An internal deletion mutant of *roX1*, *roX1*^{ex40A}, that retains only the *roX1*^{R1} and *roX1*^{R2} 5' fragments and the 3' end, supports full male survival, leading to the conclusion that these are the only regions of *roX1* necessary for function (Deng et al., 2005). However, although the 5' sequence of *roX1*^{R3} is not present in the *roX1*^{ex40A} mutation, we clearly detect the function of this region by MSL localization driven by the *roX1*^{R3} transgene. This was expected since the most active *roX1* transcription start site is situated at bp 628, slightly upstream from the beginning of the 5' fragment cloned into *roX1*^{R3}.

Although *roX1*^{ex40A} restores full male viability, co-expression of *roX1*^{R1} and

roX1^{R2} transgenes, which together contain all 5' sequences present in *roX1*^{ex40A}, does not fully rescue male survival. This supports previous studies showing that neither 3' 1 kb nor 5' 2.5kb *roX1* fragments rescues male survival or MSL localization, which suggested that *roX* RNA fulfills its function through simultaneous interaction with different molecules (Meller and Rattner, 2002; Stuckenholz et al., 2003). We do not know if the MSL complex contains more than one *roX* RNA. However, my finding that *roX1*^{R1} and *roX1*^{R2} can not complement each other suggests that the MSL complex may contain a single *roX* RNA.

The *roX1*^{R2} and *roX1*^{R3} transgenes enable MSL1 and MLE to bind to the X chromosome, suggesting that *roX1*^{R2} and *roX1*^{R3} transcripts simultaneous interact with multiple MSL proteins. However, *roX1*^{R1} partially restores MSL1 but not MLE localization to X chromosome. The helicase activity of MLE, as well as an RB1 RNA-binding domain in this protein, is reported to be important for normal spreading of the MSL complex on the X chromosome and for movement of the *roX* RNAs from their sites of synthesis (Gu et al., 2000; Meller et al., 2000; Morra et al., 2008; Morra et al., 2011). It is puzzling that *roX1*^{R1} can promote MSL spreading from its production site, but does not recruit MLE. It is possible that the relatively short *roX1*^{R1} transcript does not require the helicase activity of MLE for assembly and spreading (Figure 3-12 A). Other possible models for *roX1*^{R1} directed spreading without binding to MLE can also be postulated. It is possible that the MSL complex requires dissociation of MLE to spread from the primary binding site to different sites, and that the MLE helicase activity is required for the

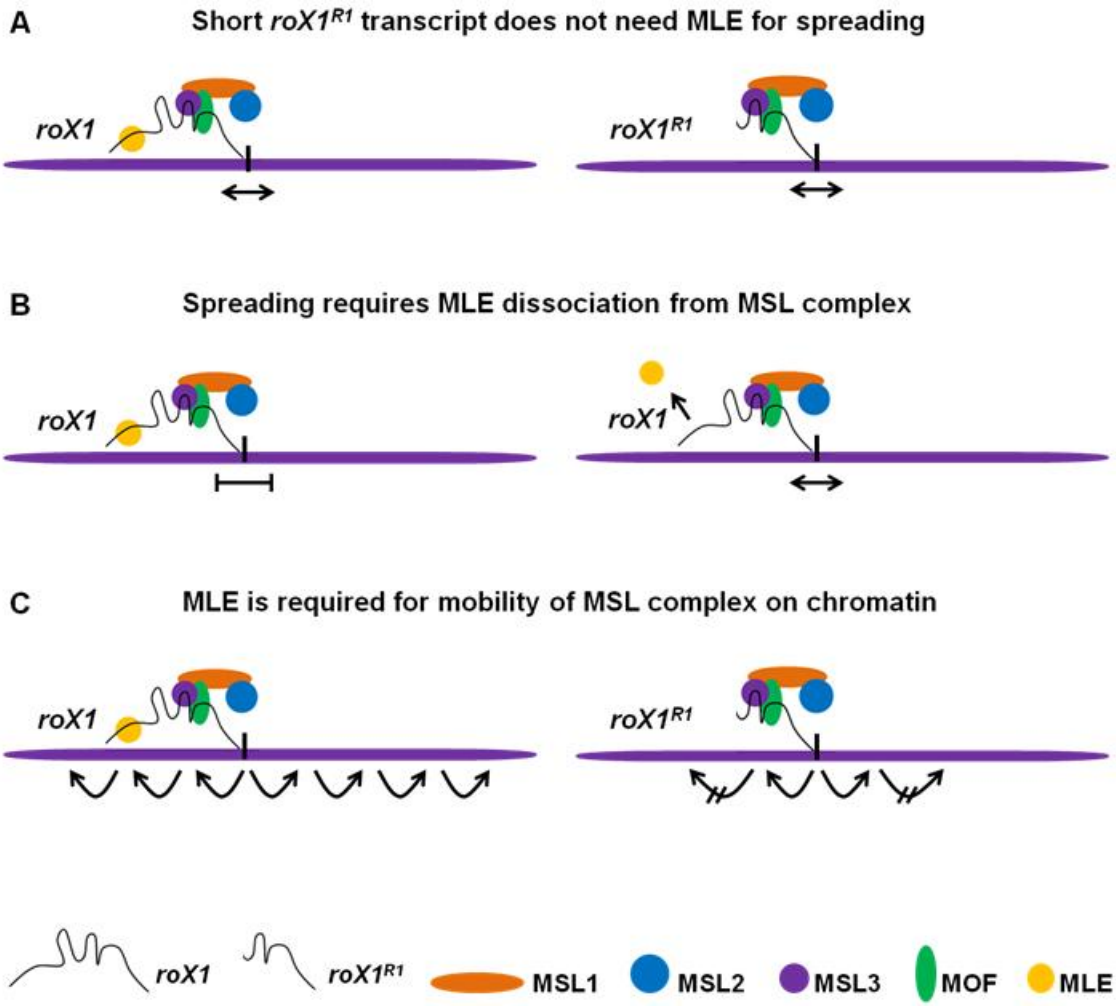

Figure 3-12. Models of *roX1*^{R1}-directed spreading of MSL complex from its transcription site. (A) Full length *roX1* requires the MLE helicase for movement from a site of transcription. Relatively short *roX1*^{R1} transcript does not need MLE. **(B)** MLE must dissociate from the MSL complex to allow spreading to a different site. Dissociation of MLE requires MLE helicase activity. Because *roX1*^{R1} never recruits MLE, it spreads *in cis* more readily. **(C)** MLE is required for repositioning of chromatin bound MSL complex. When MLE is incorporated into the MSL complex, the complex can move to other sites. However, without MLE, it is sequestered in the place of first association. This creates a high concentration of MSL complex locally, enhancing the appearance of spreading to regions surrounding the site of transcription.

dissociation. Since *roX1*^{R1} transcript never recruits MLE to the MSL complex, it can spread more readily (Figure 3-12 B). Another possible explanation is that MLE is necessary for mobility of MSL complex that has bound to chromatin. The MSL complex is thought to first bind to chromatin entry sites, or sites close to the *roX* genes, and then spread to other sites on the X chromosome. In this model, chromatin bound MSL complex can move to other sites only when MLE is present and active. Without MLE, the MSL complex is sequestered at the place of first association (Figure 3-12 C). This would create a high concentration of MSL complex surrounding the site of *roX* transcription (Figure 3-12 C). Although counterintuitive, this model suggests that what appears to be avid local spreading from a *roX1*^{R1} transgene is in fact the consequence of the inability of a partial complex lacking MLE to travel from its initial binding site. The resolution of polytene chromosome immunostaining is relatively low. Although MSL1 localization on an X chromosome rescued by autosomally produced *roX1*^{R1} is similar to that observed with *roX1*^{R2} and *roX1*^{R3}, we can not determine if the spreading of MSL complex on the X chromosome is normal. The effect of MLE on spreading of the MSL complex over short distances may only be detectable by high-resolution techniques such as ChIP-chip or ChIP-seq. For example, ChIP-seq was used to establish that the MSL3 chromo-domain is necessary for the MSL complex to spread to actively transcribed genes, in spite of the fact that immunostaining of polytene chromosome did not detect reduced MSL localization in an MSL3 chromo-domain mutant (Sural et al., 2008).

roX genes are normally located on and transcribed from the X

chromosome, and assemble with the MSL proteins to coat the X chromosome (Amrein and Axel, 1997; Meller et al., 1997). It has been suggested that X-linkage of the *roX* genes contributes to targeting MSL complex to the correct chromosome. Autosomal *roX1* transgenes enable the MSL complex to bind to the X chromosome in *trans* and rescue *roX1 roX2* males, even though *roX* RNAs are usually transcribed from X. If *roX1* is wild type, the insertion site is immaterial. However, studies on the partial loss of function $roX1^{ex7B}$ mutant suggest that partially functional *roX1* transgenes may be more active if they are inserted on the X chromosome. The $roX1^{ex7B}$ mutation has an internal deletion of 2.3 kb, leaving the same 5' sequence present in $roX1^{R1}$ (Figure 3-2). The $roX1^{ex7B}$ mutation on the X supports ~43% male survival, but an autosomal *hsp83*-driven $roX1^{ex7B}$ transgene that begins at bp 47, and is thus slightly 5' truncated, supports only ~14% male survival. However, male rescue by autosomal $roX1^{R1}$ is much higher (33%-44%). Although position effects become a confounding factor when insertions at different sites are compared, X-linked integrations of $roX1^{R1-3}$ transgenes do not consistently show higher activity. The lower male survival resulting from the autosomal $roX1^{ex7B}$ transgene thus appears to not be due to its autosomal location. The $roX1^{ex7B}$ transgene starts 120 bp downstream of the $roX1^{R1}$ start site. The sequence lacking in the $roX1^{ex7B}$ transgene is thus the likely cause of its decreased male rescue.

When $roX1^{R1-3}$ are expressed from a single integration site, the rescue of *roX1 roX2* males by $roX1^{R2}$ is highest and by $roX1^{R1}$ is lowest. The trend is the same for all integration sites. Unexpectedly, $roX1^{R1}$ showed much lower or no

rescue of *roX1 roX2* males transcribed from the X chromosome (Table 3-5**)**. This is not due to differences in expression, and is not because *roX1*^{R1} is toxic to nearby genes. Lack of effective recruitment of the MSL complex, and low support of dosage compensation is more likely responsible for low male rescue by X-linked *roX1*^{R1}. A model based on a requirement for MLE to reposition chromatin-bound MSL complex is proposed (Figure 3-13). In this model, we hypothesize that the MSL complex containing *roX1*^{R1} but no MLE can bind chromatin. However, the MSL complex needs the MLE helicase for release from chromatin and repositioning at other sites (i.e., within body of genes). Without MLE, the *roX1*^{R1}-containing MSL complex is unable to move from initial binding site, resulting in excessive accumulation at the site of transcription. When *roX1*^{R1} is transcribed from an autosomal site, because the surrounding autosome is less permissive for MSL attachment, only very closely linked sites accumulate for the MSL complex. The nucleoplasm is filled with MSL complexes that ultimately find sites on the X chromosome. However, when *roX1*^{R1} is transcribed from the X chromosome, the MSL complex immediately binds nearby X chromatin. However, subsequent redistribution to more transcribed genes on the X chromosome can not occur. This raises the question of why the similar *roX1*^{ex7B} RNA, transcribed from the X chromosome under the control of the weaker *roX1* promoter, produces reasonably good (40-50%) male survival. One possibility is that the abundant supply of *roX1*^{R1} enables most of the MSL proteins to bind chromatin, depleting the pool of free MSL proteins in the nucleoplasm. A newly transcribed gene will consequently find it extraordinarily difficult to recruit MSL complex. The

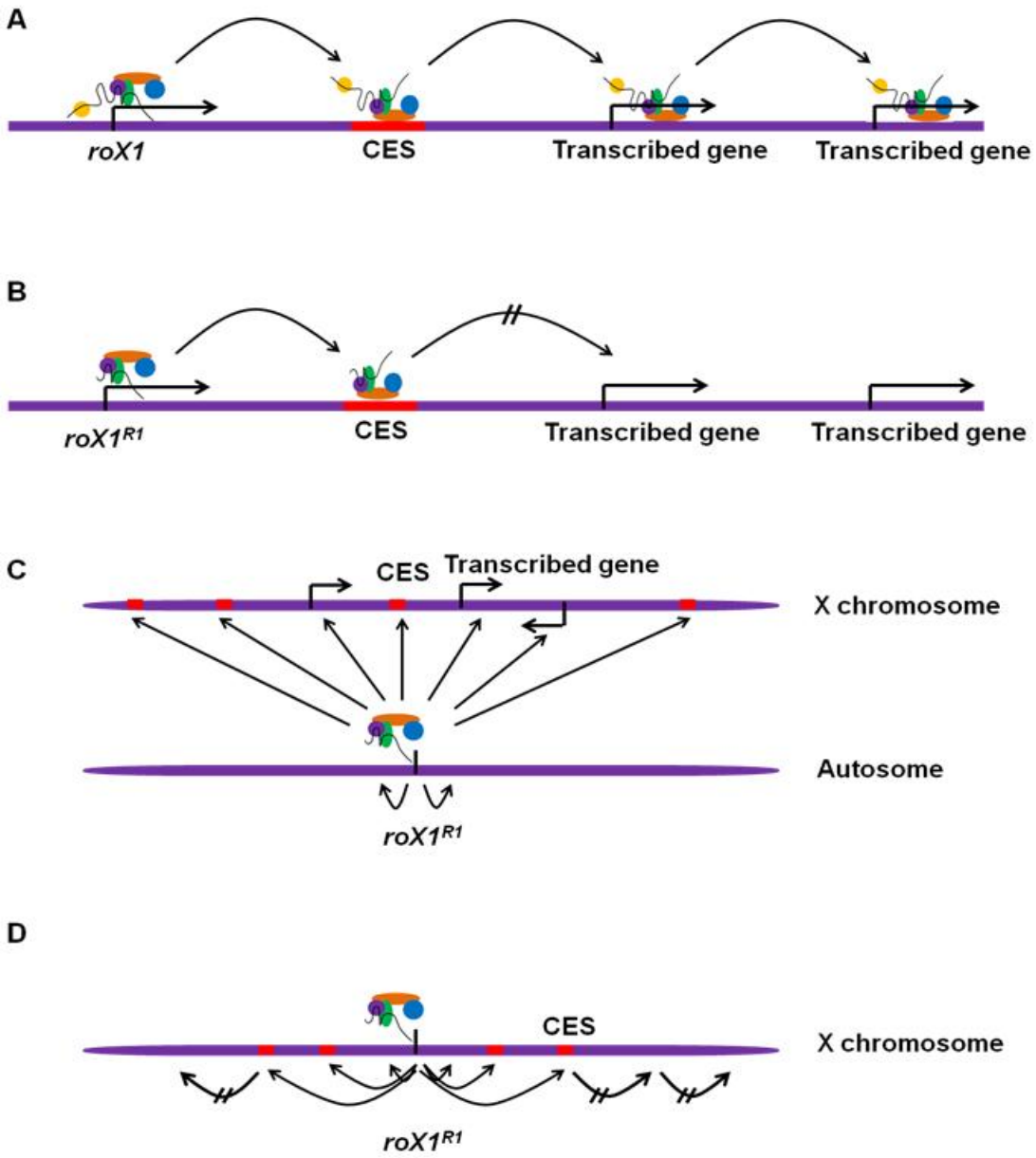

Figure 3-13. Mobility of the MSL complex depends on *roX1* and MLE. (A) After assembly of the MSL complex, chromatin entry sites are bound. The complex then spreads to transcribed genes. **(B)** MSL complex containing *roX1*^{R1} can bind X chromatin, but can not be repositioned as it lacks MLE activity. Spreading to nearby transcribed genes is inhibited. **(C)** When *roX1*^{R1} is transcribed from an autosomal site, closely linked sites accumulate the MSL complex. However, autosomal binding is inefficient and the nucleoplasm is filled with MSL complexes that ultimately finds a wide range of binding sites on the X chromosome. **(D)** When *roX1*^{R1} is transcribed from the X chromosome, the MSL complex is titrated out of the nucleoplasm by binding to nearby chromatin entry sites and transcribed genes. Bound complex is unable to redistribute to transcribed genes on the X chromosome.

fact that forced expression of MSL1 partially rescues *roX1 roX2* males carrying the *roX1*^{R1} transgene is consistent with this interpretation. Future studies testing these models may lead to a deeper understanding of how *roX* RNAs contribute to X identification, and the molecular mechanisms that underlie spreading *in cis* from a site of *roX* transcription.

In a summary, regions of *roX1* RNA with distinct functions have been identified throughout the 5' end of *roX1* by these studies (Figure 3-14). Multiple redundant elements that act co-operatively to contribute to X chromosome targeting of the MSL complex are distributed throughout the 5' end of the *roX1* transcript. The extreme 5' portion of *roX1* that is present in *roX1*^{R1} displays a unique activity, promoting MSL complex spreading from the site of transcription. However, only the sequences present in *roX1*^{R2} and *roX1*^{R3} are able to recruit MLE.

Naturally occurring *roX1* transcripts reveal minor and major transcription start sites that are developmentally regulated. The 5' ends of *roX1*$^{R1-3}$ transgenes roughly correlate with the transcription start sites of the three dominant *roX1* isoforms (Figure 3-2). The *roX1*^{R1} transgene contains a region only expressed in a minor population of *roX1* transcripts. This suggests that utilization of different transcription start sites is carefully regulated, and may contribute to normal distribution of the MSL complex. The activities I have detected are expected to be regulated by choice of transcription start site, alternative splicing and/or transcript stability.

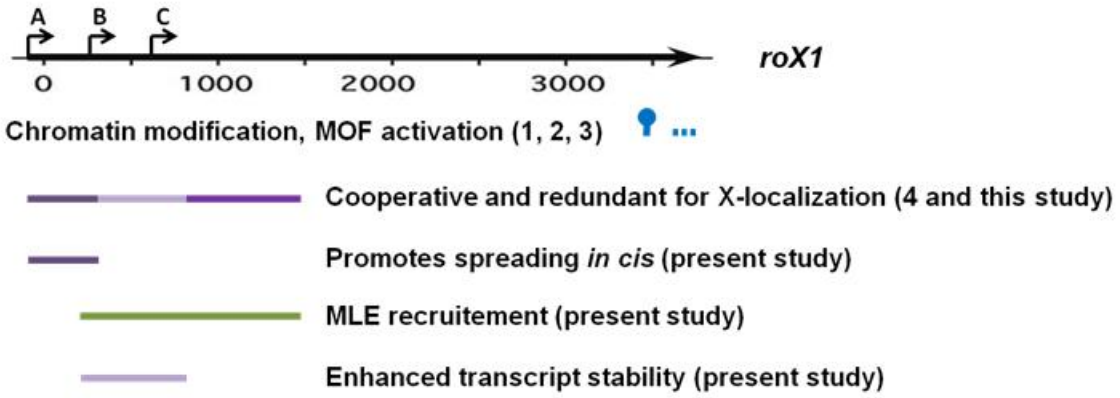

Figure 3-14. Schematic illustration of functional *roX1* domains. The top line indicates the *roX1* genomic region, numbering from (Amrein and Axel, 1997). The transcribed *roX1* region is shown as an arrow. The arrows above are: A. The start site of a minor class of *roX1* ESTs (expressed sequence tags) B. A more frequently used *roX1* transcription start site. C. The transcription start site of the majority *roX1* transcripts. Below are the *roX1* functional domains with respect to the transcribed region of *roX1*. Data is compiled from the following references. 1) Stuckenholz et al., 2003. 2) Park et al., 2008. 3) Kelley et al., 2008. 4) Deng et al., 2005.

CHAPTER 4

FUTURE DIRECTIONS

The results described in this study provide insight into chromosome-wide regulation by a large non-coding RNA. However, *roX1* function has not yet been completely elucidated. Addressing the following questions will enable a better understanding of how *roX1* carries out its function, and how this process is regulated.

What mechanism restricts *roX* regulation of heterochromatin to males?

roX RNAs participate in two distinct biological processes in *Drosophila melanogaster* males: dosage compensation and normal heterochromatin function. Dosage compensation increases expression from the male X chromosome to compensate for X chromosome monosomy. Dosage compensation is limited in males by repression of MSL2 in females (Kelley et al., 1997). However, heterochromatin is present in both sexes. All known components of the *roX*-dependent heterochromatic modulation system (*roX* RNA, MLE, MSL1, MSL3 and MOF) are present in early female embryos (Deng et al., 2009; Meller, 2003; Rastelli et al., 1995). We speculate that these molecules associate to form a complex that serves a transient role in formation of heterochromatin in male embryos. What prevents this from occuring in females is unclear. SXL binding sites in *roX1* provide a plausible point for regulation of this system by the female-

limited SXL protein (Meller et al., 1997). However, expression of *roX1* lacking these sites does not influence the expression of heterochromatic genes in females. Other possibilities exist. There could be additional factors that participate in *roX*-dependent heterochromatin modulation that are only present in males, or active in males. It is also possible that additional factors block this system in females. To investigate this possibility, *roX1* pull-down could be performed in male and female embryos, followed by a LC-MS/MS (Liquid Chromatography coupled with Tandem Mass Spectrometry) to identify *roX*-interacting factors. Different proteins associated with *roX1* in males and females could be the factors that limit this system to males. However, this interaction could be transient and weak, making it difficult to identify through a general *roX1* pull-down. Specific developmental stages and pull-down conditions are critical for this experiment.

What *roX1* regions are required for the regulation of heterochromatin?

It is possible that *roX* RNA interacts with different proteins when modulating heterochromatin and the X chromosome. These interactions could occur through distinct functional domains of *roX1*. It is of considerable interest to identify *roX1* sequences that participate in each of the known functions of this molecule. Multiple regions of *roX1*, with unique functions, are important for dosage compensation. One candidate *roX1* region suggested by our preliminary data to be required for heterochromatin regulation was found to be dispensable for this process. However, a systematic study of *roX1* mutations or transgenes

deleted for different regions could be used to identify regions necessary for heterochromatin. The *roX1* transgenes I created to test in dosage compensation can now be used for this purpose (Chapter 3). These *roX1* transgenes could be analyzed to determine their ability to restore the expression of heterochromatic genes in *roX1 roX2* males.

Is there a direct interaction between 5' *roX1* sequences and MSL proteins?

I have suggested that the function of intact *roX1* requires simultaneous interaction with different molecules. It is easy to imagine that *roX1* contacts multiply MSL proteins. Identification of sequences binding to each protein will help to elucidate how *roX1* exerts its function. My results revealed multiple redundant elements throughout the 5' region of the *roX1* transcript that cooperate in X chromosome targeting. Because $roX1^{R1\text{-}3}$ each contains a subset of these elements, they have partial function in directing MSL binding and supporting male survival. We can perform an *in vitro* transcription of the 5' portion of these transgenes and test if there is any physical interaction with any MSL proteins using gel mobility shift assay. If successful, we can further narrow down the functional elements by producing smaller RNA within the interacting region and test for interaction.

What prevents $roX1^{R1}$ from rescuing *roX1 roX2* males from X-linked insertion sites?

roX RNAs are normally located on and transcribed from the X

chromosome, and assemble with the MSL proteins to coat the X chromosome (Amrein and Axel, 1997; Meller et al., 1997). It has been suggested that X-linkage of the *roX* genes contributes to targeting of the MSL complex to the X chromosome (Kelley et al., 1999). We expected that a *roX1* transgene able to direct MSL complex localization to flanking chromatin would have higher rescue activity when inserted on the X chromosome than when inserted on an autosome. In contrast to our expectation, my *roX1* transgenes did not display higher activity when expressed from the X chromosome. Unexpectedly, one transgene, *roX1*^{R1}, showed no rescue of *roX1 roX2* males when transcribed from 16C1 or 19C4 on the X chromosome. While 19C4 is in proximal heterochromatin, perhaps lowering the effect of this site, *roX1*^{R1} is expressed at the same level when inserted at 16C1 and at 68A4 on the 3rd chromosome. My studies suggest that this insertion of *roX1*^{R1} fails to rescue males due to a deficiency of dosage compensation, rather than toxicity due to an elevated recruitment of MSL proteins to a particular site on the X chromosome. When integrated at another X-linked site, 2A3, *roX1*^{R2} and *roX1*^{R3} showed higher male rescue than the autosomal site, but *roX1*^{R1} still displayed very low rescue of *roX1 roX2* males. Future studies to understand why the transcription of *roX1*^{R1} from an autosomal site will partially rescue *roX1 roX2* males, but transcription from the X-linked sites fails to rescue these males may provide insight into the interplay between the *roX1* promoter, 5' regions of the *roX1* transcript and spreading in *cis*.

Do different *roX1* isoforms perform different functions?

roX1 displays minor and major transcription start sites that are developmentally regulated. I generated a full length *roX1* transgene that fully rescues X-lined gene expression, but only partially restores expression from fourth-linked heterochromatic genes. This incomplete rescue could be due to other aspects of the intrinsic *roX1* gene, such as correct balance of transcripts from alternative start sites or developmental regulation of transcription start sites. Different *roX1* activities may be regulated by choice of transcription start site, alternative splicing and/or transcript stability. A previous study suggested that variable alternative *roX2* splice forms are important for modulating the activity of the MSL complex (Park et al., 2005). My study and others all suggest that alternative splicing occurs in the 5' *roX1* region. A more complete analysis of alternative splicing and transcription start site utilization might provide clues to the different roles of *roX1* transcripts. My studies indicate that the $roX1^{R1}$ transgene has the ability to direct spreading *in cis* from the site of $roX1^{R1}$ transcription. The fact that the 5' portion of *roX1* included in $roX1^{R1}$ is present in very few endogenous *roX1* transcripts suggests that regulation of transcription start site may regulate this aspect of *roX1* function.

APPENDIX A

MINIMUM INFORMATION ABOUT MICROARRAY ANALYSIS OF *ROX* IN FEMALES

MIAME

This document contains minimum information for the microarray experiment of *roX1 roX2* mutant females in chapter 2.

1. Experimental design:

a) The goal of the experiment

Determine if *roX* RNAs regulate expression in *Drosophila* females.

b) Author (submitter), laboratory, contact information, links (URL), citation

Y. Kong

V. Meller

Department of Biological Sciences

5047 Gullen Mall

Wayne State University, Detroit, MI 48202

c) Type of the experiment

effect of *roX1 roX2* mutant chromosome on gene expression in females

d) Experimental factors.

mutation of both *roX1* and *roX2* genes

e) Single or multiple hybridizations

Single hybridizations:

***** Serial (yes/no)

no

***** Grouping (yes/no)

yes

Samples:

Group 1: *roX1*^{SMC17A}Df(1)52*;* [*w*$^{+}$4Δ4.3] (*roX1 roX2* mutant female larvae)

Group 2: *roX1*^{SMC17A}Df(1)52*;* [*w*$^{+}$4Δ4.3]; [hsp83-*roX1*$^{+}$] (control female larvae)

Arrays: 1, 2, 3, 4, 5, 6

Hybridizations:

Group 1,

H1: Array 1, VM_roX-1 (mutant)

H2: Array 2, VM_roX-2 (mutant)

H3: Array 3, VM_roX-3 (mutant)

H1, H2 and H3 are biological triplicates.

Group 2,

H4: Array 4, VM_roX+4 (control)

H5: Array 5, VM_roX+5 (control)

H6: Array 6, VM_roX+6 (control)

H4, H5 and H6 are biological triplicates.

f) Quality related indicators quality control steps taken:

***** Biological triplicates for each genotype

***** Other

RNA quality is analyzed on the Agilent 2100 Bioanalyzer.

Affymetrix positive and negative controls, including:

Hybridization controls: bioB, bioC, bioD from E. coli and cre from P1 bacteriophage

Poly-A controls: dap, lys, phe, thr, trp from B. subtilis

Drosophila maintenance genes: Actin (Actin 42A), GAPDH (Glyceraldehyde 3 phosphate dehydrogenase 2), Eif-4a (Eukaryotic initiation factor 4a)

Affymetrix PM-MM (Perfect match- Mismatch) comparison for unspecific binding.

g) A free text description of the experiment set or a link to a publication

The global effect of *roX* mutations on gene expression in females was measured by microarray analysis. Three independent RNA preparations for each genotype served as templates for probe synthesis. These probes were hybridized to Affymetrix *Drosophila* Genome 2.0 chips. Affymetrix Gene expression data was background corrected, normalized and summarized into a one expression value per sample and probeset using the RMA (robust multi-array average) algorithm. Changes in gene expression were determined by comparing the mean signal intensities of genes on arrays hybridized with *roX1 roX2* (mutant) probes to those hybridized with *roX1 roX2*; [hsp83-*roX1*$^{+}$] (control) probes.

2. Array design: each array used and each element (spot) on the array

2.1 Array copy: each array used and each element (spot) on the array.

***** Unique id as used in part 1

1, 2, 3, 4, 5, 6

***** Array design name

Affymetrix *Drosophila* Genome 2.0 chips

2.2 Array related information

Affymetrix eukaryotic gene expression analysis arrays

See http://www.affymetrix.com/products/arrays/specific/fly_2.affx

All transcript alignments, Gene Names and Probe set IDs were extracted

from the Affymetrix *Drosophila* 2 annotation file.

3. Samples: samples used, extract preparation and labeling

a) Bio-source properties

***** Samples

Three independent RNA preparations from *roX1 roX2* mutant female larvae and three from control female larvae (*roX1 roX2*; [hsp83-*roX1*$^{+}$]) were used. This design was adopted to minimize the genetic variation between mutant and control.

***** Organism (NCBI taxonomy)

Drosophila melanogaster

***** Gender

Female

***** Development stage

Third instar larvae

***** Genetic variation (e.g., gene knockout, transgenic variation)

Transgenic variation

b) Hybridization extract preparation laboratory protocol for extract preparation,

including: protocol description.

* Description:

Total RNA was prepared from groups of 50 third instar larvae by TRIzol (Invitrogen) extraction and purified using the RNeasy kit (Qiagen).

* Extraction method

TRIzol standard protocol (Invitrogen), purification by RNeasy kit (Qiagen)

* Whether total RNA, mRNA, or genomic DNA is extracted

total RNA

* Amplification (RNA polymerases, PCR)

none

c) Labeling: laboratory protocol for labeling, including:

* Protocol

Standard protocols as described in the Affymetrix GeneChip® Expression Analysis Technical Manual, P/N 702232, rev. 2, 2005-2006, Chapters 2-4.

cDNA synthesis was using the GeneChip® 3'-Amplification One-Cycle cDNA Synthesis Kit (Affymetrix, #900431).

100% of cDNA product is used to synthesize biotin labeled cRNA by the GeneChip® IVT Labeling Kit (Affymetrix, #900449).

* Amount of nucleic acids labeled

5 μg total RNA was used for each samples

* Label used (e.g., Cy3, Cy5, 33P)

Biotin

4. Hybridizations: procedures and parameters

* Laboratory protocol for hybridization

Standard protocols as described in the Affymetrix GeneChip® Expression Analysis Technical Manual, P/N 702232, rev. 2, 2005-2006, Chapter 3.

* Wash procedure

Wash protocol FS450_0002.

* Quantity of labeled target used

6.5 µg labeled cRNA

* Time, concentration, volume, temperature

Overnight (16h), 200 µl, at 45°C, rotating at 60 rpm

* Description of the hybridization instruments

GeneChip® Hybridization Oven 640 (Affymetrix, #800138)

5. Measurements: images, quantitation, specifications:

a) Hybridization scan raw data:

a1) the scanner image file (e.g., TIFF) from the hybridized microarray scanning

a2) Scanning information:

Arrays were scanned by the Affymetrix GeneChip® Scanner 3000 with Auto Loader (Affymetrix, #00-0090) expression analysis settings at "All Probe Sets" Scale Factor 500, Normalization = 1.0, No Masking.

b) Image analysis and quantitation

The Chip Image was analyzed using GCOS 1.4 to produce a .CEL file of probe intensities for each sample. The .CEL files were in turn modeled in PM-MM mode to produce a spreadsheet of probeset intensities for all six arrays.

The .CEL files and .CHP files were available in Gene Expression Omnibus database (Access number: GSE12076).

Mutant_1: GSM304915

Mutant_2: GSM304916

Mutant_3: GSM304917

Control_1: GSM304918

Control_2: GSM304919

Control_3: GSM304920

c) Annotation and alignment

All transcript alignments, Gene Names and Probeset IDs were extracted from the Bioconductor annotation package "Affymetrix *Drosophila* Genome 2.0 Array Annotation Data (drosophila2)". Transcripts were aligned to each chromosome.

6. Normalization

Affymetrix Gene expression data was background corrected, normalized and summarized into a one expression value per sample and probeset using the RMA (robust multi-array average) algorithm. The Affymetrix MAS5.0 Present/Absent calls were used to filter out the probesets that were not present in at least 2 out of 3 samples in each group (mutant and normal). All the probesets that could not be mapped to a known Entrez Gene ID, according to the Bioconductor (www.bioconductor.org) drosophila 2 package, were removed from further analysis. The expression values from multiple probes corresponding to the same Entrez Gene ID were averaged together in each array (sample). The chromosome assignment of each probeset was based on the drosophila 2 package in bioconductor. The probes corresponding to genes situated in heterochromatin were identified from the Fly database, assuming that all the remaining ones are euchromatic genes. The $\log_2$ fold changes were computed as the mean RMA expression values in the mutant sample minus the mean expression

values in the control samples. The resulting $\log_2$ fold changes were then compared among chromosomes using a Wilcoxon test. All the analyses were performed using the R software environment (www.r-project.org) and specialized bioconductor (www.bioconductor.org) software packages.

APPENDIX B

A *ROX1* TRANSGENE INSERTION REVEALS UNUSUAL FEMALE-SPECIFIC LETHALITY

ABSTRACT

Integration of *roX1*$^{R1-3}$ at a landing site within the X-linked *Tao-1* gene (18D3) was lethal. To obtain additional X-linked insertions of *roX1*^{R1}, the lethal *roX1*^{R1} integrant, as well as *roX1*^{R3} (as a control), was mobilized and surviving males carrying an X-linked *w*$^{+}$ marker were identified. Unexpectedly, most of these products of mobilization had strikingly reduced viability of homozygous females. Genetic and molecular analysis revealed that partial inserted sequences remaining in the original integration site cause the female-specific lethality. X chromosomes retaining different portions of the original inserted sequences complement each other and rescue female survival. The insertion of different sequences within this site may disturb *Tao-1* transcript splicing. I postulated that females are more sensitive to disruption in *Tao-1* levels or splicing.

To investigate the function of 5' *roX1* sequence, three transgenes, *roX1*^{R1}, *roX1*^{R2} and *roX1*^{R3}, were created in the pUASTB vector and inserted at 68A4 on the third chromosome and at four sites on X chromosome (2A3, 16C1, 18D3 and 19C4) using the site-specific ΦC31 integrase system. In this system, recombination occurs between an *attB* site on the plasmid (pUASTB) containing

the transgene and an *attP* site in the fly genome (Groth et al., 2004; Thorpe and Smith, 1998). The transgenes were integrated into *attP* target sites, present in the P[y^+Cary*attP*] P-element, at different sites in the fly genome. Integration of the pUASTB vector into P[y^+Cary*attP*] created a complicated P element with two 3' ends and two 5' ends (Fig A-2). These are arranged in a manner that will enable multiple modes of excision by transposase. While flies with transgenes integrated at other sites show normal viability, integration of *rox1*$^{R1-3}$ transgenes at 18D3 is lethal in males and homozygous lethal in females. The 18D3 landing site was created by mobilization of P[y^+Cary*attP*] from 64A4 (obtained from Bloomington *Drosophila* Stock Center) and selecting for insertions on X chromosome. The exact location of 18D3 X-linked insertion was determined by inverse PCR and sequencing (Inverse PCR Protocol available from Berkeley Drosophila Genome Project; http://www.fruitfly.org/about/methods/inverse.pcr.html). The landing site is located inside an intron of *Tao-1*, and, by itself, appears to be fully viable (Figure A-1). Information from Flybase (www.flybase.org) shows that this gene has 7 annotated transcripts, and 7 annotated polypeptides. The putative product of *Tao-1* is a serine/threonine kinase proposed to be involved in apoptosis. Although 8 alleles are reported, there is no null allele available and no phenotypic data available. There is also a MSL high affinity site located inside the *Tao-1* gene, which spans the insertion site (Figure A-1, (Straub et al., 2008)). It is quite possible that *Tao-1* is an essential gene and the disruption of *Tao-1* function by insertion of transgenes causes lethality. However, insertion of P[y^+Cary*attP*] P-element at this site is viable, but intergration of a pUASTB

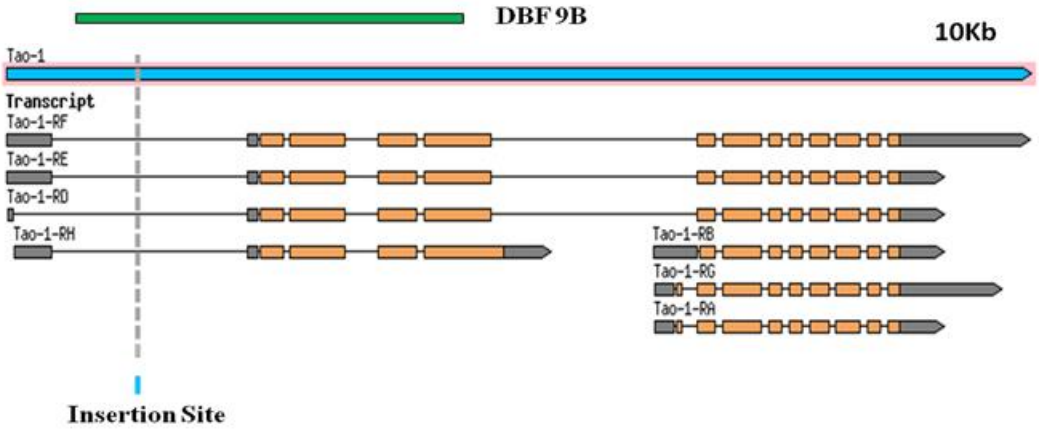

Figure A-1. Annotation of the 18D3 insertion site and the MSL High Affinity Site (HAS) located at this region. The blue arrow represents the *Tao-1* gene. Below are seven annotated *Tao-1* transcripts. The exact insertion site of the attP docking site is shown in relation to the *Tao-1* gene and the MSL HAS (green line). The annotation of *Tao-1* gene is from FlyBase (www.flybase.org).

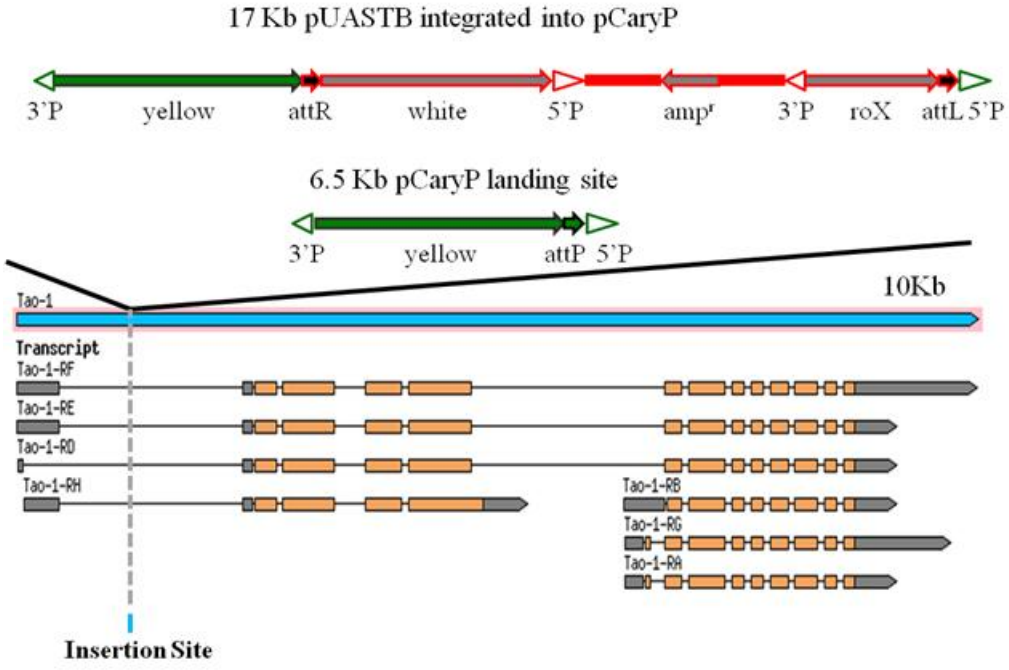

Figure A-2. Structure of viable and lethal integrations at 18D3. The structure of the viable insertion of pCaryP with attP landing site at 18D3 is shown in green. Integrated pUASTB plasmids, carrying *roX1* transgenes, are depicted in red. Arrow direction represents the direction of gene transcription and the orientation of att sequences. P-element ends are depicted by triangles. The sequence of pCaryP and pUASTB are from (Groth et al., 2004).

plasmid carrying *roX1* transgenes into the attP site is lethal (Figure A-2). The major difference of viable and lethal insertions at this site is the length of inserted sequences. Insertion of 6.5 Kb P[y^{+}Cary*attP*] P-element is viable, but pUASTB integration increases the size to 17Kb. Longer insertion size, strong transcription from the *hsp83* promoter, or the recruitment of MSL proteins by *roX1* transcription may disrupt *Tao-1* expression. Alternatively, the strong expression of *roX1* transgene may interact with the MSL high affinity site and disrupt dosage compensation of this region, which could be the reason for the lethality. However, the latter mechanism is not favored because I found no genetic interactions between the lethal insertion and MSL protein level.

My previous studies (Chapter 3) show that $roX1^{R1}$ could not rescue *roX1 roX2* males when transcribed from 16C1, but $roX1^{R2}$ and $roX1^{R3}$ could. To investigate whether $roX1^{R1}$ is toxic or non-functional when it comes from the X, $roX1^{R1}$ was mobilized from 18D3 to obtain viable insertions on the X and autosomes. This was accomplished by introducing a transposase source (TMS, sb P[$ry^{+t7.2}\Delta$2-3]99B) and recovering males with the w^{+} marker. While no lethality is observed for any autosomal reinsertion of the w^{+} marker, reinsertions on a normal X chromosome show not only normal viability, but some reinsertions also show female-specific lethality. For a control, $roX1^{R3}$ was also mobilized, X and autosomal insertions were collected. Surprisingly, $roX1^{R3}$ showed the same result, which suggested that the lethality is not due to $roX1^{R1}$ toxicity.

The most confusing aspect of this result is the male viability and female-specific lethality of X chromosomes produced by mobilization. In total nine X-

linked lines were recovered (five for *roX1*^{R1}, four for *roX1*^{R3}) that are male viable but female homozygous lethal. Homozygous females can develop into third instar larva or pupae but die in the early pupae stage (data not shown). Female specific lethality is rare in *Drosophila*. One possible reason is that the insertion interrupts a gene required only in females. Another explanation is that the X-linked *roX1* transgenes interact with a subset of MSL proteins locally to disrupt the female X chromosome. To distinguish the two possibilities, female-specific lethal lines were mated to test complementation. If transheterozygous females live, each insertion disrupts a different female essential gene. If transheterozygous females die, recruitment of MSL proteins by the transgenes is plausible. My complementation result showed that the nine female-lethal chromosomes produced by mobilization form two complementation groups. Group one contains three lines (fl(1)18D3^{TW2B,R3}, fl(1)18D3^{TW13B,R3} and fl(1)18D3^{TW14B,R3}). Group two contains six lines (fl(1)18D3^{TW5C,R3}, fl(1)18D3^{TW27B,R3}, fl(1)18D3^{3,R3}, fl(1)18D3^{9,R3}, fl(1)18D3^{11,R3} and fl(1)18D3^{12,R3}). A transheterozygous female with one X chromosome from one group and one from the other group can survive. However, a female with both X chromosomes from the same group is lethal. This made it difficult to distinguish between mutation of a female-specific essential gene and MSL recruitment as the source of lethality. To test recruitment of the MSL proteins as the source of female lethality, I attempted to rescue females by manipulation of MLE. Neither reducing MLE by half nor eliminating MLE rescued homozygous females. Therefore, disruption of the female X chromosome by recruitment of MSL proteins does not explain lethality. Next, I explored the

possibility of interruption of genes required only in females. Inverse PCR and sequencing was performed to determine the exact insertion sites in all female-specific lethal lines. Unexpectedly, all lines retained P element sequences at the original site within the *Tao-1* intron. The gene causing female-specific lethality is consequently suggested to be *Tao-1* itself.

By molecular characterization of the remaining sequences in *Tao-1*, I discovered that P element mobilization can occur multiple ways. There are alternative combinations of 5' and 3' ends of P element that can excise the inserted sequences (Figure A-3). Single fly PCR was performed to determine the remaining sequences at this site in each line. As shown in figure A-3, all members of a complementation group have excised the same way, leaving identical residual transgene material in *Tao-1*. However, by some unexpected mechanism, the excision to create group two removes material between two 5' P element ends. I conclude that each of the two modes of excision that I characterized produce male-viable and female-lethal alleles of *Tao-1*. In spite of the female lethality of these two alleles, they complement one another for female survival. It is likely that significant further effort will be required to determine the molecular basis of female lethality and complementation.

Taken together, insertion of different sequences at the same site of *Tao-1* intron displays different behaviors including full survival, lethal for both sexes and female specific lethal. Insertion of the P[y^{+}Cary*attP*] P element containing the landing site at an intron of *Tao-1* is viable (Figure A-4 A). Integration of pUASTB containing *roX1* transgenes into this landing site is lethal (Figure A-4 B). Reduction

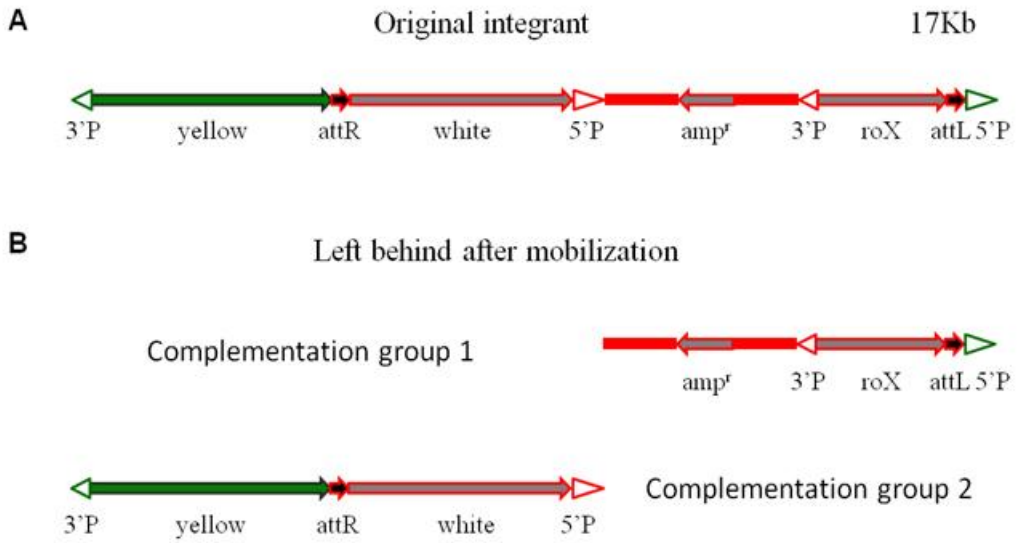

Figure A-3. Schematic illustration of insertions before and after mobilization by transposase. (A) The [y^+CaryattP] landing site at 18D3 after integration of *roX1* transgenes. Green depicts the original sequences of [y^+CaryattP], red is integrated pUASTB. **(B)** The remaining sequences at 18D3 after transposase mobilization from this site. In complementation group 1, mobilization of 3' (green) and 5' (red) P element ends removes both yellow and white. In complementation group 2, mobilization of 5' (red) and 5' (green) P element ends removes the pUAST vector backbone and the *roX1* transgene.

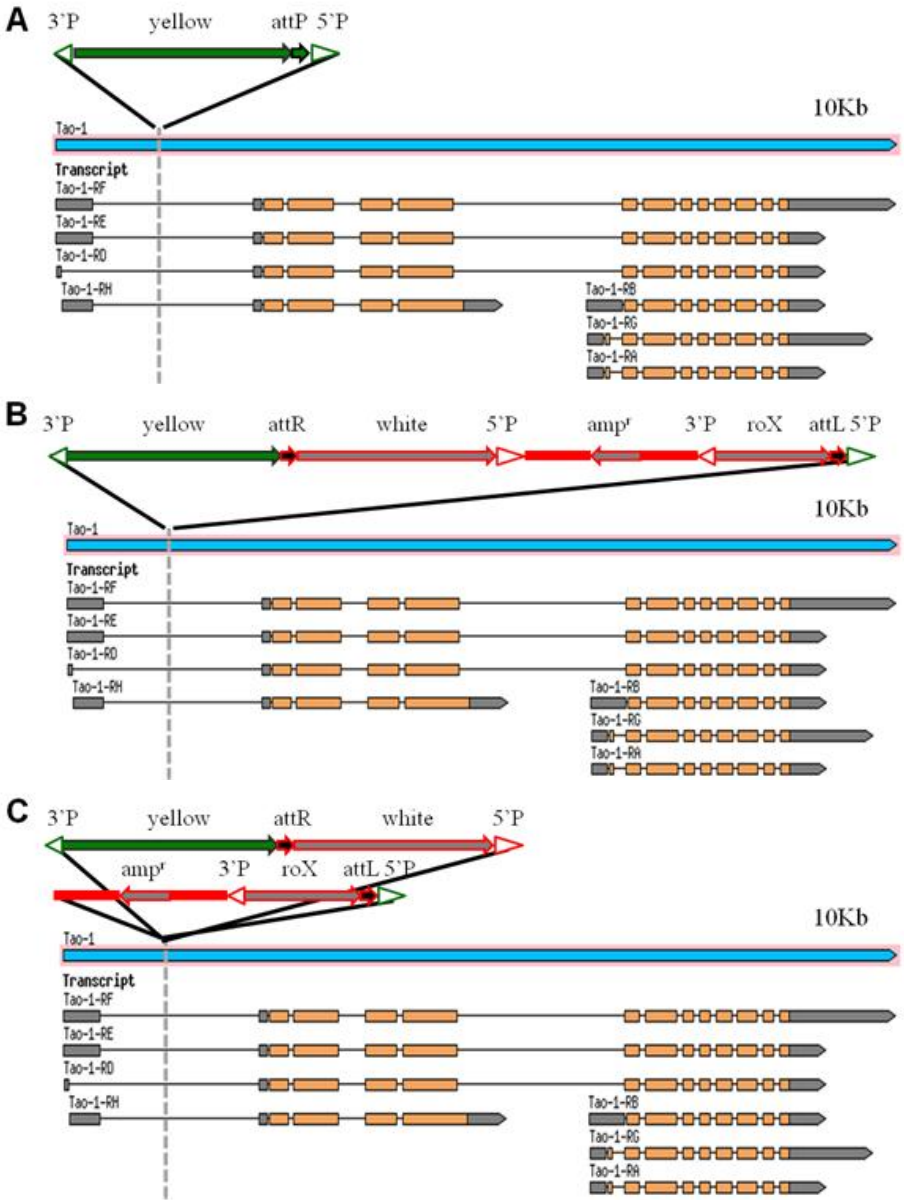

Figure A-4. Summary of insertions that show different viabilities. (A) Flies carrying only the landing site appear fully viable. (B) Flies with integration of any of the *roX1* transgenes are male and homozygous female lethal. (C) Flies with either partial deletion of the transgene integration are male viable but homozygous lethal in females. However, females are viable when they are transheterozygous, carrying both partial deletions.

of the inserted sequence by transposase-mediated mobilization produces two different male viable but female lethal alleles of *Tao-1* (Figure A-4 C). However, females that are trans-heterozygous for these two alleles are viable. The most plausible interpretation of my findings is that lethality is caused by disruption of *Tao-1* transcription or splicing. Females may be particularly sensitive to reduction of *Tao-1*. My studies suggest that this sex-specific lethality is not related to disruption of the X chromosome by *roX1* expression or inappropriate dosage compensation in females. To verify this conclusion, we ordered flies which carry duplications of the region containing the *Tao-1* gene on the third chromosome, Dp(1;3)DC363 (duplicated region is 18C8-18D3, BDSC stock# 31453) and Dp(1;3)DC364 (duplicated region is 18D1-18D13, BDSC stock# 32279), and mated these flies to my lethal transgenes and excisions. As expected, the lethal insertions of $roX1^{R1}$, $roX1^{R2}$ and $roX1^{R3}$ at 18D3 are rescued by introducing a copy of duplicated X region containing *Tao-1* (Table A-1). All female-specific lethal lines can also be rescued by the same duplications (Table A-2 and A-3).

In conclusion, our results suggest that *Tao-1* is an essential gene. Transcription or processing of *Tao-1* may occur differently in males and females. To future investigate *Tao-1* and verify my findings, a *Tao-1* deletion mutant could be created and tested for viability. The sex-specific lethality of my excision lines could be investigated by searching for alternative *Tao-1* transcripts in males and females by RT-PCR. Levels of *Tao-1* transcripts in homozygous lethal and trans-heterozygous viable females could also be determined by Quantitative RT-PCR to explore the mechanism of trans-heterozygous viability.

Table A-1. Rescue of 18D3 lethal insertions by duplications of 18D3.

X chromosomes of lethal insertion	Number of males that carry the lethal X chromosome			
	*DC363/+	+/+	#DC364/+	+/+
[*roX1*^{R1}]	13 (9)	0	22 (17)	0
[*roX1*^{R2}]	5 (12)	0	34 (34)	0
[*roX1*^{R3}]	14 (22)	0	22 (33)	0

[*roX1*$^{R1-3}$] inserted at 18D3 is fully lethal. [*roX1*$^{R1-3}$] / *yw* females were mated to *yw* males that carry one copy of Dp(1;3)DC363 or Dp(1;3)DC364 on the third chromosome. Only male offspring are recorded. The number of males carrying the lethal integration is shown; the number of brothers carrying the *yw* chromosome and the duplication is in parentheses.

* DC363 (Dp(1;3)DC363) carries a duplication of 18C8–18D3 on the third chromosome.

DC364 (Dp(1;3)DC364) carries a duplication of 18D1–18D13 on the third chromosome.

Table A-2. Rescue of female-specific lethal insertions by Dp(1;3) DC363.

Female-specific lethal lines	*fl(1)18D3 / fl(1)18D3		fl(1)18D3 / +	
	#DC363/+	+/+	DC363/+	+/+
$fl(1)18D3^{3,R3}$	18	0	11	20
$fl(1)18D3^{9,R3}$	22	0	21	22
$fl(1)18D3^{11,R3}$	13	0	16	13
$fl(1)18D3^{12,R3}$	13	0	15	20
$fl(1)18D3^{TW14B,R1}$	21	0	20	17
$fl(1)18D3^{TW27B,R1}$	19	0	17	14
$fl(1)18D3^{TW2B,R1}$	29	0	30	27
$fl(1)18D3^{TW5C,R1}$	25	0	18	22

Heterozygous females carrying a female-lethal X chromosome fl(1)18D3 were mated to males with the same X chromosome and one copy of Dp(1;3)DC363 on the third chromosome. The adults of the illustrated genotypes from each cross were counted.

* the X chromosome genotype, fl(1)18D3 chromosome is the *yw* chromosome with insertions of the leftover sequences at 18D3 after mobilization of P element from this site.

the third chromosome genotype, DC363 is short for Dp(1;3)DC363, which carries a duplication region of 18C8 – 18D3 on third chromosome. + means wildtype.

Table A-3. Rescue of female-specific lethal insertions by Dp(1;3) DC364.

Female-specific lethal lines	*18D3 X / 18D3 X		18D3 X / *yw*	
	#DC364/+	+/+	DC364/+	+/+
fl(1)18D3^{3,R3}	17	0	15	14
fl(1)18D3^{9,R3}	21	0	19	14
fl(1)18D3^{11,R3}	25	0	15	25
fl(1)18D3^{12,R3}	12	0	12	20
fl(1)18D3^{TW14B,R1}	20	0	20	22
fl(1)18D3^{TW27B,R1}	10	0	20	17
fl(1)18D3^{TW2B,R1}	16	0	19	15
fl(1)18D3^{TW5C,R1}	25	0	12	9

Heterozygous females carrying a female-lethal X chromosome fl(1)18D3 were mated to males with the same X chromosome and one copy of Dp(1;3)DC363 on the third chromosome. The adults of the illustrated genotypes from each cross were counted.

* the X chromosome genotype, fl(1)18D3 chromosome is the *yw* chromosome with insertions of the leftover sequences at 18D3 after mobilization of P element from this site.

the third chromosome genotype, DC364 is short for Dp(1;3)DC364, which carries a duplication region of 18C8 – 18D3 on third chromosome. + means wildtype.

APPENDIX C

JIL-1 ENRICHMENT ON THE X CHROMOSOME IS BLOCKED IN *ROX1 ROX2* MALES

Two types of histone modifications are enriched on the male X chromosome: H4K16 acetylation (Brownell and Allis, 1996) and histone H3S10 phosphorylation (Jin et al., 2000). The JIL-1 histone kinase is responsible for phosphorylation of H3 on serine 10 during interphase. JIL-1 binds to chromatin throughout the genome but is enriched on the male X chromosome (Jin et al., 2000). JIL-1 is also suggested to be involved in dosage compensation. Proper dosage compensation of the X-linked *white* gene requires normal JIL-1 function (Lerach et al., 2005). However, JIL-1 does not appear to be a component of the MSL complex (Mendjan et al., 2006; Smith et al., 2000). JIL-1 also plays a general role in limiting heterochromatin spreading, and is essential in both sexes (Ebert et al., 2004; Wang et al., 2001).

Simultaneous antibody staining of JIL-1 and MSL3 revealed that JIL-1 colocalized with MSL proteins on the male X chromosome (Figure A-5 A-C). The X-specific binding of MSL is disrupted in *roX1 roX2* males, which showed deceased binding on X chromosome but a lot of ectopic binding on autosomal sites (Figure A-5 E). JIL-1 enrichment on the X chromosome is also blocked (Figure A-5 D). However, colocalization of JIL-1 and MSL3 occurred even in the *roX1 roX2* double mutant (Figure A-5 D-F). The colocalization of JIL-1 and MSL3 occurs at both X-linked and ectopic autosomal sites that are only observed in

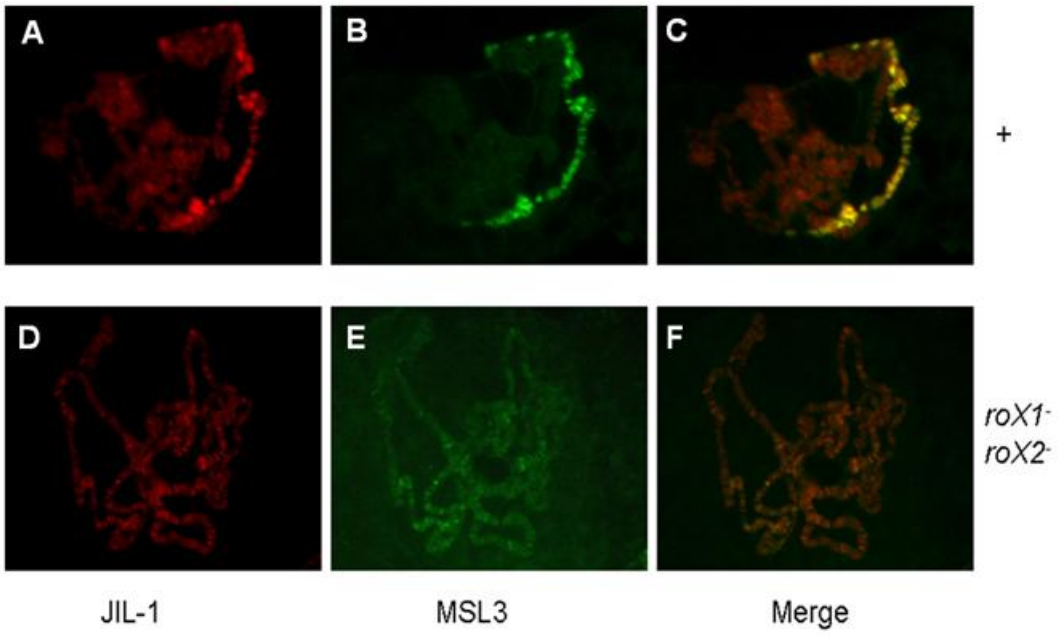

Figure A-5. JIL-1 is mislocalized in *roX1 roX2* males. Antibodies against JIL-1 and MSL3 were applied simultaneously to polytene chromosome preparations. JIL-1 is detected with FITC conjugated anti-rabbit, and MSL3 with TR conjugated anti-goat antibodies. JIL-1 colocalizes with the MSL proteins to many sites in *roX1 roX2* double mutants. The full genotype of *roX1 roX2* is $roX1^{SMC17A}$ *Df(1)52;* [4Δ4.3].

roX1 roX2 males (Figure A-5 F). This indicates that *roX* RNA influences JIL-1 localization. This might occur either directly, by interaction with the MSL complex, or indirectly, for example, by recruitment by the H4K16ac modification that is deposited by the MSL complex.

REFERENCES

Adams, R.R., H. Maiato, W.C. Earnshaw, and M. Carmena. 2001. Essential roles of Drosophila inner centromere protein (INCENP) and aurora B in histone H3 phosphorylation, metaphase chromosome alignment, kinetochore disjunction, and chromosome segregation. *J Cell Biol*. 153:865-880.

Akhtar, A., and P.B. Becker. 2000. Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in Drosophila. *Mol Cell*. 5:367-375.

Akhtar, A., D. Zink, and P.B. Becker. 2000. Chromodomains are protein-RNA interaction modules. *Nature*. 407:405-409.

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*. 20:848-857.

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*. 88:459-469.

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*.

23:2853-2861.

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*. 176:1491-1499.

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

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*. 89:789-798.

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 Dev*. 8:96-104.

Brownell, J.E., and C.D. Allis. 1996. Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. *Curr Opin Genet Dev*. 6:176-184.

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. *Mol Cell*. 11:1265-1277.

Buscaino, A., G. Legube, and A. Akhtar. 2006. X-chromosome targeting and dosage compensation are mediated by distinct domains in MSL-3. *EMBO Rep*. 7:531-538.

Cabal, G.G., A. Genovesio, S. Rodriguez-Navarro, C. Zimmer, O. Gadal, A.

Lesne, H. Buc, F. Feuerbach-Fournier, J.C. Olivo-Marin, E.C. Hurt, and U. Nehrbass. 2006. SAGA interacting factors confine sub-diffusion of transcribed genes to the nuclear envelope. *Nature*. 441:770-773.

Casolari, J.M., C.R. Brown, D.A. Drubin, O.J. Rando, and P.A. Silver. 2005. Developmentally induced changes in transcriptional program alter spatial organization across chromosomes. *Genes Dev*. 19:1188-1198.

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.

Charlesworth, B. 1991. The evolution of sex chromosomes. *Science*. 251:1030-1033.

Chaumeil, J., P. Le Baccon, A. Wutz, and E. Heard. 2006. A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. *Genes Dev*. 20:2223-2237.

Csankovszki, G., P. McDonel, and B.J. Meyer. 2004. Recruitment and spreading of the C. elegans dosage compensation complex along X chromosomes. *Science*. 303:1182-1185.

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.

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*.

112:103-115.

Deng, X., S.K. Koya, Y. Kong, and V.H. Meller. 2009. Coordinated regulation of heterochromatic genes in Drosophila melanogaster males. *Genetics*. 182:481-491.

Deng, X., and V.H. Meller. 2006. roX RNAs are required for increased expression of X-linked genes in Drosophila melanogaster males. *Genetics*. 174:1859-1866.

Deng, X., B.P. Rattner, S. Souter, and V.H. Meller. 2005. The severity of roX1 mutations is predicted by MSL localization on the X chromosome. *Mech Dev*. 122:1094-1105.

Dieci, G., and A. Sentenac. 2003. Detours and shortcuts to transcription reinitiation. *Trends Biochem Sci*. 28:202-209.

Ebert, A., G. Schotta, S. Lein, S. Kubicek, V. Krauss, T. Jenuwein, and G. Reuter. 2004. Su(var) genes regulate the balance between euchromatin and heterochromatin in Drosophila. *Genes Dev*. 18:2973-2983.

Eisen, A., R.T. Utley, A. Nourani, S. Allard, P. Schmidt, W.S. Lane, J.C. Lucchesi, and J. Cote. 2001. The yeast NuA4 and Drosophila MSL complexes contain homologous subunits important for transcription regulation. *J Biol Chem*. 276:3484-3491.

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

Fauth, T., F. Muller-Planitz, C. Konig, T. Straub, and P.B. Becker. 2010. The DNA

binding CXC domain of MSL2 is required for faithful targeting the Dosage Compensation Complex to the X chromosome. *Nucleic Acids Res*. 38:3209-3221.

Feuerbach, F., V. Galy, E. Trelles-Sticken, M. Fromont-Racine, A. Jacquier, E. Gilson, J.C. Olivo-Marin, H. Scherthan, and U. Nehrbass. 2002. Nuclear architecture and spatial positioning help establish transcriptional states of telomeres in yeast. *Nat Cell Biol*. 4:214-221.

Franke, A., and B.S. Baker. 1999. The rox1 and rox2 RNAs are essential components of the compensasome, which mediates dosage compensation in Drosophila. *Mol Cell*. 4:117-122.

Gautier, L., L. Cope, B.M. Bolstad, and R.A. Irizarry. 2004. affy--analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics*. 20:307-315.

Gelbart, M.E., and M.I. Kuroda. 2009. Drosophila dosage compensation: a complex voyage to the X chromosome. *Development*. 136:1399-1410.

Gelbart, M.E., E. Larschan, S. Peng, P.J. Park, and M.I. Kuroda. 2009. Drosophila MSL complex globally acetylates H4K16 on the male X chromosome for dosage compensation. *Nat Struct Mol Biol*. 16:825-832.

Giet, R., and D.M. Glover. 2001. Drosophila aurora B kinase is required for histone H3 phosphorylation and condensin recruitment during chromosome condensation and to organize the central spindle during cytokinesis. *J Cell Biol*. 152:669-682.

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*. 35:3561-3572.

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*. 20:858-870.

Grimaud, C., and P.B. Becker. 2009. The dosage compensation complex shapes the conformation of the X chromosome in Drosophila. *Genes Dev*. 23:2490-2495.

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*. 166:1775-1782.

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 J*. 19:5202-5211.

Gupta, A., G.G. Sharma, C.S. Young, M. Agarwal, E.R. Smith, T.T. Paull, J.C. Lucchesi, K.K. Khanna, T. Ludwig, and T.K. Pandita. 2005. Involvement of human MOF in ATM function. *Mol Cell Biol*. 25:5292-5305.

Gupta, R.A., N. Shah, K.C. Wang, J. Kim, H.M. Horlings, D.J. Wong, M.C. Tsai, T. Hung, P. Argani, J.L. Rinn, Y. Wang, P. Brzoska, B. Kong, R. Li, R.B. West, M.J. van de Vijver, S. Sukumar, and H.Y. Chang. 2010. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature*. 464:1071-1076.

Henikoff, S., and P.M. Meneely. 1993. Unwinding dosage compensation. *Cell*. 72:1-2.

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 J*. 16:2054-2060.

Hoskins, R.A., J.W. Carlson, C. Kennedy, D. Acevedo, M. Evans-Holm, E. Frise, K.H. Wan, S. Park, M. Mendez-Lago, F. Rossi, A. Villasante, P. Dimitri, G.H. Karpen, and S.E. Celniker. 2007. Sequence finishing and mapping of Drosophila melanogaster heterochromatin. *Science*. 316:1625-1628.

Irizarry, R.A., B. Hobbs, F. Collin, Y.D. Beazer-Barclay, K.J. Antonellis, U. Scherf, and T.P. Speed. 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*. 4:249-264.

Jaenisch, R., C. Beard, J. Lee, Y. Marahrens, and B. Panning. 1998. Mammalian X chromosome inactivation. *Novartis Found Symp*. 214:200-209; discussion 209-213, 228-232.

Jeppesen, P., and B.M. Turner. 1993. The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. *Cell*. 74:281-289.

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. *J Cell Biol*. 149:1005-1010.

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. *Mol 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 J*. 20:2236-2245.

Kelley, R.L. 2004. Path to equality strewn with roX. *Dev Biol*. 269:18-25.

Kelley, R.L., O.K. Lee, and Y.K. Shim. 2008. Transcription rate of noncoding roX1 RNA controls local spreading of the Drosophila MSL chromatin remodeling complex. *Mech Dev*. 125:1009-1019.

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*. 98:513-522.

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*. 81:867-877.

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*. 387:195-199.

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*. 133:813-828.

Kong, Y., and V. Meller. 2007. How to get extra performance from a chromosome: recognition and modification of the X chromosome in male Drosophila melanogaster. *In* The Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics. Vol. Part I Genetics. Section 1.3 Epigenetics. .

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*. 172:963-974.

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

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*. 28:121-133.

Larschan, E., E.P. Bishop, P.V. Kharchenko, L.J. Core, J.T. Lis, P.J. Park, and M.I. Kuroda. 2011. X chromosome dosage compensation via enhanced transcriptional elongation in Drosophila. *Nature*. 471:115-118.

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 J*. 16:2671-2681.

Lee, J.T., and R. Jaenisch. 1997. Long-range cis effects of ectopic X-inactivation centres on a mouse autosome. *Nature*. 386:275-279.

Lee, J.T., W.M. Strauss, J.A. Dausman, and R. Jaenisch. 1996. A 450 kb transgene displays properties of the mammalian X-inactivation center. *Cell*. 86:83-94.

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*. 20:871-883.

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*. 43:213-215.

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*. 25:8913-8924.

Lindsley, D.L., and G.G. Zimm. 1992. The Genome of Drosophila melanogaster. Academic Press, San Diego.

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*. 37:1361-1366.

Liu, T.Y., C.W. Lin, S. Falcon, J. Zhang, and J.W. MacDonald. 2007. drosophila2: Affymetrix Drosophila Genome 2.0 ArrayAnnotation Data (drosophila2). R

packager version 2.2.0.

Locke, J., L.T. Howard, N. Aippersbach, L. Podemski, and R.B. Hodgetts. 1999. The characterization of DINE-1, a short, interspersed repetitive element present on chromosome and in the centric heterochromatin of Drosophila melanogaster. *Chromosoma*. 108:356-366.

Lucchesi, J.C., W.G. Kelly, and B. Panning. 2005. Chromatin remodeling in dosage compensation. *Annu Rev Genet*. 39:615-651.

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*. 147:1743-1753.

Mahadevan, L.C., A.L. Clayton, C.A. Hazzalin, and S. Thomson. 2004. Phosphorylation and acetylation of histone H3 at inducible genes: two controversies revisited. *Novartis Found Symp*. 259:102-111; discussion 111-104, 163-109.

Marin, I., and B.S. Baker. 2000. Origin and evolution of the regulatory gene male-specific lethal-3. *Mol Biol Evol*. 17:1240-1250.

Mattick, J.S. 2003. Challenging the dogma: the hidden layer of non-protein-coding RNAs in complex organisms. *Bioessays*. 25:930-939.

Mattick, J.S., and I.V. Makunin. 2005. Small regulatory RNAs in mammals. *Hum Mol Genet*. 14 Spec No 1:R121-132.

McDonel, P., J. Jans, B.K. Peterson, and B.J. Meyer. 2006. Clustered DNA motifs mark X chromosomes for repression by a dosage compensation complex.

Nature. 444:614-618.

Meller, V.H. 2003. Initiation of dosage compensation in Drosophila embryos depends on expression of the roX RNAs. *Mech Dev*. 120:759-767.

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. *Curr Biol*. 10:136-143.

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*. 21:1084-1091.

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*. 88:445-457.

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*. 21:811-823.

Menon, D.U., and V.H. Meller. 2009. Imprinting of the Y chromosome influences dosage compensation in roX1 roX2 Drosophila melanogaster. *Genetics*. 183:811-820.

Mercer, T.R., M.E. Dinger, and J.S. Mattick. 2009. Long non-coding RNAs:

insights into functions. *Nat Rev Genet*. 10:155-159.

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*. 23:2258-2268.

Morra, R., E.R. Smith, R. Yokoyama, and J.C. Lucchesi. 2008. The MLE subunit of the Drosophila MSL complex uses its ATPase activity for dosage compensation and its helicase activity for targeting. *Mol Cell Biol*. 28:958-966.

Morra, R., R. Yokoyama, H. Ling, and J.C. Lucchesi. 2011. Role of the ATPase/helicase maleless (MLE) in the assembly, targeting, spreading and function of the male-specific lethal (MSL) complex of Drosophila. *Epigenetics & chromatin*. 4:6.

Nielsen, P.R., D. Nietlispach, A. Buscaino, R.J. Warner, A. Akhtar, A.G. Murzin, N.V. Murzina, and E.D. Laue. 2005. Structure of the chromo barrel domain from the MOF acetyltransferase. *J Biol Chem*. 280:32326-32331.

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*. 14:481-487.

Oh, H., Y. Park, and M.I. Kuroda. 2003. Local spreading of MSL complexes from roX genes on the Drosophila X chromosome. *Genes Dev*. 17:1334-1339.

Osborne, C.S., L. Chakalova, K.E. Brown, D. Carter, A. Horton, E. Debrand, B. Goyenechea, J.A. Mitchell, S. Lopes, W. Reik, and P. Fraser. 2004. Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat*

Genet. 36:1065-1071.

Park, S.W., M.I. Kuroda, and Y. Park. 2008. Regulation of histone H4 Lys16 acetylation by predicted alternative secondary structures in roX noncoding RNAs. *Mol Cell Biol*. 28:4952-4962.

Park, S.W., H. Oh, Y.R. Lin, and Y. Park. 2010. MSL cis-spreading from roX gene up-regulates the neighboring genes. *Biochemical and biophysical research communications*. 399:227-231.

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. *Mol Cell*. 11:977-986.

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*. 2:157-164.

Penalva, L.O., and L. Sanchez. 2003. RNA binding protein sex-lethal (Sxl) and control of Drosophila sex determination and dosage compensation. *Microbiology and molecular biology reviews : MMBR*. 67:343-359, table of contents.

Pfaffl, M.W. 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res*. 29:e45.

Plath, K., S. Mlynarczyk-Evans, D.A. Nusinow, and B. Panning. 2002. Xist RNA

and the mechanism of X chromosome inactivation. *Annu Rev Genet*. 36:233-278.

Plath, K., D. Talbot, K.M. Hamer, A.P. Otte, T.P. Yang, R. Jaenisch, and B. Panning. 2004. Developmentally regulated alterations in Polycomb repressive complex 1 proteins on the inactive X chromosome. *J Cell Biol*. 167:1025-1035.

Prasanth, K.V., and D.L. Spector. 2007. Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum. *Genes Dev*. 21:11-42.

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. *Mech Dev*. 53:223-233.

Rattner, B.P., and V.H. Meller. 2004. Drosophila male-specific lethal 2 protein controls sex-specific expression of the roX genes. *Genetics*. 166:1825-1832.

Rice, W.R. 1996. Evolution of the Y sex chromosome in animals. *Bioscience*. 46:331–343.

Richter, L., J.R. Bone, and M.I. Kuroda. 1996. RNA-dependent association of the Drosophila maleless protein with the male X chromosome. *Genes Cells*. 1:325-336.

Riddle, N.C., and S.C. Elgin. 2006. The dot chromosome of Drosophila: insights into chromatin states and their change over evolutionary time. *Chromosome Res*. 14:405-416.

Salz, H.K., and J.W. Erickson. 2010. Sex determination in Drosophila: The view

from the top. *Fly (Austin)*. 4:60-70.

Sanjuan, R., and I. Marin. 2001. Tracing the origin of the compensasome: evolutionary history of DEAH helicase and MYST acetyltransferase gene families. *Mol Biol Evol*. 18:330-343.

Schuettengruber, B., D. Chourrout, M. Vervoort, B. Leblanc, and G. Cavalli. 2007. Genome regulation by polycomb and trithorax proteins. *Cell*. 128:735-745.

Shogren-Knaak, M., H. Ishii, J.M. Sun, M.J. Pazin, J.R. Davie, and C.L. Peterson. 2006. Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science*. 311:844-847.

Shoresh, M., S. Orgad, O. Shmueli, R. Werczberger, D. Gelbaum, S. Abiri, and D. Segal. 1998. Overexpression Beadex mutations and loss-of-function heldup-a mutations in Drosophila affect the 3' regulatory and coding components, respectively, of the Dlmo gene. *Genetics*. 150:283-299.

Silva, J., W. Mak, I. Zvetkova, R. Appanah, T.B. Nesterova, Z. Webster, A.H. Peters, T. Jenuwein, A.P. Otte, and N. Brockdorff. 2003. Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. *Dev Cell*. 4:481-495.

Smith, C.D., S. Shu, C.J. Mungall, and G.H. Karpen. 2007. The Release 5.1 annotation of Drosophila melanogaster heterochromatin. *Science*. 316:1586-1591.

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. *J Biol Chem*. 276:31483-31486.

Smith, E.R., C. Cayrou, R. Huang, W.S. Lane, J. Cote, and J.C. Lucchesi. 2005. A human protein complex homologous to the Drosophila MSL complex is responsible for the majority of histone H4 acetylation at lysine 16. *Mol Cell Biol*. 25:9175-9188.

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. *Mol Cell Biol*. 20:312-318.

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.

Stuckenholz, C., V.H. Meller, and M.I. Kuroda. 2003. Functional redundancy within roX1, a noncoding RNA involved in dosage compensation in Drosophila melanogaster. *Genetics*. 164:1003-1014.

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*. 15:1318-1325.

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*. 441:774-778.

Taft, R.J., K.C. Pang, T.R. Mercer, M. Dinger, and J.S. Mattick. 2010. Non-coding RNAs: regulators of disease. *The Journal of pathology*. 220:126-139.

Taipale, M., S. Rea, K. Richter, A. Vilar, P. Lichter, A. Imhof, and A. Akhtar. 2005. hMOF histone acetyltransferase is required for histone H4 lysine 16 acetylation in mammalian cells. *Mol Cell Biol*. 25:6798-6810.

Talbert, P.B., and S. Henikoff. 2006. Spreading of silent chromatin: inaction at a distance. *Nat Rev Genet*. 7:793-803.

Thorpe, H.M., and M.C. Smith. 1998. In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. *Proc Natl Acad Sci U S A*. 95:5505-5510.

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*. 69:375-384.

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.

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

Wu, L., B.M. Zee, Y. Wang, B.A. Garcia, and Y. Dou. 2011. The RING finger protein MSL2 in the MOF complex is an E3 ubiquitin ligase for H2B K34 and is involved in crosstalk with H3 K4 and K79 methylation. *Mol Cell*. 43:132-144.

Wutz, A. 2003. RNAs templating chromatin structure for dosage compensation in animals. *Bioessays*. 25:434-442.

Wutz, A., and R. Jaenisch. 2000. A shift from reversible to irreversible X

inactivation is triggered during ES cell differentiation. *Mol Cell*. 5:695-705.

Wutz, A., T.P. Rasmussen, and R. Jaenisch. 2002. Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat Genet*. 30:167-174.

Yasuhara, J.C., and B.T. Wakimoto. 2006. Oxymoron no more: the expanding world of heterochromatic genes. *Trends Genet*. 22:330-338.

Zhang, L.F., K.D. Huynh, and J.T. Lee. 2007. Perinucleolar targeting of the inactive X during S phase: evidence for a role in the maintenance of silencing. *Cell*. 129:693-706.

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*. 133:229-235.

ABSTRACT

roX1 FUNCTION IN DOSAGE COMPENSATION: STRUCTURAL / FUNCTIONAL ANALYSIS OF A NON-CODING RNA

by

YING KONG

May 2012

Advisor: Dr. Victoria H. Meller

Major: Biological Sciences

Degree: Doctor of Philosophy

roX1 is a long non-coding RNA involved in the chromosome-wide gene regulation that occurs during dosage compensation in *Drosophila*. Dosage compensation in *Drosophila melanogaster* occurs by a global two-fold increase of transcription from the single male X chromosome. This essential process compensates for X chromosome monosomy. The male-specific lethal (MSL) complex, containing five proteins, localizes to the male X chromosome and alters chromatin to modify gene expression. *roX1* and *roX2* RNAs are redundant components of MSL complex that are required for its exclusive X-localization. Recent studies in our lab have revealed a second role of *roX* RNAs in heterochromatic gene expression in males. The *roX*-dependent heterochromatic regulation system involves some, but not all, MSL proteins. Although all components of this system discovered now are expressed in females, microarray analysis showed that the *roX1* RNA has no detectable affect on expression of both X-linked and heterochromatic genes in females. Therefore, like dosage compensation, the *roX*-dependent heterochromatic regulation system is also

limited to males. The differential regulation of heterochromatic genes in males and females may reflect the differences of heterochromatin between them. Previous studies of *roX1* mutants and transgenes have identified a large region at the 5' end of *roX1* that is necessary for X-localization of the MSL complex. To dissect the function of this region, *roX1* transgenes containing portions of the 5' end were generated and analyzed. Multiple redundant elements contributing to X chromosome targeting were found to be present throughout the 5' end of *roX1*. These *roX1* transgenes display different stability, but all can partially restore X-linked gene expression in a *roX1 roX2* mutant. One portion of this region is uniquely able to promote MSL complex spreading from sites of transcription. Previous model of MSL spreading suggested a verse relevance of *roX1* transcript abundance and the ability to spread. However, the ability of this region to direct MSL spreading is not relevant to its abundance. The activities I have detected are hypothesized to be regulated by choice of transcription start site, alternative splicing and/or transcript stability.

AUTOBIOGRAPHICAL STATEMENT

Ying Kong

Education

2004-2011 Doctor of Philosophy, Wayne State University, Detroit, MI, USA

1999-2003 Bachelor of Science, Nankai University, Tianjin, P. R. China

Honors and Awards

Enhancement Graduate Research Assistant Award, Wayne State University 2007-2008, 2009-2010

Thomas C. Rumble Graduate Fellowship, Wayne State University 2004-2005

Publications

Kong Y. and Meller V.H. (In preparation) Noncoding *roX1* RNA 5' sequences contribute to X chromosome localization and local spreading of the MSL complex in *Drosophila*.

Deng X., Koya S.K., **Kong Y.**, and Meller V.H. (2009) Coordinated regulation of heterochromatic genes in *Drosophila melanogaster* males. Genetics. 182(2): 481-491.

Kong Y. and Meller V.H. (2007) How to get extra performance from a chromosome: recognition and modification of the X chromosome in male *Drosophila melanogaster*. The Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics. Part I Genetics. Section 1.3 Epigenetics.