ANALYZING THE EFFECT ON *DROSOPHILA* SEX DETERMINATION BY TRANSCRIPTION FACTOR BINDING SITE MUTATIONS IN THE SXL ENHANCER

A Dissertation

by

YOUNG-HO JUNG

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	James W. Erickson
Committee Members,	L. Rene Garcia
	Paul E. Hardin
	Gregory T. Reeves
Head of Department,	Alex C. Keene

May 2022

Major Subject: Biology

ABSTRACT

Cell fate decisions in eukaryotic organisms can be altered in response to small dose changes of transcription factors. *Drosophila melanogaster* sex determination is such an example. *Sex-lethal (Sxl)* is the master regulatory gene of fly sex determination. The goal of my work is to understand the molecular mechanism of how *Sxl* reads and responds to the X chromosome signal. Early establishment *Sxl* promoter, *SxlPe* is activated responding to two dose XSEs, but a single XSE dose does not activate *SxlPe*. Previous efforts to understand sex specific expression of *SxlPe* relied on *Sxl* transgenes. Although we learned a lot from the transgene experiments, our knowledge from transgenes was limited because genetic background of the transgene is different from the endogenous *Sxl* environment. To overcome this limitation, I engineered endogenous *Sxl* mutant lines by CRISPR/Cas9. The new endogenous *Sxl* mutants allowed precise quantification of *SxlPe* expression without the genetic background issue.

Negative regulators such as zygotic *deadpan (dpn)* and maternal *groucho (gro)* are the critical element in fly sex determination by establishing X chromosome signal threshold. Analyzing the effect of repressor binding sites showed that all the repressor sites were important for sex specific expression of *SxlPe*. Mutant repressor sites caused ectopic expression of *SxlPe* in male embryos. I observed that the non-canonical repressor site, which was expected to be less efficient for repressor Dpn binding, induced strong ectopic *SxlPe* expression in male embryos. To provide full constitutive activity of *SxlPe*, I mutated all the three repressor sites. As expected, the mutant allele induced strong ectopic expression of *SxlPe*. Interestingly, the strong constitutive allele was perfectly countered by loss of *sisB*.

To assess the direct contribution of transcription activators in sex specific expression of *SxlPe*, new *Sxl* alleles with mutant activator binding sites were created. Genetic testing and analysis of the *SxlPe* expression pattern showed that all the transcription activator binding sites were important. Surprisingly, a non-canonical SisB/Da activator site had a predominating effect in *SxlPe* expression, suggesting that the activator site may interact with nearby activator sites.

Recently, I inserted epitope tag *Llama* to N-terminus of endogenous *Sxl*. The *Llama-Sxl* allele was ectopically expressed in male embryos which could be attributed to the presence of two SisB/Da activator sites in the tag. Removing these two activator sites eliminated ectopic *Sxl* expression in male embryo, suggesting that the current balance between transcription activator and repressor binding sites is an evolutionary prerequisite for sex specific expression of *SxlPe*.

My work showed that all transcription factor sites in the 400bp proximal enhancer are important for sex specific expression of *SxlPe*, but the contribution of each transcription factor binding site varies in context dependent manner. Future work will require identification of *sisA* and *runt* binding sites and characterization of all the transcription factor binding sites.

DEDICATION

To my parents and sister who always supported me

To love of my life, Ye-Lin for everything she has done for me.

ACKNOWLEDGEMENTS

I would like to thank Dr. Jim Erickson for his financial, mental, scientific supports and his guidance to develop me as a scientist. I also appreciate my previous and current committee members Dr. Rene Garcia, Dr. Paul Hardin, Dr. Gregory Reeves, Dr. Arne Lekven and Dr. Ji Jun-yuan for their encouragements and insightful advice to direct my research to be productive.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professor Jim Erickson, Rene Garcia and Paul Hardin of Department of Biology and Professor Greg Reeves of Department of Chemical Engineering.

The analysis depicted in FIG. 1.5 and FIG. 2.1 were provided by Jayashre Rajendren, (unpublished). All other work conducted for dissertation was completed by the student independently.

Funding sources

Graduate study was supported by research fellowship from NSF.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I INTRODUCTION

Sxl reads and distinguish one X and two X chromosome signal

Cell fate decisions in eukaryotic organisms can change in response to subtle dose change of transcription factors (Nusslein-Volhard & Wieschaus, 1980). Sex determination of the fruit fly *Drosophila melanogaster* is such an example. *Sex-lethal (Sxl)* is the master regulator gene of fly sex determination. A transient molecular signal from two X chromosomes activates *Sxl* but the signal from one X chromosome does not activate *Sxl (Cline, 1984, 1988; Erickson & Quintero, 2007; Salz & Erickson, 2010).*

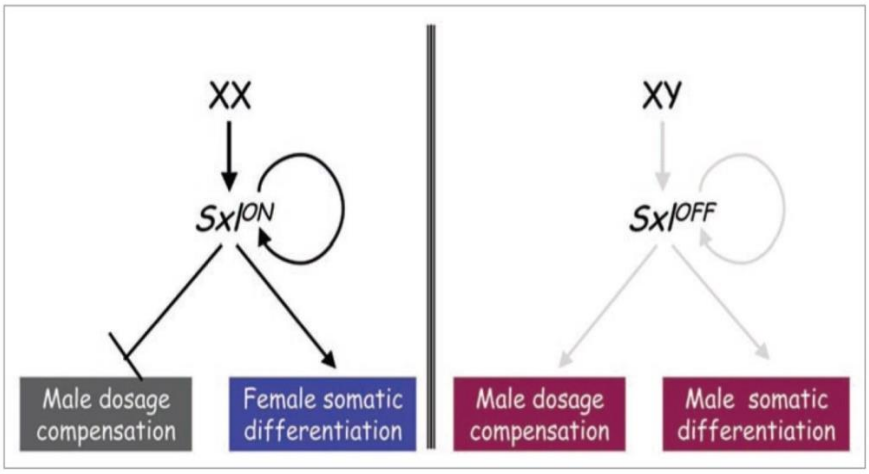

FIG. 1.1 : Sexually dimorphic switch *Sxl*

Sxl is expressed by two X chromosome signal in XX embryo. Once active, *Sxl* keeps itself on in positive auto regulatory loop. The consequence of *Sxl* activation in XX embryo is to activate female somatic differentiation and inhibit male dosage compensation. On the other hand, one X

chromosome signal does not activate *Sxl*. As a consequence, XY embryo achieves male somatic differentiation and activate male dosage compensation.

Sxl autoregulation in XX embryo

Sxl spans more than 20kbp and it has two promoters, *SxlPe* and *SxlPm*. The early *Sxl* promoter, *SxlPe* is first expressed at nuclear cycle 12 and the transcription is maintained until 5-10 minutes after entry to nuclear cycle 14 (Gonzalez et al., 2008; Lu et al., 2008; Mahadeveraju et al., 2020) (FIG. 1.5). The late *Sxl* promoter, *SxlPm* begins to be expressed at nuclear cycle 13 in XX embryo and late transcription is maintained until the completion of embryonic development (Lu et al., 2008). SXL is RNA binding protein which is known to bind poly (U) sequence with high affinity. Early SXL binds to poly (U) sequence in late *Sxl* pre-mRNA (Flickinger & Salz, 1994; Horabin & Schedl, 1993). SXL protein antagonizes spliceosome assembly near exon 3, resulting in the omission of exon 3 in mature female *Sxl* mRNAs (Flickinger & Salz, 1994; Lallena et al., 2002; Nagengast et al., 2003). Because exon 3 has a stop codon, exclusion of the exon 3 is critical to express functional late SXL in XX embryo. Once expressed, the late *Sxl* proteins bind to late *Sxl* pre-mRNA, establishing a stable positive autoregulatory loop (Cline, 1984) (FIG. 1.2). XY male embryos manage to activate *SxlPm* at early nuclear cycle 14 (Gonzalez et al., 2008). However, XY male embryos cannot produce functional late SXL proteins because one X chromosome signal does not activate *SxlPe* initially. Lack of early SXL causes inclusion of premature stop codon of exon 3 in the *Sxl* late mRNA transcripts, not producing functional late SXL.

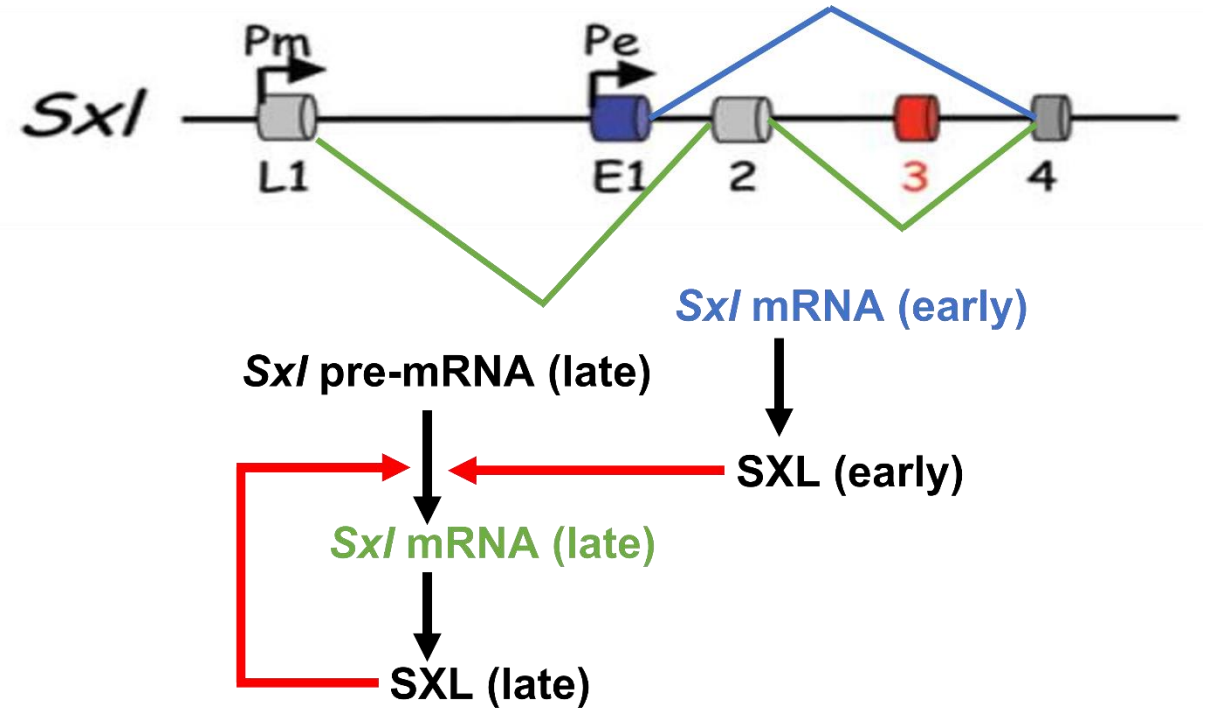

FIG. 1.2 : Autoregulation of *Sxl* in female embryos

From nuclear cycle 12 to early nuclear cycle 14 early *Sxl* promoter, *SxlPe* is expressed in response to two X chromosome signal. From nuclear cycle 13, *SxlPm* begins to be expressed in XX female embryo, transcribing *Sxl* late pre-mRNA transcripts. *Sxl* early proteins bind to *Sxl* late pre-mRNA transcripts, shifting splicing pattern of the *Sxl* late mRNA and the functional late *Sxl* proteins are expressed (Samuels et al., 1991). Late *Sxl* proteins feed back and establish a positive autoregulatory loop that maintains throughout the embryonic development.

Sxl target genes and dosage compensation

Functional late SXL post transcriptionally regulate the two direct target genes, *transformer (tra)* and *male specific lethal-2 (msl-2)* (Sanchez et al., 1994)(FIG. 1.3). In male embryo that does not express functional SXL, *tra* mRNA transcripts include a premature stop codon, producing non-

functional *tra* protein. In female embryo active late SXL antagonizes spliceosome assembly and the stop codon is spliced out in *tra* mRNA, producing functional *tra* protein products (Valcarcel et al., 1993). TRA itself is also a RNA binding protein, regulating splicing pattern of *doublesex (dsx)* and *fruitless* (*fru*) to produce female isoform proteins DSX^F and FRU^F (Burtis & Baker, 1989; Heinrichs et al., 1998; Rideout et al., 2007; Rideout et al., 2010). On the other hand, XY male embryos do not produce functional TRA so they express male isoform proteins DSX^M and FRU^M. DSX^F activates genes involved in female somatic cell development and represses genes involved in male differentiation (Burtis & Baker, 1989; Cline, 1979). FRU^M regulates male courtship behavior (Rideout et al., 2007; Rideout et al., 2010).

The other SXL target gene, *msl-2* is post transcriptionally regulated as well (Bashaw & Baker, 1997; Gergen, 1987). SXL binds to the 5'UTR of the *msl-2* pre-mRNA, preventing a small, male-specific intron from being spliced out (Gelbart & Kuroda, 2009). Because the full *msl-2* exon 1 inhibits translation, no MSL-2 protein will be produced in the presence of SXL. The MSL-2 protein is the core element to assemble male dosage compensation complex, so female embryo will not assemble the dosage compensation complex (Bashaw & Baker, 1997). In contrast, males do not express SXL and male dosage compensation complex can be assembled. The dosage compensation complex increases expression of the genes located on X chromosome. Failure to activate *Sxl* in XX female embryo or ectopic expression of *Sxl* in XY male embryo causes a sex specific lethal effect because of genetic imbalance.

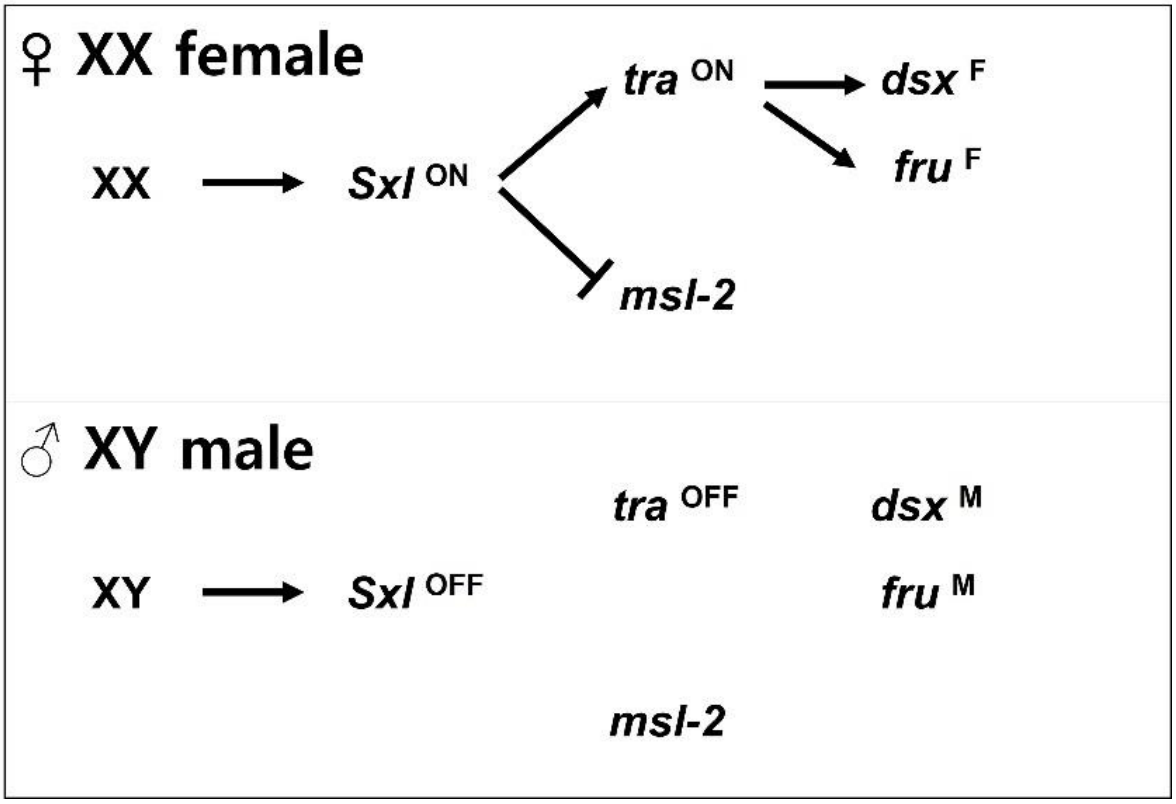

FIG. 1.3 : Regulation of *Sxl* target genes

Functional SXL in XX embryo activates *tra* and inhibits *msl-2* by shifting splicing pattern of RNA transcripts. Active *tra* induces expression of female isoform DSX^F and FRU^F that are involved in female differentiation. Lack of MSL-2 in XX embryo inhibits male dosage compensation complex assembly. XY males, on the other hand, does not express functional late *Sxl* protein and downstream target *tra* is inactive, producing male isoform DSX^M and FRU^M for male differentiation. Males express MSL-2 and they assemble dosage compensation complex.

X-Signaling Elements (XSEs) and *SxlPe* activation

The X chromosome signal in fly sex determination is delivered as transcription activators or signaling molecule that activates transcription. These X signal factors are called X-Signaling Elements (XSEs) and *sisterless A (sis A), sisterless B (sisB* or *sc), sisterless C (sisC)* and *runt* are the known XSEs (Cline, 1988; Cline & Meyer, 1996) (FIG. 1.4). *sisB* encodes class A basic helix-loop-helix protein that binds to E-box CANNTG sequence (Hoshijima et al., 1995; Yang et al., 2001). *sisB* RNA is first detected at nuclear cycle 9 and peaks at nuclear cycle 13, then quickly diminish at nuclear cycle 14 (Deshpande et al., 1995). *sisB* protein dimerizes with bHLH maternal activator Daughterless (Da) and activate *Sxl* transcription (Cline, 1976, 1978, 1988). Both canonical and non-canonical SisB/Da binding sites in *Sxl* enhancer were identified (Yang et al., 2001). *sisA* appears to encode a basic leucine zipper protein; however, it's binding sites are not known. The isolated protein is insoluble and structural predictions from the Charles Vinson lab suggest it requires a dimerization partner that has not been identified. (Erickson & Cline, 1993), (Erickson & Cline 1993, C.
Vinson personal communication.) *sisA* RNA expression is first observed at nuclear cycle 8 and the expression peaks at nuclear cycle 13 and quickly disappear at nuclear cycle 14. Thereafter *sisA* expression is limited to yolk nuclei and maintained. Unlike *sisB* and *sisA*, the two strongest XSEs that initiate *SxlPe* expression (Cline, 1988), *sisC* and *runt* have a relatively weak effect in *SxlPe* expression and seem to maintain *SxlPe* expression. *sisC* encodes secreted ligand that signals through Janus Kinase, Hopscotch, activating STAT92E transcription activator (Avila & Erickson, 2007; Jinks et al., 2000; Sefton et al., 2000). *sisC* RNA expression is first observed in nuclear cycle 13. During mid to late nuclear cycle 14 *sisC* expression is distinguished as faint

stripe patterns and the stripe pattern becomes clear in germ band extension stage (Avila & Erickson, 2007). Mutant *sisC* germline clones showed a defective *SxlPe* expression pattern, primarily in the central region of the embryo. *runt* is the founding member of the RUNX family transcription factors (Duffy & Gergen, 1991; Golling et al., 1996). While runt seems to be less dose sensitive than *sisA* or *sisB*, complete loss of *runt* causes no *SxlPe* expression in central region of female embryo and the defective phenotype is evident during nuclear cycle 13 and 14. The current evidence strongly suggests that Runt dimerizes with CBF- β protein Brother or Bigbrother to antagonize corepressor Groucho (Gro) (Mahadeveraju et al., 2020).

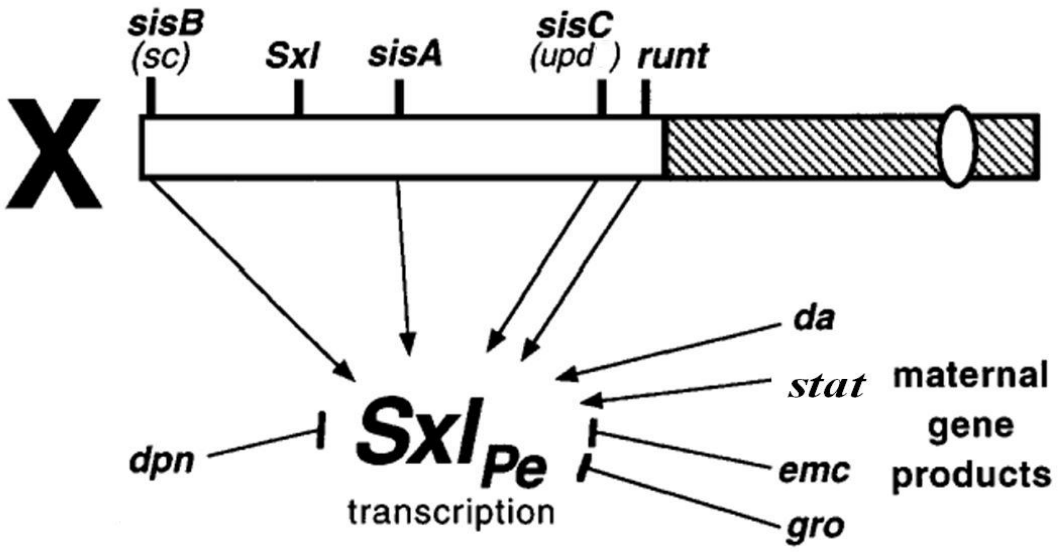

FIG. 1.4 : *SxlPe* regulation by XSEs and other zygotic and maternal regulators

Sxl and the four known XSEs (*sisB, sisA, sisC and runt*) are located on the X chromosome. Maternal *da* encodes transcription activator that dimerizes with SisB proteins and Stat92E is activated by signaling molecule SisC through JAK-STAT pathway. Zygotic *dpn* and *emc* repressor and maternal co-repressor *gro* represses *SxlPe* expression.

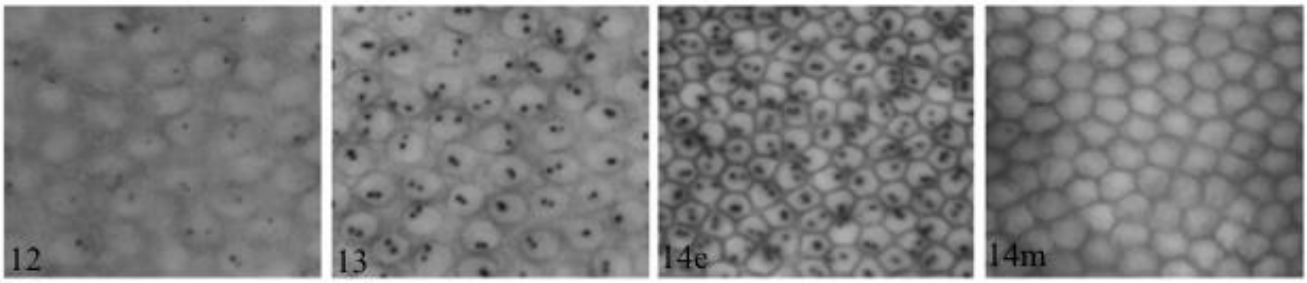

FIG. 1.5 : *SxlPe* expression in XX female embryos

In response to two XSE dose *SxlPe* is expressed in nuclear cycle 12 to early nuclear cycle 14. In situ hybridization images showing *Sxl* early nascent transcript expression.

Negative regulators of *SxlPe*

The XSEs as well as other zygotic and maternal transcription activators play critical function for *SxlPe* expression in XX female embryo. Negative regulators have a critical function to distinguish the XSE dose by establishing a threshold to activate *SxlPe* (Lu et al., 2008). One such negative regulator is autosomal *deadpan* (*dpn*). *dpn* is a member of Hairy-Enhancer of Split (HES) family, which is known to recruit co-repressor Groucho (Gro) (Barbash & Cline, 1995; Paroush et al., 1994). *dpn* protein is a basic helix-loop-helix transcription repressor that binds to E-box CACGTG and related CACGCG sequences. There are several canonical and non-canonical Dpn repressor binding sites in *Sxl* enhancer region (Lu et al., 2008). The current model is that zygotic Dpn and the co-repressor Gro establish threshold of *SxlPe* activation (Lu et al., 2008). XX females that have two doses of XSE, which is enough to activate *SxlPe* (FIG. 1.6). On

the contrary, one dose XSE in males cannot activate *SxlPe* because of transcription repression by the negative regulators.

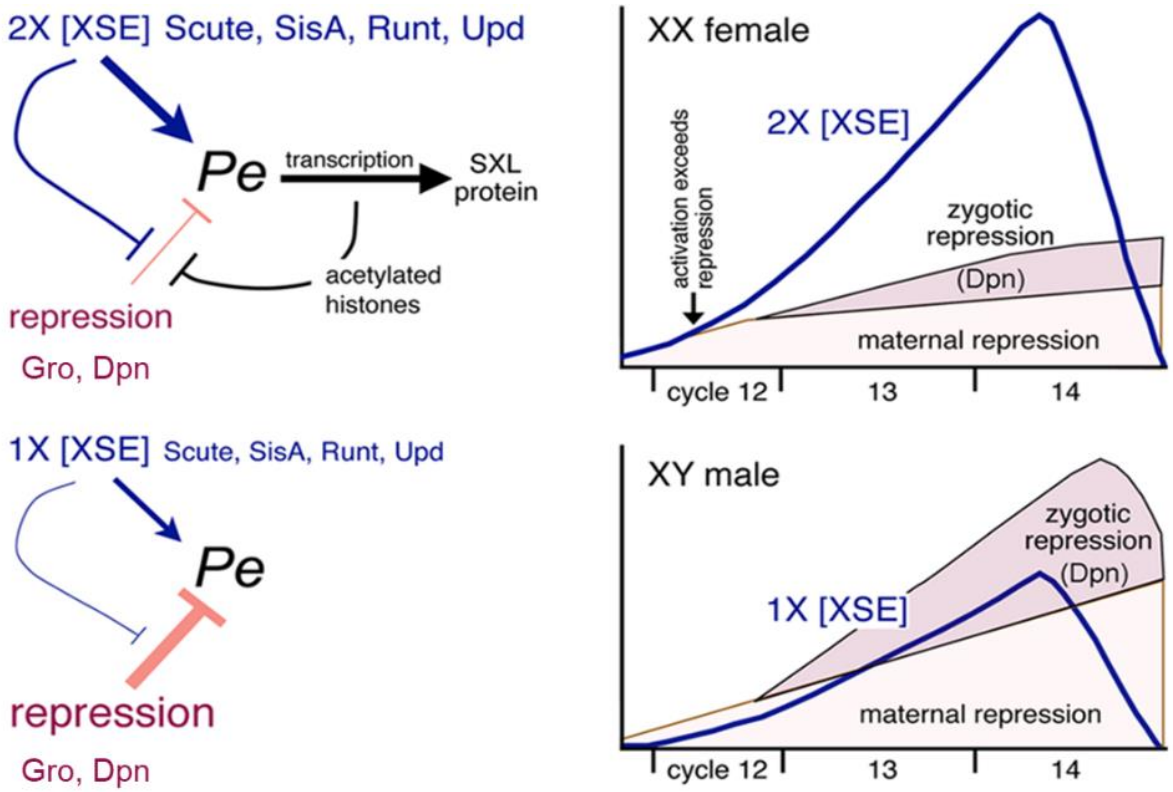

FIG. 1.6 : Current model of *SxlPe* regulation in XX female and XY male

Transcription repression by zygotic Dpn and maternal Gro are equal in XX and XY male embryo. Transcription repression signal rises as more zygotic Dpn accumulates through nuclear cycle 13 and 14. Starting nuclear cycle 12, XX female embryos produce enough XSE and XSE expression is maintained until early nuclear cycle 14. 2X dose XSEs antagonizes transcription repression from nuclear cycle 12 to early 14. However, XY embryos express a single dose XSE, which is continuously countered by the negative regulators.

Transgenic tools for studying *Sxl* regulation

Previous efforts to understand how sex specific expression of *SxlPe* relies on transgenes to mimic endogenous *Sxl*. *Sxl-LacZ* fusion transgene was commonly used for this purpose. Two of the most important discoveries from analyzing *Sxl-LacZ* expression pattern are that the 400bp proximal enhancer is sufficient for sex specific expression and a 1.4kb or bigger enhancer is required for normal *SxlPe* expression in female embryos (Estes et al., 1995). In fact, multiple transcription factor binding sites are clustered in the 400bp proximal enhancer region, supporting the idea that the 400bp proximal enhancer is important for *SxlPe* regulation (FIG.1.7). Recently, a full length *Sxl* transgene was developed, and we expected that the new transgene would better represent endogenous *Sxl* regulation. The new *Sxl* transgene was very informative and it seemed to better represent endogenous *Sxl* than the previous *Sxl-LacZ* transgene. Unexpectedly, I found that the new *Sxl* transgene was somewhat leaky in *SxlPe* expression and the experimental results from the new transgene were often inconsistent. Because precise quantification of *SxlPe* expression required endogenous *Sxl* genetic background, I decided to engineer endogenous *Sxl* mutants. Multiple endogenous *Sxl* mutant alleles were made by CRISPR/Cas9 and some of the previous works with transgenes were re-analyzed with the endogenous mutant lines. In chapter two, I will discuss how transcription repressor binding sites regulate sex specific expression of *SxlPe* and transactivation test in endogenous *Sxl* loci. In chapter three, I will explain how transcription activator sites contribute the primary sex determination. The last chapter will summarize my findings and discuss the implications of the results.

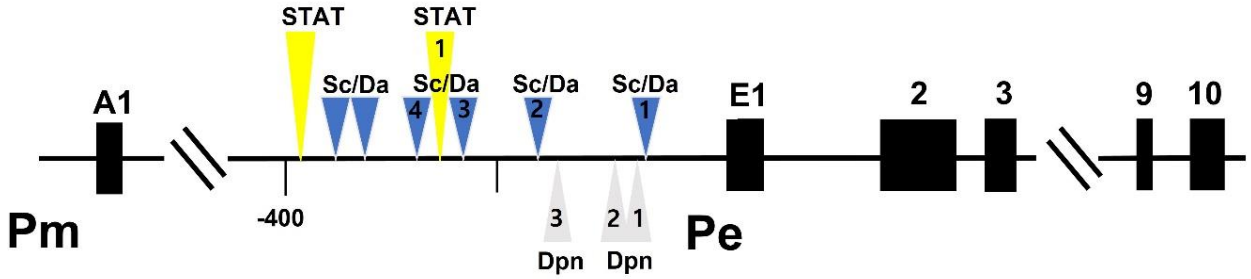

FIG. 1.7 : Map of the 400bp proximal *Sxl* enhancer and identified transcription factor binding sites

The map describes exons as black boxes and the two promoter *SxlPe* and *SxlPm*. Relevant transcription factor binding sites are described in the map.

CHAPTER II ANALYZING THE EFFECT OF REPRESSOR BINDING SITE MUTATIONS ON SXLPE EXPRESSION

Endogenous *SxlPe* is constitutively expressed by repressor binding site mutation

Negative regulators have an important role in primary sex determination of *Drosophila* because they help establish the XSE concentration threshold needed to activate *SxlPe* (Lu et al., 2008). Three repressor binding sites are clustered within 200bp of the transcription start site (FIG. 2.1 schematic). The three repressor binding sites interact with zygotic bHLH repressor Deadpan (Dpn), and likely other bHLH repressors such as Hairy, and maternally provided Hey. Repressor sites 1 and 2 are canonical CACGCG sequences, which are optimal binding sequence for HES (Hairy-Enhancer of Split) class repressors like Dpn. Repressor site 3 has a non-canonical CACACT sequence that binds Dpn in vitro with somewhat less affinity than canonical sequences (Lu et al., 2008). Previous studies in our lab showed mutating any of these repressor binding sites in the context of a 1.4 kb *SxlPe-lacZ* transgene caused ectopic expression of *lacZ* mRNA in male embryos with alterations of sites 1 and 2 exhibiting a somewhat stronger effect than site 3 (Kappes et al., 2011; Lu et al., 2008). However, Kappes et al. (2011) observed that mutating repressor sites 1 and 2 did not cause ectopic expression from smaller 0.4 kb *SxPe-LacZ* transgenes. One explanation for their observation was that a positive transcription regulator, dMyc, may share the binding sites with repressors and the 0.4 kb *SxlPe-lacZ* transgenes were more dependent on dMyc than the longer transgenes (Kappes et al., 2011)(Jung et al. ref, in preparation). More recently, we found that mutating repressor sites 1 or 2 induced ectopic expression of *SxlPe* in male embryos in the full-length 40kb *Sxl* transgene (FIG 2.1).

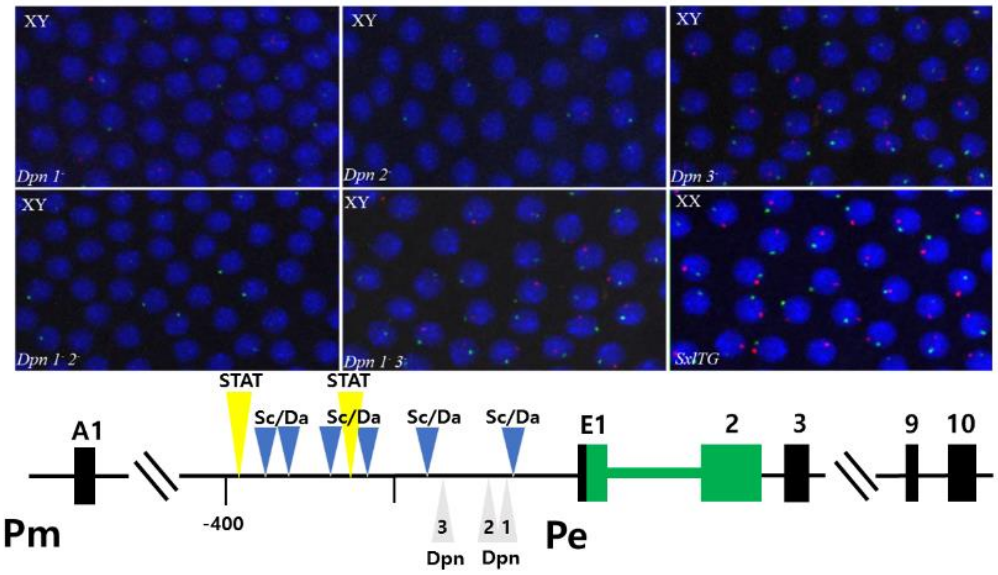

FIG. 2.1 Constitutive *SxlPe* expression of *Sxl* transgenes by repressor site mutations

Schematic map of *Sxl* transgenes with a *D. eugracilis* exon-intron sequence swap (green) and mutated Dpn repressor binding sites. Fluorescent in-situ hybridization images by confocal microscopy with 20X objective showing male (XY) or female (XX) embryos that are deleted for the endogenous *Sxl* locus (Sxl^{f7bo}) and heterozygous for the indicated Dpn site mutation transgene (green) and a wildtype transgene (red). Male embryos with Dpn site 1 and 2 mutant transgenes express nascent transcripts almost exclusively from the mutant (red) transgenes. Embryos with the Dpn site 3, Dpn sites 1 and 2, and 1 and 3 express both mutant (red) and wild-type (green) transgenes in many nuclei, indicating that the constitutive allele trans activates the normally silent wild type *Sxl* transgene. Females express both mutant and wild-type transgenes in all nuclei.

Further analyses of the repressor site mutations in transgenes provided two unexpected results. First, we found that the non-canonical repressor site 3, which is known to bind Dpn less efficiently in vitro than the canonical sequence (Lu et al., 2008), caused strong ectopic *SxlPe* expression in male embryos. Second, as detailed in the last section of this chapter, we found evidence that the constitutive repressor binding site mutant *Sxl* transgenes were capable of activating a wild-type *Sxl* transgene in trans, suggesting that a transvection-like phenomenon might be associated with the *Sxl* locus. To address both these questions I developed CRISPR/casp9 technology to mutate the repressor binding sites at the endogenous *Sxl* locus.

To see if the non-canonical repressor site 3 has a strong effect in *SxlPe* repression, repressor site 3 was targeted by CRISPR/Cas9 mutagenesis (FIG. 2.2). My initial genetic screening of site 3 mutant candidates relied on the assumption that an ineffective Dpn binding site should cause ectopic *Sxl* expression in males and result in at partial male-lethality. As expected, my initial genetic screen identified several candidates from the CRISPR injected lines with male-lethal effects. The magnitude of the male-lethal effect observed during the screen was variable, with male viability ranging from 50-90%. The putative mutant lines were tested by PCR amplification of the *Sxl* loci followed by HindIII restriction digestion. Twelve positive mutant lines were recovered in the screen and one was named as Sxl^{MD3}, following the convention of a superscript M denoting male-lethal *Sxl* alleles.

To see if the Sxl^{MD3} allele is also partially constitutive in females, I asked if Sxl^{MD3} could suppress the female-lethal effects of reduced XSE gene dose. The *sisA* and *sisB* genes are the two strongest XSE elements and reducing their doses from two to one in females greatly reduces their viability (Cline, 1988). As expected, the control experiment showed that none of the $sc^{sisB3\text{-}1}$ $sisA^{+}$/+ + females were viable when they carried two wild type *Sxl* alleles. (Table 2.1). In

contrast, the presence of a single copy of Sxl^{MD3} rescued at least 54% of $sc^{sisB3\text{-}1}$ $sisA^{+}$/+ + females demonstrating constitutive expression of Sxl^{MD3}. A smaller female-lethal effect was observed with a weaker *sisB* allele, sc^{sisB3}, and $sisA^{1}$ (50% female viability), however, Sxl^{MD3} suppressed the female-lethal effect resulting in at least 78% female-viability. It is important to note that female viability was likely underestimated in crosses with Sxl^{MD3} as the partially male-lethal mutation likely reduced the number of reference males. To summarize, the repressor site 3 mutation (Sxl^{MD3}) causes partially constitutive *SxlPe* expression in both males and females, implicating the importance of the non-canonical repressor site in *Drosophila* primary sex determination.

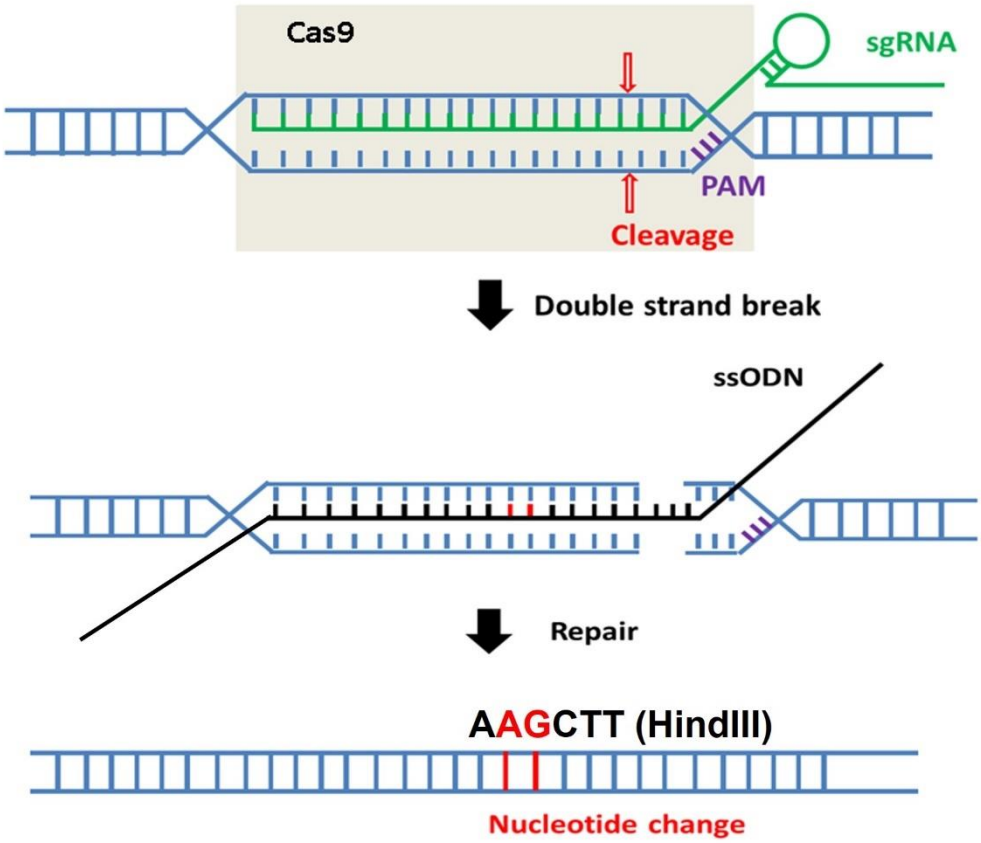

FIG. 2.2 CRISPR/Cas9 mutagenesis engineering for *Sxl*^{MD3} allele

The endogenous *Sxl* loci (blue) is targeted by guide RNA (green) that hybridizes to repressor site 3. The guide RNA recruits maternally expressed Cas9 proteins which then cleave the target site, leaving a double-stranded break. A single-stranded oligonucleotides (ssODN) that contains a mutated repressor site 3 is used to repair the damage via homologous recombination. The site 3 change converts the Dpn binding sequence from CACACTt to the non-functional CAagCTt creating a AAGCTT HindIII restriction site, that allows efficient screening via PCR and restriction analysis. A similar strategy was used to create mutations in repressor sites 1 and 2 as detailed in Material and Methods (Chapter IV).

Genotype	♀ Viability % (#)
♀ $\frac{+ \quad sxl^{+} \quad +}{sc^{sisB3-1} + \quad sisA^{1}}$	0 (150)
♀ $\frac{+ \quad sxl^{MD3} \quad +}{sc^{sisB3-1} + \quad sisA^{1}}$	53.63 (110*)
♀ $\frac{+ \quad sxl^{+} \quad +}{sc^{sisB3} + \quad sisA^{1}}$	50 (100)
♀ $\frac{+ \quad sxl^{MD3} \quad +}{sc^{sisB3} + \quad sisA^{1}}$	77.97 (110*)

Table 2.1 Constitutive *Sxl* allele, Sxl^{MD3}, partially rescues females from lethal effects of reduced XSE dose

Sxl^{MD3} carries a Dpn repressor binding site 3 mutation. $sc^{sisB3-1}$ and sc^{sisB3} are hypomorphic alleles of the XSE gene *sisB*. $sisA^{1}$ is a hypomorphic allele of the XSE *sisA*. Data are expressed as % female viability relative to control male siblings. Crosses were: ♀♀ *y w cas9* Sxl^{+} x ♂♂$sc^{sisB3-1}sisA^{1}$ (29°), ♀♀ *y w cas9* Sxl^{MD3} x ♂♂$sc^{sisB3-1}sisA^{1}$ (29°), ♀♀ *y w cas9* Sxl^{+} x ♂♂ sc^{sisB3} $sisA^{1}$ (25°) and ♀♀ *y w cas9* Sxl^{MD3} x ♂♂$sc^{sisB3}sisA^{1}$ (25°). The number of male progeny from each cross (parentheses) served as the viability reference. Apostrophes indicate crosses where male viability was likely reduced due to the Sxl^{MD3} allele meaning that % female-viability was probably underestimated.

All three repressor sites contribute to *SxlPe* regulation

The genetic evidence indicates that mutant repressor site 3 induces partially constitutive *SxlPe* expression. To confirm that repressor sites 1 and 2 also contribute to *SxlPe* regulation, I designed a CRISPR/Cas9 strategy to introduce mutations in repressor sites 1 and 2 into flies carrying the Sxl^{MD3} allele. As expected, I recovered flies with all three repressor sites mutated, Sxl^{MD123}, but unexpectedly, also found a variant with mutant repressor sites 2 and 3, but an unaltered, wild-type, repressor site 1, Sxl^{MD23}. Because ectopic *SxlPe* expression in males causes lethal effect, I expected that the constitutive alleles generated from the CRISPR/Cas9 mutagenesis will kill the males. As expected, I found that Sxl^{MD23} males were 30% viable and Sxl^{MD123} were 7-8% viable indicating there is an additive effect of the three repressor site mutations. To confirm that Sxl^{MD3}, Sxl^{MD23} and Sxl^{MD123} cause increased *SxlPe* expression, I performed in-situ hybridization experiments to detect both nascent and mature transcripts from *SxlPe* (FIG. 2.3). As expected, all three mutant alleles were constitutively expressed. Mutant male embryos begin to express *Sxl* at nuclear cycle 13 (data not shown) and the expression was maintained through early nuclear cycle 14. Comparing the frequency of nuclei expressing *Sxl* clearly showed a positive correlation between the number of mutated repressor sites and number of nuclei actively expressing *SxlPe*.

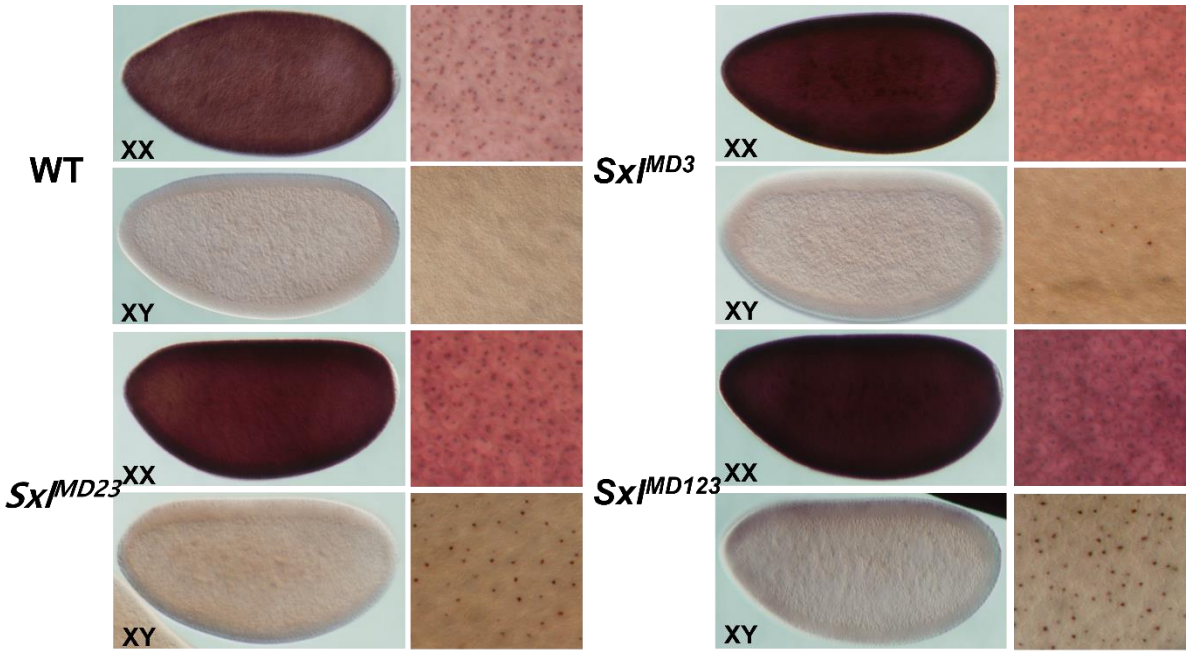

FIG. 2.3 Repressor site mutations induce constitutive *SxlPe* expression

Wild-type and repressor mutant embryos were stained following *in situ* hybridization. RNA probes complementary to *Sxl* exon E1 and downstream intronic sequences were used to allow detection of both nascent and mature transcripts from *SxlPe*. Embryos in early (< 10 min) nuclear cycle 14 are shown. Whole embryo images illustrate the accumulation of *SxlPe*-derived mRNA as evidenced by overall purple staining. Expanded surface views, of the central regions of the same embryos, show nascent transcripts from the X-linked *SxlPe* promoter as dots of staining in nuclei. Embryos were collected from homozygous *y w cas9 Sxl*$^{+}$(wild-type *WT), y w cas9 Sxl*^{MD3}, *y w cas9 Sxl*^{MD23} and *y w cas9 Sxl*^{MD123} lines. Female and male embryos were distinguished by the presence or absence of staining (WT), or by the presence of one (male) or two (female) dots in expressing nuclei.

Sxl^{MD3} males showed sporadic *SxlPe* expression, visible as nuclear dots representing the nascent transcripts. *Sxl*^{MD23} males exhibited more frequent *SxlPe* expression while most *Sxl*^{MD123} males exhibited *SxlPe* activity in most of their nuclei (Fig. 2.3). The correlation between the number of mutated repressor sites and *SxlPe* activity was also evident in female embryos. All females expressed both copies *Sxl* but the intensity of staining appeared strongest in *Sxl*^{MD123} females followed by *Sxl*^{MD23}, *Sxl*^{MD3}, and wild-type females (Fig. 2.3). These results strongly suggest that each of the three repressor sites is important for normal *SxlPe* regulation in both sexes and that the three sites work in an additive manner. The elevated *SxlPe*-derived mRNA levels in mutants defective in repressor sites 1 and 2 offers no support for the proposal that these sites are also bound and regulated by the transcription activator dMyc (Kappes et al., 2011). A finding consistent with our lab's failure to find genetic evidence for an involvement of dMyc in sex determination (J. W. Erickson, unpublished data.).

Next, I asked if the constitutive expression of *Sxl* early transcripts caused increased *Sxl* protein expression. Analysis of SXL staining pattern in nuclear cycle 14 embryos showed that mutating the repressor sites caused ectopic SXL expression in at least some males in each of the genotypes (FIG. 2.4). *Sxl*^{MD3} and *Sxl*^{MD23} males were mixed population of stained embryos and non-stained embryos. On the other hand, virtually all *Sxl*^{MD123} males (43/45) showed some level of SXL expression with variable intensity. Comparing SXL expression in wildtype and *Sxl*^{MD123} females showed that the intensity of *Sxl*^{MD123} female staining was greater, consistent with increased mRNA levels observed in mutant females.

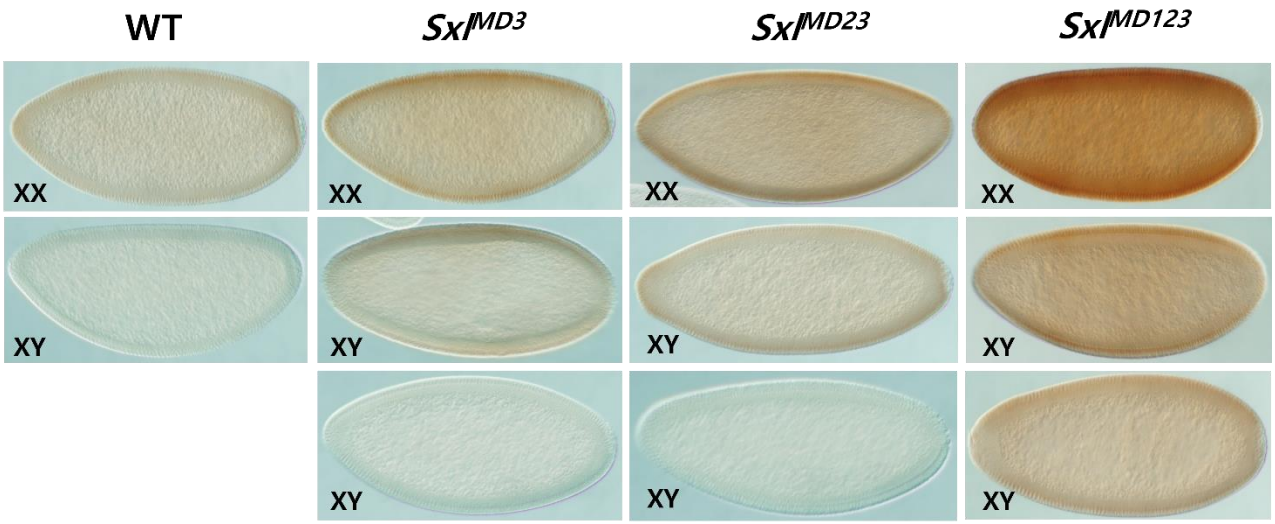

FIG. 2.4. Repressor site mutations induce constitutive *Sxl* protein expression

Sxl protein in wild-type and repressor binding site mutant embryos. Mid to late nuclear cycle 14 embryos are shown. Wild-type female embryos express SXL and wild-type males do not. For Sxl^{MD3}, Sxl^{MD23}, and Sxl^{MD123} one half of embryos observed stained darkly and were taken to be XX females. The remainder exhibited SXL staining ranging from unstained to moderately stained and were taken to be male. For Sxl^{MD3} and Sxl^{MD23} the presumed XY male embryos showed variable expression of SXL with those shown representing the strongest and weakest staining observed. For Sxl^{MD123} occasional unstained embryos were observed, but the two embryos shown were judged to be representative of the range of staining most typically seen. Embryos were collected from homozygous *y w cas9* Sxl^{+}*(WT), y w cas9* Sxl^{MD3}*, y w cas9* Sxl^{MD23} and *y w cas9* Sxl^{MD123} flies at 25˚C.

Does loss of repressor binding sites render the strong XSEs unnecessary for sex determination?

Data in Table 2.1 show that the partially constitutive allele Sxl^{MD3}, suppresses the female-lethal effects of reduced *sisA* and *sisB* dose. We wondered whether the stronger constitutive Sxl^{MD123} allele might also suppress the complete loss of *sisB* function in females, and conversely, if loss of *sisB* function would rescue males from the male-lethal effect of Sxl^{MD123}. Data in table 2.2 show the answer to both questions is yes. Females with no functional copies of *sisB*, but two copies of Sxl^{MD123}, are fully viable, and the strong male-lethality of Sxl^{MD123} is fully suppressed by the loss of *sisB* function. Stocks of sc^{M6} Sxl^{MD123} flies are stable and are maintained with an approximately 50/50 sex ratio indicating that the cis-acting Sxl^{MD123} allele mutation renders the X-chromosome counting mechanism independent of the normally essential XSE *sisB*. I reasoned that mutating the three Dpn repressor binding sites allows *SxlPe* to be active, not requiring transcription activator binding.

To test the idea further we found that Sxl^{MD123} flies defective for the two strong XSE elements *sisB* and *sisA* can be maintained in a stable stock that produces an approximately equal sex ratio (Preliminary data suggests that males may be slightly underrepresented, J. W. Erickson unpublished.). In effect, this indicates that proper X-chromosome counting can occur in the almost total absence of the two most potent XSE activators when cis-acting changes in repressor binding sites render *SxlPe* partially constitutive. To see what happens to *Sxl* protein expression in Sxl^{MD123} embryos that lack *sisB*, or both *sisB* and *sisA*, functions I stained embryos collected from the two stocks. As shown in FIG. 2.5 the stocks produce two types of embryos in equal proportions. Moderately-stained embryos that express *Sxl* protein at levels that appear to be slightly reduced compared to wild-type, and lighter-stained embryos that express low-levels of

SXL. Given the high male viability this suggests that the male embryos either do not express SXL or they express low levels of SXL, which is insufficient to establish or maintain late SXL protein expression. The significance of these findings is not yet clear. At one extreme it could be argued that the triply mutant X chromosome functions effectively as novel sex chromosome with a different X-counting mechanism. At the other extreme, it could be argued that that the viable stocks simply represent a predictable example of genetic suppression that, by happenstance, delivers an approximately 50/50 sex ratio. Mechanistically, our findings do suggest that the simple conventional view that *sisB* and *sisA* encode essential direct transcriptional activators of *SxlPe* requires revision.

Genotype	Viability % (#)	Genotype	Viability % (#)
♀ *sc^{M6}Sxl^{MD123}*	101.08 % (94)	♀ *sc^{sisB3}Sxl^{MD123}*	103.05 % (135)
♀ *sc^{M6}Sxl^{MD123}/FM6, ct*	100 % (93)	♀ *sc^{sisB3}Sxl^{MD123}/FM6, ct*	100 % (131)
♂ *sc^{M6}Sxl^{MD123}*	95.70 % (89)	♂ *sc^{sisB3}Sxl^{MD123}*	96.95 % (127)
♂ *FM6, ct*	61.29 % (57)	♂ *FM6, ct*	66.41 % (87)

Table 2.2 Constitutive *Sxl* allele, *Sxl^{MD123}* counters loss of *sisB*

Percentage viability of female and male progeny from the crosses: ♀♀ *sc^{M6} w Sxl^{MD123} ct sn/FM6, ct* x ♂♂ *sc^{M6} w Sxl^{MD123} ct sn/Y* and ♀♀ *sc^{sisB3}Sxl^{MD123}*/FM6, ct x ♂♂ *sisB^{3}Sxl^{MD123}/Y*. Raw numbers of flies observed are in parentheses. *sc^{M6}* is an early nonsense mutation that nearly eliminates *sisB* function and *sc^{sisB3}* is a strong hypomorphic *sisB* allele.

FIG. 2.5 Loss of the XSE function reduces SXL expression from the constitutive *Sxl*^{MD123} allele in both sexes

Sxl protein staining in wild-type (WT) and in homozygous XSE mutants carrying the constitutive *Sxl*^{MD123} allele. Mid to late nuclear cycle 14 embryos are shown. Wild-type female embryos express SXL and wild-type males do not. *sc*^{M6} and *sisA*^{1} are strong hypomorphic alleles of the XSE genes *sisB* and *sisA.* Half the homozygous *sc*^{M6}*Sxl*^{MD123} and *sc*^{M6}*Sxl*^{MD123} *sisA*^{1} embryos showed relatively normal SXL expression patterns, with a small decrease in staining intensity compared to wild-type, and were presumed to be female. The remainder of the *sc*^{M6} *Sxl*^{MD123} and *sc*^{M6} *Sxl*^{MD123} *sisA*^{1} embryos exhibited background or low-level SXL staining and are presumed to be male.

Transactivation/Transvection does not appear to be involved in regulating endogenous *SxlPe* expression

Previous results with our *SxlPe* transgenes showed that *Sxl* transgenes rendered partially constitutive by cis-acting repressor site mutations were capable of activating *SxlPe* expression from a wild type homolog in trans (Rajendren dissertation, 2015). This raised the exciting possibility that a kind of cooperative phenomenon whereby the activation of *SxlPe* on one X chromosome facilitated that activation on the other X chromosome could play an important role in ensuring the robust establishment of female-specific *Sxl* transcription. The embryos shown in FIG. 2.1 were males deleted for the endogenous *Sxl* locus, but that carried two different *Sxl* transgenes. One transgene was wild-type and the other was expressed constitutively due to repressor site mutations. In such males, one would expect to see *Sxl* expression from the constitutive allele but not from the wild-type allele as only one X chromosome was present. Instead, we found that the wild-type *Sxl* transgene was expressed in the presence of a constitutive transgene (but not in the presence of a 2nd wild-type transgene). Furthermore, that level of expression of the wild-type allele was correlated with the strength of constitutive transgene. Specifically, the greater the number of repressor sites mutated in the constitutive allele, the greater the number of nuclei that also activated the wild-type transgene. These results strongly suggested that the constitutive *Sxl* transgene somehow activated the wildtype *Sxl* transgene in trans. A phenomenon reminiscent of transvection, a chromosome pairing-dependent transactivation phenomenon well known in flies (Lee & Wu, 2006; LEWIS, 1954).

One possible explanation of transactivation by the constitutive transgenes was that the *Sxl* protein might possess a previously unrecognized autoregulatory transcription factor function that

could bind to, and activate, the *SxlPe* promoter. I excluded this possibility by showing that an endogenous wild-type X-linked *Sxl* allele was not activated by the presence of the strongest constitutive *Sxl* transgenes (FIG. 2.7).

A more intriguing alternative was that the transactivation could require chromosome pairing (classical transvection) or at least that the alleles be in close proximity. Because *Sxl* is located on X chromosome, this idea could only be tested in female embryos. The issue with females is that normal females activate both alleles of *SxlPe* during within a few minutes during nuclear cycle 12, making it extremely challenging quantify the effects of transactivation, if any. My strategy to overcome this problem was to make females mimic males by reducing XSE dose. As seen in Table 2.1, *SxlPe* activation is sensitive to XSE dose. In this reduced XSE background, the wild-type *SxlPe* would not normally expressed. However, the constitutive *Sxl* allele, Sxl^{MD3} or Sxl^{MD123}, would be still active. If the constitutive allele were capable of activating the normal allele in trans, then expression from the normal *SxlPe* allele could be observed using in situ hybridization. My initial plan was simply to count the number of nuclei expressing both the constitutive and wild-type alleles in a reduced XSE dose genotype. Two findings made that approach impractical. The first was my finding that upon close inspection *SxlPe* expression was leaky and highly variable in embryos with reduced XSE dose. Some embryos barely expressed any *Sxl*, while others expressed one or two alleles in substantial numbers of nuclei. There was also considerable regional variation of expression within individual embryos with no consistent pattern. My second finding was that while I could easily detect expression from the constitutive alleles in males (FIG. 2.6), expression was also somewhat variable with some nuclei failing to express the constitutive alleles. To avoid the complications of having to assess two variable signals in each nucleus, I inserted an epitope tag, *Llama* (Bothma et al., 2018), in an otherwise

normal *Sxl* gene so, that I would detect wild-type *Sxl* expression using *Llama* sequence-specific RNA probes (FIG. 2.7 A). The approach allowed me to observe activation of only the normal, Llama-tagged, *Sxl* allele, making quantification of the *Sxl* dots a binary task of counting of activation or inactivation, without considering constitutive *Sxl* alleles.

Two groups of embryos were tested in the transactivation tests. The control group has two copies of normal *Sxl*, one marked by the Llama tag, and the experimental group a normal Llama-tagged *Sxl* allele and an un-tagged constitutive Sxl^{MD123} allele. The results showed that there was no difference in the fraction of nuclei that express the normal *Sxl* in these two groups, suggesting that there is no transactivation of endogenous *Sxl* allele (two-tailed t test, p=0.46, FIG. 2.7B).

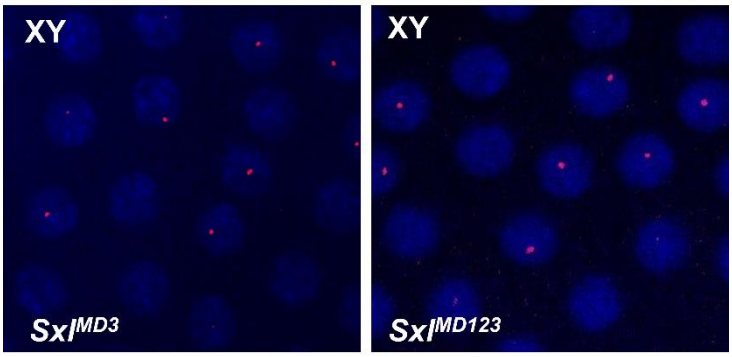

FIG. 2.6 FISH of nascent *SxPe* transcripts in constitutive males

Sxl^{MD3}, a *Sxl* allele defective for repressor site 3 and Sxl^{MD123}, a *Sxl* allele with repressor sites 1, 2 and 3 mutations show ectopic *SxlPe* expression in nuclear cycle 13 male embryos. Red dots represent *SxlPe* transcripts. Nuclei are stained with DAPI (blue). Brightness and contrasts of the images were adjusted.

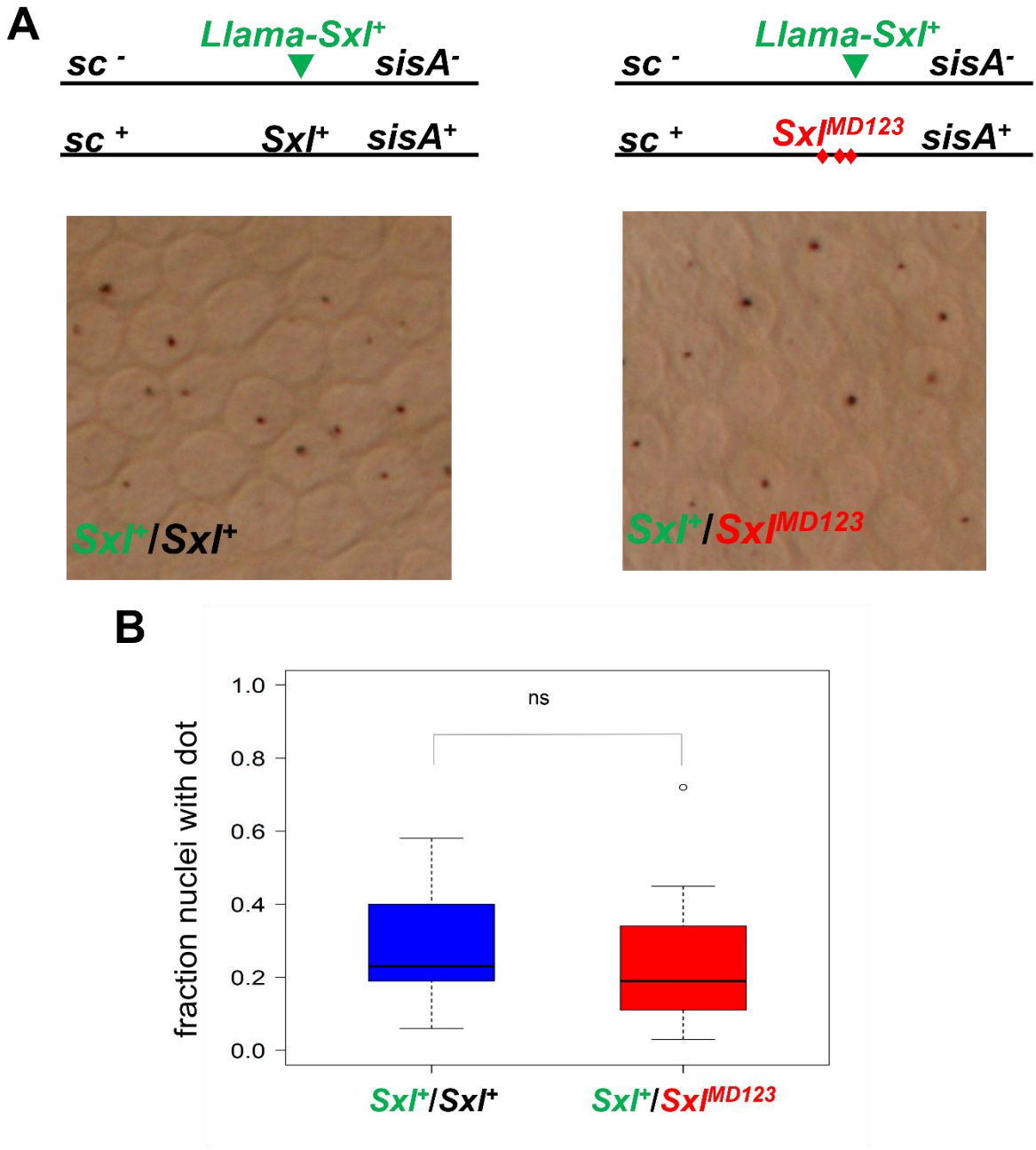

FIG. 2.7 Endogenous *Sxl* transactivation test

Transactivation test at the endogenous *Sxl* locus with reduced XSE signal. (A) Schematic of the *Sxl* alleles used. Green triangles represent the *llama*-tag inserted at the normal *Sxl* locus. Red diamonds represent the mutant repressor binding sites in the constitutive Sxl^{MD123} allele.

Activation of the normal *Sxl* allele is reduced in an XSE deficient background whereas expression from the constitutive Sxl^{MD123} allele is minimally affected. Images show surface views of nuclear cycle 13 female embryo after in situ hybridization to detect the Llama containing transcripts. Embryos were from the crosses: ♀♀ *y w cas9* Sxl^{+} x ♂♂ $sc^{sisB3\text{-}1}$ *Llama-*Sxl^{+} $sisA^{1}$/Y and ♀♀ Sxl^{MD123} x ♂♂ $sc^{sisB3\text{-}1}$ *Llama-*Sxl^{+} $sisA^{1}$/Y. $sc^{sisB3\text{-}1}$ is a strong hypomorphic *sisB* allele and $sisA^{1}$ is a strong hypomorphic *sisA* allele. (C) Box plot comparing fraction of nuclei expressing *Sxl* in the presence or absence of Sxl^{MD123}. A total of 9 embryos were analyzed for Sxl^{+}/Sxl^{MD123} group and 14 embryos were analyzed for Sxl^{+}/Sxl^{+} group. p=0.4567, two-sampled t test.

To provide a compelling answer for *Sxl* transactivation, a stringent genetic test was performed. Most female embryos who lose one copy of *sisA* and *sisB* die, however, I showed that Sxl^{MD3} restored viability to females with reduced XSE dose (Table 2.1). If rescue were due only to Sxl^{M3} producing *Sxl* protein, then a constitutive allele that produced no SXL would fail to rescue the females. On the other hand, if transactivation occurs, the constitutive *SxlPe* should be able to activate the wildtype *Sxl* homolog and rescue at least some of the affected females. To engineer a new *Sxl* double mutant that has constitutive promoter but produces non-functional proteins, a frameshift mutation was introduced in exon 5 in flies carrying the Sxl^{MD3} and Sxl^{MD123} alleles by CRISPR/Cas9. The engineered alleles are $Sxl^{f34,\ MD3}$ and $Sxl^{f34,MD123}$, following the standard nomenclature for *Sxl* alleles. The same frameshift mutation was also introduced into the normal *Sxl* gene as a control the allele Sxl^{f34}. Genetic test showed that females carrying a Sxl^{f34} allele in the $sis^{B3}sisA^{1}$ background were fully lethal as expected. The experimental females

carrying $Sxl^{f34,MD3}$ or $Sxl^{f34,MD123}$ were also completely lethal in the same genetic background, suggesting that there is no transactivation at endogenous *Sxl* loci (data not shown).

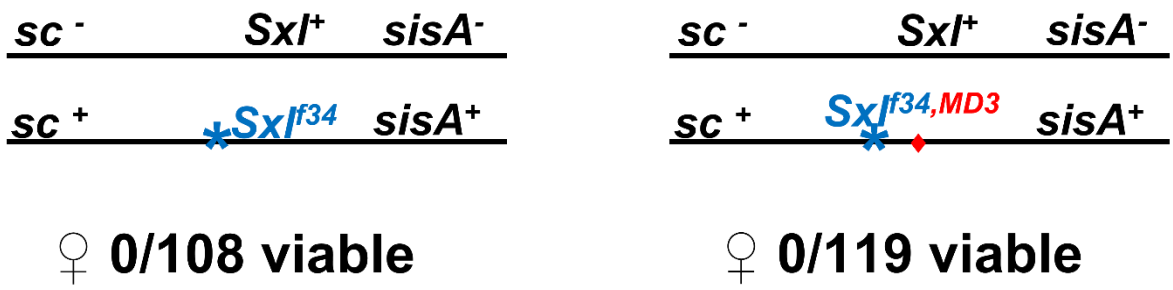

FIG. 2.8 Genetic test finds no evidence for *Sxl* transactivation

If a constitutively active *SxlPe* is sufficient to activate a second copy of *SxlPe* in trans, then transactivation should occur even when the constitutive allele produces no *Sxl* protein. Sxl^{f34} is a non-functional allele causing a frameshift in exon A5. The double mutant, $Sxl^{f34,MD123}$, retains constitutive *SxlPe* promoter activity (data not shown), but does not produce functional *Sxl* protein. The double mutant fails to activate Sxl^{+} in trans as illustrated by its inability to restore female viability. Data are number of females of indicated genotypes recovered compared to control sc^{-} Sxl^{+} $sisA^{-}$ */Binsinscy* females. Crosses were: ♀♀ *y w cas9* Sxl^{f34}/Binsinscy x ♂♂ $sc^{sisB3\text{-}1}sisA^{1}$/Y and ♀♀ $Sxl^{f34,MD123}$/Binsinscy x ♂♂ $sc^{sisB3\text{-}1}sisA^{1}$/Y. No viable $Sxl^{f34}/sc^{sisB3\text{-}1}sisA^{1}$ or $Sxl^{f34,MD123}/sc^{sisB3\text{-}1}sisA^{1}$ females were observed. Under similar conditions, both Sxl^{MD3} and Sxl^{MD123} effectively rescue female lethality (Table 2.1, data not shown).

I did not further address the mechanism responsible for the transactivation phenomenon seen with the constitutive *Sxl* transgenes integrated at the attP40 site on the 2nd chromosome. We

presume it is related to the site of integration as there is growing evidence for transvection-like phenomena at various chromosomal locations (King et al., 2019). I note, that the mechanism proposed by King et al. (2019), that transvection occurs at chromosomal loci that promote high-level expression in cis, does not fit perfectly with our observations of the constitutive *Sxl* transgenes as the transgenes and the endogenous X-linked locus appear to be expressed at comparable levels as measured by both genetic and in situ hybridization analysis.

CHAPTER III ANALYZING THE EFFECT OF TRANSCRIPTION ACTIVATOR BINDING SITE MUTATIONS ON SXLPE REGULATION

A Predominant effect of SisB/Da activator site 3 mutation in *SxlPe* expression

The current model of fly sex determination is that the twofold dose difference in XSE transcription factors is the critical element for *SxlPe* expression. To see how transcription activator binding sites regulate *SxlPe* expression, our lab previously analyzed the effects of mutating the activator binding sites in our *Sxl* transgene system, which better represents endogenous *Sxl* regulation than previous *Sxl-lacZ* fusion transgene. The results showed that all the transcription activator sites had measurable importance for primary sex determination although most exhibited somewhat modest effects. Surprisingly, a E-box binding site (site 3) for the SisB/Da activator appears particularly important for *SxlPe* activation. A mutation that in vitro binding sites of SisB/Da at the activator site 3 greatly reduced *SxlPe* expression from our full length *Sxl* transgene, suggesting that somehow this single binding site has a critical role for transcription ; a finding quite unusual in eukaryotic transcriptional control. The result was directly contradictory to previous observations that mutating any single SisB/Da activator site, including the site 3 had no substantial effect on *SxlPe* expression in lacZ transgenes—a finding by both our lab and Paul Shedl's (Yang et al., 2001). To answer if the activator site 3 plays an important role in *SxlPe* expression at the normal *Sxl* locus, I generated multiple mutant *Sxl* alleles by CRISPR/Cas9. Building on the techniques I developed to create Dpn binding site mutations (FIG. 2.2). I developed a modified and highly efficient CRISPR/Cas9 protocol that allows multiple mutant *SxlPe* alleles to be generated in a single CRISPR injection (FIG. 3.2). The strategy relies on using two efficient guide RNAs. One targets at -500 relative to the *SxlPe*

transcription start site, the other targets intronic sequence between exon E1 and A2. Because both guide RNAs are very efficient, the entire 400bp proximal *Sxl* enhancer could be effectively deleted by recruited Cas9 enzyme and then replaced from a segment donated from co-injected plasmid constructs with desired mutations. Successful integration of the mutation can be quickly identified by the presence of EcoRI site in genomic *Sxl* locus. (The change creating the EcoRI site produces no detectable effects on *Sxl* function. The EcoR1 variant serves as the wild-type control in the experiments presented in this chapter). Because guide RNA targeting, and Cas9 cleavage are so efficient, multiple different plasmids were co-injected together and all 3 of the desired mutant alleles were successfully recovered.

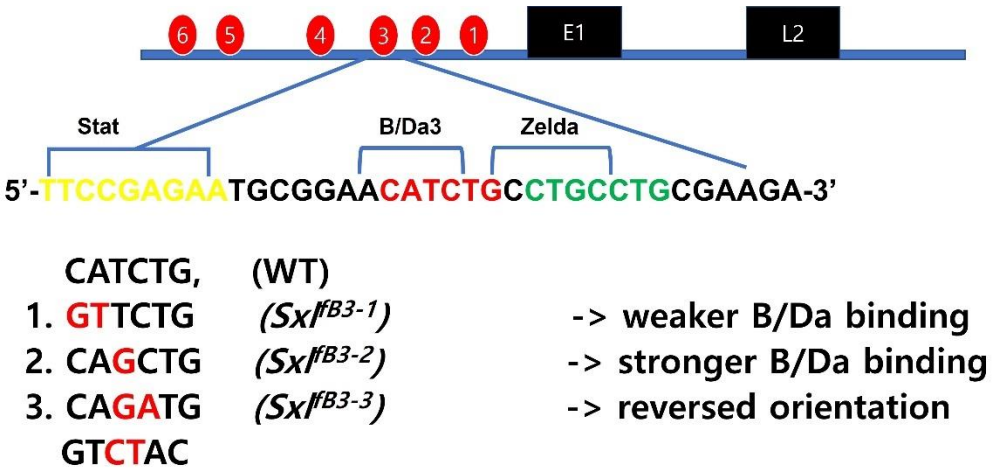

FIG. 3.1 Map of SisB/Da activator binding sites in the 400bp sex specific *Sxl* enhancer

Location of SisB/Da binding site 1 to 6 are described in the map. The three mutant *Sxl* alleles created by CRISPR/Cas9 ($Sxl^{fB3\text{-}1}$, $Sxl^{fB3\text{-}2}$ and $Sxl^{fB3\text{-}3}$), their sequence changes and expected phenotypes are described.

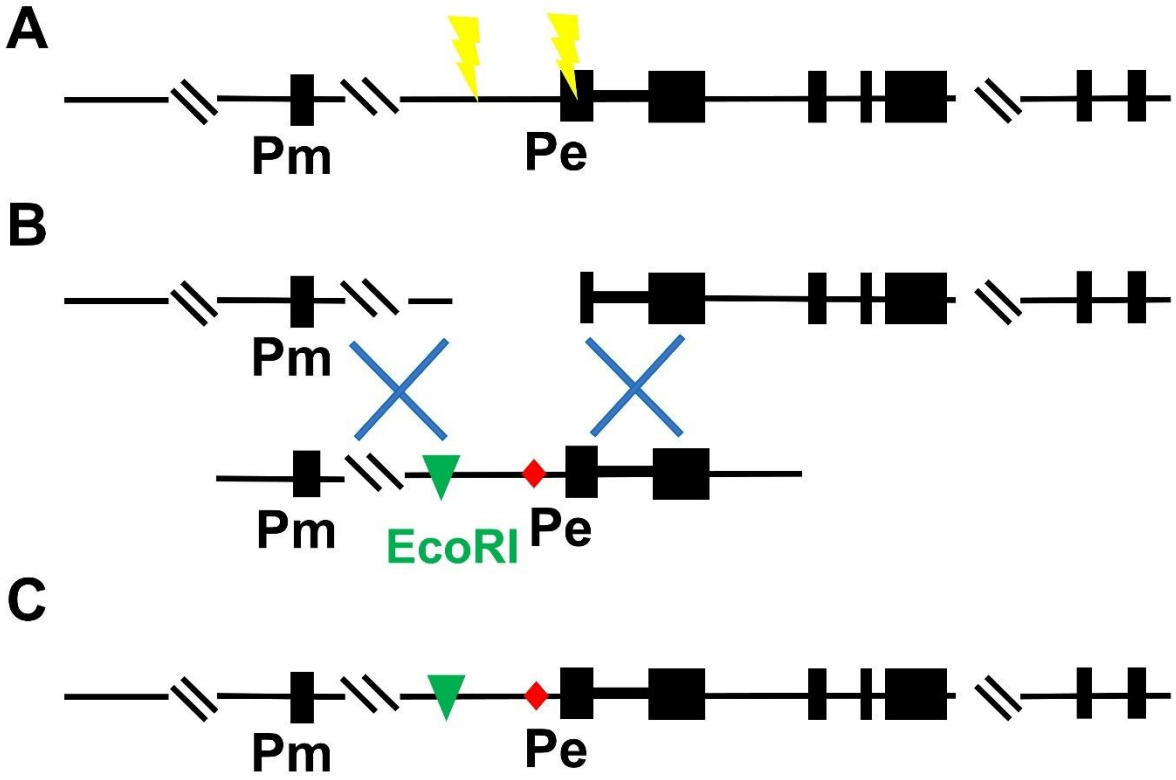

FIG. 3.2 Quick and efficient CRISPR/Cas9 mutagenesis of mutant *Sxl* alleles

(A) Two guide RNAs (gRNA) target endogenous *Sxl* loci. One targets -500bp, just upstream of the 400bp sex specific enhancer and the other targets 5'-UTR of *Sxl* exon E1. (B) Cas9 enzyme causes double stranded break to the two target sites, resulting deletion of the 400bp sex specific enhancer. Repair donor DNA oligo that contains mutant transcription factor binding site integrates by homologous recombination. (C) Endogenous mutation is successfully integrated and screened by EcoRI restriction enzyme digestion

Three different mutations at SisB/Da site 3 were engineered. $Sxl^{fB3\text{-}1}$ carried a change from the wildtype CATCTG to gtTCTG, that based on in vitro binding experiments (Yang et al.

2001), should eliminate or drastically reduce binding of the SisB/Da activator. A second mutant allele, $Sxl^{fB3\text{-}2}$ had CAgCTG mutation, creating a canonical SisB/Da binding site E-box, which was previously shown to bind more tightly to SisB/Da protein in vitro (Yang et al.2001). The third mutant allele, $Sxl^{fB3\text{-}3}$ had CAgaTG mutation that flips the orientation of the normal CATCTG sequence. In effect, testing whether the asymmetry present in the normal CATCTG E-box may be important for its function at *SxlPe*. Genetic complementation experiments showed significant viability reduction for $Sxl^{fB3\text{-}1}$ homozygous females (55.56%), suggesting a substantial defect in *SxlPe* expression (Table 3.1). Females homozygous for the other two alleles $Sxl^{fB3\text{-}2}$ and $Sxl^{fB3\text{-}3}$ were fully viable, and the male progeny of all 3 alleles were viable.

To see if $Sxl^{fB3\text{-}2}$ and $Sxl^{fB3\text{-}3}$ are mildly defective in *SxlPe* expression, a more stringent genetic complementation test was performed in which the mutant promoter alleles were present in once copy as the other homolog carried a deletion of the entire *Sxl* locus. Using this test, the defective allele, $Sxl^{fB3\text{-}1}$ was seen to have a dramatic effect on female viability as only 1 % of hemizgyous $Sxl^{fB3\text{-}1}$ survived (Table 3.1). Surprisingly, the other two mutant alleles showed mild defects in *Sxl* function when carried in single copy in Sxl^{f7bO} background. Female viability of $Sxl^{fB3\text{-}2}$ and $Sxl^{fB3\text{-}3}$ hemizygotes was approximately 80%. To exaggerate the apparent defects in the $Sxl^{fB3\text{-}2}$ and $Sxl^{fB3\text{-}3}$ alleles, XSE dose was reduced by introducing *sisB* and *sisA* mutations. Under these conditions, the *sisB sisA*/++ genotype reduced the viability of control females to 49%. The $Sxl^{fB3\text{-}1}/ Sxl^{+}$ heterozygous females were 100% lethal in the reduced XSE background offering strong evidence that sisB/Da site 3 is critical for *SxlPe* function. Viability of heterozygous $Sxl^{fB3\text{-}2}/Sxl^{+}$ and $Sxl^{fB3\text{-}3}/Sxl^{+}$ males were 18.05% and 15.96%, significantly lower than control female viability, again suggesting that both alleles with the canonical E-Box at site 3 and the reversed sisB/Da site 3 are somewhat defective in *SxlPe* expression.

Sxl genotype	Sxl^{+} CATCTG	$Sxl^{fB3\text{-}1}$ GTTCTG	$Sxl^{fB3\text{-}2}$ CAGCTG	$Sxl^{fB3\text{-}3}$ CAGATG
♀ *Sxl* / *Sxl*	120 (60)	55.56 (45)	100 (61)	100 (49)
♂ *Sxl*	105 (60)	106.67 (45)	98.36 (61)	93.88 (49)
♀ Sxl/Sxl^{f7bO}	104.30 (93)	1.01 (108)	78.67 (136)	78.99 (119)
♀ $\frac{+\ Sxl\ +}{sc^{sisB3}+\ sisA^{1}}$	48.58 (177)	0 (138)	18.05 (72)	15.96 (119)

TABLE 3.1 Genetic complementation tests of mutant $Sxl^{fB3\text{-}}$ alleles

Progenies from the cross ♀Sxl^{+}/*Binsinscy* x ♂Sxl^{+} and the cross ♀$Sxl^{fB3\text{-}}$/*Binsinscy* x ♂ $Sxl^{fB3\text{-}}$ were counted to measure the viability of males and females that are homozygous for each *Sxl* allele. Sisters that had *Binsinscy* balancer chromosome served a control for viability measurement. Progenies from the cross ♀*Sxl* /*Binsinscy* x ♂Sxl^{f7bO} and the cross ♀*Sxl* /*Binsinscy* x ♂ $sc^{sisB3}sis^{A1}$ were counted to measure the female viability in Sxl^{f7bO} and $sc^{sisB3}sis^{A1}$ background, using male siblings as viability reference.

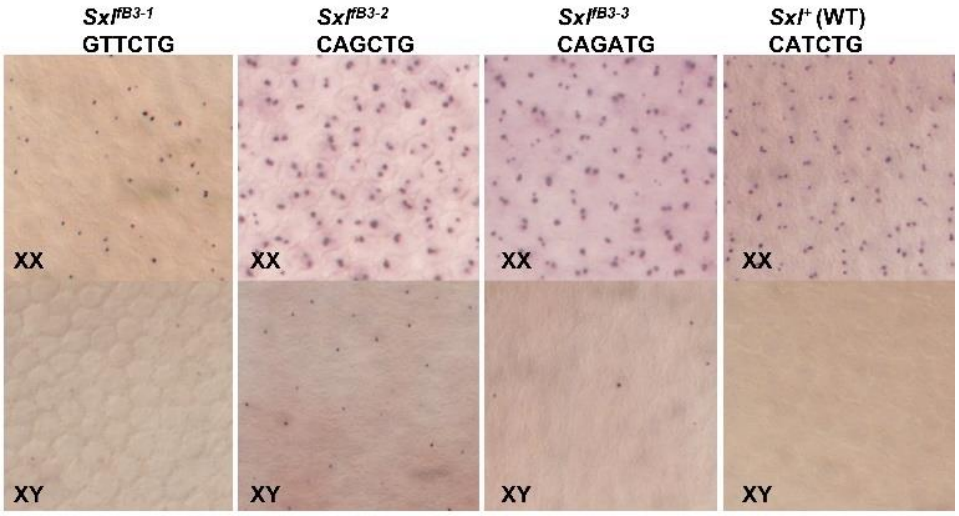

FIG. 3.3 *SxlPe* expression of Sxl^{fB3-} alleles

In situ hybridization images of nuclear cycle 13 embryos that are homozygous (XX) or hemizygous (XY) for each Sxl^{fB3-} mutant alleles and mutated sequence are shown. Brightness and contrasts are adjusted.

To more directly test if the site 3 mutations directly affected *SxlPe* activity, expression from three mutant alleles were observed by *in situ* hybridization. Homozygous Sxl^{fB3-1} female embryos showed non-uniform and patchy *SxlPe* expression, indicative of a defect in *SxlPe* activation. On the other hand, homozygous Sxl^{fB3-2} or Sxl^{fB3-3} embryos showed no significant alterations in *SxlPe* expression, consistent with the genetic experiments and expectations of in vitro binding experiments. (FIG. 3.1 and Table 3.1). Unexpectedly, we observed abnormal *SxlPe* expression from the Sxl^{fB3-2} and Sxl^{fB3-3} alleles in male embryos. The *Sxl* early transcripts from Sxl^{fB3-2} and Sxl^{fB3-3} males were clearly distinguished from occasional nuclear dots that may represent non-specific DNA-RNA binding. Together, our data showed that SisB/Da site 3 on its own has a

significant effect on *SxlPe* expression, suggesting that the activator site is particularly important for primary sex determination. The results showing that the mildly defective *Sxl* alleles $Sxl^{fB3\text{-}2}$ and $Sxl^{fB3\text{-}3}$ show ectopic expression in males is both curious and unexpected as both alleles carry what should be functional E-box sequences. The finding is somewhat paradoxical as the two alleles appear to be slightly defective in females but partially constitutive in males. The finding hints that binding affinity and binding site orientation may be subtly important for proper sisB/Da site 3 function. While it is difficult to offer a completely consistent explanation of the effects of the two mutations, with altered, but still functional, E-box sites, the ectopic expression in the high affinity CAGCTG E-box suggests that perhaps having less than maximal affinity for sisB/Da sites at *SxlPe* may be important for function. I note that none of 5 other sisB/Da sites in the 400 bp *SxlPe* regulatory region contain E-box sequences. An additional hint that affinity and binding site number comes from my analysis of the expression of a Llama-tagged variant of *Sxl* in the next section.

Addition of extra E-box sites induces constitutive *SxlPe* activation

Recently I engineered an endogenous *Sxl* allele that has the epitope tag *Llama* inserted at the N-terminus of exon E1 for the *Sxl* transactivation test (Chapter II). To see if *Llama-Sxl* complements *Sxl* function, a genetic test was performed over defective *Sxl* allele, Sxl^{f1}. The result showed that a single *Llama-Sxl* allele rescued females (*Llama* B/Da^{+}, 129%), but unexpectedly males seemed to be partially lethal (Table 3.2). (Because males served as the viability control female viability appeared higher than 100%.) A replicate genetic test with increased number of progeny also showed the male lethal effect of *Llama-Sxl* allele. The *Llama-Sxl* males were only

61% viable, suggesting that abnormal expression of *SxlPe* was responsible for the lethality. Analyzing the 351bp *Llama* sequence revealed two potential SisB/Da activator sites. One is a CAGCTG canonical E-box, located near the N-terminus of Llama and the other is a CAGATG E-box, located in the center of the Llama coding sequence. I hypothesized that these exogenous SisB/Da activator sites induced ectopic expression of *SxlPe* in male embryo. To test the hypothesis, two mutant *Llama-Sxl* alleles were engineered by CRISPR/Cas9. One allele had the N-terminal canonical CAGCTG site mutated to a sequence that should not bind SisB/Da (E-box 1^-). The other allele created non-binding mutations in both of the E-box sequences (E-box 1^-2^-). Genetic complementation tests showed that males with the original Llama were 82% viable in this experiment. Males were apparently fully viable with either one or both E-boxes mutated (129 % and 124%, Table 3.2), suggesting that removal of the E-box sequences reduced or eliminated ectopic *SxlPe* expression in male embryos. To further test whether there was ectopic *Sxl* expression in the Llama lines genetic tests were performed in two lines where male viability was sensitized by two different *sisB+* duplications, one on an autosome, the other on a Y chromosome. As expected, males with the original Llama allele (E-box$^+$) were 55-56% viable. Mutating E-box site 1 (E-box1$^-$) raised male viability to 72%. When both E-box sites were removed (E-box1^-2^-) male viability was rescued to 99% and 89 %, suggesting normal *Sxl* activity. Consistent with the genetic tests, male embryos with normal *Llama-Sxl* allele showed ectopic *SxlPe* expression when assayed for nascent *SxlPe* transcripts by in situ hybridization (Fig 3.2). The loss of one E-box reduced the *SxlPe* in males and the loss of both eliminated ectopic *SxlPe* expression in males.

	Llama **B/Da^{+}**	***Llama*** **B/Da1^{-}**	***Llama*** **B/Da1^{-}2^{-}**
♀$\frac{\boldsymbol{Llama-Sxl}}{sxl^{f1}}$	**128.57 (65)**	**NA**	**NA**
♂ ***Llama-Sxl***	**61.17 (103)**	**NA**	**NA**
♂ ***Llama-Sxl***	**82.35 (34)**	**129.27 (41)**	**124.14 (29)**
♂ ***Llama-Sxl*** **; Dp(*sisB*$^{+}$)/+**	**55.06 (89)**	**71.57 (102)**	**98.98 (98)**
♂ ***Llama-Sxl*/Y$_{sc+}$**	**55.88 (102)**	**72.36 (123)**	**89.39 (132)**

Table 3.2 Genetic complementation test of wildtype and mutant *Llama-Sxl* alleles

Progenies from the cross ♀*Llama-Sxl* x♂*Sxl^{f1}*, ♀*Llama-Sxl* x♂*w^{1118}*, ♀*Llama-Sxl* x♂*Llama-Sxl*, the cross ♀ *Llama-Sxl* x ♂*sisA^{1}*/Y;Dp(*sisB*$^{+}$) and the cross ♀ *Llama-Sxl* x ♂ +/Y$_{sc+}$ were counted for the viability tests. *Llama* B/Da^{+} denotes wildtype Llama tag sequence that has two SisB/Daughterless (B/Da) activator binding sites. *Llama* B/Da1^{-} or 1^{-}2^{-} denotes mutation in B/Da binding site 1 or both 1 and 2. The number of reference siblings is in parentheses.

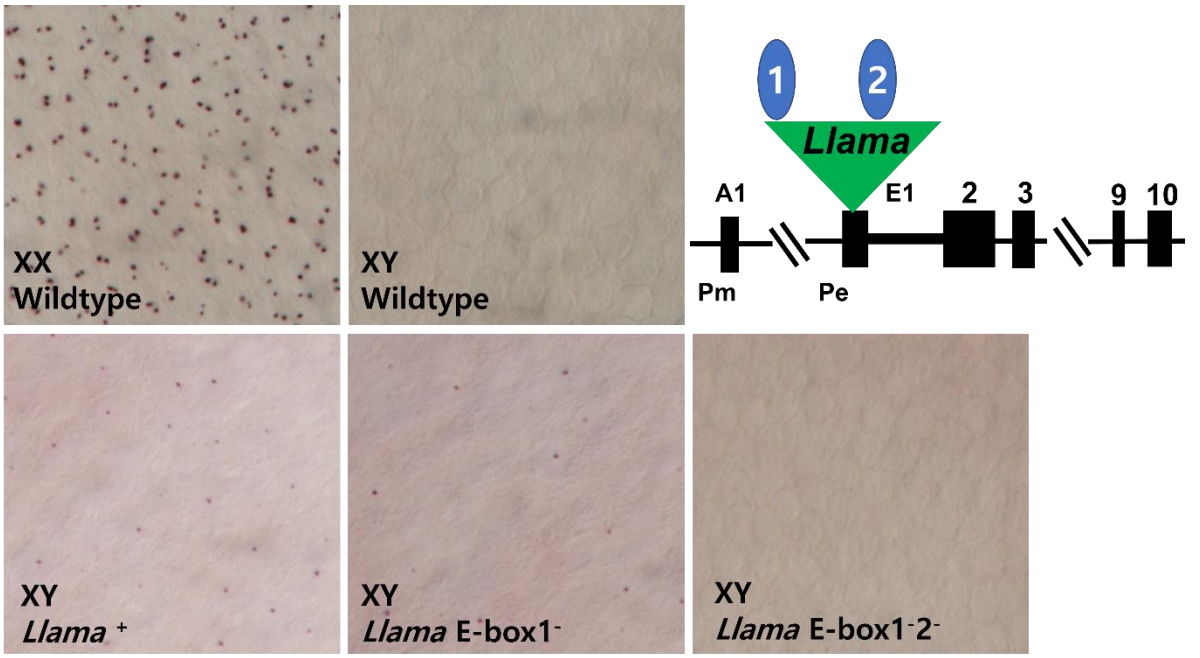

FIG. 3.4 *Llama-Sxl* early transcript expression

Nuclear cycle 13 embryos were observed under 40X brightfield microscopy. SisB/Daughterless activator binding sites in *Llama* sequence were described in the schematic.

CHAPTER IV CONCLUSIONS AND METHODS

Sexually dimorphic regulation of *Sex-lethal* depends on the ability of the *SxlPe* enhancer to distinguish the XSE dose from one X and two X chromosomes. Previous efforts directed to understanding sex-specific regulation of *SxlPe* relied on relatively small *SxlPe-lacZ* transgenes. Although *SxlPe-lacZ* transgenes were useful tools, experimental results from the transgenes were variable. A previous PhD student in our lab, Jayashre Rajendren, created full-length *Sxl* transgenes, that could be experimentally manipulated by recombineering technology, to try to more accurately mimic the endogenous *Sxl* locus. Although the full-length *Sxl* transgenes produced results that clarified, and in some cases corrected, earlier findings they too represent an artificial system that may be influenced the local chromatin environment they are inserted in. The transactivation phenomenon described in Chapter II appears to be an example of this. The rapid development of CRISPR/Cas9 technology to manipulate the Drosophila genome has rendered the transgene system somewhat obsolete. I decided to exploit CRISPR/Cas9 to manipulate the endogenous *Sxl* locus to test the roles of specific transcription factor binding sites in the regulation of *SxlPe*, the critical switch element promoter, that initiates female-specific expression of *Sxl*. In the process, I developed highly efficient Crispr/cas9 based tools to engineer changes throughout the *Sxl* locus.

Mutating repressor binding sites activates *SxlPe* constitutively

To determine the effect of mutating the non-canonical repressor site 3, Sxl^{MD3} was engineered. Sxl^{MD3} showed variable male-lethal effect in a range from 50% to 90%.

Staining *Sxl* nascent early transcripts showed ectopic *SxlPe* expression in Sxl^{MD3} male embryos. A striking feature of this finding is that a mutation in a single transcription factor binding site in an enhancer has a measurable effect on *Sxl* expression with real biological consequences. This is unusual in studies of eukaryotic regulation where multiple sites are typically inactivated before significant effects on transcription are observed.

The fact that the non-canonical repressor site 3 had significant effect on transcription repression does not imply it is the only, or the most important, repressor binding site in the *SxlPe* enhancer. A second mutant allele, Sxl^{MD123} had all three known repressor sites mutated. Sxl^{MD123} showed much a stronger male-lethal effect with less variability, and ectopic male *SxlPe* expression was more often observed. This argues that the individual repressor binding sites interact in a generally additive way to control *SxlPe* activity. My findings that repressor binding sites lead to elevated expression are in contradiction with a report that mutations affecting repressor site 1 and 2 had no effect in *SxlPe-lacZ* expression because these binding sites are shared by a positive bHLH regulator dMyc (Kappes et al.,2011). While formally possible that dMyc regulates *SxlPe* through the repressor sites, our lab has been unable to find genetic evidence for an effect of dMyc on *Sxl* using a range of dMyc alleles and sensitized genetic backgrounds. An unexplained aspect of my work is what accounts for the variability in the magnitude of the male-lethal effects I observed with the repressor site mutants. While X-chromosome counting process is known to

be highly variable in different genetic backgrounds (Cline, 1988) the variability I observed was in and between experiments using closely related lines.

One possible contribution to the variable magnitude of the male-lethal effects could be that the *SxlPe* enhancer also regulates *SxlPm* (Gonzalez et al., 2008). During nuclear cycle 13 to early 14 XX embryos express both *SxlPe* and *SxlPm* with expression of *SxlPm* depending in part on the *SxlPe* enhancer. XY embryos, in contrast, do not express *SxlPe* at all, and activate *SxlPm* later in cycle 14 than do females. The overlapping expression in females is thought to facilitate transition to autoregulatory splicing control whereas the delayed *SxlPm* expression in males may guard against occasional misfiring of *SxlPe* triggering stable splicing control. It may be the case that variation in the level or timing of *SxlPe* activation in Sxl^{MD3} and Sxl^{MD123} mutants compounds variation in the activation of *SxlPm* causing a variable number of them to stably engage *Sxl* splicing control. The less variable phenotype of Sxl^{MD123} is consistent with this notion.

Endogenous *Sxl* does not facilitate the activation of its homologous allele

A striking and unexpected aspect of our lab's studies with full-length *Sxl* transgenes located on the 2nd chromosome was the ability of constitutively active versions to activate a normal transgenic *Sxl* allele in male embryos (FIG. 2.1). While reminiscent and analogous to classical transvection, the ability to activate in trans did not seem to require chromosome pairing as the two alleles were visible far apart in the nuclei. We considered three general kinds of explanations for *Sxl* transactivation: First, that *Sxl* protein could possess a previously unknown transcriptional activation function amounting to a 2nd mode of *Sxl* autoregulation. Second, that transactivation is

a normal part of *Sxl* regulation facilitating rapid activation of both *Sxl* alleles, in effect creating a kind of cooperative response to XSE protein concentrations in females that could not occur in males. The third explanation is that *Sxl* transactivation is a product of the chromosomal site of insertion.

I was able to eliminate transcriptional autoregulation of *Sxl* by showing that a constitutive *Sxl* transgene was not capable of activating the endogenous *Sxl* locus (*Sxl* protein is known to act in trans in regulating splicing.). This finding also indicated that, whatever the mechanism, the alleles likely had to be in relatively close proximity. I chose to examine the possibility that transactivation was a normal part of *Sxl* regulation for several reasons. First, it was more biologically interesting than the alternative. Second, we had not proven the efficiency of transgene integration at alternative landing sites, and even if transgenes could be efficiently isolated in new locations, one could argue that the novel integration site disrupted the normal process. Third, that in manipulating the endogenous locus, I would create alleles that would be useful whether or not transactivation was a normal aspect of *Sxl* regulation.

As detailed in Chapter II, neither *SxlPe* expression patterns nor genetic complementation test provided evidence of transactivation, suggesting that the phenomenon seen with the *Sxl* transgenes may well have been caused by insertion into the attP40 landing site on the 2nd chromosome. In fact, a previous literature pointed out that the same attP40 landing as used in this study, caused transvection in their transgenic system (Mellert & Truman, 2012). Alternatively, it is also possible that the attB-P[acman]-CmR-BW vector could have caused transactivation. The *Sxl* transgenes were a useful and convenient tool to study fly sex determination, but my experimental results clearly highlight the importance of examining the endogenous *Sxl* locus,

especially when definitive experimental result and precise quantification of gene expression is required.

Quick and efficient CRISPR/Cas9 mutagenesis in endogenous *Sxl*

My initial CRISPR/Cas9 mutagenesis of repressor binding sites used a single CRISPR target site and a single stranded DNA oligo for homologous recombination and repair. The mutagenesis successfully produced mutant lines but the screening process took a lot of time because of its relatively low efficiency. Moreover, this method required the presence of appropriate CRISPR target sites very close to the site to be changed making it difficult to apply to other binding sites. To overcome these problems and engineer multiple mutant lines efficiently, I developed a CRISPR/Cas9 mutagenesis protocol to mutate any of the sequences in the female-specific *SxlPe* enhancer. My protocol depends on two efficient guide RNA target sites that flank the *SxlPe* enhancer generating a deletion of the enhancer that can be replaced by modified homologous sequences from a co-injected plasmid donor. I introduced an upstream EcoRI restriction site in the repair plasmid to allow for quick screening of positive mutant lines via PCR (FIG. 3.2). By preparing *Sxl* plasmids that contain multiple different mutations, and injecting several repair plasmids at the same time, different mutant lines were quickly and easily recovered from a single CRISPR injection. This efficient and fast CRISPR/Cas9 tool will allow us to mutate other transcription factor sites or to swap related enhancers from different species in our continuing effort to understand the operation of the *SxlPe* switch.

A single SisB/Da activator site has a predominate effect in *SxlPe* activation

Previous observation with the full length *Sxl* transgenes suggested that SisB/Da site 3 had a remarkably strong effect on *SxlPe* expression. Because this result was different from what had been reported with *SxlPe-lacZ* fusions and because it is unusual for a single transcription factor binding site mutation to strongly affect transcription in higher eukaryotes, I engineered mutant *Sxl* allele in endogenous loci to replicate the result. The $Sxl^{fB3\text{-}1}$ allele, which carries a mutation that prevents SisB/Da binding in vitro, showed a substantial reduction of *SxlPe* expression that resulted in decreased female viability. The SisB/Da site 3 is located in between adjacent binding sites for the STAT92E activator and the Zelda pioneering factor, which were shown to be important in *SxlPe* activation using *SxlPe-lacZ* fusions (Avila & Erickson, 2007) and in our previous transgene experiments. It may be that the particular importance of site 3 arises because SisB/Da activator binding to site 3 facilitates interactions with nearby STAT92E and Zelda activators. The other SisB/Da site 3 mutations, $Sxl^{fB3\text{-}2}$ and $Sxl^{fB3\text{-}3}$ were created to test whether binding site affinity or directionality might be important for proper enhancer function. The $Sxl^{fB3\text{-}2}$ mutation changes the site 3 sequence into a consensus binding site with a higher in vitro binding affinity (Yang et al.). The change has another difference in that its consensus CAGCTG binding site is symmetric, with two CAG half sites, whereas the normal site 3 CATCTG has asymmetric half sites. If the SisB/Da heterodimer has a preference in binding orientation the direction of the site could have importance in regulation. The results with $Sxl^{fB3\text{-}2}$ and $Sxl^{fB3\text{-}3}$ are difficult to interpret. Since they should bind SisB/Da, it was not surprising that both alleles drove enough *Sxl* expression to make viable females. Both, however, appeared somewhat defective under conditions that sensitize female viability to loss of *Sxl* function such as being hemizgous,

or when *sisB* and *sisA* doses are reduced (Table 3.1). This suggests that the particular wild type site 3 is important for proper function. Another indication that both the $Sxl^{fB3\text{-}2}$ and $Sxl^{fB3\text{-}3}$ alleles have altered function is that both are expressed ectopically in males, albeit at low levels (FIG. 3.3). Since the $Sxl^{fB3\text{-}2}$ changes the E-box to the high-affinity consensus sequence it's tempting to speculate that binding site affinity is important to *SxlPe* regulation, however, $Sxl^{fB3\text{-}3}$ carries the wild-type site 3 sequence (in inverted orientation) indicating that simple reliance on predicted binding affinities is insufficient. Of course, what may be critical is binding affinity in vivo and how that affects SisB/Da concomitant interactions with other regulatory proteins.

Extra copy of SisB/Da activator site induces constitutive *SxlPe* expression

Another indication that binding site number (and/or affinity) is important for proper *SxlPe* regulation was my finding that adding a *Llama*-tag sequence to exon E1 that contained two E-box sequences predicted to bind SisB/Da, caused an ectopic expression of *SxlPe* in males resulting is some male-lethality. Mutating one of the E-boxes (*Llama* E-box 1^{-}) partially rescued male-viability (and ectopic expression) while mutating both E-box sites (*Llama* E-box $1^{-}2^{-}$) completely rescued the male viability and eliminated ectopic expression, indicating that both likely bind SisB/Da in vivo. Together our observations imply that the current *SxlPe* regulation is the result of evolution keeping a balance between transcription repressors and activators to tightly regulate female-specific activation of *SxlPe*. The results with the Llama tag sequence caution that a regulatory sequence outside the known *Sxl* enhancer sequence could have important function for primary sex determination. Sequences, downstream of exon E1 have not been rigorously examined to see if they function in regulating *SxlPe*. I do note that only two

potential SisB/Da binding sequences map nearby in the downstream of *SxlPe*. One in exon A2, ~1,000 bp downstream and another in an intron ~2,400 bp downstream.

Ultimately an explanation of how sex specific regulation of *SxlPe* occurs will require the identification of both the locations and identities of the binding sites for the key regulatory proteins. Conventionally this is done with chromatin immunoprecipitation (CHIP) experiments, however, CHIP is not well suited to studying *SxlPe* for two reasons. The first, and less important, reason is the mixed population of males and females in any simple. More critical is the extremely narrow time window (~ 30-40 min) in which *SxlPe* is active. Since Drosophila egg laying cannot be synchronized obtaining a sufficient number of properly staged samples is a daunting task. An alternative method that can overcome these problems is the recently developed in vivo footprinting technology that exploits the ability of bound proteins to block in vitro methylation of DNA in isolated nuclei (Rao et al., 2021; Sonmezer et al., 2021). The technique allows high resolution of transcription factor binding (albeit it cannot identify the specific factors) from single DNA molecules and its sensitivity is such that it is amenable to the analysis of nuclei isolated from single, precisely staged, embryos from cycle 12 and beyond (Blythe & Wieschaus, 2016) —exactly the period needed to examine *SxlPe*. This can be done by recently developed in vivo footprinting technology and the efficient CRISPR/Cas9 mutagenesis protocol that I established will help characterize the *Sxl* regulatory elements.

Methods

Fly culture

Flies were grown in standard cornmeal and molasses medium at 25° in non-crowded condition. *Sxl*^{f1}, *Sxl*^{f9}, *Sxl*^{f7BO}, *sisB*$^{sc3\text{-}1}$, *sc*^{sisB3}, *sc*^{M6}, *sisA*^{1} and *w*^{1118} fly stocks were obtained from Bloomington Stock Center. *sc*^{M6} and *Sxl*^{f9} alleles were recombined to construct *sc*^{M6}*Sxl*^{f9} chromosome.

Molecular cloning of plasmids

Conventional PCR-based mutagenesis was performed to prepare normal and mutant *Sxl* plasmids. To prepare *Sxl* plasmid a forward primer 5'-CGGCATTTGCTGTATATTGTG-3' and a reverse primer 5'-TTTCTTTGCCAGTGACATCG-3' were used to PCR amplify the *Sxl* DNA using *w*^{1118} wildtype fly genomic DNA as template DNA. The PCR product is cloned into pCR II-TOPO TA vector by using TA cloning kit (ThermoFisher). To introduce DNA sequence change in circular DNA, the protocol below was performed. Acuprime Pfx DNA polymerase (Invitrogen) was used to synthesize PCR products for 18 cycles of denaturation, anneal and extension step. The PCR products were digested with DpnI restriction enzyme to degrade the template plasmids and gel purified. Gel purified PCR products were incubated with T4 PNK and T4 DNA ligase at 37° for 2 hours to re-circularize the products and directly transformed to competent *E. coli* DH5α strain. Following the site directed mutagenesis protocol above, the *Sxl* plasmid was mutated by two consecutive PCR-based mutagenesis to insert EcoRI restriction site for screen and a mutation in Exon E1, preventing undesirable digestion of the repair DNA sequence. The first PCR mutagenesis was performed with a forward primer 5'-

GaatTcGTATTGCTTCAGAAACGGATAACG-3’ and a reverse primer 5-AATACTATTTCCATATCTAATCTATTACTGTTAAATGTC-3’ to insert EcoRI restriction site for screen and followed by the second PCR mutagenesis with a forward primer 5’-tCCATaATGGATTTCAATTTTGATAC-3’ and a reverse primer 5’-ATAATCACAAAGTTACGTTTTTCAAC-3’ to engineer the mutation in exon E1.The PCR product was named, pYH44 and served as a template for creating mutant *Sxl* plasmids for CRISPR/Cas9 mutagenesis.

To introduce B/Da site 3 mutations a common forward primer 5’-TGCCTGCCTGCGAAGATC-3’ was used and three different reverse primers contained different mutations. The reverse primer sequences were 5’-GcTGTTCCGCATTCTCGGAA-3’, 5’-GAacTTCCGCATTCTCGGAA-3’ and 5’-tcTGTTCCGCATTCTCGGAA-3’ to create the B/Da site 3 mutations.To engineer B/Da site 2 mutations, a common reverse primer 5’-GCATTTCGCGGATCCCCGATTC-3’ and forward primers 5’- AGatGGCCACCCAAGAAAGTACGC-3’, 5’-AGCtGGCCACCCAAGAAAGTACGC-3’ and 5’-taCCGGCCACCCAAGAAAGTACG-3’ were used. To engineer STAT92E binding site mutation forward primer 5’-tCGAGAATGCGGAACATCTG-3’ and reverse primer 5’-AAGCTGGATGCCGGCAAGG-3’ were used, changing TTCCGAGAA normal sequence to TTtCGAGAA.

Llama sequence was amplified by a series of PCR reaction and flanked by *Sxl* sequences. Llama tag was inserted to N-termini of SXL in frame and 5’-TGT GAT TAT CCC ATT-3’ *Sxl* 5’ UTR sequence is mutated to 5’-TGT GAT TAT CCC <u>CAT</u>-3’ to include NdeI restriction site for screen. To engineer mutant *Llama-Sxl* alleles, normal *Llama-Sxl* plasmid was used as PCR template. To mutate B/Da binding site 1 in *Llama*, forward primer 5’-GTCGAGTCCGGTGGTGCCCT-3’ and reverse primer 5’-

tAGtTGAACCTGGGCCATATGGGG-3' were used for shifting CAGCTG sequence to CAaCTa. To mutate B/Da binding site 2 in *Llama*, forward primer 5'-ATGAACTCCCTGAAGCCCGAG-3' and reverse primer 5'-tTGCAAATAGACGGTGTTGCGG-3' were used for changing CAGATG normal sequence to CAaATG. Both B/Da site 1 and 2 mutations were designed to be silent mutations.

CRISPR/Cas9 mutagenesis and Screen

The overall CRISPR/Cas9 mutagenesis process is the same as described in Gratz et al. 2013. Bloomington 51323 stock that maternally expresses Cas9 enzyme was used for CRISPR injection. pU6-BbsI-chiRNA vector that contains gRNA scaffold was obtained from Addgene. The plasmid vector was digested with BbsI restriction enzyme and 3kb fragment was purified by Wizard SV Gel and PCR Clean-Up System (Promega). To mutate Dpn repressor site 3, sense oligo 5'-CTTCGACGCCTGGCACACTTCCTAG-3' and antisense oligo 5'-AAACCTAGGAAGTGTGCCAGGCGTC-3' were annealed by heating up in the boiling water bath for 5 minutes followed by gradual cool down in the water bath for an hour. The annealed double stranded oligo was ligated into the BbsI digested fragment of the pU6-BbsI-chiRNA vector and this served a guide RNA encoding plasmid. Single-stranded deoxyoligonucleotide (ssODN) was used as repair donor. The sequence of the ssODN was 5'- ATG CGG AAC ATC TGC CTG …. TGG CAA GCT TCC TAG ….. AGC CAC CGC CCA CTC GC -3' that spans a total of 200bp in *Sxl* enhancer sequence that mutates Dpn repressor binding site 3 CACACT sequence to CAAGCT sequence. The injection of the DNA was performed by BestGene Inc. Injected fly lines were tested by PCR amplification of *Sxl* fragment followed by HindIII restriction digestion at mutant Dpn repressor site 3 and positive allele, Sxl^{MD3} was recovered. To

mutate all the three Dpn repressor site 1, 2 and 3, Sxl^{MD3} was used for the injection stock. Two oligos, sense 5'–CTTCGTCTTAGGTAGCCCACGCGAC-3' and antisense 5'-AAACGTCGCGTGGGCTACCTAAGAC-3' were annealed and ligated into the BbsI digested pU6-BbsI-chiRNA vector as described above. ssODN sequence was 5'- GGC CAT CGA TCT ATT …. CAC GCT ACTGG CTA GCG …. GGA TCC CCG ATT CC-3', mutating Dpn repressor site 2 CAC GCG sequence to CAC GCT and site 1 CAC GCG sequence to CTA GCG. Positive mutant lines were recovered from Nhe I digestion at Dpn site 1 and named as Sxl^{MD123}. To engineer non-functional frameshift mutation in *Sxl* exon 5 Bloomington 51323 stock is used as injection stock. The two oligos 5'- CTTCGTACAATGATTTCCCCGGCTG-3' was annealed to another oligo 5'- AAACCAGCCGGGGAAATCATTGTAC-3' and ligated into BbsI digested pU6-BbsI-chiRNA vector. 198bp ssODN that has 5'- GCG AGT TGT TGC TGG …. GG GGC …. GAT TAA CCA TAC AAA-3' was used as repair donor, deleting a single C from GGC GGC sequence. Genetic complementation test was performed to screen and confirm non-functional frameshift mutation and the allele was formally named Sxl^{fD3}. The constitutive non-functional *Sxl* alleles were made by injecting the same gRNA plasmid and ssODN to mutate *Sxl* exon 5 using Sxl^{MD3} and Sxl^{MD123} as injection stock. Nonfunctional constitutive allele Sxl^{fMD3} was derived from Sxl^{MD3} and Sxl^{fMD123} was derived from Sxl^{MD123}, respectively. To engineer Llama tag inserted *Sxl* allele, 5'-CTTCGACTTTGTGATTATCCCATTA-3' oligo was annealed to 5'-AAACTAATGGGATAATCACAAAGTC-3' and ligated to BbsI digested pU6-BbsI-chiRNA vector. pBPhi-Llama-Hb was obtained from Hernan Garcia Lab (Bothma et al. 2018). Wildtype and mutant *Llama-Sxl* plasmids were injected with the gRNA encoding plasmid and the positive mutants were screened by NdeI restriction enzyme introduced in exon E1.

Multiple SisB/Da site mutant plasmids and the STAT site 1 mutant plasmid were

Injected together to obtain multiple different lines in a single CRISPR injection. The repair plasmids for each mutant binding site were prepared by PCR mutagenesis as described and they were co-injected with the two guide RNA encoding plasmids. One guide RNA encoding plasmid was engineered by annealing 5'- CTTCGATATGGAAATAGTATTGTGC-3' oligo and 5'-AAACGCACAATACTATTTCCATATC-3'and ligating the annealed double stranded oligo into BbsI digested gRNA scaffold plasmid. The other guide RNA encoding plasmid was engineered by annealing 5'- CTTCGACTTTGTGATTATCCCATTA-3' oligo and 5'-AAACTAATGGGATAATCACAAAGTC-3' and ligating the annealed oligo into BbsI digested gRNA scaffold plasmid

In situ hybridization

Embryos were collected 0-3.5 hours after egg laying at 25°. In-situ hybridization was performed as described (Lu et al. 2008). Staging embryo was performed by number and density of nuclei and cellular furrow that emerges at nuclear cycle 14 (Mahadevaraju et al. 2021 and Lu et al. 2008). NBT/BCIP stained embryos were mounted in 70% glycerol and fluorescent in situ hybridization embryos were mounted in Vectashield medium (Vector Laboratories). To synthesize RNA probes, DNA templates were PCR amplified by primer pairs and T7 promoter sequence was fused to the reverse primer for in vitro transcription of RNA probes with (MAXISCRIPT T7 kit, Ambion). Probe was detected by anti-digoxygenin antibody (Roche) through NBT-BCIP staining or by fluorescent antibodies. The fluorescent labeling of nascent *Sxl* transcripts used sheep anti-digoxygenin primary antibody (Roche) and Donkey anti-sheep Alexa 555 secondary antibody (Molecular probes). Primers used for in-vitro templates were *D. melanogaster Sxl* forward 5'-AACCCTGGATAAGCGACTGT-3', *D. melanogaster Sxl* reverse

5’-TAATACGACTCACTATAGGG-GGTGTAATGGTGGTTCTAC-3’. To design *Llama* sequence specific RNA probe, a forward primer 5’-AGGTTCAGCTGGTCGAGTC-3’ and a reverse primer 5’-TAATACGACTCACTATAGGG CGACGAGACAGTGACCTGAG-3’ were used. *Sxl* RNA probe synthesis was done by MAXISCRIPT T7 Kit (Ambion) using PCR products with T7 promoter sequence on one end.

Confocal microscopy imaging

In-situ hybridization embryos with fluorescent probes were observed using ZEISS. Maximum Z projection around 10 sections in 1 μm intervals were used to observe embryos using 20X objective lens and the images were processed by Image J. Brightness and contrasts were adjusted to optimize observing *Sxl* transcripts.

REFERENCES

Avila, F. W., & Erickson, J. W. (2007). Drosophila JAK/STAT pathway reveals distinct initiation and reinforcement steps in early transcription of Sxl. *Curr Biol*, *17*(7), 643-648. https://doi.org/10.1016/j.cub.2007.02.038

Barbash, D. A., & Cline, T. W. (1995). Genetic and molecular analysis of the autosomal component of the primary sex determination signal of Drosophila melanogaster. *Genetics*, *141*(4), 1451-1471. https://doi.org/10.1093/genetics/141.4.1451

Bashaw, G. J., & Baker, B. S. (1997). The regulation of the Drosophila msl-2 gene reveals a function for Sex-lethal in translational control. *Cell*, *89*(5), 789-798. https://doi.org/10.1016/s0092-8674(00)80262-7

Blythe, S. A., & Wieschaus, E. F. (2016). Establishment and maintenance of heritable chromatin structure during early Drosophila embryogenesis. *Elife*, *5*. https://doi.org/10.7554/eLife.20148

Bothma, J. P., Norstad, M. R., Alamos, S., & Garcia, H. G. (2018). LlamaTags: A Versatile Tool to Image Transcription Factor Dynamics in Live Embryos. *Cell*, *173*(7), 1810-1822 e1816. https://doi.org/10.1016/j.cell.2018.03.069

Burtis, K. C., & Baker, B. S. (1989). Drosophila doublesex gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. *Cell*, *56*(6), 997-1010. https://doi.org/10.1016/0092-8674(89)90633-8

Cline, T. W. (1976). A sex-specific, temperature-sensitive maternal effect of the daughterless mutation of Drosophila melanogaster. *Genetics*, *84*(4), 723-742. https://doi.org/10.1093/genetics/84.4.723

Cline, T. W. (1978). Two closely linked mutations in Drosophila melanogaster that are lethal to opposite sexes and interact with daughterless. *Genetics*, *90*(4), 683-698. https://doi.org/10.1093/genetics/90.4.683

Cline, T. W. (1979). A male-specific lethal mutation in Drosophila melanogaster that transforms sex. *Dev Biol*, *72*(2), 266-275. https://doi.org/10.1016/0012-1606(79)90117-9

Cline, T. W. (1984). Autoregulatory functioning of a Drosophila gene product that establish es and maintains the sexually determined state. *Genetics*, *107*(2), 231-277. https://doi.org/10.1093/genetics/107.2.231

Cline, T. W. (1988). Evidence that sisterless-a and sisterless-b are two of several discrete "numerator elements" of the X/A sex determination signal in Drosophila that switch Sxl between two alternative stable expression states. *Genetics*, *119*(4), 829-862. https://doi.org/10.1093/genetics/119.4.829

Cline, T. W., & Meyer, B. J. (1996). Vive la difference: males vs females in flies vs worms. *Annu Rev Genet*, *30*, 637-702. https://doi.org/10.1146/annurev.genet.30.1.637

Deshpande, G., Stukey, J., & Schedl, P. (1995). scute (sis-b) function in Drosophila sex determination. *Mol Cell Biol*, *15*(8), 4430-4440. https://doi.org/10.1128/MCB.15.8.4430

Duffy, J. B., & Gergen, J. P. (1991). The Drosophila segmentation gene runt acts as a position-specific numerator element necessary for the uniform expression of the sex-determining gene Sex-lethal. *Genes Dev*, *5*(12A), 2176-2187. https://doi.org/10.1101/gad.5.12a.2176

Erickson, J. W. (2016). Primary Sex Determination in Drosophila melanogaster Does Not Rely on the Male-Specific Lethal Complex. *Genetics*, *202*(2), 541-549. https://doi.org/10.1534/genetics.115.182931

Erickson, J. W., & Cline, T. W. (1993). A bZIP protein, sisterless-a, collaborates with bHLH transcription factors early in Drosophila development to determine sex. *Genes Dev*, *7*(9), 1688-1702. https://doi.org/10.1101/gad.7.9.1688

Erickson, J. W., & Quintero, J. J. (2007). Indirect effects of ploidy suggest X chromosome dose, not the X:A ratio, signals sex in Drosophila. *PLoS Biol*, *5*(12), e332. https://doi.org/10.1371/journal.pbio.0050332

Estes, P. A., Keyes, L. N., & Schedl, P. (1995). Multiple response elements in the Sex-lethal early promoter ensure its female-specific expression pattern. *Mol Cell Biol*, *15*(2), 904-917. https://doi.org/10.1128/MCB.15.2.904

Flickinger, T. W., & Salz, H. K. (1994). The Drosophila sex determination gene snf encodes a nuclear protein with sequence and functional similarity to the mammalian U1A snRNP protein. *Genes Dev*, *8*(8), 914-925. https://doi.org/10.1101/gad.8.8.914

Gelbart, M. E., & Kuroda, M. I. (2009). Drosophila dosage compensation: a complex voyage to the X chromosome. *Development*, *136*(9), 1399-1410. https://doi.org/10.1242/dev.029645

Gergen, J. P. (1987). Dosage Compensation in Drosophila: Evidence That daughterless and Sex-lethal Control X Chromosome Activity at the Blastoderm Stage of Embryogenesis. *Genetics*, *117*(3), 477-485. https://doi.org/10.1093/genetics/117.3.477

Gladstein, N., McKeon, M. N., & Horabin, J. I. (2010). Requirement of male-specific dosage compensation in Drosophila females--implications of early X chromosome gene expression. *PLoS Genet*, *6*(7), e1001041. https://doi.org/10.1371/journal.pgen.1001041

Golling, G., Li, L., Pepling, M., Stebbins, M., & Gergen, J. P. (1996). Drosophila homologs of the proto-oncogene product PEBP2/CBF beta regulate the DNA-binding properties of Runt. *Mol Cell Biol*, *16*(3), 932-942. https://doi.org/10.1128/MCB.16.3.932

Gonzalez, A. N., Lu, H., & Erickson, J. W. (2008). A shared enhancer controls a temporal switch between promoters during Drosophila primary sex determination. *Proc Natl Acad Sci U S A*, *105*(47), 18436-18441. https://doi.org/10.1073/pnas.0805993105

Heinrichs, V., Ryner, L. C., & Baker, B. S. (1998). Regulation of sex-specific selection of fruitless 5' splice sites by transformer and transformer-2. *Mol Cell Biol*, *18*(1), 450-458. https://doi.org/10.1128/MCB.18.1.450

Horabin, J. I., & Schedl, P. (1993). Regulated splicing of the Drosophila sex-lethal male exon involves a blockage mechanism. *Mol Cell Biol*, *13*(3), 1408-1414. https://doi.org/10.1128/mcb.13.3.1408-1414.1993

Hoshijima, K., Kohyama, A., Watakabe, I., Inoue, K., Sakamoto, H., & Shimura, Y. (1995). Transcriptional regulation of the Sex-lethal gene by helix-loop-helix proteins. *Nucleic Acids Res*, *23*(17), 3441-3448. https://doi.org/10.1093/nar/23.17.3441

Jinks, T. M., Polydorides, A. D., Calhoun, G., & Schedl, P. (2000). The JAK/STAT signaling pathway is required for the initial choice of sexual identity in Drosophila melanogaster. *Mol Cell*, *5*(3), 581-587. https://doi.org/10.1016/s1097-2765(00)80451-7

Kaminker, J. S., Singh, R., Lebestky, T., Yan, H., & Banerjee, U. (2001). Redundant function of Runt Domain binding partners, Big brother and Brother, during Drosophila development. *Development*, *128*(14), 2639-2648. https://www.ncbi.nlm.nih.gov/pubmed/11526071

Kappes, G., Deshpande, G., Mulvey, B. B., Horabin, J. I., & Schedl, P. (2011). The Drosophila Myc gene, diminutive, is a positive regulator of the Sex-lethal establishment promoter, Sxl-Pe. *Proc Natl Acad Sci U S A*, *108*(4), 1543-1548. https://doi.org/10.1073/pnas.1017006108

King, T. D., Johnson, J. E., & Bateman, J. R. (2019). Position Effects Influence Transvection in Drosophila melanogaster. *Genetics*, *213*(4), 1289-1299. https://doi.org/10.1534/genetics.119.302583

Lallena, M. J., Chalmers, K. J., Llamazares, S., Lamond, A. I., & Valcarcel, J. (2002). Splicing regulation at the second catalytic step by Sex-lethal involves 3' splice site recognition by SPF45. *Cell*, *109*(3), 285-296. https://doi.org/10.1016/s0092-8674(02)00730-4

Lee, A. M., & Wu, C. T. (2006). Enhancer-promoter communication at the yellow gene of Drosophila melanogaster: diverse promoters participate in and regulate trans interactions. *Genetics*, *174*(4), 1867-1880. https://doi.org/10.1534/genetics.106.064121

LEWIS, E. B. (1954). THE THEORY AND APPLICATION OF A NEW METHOD OF DETECTING CHROMOSOMAL REARRANGEMENTS IN DROSOPHILA MELANOGASTER. *The American Naturalist*, *LXXXVIII*(841).

Lu, H., Kozhina, E., Mahadevaraju, S., Yang, D., Avila, F. W., & Erickson, J. W. (2008). Maternal Groucho and bHLH repressors amplify the dose-sensitive X chromosome signal in Drosophila sex determination. *Dev Biol*, *323*(2), 248-260. https://doi.org/10.1016/j.ydbio.2008.08.012

Mahadeveraju, S., Jung, Y. H., & Erickson, J. W. (2020). Evidence That Runt Acts as a Counter-Repressor of Groucho During Drosophila melanogaster Primary Sex Determination. *G3 (Bethesda)*, *10*(7), 2487-2496. https://doi.org/10.1534/g3.120.401384

Nagengast, A. A., Stitzinger, S. M., Tseng, C. H., Mount, S. M., & Salz, H. K. (2003). Sex-lethal splicing autoregulation in vivo: interactions between SEX-LETHAL, the U1 snRNP and U2AF underlie male exon skipping. *Development*, *130*(3), 463-471. https://doi.org/10.1242/dev.00274

Nusslein-Volhard, C., & Wieschaus, E. (1980). Mutations affecting segment number and polarity in Drosophila. *Nature*, *287*(5785), 795-801. https://doi.org/10.1038/287795a0

Paroush, Z., Finley, R. L., Jr., Kidd, T., Wainwright, S. M., Ingham, P. W., Brent, R., & Ish-Horowicz, D. (1994). Groucho is required for Drosophila neurogenesis, segmentation, and sex determination and interacts directly with hairy-related bHLH proteins. *Cell*, *79*(5), 805-815. https://doi.org/10.1016/0092-8674(94)90070-1

Rabinow, L., & Samson, M. L. (2010). The role of the Drosophila LAMMER protein kinase DOA in somatic sex determination. *J Genet*, *89*(3), 271-277. https://doi.org/10.1007/s12041-010-0038-6

Rao, S., Ahmad, K., & Ramachandran, S. (2021). Cooperative binding between distant transcription factors is a hallmark of active enhancers. *Mol Cell*, *81*(8), 1651-1665 e1654. https://doi.org/10.1016/j.molcel.2021.02.014

Rideout, E. J., Billeter, J. C., & Goodwin, S. F. (2007). The sex-determination genes fruitless and doublesex specify a neural substrate required for courtship song. *Curr Biol*, *17*(17), 1473-1478. https://doi.org/10.1016/j.cub.2007.07.047

Rideout, E. J., Dornan, A. J., Neville, M. C., Eadie, S., & Goodwin, S. F. (2010). Control of sexual differentiation and behavior by the doublesex gene in Drosophila melanogaster. *Nat Neurosci*, *13*(4), 458-466. https://doi.org/10.1038/nn.2515

Salz, H. K., & Erickson, J. W. (2010). Sex determination in Drosophila: The view from the top. *Fly (Austin)*, *4*(1), 60-70. https://doi.org/10.4161/fly.4.1.11277

Samuels, M. E., Schedl, P., & Cline, T. W. (1991). The complex set of late transcripts from the Drosophila sex determination gene sex-lethal encodes multiple related polypeptides. *Mol Cell Biol*, *11*(7), 3584-3602. https://doi.org/10.1128/mcb.11.7.3584-3602.1991

Sanchez, L., Granadino, B., & Torres, M. (1994). Sex determination in Drosophila melanogaster: X-linked genes involved in the initial step of sex-lethal activation. *Dev Genet*, *15*(3), 251-264. https://doi.org/10.1002/dvg.1020150307

Sefton, L., Timmer, J. R., Zhang, Y., Beranger, F., & Cline, T. W. (2000). An extracellular activator of the Drosophila JAK/STAT pathway is a sex-determination signal element. *Nature*, *405*(6789), 970-973. https://doi.org/10.1038/35016119

Sonmezer, C., Kleinendorst, R., Imanci, D., Barzaghi, G., Villacorta, L., Schubeler, D., Benes, V., Molina, N., & Krebs, A. R. (2021). Molecular Co-occupancy Identifies Transcription Factor Binding Cooperativity In Vivo. *Mol Cell*, *81*(2), 255-267 e256. https://doi.org/10.1016/j.molcel.2020.11.015

Valcarcel, J., Singh, R., Zamore, P. D., & Green, M. R. (1993). The protein Sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of transformer pre-mRNA. *Nature*, *362*(6416), 171-175. https://doi.org/10.1038/362171a0

Yang, D., Lu, H., Hong, Y., Jinks, T. M., Estes, P. A., & Erickson, J. W. (2001). Interpretation of X chromosome dose at Sex-lethal requires non-E-box sites for the basic helix-loop-helix proteins SISB and daughterless. *Mol Cell Biol*, *21*(5), 1581-1592. https://doi.org/10.1128/MCB.21.5.1581-1592.2001

APPENDIX A

Identifying an unknown maternal factor regulating sex determination

When I joined the Erickson lab, my initial project was to try to identify a maternally deposited positive regulator of *Sxl* that maps on the 3rd chromosome (Erickson, 2016). The regulator, referred to here as, X-7, was identified as being responsible for a recessive female-lethal maternal effect interaction with XSE mutants that had been misidentified as being caused by a mutation in the dosage compensation gene, *msl-3* (Gladstein et al., 2010).

To identify the novel maternal *Sxl* regulator, I performed deletion mapping. Because the X-7 mutation is recessive, it does not show its female-lethal phenotype when paired with a normal allele. However, maternal-effect female-lethality should be revealed when the X-7 allele paired with deletions that remove the regulator (FIG. A1). To carry out the screen deletion deletion-bearing X-7 heterozygous females were crossed to males carrying the hypomorphic $sisA^{1}$ allele and the early defective Sxl^{f9} allele to reveal the maternal effect.

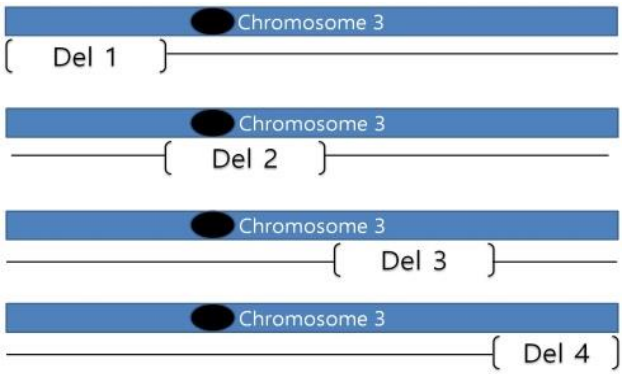

FIG. A1. Deletion mapping strategy to identify maternal *Sxl* regulator

Because the unidentified maternal factor mutation (black oval) was recessive, its female-lethal effect can only be observed if the unidentified mutation is uncovered by an overlapping deletion in the homologous chromosome pair (Del 2 in this example). Any deletion showing a recessive female lethal effect was further analyzed using smaller deletions in the region to narrow down and to attempt to identify the unknown maternal regulator.

My initial round of screening showed a positive result for large deletion (Df(3R)BSC874, 98E1-99A1) and I then used smaller deletions to narrow down the region encoding the maternal regulator. My deletion mapping suggested that the maternal regulator mapped near *darkener of apricot (doa)* gene, which genetically interacts with the zygotic Sex determination hierarchy later in development (Rabinow & Samson, 2010). The next round of deletion mapping showed strong female-lethal interaction with the two deletions *Exel*6211 and *BSC*806, excluding the possibility that *doa* is the maternal *Sxl* regulator (Table A1 and FIG. A2).

Deletions	*Df(3R)Exel*6211	*Df(3R)BSC*806	*Df(3R)BSC*789	*Df(3R)Exel*6210	$doa^{MI14992}$	doa^{01705b}
% ♀VIABILITY	22 (65)	0 (71)	100 (73)	83 (81)	86 (142)	109 (93)

TABLE A1. Deletion mapping to identify the maternal *Sxl* regulator, X-7.

Progeny from the crosses: ♀♀ X-7/Deletion x ♂♂ $sisA^{1}Sxl^{f9}$ were counted. X-7 represents the unidentified regulator mutation. Percentage female viability was measured relative to the viability of reference male siblings (parentheses). $doa^{MI14992}$ and doa^{01705b} are point mutation alleles of *doa*.

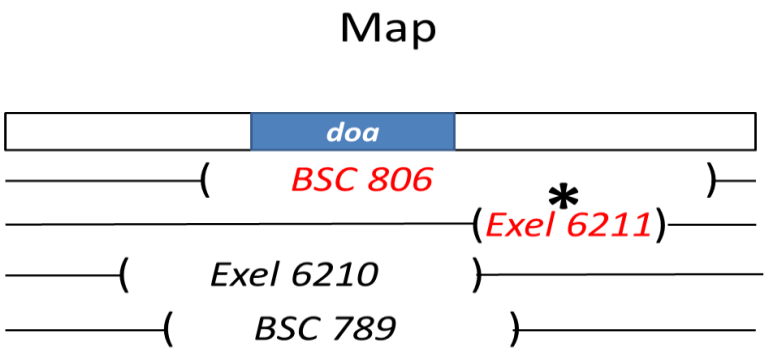

FIG. A2. Deletion mapping of maternal *Sxl* regulator, X-7 excluded *doa* as the candidate

BSC806 and *Exel*6211 showed strong lethal effect and the maternal *Sxl* regulator is likely located in the *Exel*6211 deleted loci.

Exel6211 deletion, which was positive for the lethal interaction in the deletion mapping, encompasses 6 known genes. The same mapping strategy was applied with point mutations, or RNAi lines available for the candidate genes: *sirt7*^{f07159} (viable, fertile allele), *cul5*^{Y00051} (lethal allele), *Cpsf100*^{f00376} (lethal allele), *Pkc98E*^{f06221} (lethal allele), *CG11837* and CG 11873 (Trip RNAi lines). None of the candidate alleles showed positive results when paired with, or when knocked down in combination with, the X-7 mutation. While my experiments likely excluded *cul5*, *CPSF100*, and *Pkc98E* from consideration, the results with the other tested genes must be considered as inconclusive given that we had no evidence that the mutations or RNAi lines had any effect on those genes. At that point, given that there was no easily identifiable way to make further progress, and no certainty that the X-7 gene would have an interesting or compelling link to *SxlPe* regulation, I switched to work on the other projects recorded in this thesis.

I do note that updates to the genomic annotations for this region made since I ended this project have identified a plausible candidate gene, *Slu7*. *Slu7* encodes a component of the spliceosome that is involved in joining the upstream and downstream splice sites and in lariat release. Its mammalian homolog has been suggested to be involved in alternative splicing. No useful *Slu7* alleles are available at this time, but the protein's function suggests that it could be involved in the unknown mechanism by which exon E1 is joined to exon 4--while skipping exons A2 and 3--that is required in order to express the early *Sxl* proteins that initiate autoregulatory splicing.

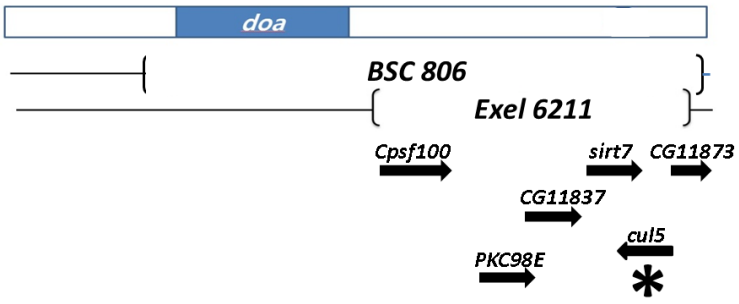

FIG. A3. Point mutations of the 6 candidate genes tested for lethal interaction

APPENDIX B

Analysis of the functions of the CBF-beta proteins Bro and Bgb in sex determination

The XSE *runt* has diverse functions in development including segmentation, neuronal, and eye development and *Sxl* activation. Runt-class (Runx) transcription factor bind DNA in vitro, but appear to require dimerization partners, known as a CBF-beta proteins, to bind DNA with high affinity and to carry out their in vivo functions (Mahadeveraju et al., 2020). There are two CBF-β proteins in Drosophila, encoded on the 3rd chromosome by the *Brother (Bro)* and *Bigbrother (Bgb)* genes. *Bro* and *Bgb* are considered to be partially redundant (Kaminker et al., 2001). Bro is exclusively zygotic and Bgb is exclusively maternal in the early embryo when the *SxlPe* promoter is active. Previous work in our lab showed that *Bro* deletions are viable and fertile but that *Bgb* is essential for both viability and germ line development. Complete loss of Bro causes no noticeable defects in *Sxl* expression nor does loss of one maternal copy of *Bgb* (Two copies of *Bgb* are required for germline development). To determine if *bro* and *bgb* have partially redundant functions in *SxlPe* activation, or if they might be dispensable for Runt's XSE function at *SxlPe*, I analyzed *SxlPe* expression by *in situ* hybridization in embryos with a variety of defects in *Bro*, and/or *Bgb* and sensitized by a reduction in *runt* gene dose (FIG. B1, and data not shown). Expression of *SxlPe* was normal in female embryos with one copy *runt* and *Bro* and a normal maternal Bgb contribution (FIG. B1, panel A). Mild *SxlPe* expression defects were seen in many embryos lacking one copy of *runt* and all zygotic *Bro* but that received a normal maternal Bgb contribution. Overall staining was almost normal (the difference in intensity of purple color is not meaningful) but small regions with non-expressing nuclei were observed

(Panel B). The *Sxl* expression defects had no discernable effect on viability as females and males were equally viable from these crosses (data not shown). Somewhat more severe defects were observed when the maternal Bgb contribution was halved and zygotic *runt* and *bro* doses reduced. Under these conditions the areas of reduced *SxlPe* expression were both increased in number and in size (panel C). These defects seem likely to have had an impact on female survival as males were somewhat more viable than their sisters in these crosses (data not shown.). Much more severe impacts on *SxlPe* activity were seen in females with reduced *runt* dose when both the maternal Bgb contribution was halved and zygotic Bro function was eliminated (panel D). Overall staining intensity was reduced and large sections of the embryos were devoid of *Sxl* nascent transcripts. No females were recovered from these crosses in any of our experiments, but the impact of the sex determination defect was impossible to quantify because most male progeny also died because of the pleiotropic effects of the CBF-beta genes in development. The simplest interpretation of these data is that both maternally supplied Bgb, and zygotically produced Bro, interact in a partially redundant manner with Runt to regulate *SxlPe*. The zygotic contribution of Bro is dispensable so long as sufficient maternal Bgb is present but Bgb is necessary for proper *SxlPe* expression when maternal Bgb and zygotic *runt* doses are reduced. Because Bgb^{-} germline clones are not viable we were unable to assess whether a full complement of zygotic Bro can substitute for a complete loss of maternal Bgb.

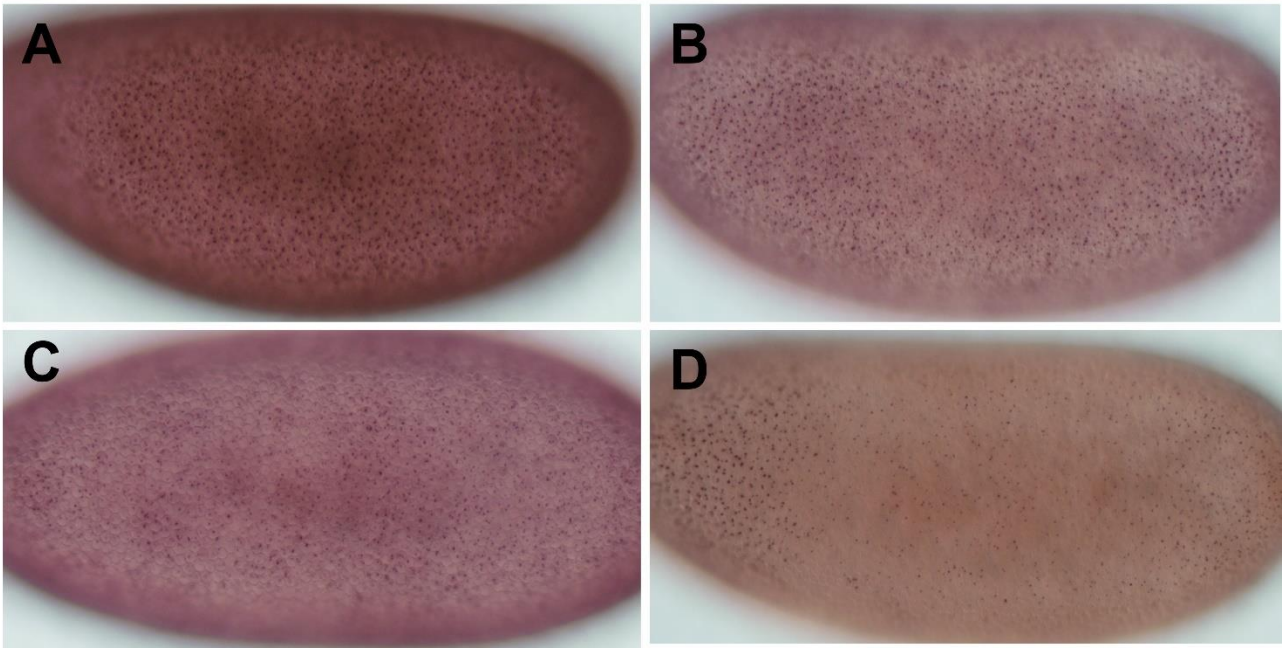

FIG. B1. *runt* and *Bro/Bgb* have genetic interactions in *SxlPe* expression.

SxlPe expression of nuclear cycle 13 female embryos are observed by in situ hybridization. No defects in *SxlPe* expression are seen in homozygous *ΔBro* mutants derived from homozygous *ΔBro* mothers, or *Bgb ΔBro* /+ *ΔBro* mothers. (A) Normal *SxlPe* expression in female $runt^3$/+ *ΔBro*/+ embryo. Cross ♀♀*ΔBro* x ♂♂ $runt^+$ /$Yrunt^+$. (B) Mild defect in *SxlPe* expression in $runt^3$/+ female lacking all *Bro* function. Cross: ♀♀*ΔBro* x ♂♂ $runt^3$ /$Yrunt^+$; *ΔBro*. (C) Mild to moderate *SxlPe* defect in $runt^3$/+ female embryo with reduced maternal Bgb protein and reduced zygotic *Bro*. Cross: ♀♀ Bgb^-*ΔBro* /+ *ΔBro* x ♂♂ $runt^3$ /$Yrunt^+$. (D) Strong *SxlPe* defect in $runt^3$/+ female embryo with reduced maternal Bgb and no zygotic *Bro*. Cross: ♀♀ Bgb^- *ΔBro* /+ *ΔBro* x ♂♂ $runt^3$ /$Yrunt^+$; *ΔBro.*

APPENDIX C

Live imaging of SXL using *Llama* tag

For the endogenous *Sxl* transactivation test, a normal *Sxl* allele that has an exogenous epitope tag was needed to design sequence specific RNA probes. Currently, commercially available monoclonal SXL antibodies are not sensitive enough to answer the question of when SXL is first expressed. Because of the need to observe early SXL derived from *SxlPe*, I decided to insert an epitope tag for live imaging of SXL. Initially, I considered conventional eGFP, however, eGFP takes more than 40 minutes to fold and its turnover rate is fast, making eGFP unsuitable given the short time window in which primary sex determination occurs. The recently developed *Llama* tag method (Bothma et al., 2018) solves these issues, so I decided to engineer a *Llama-Sxl* variant with tag inserted in exon E1 to tag the N-terminus of the early *Sxl* protein. The core idea of *Llama* tag live imaging is to load maternally deposited cytoplasmic GFP protein (or mRNA) into the fly embryo. GFP then has sufficient time to be translated and to fold by the time *Sxl* is first expressed. The *Llama* tag itself consists of an anti-GFP nanobody. Once Llama-SXL is expressed, a nuclear localized fusion protein recruits cytoplasmic GFP into nuclei, resulting in a large increasing in fluorescent signal intensity in the nuclei. Using the modified Llama tag line that lacks the two E-boxes that cause ectopic *SxlPe* activity (Chapter III), I observed nuclear GFP in mid nuclear cycle 14 (>30min) and the signal was detectable at least until the germband extension stage. No signal was detected in nuclear cycle 13 embryos or in cycle 14 males. It is important to note that this does not exclude the possibility of earlier SXL expression because a single Llama tag segment may not be sensitive enough to pick low-level expression of SXL.

Indeed, in my experiments with the original, partially constitutive *Llama-Sxl* (E-box^{+}), I observed nuclear GFP in cycle 13 female embryos (Fig. C2.). In situ-hybridization experiments to identify nascent Llama-Sxl transcripts revealed that the *Llama-Sxl* (E-box^{+}) line was expressed at low-levels in cycle 13 male embryos and, possibly, at higher levels in cycle 13 females. I found no evidence that the *Llama-Sxl* (E-box^{+}) line was expressed prior to cycle 12, that cycle when transcription begins from the normal *Sxl* locus. These results demonstrate that there is sufficient time to transcribe and translate *Sxl* during cycle 13 suggesting that our failure to observe nuclear GFP prior to cycle 14 with the E-box 1^{-}2^{-} *Llama-Sxl* variant may be due to a comparatively low signal to noise ratio. One approach to address this would be to engineer multiple tandem Llama tags to the N-terminus of SXL to increase the amount of GFP that can be delivered to the nuclei.

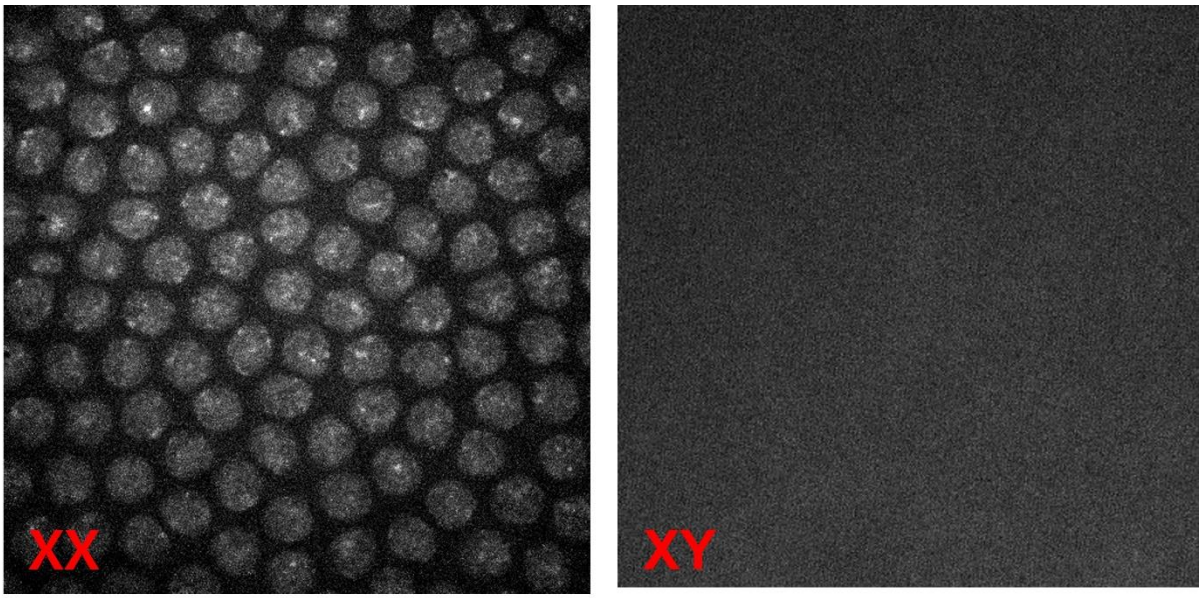

FIG. C1. Llama-SXL live imaging with E-box $1^{-}2^{-}$ Llama tag.

Living nuclear cycle 14 XX embryos and XY embryos were observed by confocal microscopy image. I have been unable to detect nuclear GFP prior to cycle 14 using this variant.

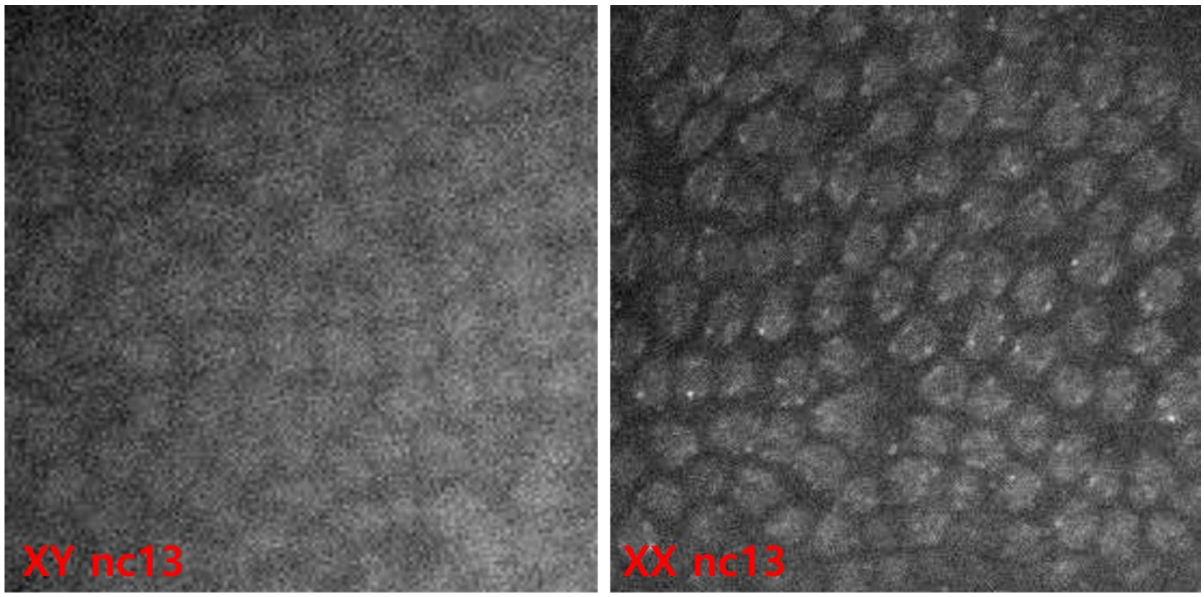

FIG. C2. Llama-SXL live imaging with original, E-box^{+}) Llama tag.

Living nuclear cycle 13 XX embryos and XY embryos were observed by confocal microscopy. As measured by in situ hybridization of nascent transcripts, the original *Llama-Sxl* line (E-box^{+}) is expressed at low levels in male cycle 13 embryos and may be more strongly expressed in female embryos in cycles 12-14. We found no evidence that *Llama-Sxl* line (E-box^{+}) is expressed earlier than wild-type *Sxl* in female embryos.