UNIVERSITY OF CALIFORNIA
RIVERSIDE

Mobilization of Newly Identified Transposon *Muta1* in *Aedes aegypti* and *Drosophila melanogaster*

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Biochemistry and Molecular Biology

by

Presha Vijaykumar Shah

June 2015

Dissertation Committee:
Dr. Peter W. Atkinson, Chairperson
Dr. Jason Stajich
Dr. Jikui Song

The Dissertation of Presha Vijaykumar Shah is approved:

Chairperson

University of California, Riverside.

ACKNOWLEDGEMENTS

I would like to begin thanking my advisor Dr. Peter Atkinson, for granting me the opportunity to achieve my dream to obtain a doctorate degree. I am very thankful for his tireless efforts in guiding, and supporting this research. Apart from serving as my academic advisor, Dr. Peter Atkinson has also nurtured and mentored me to develop skills as scientist and critically by allowing me to design projects.

I would also like to thank my committee members Dr. Jason Stajich and Dr. Jikui Song for accepting my thesis and being part of my graduate committee. Both of them have provided excellent input, which has greatly enhanced the quality of my dissertation research. I would also like to thank our collaborator Dr. Susan Wessler.

I am very grateful to Robert Hice, without him this journey would have been extremely difficult. He has taught me to think critically, be patient and to troubleshoot the difficulties related to molecular biology. I would like to thank other Atkinson lab member, Michael Han, Anna-Louise Doss and Patrick Schreiner for their friendship and support. The time spent in the lab could not be have been enjoyable without them. I am highly thankful to Patrick Schreiner for bioinformatics guidance.

I would like to thank my parents Vijaykumar Shah and Saroj shah, without their support and motivation it would have been difficult to pursue my Ph.D. I would also like to thank my brother Jainik Shah for his support during the years I have spent in Riverside.

Friends have played a critical role in the completion of my dissertation Jui Dave, Shaily

Shah, Umesh Patel, Deepanshu Madan, Sonali Chaturvedi, Paurav Desai, Pankti Parmar, Priyanka Khire, and Neerja Katiyar have all provided me strength and encouragement that made getting my Ph.D. possible. I am very thankful to Kavita Sharma, Prithwiraj Das and Arun Kumar for helping for supporting me in the stressful time.

ABSTRACT OF THE DISSERTATION

Mobilization of Newly Identified Transposon *Muta1* in
Aedes aegypti and *Drosophila melanogaster*

by

Presha Vijaykumar Shah

Doctor of Philosophy, Graduate Program in Biochemistry and Molecular Biology
University of California, Riverside, June 2015
Dr. Peter Atkinson, Chairperson

My research aims to identify and test endogenous transposons in *Ae. aegypti* that remain active. In order to address this, the bioinformatics pipelines MITE-Hunter and TARGeT (Tree Analysis of Related Genes and Transposon) were used for identifying new active transposons in *Ae. aegypti* genome. Concurrent bioinformatics analysis performed discovered an interesting *Mutator* superfamily element called *Muta1*. To verify the activity of the *Muta1* element it was tested for it somatic activity in *D. melanogaster and Ae. aegypti* through excision and transposition assays. The somatic assays revealed that the *Muta1* element was active in *D. melanogaster* and *Ae. aegypti*. I also determined the germline and remobilization activity of the *Muta1* element post integration in *D. melanogaster* and *Ae. aegypti.* The experiment determined that the *Muta1* element was

able to remobilized post integration in *D. melanogaster* and *Ae. aegypti.* Another goal of this project focused on determining any relationship that might exist between the *Muta1* element and piwi-interacting RNAs (piRNAs) in *D. melanogaster*, which was naïve to the *Muta1* element. Based on the studies in *D. melanogaster* it is believed that piRNAs play role in protecting the genome from invasions of transposons. My goal here was to generate small RNA libraries to determine at which generation following transformation; the autonomous *Muta1* element in *D. melanogaster* will be silenced by detection of sequence specific piRNAs to the *Muta1* element. The library analysis revealed production of abundance of sense piRNAs to the *Muta1* transposase. Despite production of piRNAs, the *Muta1* element was still active in the transgenic flies. This suggests that the *Muta1* element might be highly active such that it was able to evade detection for suppression by host silencing machinery. Together all of these studies present first evidence of a *Mutator* superfamily transposon activity in insects specifically *D. melanogaster* and *Ae. aegypti*.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

Mobilization of Newly Identified Transposon *Muta1* in *Ae. aegypti* and *D. melanogaster*

1.1 Introduction to Transposable Elements

Transposable elements (TEs), also known as "jumping genes", are DNA sequences that move from one location in the genome to another. Barbara McClintock found the first evidence for mobile genetic elements in the form of chromosomal abnormalities within *Zea mays* in 1953, which she later named the *Ac* and *ds* transposons (McClintock, 1953). TEs are ubiquitous across the genomes of prokaryotes (Egner & Berg, 1981) and eukaryotes; almost half of the human genome is composed of TEs (Smit, 1999). As TEs can make up a large proportion of genomes, it is hypothesized that they have participated in changes of genome size during speciation and evolution, as reported in plants (SanMiguel, Gaut, Tikhonov, Nakajima, & Bennetzen, 1998), *Drosophila* (Sheen & Levis, 1994), and primates (Locke et al., 2003). One of the factors for genome evolution in *Drosophila virilis* is believed to be chromosomal inversions and fusions caused by the presence of *Penelope* and *Ulysses* TEs insertions (Evgen'ev et al., 2000). As a consequence, these changes led to speciation due to the incompatibility between different breeding populations (Evgen'ev et al., 2000). TEs are able to produces various genetic alterations, upon insertion as a consequence of the transposition process. TEs can inactivate, cause deleterious mutation or alter gene expression by insertion within exons, introns, or regulatory regions (Feschotte & Pritham, 2007).

1.2 Types of Transposons

There are two classes of transposable elements. Class I TEs are RNA transposons, or retroelements, which use copy and paste mechanism for their transposition in the genome. In the copy and paste mechanism, RNA is reverse transcribed into complementary DNA (cDNA) by an element-encoded reverse transcriptase. The integrase then inserts cDNA into a new region in the genome (Beauregard, Curcio, & Belfort, 2008). Retroelements can be further divided into two groups, LTR (long terminal repeats) and non-LTR elements. The non-LTR elements do not have long terminal repeats flanking the element and can be further divided into autonomous LINEs (long interspersed nuclear elements) and non-autonomous SINEs (short interspersed nuclear elements). Transposition of SINEs is dependent on the reverse transcription machinery of other retroelements (Beauregard et al., 2008) (Schmidt, n.d.).

In contrast, class II DNA transposons move via cut and paste mechanism, where they excise from one genomic region and integrate into a new location. There are two forms of class II elements, autonomous and non-autonomous DNA transposons. The autonomous DNA transposons contains a transposase gene to catalyze their movement in the genome while non-autonomous DNA transposons lack a transposase gene and require *trans* acting transposase for transposition. These DNA transposons have terminal inverted repeats (TIRs) and are grouped into widely divergent families based on their sequences and mobility properties. MITEs (miniature inverted terminal repeat elements) are deletion

derivatives of non-autonomous DNA transposons and are characterized by their small size and high copy number in the genome (Wessler, 1995) (Figure 1.1).

The class II superfamilies can be divided into three major classes; (1) those that use a cut and paste mechanism of transposition, (2), those that use a rolling circle mechanism, like *Helitrons* (Kapitonov & Jurka, 2007); and (3) *Maverick*, whose mechanism is not well understood, but likely replicate using a self encoded DNA polymerase (Pritham, Putliwala, & Feschotte, 2007). Within these major classes, there are 10 superfamilies of cut and paste DNA transposons, which are classified based on consensus sequences of their transposases (Han, Qin, & Wessler, 2013). These superfamilies are *Tc1/Mariner*, *hAT, P*-element, CACTA, *PiggyBac, MuDR/Foldback, PIF/Harbinger, Merlin, Transib,* and *Banshee,* (Feschotte & Pritham, 2007).

Transposons from the superfamilies; *piggyBac, hAT, Tc1/mariner* and *P*-element superfamilies are discussed here in more detail as these elements have been widely used for insect transgenesis.

1.2.1 The *P*-element

P-element was discovered as a result of a phenomenon called hybrid dysgenesis in *D. melanogaster* (Kidwell, Kidwell, & Nei, 1973). Hybrid dysgenesis was observed when males derived from a newly collected strain (P cytotype) were mated with females from a laboratory strain (M cytotype), which produced sterile offspring with abnormal gonads. Later, it was discovered that this abnormality was due to *P*-element invasion in flies.

P-element has been widely used for germline transformation of *D. melanogaster* (Daniels, Clark, Kidwell, & Chovnick, 1987). The full-length element is 2.9 kb in length, with 31 bp TIRs and inverted 11 bp repeats approximately 100 bp from each end (O'Hare & Rubin, 1983). It encodes a single ORF with four exons which produces a 87 KDa transposase (Beall & Rio, 1997). Germline transformation using the *P*-element has led to its use in the Genome Wide Gene Disruption Project, where *P*-element insertion disrupts regulatory regions of the genes, allowing mutant stocks to be created for each gene in the fly (Spradling et al., 1999).

Surveys of a large number of strains revealed that the *P*-element has spread rapidly in the natural population of *D. melanogaster* in last few decades (Daniels, Peterson, Strausbaugh, Kidwell, & Chovnick, 1990). Studies have shown that the *P*-element has invaded a related species, *D. simulans,* through horizontal transfer and it has been transferred from the distantly related species *D. willistoni* to *D. melanogaster* (Daniels et al., 1990).

1.2.2 The *piggyBac* element

The *piggyBac* element was first discovered in the *Tricoplusia ni* cell line as a repetitive element (M J Fraser, Smith, & Summers, 1983). The *piggyBac* element is 2.4 kb in length with 19-bp short TIRs, has an asymmetrical terminal repeat (TR) structure with a 3-bp spacer between the 5' 13-bp TR and a 19-bp TIR and a 31-bp spacer between 3' TR and TIR (Handler, McCombs, Fraser, & Saul, 1998). It has a single 2.1 kb open reading frame that encodes a functional transposase (Elick, Bauser, Principe, & Fraser, 1996) .

The *piggyBac* transposase uses a cut and paste mechanism to integrate into TTAA nucleotide sites (Wang and Frasier 1993). The excision site is repaired to original sequence after excision and thus excision of the *piggyBac* element leaves no footprints (M J Fraser, Ciszczon, Elick, & Bauser, 1996). Some of the advantages of the *piggyBac* element are the ability to carry large DNA fragments; element is not sensitive to overproduction inhibition in which elevated level of transposase causes decreased transposition (M. H. Wilson, Coates, & George, 2007), and the ability to integrate up to four transgene concurrently from a single transfection in human cell line. (Kim & Pyykko, 2011). The *piggyBac* element is useful for creating cell lines with stable expression of multiprotein complexes useful for drug discovery, and for therapeutic gene transfer where more than one transgene is required (Kahlig et al., 2010)

The *piggyBac* element is used for non-drosophilid transformation in Diptera, Lepidoptera, and Coleoptera species (Lorenzen et al., 2003). In the Genome Wide Disruption Project, *piggyBac* had been used along with the *P*-element in enhancing gene disruption in order to complete gene knockout collections (Thibault et al., 2004). The *piggyBac* element has been used for enhancer trapping in *D. melanogaster* (Horn, Offen, Nystedt, Häcker, & Wimmer, 2003) , in red flour beetle *Tribolium castaneum* (Lorenzen et al., 2003) and in *Anopheles stephensi* (David A O'Brochta, Alford, Pilitt, Aluvihare, & Harrell, 2011). In addition, this element has also showed its application for gene therapy in mouse and human cell lines, where the *piggyBac* element carried multiple genes and was shown to efficiently transpose in these cell lines (Ding et al., 2005).

1.2.3 The *Hermes* element

Hermes is a *hAT* (*hobo, Ac* and *Tam3*) transposon, originally identified from the housefly *Musca domestica* (Sarkar, Coates, et al., 1997). It is closely related to the *hobo* element from *D. melanogaster* (Sundararajan, Atkinson, & O'Brochta, 1999). It is 2,749 bp in length, has 17 bp imperfect TIR repeats, and has an 1800 bp open reading frame that encodes the *Hermes* transposase (Subramanian, Cathcart, Krafsur, Atkinson, & O'Brochta). The *Hermes* element has a wide host range and is active in fifteen different species including mosquitoes, fruit flies, moths, beetles among others (Michel et al., 2001) (Sarkar, Yardley, Atkinson, James, & O'Brochta, 1997) (Guimond, Bideshi, Pinkerton, Atkinson, & O'Brochta, 2003). The *Hermes* transposase crystal structure reveals that it forms an octamer during integration with the host DNA (Hickman et al., 2014). Studies have shown that the transposition of the *Hermes* in mosquitos leaves flanking plasmid sequence along with the gene of interest and thus moves via non-canonical cut and paste mechanism (N Jasinskiene, Coates, & James, 2000).

1.2.4 The *Mos1* element

Mos1, is a *Tc1/mariner* element, was first isolated from *Drosophila mauritiana* as a somatically unstable insertion into the *white-peach* gene, excision in somatic cells was recognized phenotypically by the occurrence of eye-color mosaicism in flies (Medhora, Maruyama, & Hartl, 1991). The *Mos1* element is 1.3 kb in length with 28 bp TIRs and contains a single ORF which codes for a 345 amino acid (Medhora et al., 1991). *Mos1* has a target site preference for TA and duplicates this site upon transposition. The *Mos1*

element has been successfully used for transformation of several insect species including *D. melanogaster* (Bryan, Garza, & Hartl, 1990), *D. hydei* (Lidholm, Lohe, & Hartl, 1993) and *Aedes aegypti* (Coates, Jasinskiene, Miyashiro, & James, 1998). The *Mos1* element transposes by forming a paired-end complex in which the *Mos1* transposase binds to a single end of the *Mos1* element as a monomer or a dimer. The ends are brought together to form a paired end complex and the transposon is then excised from the flanking DNA, and is subsequently integrated in a TA sequence elsewhere in the genome (Richardson, Colloms, Finnegan, & Walkinshaw, 2009).

1.2.5 The *Minos* element

The *Minos* element belongs to the *Tc1/mariner* superfamily. It was discovered by screening clones for non-ribosomal DNA in *D. hydei* (G. Franz & Savakis, 1991). It is 1.8 kb in length, possesses 254 bp TIRs, and encodes a single transposase (G. Franz, Loukeris, Dialektaki, Thompson, & Savakis, 1994). The target site preference for the *Minos* element is a TA dinucleotide, which is similar to other *Tc1/mariner* superfamily transposons. The first transformation using the *Minos* element was in *Ceratitis capitata* (Loukeris, Arcà, Livadaras, Dialektaki, & Savakis, 1995). Apart from the Mediterranean fruit fly, *Minos* has transformed *D. melanogaster* (Arcà, Zabalou, Loukeris, & Savakis, 1997), *An. Stephensi* (Catteruccia, Nolan, Blass, et al., 2000), *Tribolium castaneum* (Pavlopoulos, Oehler, Kapetanaki, & Savakis, 2007), *Bactrocera oleae* (Koukidou et al., 2006), and *Bombyx mori* (Uchino, Imamura, Shimizu, Kanda, & Tamura, 2007). Genome wide insertional mutagenesis and gene tagging in mammal cells has been

demonstrated using the *Minos* element (Klinakis, Zagoraiou, Vassilatis, & Savakis, 2000).

The *Minos* mediated integration cassette called, MiMIC system has been used for mutagenesis and genome manipulation in *D. melanogaster* (Venken et al., 2011). In this system, the *Minos* transposon carries a dominant marker and a gene trap cassette flanked by two inverted ΦC31 *attP* sites. The *attP* sites allow for replacement of the intervening sequence in the transposon with other sequences using a recombinase mediated cassette exchange (RCME); it can also revert insertions that function as a gene trap and convert a mutant phenotype back to wild-type (Venken et al., 2011).

1.3 Transposons in genetic modification of mosquitoes

Pest insects carry pathogens that cause human, plant, and animal diseases. Mosquitoes in particular are important vectors of several human pathogens. The maintenance and transmission of pathogens that cause infectious diseases like malaria, lymphatic filariasis, dengue and numerous other viral infectious diseases are dependent on the competence of the mosquito vector (Beerntsen, James, & Christensen, 2000). Reducing or even eliminating mosquito populations as a method of controlling malaria was first proposed by Ronald Ross (Prevention, n.d.). Insecticides, such as DDT, were used effectively under a campaign of malaria eradication initiated by the World Health Organization, where by 1965, 60% of the people previously living in malarial regions were subsequently living in areas where malaria was no longer endemic. One part of the world that did not benefit even at the peak of the eradication of campaign was sub-

saharan Africa. Efforts to control these diseases have led to insecticide resistant mosquito strain and drug resistant plasmodium strains. The World Health Organization reinforces prohibition on the use of DDT, which was widely used as an insecticide spray in tropical areas to control the incidence of malaria (United Nations Environment Programme Persistent Organic Pollutants Website; http://www.chem.unep.ch/pops), due to potential hazards associated with the use of DDT. These hazards include the poisoning of wildlife, environmental perturbation and the endangerment to human health by causing birth defect, reduced fertility, breast cancer, and diabetes (Davies, 2006) (van den Berg, 2009) (Kabasenche & Skinner, 2014). Vaccine development remains a viable approach, but requires extensive research to overcome the problems associated with the complexity of the pathogen (Miller & Hoffman, 1998). *Plasmodium berghei* expresses more than 5000 proteins throughout its different life stages, which makes it very difficult to determine potential effective protein targets of vaccine development (Herrington et al, 1997.).

Recent advances in molecular biology and genetics have allowed the study of mosquito vectors and pathogens in great depth. For the purposes of insect control, creating the technologies to produce transgenic insects has been a goal of molecular geneticists for years (David A O'Brochta et al., 2003). Efficient genetic manipulation of *Ae. aegypti*, *Anopheles gambiae* and *Culex quinquefasciatus* awaits the development of an array of powerful genetic tools like those readily available for well-studied model organisms such as fruit flies, yeast, and mice.

Ae. aegypti is a vector of several pathogens such as *P. gallinaceum*, and flaviviruses that cause dengue and yellow fever. Several other strains have been selected that support the complete development of the filarial worms *Brugia malayi*, *Brugai pahangi*, and *Dirofiaria immitis* that causes lymphatic filariasis (Erickson et al., 2009). *Ae. aegypti* currently supports genetic technologies such as transposon mediated transformation, site specific recombination and RNAi (Franz et al., 2011). The release of mosquitoes carrying a dominant lethal gene (RIDL), which renders males sterile, is another approach to control *Ae. aegypti* populations. In this approach, sterile males are released in the field where they compete with the wild-type males for mating partners (Winskill et al., 2014). A number of exogenous transposons such as *piggyBac*, *Mos1* and *Hermes* have been used for genetic transformation of *Ae. aegypti* (Kokoza, Ahmed, Wimmer, & Raikhel, 2001), but most of these transposon are not able to remobilize (D. A. O'Brochta, 2003).

The genetic transformation of *An. gambiae* was believed to be achieved through the use of *P* element, resulting in the integration of the element near a telomeric region, and integration was independent of the transposase (Miller et al., 1987). Later it was shown that this transformation event was due to non-homologous integration and not by *P* element mediated transposition (*Advances in Genetics, Volume 47*, 2002) (Oliveira de Carvalho, Silva, & Loreto, 2004). Thus, the *P*- element transformation is confined to drosophilid species (O'Brochta & Handler, 1988). The class II elements *Hermes* (N Jasinskiene et al., 1998) and *mariner* (Coates et al., 1998) have been used successfully to genetically transform the yellow fever mosquito, *Ae. aegypti*, but neither of these elements have been shown to transpose at high frequencies in *An. gambiae*

(Subramanian, Akala, Adejinmi, & O'Brochta, 2008) (Zhao & Eggleston, 1998). Catteruccia have demonstrated that the *Minos* element is capable of germline transformation of *An. stephensi* (Catteruccia, Nolan, Loukeris, et al., 2000).

Understanding gene vectors is critical to the success of insect transgenesis. The major constraint with the sterile insect technique (SIT) is the fitness of the transgenic strain that is used to release in the field, and their mating compatibility with the wild type population. The other problem with the transgenic strain is the instability of the transgene due to possible remobilization by closely related TEs in the different field population. The remobilization can lead to change in gene expression and loss of specific phenotype for which the strain was being used. To minimize instability, transformation construct should be developed using TEs that are not closely related to endogenous elements in the genome of insect being transformed. However, TEs are required to have high transformation efficiency so that many transgenic strains can be evaluated. This required the recipient strain possess the host factors necessary for transformation, and use of endogenous TEs can have an advantage (Malcolm J Fraser, 2012). However, the field is currently limited by the availability of efficient genetic tools that can be used for generating transgenic mosquitoes. An increase in transformation efficiency and a wider range of target site preferences would greatly propel the field of insect transgenesis allowing for the acceleration of important gene integration studies that could incidentally save the lives of millions of people by creating mosquitoes refractory to a disease

1.4 Transposable Elements as Genetic Tools

Forward genetics is an important approach for the discovery of new genes using an unbiased look into the genome. In forward genetics, a transposon that is used for insertional mutagenesis serves as a molecular tag to identify the affected gene. The first transposon used for this kind of approach was *Sleeping Beauty* (SB) (Dupuy, Fritz, & Largaespada, 2001). The SB transposon system has a large cargo capacity; it can efficiently move inserts >8 kb (de Silva et al., 2010), and seems to have a random integration pattern that is favorable for use in gene therapy, where high transposition frequency is required that can lead to increase in expression of the therapeutic gene used for gene therapy (Yant et al., 2000) (Vigdal, Kaufman, Izsvák, Voytas, & Ivics, 2002). SB was successfully used in cancer genetics in order to identify genes involved in malignant phenotypes, where SB acts as a insertional mutagen (Collier, Carlson, Ravimohan, Dupuy, & Largaespada, 2005).

The *P-element* from *Drosophila* has also been used in for forward genetics techniques. A number of *Drosophila* lines have been developed with non-autonomous *P*-element carrying selectable markers such as the eye color markers white and rosy (Zhai et al., 2003). In these lines, insertion of the *P*-element into a gene or next to a promoter can disrupt gene function producing visible or lethal phenotypes. The Gene Disruption Project, which aimed to disrupt every gene in *Drosophila*, had difficulties due to certain drawbacks such as insertion preference into specific sites. *P*-elements prefer to integrate

into 5' regions of the genes and near existing *P*-elements integrations (Spradling et al., 1999). *P-elements* have also been used to study the pattern and timing of gene expression by enhancer trapping. Here, a *P-element* carrying a reporter gene is linked to a weak basal promoter, which randomly mobilizes in the genome and eventually inserts near an endogenous enhancer. Subsequently, this activates the weak basal promoter and expresses the reporter gene under the control of the enhancer (C. Wilson et al., 1989).

Another widely used enhancer trap strategy is the GAL4-UAS binary system (Andrea Brand and Norbert Perrimon, 1953). This system utilizes enhancer trapping with the construct carrying the *Saccharomyces cerevisiae* transcriptional activator, GAL4, as a reporter gene. The binary system has two components, the GAL4 gene and the upstream activator sequence (UAS), which is bound by GAL4. TEs carrying GAL4 allows for the insertion of GAL4 next to a weak promoter, which can then express and bind to the upstream activator sequence linked to a reporter gene such as green or red fluorescent protein (Sepp & Auld, 1999) (D. A. O'Brochta, Pilitt, Harrell, Aluvihare, & Alford, 2012).

1.5 Genetic Technologies used for Insect Transformation

Site-specific transgene integration systems such as FLP- *FRT* from the *Saccharomyces cerevisiae* (Park, Masison, Eisenberg, & Greene, 2011) and CRE-*lox* from the Bacteriophage P1 (Sauer & Henderson, 1988) are used for gene knockout and

transgene expression. In the Cre-*loxP* or Flp-*FRT* system, Cre or Flp recombinase recognizes the 34-bp nucleotide sequence named *loxP* or *FRT* and precisely catalyzes the homologous exchange between the two *loxP* or the two *FRT* sites, respectively. The Cre-*loxP* and Flp-*FRT* systems (Hoess, Abremski, & Sternberg, 1984) enable the specific manipulation of DNA based on the direction and location of the two *loxPs* or *FRTs* sites. Cre catalyzes the deletion of the DNA between the two *loxPs* sites, when the two *loxPs* are in the same direction on one DNA molecule. When one *loxP* is on a linear DNA molecule and another *loxP* site is on a circular DNA molecule, the circular DNA integrates into the linear DNA at the target. If two *loxPs* sites are oriented in opposite directions, the fragment between them inverts. When one *loxP* site is on a linear strand and a second *loxP* is position on a different linear DNA strand, the two linear DNA molecules exchange a segment (similar to chromosomal rearrangement). This system has been used in *D. melanogaster* as a tool kit to inactivate genes (Frickenhaus, Wagner, Mallik, Catinozzi, & Storkebaum, 2015). The *cre-loxP* site-specific recombination system has also been used in *Ae. aegypti* (Nijole Jasinskiene, Coates, Ashikyan, & James, 2003). The *cre-loxP* system was used to remove a marker gene from transgenic mosquitoes. The *cre* recombinase precisely recognizes *loxP* sites in the genome and was able to catalyze excision, resulting in a excision frequency of 99.4% (Nijole Jasinskiene et al., 2003)..

The bacterial chromosome phiC31 from a *Streptomyces* Bacteriophage is used for stable integration of foreign DNA (Thorpe & Smith, 1998). Integrase from phage phiC31 catalyzes the homologous recombination between the *Streptomyces* attachment site *attB*

and the phage attachment site *attP*. The advantage of phiC31 integrase is that recombination is unidirectional, with interaction between *attB* and *attP* sites creating *attL* and *attR* junctions, that are no longer recognized by the integrase, making the integration both stable and efficient (Thyagarajan, Olivares, Hollis, Ginsburg, & Calos, 2001). The phiC31 system has been used in *D. melanogaste*r S2 cell culture, where intermolecular recombination occurred at a 47% rate. Transgenic lines were created using *P*-element to integrate *attP* sites and phiC31 integrase mRNA was injected into embryos to promote integration of plasmid with the *attB* site; 55% transgenic offspring were produced in *D. melanogaster* with precise integration of the *attP* site (Groth, Fish, Nusse, & Calos, 2004). In order for precise targeting of transgenic constructs to a predetermined position in the genome of *D. melanogaster*, phiC31 integrase system was used in conjunction with recombinase mediated cassette exchange. Here, the two-donor cassettes were marked with different marker genes. The exchange of the donor cassette with the integrated cassette carrying another marker results in loss of phenotype due to integrated cassette. Thus, exchange of sequences takes place as opposed to insertion only. This results in a change of phenotype that makes selection convenient (Bateman, Lee, & Wu, 2006). Germline transformation of *Ae. albopictus*, the Asian tiger mosquito, was achieved by microinjection of *piggyBac* with 3xp3-ECFP marker and an *attP* site combined with *piggyBac* transposase mRNA. Transformation frequency obtained with this injection was 2-3%. The transgenic lines were injected with a second plasmid containing an *attB* site and 3xP3-DsRed marker combined with phiC31 integrase mRNA. All three lines were successfully transformed with a transformation efficiency of 2-6% (Labbé, Nimmo, &

Alphey, 2010). The phiC31 integrase system has been efficiently used in human cell lines where the risk of integration into pseudo *attP* site has been assessed and it has been concluded that this system is safe for use in the human gene therapy (Chalberg et al., 2006).

Despite the availability of other genetic technologies for the transformation of insects, transposons typically are used for insertion of transgenes as seen in FLP/FRT system used in *D. melanogaster* where the enzyme FRT had to be introduced into the *D. melanogaster* using *P*-element mediated transformation (Rong and Golic, 2000). Furthermore, transformation is not always achieved using this recombination system in mosquitoes, which was observed in *Ae. aegypti* with the use of FLP/FRT recombination system (Nijole Jasinskiene et al., 2003).

1.6 Transposon Control System

Transposons can alter the gene expression of nearby genes and potentially large chromatin domains, triggering coordinated changes in gene transcription that could disrupt development or drive evolution (Feschotte, 2008). Thus it is important to have a control system for transposons that will balance the beneficial and maladaptive affects transposons have on the genome.

In order to control transposon load in the genome, organisms need to be able to distinguish between host transcript and those of transposable elements. Looking at the structure of transposons based on class I retrotransposons and class II DNA transposons,

there are multiple families and sub-families with elements that contains unique TIRS and vary greatly in length. Endogenous retroviruses, such as gypsy and ZAM, are another group of elements that have the ability to move in a fashion very similar to retrotransposons (Desset, Meignin, Dastugue, & Vaury, 2003). Thus, a genome has to defend against a wide variety of invaders.

1.6.1 Small RNA Biogenesis and Function

RNA interference (RNAi) was first discovered in 1993. There are different types of small RNAs including siRNAs, microRNAs (miRNAs) and Piwi interacting RNAs or piRNAs (Wightman, Burglin, Gatto, Arasu, & Ruvkun, 1991) (Zamore, Tuschl, Sharp, & Bartel, 2000) (Aravin et al., 2001). Each of these classes differs in their biogenesis and target regulation, yet together these three classes of small interfering RNA collaborate to regulate gene expression and genome defense.

In 1998, Fire and Mello established double stranded RNA (dsRNA) as the silencing mechanism in *Caenorhabditis elegans* (Fire et al., 1998). RNAi is triggered by long dsRNAs, which are cleaved by an RNase III family member, Dicer, into short RNA sequences. These cleaved short interfering RNAs (siRNAs) are 21-23 nt in length with 5' phosphorylated ends. Dicer produces small RNAs then interacts with a second complex composed of Argonaute proteins which together form the RNA-induced silencing complex (RISC), that target mRNAs for silencing (Djikeng, Shi, Tschudi, & Ullu, 2001). In the miRNA pathway, small RNAs inhibit mRNA translation. The miRNAs that are partially complementary to mRNAs are generated by Dicer, which acts with a dsRNA-

binding protein partner. siRNAs and miRNAs are derived from different sources of dsRNA and act in common pathways interchangeably depending on the degree of complementarity with their target RNA sequence (Aravin et al., 2003).

1.6.2 piRNAs in *Drosophila*

Aravin and Brennecke first characterized repeat associated small interfering RNAs (rasiRNAs) in *Drosophila* in 2003 (Aravin et al., 2003). In 2006, the rasiRNAs were reevaluated and renamed piRNAs (Aravin et al., 2006). The *Drosophila melanogaster* genome is comprised of ~15%–20% of TEs (Kaminker et al., 2002). Uncontrolled activity of TEs triggers defects in genome integrity due to DNA strand breaks, insertional mutagenesis, and illegitimate recombination (Levin & Moran, 2011) (Slotkin & Martienssen, 2007). The piRNA pathway is a small RNA silencing system that includes the PIWI family of proteins bound to 23-30 nt piRNAs (Malone & Hannon, 2009) (Siomi, Sato, Pezic, & Aravin, 2011a). The piRNA pathway is dicer-independent compared to the siRNA and miRNA pathways. The siRNA and miRNA pathways are well studied compared to the piRNA pathway. Studies in *Drosophila* have shown that most of the piRNAs are derived from piRNA clusters, which are large loci that are filled with TE sequences (Brennecke et al., 2007)

The *Drosophila* ovary contains two major cell types: germline cells derived from primordial germ cells, and somatic support cells derived from the mesodermal. Both cell types silence TEs via the piRNA pathway, but the respective pathways differ considerably (Senti & Brennecke, 2010) (Siomi, Sato, Pezic, & Aravin, 2011b). The

somatic cells expresses only the Piwi protein, and the germline cells express Aubergine (AUB) and Argonaute (AGO3) (Malone et al., 2009). There are two biogenesis pathways: the primary pathway, and the secondary pathway that involves the ping-pong amplification mechanism. Both pathways are important for mounting an effective defense against transposons.

First, primary piRNA biogenesis provides an initial pool of piRNAs that target multiple TEs, which are then channeled into the ping-pong pathway in the germline (Siomi et al., 2011b). A putative nuclease encoded by zucchini is required for piRNA production in somatic cells, which is believed to form the 5' end of the primary piRNAs (Gunawardane et al., 2007) (Malone et al., 2009). Moreover, primary piRNAs have a strong bias for uridine at position 1 (U1) and are produced from only one genomic strand (Brennecke et al., 2007).

The ping-pong model requires preexisting primary piRNA, which are associated with PIWI or AUB; and are mostly antisense to transposon mRNAs (Nishida et al., 2007) (Saito et al., 2006). These primary piRNAs target complementary transcripts, which are processed and loaded onto AGO3. In addition, it is believed that the nuclease activity of AUB and AGO3 cleaves target RNAs between their 10^{th} and 11^{th} nucleotides relative to the 'guide' small RNAs. Cleavage of complementary transcripts targeted by AUB-bound primary piRNA leads to the generation of the 5' end of new secondary piRNAs that have an adenine bias at 10^{th} nucleotide, and are in the sense orientation. (Gunawardane et al., 2007) (Brennecke et al., 2007). (Figure 1.2)

1.6.4 Silencing of Transposable elements

In *D. melanogaster*, mutations in the members of the PIWI family; Piwi, Aubergine and Ago3 lead to transposon derepression in the germ line, indicating that they act non-redundantly during TE silencing (Aravin et al., 2001) (Kalmykova, Klenov, & Gvozdev, 2005) (Savitsky, Kwon, Georgiev, Kalmykova, & Gvozdev, 2006) (Li et al., 2009). The PIWI proteins have crucial roles in gonad development: both PIWI and AUB are required for male and female fertility and AGO3 is required for female fertility, but only partially for male fertility (Cox et al., 1998) (Lin & Spradling, 1997).

1.6.5 Epigenetic role of piRNAs

The importance of the amplification loop in defense against transposons is underlined by studies of hybrid dysgenesis. In this phenomenon, maternally deposited piRNAs protect eggs from transposon invasion, as piRNAs are present before zygotic transcription. Embryos that do not have maternally deposited piRNAs targeted against invasive transposons and are found to be sterile due to the over abundant integration of transposons in the genome (Brennecke et al., 2008).

Mutation in PIWI proteins also leads to defects in maintenance of germline stem cells (Lin & Spradling, 1997) (Cox et al., 1998). Deletion mutants of *flamenco,* which is a piRNA cluster in somatic cells, leads to female sterility as well as transposon derepression (Mével-Ninio, Pelisson, Kinder, Campos, & Bucheton, 2007). In addition, AUB mutants resemble spindle class mutants that are defective in mitotic progression due

to kinase dependent mitotic checkpoint activation (J. E. Wilson, Connell, & Macdonald, 1996) (Klattenhoff & Theurkauf, 2008). piRNAs act at several levels; Piwi, localized to the nucleus, binds to HP1a protein and has been implicated in heterochromatin assembly in somatic cells (Pal-Bhadra et al., 2004) (Brower-Toland et al., 2007). In addition, mutation in spn-E, which encodes a putative helicase required for piRNA production, reduces HP1a binding to the telomere specific transposon TART (Klenov et al., 2007). These findings suggest that piRNA bound to Piwi guide heterochromatin assembly, and thus impose transcriptional silencing.

Recent studies have shown the presence of piRNAs in neurons, these piRNAs are abundant in neurons, and are called *Aplysia* piRNAs. The *Aplysia* piRNAs associate with a neuronal PIWI protein, and are distinctly regulated by neuromodulators important for learning and memory. The Piwi-piRNA complex can methylate the CpG island in the promoter of *Creb2* gene in neurons, which is a major inhibitor constraint of memory in *Aplysia*, leading to long-term synaptic facilitation. These finding suggest the small RNA mediated gene regulation mechanism for establishing changes in gene expression in neurons involved in long-term memory storage (Rajasethupathy et al., 2012).

1.6.7 piRNAs in *Ae. aegypti*

The low transformation efficiency of *Ae. aegypti* suggest that the piRNA pathway might be even more complex and robust than in *D. melanogaster* (Arensburger, Hice, Wright, Craig, & Atkinson, 2011b). In addition, piRNAs with ping- pong signature have been found in *Ae. aegypti,* directed against viral sequences (Morazzani, Wiley, Murreddu,

Adelman, & Myles, 2012). The mechanism of the piRNA biogenesis in *Ae. aegypti* is not well understood and fairly recent bioinformatics analysis in the *Ae. aegypti* genome shows that it contains an expansion of the Piwi gene family, there is a single Ago3 homolog and six different Piwi genes (Arensburger, Hice, Wright, Craig, & Atkinson, 2011a). Piwi2 is similar to *Drosophila* Aub based on sequence similarity. These data were obtained by high throughput sequencing of seven libraries from *Ae. aegypti* (Arensburger et al., 2011a). A number of piRNA clusters were identified, which covered 20.6% of the *Ae. aegypti* genome. Additionally, 84% of these piRNAs were antisense in consistence with the pattern observed in *D. melanogaster*. In *Ae. aegypti*, 19% of the sequenced piRNAs mapped to transposons in contrast to *D. melanogaster* libraries where 50% of the sequenced piRNAs mapped to the TEs. 81% of the piRNAs that did not map to transposons were analyzed for their possible association with protein coding genes. Many piRNAs were observed to be mapped to the sense strand of the 3'-UTR, which suggest that piRNAs may also be involved in the regulation of downstream genes (Arensburger et al., 2011b).

1.6.8 Transposon Defense in *Ae. aegypti*

The genetic transformation of *Ae. aegypti* with TEs is possible, but efforts to remobilize a transposon once integrated have failed in the past. This is possibly due to silencing mechanisms in *Ae. aegypti*. *piggyBac*, *Mos1* and *Hermes* have been used to transform *Ae. aegypti,* but their inefficiency to remobilize have rendered them unfavourable as tools for gene tagging and enhancer trapping (Smith & Atkinson, 2011) (David A O'Brochta et al.,

2003) (Guimond et al., 2003). Approximately 47% of the *Ae. aegypti* genome is composed of transposons; the diversity of endogenous transposons together with the diversity of piRNAs mapped to them suggest that piRNA pathway may be involved in maintaining the integrity of the genome against the high transposon load (Arensburger et al., 2011b).

1.7 Thesis objectives and Aims

This thesis demonstrates identification and characterization of a new endogenous DNA transposons in *Ae, aegypti* using bioinformatics pipelines. The bioinformatics analysis discovered many new potentially active transposons that include the *Muta1* transposon. Chapter three and four investigates transposition, excision, target site preference, and transformation efficiency of newly discovered transposon *Muta1* in *D. melanogaster* and *Ae. aegypti.* Chapter five demonstrates relationship between piRNAs and the *Muta1* element in *D. melanogaster.* Overall, the research provides valuable information about the activity of the *Muta1* element, which can have a wide variety of application for insect transgenesis and human gene therapy.

1.8 Reference

Advances in Genetics, Volume 47. (2002). Academic Press. Aravin, A. A., Lagos-Quintana, M., Yalcin, A., Zavolan, M., Marks, D., Snyder, B., … Tuschl, T. (2003). The Small RNA Profile during Drosophila melanogaster Development. *Developmental Cell*, *5*(2), 337–350. http://doi.org/10.1016/S1534-5807(03)00228-4

Aravin, A. A., Naumova, N. M., Tulin, A. V., Vagin, V. V., Rozovsky, Y. M., & Gvozdev, V. A. (2001). Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline. *Current Biology*, *11*(13), 1017–1027. http://doi.org/10.1016/S0960-9822(01)00299-8

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., … Tuschl, T. (2006). A novel class of small RNAs bind to MILI protein in mouse testes. *Nature*, *442*(7099), 203–7. http://doi.org/10.1038/nature04916

Arcà, B., Zabalou, S., Loukeris, T. G., & Savakis, C. (1997). Mobilization of a Minos transposon in Drosophila melanogaster chromosomes and chromatid repair by heteroduplex formation. *Genetics*, *145*(2), 267–79.

Arensburger, P., Hice, R. H., Wright, J. A., Craig, N. L., & Atkinson, P. W. (2011a). The mosquito Aedes aegypti has a large genome size and high transposable element load but contains a low proportion of transposon-specific piRNAs. *BMC Genomics*.

Arensburger, P., Hice, R. H., Wright, J. A., Craig, N. L., & Atkinson, P. W. (2011b). The mosquito Aedes aegypti has a large genome size and high transposable element load but contains a low proportion of transposon-specific piRNAs. *BMC Genomics*, *12*, 606. http://doi.org/10.1186/1471-2164-12-606

Bateman, J. R., Lee, A. M., & Wu, C. (2006). Site-specific transformation of Drosophila via phiC31 integrase-mediated cassette exchange. *Genetics*, *173*(2), 769–77. http://doi.org/10.1534/genetics.106.056945

Beall, E. L., & Rio, D. C. (1997). Drosophila P-element transposase is a novel site-specific endonuclease. *Genes & Development*, *11*(16), 2137–51. Beauregard, A., Curcio, M. J., & Belfort, M. (2008). The take and give between retrotransposable elements and their hosts. *Annual Review of Genetics*, *42*, 587–617. http://doi.org/10.1146/annurev.genet.42.110807.091549

Beerntsen, B. T., James, A. A., & Christensen, B. M. (2000). Genetics of mosquito vector competence. *Microbiology and Molecular Biology Reviews : MMBR*, *64*(1), 115–37. Brennecke, J., Aravin, A. A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., &

Hannon, G. J. (2007). Discrete Small RNA-Generating Loci as Master Regulators of Transposon Activity in Drosophila. *Cell*, *128*(6), 1089–1103.

Brennecke, J., Malone, C. D., Aravin, A. A., Sachidanandam, R., Stark, A., & Hannon, G. J. (2008). An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science (New York, N.Y.)*, *322*(5906), 1387–92. http://doi.org/10.1126/science.1165171

Brower-Toland, B., Findley, S. D., Jiang, L., Liu, L., Yin, H., Dus, M., … Lin, H. (2007). Drosophila PIWI associates with chromatin and interacts directly with HP1a. *Genes & Development*, *21*(18), 2300–11. http://doi.org/10.1101/gad.1564307

Bryan, G., Garza, D., & Hartl, D. (1990). Insertion and excision of the transposable element mariner in Drosophila. *Genetics*, *125*(1), 103–14.

Catteruccia, F., Nolan, T., Blass, C., Muller, H. M., Crisanti, A., Kafatos, F. C., & Loukeris, T. G. (2000). Toward Anopheles transformation: Minos element activity in anopheline cells and embryos. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(5), 2157–62. http://doi.org/10.1073/pnas.040568397

Catteruccia, F., Nolan, T., Loukeris, T. G., Blass, C., Savakis, C., Kafatos, F. C., & Crisanti, A. (2000). Stable germline transformation of the malaria mosquito Anopheles stephensi. *Nature*, *405*(6789), 959–62. http://doi.org/10.1038/35016096

Chalberg, T. W., Portlock, J. L., Olivares, E. C., Thyagarajan, B., Kirby, P. J., Hillman, R. T., … Calos, M. P. (2006). Integration specificity of phage phiC31 integrase in the human genome. *Journal of Molecular Biology*, *357*(1), 28–48. http://doi.org/10.1016/j.jmb.2005.11.098

Coates, C. J., Jasinskiene, N., Miyashiro, L., & James, A. A. (1998). Mariner transposition and transformation of the yellow fever mosquito, Aedes aegypti. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(7), 3748–51.

Collier, L. S., Carlson, C. M., Ravimohan, S., Dupuy, A. J., & Largaespada, D. A. (2005). Cancer gene discovery in solid tumours using transposon-based somatic mutagenesis in the mouse. *Nature*, *436*(7048), 272–6. http://doi.org/10.1038/nature03681

Cox, D. N., Chao, A., Baker, J., Chang, L., Qiao, D., & Lin, H. (1998). A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. *Genes & Development*, *12*(23), 3715–27.

Daniels, S. B., Clark, S. H., Kidwell, M. G., & Chovnick, A. (1987). Genetic transformation of Drosophila melanogaster with an autonomous P element: phenotypic and molecular analyses of long-established transformed lines. *Genetics*, *115*(4), 711–23.

Daniels, S. B., Peterson, K. R., Strausbaugh, L. D., Kidwell, M. G., & Chovnick, A. (1990). Evidence for horizontal transmission of the P transposable element between Drosophila species. *Genetics*, *124*(2), 339–55.

Davies, K. (2006). Strategies for eliminating and reducing persistent bioaccumulative toxic substances: common approaches, emerging trends, and level of success. *Journal of Environmental Health*, *69*(5), 9–15, 36, 38.

De Silva, S., Mastrangelo, M. A., Lotta, L. T., Burris, C. A., Izsvák, Z., Ivics, Z., & Bowers, W. J. (2010). Herpes simplex virus/Sleeping Beauty vector-based embryonic gene transfer using the HSB5 mutant: loss of apparent transposition hyperactivity in vivo. *Human Gene Therapy*, *21*(11), 1603–13. http://doi.org/10.1089/hum.2010.062

Desset, S., Meignin, C., Dastugue, B., & Vaury, C. (2003). COM, a heterochromatic locus governing the control of independent endogenous retroviruses from Drosophila melanogaster. *Genetics*, *164*(2), 501–9.

Ding, S., Wu, X., Li, G., Han, M., Zhuang, Y., & Xu, T. (2005). Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. *Cell*, *122*(3), 473–83. http://doi.org/10.1016/j.cell.2005.07.013

Djikeng, A., Shi, H., Tschudi, C., & Ullu, E. (2001). RNA interference in Trypanosoma brucei: Cloning of small interfering RNAs provides evidence for retroposon-derived 24-26-nucleotide RNAs. *RNA*, *7*(11), 1522–1530.

Dupuy, A. J., Fritz, S., & Largaespada, D. A. (2001). Transposition and gene disruption in the male germline of the mouse. *Genesis (New York, N.Y. : 2000)*, *30*(2), 82–8.

Egner, C., & Berg, D. E. (1981). Excision of transposon Tn5 is dependent on the inverted repeats but not on the transposase function of Tn5. *Proceedings of the National Academy of Sciences of the United States of America*, *78*(1), 459–63.

Elick, T. A., Bauser, C. A., Principe, N. M., & Fraser, M. J. (1996). PCR analysis of insertion site specificity, transcription, and structural uniformity of the Lepidopteran transposable element IFP2 in the TN-368 cell genome. *Genetica*, *97*(2), 127–39.

Erickson, S. M., Xi, Z., Mayhew, G. F., Ramirez, J. L., Aliota, M. T., Christensen, B. M., & Dimopoulos, G. (2009). Mosquito infection responses to developing filarial worms. *PLoS Neglected Tropical Diseases*, *3*(10), e529. http://doi.org/10.1371/journal.pntd.0000529

Evgen'ev, M. B., Zelentsova, H., Poluectova, H., Lyozin, G. T., Veleikodvorskaja, V., Pyatkov, K. I., … Kidwell, M. G. (2000). Mobile elements and chromosomal evolution in

the virilis group of Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(21), 11337–42. http://doi.org/10.1073/pnas.210386297

Feschotte, C. (2008). Transposable elements and the evolution of regulatory networks. *Nature Reviews. Genetics*, *9*(5), 397–405. http://doi.org/10.1038/nrg2337

Feschotte, C., & Pritham, E. J. (2007). DNA transposons and the evolution of eukaryotic genomes. *Annual Review of Genetics*, *41*, 331–68. http://doi.org/10.1146/annurev.genet.40.110405.090448

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., & Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, *391*(6669), 806–11. http://doi.org/10.1038/35888

Franz, A. W. E., Jasinskiene, N., Sanchez-Vargas, I., Isaacs, A. T., Smith, M. R., Khoo, C. C. H., … Olson, K. E. (2011). Comparison of transgene expression in Aedes aegypti generated by mariner Mos1 transposition and ΦC31 site-directed recombination. *Insect Molecular Biology*, *20*(5), 587–98. http://doi.org/10.1111/j.1365-2583.2011.01089.x

Franz, G., Loukeris, T. G., Dialektaki, G., Thompson, C. R., & Savakis, C. (1994). Mobile Minos elements from Drosophila hydei encode a two-exon transposase with similarity to the paired DNA-binding domain. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(11), 4746–50.

Franz, G., & Savakis, C. (1991). Minos, a new transposable element from Drosophila hydei, is a member of the Tc1-like family of transposons. *Nucleic Acids Research*, *19*(23), 6646.

Fraser, M. J. (2012). Insect transgenesis: current applications and future prospects. *Annual Review of Entomology*, *57*, 267–89. http://doi.org/10.1146/annurev.ento.54.110807.090545

Fraser, M. J., Ciszczon, T., Elick, T., & Bauser, C. (1996). Precise excision of TTAA-specific lepidopteran transposons piggyBac (IFP2) and tagalong (TFP3) from the baculovirus genome in cell lines from two species of Lepidoptera. *Insect Molecular Biology*, *5*(2), 141–51.

Fraser, M. J., Smith, G. E., & Summers, M. D. (1983). Acquisition of Host Cell DNA Sequences by Baculoviruses: Relationship Between Host DNA Insertions and FP Mutants of Autographa californica and Galleria mellonella Nuclear Polyhedrosis Viruses. *Journal of Virology*, *47*(2), 287–300.

Frickenhaus, M., Wagner, M., Mallik, M., Catinozzi, M., & Storkebaum, E. (2015). Highly efficient cell-type-specific gene inactivation reveals a key function for the

Drosophila FUS homolog cabeza in neurons. *Scientific Reports*, *5*, 9107. http://doi.org/10.1038/srep09107

Groth, A. C., Fish, M., Nusse, R., & Calos, M. P. (2004). Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31. *Genetics*, *166*(4), 1775–82.

Guimond, N., Bideshi, D. K., Pinkerton, A. C., Atkinson, P. W., & O'Brochta, D. A. (2003). Patterns of Hermes transposition in Drosophila melanogaster. *Molecular Genetics and Genomics : MGG*, *268*(6), 779–90. http://doi.org/10.1007/s00438-002-0800-4

Gunawardane, L. S., Saito, K., Nishida, K. M., Miyoshi, K., Kawamura, Y., Nagami, T., … Siomi, M. C. (2007). A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. *Science (New York, N.Y.)*, *315*(5818), 1587–90. http://doi.org/10.1126/science.1140494

Han, Y., Qin, S., & Wessler, S. R. (2013). Comparison of class 2 transposable elements at superfamily resolution reveals conserved and distinct features in cereal grass genomes. *BMC Genomics*, *14*(1), 71. http://doi.org/10.1186/1471-2164-14-71

Handler, A. M., McCombs, S. D., Fraser, M. J., & Saul, S. H. (1998). The lepidopteran transposon vector, piggyBac, mediates germ-line transformation in the Mediterranean fruit fly. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(13), 7520–5.

Herrington, D. A., Clyde, D. F., Losonsky, G., Cortesia, M., Murphy, J. R., Davis, J., … Gillessen, D. Safety and immunogenicity in man of a synthetic peptide malaria vaccine against Plasmodium falciparum sporozoites. *Nature*, *328*(6127), 257–9. http://doi.org/10.1038/328257a0

Hickman, A. B., Ewis, H. E., Li, X., Knapp, J. A., Laver, T., Doss, A.-L., … Dyda, F. (2014). Structural basis of hAT transposon end recognition by Hermes, an octameric DNA transposase from Musca domestica. *Cell*, *158*(2), 353–67. http://doi.org/10.1016/j.cell.2014.05.037

Hoess, R., Abremski, K., & Sternberg, N. (1984). The nature of the interaction of the P1 recombinase Cre with the recombining site loxP. *Cold Spring Harbor Symposia on Quantitative Biology*, *49*, 761–8.

Horn, C., Offen, N., Nystedt, S., Häcker, U., & Wimmer, E. A. (2003). piggyBac-based insertional mutagenesis and enhancer detection as a tool for functional insect genomics. *Genetics*, *163*(2), 647–61.

Jasinskiene, N., Coates, C. J., Ashikyan, A., & James, A. A. (2003). High efficiency, site-specific excision of a marker gene by the phage P1 cre-loxP system in the yellow fever mosquito, Aedes aegypti. *Nucleic Acids Research*, *31*(22), e147.

Jasinskiene, N., Coates, C. J., Benedict, M. Q., Cornel, A. J., Rafferty, C. S., James, A. A., & Collins, F. H. (1998). Stable transformation of the yellow fever mosquito, Aedes aegypti, with the Hermes element from the housefly. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(7), 3743–7.

Jasinskiene, N., Coates, C. J., & James, A. A. (2000). Structure of hermes integrations in the germline of the yellow fever mosquito, Aedes aegypti. *Insect Molecular Biology*, *9*(1), 11–8.

Kabasenche, W. P., & Skinner, M. K. (2014). DDT, epigenetic harm, and transgenerational environmental justice. *Environmental Health : A Global Access Science Source*, *13*, 62. http://doi.org/10.1186/1476-069X-13-62

Kahlig, K. M., Saridey, S. K., Kaja, A., Daniels, M. A., George, A. L., & Wilson, M. H. (2010). Multiplexed transposon-mediated stable gene transfer in human cells. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(4), 1343–8. http://doi.org/10.1073/pnas.0910383107

Kalmykova, A. I., Klenov, M. S., & Gvozdev, V. A. (2005). Argonaute protein PIWI controls mobilization of retrotransposons in the Drosophila male germline. *Nucleic Acids Research*, *33*(6), 2052–9. http://doi.org/10.1093/nar/gki323

Kaminker, J., Bergman, C., Kronmiller, B., Carlson, J., Svirskas, R., Patel, S., … Celniker, S. (2002). The transposable elements of the Drosophila melanogaster euchromatin: a genomics perspective. *Genome Biology*, *3*(12), research0084.1–0084.20. http://doi.org/10.1186/gb-2002-3-12-research0084

Kapitonov, V. V, & Jurka, J. (2007). Helitrons on a roll: eukaryotic rolling-circle transposons. *Trends in Genetics : TIG*, *23*(10), 521–9. http://doi.org/10.1016/j.tig.2007.08.004

Kidwell, M. G., Kidwell, J. F., & Nei, M. (1973). A case of high rate of spontaneous mutation affecting viability in Drosophila melanogaster. *Genetics*, *75*(1), 133–53. Kim, A., & Pyykko, I. (2011). Size matters: versatile use of PiggyBac transposons as a genetic manipulation tool. *Molecular and Cellular Biochemistry*, *354*(1-2), 301–9. http://doi.org/10.1007/s11010-011-0832-3

Klattenhoff, C., & Theurkauf, W. (2008). Biogenesis and germline functions of piRNAs. *Development (Cambridge, England)*, *135*(1), 3–9. http://doi.org/10.1242/dev.006486

Klenov, M. S., Lavrov, S. A., Stolyarenko, A. D., Ryazansky, S. S., Aravin, A. A., Tuschl, T., & Gvozdev, V. A. (2007). Repeat-associated siRNAs cause chromatin silencing of retrotransposons in the Drosophila melanogaster germline. *Nucleic Acids Research*, *35*(16), 5430–8. http://doi.org/10.1093/nar/gkm576

Klinakis, A. G., Zagoraiou, L., Vassilatis, D. K., & Savakis, C. (2000). Genome-wide insertional mutagenesis in human cells by the Drosophila mobile element Minos. *EMBO Reports*, *1*(5), 416–21. http://doi.org/10.1093/embo-reports/kvd089

Kokoza, V., Ahmed, A., Wimmer, E. A., & Raikhel, A. S. (2001). Efficient transformation of the yellow fever mosquito Aedes aegypti using the piggyBac transposable element vector pBac[3xP3-EGFP afm]. *Insect Biochemistry and Molecular Biology*, *31*(12), 1137–43.

Koukidou, M., Klinakis, A., Reboulakis, C., Zagoraiou, L., Tavernarakis, N., Livadaras, I., … Savakis, C. (2006). Germ line transformation of the olive fly Bactrocera oleae using a versatile transgenesis marker. *Insect Molecular Biology*, *15*(1), 95–103. http://doi.org/10.1111/j.1365-2583.2006.00613.x

Labbé, G. M. C., Nimmo, D. D., & Alphey, L. (2010). piggybac- and PhiC31-mediated genetic transformation of the Asian tiger mosquito, Aedes albopictus (Skuse). *PLoS Neglected Tropical Diseases*, *4*(8), e788. http://doi.org/10.1371/journal.pntd.0000788

Levin, H. L., & Moran, J. V. (2011). Dynamic interactions between transposable elements and their hosts. *Nature Reviews Genetics*, *12*(9), 615–627. http://doi.org/10.1038/nrg3030

Li, C., Vagin, V. V., Lee, S., Xu, J., Ma, S., Xi, H., … Zamore, P. D. (2009). Collapse of Germline piRNAs in the Absence of Argonaute3 Reveals Somatic piRNAs in Flies. *Cell*, *137*(3), 509–521. http://doi.org/10.1016/j.cell.2009.04.027

Lidholm, D. A., Lohe, A. R., & Hartl, D. L. (1993). The Transposable Element mariner Mediates Germline Transformation in Drosophila melanogaster. *Genetics*, *134*(3), 859–868.

Lin, H., & Spradling, A. C. (1997). A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary. *Development (Cambridge, England)*, *124*(12), 2463–76.

Locke, D. P., Segraves, R., Carbone, L., Archidiacono, N., Albertson, D. G., Pinkel, D., & Eichler, E. E. (2003). Large-scale variation among human and great ape genomes determined by array comparative genomic hybridization. *Genome Research*, *13*(3), 347–57. http://doi.org/10.1101/gr.1003303

Lorenzen, M. D., Berghammer, A. J., Brown, S. J., Denell, R. E., Klingler, M., & Beeman, R. W. (2003). piggyBac-mediated germline transformation in the beetle Tribolium castaneum. *Insect Molecular Biology*, *12*(5), 433–40.

Loukeris, T. G., Arcà, B., Livadaras, I., Dialektaki, G., & Savakis, C. (1995). Introduction of the transposable element Minos into the germ line of Drosophila melanogaster. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(21), 9485–9.

Malone, C. D., Brennecke, J., Dus, M., Stark, A., McCombie, W. R., Sachidanandam, R., & Hannon, G. J. (2009). Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. *Cell*, *137*(3), 522–35. http://doi.org/10.1016/j.cell.2009.03.040

Malone, C. D., & Hannon, G. J. (2009). Small RNAs as Guardians of the Genome. *Cell*, *136*(4), 656–668. http://doi.org/10.1016/j.cell.2009.01.045

McClintock, B. (1953). Induction of Instability at Selected Loci in Maize. *Genetics*, *38*(6), 579–99.

Medhora, M., Maruyama, K., & Hartl, D. L. (1991). Molecular and functional analysis of the mariner mutator element Mos1 in Drosophila. *Genetics*, *128*(2), 311–8.

Mével-Ninio, M., Pelisson, A., Kinder, J., Campos, A. R., & Bucheton, A. (2007). The flamenco locus controls the gypsy and ZAM retroviruses and is required for Drosophila oogenesis. *Genetics*, *175*(4), 1615–24. http://doi.org/10.1534/genetics.106.068106

Michel, K., Stamenova, A., Pinkerton, A. C., Franz, G., Robinson, A. S., Gariou-Papalexiou, A., … Atkinson, P. W. (2001). Hermes-mediated germ-line transformation of the Mediterranean fruit fly Ceratitis capitata. *Insect Molecular Biology*, *10*(2), 155–62.

Miller, L. H., & Hoffman, S. L. (1998). Research toward vaccines against malaria. *Nature Medicine*, *4*(5 Suppl), 520–4.

Miller, L. H., Sakai, R. K., Romans, P., Gwadz, R. W., Kantoff, P., & Coon, H. G. (1987). Stable integration and expression of a bacterial gene in the mosquito Anopheles gambiae. *Science (New York, N.Y.)*, *237*(4816), 779–81.

Morazzani, E. M., Wiley, M. R., Murreddu, M. G., Adelman, Z. N., & Myles, K. M. (2012). Production of virus-derived ping-pong-dependent piRNA-like small RNAs in the mosquito soma. *PLoS Pathogens*, *8*(1), e1002470. http://doi.org/10.1371/journal.ppat.1002470

Nishida, K. M., Saito, K., Mori, T., Kawamura, Y., Nagami-Okada, T., Inagaki, S., … Siomi, M. C. (2007). Gene silencing mechanisms mediated by Aubergine piRNA

complexes in Drosophila male gonad. *RNA (New York, N.Y.)*, *13*(11), 1911–22. http://doi.org/10.1261/rna.744307

O'Brochta, D. A. (2003). Gene vector and transposable element behavior in mosquitoes. *Journal of Experimental Biology*, *206*(21), 3823–3834. http://doi.org/10.1242/jeb.00638

O'Brochta, D. A., Alford, R. T., Pilitt, K. L., Aluvihare, C. U., & Harrell, R. A. (2011). piggyBac transposon remobilization and enhancer detection in Anopheles mosquitoes. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(39), 16339–44. http://doi.org/10.1073/pnas.1110628108

O'brochta, D. A., & Handler, A. M. (1988). Mobility of P elements in drosophilids and nondrosophilids. *Proceedings of the National Academy of Sciences of the United States of America*, *85*(16), 6052–6.

O'Brochta, D. A., Pilitt, K. L., Harrell, R. A., Aluvihare, C., & Alford, R. T. (2012). Gal4-based Enhancer-Trapping in the Malaria Mosquito Anopheles stephensi. *G3: Genes|Genomes|Genetics*, *2*(11), 1305–1315. http://doi.org/10.1534/g3.112.003582

O'Brochta, D. A., Sethuraman, N., Wilson, R., Hice, R. H., Pinkerton, A. C., Levesque, C. S., … Atkinson, P. W. (2003). Gene vector and transposable element behavior in mosquitoes. *The Journal of Experimental Biology*, *206*(Pt 21), 3823–34.

O'Hare, K., & Rubin, G. M. (1983). Structures of P transposable elements and their sites of insertion and excision in the Drosophila melanogaster genome. *Cell*, *34*(1), 25–35.

Oliveira de Carvalho, M., Silva, J. C., & Loreto, E. L. S. (2004). Analyses of P-like transposable element sequences from the genome of Anopheles gambiae. *Insect Molecular Biology*, *13*(1), 55–63. http://doi.org/10.1111/j.1365-2583.2004.00461.x

Pal-Bhadra, M., Leibovitch, B. A., Gandhi, S. G., Chikka, M. R., Rao, M., Bhadra, U., … Elgin, S. C. R. (2004). Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. *Science (New York, N.Y.)*, *303*(5658), 669–72. http://doi.org/10.1126/science.1092653

Park, Y.-N., Masison, D., Eisenberg, E., & Greene, L. E. (2011). Application of the FLP/FRT system for conditional gene deletion in yeast Saccharomyces cerevisiae. *Yeast (Chichester, England)*, *28*(9), 673–81. http://doi.org/10.1002/yea.1895

Pavlopoulos, A., Oehler, S., Kapetanaki, M. G., & Savakis, C. (2007). The DNA transposon Minos as a tool for transgenesis and functional genomic analysis in vertebrates and invertebrates. *Genome Biology*, *8 Suppl 1*, S2. http://doi.org/10.1186/gb-2007-8-s1-s2

Prevention, C.-C. for D. C. and. (n.d.). CDC - Malaria - About Malaria - History - Ross and the Discovery that Mosquitoes Transmit Malaria Parasites.

Pritham, E. J., Putliwala, T., & Feschotte, C. (2007). Mavericks, a novel class of giant transposable elements widespread in eukaryotes and related to DNA viruses. *Gene*, *390*(1-2), 3–17. http://doi.org/10.1016/j.gene.2006.08.008

Rajasethupathy, P., Antonov, I., Sheridan, R., Frey, S., Sander, C., Tuschl, T., & Kandel, E. R. (2012). A role for neuronal piRNAs in the epigenetic control of memory-related synaptic plasticity. *Cell*, *149*(3), 693–707. http://doi.org/10.1016/j.cell.2012.02.057

Richardson, J. M., Colloms, S. D., Finnegan, D. J., & Walkinshaw, M. D. (2009). Molecular architecture of the Mos1 paired-end complex: the structural basis of DNA transposition in a eukaryote. *Cell*, *138*(6), 1096–108. http://doi.org/10.1016/j.cell.2009.07.012

Saito, K., Nishida, K. M., Mori, T., Kawamura, Y., Miyoshi, K., Nagami, T., … Siomi, M. C. (2006). Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. *Genes & Development*, *20*(16), 2214–22. http://doi.org/10.1101/gad.1454806

SanMiguel, P., Gaut, B. S., Tikhonov, A., Nakajima, Y., & Bennetzen, J. L. (1998). The paleontology of intergene retrotransposons of maize. *Nature Genetics*, *20*(1), 43–5. http://doi.org/10.1038/1695

Sarkar, A., Coates, C. J., Whyard, S., Willhoeft, U., Atkinson, P. W., & O'Brochta, D. A. (1997). The Hermes element from Musca domestica can transpose in four families of cyclorrhaphan flies. *Genetica*, *99*(1), 15–29.

Sarkar, A., Yardley, K., Atkinson, P. W., James, A. A., & O'Brochta, D. A. (1997). Transposition of the Hermes element in embryos of the vector mosquito, Aedes aegypti. *Insect Biochemistry and Molecular Biology*, *27*(5), 359–63.

Sauer, B., & Henderson, N. (1988). Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *Proceedings of the National Academy of Sciences of the United States of America*, *85*(14), 5166–70.

Savitsky, M., Kwon, D., Georgiev, P., Kalmykova, A., & Gvozdev, V. (2006). Telomere elongation is under the control of the RNAi-based mechanism in the Drosophila germline. *Genes & Development*, *20*(3), 345–54. http://doi.org/10.1101/gad.370206

Schmidt, T. (n.d.). LINEs, SINEs and repetitive DNA: non-LTR retrotransposons in plant genomes. *Plant Molecular Biology*, *40*(6), 903–910. http://doi.org/10.1023/A:1006212929794

Senti, K.-A., & Brennecke, J. (2010). The piRNA pathway: a fly's perspective on the guardian of the genome. *Trends in Genetics : TIG*, *26*(12), 499–509. http://doi.org/10.1016/j.tig.2010.08.007

Sepp, K. J., & Auld, V. J. (1999). Conversion of lacZ enhancer trap lines to GAL4 lines using targeted transposition in Drosophila melanogaster. *Genetics*, *151*(3), 1093–101.

Sheen, F. M., & Levis, R. W. (1994). Transposition of the LINE-like retrotransposon TART to Drosophila chromosome termini. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(26), 12510–4.

Siomi, M. C., Sato, K., Pezic, D., & Aravin, A. A. (2011a). PIWI-interacting small RNAs: the vanguard of genome defence. *Nature Reviews. Molecular Cell Biology*, *12*(4), 246–58. http://doi.org/10.1038/nrm3089

Siomi, M. C., Sato, K., Pezic, D., & Aravin, A. A. (2011b). PIWI-interacting small RNAs: the vanguard of genome defence. *Nature Reviews. Molecular Cell Biology*, *12*(4), 246–58. http://doi.org/10.1038/nrm3089

Slotkin, R. K., & Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. *Nature Reviews. Genetics*, *8*(4), 272–85. http://doi.org/10.1038/nrg2072

Smit, A. F. (1999). Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Current Opinion in Genetics & Development*, *9*(6), 657–63.

Smith, R. C., & Atkinson, P. W. (2011). Mobility properties of the Hermes transposable element in transgenic lines of Aedes aegypti. *Genetica*, *139*(1), 7–22. http://doi.org/10.1007/s10709-010-9459-7

Spradling, A. C., Stern, D., Beaton, A., Rhem, E. J., Laverty, T., Mozden, N., … Rubin, G. M. (1999). The Berkeley Drosophila Genome Project gene disruption project: Single P-element insertions mutating 25% of vital Drosophila genes. *Genetics*, *153*(1), 135–77.

Subramanian, R. A., Akala, O. O., Adejinmi, J. O., & O'Brochta, D. A. (2008). Topi, an IS630/Tc1/mariner-type transposable element in the African malaria mosquito, Anopheles gambiae. *Gene*, *423*(1), 63–71. http://doi.org/10.1016/j.gene.2008.06.023

Subramanian, R. A., Cathcart, L. A., Krafsur, E. S., Atkinson, P. W., & O'Brochta, D. A. Hermes transposon distribution and structure in Musca domestica. *The Journal of Heredity*, *100*(4), 473–80. http://doi.org/10.1093/jhered/esp017

Sundararajan, P., Atkinson, P. W., & O'Brochta, D. A. (1999). Transposable element interactions in insects: crossmobilization of hobo and Hermes. *Insect Molecular Biology*, *8*(3), 359–68.

Thibault, S. T., Singer, M. A., Miyazaki, W. Y., Milash, B., Dompe, N. A., Singh, C. M., … Margolis, J. (2004). A complementary transposon tool kit for Drosophila melanogaster using P and piggyBac. *Nature Genetics*, *36*(3), 283–7. http://doi.org/10.1038/ng1314

Thorpe, H. M., & Smith, M. C. (1998). In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(10), 5505–10.

Thyagarajan, B., Olivares, E. C., Hollis, R. P., Ginsburg, D. S., & Calos, M. P. (2001). Site-specific genomic integration in mammalian cells mediated by phage phiC31 integrase. *Molecular and Cellular Biology*, *21*(12), 3926–34. http://doi.org/10.1128/MCB.21.12.3926-3934.2001

Uchino, K., Imamura, M., Shimizu, K., Kanda, T., & Tamura, T. (2007). Germ line transformation of the silkworm, Bombyx mori, using the transposable element Minos. *Molecular Genetics and Genomics : MGG*, *277*(3), 213–20. http://doi.org/10.1007/s00438-006-0176-y

Van den Berg, H. (2009). Global status of DDT and its alternatives for use in vector control to prevent disease. *Environmental Health Perspectives*, *117*(11), 1656–63. http://doi.org/10.1289/ehp.0900785

Venken, K. J. T., Schulze, K. L., Haelterman, N. A., Pan, H., He, Y., Evans-Holm, M., … Bellen, H. J. (2011). MiMIC: a highly versatile transposon insertion resource for engineering Drosophila melanogaster genes. *Nature Methods*, *8*(9), 737–43.

Vigdal, T. J., Kaufman, C. D., Izsvák, Z., Voytas, D. F., & Ivics, Z. (2002). Common physical properties of DNA affecting target site selection of sleeping beauty and other Tc1/mariner transposable elements. *Journal of Molecular Biology*, *323*(3), 441–52.

Wessler, S. (1995). LTR-retrotransposons and MITEs: important players in the evolution of plant genomes. *Current Opinion in Genetics & Development*, *5*(6), 814–821. http://doi.org/10.1016/0959-437X(95)80016-X

Wightman, B., Burglin, T. R., Gatto, J., Arasu, P., & Ruvkun, G. (1991). Negative regulatory sequences in the lin-14 3'-untranslated region are necessary to generate a temporal switch during Caenorhabditis elegans development. *Genes & Development*, *5*(10), 1813–1824. http://doi.org/10.1101/gad.5.10.1813

Wilson, C., Pearson, R. K., Bellen, H. J., O'Kane, C. J., Grossniklaus, U., & Gehring, W. J. (1989). P-element-mediated enhancer detection: an efficient method for isolating and characterizing developmentally regulated genes in Drosophila. *Genes & Development*, *3*(9), 1301–13.

Wilson, J. E., Connell, J. E., & Macdonald, P. M. (1996). aubergine enhances oskar translation in the Drosophila ovary. *Development (Cambridge, England)*, *122*(5), 1631–9.

Wilson, M. H., Coates, C. J., & George, A. L. (2007). PiggyBac transposon-mediated gene transfer in human cells. *Molecular Therapy : The Journal of the American Society of Gene Therapy*, *15*(1), 139–45. http://doi.org/10.1038/sj.mt.6300028

Winskill, P., Harris, A. F., Morgan, S. A., Stevenson, J., Raduan, N., Alphey, L., … Donnelly, C. A. (2014). Genetic control of Aedes aegypti: data-driven modelling to assess the effect of releasing different life stages and the potential for long-term suppression. *Parasites & Vectors*, *7*, 68. http://doi.org/10.1186/1756-3305-7-68

Yant, S. R., Meuse, L., Chiu, W., Ivics, Z., Izsvak, Z., & Kay, M. A. (2000). Somatic integration and long-term transgene expression in normal and haemophilic mice using a DNA transposon system. *Nature Genetics*, *25*(1), 35–41. http://doi.org/10.1038/75568

Zamore, P. D., Tuschl, T., Sharp, P. A., & Bartel, D. P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell*, *101*(1), 25–33. http://doi.org/10.1016/S0092-8674(00)80620-0

Zhai, R. G., Hiesinger, P. R., Koh, T.-W., Verstreken, P., Schulze, K. L., Cao, Y., … Bellen, H. J. (2003). Mapping Drosophila mutations with molecularly defined P element insertions. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(19), 10860–5. http://doi.org/10.1073/pnas.1832753100

Zhao, Y., & Eggleston, P. (1998). Stable transformation of an Anopheles gambiae cell line mediated by the Hermes mobile genetic element. *Insect Biochemistry and Molecular Biology*, *28*(4), 213–219. http://doi.org/10.1016/S0965-1748(98)00020-4

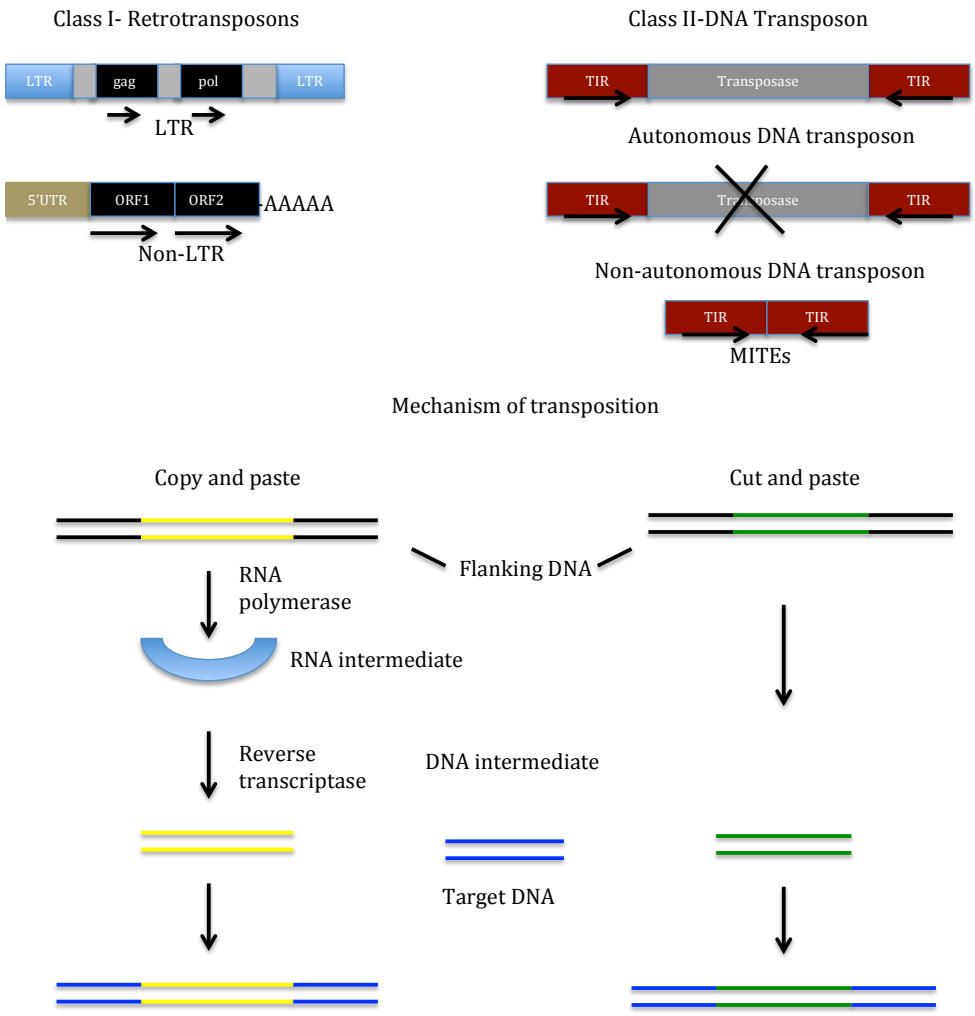


Figure 1.1 Types of transposons and their mechanism of transposition; Class I-Reteroelements, and Class II-DNA transposons. Adapted from (Slotkin & Martienssen, 2007).

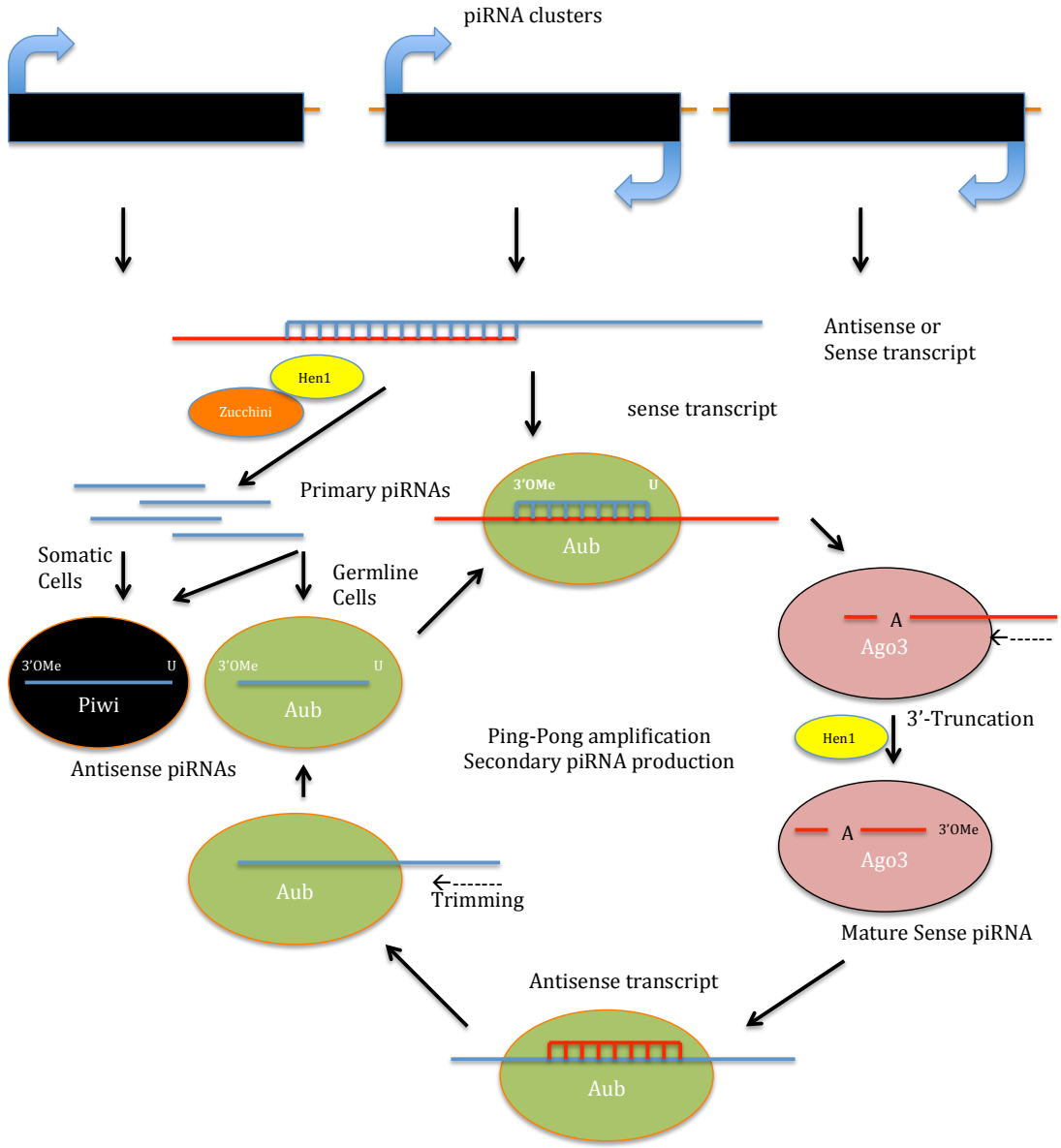


Figure 1.2 piRNA biogenesis and transposon silencing in somatic and germline cells. Adapted from Jaspreet Khurana and William Theurkauf, Cell Biology 2010

Chapter 2

Identification of new class II DNA transposons in *Ae. aegypti*

2.1 Abstract

The most common problem of using exogenous TEs in *Ae. aegypti* genetics has been lack of activity of these TEs relative to other species and the inability to remobilize them once they have integrated into the genome (O'Brochta, 2003). The aim of this project is the identification of potentially active transposons in the *Ae. aegypti* genome through computational analysis and the subsequent determination of whether they are biologically active. The rationale is that an active endogenous element may have remained active because it has overcome or evaded any suppression systems that may lead to transposon inactivation. To find potential active candidates from the genome sequence, consensus sequences with a conserved catalytic domain from many different transposon superfamilies were used as a query to search in the *Ae. aegypti* genome. A phylogenetic tree was constructed and sequences showing identity greater than 99% were considered as potentially active elements. Two different computational pipelines MITE-Hunter (Han & Wessler, 2010) and TARGeT (Han, Burnette, & Wessler, 2009) were used to identify new active transposable elements in *Ae. aegypti*. With this analysis *Muta1* was identified, which was subsequently shown to be capable of transforming *Ae. aegypti* supports this as could be definitive proof of remobilization.

2.2 Introduction

Computational approaches have been used to characterize and discover TEs in various genomes (Xu et al., 2012), TE discovery is suitable for large scale bioinformatics analysis: they are repetitive and have structural signatures such as TIRs, and conserved transposases domains (Janicki, Rooke, & Yang, 2011). The major tools that are used for the discovery of TEs can be grouped into three categories; homology-based, signature-based and *de novo* methods (Lerat, 2010). Homology-based methods search for sequences that relate to known TEs, that have been deposited and annotated in databases such as RepBase, ENSEMBL, GENBANK and TEfam, but these approaches are not useful for discovery of non-autonomous elements and miniature inverted repeat terminal elements (MITEs), which lack consensus transposase sequences used for homology-based searches of TEs (Han & Wessler, 2010).

Signature-based approaches discover TEs based on known conserved regions, amino acid sequences and motifs (Saha, Bridges, Magbanua, & Peterson, 2008). MITEs, as they have same signature set as DNA TEs, have been discovered using signature-based identification (Feschotte, Swamy, & Wessler, 2003). However, signature-based identification of MITEs has resulted in high false positive rates due to the complexity of the higher eukaryotes and the abundance of MITEs in their genomes leading to the necessity of manually annotation (Saha et al., 2008). *De novo* approaches discover new repeats or repeat families based on the repetitive nature of TEs. The output of *de novo*

methods is mixture of TEs from all superfamilies and non-TE repeats, making annotation of putative TEs difficult (Flutre, Duprat, Feuillet, & Quesneville, 2011).

MITEs are a special class of non-autonomous class II element that are abundant in many eukaryotes; they are found in non-coding region of plants and many animal species, and have very high copy numbers in the genome (Han & Wessler, 2010). MITEs have no coding potential, are relatively short (50-800 bp), are AT rich, possess TIRs less than 20 bp, and share very little sequence similarity to DNA TEs, (Bureau, 1994). Studies on *TOURIST* type MITEs led to the discovery of a new type of TE superfamily called *PIF/Harbinger* in plants (X. Zhang et al., 2001). The *STOWAWAY* MITEs were found to be related to *Tc1/mariner* superfamily elements that are found in plants and animal species (Bureau, 1994). The *Stowaway*-like MITEs and *Mariner*-like elements (MLEs) share similarities in their TIRs and TSDs that strongly suggested that *Stowaway* MITEs were mobilized by transposase encoded by MLEs (Feschotte et al., 2003). MITEs also have the ability to be cross mobilized. For example, MITEs from the *STOWAWAY* family were cross mobilized by the *Osmar* transposase (Guojun Yang, Nagel, Feschotte, Hancock, & Wessler, 2009a). Similarly, in rice, *mPing* can be remobilized by the related element *Pong* (Jiang et al., 2003). MITEs can be a powerful factor for promoting intra and inter species variability through insertion into genes by which they can cause changes in expression, such as with the *Stowaway* MITE named *dTstu*1 insertion into a flavonoid 3', 5' hydroxylase gene in potato which leads to red pigmentation (Momose, Abe, & Ozeki, 2010), and *mPing* MITE insertion into a *H*d1 gene which causes change in

flowering time (Yano et al., 2000) (Jiang et al., 2003). MITEs have been frequently associated with host genes in mosquitoes (Tu, 1997), where they show the bias for integration into non-coding regions same as plants.

A direct relationship between MITEs and class II elements was discovered through studies in plants (Feschotte et al., 2003) (Xiaoyu Zhang, Jiang, Feschotte, & Wessler, 2004) (Feschotte, Osterlund, Peeler, & Wessler, 2005) (G. Yang, Zhang, Hancock, & Wessler, 2007) (Guojun Yang, Nagel, Feschotte, Hancock, & Wessler, 2009b). An active, full length autonomous element that encodes a transposase might be required for mobilization of a MITE (Deprá, Ludwig, Valente, & Loreto, 2012). Studies in *An. gambiae* have identified *P* TEs sequences with related MITEs families, based on similarities within TIRs of the *P* TEs and the associated MITEs (Quesneville, Nouaud, & Anxolabéhère, 2006). Determining the abundance of MITEs in a genome and identifying the transposons are required for their movement may reveal novel active (or recently active) transposons that could be used as foundation for the development of more robust transposable element-based genetic technologies in mosquitoes.

2.3 Material and Methods

2.3.1 The TARGeT pipeline

The *Ae. aegypti* whole genome scaffolds were used for the discovery of consensus sequences of TEs using MITE-Hunter pipeline by Kun Liu and Susan R. Wessler, (Plant

Biology graduate Program, University of California Riverside). I then used *hAT, Mule and Novosib* superfamily consensus sequences for further analysis.

TARGeT is a program that streamlines the process of retrieving, annotating and analyzing transposase superfamilies from genomic databases (Han et al., 2009). TARGeT can use either protein or DNA sequences as the query. BLASTN searches are used for DNA queries, while TBLASTN is used for protein queries. In this analysis, TBLASTN was used for protein queries generated by the MITE-Hunter program. TARGeT uses MUSCLE (Edgar, 2004) to calculate the multiple alignment and the TreeBest (Vilella et al., 2009) program, which correlate known gene families to generate a phylogenetic tree of the putative homologs with the neighbor-joining method (Han et al., 2009). The TARGeT pipeline has four main steps (Figure 2.2) (I). The output consensus sequences from the MITE-Hunter program generated by Kun Liu were used for identifying potential autonomous elements, which gave rise to short non-autonomous TE families. The BLAST searches were performed with specified parameters, which were based on identification of homologs that show high similarity to the *Ae. aegypti* genome. The following BLAST parameters were modified: expectation value to 0.1, number of hits required to 1000, and the input query that used to search the *Ae. aegypti* genome was filtered out from the output result. For PHI, minimum percentage of query was changed from 0.7 to 0.3, the maximum number of output was changed to 1000, the expectation value set to 0.1, the length of flanking sequence to 10000 and the maximum number of output homologs to 1000. The default values for any unspecified parameters were used.

The BLAST output represents the match of a query to the genome. The darker region was indicative of a good match, (II) the PHI output shows number of hits to the query, and the length of the transposase. Flanking sequence to homologs were obtained from the view tab. Long alignments were plus to plus strand, since PHI converts all reads to the plus strand if the TE is located on minus strand. TIRs and TSDs were discovered through manual screening of flanking sequences of the homologs from the output file, TIRs were discovered 1000 bp from the match and TSDs 100 bp from the ends, (III) homologs obtained from the PHI output were then used for carrying out multiple sequence alignments (MSA) to determine the percentage match. I reasoned that, if homologs show greater than 99% identity, then they, might be potentially active TEs, (IV) a phylogenetic tree is built, which represents the distribution of the hits for a particular query based on the match. Smaller branches represent greater homology to the match. The final step to determine the coding ability of a newly discovered TE from the TARGeT pipeline was performed using GENESCAN (Burge & Karlin, 1997). Active TEs must have a transposase gene with at least several hundred amino acids long that has a complete and intact ORF. These homologs generated from the TARGeT pipeline were manually examined for TSDs and TIRs to discover new TEs.

2. 4 Results

2.4.1 TEs discovered through TARGeT pipeline.

TARGeT analysis was performed for the *hAT*, *Mule* and *Novosib* superfamilies; seven intact autonomous elements were identified (Table 2.1). From the analysis of TEs

that belonged to the *hAT* superfamily, I was able to discover autonomous elements that contained characteristics of a *hAT* element, formed 8 bp TSDs and short TIRs (Table 2.1) (Rubin, Lithwick, & Levy, 2001). The genomic DNA of the transposase of these elements was analyzed using GENESCAN, which predicts gene structures such as exons and introns (Burge & Karlin, 1997). The NCBI Blast tool was used to detect conserved regions such as the BED zinc finger binding domain, the catalytic domain, and the triad containing at least first two carboxylates amino acid residues (Arensburger et al., 2011). A sequence with the high similarity to the Tam3 consensus sequences produced 8 bp TSDs and TIRs of greater than 13 bp in length, and was later determined to be *AeHerves2*, which had previously been identified and deposited in RepBase (Table 2.1) (Jurka, 1998) (Jurka et al., 2005). A number of other conserved sequences obtained from the MITE-Hunter output discovered the same element with a different 8 bp TSD in the genome of *Ae. aegypti* (Table 2.1). Another *hAT* element that was discovered and determined to be present in RepBase database was *AeBuster1*, which had been tested for its transposition activity in *D. melanogaster* and *Ae. aegypti* (Figure 2.5) (Arensburger et al., 2011). Two other *hAT* elements that were discovered were identified from the consensus sequences of the hATx3_SM and hAT-1_AA (Table 2.1). The *hAT* element discovered using consensus sequence hATx3_SM formed 8 bp TSDs, and had 19 bp TIRs. The analysis of the transposase domains using the NCBI Blast tool showed it contained *hAT* superfamily BED zinc finger DNA binding domain (Figure 2.6). The element discovered using hAT-1_AA consensus sequence formed 8 bp TSDs, and contained 11 bp TIRs. The analysis of the transposase domains showed it contained BED

zinc finger DNA binding domain (Figure 2.7). The analysis of the TSDs formed by this element suggests that *hAT-1_AA* is a member of *Ac* superfamily of *hAT* transposons, as it contains 5' nTnnnAn 3' consensus sequence (Arensburger et al., 2011). The analysis of the DDE motif for *hATx3_SM* and *hAT-1_AA* transposons was based on the DDE motif identified in other *hAT* transposons (Atkinson PW (2015) *hAT* Transposable Elements, in Mobile DNA III, Craig N. L. (senior editor) ASM Press, Washington D. C., 27 PP.) (Figure 2.5).

Three *Mule* superfamily elements were discovered. The *Mutator* element was discovered by Kun Liu and Susan R. Wessler, University of California, Riverside (Table 2.1), and working concurrently, I performed analysis and characterization of this transposon. The *Muta1* element was discovered using consensus sequences of MuDR_1_TV, and 8 copies with greater than 99% identity in the genome were identified. Two out of eight elements had an 8 bp TSD and other six elements formed 9 bp TSD (Figure 2.4). All eight copies formed 146 bp TIRs. The transposase of this element was analyzed using GENESCAN and NCBI Blast tool. It contained two exons and an intron with the FLYWCH domain and a catalytic domain. The FLYWCH domain is found in most *Mutator* transposases (Babu, Iyer, Balaji, & Aravind, 2006) and functions in DNA binding for transposition (Figure 2.3). *Muta12* was discovered using the consensus sequences of MuDR7x_AP, it has 59 bp TIRs and forms 9 bp TSDs (Figure 2.8), while *Muta3* was discovered using consensus sequences of MuDR4x_SM, has 16 bp TIRs and 8 bp TSDs (Figure 2.9). The analysis of the DDE motif for *Muta12* and *Muta3* was based on the DDE motif identified

in other *Mutator* transposons (Yuan & Wessler, 2011). I did not discover TEs using the consensus sequences for *Novosib* superfamily in *Ae. aegypti*.

2.5 Discussion

The bioinformatics analysis identified five new DNA transposons excluding *AeBuster1* and *AeHerves2* whose activities have already been determined in *Ae. aegypti* (Arensburger et al., 2005; 2011). Two of the TEs belonged to the *hAT* superfamily and the other three belonged to the *Mule/Mutator* superfamily. Most of the novel TEs discovered had more than one copy in the genome of *Ae. aegypti,* and these TEs have been discovered and deposited into RepBase. However none have been experimentally tested to determine their transposition frequency in *Ae. aegypti*.

Previous studies for discovery of *hAT* elements was based on transposons and sequences related to *hobo*, *Tol2*, *Ac3* and *Hermes* (Kempken & Windhofer, 2001) (Rubin et al., 2001). In this study, I have focused on the discovery of consensus sequence based on the TIRs and TSDs of a MITE, as MITEs are found in abundance in the genome of *Ae. aegypti* and this characteristic can be applied for generating consensus sequences for various superfamily. Computational approaches to discover new *hAT* elements in *Ae. aegypti* resulted in large number of consensus sequences, which is indicative that new *hAT* elements are present in the *Ae. aegypti,* which are yet to be discovered. The *hAT* superfamily of DNA transposons have showed to play a role in genome evolution

through domestication within the host genes retaining their catalytic and DNA binding domains (Sinzelle, Izsvák, & Ivics, 2009). The identification of *hAT* elements *AeHerves2* and *AeBuster1* that were discovered previously suggests that the two bioinformatics pipeline were efficient in discovery of novel TEs in the genome of *Ae. aegypti.* The *hAT-1_AA* element was determined to be a member of a *Ac* family of TEs based on the consensus sequence of TSDs, moreover further analysis of the transposase domain confirmed that it contained catalytic domain with DDE motif and a Zinc finger binding, which is a DNA binding domain found in *hAT* elements (Arensburger et al., 2011). The *hATx3_SM* TE has 5'nnnnAn 3' TSDs, due to the unusual TSD it could not be assigned to the *Ac* or *Buster* family of *hAT* transposons. Moreover, there was a 19 bp gap between the left end TIR and the TSD. The study on the identification of *Hobo* element in *Drosophila*, which is a *hAT* element, was found to have an additional 12 bp inverted repeat of left end in reverse orientation, 250 bp downstream of the first left end TIR (McGinnis, Shermoen, & Beckendorf, 1983). Thus, it is possible that the *hATx3_SM* might be a potentially active *hAT* element, further analysis need to be performed to determine its activity through somatic transposition assays.

Mutator elements are wide-spread in plants (Talbert & Chandler, 1988), but they have also been discovered in other eukaryotes (Chalvet, Grimaldi, Kaper, Langin, & Daboussi, 2003). A group of *Phantom* elements, which belongs to a *Mutator* superfamily, have been discovered in wide variety of organisms including two insect viruses that are known to infect wasps (Marquez & Pritham, 2010). Moreover, *Phantom*-like proteins have also

been identified in *Ae. aegypti* (Marquez & Pritham, 2010). The identification of *Phantom* elements in these insect viruses was taken to be the evidence of horizontal movement of TEs in eukaryotes (Piskurek & Okada, 2007), (Lerch & Friesen, 1992). In the analysis of *Mutator/Mule* superfamily of transposons Kun Liu and Susan R. Wessler discovered *Muta1* elements in *Ae. aegypti.* I also found some of these elements in studies performed concurrently. Moreover, Kun Liu also determined that there are eight full length copies of *Muta1* element that were similar to each other suggesting that the *Muta1* element could be potentially active in the genome of *Ae. aegypti*. These data thus identify the *Mutator* superfamily TEs in *Ae. aegypti.*

Analysis of the overall structure of *Muta1* element resembles a *Foldback (*FB) element, *Galileo;* with seven 11 bp repeats in the left end and six 11 bp repeats in the right end (Figure 2.3) A similar structure was identified in *Phantom* elements. In fact they have three structural variants for TIRs: TIRs, sub-TIRs, and FB-TIRs (Marquez & Pritham, 2010). *Phantom* elements with TIR have a structure typical to *Mutator* elements with inverted terminal repeat flanking to the ends, sub-TIR *Phantom* elements have non-repetitive structure flanking to the TIRs at the end and FB-TIR *Phantom* elements have long complicated TIRs with lots of repeats in the ends (Marquez & Pritham, 2010). TIRs of a TE are required for the cleavage and integration of DNA transposons through transposase binding. Increasing the number of transposase binding sites might eventually increase the transposition activity of a transposon (Potter, 1982) and its is hypothesized that variation in structure of TIR could be to avoid detection by host for silencing through

formation of hairpin loop structure that would evade its detection by proteins of a host silencing machinery for recognition (Marquez & Pritham, 2010) (Gross & Williamson, 2011).

2.6 References

Arensburger, P., Hice, R. H., Zhou, L., Smith, R. C., Tom, A. C., Wright, J. A., … Atkinson, P. W. (2011). Phylogenetic and functional characterization of the hAT transposon superfamily. *Genetics*, *188*(1), 45–57. http://doi.org/10.1534/genetics.111.126813

Babu, M. M., Iyer, L. M., Balaji, S., & Aravind, L. (2006). The natural history of the WRKY-GCM1 zinc fingers and the relationship between transcription factors and transposons. *Nucleic Acids Research*, *34*(22), 6505–20. http://doi.org/10.1093/nar/gkl888

Bureau, T. E. (1994). Stowaway: A New Family of Inverted Repeat Elements Associated with the Genes of Both Monocotyledonous and Dicotyledonous Plants. *THE PLANT CELL ONLINE*, *6*(6), 907–916. http://doi.org/10.1105/tpc.6.6.907

Burge, C., & Karlin, S. (1997). Prediction of complete gene structures in human genomic DNA. *Journal of Molecular Biology*, *268*(1), 78–94. http://doi.org/10.1006/jmbi.1997.0951

Chalvet, F., Grimaldi, C., Kaper, F., Langin, T., & Daboussi, M.-J. (2003). Hop, an active Mutator-like element in the genome of the fungus Fusarium oxysporum. *Molecular Biology and Evolution*, *20*(8), 1362–75. http://doi.org/10.1093/molbev/msg155

Deprá, M., Ludwig, A., Valente, V. L., & Loreto, E. L. (2012). Mar, a MITE family of hAT transposons in Drosophila. *Mobile DNA*, *3*(1), 13. http://doi.org/10.1186/1759-8753-3-13

Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, *5*, 113. http://doi.org/10.1186/1471-2105-5-113

Feschotte, C., Osterlund, M. T., Peeler, R., & Wessler, S. R. (2005). DNA-binding specificity of rice mariner-like transposases and interactions with Stowaway MITEs. *Nucleic Acids Research*, *33*(7), 2153–65. http://doi.org/10.1093/nar/gki509

Feschotte, C., Swamy, L., & Wessler, S. R. (2003). Genome-wide analysis of mariner-like transposable elements in rice reveals complex relationships with stowaway miniature inverted repeat transposable elements (MITEs). *Genetics*, *163*(2), 747–58.

Flutre, T., Duprat, E., Feuillet, C., & Quesneville, H. (2011). Considering transposable element diversification in de novo annotation approaches. *PloS One*, *6*(1), e16526. http://doi.org/10.1371/journal.pone.0016526

Gross, S. M., & Williamson, V. M. (2011). Tm1: a mutator/foldback transposable element family in root-knot nematodes. *PloS One*, *6*(9), e24534. http://doi.org/10.1371/journal.pone.0024534

Han, Y., Burnette, J. M., & Wessler, S. R. (2009). TARGeT: a web-based pipeline for retrieving and characterizing gene and transposable element families from genomic sequences. *Nucleic Acids Research*, *37*(11), e78. http://doi.org/10.1093/nar/gkp295

Han, Y., & Wessler, S. R. (2010). MITE-Hunter: A program for discovering miniature inverted-repeat transposable elements from genomic sequences. *Nucleic Acids Research*, *38*(22).

Janicki, M., Rooke, R., & Yang, G. (2011). Bioinformatics and genomic analysis of transposable elements in eukaryotic genomes. *Chromosome Research*, *19*(6), 787–808. http://doi.org/10.1007/s10577-011-9230-7

Jiang, N., Bao, Z., Zhang, X., Hirochika, H., Eddy, S. R., McCouch, S. R., & Wessler, S. R. (2003). An active DNA transposon family in rice. *Nature*, *421*(6919), 163–167. http://doi.org/10.1038/nature01214

Jurka, J. (1998). Repeats in genomic DNA: mining and meaning. *Current Opinion in Structural Biology*, *8*(3), 333–7.

Jurka, J., Kapitonov, V. V, Pavlicek, A., Klonowski, P., Kohany, O., & Walichiewicz, J. (2005). Repbase Update, a database of eukaryotic repetitive elements. *Cytogenetic and Genome Research*, *110*(1-4), 462–7. http://doi.org/10.1159/000084979

Kempken, F., & Windhofer, F. (2001). The hAT family: a versatile transposon group common to plants, fungi, animals, and man. *Chromosoma*, *110*(1), 1–9.

Lerat, E. (2010). Identifying repeats and transposable elements in sequenced genomes: how to find your way through the dense forest of programs. *Heredity*, *104*(6), 520–33. http://doi.org/10.1038/hdy.2009.165

Lerch, R. A., & Friesen, P. D. (1992). The baculovirus-integrated retrotransposon TED encodes gag and pol proteins that assemble into viruslike particles with reverse transcriptase. *Journal of Virology*, *66*(3), 1590–601.

Marquez, C. P., & Pritham, E. J. (2010). Phantom, a new subclass of Mutator DNA transposons found in insect viruses and widely distributed in animals. *Genetics*, *185*(4), 1507–17. http://doi.org/10.1534/genetics.110.116673

McGinnis, W., Shermoen, A. W., & Beckendorf, S. K. (1983). A transposable element inserted just 5′ to a Drosophila glue protein gene alters gene expression and chromatin structure. *Cell*, *34*(1), 75–84. http://doi.org/10.1016/0092-8674(83)90137-X

Momose, M., Abe, Y., & Ozeki, Y. (2010). Miniature inverted-repeat transposable elements of Stowaway are active in potato. *Genetics*, *186*(1), 59–66. http://doi.org/10.1534/genetics.110.117606

O'Brochta, D. A. (2003). Gene vector and transposable element behavior in mosquitoes. *Journal of Experimental Biology*, *206*(21), 3823–3834. http://doi.org/10.1242/jeb.00638

Piskurek, O., & Okada, N. (2007). Poxviruses as possible vectors for horizontal transfer of retroposons from reptiles to mammals. *Proceedings of the National Academy of Sciences*, *104*(29), 12046–12051. http://doi.org/10.1073/pnas.0700531104

Potter, S. S. (1982). DNA sequence of a foldback transposable element in Drosophila. *Nature*, *297*(5863), 201–4.

Quesneville, H., Nouaud, D., & Anxolabéhère, D. (2006). P elements and MITE relatives in the whole genome sequence of Anopheles gambiae. *BMC Genomics*, *7*, 214. http://doi.org/10.1186/1471-2164-7-214

Rubin, E., Lithwick, G., & Levy, A. A. (2001). Structure and evolution of the hAT transposon superfamily. *Genetics*, *158*(3), 949–57.

Saha, S., Bridges, S., Magbanua, Z. V., & Peterson, D. G. (2008). Computational Approaches and Tools Used in Identification of Dispersed Repetitive DNA Sequences. *Tropical Plant Biology*, *1*(1), 85–96. http://doi.org/10.1007/s12042-007-9007-5

Sinzelle, L., Izsvák, Z., & Ivics, Z. (2009). Molecular domestication of transposable elements: from detrimental parasites to useful host genes. *Cellular and Molecular Life Sciences : CMLS*, *66*(6), 1073–93. http://doi.org/10.1007/s00018-009-8376-3

Talbert, L. E., & Chandler, V. L. (1988). Characterization of a highly conserved sequence related to mutator transposable elements in maize. *Molecular Biology and Evolution*, *5*(5), 519–529.

Tu, Z. (1997). Three novel families of miniature inverted-repeat transposable elements are associated with genes of the yellow fever mosquito, Aedes aegypti. *Proceedings of the National Academy of Sciences*, *94*(14), 7475–7480. http://doi.org/10.1073/pnas.94.14.7475

Vilella, A. J., Severin, J., Ureta-Vidal, A., Heng, L., Durbin, R., & Birney, E. (2009). EnsemblCompara GeneTrees: Complete, duplication-aware phylogenetic trees in vertebrates. *Genome Research*, *19*(2), 327–35. http://doi.org/10.1101/gr.073585.107

Xu, H.-E., Zhang, H.-H., Han, M.-J., Shen, Y.-H., Huang, X.-Z., Xiang, Z.-H., & Zhang, Z. (2012). [Computational approaches for identification and classification of transposable elements in eukaryotic genomes]. *Yi Chuan = Hereditas / Zhongguo Yi Chuan Xue Hui Bian Ji*, *34*(8), 1009–19.

Yang, G., Nagel, D. H., Feschotte, C., Hancock, C. N., & Wessler, S. R. (2009a). Tuned for transposition: molecular determinants underlying the hyperactivity of a Stowaway MITE. *Science (New York, N.Y.)*, *325*(5946), 1391–4. http://doi.org/10.1126/science.1175688

Yang, G., Nagel, D. H., Feschotte, C., Hancock, C. N., & Wessler, S. R. (2009b). Tuned for transposition: molecular determinants underlying the hyperactivity of a Stowaway MITE. *Science (New York, N.Y.)*, *325*(5946), 1391–4. http://doi.org/10.1126/science.1175688

Yang, G., Zhang, F., Hancock, C. N., & Wessler, S. R. (2007). Transposition of the rice miniature inverted repeat transposable element mPing in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences*, *104*(26), 10962–10967. http://doi.org/10.1073/pnas.0702080104

Yano, M., Katayose, Y., Ashikari, M., Yamanouchi, U., Monna, L., Fuse, T., … Sasaki, T. (2000). Hd1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the Arabidopsis flowering time gene CONSTANS. *The Plant Cell*, *12*(12), 2473–2484.

Yuan, Y.-W., & Wessler, S. R. (2011). The catalytic domain of all eukaryotic cut-and-paste transposase superfamilies. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(19), 7884–9. http://doi.org/10.1073/pnas.1104208108

Zhang, X., Feschotte, C., Zhang, Q., Jiang, N., Eggleston, W. B., & Wessler, S. R. (2001). P instability factor: An active maize transposon system associated with the amplification of Tourist-like MITEs and a new superfamily of transposases. *Proceedings of the National Academy of Sciences*, *98*(22), 12572–12577. http://doi.org/10.1073/pnas.211442198

Zhang, X., Jiang, N., Feschotte, C., & Wessler, S. R. (2004). PIF- and Pong-like transposable elements: distribution, evolution and relationship with Tourist-like miniature inverted-repeat transposable elements. *Genetics*, *166*(2), 971–86.

Table 2.1 Transposons discovered using the TARGeT pipeline.

Superfamily	Consensus sequence	TSDs	TIRs
hAT	TAHAT7 (*AeHerves2*)	GTTGAGTA 8 bp	TAGAGATGGGCAA
	hAT_29_HM (*AeBuster1*)	GCTTATGG 8 bp	CATAGATTCCCA
	hATx-3_SM	CCCGGGAA 8 bp	AAATCCCGGGAATCC
	hAT-1_AA	CATCGAAT 8 bp	TAGAGTGTCCATGGA
Mule/Mutator	MuDr_1_TV (*Muta1*)	GACGCAGT 8/9 bp	GGGTCTACCCCGT
	MuDr7x_AP *(Muta12)*	TTCCCCCAT 9 bp	TAGGGCGGTTCACAA
	MuDR4x_SM *(Muta3*)	TCTAGAAG 8 bp	TAAGGGTATGCGAAAT

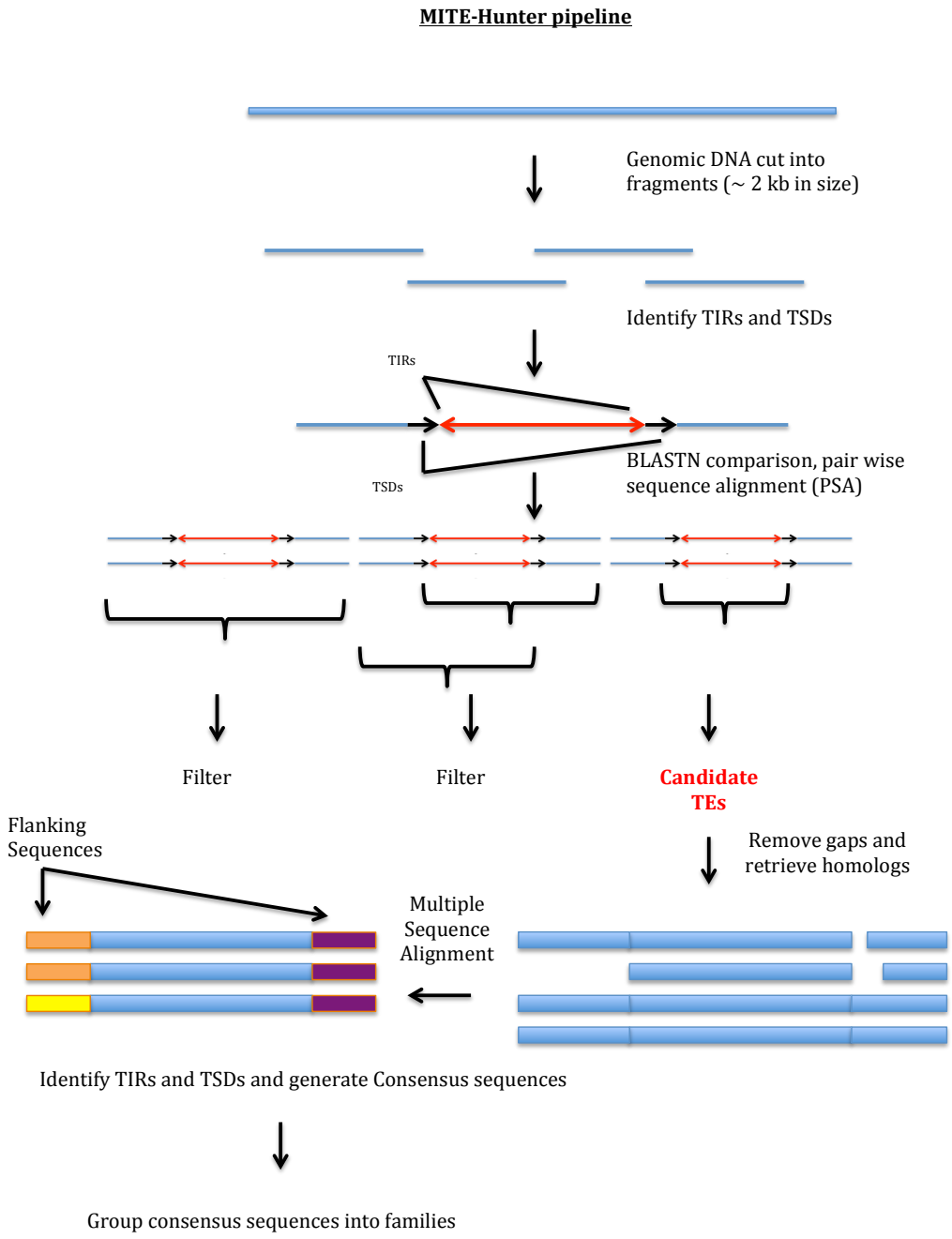


Figure 2.1 Steps involved in discovery of consensus sequences from a MITE-Hunter pipeline 1) The *Ae. aegypti* genome scaffolds were fragmented into approximately 2 kb fragments, 2) fragments were looked for TSDs and TIRs, 3) Candidate TE were identified based on pair wise sequence alignment, 3) gaps were removed from the alignments, 4) multiple alignment of homologs was performed, and 5) consensus sequences were generated, which were then grouped into superfamilies. (Adapter from Han & Wessler, 2010).

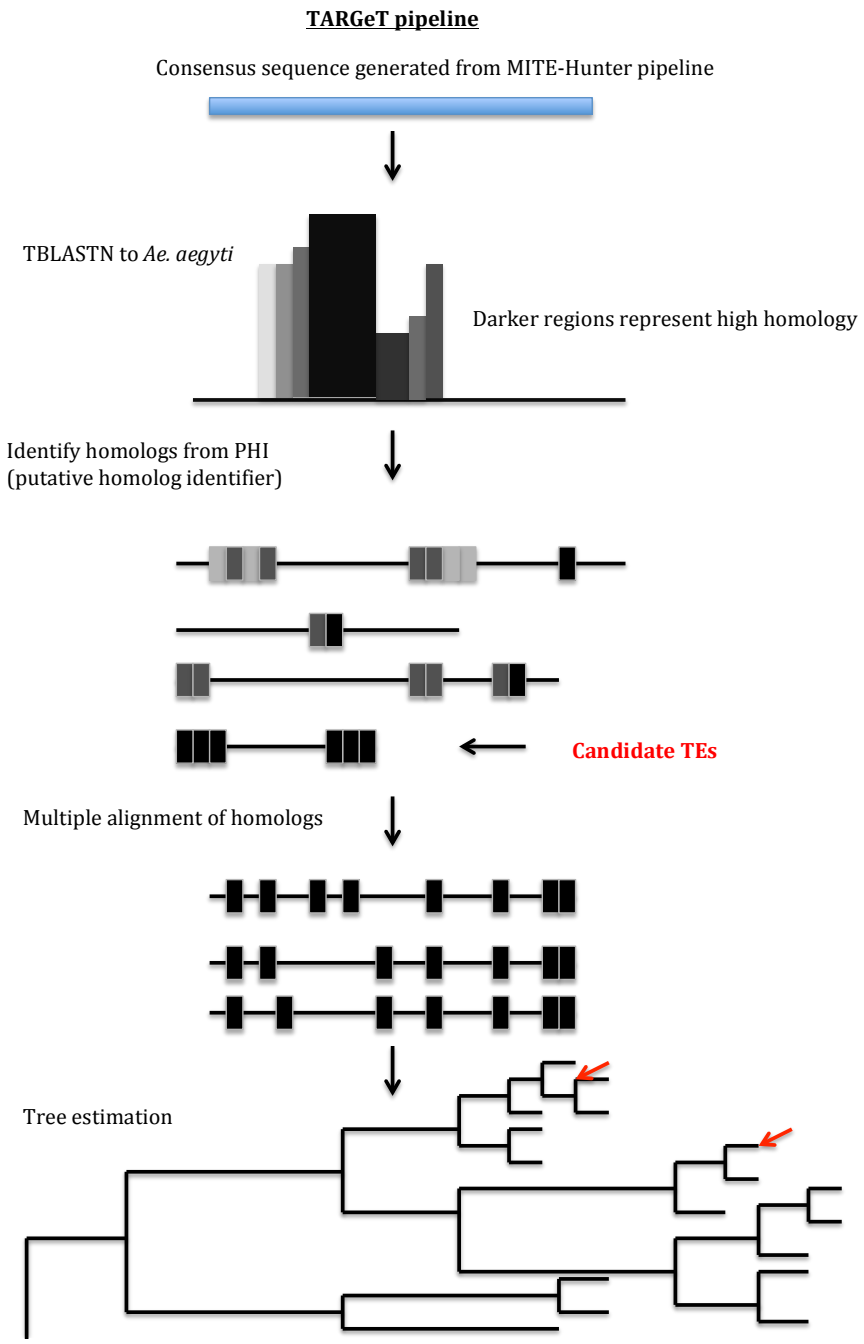


Figure 2.2 Steps involved in discovery of active TEs using TARGeT pipeline. 1) The *Ae. aegypti* genome was used to blast the consensus sequences, 2) darker region resembles a good match to the genome, 3) homologs were identified, 4) multiple alignment was carried out to determine similarities between the homologs, and 5) phylogenetic tree was built, smaller branches resembles a good match to the genome (Adapted from Han et al., 2009).

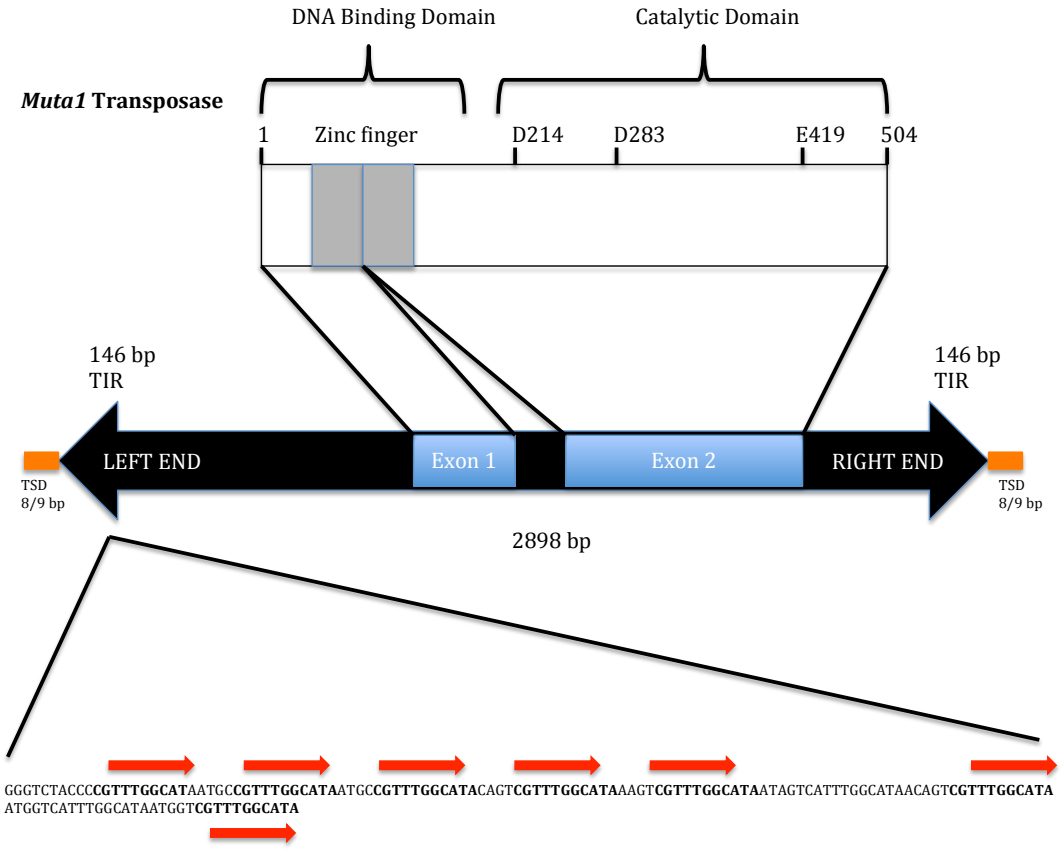


Figure 2.3 Structure of the *Muta1* element discovered using the TARGeT pipeline. The figure shows the TSDs and TIRs for the *Muta1* element. The red arrows in the left end represent 11 bp direct repeats. The left end has seven 11 bp repeats and the right end has six 11 bp repeats (not shown in the figure). Kun Liu and Susan R. Wessler identified the DDE motif in the catalytic domain of the *Muta1* transposase.

Supercontig	Flanking sequence	**TSD**	TIR
Supercontig 1.223	ATATACTACAATAGTTCTAAATAATG	**GACGCAGT**	GGGTCTACCCCGTTTGGCATAATGCCG
Supercontig 1.869	TTTTTAACATTTCAGTTTCAGCAAC	**ATGTTCGTT**	GGGTCTACCCCGTTTGGCATAATGCCG
Supercontig 1.873	TTTACTACAAAGCTACTGGTAGTTAG	**CTTCAGAG**	GGGTCTACCCCGTTTGGCATAATGCCG
Supercontig 1.801	TGCCCGAGTACCTCATCACTTAAGT	**CATTTGAAG**	GGGTCTACCCCGTTTGGCATAATGCCG
Supercontig 1.509	TGGTAAGGTAATAGGTATTTATTAT	**TTCGTTTTT**	GGGTCTACCCCGTTTGGCATAATGCCG
Supercontig 1.304	GCATTCGAGAATGGAAGTCGAGGTA	**ATGGTACAT**	GGGTCTACCCCGTTTGGCATAATGCCG
Supercontig 1.3	TTGTTTGTAATCCCTCTACCGGTAG	**CTTTATTAT**	GGGTCTACCCCGTTTGGCATACAGTCG
Supercontig 1.922	ACCCTAATCAAAAAATCCAAAAATA	**CGTGTATCC**	GGGTCTACCCCGTTTGGCATAATGCCG

Figure 2.4 Supercontig locations of eight wild-type copies of *Muta1* in *Ae. aegypti* forming 8 or 9 bp TSD (bold).

***hAT* Transposase domain analysis**

								D				**D**	
hAtx-3_SM	19	CYK—C	(10)	HMKN--KH	91	ARLAAKDGFSI	105	TRS	105	S**D**FIR	126	VMKLV	131 TDF
hAT-1_AA	23	CEQI--C	(40)	HMK----KH	93	ARLAAKDGFSI	108	TRS	165	L**D**EWT	211	RNMVQ	229 T**D**G
Tol2	47	CVL—C	(19)	HIER---MH	127	YIIQGLKPFST	163	RSK	191	T**D**CWT	234	EVLAS	258 T**D**S
Herves	26	CLY—C	(16)	HLNL---VH	115	LICKECLPFNL	151	SNA	180	S**D**GWT	222	RNIAN	246 T**D**N
Hermes	51	CRK---C	(16)	HKCC---AS	101	WVVRDCRPFSA	147	SRK	179	I**D**LWT	222	ENIYK	247 T**D**R
TcBuster1	77	CVI—C	(16)	HLDT-KH	147	RIAKQGEAYTI	193	SRR	222	M**D**EST	263	EEIFN	288 T**D**G
AeBuster1	75	CVI—C	(16)	HLET-KH	145	KIAKSGKAHTI	191	ARR	219	L**D**EST	259	DEIFD	284 T**D**G

Domain BED DNA binding CATALYTIC

--Block A--

hATx3_SM	148	ICTHHN	243	KLFSQLDLHRRLWVDEMLL	273	SCMQNRRP
hAT-1_AA	249	QCYNHA	287	DDLYASQFGEVLKRIRKDVKI	325	KKFQKTLQ
Tol2	325	KCACHL	382	LQLLRPNQTRWNSTFMAVD	449	DILQAETN
Herves	265	PCFAHT	319	LKMTQEVSTRWNNSGYDML	382	NIVSAQKY
Hermes	264	NCSSHL	309	SSLKSECRTRWNSTYTMLR	372	KELQTCSS
TcBuster1	316	HCCLHR	369	KNLLLHTEVRWLSRGKVLT	433	LSLQGPNS
AeBuster1	312	HCSLHR	365	TSLLLHTEVRWLSRGKVLT	420	TYLQGSTS

Insertion

Domain --Block B-- --Block C--

E

hATx-3_SM	303	----------HQFQSLRTH**P**SLRRSSLLRSK**E**DFRIFIRTTEDPWIMLNLNRPVYVSWNRLSKSFRIQY
hAT-1_AA	338	KTRWNS --TEVMLERFIQTFDCLR**E**ALEEIGLPDLVLNTDLEVAKSLHEALQPVRLASEALGRRDAT
Tol2	583	--TRE-------SLLTF**P**AICSLAIKTNTPLPASAAC**E**RLFSTAGLLFSPKRARLDTNNFENQLLLKLNLRFY
Herves	541	LWWKE----HQVLY**P**SLYTLAMSTLCIPGTSVPC**E**RLFSKAGQIYSEKRSRLAPKKLQEILFIQQN-------

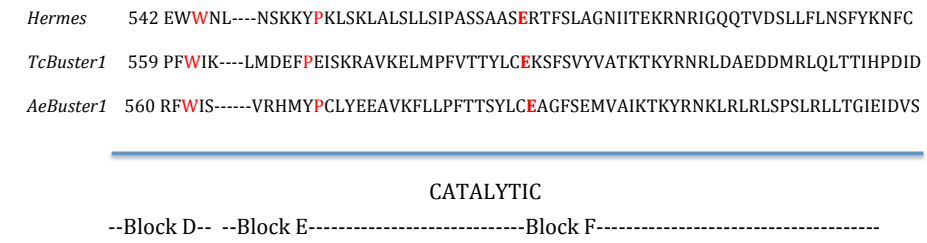


Figure 2.5 The analysis of BED domain and the catalytic domain of two new *hAT* transposons *hATx3_SM* and *hAT-1_AA*. The domains identified in these two transposons were based on alignment to other known *hAT* transposons. (Atkinson PW (2015) *hAT* Transposable Elements, in Mobile DNA III, Craig N. L (senior editor) ASM Press, Washington D. C., 27 pp.)

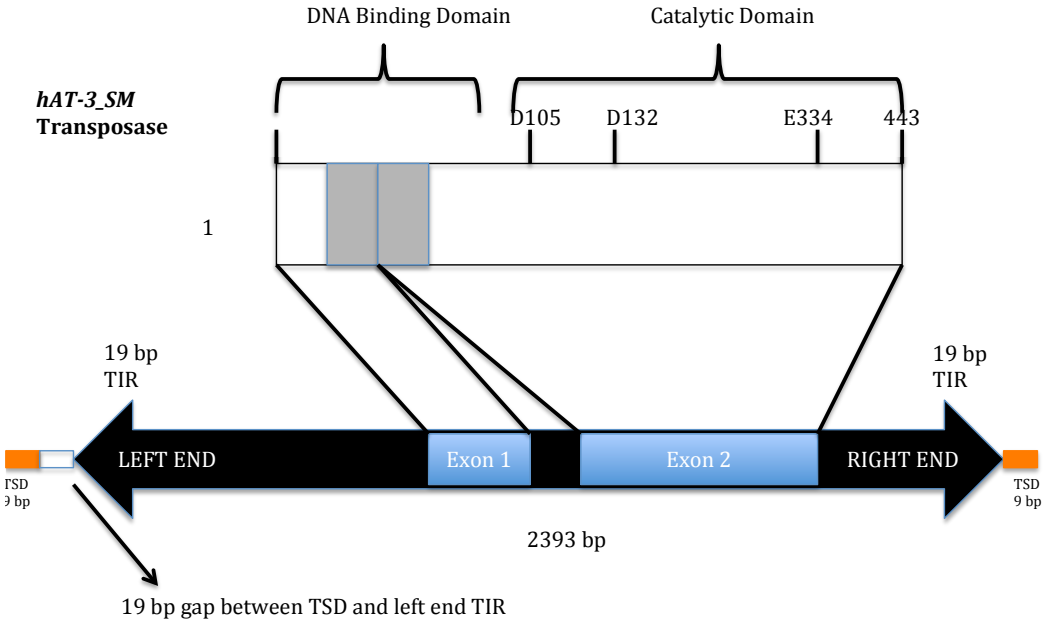


Figure 2.6. The TIRs and TSDs length, BED zinc finger DNA binding domain and catalytic DDE motif identified in the *hAT-3_SM* transposon. I predicted the DDE motif and catalytic domain based on the alignment from other *hAT* transposons. (Atkinson PW (2015) *hAT* Transposable Elements, in Mobile DNA III, Craig N. L (senior editor) ASM Press, Washington D. C., 27 pp.)

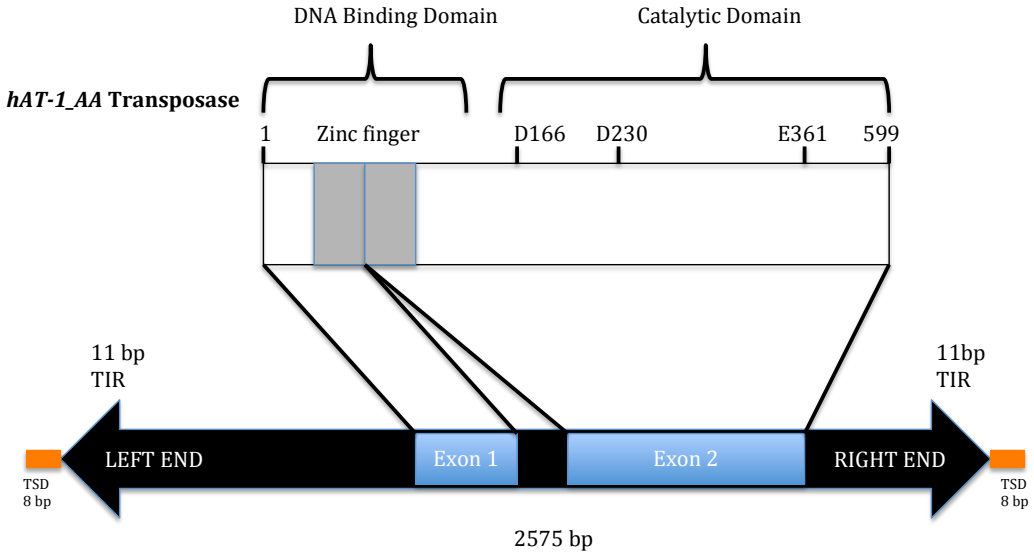


Figure 2.7. The TIRs and TSDs length, BED zinc finger DNA binding domain and catalytic DDE motif identified in *hAT-1_AA* transposon. I predicted the DDE motif and catalytic domain based on the alignment from other *hAT* transposons. (Atkinson PW (2015) *hAT* Transposable Elements, in Mobile DNA III, Craig N. L (senior editor) ASM Press, Washington D. C., 27 pp.)

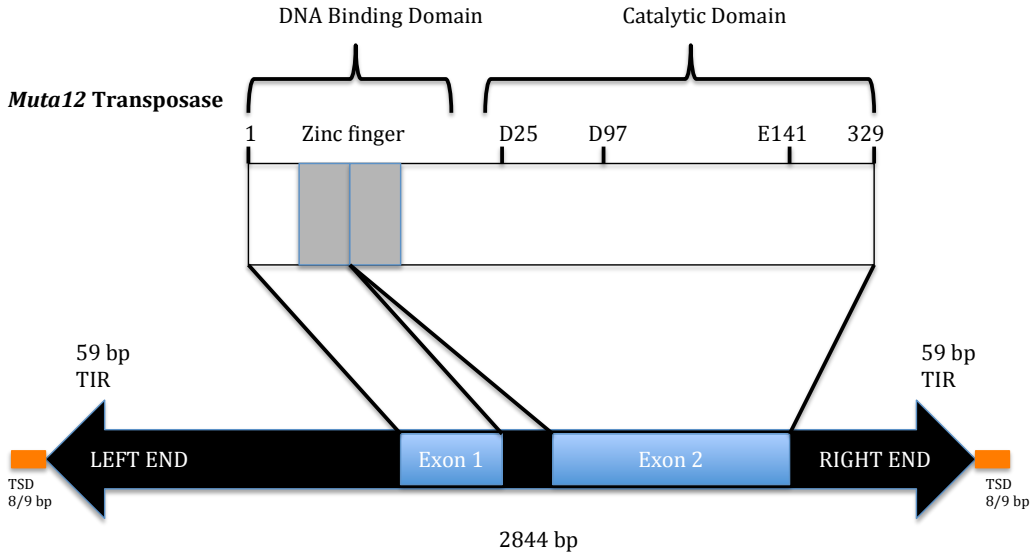


Figure 2.8 The figure shows the TIRs and TSDs length, BED zinc finger DNA binding domain and catalytic DDE motif identified in *Muta12* transposon. I predicted the DDE motif based on the DDE motif identified in other *Mutator* transposons. (Yuan & Wessler, 2011)

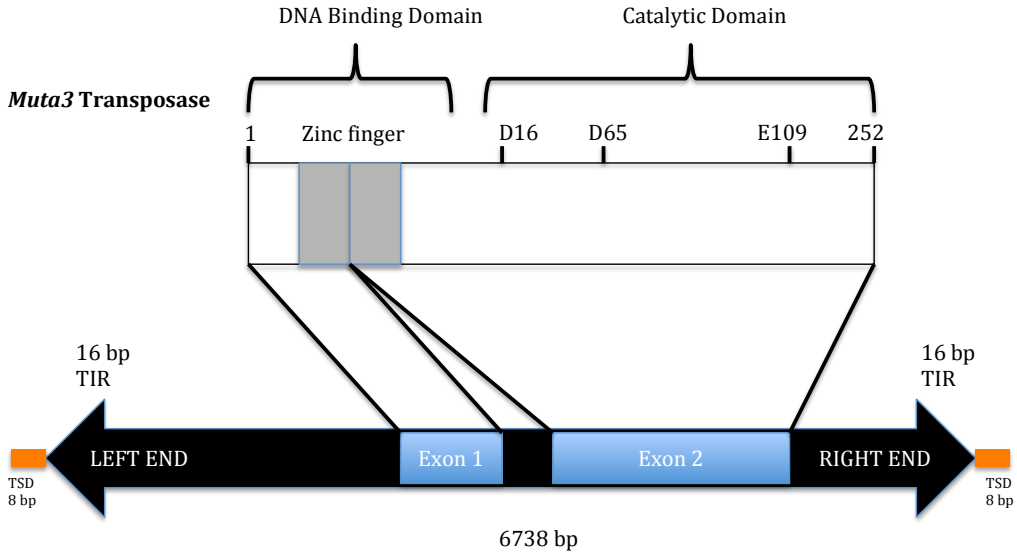


Figure 2.9 The TIRs and TSDs length, BED zinc finger DNA binding domain and catalytic DDE motif identified in *Muta3* transposon. I predicted the DDE motif based on the DDE motif identified in other *Mutator* transposons. (Yuan & Wessler, 2011)

Chapter Three. The *Muta1* transposon is active in *D. melanogaster* and *Ae. aegypti.*

3.1 Abstract

A somatic transposition activity using a five-plasmid assay showed the activity of the *Muta1* transposon in *D. melanogaster* and *Ae. aegypti,* this being the first time a member of this family has been shown to function in insects. *Muta1* showed 1.3 times higher activity in somatic cells of *D. melanogaster* and 16 times higher activity in *Ae. aegypti* than the *piggyBac* transposon which is widely used a gene vector in these species. This study also revealed the excision rate and footprint, and a pattern of integration for the *Muta1* element. *Muta1* precisely excised in 67% to 73% of the events in *D. melanogaster* and in 25% to 60% of the events in *Ae. aegypti*. *Muta1* has a broad target site preference and prefers to form a 9 bp TSDs in 86% to 100% of integrations.

3.2 Introduction

The *Mutator* (*Mu*) superfamily of transposable elements was first identified by Donald Robertson in a line of maize possessing an extremely high forward mutation rate (Robertson, 1980). A two-component system, one autonomous *MuDR* and many non-autonomous *Mu* elements, was identified as a source of these mutations, which are unstable in somatic and germline tissues (Barker, Thompson, Talbot, Swanson, & Bennetzen, 1984) (D Lisch, Chomet, & Freeling, 1995). The autonomous element *MuDR* encodes a Mutator Regulator A (MURA) protein that acts as a transposase and a MURB protein that is implicated in element reinsertion (Bennetzen, 1996). However, its exact function remains is unknown (Hua-Van & Capy, 2008). *Mutator* elements have an integration preference into gene-rich regions, thereby causing the high forward mutation

rate (Bennetzen, 1996). In addition, this study also revealed that *Mu* transposons may have a preference for insertion into promoters, because a number of insertions were isolated from promoter regions (Bennetzen, 1996). The *Mu* transposon contains TIRs of approximately 215 bp in length that are highly conserved, and forms a 9 bp TSD upon transposition (Dietrich et al., 2002). *Mutator* autonomous transposons that have been discovered in other organisms are *Hop1* in *Fusarium oxysporum* (Chalvet, Grimaldi, Kaper, Langin, & Daboussi, 2003a), *Jittery* in maize (Xu et al., 2004), (Li, Harris, & Dooner, 2013), and *Os3378* in rice (Gao, 2012). These TEs encode an intact protein homologous to MURA but lack the gene encoding the MURB protein, indicating that MURA is necessary for transposition of these TEs (Damon Lisch, 2002) (Tan et al., 2011).

In plants, animals, and fungi, similar *Mutator* TEs were later identified and referred to as *Mutator* (*Mu)*-like elements (*MULEs*) (Yu, Wright, & Bureau, 2000) (Xu et al., 2004) (Chalvet, Grimaldi, Kaper, Langin, & Daboussi, 2003b). *MULEs* are typically characterized by an 8 to 11 bp TSDs, with a 9 bp TSD being the most frequent form. There are two types of *MULEs*, one with long TIRs, which range from 100 to 500 bp and are known as TIR-*MULEs* (Benito & Walbot, 1997) and another with short TIRs, which are known as non-TIR *MULEs.* Non-TIR *MULEs* have low similarity between the TIRs, and were reported in *Arabidopsis* (Yu et al., 2000), *Lotus japonicas* (Holligan, Zhang, Jiang, Pritham, & Wessler, 2006), *Zea maize* (Wang & Dooner, 2006) and

Saccharomyces cerevisiae (Neuvéglise, Chalvet, Wincker, Gaillardin, & Casaregola, 2005).

Pack-MULEs are non-autonomous *Mutato*r and *MULE* TEs that carry genes or gene fragments (Jiang, Bao, Zhang, Eddy, & Wessler, 2004). These elements are known to acquire sequences from the genome, for example 2,853 *Pack-MULEs* have transduced 1500 parental genes in rice (Jiang, Ferguson, Slotkin, & Lisch, 2011). The first *Mutator* element discovered, *Mu1*, is an example of a *Pack-MULE* (Talbert & Chandler, 1988). Pack-MULEs can carry genes from multiple loci, therefore forming new ORFs (Jiang et al., 2011). They can also serve as a part of untranslated regions that fuse with adjacent sequences to form chimeric transcripts (Jiang et al., 2011). These transcripts are found in either orientation, which may suggest a role in regulation by generation of small RNAs. In the siRNA pathway, transcripts that are produced from either orientation can form double stranded RNAs (dsRNAs), which serve as a precursor for production of endo-siRNA (Ghildiyal & Zamore, 2009). The genes that have shared small RNAs with *Pack-MULEs* show lower expression levels compared to genes without association with small RNAs (Hanada et al., 2009). Thus, these elements may play a crucial role in gene evolution and regulation.

Recently, a member of the *Mutator* superfamily have been identified in animals (Marquez & Pritham, 2010). *Phantom* is a group of *Mutator* DNA transposons identified in the genomes of wide range of eukaryotes (Marquez & Pritham, 2010). *Phantom*

proteins were also identified in two insect viruses, *Chelono bracovirus* and *Glypta fumiferanae ichnovirus* (Marquez & Pritham, 2010). *Phantom* elements were characterized into three types- (I) TEs with TIRs like *Mutator* elements, (II) TEs with sub-terminal inverted repeats (sub-TIR) and (III) TEs that have complex termini reminiscent of *Foldback* transposons. *Foldback* TEs were identified in *D. melanogaster* (Silber, Bazin, Lemeunier, Aulard, & Volovitch, 1989), through southern blots analysis it was shown that TIRs of *Foldback* elements have repetitive structures that sometime fold back to central region. TSDs for *Phantom* are variable in sequence and range from 7 to 12 bp, which is in consistent with other *Mutator* transposons. Also, *Phantom* elements possess a conserved FLYWCH DNA binding domain and a catalytic domain similar to MULEs and *MuDR* transposase (Marquez & Pritham, 2010).

The *Mutator* elements transpositional activity has not been demonstrated in mosquitoes and *Drosophila*. Here, I describe the behavior of *Muta1* transposon, an autonomous class II DNA transposon, which belongs to the *Mutator* superfamily and is endogenous to one of the important vector for human disease *Ae. aegypti*. *Muta1* was previously found to be active in yeast in independent studies (Kun Liu and Susan R. Wessler, Plant Biology Graduate Program, University of California, Riverside).

3.3 Materials and Method

3.3.1 Fly stocks

A strain of *D. melanogaster* Canton-S white was used for transposition assays. The flies were raised on fruit fly media and supplemented with dry active yeast.

The laboratory strain is maintained in the Atkinson laboratory at University of California, Riverside.

3.3.2 Mosquito rearing

The Liverpool strain of *Ae. aegypti* was maintained under standard insectary conditions (Munstermann, 1997) with the exception that larvae were fed on a modified diet consisting of ground Milkbone Original Dog Biscuits mixed with Red star specialty Nutrex 55 (Red Star, Milwaukee, WI). Adult cages were fed on a 10% sucrose solution and blood fed on artificial membrane feeding using sheep blood. The mosquitoes are maintained at 26° C with a photoperiod of 16:8 (L:D).

3.3.3 Plasmid Constructs

All of the plasmids made for the transposition assays were constructed using a standard protocol as determined in *Sambrook et al. (*1989).

3.3.3.1 *Muta1* Donor plasmid

Muta1 was amplified in sections from *Ae. aegypti* (Liverpool). Genomic DNA was purified using a Blood DNAeasy kit (Promega), and 360 ng of DNA was used as template in 50-µl PCR reactions using a TripleMaster PCR System (Eppendorf). The pGT-Muta1LE clone was made by PCR amplification using primers that encompassed the region from the TSD of one copy of *Muta1* to the sequence 350 bp upstream of the *Muta1* ORF. The primers used were Muta1LE Forward (5′-AATGGTACCGCTTATGG**CATAGATTCCCAAACT**GTG-3′), and *Muta1* LE Reverse (5′-GATCTCGAGATCTGAAATTATCAAATAATGAATCGCATATTCTG-3′), with the following PCR program: 94° 2′, 4 × (94° 20″, 60° 20″, 72° 30″), 25 × (94°

20″, 69° 20″, 72° 30″), 72° 5′, 4°. Following amplification, the DNA was purified using a Qiaquick PCR purification kit (Qiagen), quantified on an agarose gel, and ligated into the pJet2-1 vector. Inserts were sequenced, and then clones were digested with *Kpn*I and *Xho*I (New England Biolabs). Gel-purified fragments (Zymoclean Gel DNA Recovery Kit, Zymo Research) were cloned into pBluescript SK+ digested with the same enzymes to make the clone pBSMuta1LE. The right end of *Muta1* was amplified in a similar manner using the primers *Muta1* RE forward (5′-GATTCTAGATGCGCATCGAACAACATTTTTAGTGAG-3′) and *Muta1* RE Reverse (5′-AATGAGCTCCCATAAGCCATAGGTTCCCAAACTTTTC-3′), which encompassed the region 350 bp 3′ of the stop codon through the target-site duplication. The PCR program was: 94° 2′, 4 × (94° 20″, 60° 20″, 72° 30″), 25 × (94° 20″, 72° 30″), 72° 5′, 4°. The right end PCR product was cloned as above, first into pJet2-1 and then into the left end clone following digestion with *Sac*I and *Xba*I (New England Biolabs, NEB) to yield the clone pBSMuta1LR. *Xba1*-digested pBSMuta1LR was ligated with *Nhe1*-digested pGENToriAlpha, which was derived from pK19 and in which the kanamycin^R gene was replaced with the gentamycin^R gene from pFastBac HTb (Invitrogen) to generate a donor element, which has a replication origin, a gentamycin^R gene, and a lacZ-alpha gene.

3.3.3.2 *Muta1* Helper plasmid

The *Muta1* helper plasmid was made as follows: PCR was performed with the primer primers *Muta1* ORF Forward

(5’-AATGATATCAGAAATATGATGCTGAATTGGCTCAAAAGTGG-3’) and *Muta1* ORF Rev (5’-GATGATATCTTAATGACTTTTTTGCGCTTGCTTATTATTGCAC-3’). The PCR program was: 94° for 2 min, 3 × (94° for 20 sec, 67.5° for 20 sec, 72° for 2 min), 26 × (94° for 20 sec, 70° for 20 sec, 72° for 2 min), 72° for 5 min, 4°. PCR product was purified as above, digested with *Spe1 and Xho1,* cloned into the plasmid pKhsp70 (Arensburger et al., 2005).

3.3.3.3 *piggyBac* Donor plasmid

The pBacGoEGFP plasmid used in the normalized transposition assays was constructed by first PCR amplifying the left and right arms of *piggyBac* using the pBac[3xP3-EGFPafm] plasmid as a template. The left end, including a small portion of flanking plasmid, TSD, and a short region of the ORF was amplified with the primers *SacII*-pBac (5’- ARACCGCGGTCTTTTTTAACCCYAGAAAGATAGTCGCC-3’) and *XbaI*-pBac LE R + ORF (5’-ATATCTAGAGCTCATCGTCTAAAGAACTACCC-3’). The right end, flanking plasmid, and TSD were amplified with the restriction site primers *KpnI*-pBac RE F 5’- ATAGGTACCCTATTATTAACCCTAGAAAGATAATCATATTGTG-3’ and *Xho1*-pBac RE F

5’-ATACTCGAGGACTAATAAGTATAATTTGTTTCTATTATGTATAAGTTAAGC-3’). Using restriction sites incorporated into the PCR products, the left and right end samples were digested and inserted into pBluescript II KS+ using the *SacII-XbaI* and *KpnI-XhoI* sites respectively. The resultant plasmid, pBSpBacLR, was then linearized by digestion with *XbaI* to incorporate a *NheI* linearized pGENToriEGFP plasmid to create the pBacGoEGFP plasmid. The pGENToriEGFP plasmid was created by excising the

LacZα ORF from the pGENToriα plasmid by digestion with *BamHI* and *NheI*, and replaced by a *BamHI*/*NheI* fragment containing the EGFP ORF. The *BamHI*/*NheI* EGFP fragment was generated through the PCR amplification of the ORF using the primers BamHI EGFP and NheI EGFP and the pBac[3xP3-EGFPafm] plasmid as a template. The PCR product was then digested with *BamHI* and *NheI* and ligated into the pBacGoEGFP (Smith & Atkinson, 2011).

3.3.3.4 *PiggyBac* helper plasmid

The *piggyBac* helper was phsp70-Bac (formerly pBhsΔSa) (Handler *et al*. 1998). phsp70-Bac contains the gene AmpR within the plasmid backbone. The transposase helper plasmid expresses the *piggyBac* ORF under the control of the *D. melanogaster Hsp70* promoter.

3.3.3.5 Target Plasmid

pGDV1 was used as a target plasmid. pGDV1 is a *Bacillus subtilus* low copy plasmid incapable of replicating in *Escherichia coli* and used as a target plasmid in plasmid-based transposition assays (Sarkar et al., 1997).

3.3.4 Plasmid mixes

Injection mix was prepared for transposition assay through microinjection into embryos of D. melanogaster and Ae. aegypti. Normalization of the data is required when comparing relative rate of transposition frequency across experiments. This is achieved through the use of a five-plasmid inter-plasmid transposition assay as developed in Smith (2007). In this assay, *Muta1* transposon is tested by co-injecting the appropriate

helper and donor plasmids along with pGDV1 as a target. This injection mix contains the following: 250 ng/ul of *Muta1* donor, 250 ng/ul of the *Muta1* helper-transposase, 250 ng/ul of *piggyBac* donor, 250 ng/ul of wild-type *piggyBac* helper-transposase and 1000 ng/ul of *pGDV1* as a donor plasmid. By using *piggyBac* in the assay along with *Muta1,* we can compare the number of transposition events of the *Muta1* transposase relative to the number of *piggyBac* transposition events. Thus, transposition rates can be normalized between experiments. In each of the experiments the transposition events produced from different transposases can be distinguished by the markers present in the donor plasmids. The *Muta1* donor plasmid has the *LacZ* gene, which encodes the *β*-Galactosidase enzyme that catalyze hydrolysis of X-gal producing a blue precipitate that can be easily visualized. The *piggyBac* donor plasmid contains the GFP marker that can be visualized under a fluorescent microscope.

3.3.5 Microinjection into *D. melanogaster* embryos.

D. melanogaster females were placed on pineapple agar plates that induce oviposition. Pre- blastoderm embryos were collected within 45 minutes of oviposition and dechorionated using a 10% bleach solution. The embryos were placed on a slide and then coated in a layer of halocarbon 700 oil (Poly Sciences, Inc) to protect from desiccation. Borosilicate glass capillaries 0.7 mm in diameter was pulled into needles using a Flaming/Brown Micropipette Puller. An Eppendorf Femtojet was used to deliver the plasmid mix into the embryo. Injected embryos were placed in a humidity chamber under 100% oxygen for 15-20 hours. The embryos were collected from the glass slides

and placed into a proteinase buffer. The embryos were then lysed and incubated for 30°C minutes at 55°C. Next the plasmids were recovered and electroporated into the DH10B *E.coli* strain (Gibco-BRL). Of these cells, 1/200th were plated on LB plates containing the appropriate antibiotics for 24 hours at 37° to screen for donor titer. The remaining cells were then placed onto plates containing a mixture of antibiotics for three days at 37°C to screen for transposition events as described in Smith (2007). After 3 days incubation at 37°C, resistant colonies were picked and grown in LB media containing only gentamycin and plating gentamycin resistant colonies on LBplates containing ampicillin as negative test for transposition events. Plasmid DNA was purified from these cells of colonies that do not grow on LBplates containing ampicillin using the Wizard Plus miniprep kit (Promega). The presence of a recombinant plasmid arising from transposition was verified by digesting the plasmid DNA with *HindIII* to check for a diagnostic pattern of bands (1.1, 1.5, and 3.2 kb). Plasmids passing this initial test were confirmed as transposition events by DNA sequencing. The transposition events were sequenced at the University of California, Riverside Genomic Core Facility to determine the integration sites of *Muta1* into the pGDV1target plasmid (Sarkar et al., 1997) (Arensburger et al., 2005).

3.3.6 Microinjection into *Ae. aegypti* embryos.

Ae. aegypti adults were blood fed on artificial membrane feeding using sheep blood. Females were provided with a small tub of water with a Whatman paper filter in it and allowed to lay eggs for 30 minutes. The embryos were then aligned on double stick tape

and mounted on a plastic slide and covered with halocarbon oil 27. The embryos are injected in the same manner as described for *D. melanogaster*. After microinjection the oil was washed away and the slides were placed in a pan of diH_2O. After 16-18 hours post injection the embryos were then processed using the transposition assay described above.

3.3.7 Excision assays performed in insect embryos.

Alternatively, *Muta1* excision assays were performed on plasmids recovered from transposition assays performed in *D. melanogaster* and *Ae. aegypti* embryos. Recovered plasmid DNA was digested with *Eco*RV and the resulting DNA was used to transform DH10 cells by electroporation followed by selection on LB plates containing ampicillin and X-gal (20 mg/liter). Because *Eco*RV cuts only within the *Muta1* donor element, plasmids arising as a result of excision are resistant to *Eco*RV linearization. Uncut excision products efficiently transform *E. coli* while linearized donor plasmids do not. Putative excision events were confirmed by restriction digestion and DNA sequencing. $Ampicillin^R$, $LacZ^-$ colonies were selected, mapped and then sequenced across the empty excision site using the primers 5′-CGTCCCATTCGCCATTCAGG-3′.

3.4 Results

3.4.1 *Muta1* is somatically active in *D. melanogaster* and *Ae. aegypti*

Muta1 was co-injected with the *piggyBac* element into *D. melanogaster* embryos as a means of normalizing the resulting frequencies between experiments (Table

3.1). *Muta1* flanked by an 8 bp TSD had a transposition frequency of 5.7 x 10^{-4} in *D. melanogaster* and 8.3 x 10^{-3} in *Ae. aegypti* (Table 3.1). *Muta1* was 1.5 times more active compared to *piggyBac* in *D. melanogaster* and 16 times more active in *Ae, aegypti* (Table 3.1). *Muta1* flanked by a 9 bp TSD had a transposition frequency of 3.96 x 10^{-4} in *D. melanogaster* and 9.45 x 10^{-2} in *Ae. aegypti* (Table 3.1). *Muta1* was therefore approximately1.3 times more active than *piggyBac* in *D. melanogaster* and 13 times more active in *Ae, aegypti* compared to *piggyBac* (Table 3.1). Zero control transposition assays showed no activity of *Muta1* in absence of wild-type transposase in *D. melanogaster* and *Ae, aegypt*i respectively (Table 3.1). Lack of activity of *Muta1* element in the absence of *Muta1* transposase from the helper plasmid demonstrated that the endogenous wild-type copies of *Muta1* element in *Ae. aegypti* were either not active or that the transposase expression was insufficient to mobilize *Muta1* to detectable levels in the assays used.

3.4.2 *Muta1* has different excision pattern in *D. melanogaster* and *Ae. aegypti*

Muta1 flanked by an 8 bp TSD had an excision frequency of 1.4 x 10^{-2} in *D. melanogaster* and 1.9 x 10^{-3} in *Ae. aegypti* (Table 3.2). Approximately 67% of the recovered events from excision assays indicated precise excision of the *Muta1* transposon from the donor plasmid flanked by an 8 bp TSD, forming a single TSD with no foot print of integration in *D. melanogaster* (Figure 3.2), Other patterns of excision that were observed were the *Muta1* element cleaving inside the plasmid DNA with no TSD, one

event with an extra nucleotide flanking to the TSD; two events with a deletion of nucleotides in the TSD and a restriction site, respectively (Figure 3.2). Precise excision was observed in 25% of the events characterized in *Ae. aegypti* performed using a donor plasmid with an 8 bp TSD (Figure 3.3). One event acquired five extra nucleotides flanking to the restriction site, another showed cleavage approximately 100 bp inside the right end, one event had two TSDs, and two events were found to have an excision cleaving inside the restriction site on either side of the end (Figure 3.3).

The same *Muta1* TE flanked by a 9 bp TSD had an excision frequency of 2.24×10^{-2} in *D. melanogaster* and 8.73×10^{-3} in *Ae. aegypti* (Table 3.2). The analysis of the excision pattern for the events recovered using the *Muta1* element flanked by a 9 bp TSD revealed, that *Muta1* element excised precisely approximately 73% of the time in *D. melanogaster* and 60% of the time in *Ae. aegypti*. An event was characterized to have an extra nucleotide flanking to the TSD, another was found to have excision in the *Kpn1* restriction site cleaving two nucleotides, and a third to have excised outside the ends causing deletion of approximately 171 bp nucleotides in *D. melanogaster* (Figure 3.4). In *Ae. aegypti,* one excision event was characterized to have an extra nucleotide flanking to the TSD, another had two TSDs, and one had a break point in the TIR leaving approximately 41 bp nucleotides from the right end TIR along with a TSD (Figure 3.4). The increase in frequency of precise excision with *Muta1* flanked by 9 bp TSD was obtained from analysis of 10 events.

3.4.3 *Muta1* target site preference

Transposition events recovered from *D. melanogaster* and *Ae. aegypti* embryos were sequenced from both left and right end outwards to determine the frequency of 8 or 9 bp TSDs (Table 3.3). The *Muta1* element flanked by an 8 bp TSD generated 9 bp TSD 86% of the time in *D. melanogaster* (Table 3.3, section A) and 90% of the time in *Ae. aegypti* (Table 3.3, section B). The *Muta1* element flanked by a 9 bp TSDs, showed 90-100% of the events formed 9 bp TSDs in *D. melanogaster* (Table 3.3, section C) and *Ae. aegypti* (Table 3.3, section D). Clearly, there is a very strong preference for the generation of 9 bp TSDs.

The TSDs were analyzed to construct a consensus sequence depicting the preference for certain nucleotides in *D. melanogaster* and *Ae. aegypti* (Table 3.4). The *Muta1* element flanked by an 8 bp TSD generated a 'T' at second position in 52% of the events, 'A' at sixth position in 57% of the events, and 'A' at seventh position in 71% of the sequenced events in *D. melanogaster* (Figures 3.6), In *Ae. aegypti,* 50% of the events showed to have 'T' at second position, 'T' at third position in 56% of the events and 'A' at eighth position in 50% of the events (Figures 3.7) In addition, analysis of the consensus sequence from *Muta1* element flanked by a 9 bp TSD revealed to have 'T' at second position in 50% of the events, 'T' at sixth position in 50% of the events and 'T' at seventh position 50% of the sequenced events in *D. melanogaster* (Figures 3.8) and in

Ae. aegypti, 60% of the events had 'T' at second position, 'T' at third position in 70% of the events and 'A' at sixth position 50% of the events (Figures 3.9)

3.5 Discussion

Transposition assays are useful for detecting somatic activity of a transposon, but these assays are subject to great deal of variation within and between experiments (Brust-Mascher & Scholey, 2009). Factors that affect the success of somatic transposition assays are embryo age, embryo desiccation, needle shape and sharpness, injection pressure, atmospheric temperature, humidity among many other variables that cannot be kept constant between experiments (Adelman, Jasinskiene, & James, 2002). Two key parameters, the transposition rate and the excision rate are fundamental to revealing the possible mechanism of transposition (Sousa, Bourgard, Wahl, & Gordo, 2013).

Muta1 transposition assays performed in *D. melanogaster* demonstrated that this TE is active in somatic tissues with 1.5 times higher activity compared to the *piggyBac* element and 16 times higher activity in *Ae. aegypti* (Table 3.1). Bioinformatics analysis was performed to determine presence of *Muta1* and *Muta1*-like sequences in *D. melanogaster.* This analysis showed no evidence of this TE or related sequences in *D. melanogaster*. Zero-control transposition assays confirmed absence of *Muta1* activity in in absence of the helper plasmid. Thus, the *D. melanogaster* genome is naïve to the *Muta1* element and appears to be naïve to *Mutator* elements in general. Eight wild-type

copies of full length *Muta1* element are present in the *Ae*. *aegypti* genome, however no evidence of *Muta1* element movement in absence of introduced *Muta1* transposase was detected in these assays. A possible explanation for this observation could be that the endogenous copies of the *Muta1* elements are inactive in the stage and tissue examined (embryos). An alternative explanation could be that the somatic transposition assays were not sensitive enough to detect activity of the *Muta1* element from the transposase contained in the wild-type copies in *Ae. aegypti*. Furthermore, it is possible that the *Muta1* transposase expression from the endogenous copies was restricted to the germline and thus no activity was detected in the somatic transposition assays.

Transposon stability is important for the use of transposons for human gene therapy, where a therapeutic gene can be inserted in between the TIRs of a non-autonomous transposon and its integration into the genome can be achieved using the transposase encoded in a separate vector. Long term expression of the transgene thus can be maintained within the cells, if neither of these constructs are autonomous (Davidson, Gratsch, Morell, O'Shea, & Krull, 2009). Stability of an integrated transposon is also important in generation of insects for genetic control strategies such as sterile insect techniques (SIT), where these transgenic insects are released into the field, albeit for a short period of a time. A major concern with the release of these insects is the stability of the transgene and its long-term expression within these insects (Handler, 2004), because instability of the transgene (Handler, 2004). Therefore, the *Muta1* transposon might serve as a useful genetic tool where we can obtain long-term stability of a transgene using a

non-autonomous *Muta1* element. Its interaction with genome in which there are *Muta1*-like elements present will be a form of further study.

The analysis of target site preference for *Muta1* revealed that it creates 8 or 9 TSDs upon integration. The *Muta1* element flanked by an 8 bp or a 9 bp TSD forms 9 bp TSDs when integrating into a new location in approximately 86%-100% of the transposition events characterized in *D. melanogaster* and *Ae. aegypti* (Table 3.2). Furthermore, analysis of the TSD consensus sequence showed bias for 'T' at the 2nd position, A/T at the 6th and 'T' at the 7th position in the TSD sequence. Previously, it has been reported from analysis of nine MULE groups identified in *Arabidopsis thaliana* to be AT-rich in their consensus TSD sequences (Yu et al., 2000). The analysis of TSD sequence from other *Mutator* TEs such as *Mu* element showed that it targets 5' UTR of the gene, which are GC rich as opposed to AT rich region (Dietrich et al., 2002). The TSDs sequence for *Muta1* element did not show a strong bias towards GC rich or AT rich region that suggest that it might have broad target site specificity, which can be useful for gene tagging and mutagenesis.

The *Muta1* excision assay in *D. melanogaster* revealed that excision of *Muta1* element flanked by an 8 bp TSD is precise in 67% of the events and in 73% of the events with *Muta1* element flanked by a 9 bp TSD. In *Ae. aegypti,* excision was precise in 25% of the events for a *Muta1* element flanked by an 8 bp TSD and 60% of the time for a *Muta1* element flanked by a 9 bp TSD. The number of precise excisions obtained was high with

the *Muta1* element flanked by a 9 bp TSD compared to an 8 bp TSD, which suggests that a 9 bp is the preferred TSD for transposition of *Muta1* element. This is also supported by observation of other *Mutator* elements such as *Mu1, Phantom, Jittery* and *Hop,* where 9 bp TSDs were observed in most of the recovered events (Doseff, Martienssen, & Sundaresan, 1991) (Marquez & Pritham, 2010) (Xu et al., 2004) (Chalvet et al., 2003b). Two excision events were found to have acquired an extra nucleotide flanking the TSD that represents approximately 9% of the events in *D. melanogaster* (Figure 3.4) and approximately 10% of the events in *Ae. aegypti* (Figure 3.5). One of the event was characterized to have acquired five extra bases, which is 12.5% of the events in *Ae. aegypti* (Figure 3.3); excision of the *Muta1* element inside and outside of the ends have also been recorded (Figure 3.3, 3.4, and 3.5). Similar excision pattern have been observed in other *Mutator* elements, for example the *Mu1* element in *maize*, there are excision products observed to have short sequences of the *Mu1* elements remaining after the excision (Doseff et al., 1991). Another excision product was interpreted to have an recombination event from a cross over with a *Bronze* gene where the element was previously characterized to be integrated (Doseff et al., 1991). Imprecise excision of *Muta1* could be due to transposase behavior in somatic cells of *D. melanogaster* and *Ae. aegypti*, and its interaction with other factors such as protein involved in TE silencing and DNA methylation (Pritham, 2009) that might be influencing binding of the *Muta1* transposase to the ends of transposon for the excision. The *Muta1* element has 11 bp direct repeats in the ends (Figure 2.1, Chapter 2), this types of repeats have previously been identified in a *Foldback* like element- *Galileo*, these repeats are capable of forming

harping structure and were identified to be the cause of chromosomal inversions in *Drosophila buzzati* (Casals, Cáceres, & Ruiz, 2003). Thus, there is a possibility that the *Foldback* like structure of the *Muta1* element might have an influence on the excision patterns observed in the characterized events. The direct repeats in the ends could loop back to form secondary structure that might need to be resolved during excision of the *Muta1* element causing imprecise excision.

3.6 Reference

Adelman, Z. N., Jasinskiene, N., & James, A. A. (2002). Development and applications of transgenesis in the yellow fever mosquito, Aedes aegypti. *Molecular and Biochemical Parasitology*, *121*(1), 1–10. http://doi.org/10.1016/S0166-6851(02)00028-2

Arensburger, P., Kim, Y.-J., Orsetti, J., Aluvihare, C., O'Brochta, D. A., & Atkinson, P. W. (2005). An active transposable element, Herves, from the African malaria mosquito Anopheles gambiae. *Genetics*, *169*(2), 697–708. http://doi.org/10.1534/genetics.104.036145

Barker, R. F., Thompson, D. V, Talbot, D. R., Swanson, J., & Bennetzen, J. L. (1984). Nucleotide sequence of the maize transposable element Mul. *Nucleic Acids Research*, *12*(15), 5955–67.

Benito, M., & Walbot, V. (1997). Characterization of the maize Mutator transposable element MURA transposase as a DNA-binding protein. *Mol. Cell. Biol.*, *17*(9), 5165–5175.

Bennetzen, J. L. (1996). The Mutator transposable element system of maize. *Current Topics in Microbiology and Immunology*, *204*, 195–229.

Brust-Mascher, I., & Scholey, J. M. (2009). Microinjection techniques for studying mitosis in the Drosophila melanogaster syncytial embryo. *Journal of Visualized Experiments : JoVE*, (31). http://doi.org/10.3791/1382

Casals, F., Cáceres, M., & Ruiz, A. (2003). The foldback-like transposon Galileo is involved in the generation of two different natural chromosomal inversions of Drosophila buzzatii. *Molecular Biology and Evolution*, *20*(5), 674–85. http://doi.org/10.1093/molbev/msg070

Chalvet, F., Grimaldi, C., Kaper, F., Langin, T., & Daboussi, M.-J. (2003a). Hop, an active Mutator-like element in the genome of the fungus Fusarium oxysporum. *Molecular Biology and Evolution*, *20*(8), 1362–75. http://doi.org/10.1093/molbev/msg155

Chalvet, F., Grimaldi, C., Kaper, F., Langin, T., & Daboussi, M.-J. (2003b). Hop, an active Mutator-like element in the genome of the fungus Fusarium oxysporum. *Molecular Biology and Evolution*, *20*(8), 1362–75. http://doi.org/10.1093/molbev/msg155

Davidson, A. E., Gratsch, T. E., Morell, M. H., O'Shea, K. S., & Krull, C. E. (2009). Use of the Sleeping Beauty transposon system for stable gene expression in mouse embryonic stem cells. *Cold Spring Harbor Protocols*, *2009*(8), pdb.prot5270. http://doi.org/10.1101/pdb.prot5270

Dietrich, C. R., Cui, F., Packila, M. L., Li, J., Ashlock, D. A., Nikolau, B. J., & Schnable, P. S. (2002). Maize Mu transposons are targeted to the 5' untranslated region of the gl8 gene and sequences flanking Mu target-site duplications exhibit nonrandom nucleotide composition throughout the genome. *Genetics*, *160*(2), 697–716.

Doseff, A., Martienssen, R., & Sundaresan, V. (1991). Somatic excision of the Mu1 transposable element of maize. *Nucleic Acids Research*, *19*(3), 579–84.

Gao, D. (2012). Identification of an active Mutator-like element (MULE) in rice (Oryza sativa). *Molecular Genetics and Genomics : MGG*, *287*(3), 261–71. http://doi.org/10.1007/s00438-012-0676-x

Ghildiyal, M., & Zamore, P. D. (2009). Small silencing RNAs: an expanding universe. *Nature Reviews. Genetics*, *10*(2), 94–108. http://doi.org/10.1038/nrg2504

Hanada, K., Vallejo, V., Nobuta, K., Slotkin, R. K., Lisch, D., Meyers, B. C., … Jiang, N. (2009). The functional role of pack-MULEs in rice inferred from purifying selection and expression profile. *The Plant Cell*, *21*(1), 25–38. http://doi.org/10.1105/tpc.108.063206

Handler, A. M. (2004). Understanding and improving transgene stability and expression in insects for SIT and conditional lethal release programs. *Insect Biochemistry and Molecular Biology*, *34*(2), 121–30. http://doi.org/10.1016/j.ibmb.2003.08.005

Holligan, D., Zhang, X., Jiang, N., Pritham, E. J., & Wessler, S. R. (2006). The transposable element landscape of the model legume Lotus japonicus. *Genetics*, *174*(4), 2215–28. http://doi.org/10.1534/genetics.106.062752

Hua-Van, A., & Capy, P. (2008). Analysis of the DDE motif in the Mutator superfamily. *Journal of Molecular Evolution*, *67*(6), 670–81. http://doi.org/10.1007/s00239-008-9178-1

Jiang, N., Bao, Z., Zhang, X., Eddy, S. R., & Wessler, S. R. (2004). Pack-MULE transposable elements mediate gene evolution in plants. *Nature*, *431*(7008), 569–73. http://doi.org/10.1038/nature02953

Jiang, N., Ferguson, A. A., Slotkin, R. K., & Lisch, D. (2011). Pack-Mutator-like transposable elements (Pack-MULEs) induce directional modification of genes through biased insertion and DNA acquisition. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(4), 1537–42. http://doi.org/10.1073/pnas.1010814108

Li, Y., Harris, L., & Dooner, H. K. (2013). TED, an autonomous and rare maize transposon of the mutator superfamily with a high gametophytic excision frequency. *The Plant Cell*, *25*(9), 3251–65. http://doi.org/10.1105/tpc.113.116517

Lisch, D. (2002). Mutator transposons. *Trends in Plant Science*, *7*(11), 498–504. http://doi.org/10.1016/S1360-1385(02)02347-6

Lisch, D., Chomet, P., & Freeling, M. (1995). Genetic characterization of the Mutator system in maize: behavior and regulation of Mu transposons in a minimal line. *Genetics*, *139*(4), 1777–96.

Marquez, C. P., & Pritham, E. J. (2010). Phantom, a new subclass of Mutator DNA transposons found in insect viruses and widely distributed in animals. *Genetics*, *185*(4), 1507–17. http://doi.org/10.1534/genetics.110.116673

Neuvéglise, C., Chalvet, F., Wincker, P., Gaillardin, C., & Casaregola, S. (2005). Mutator-like element in the yeast Yarrowia lipolytica displays multiple alternative splicings. *Eukaryotic Cell*, *4*(3), 615–24. http://doi.org/10.1128/EC.4.3.615-624.2005

Pritham, E. J. (2009). Transposable elements and factors influencing their success in eukaryotes. *The Journal of Heredity*, *100*(5), 648–55. http://doi.org/10.1093/jhered/esp065

Robertson, D. S. (1980). The timing of mu activity in maize. *Genetics*, *94*(4), 969–78.

Sarkar, A., Coates, C. J., Whyard, S., Willhoeft, U., Atkinson, P. W., & O'Brochta, D. A. (1997). The Hermes element from Musca domestica can transpose in four families of cyclorrhaphan flies. *Genetica*, *99*(1), 15–29.

Silber, J., Bazin, C., Lemeunier, F., Aulard, S., & Volovitch, M. (1989). Distribution and conservation of the foldback transposable element in Drosophila. *Journal of Molecular Evolution*, *28*(3), 220–4.

Smith, R. C., & Atkinson, P. W. (2011). Mobility properties of the Hermes transposable element in transgenic lines of Aedes aegypti. *Genetica*, *139*(1), 7–22. http://doi.org/10.1007/s10709-010-9459-7

Sousa, A., Bourgard, C., Wahl, L. M., & Gordo, I. (2013). Rates of transposition in Escherichia coli. *Biology Letters*, *9*(6), 20130838. http://doi.org/10.1098/rsbl.2013.0838

Talbert, L., & Chandler, V. (1988). Characterization of a highly conserved sequence related to mutator transposable elements in maize. *Mol. Biol. Evol.*, *5*(5), 519–529.

Tan, B.-C., Chen, Z., Shen, Y., Zhang, Y., Lai, J., & Sun, S. S. M. (2011). Identification of an active new mutator transposable element in maize. *G3 (Bethesda, Md.)*, *1*(4), 293–302. http://doi.org/10.1534/g3.111.000398

Wang, Q., & Dooner, H. K. (2006). Remarkable variation in maize genome structure inferred from haplotype diversity at the bz locus. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(47), 17644–9. http://doi.org/10.1073/pnas.0603080103

Wright, J. A., Smith, R. C., Li, X., Craig, N. L., & Atkinson, P. W. (2013). IPB7 transposase behavior in Drosophila melanogaster and Aedes aegypti. *Insect Biochemistry and Molecular Biology*, *43*(10), 899–906. http://doi.org/10.1016/j.ibmb.2013.06.009

Xu, Z., Yan, X., Maurais, S., Fu, H., O'Brien, D. G., Mottinger, J., & Dooner, H. K. (2004). Jittery, a Mutator distant relative with a paradoxical mobile behavior: excision without reinsertion. *The Plant Cell*, *16*(5), 1105–14. http://doi.org/10.1105/tpc.019802

Yu, Z., Wright, S. I., & Bureau, T. E. (2000). Mutator-like Elements in Arabidopsis thaliana: Structure, Diversity and Evolution. *Genetics*, *156*(4), 2019–2031.

Adelman, Z. N., Jasinskiene, N., & James, A. A. (2002). Development and applications of transgenesis in the yellow fever mosquito, Aedes aegypti. *Molecular and Biochemical Parasitology*, *121*(1), 1–10. http://doi.org/10.1016/S0166-6851(02)00028-2

Arensburger, P., Kim, Y.-J., Orsetti, J., Aluvihare, C., O'Brochta, D. A., & Atkinson, P. W. (2005). An active transposable element, Herves, from the African malaria mosquito Anopheles gambiae. *Genetics*, *169*(2), 697–708. http://doi.org/10.1534/genetics.104.036145

Barker, R. F., Thompson, D. V, Talbot, D. R., Swanson, J., & Bennetzen, J. L. (1984). Nucleotide sequence of the maize transposable element Mul. *Nucleic Acids Research*, *12*(15), 5955–67.

Benito, M., & Walbot, V. (1997). Characterization of the maize Mutator transposable element MURA transposase as a DNA-binding protein. *Mol. Cell. Biol.*, *17*(9), 5165–5175.

Bennetzen, J. L. (1996). The Mutator transposable element system of maize. *Current Topics in Microbiology and Immunology*, *204*, 195–229.

Brust-Mascher, I., & Scholey, J. M. (2009). Microinjection techniques for studying mitosis in the Drosophila melanogaster syncytial embryo. *Journal of Visualized Experiments : JoVE*, (31). http://doi.org/10.3791/1382

Casals, F., Cáceres, M., & Ruiz, A. (2003). The foldback-like transposon Galileo is involved in the generation of two different natural chromosomal inversions of Drosophila buzzatii. *Molecular Biology and Evolution*, *20*(5), 674–85. http://doi.org/10.1093/molbev/msg070

Chalvet, F., Grimaldi, C., Kaper, F., Langin, T., & Daboussi, M.-J. (2003a). Hop, an active Mutator-like element in the genome of the fungus Fusarium oxysporum. *Molecular Biology and Evolution*, *20*(8), 1362–75. http://doi.org/10.1093/molbev/msg155

Chalvet, F., Grimaldi, C., Kaper, F., Langin, T., & Daboussi, M.-J. (2003b). Hop, an active Mutator-like element in the genome of the fungus Fusarium oxysporum. *Molecular Biology and Evolution*, *20*(8), 1362–75. http://doi.org/10.1093/molbev/msg155

Davidson, A. E., Gratsch, T. E., Morell, M. H., O'Shea, K. S., & Krull, C. E. (2009). Use of the Sleeping Beauty transposon system for stable gene expression in mouse embryonic stem cells. *Cold Spring Harbor Protocols*, *2009*(8), pdb.prot5270. http://doi.org/10.1101/pdb.prot5270

Dietrich, C. R., Cui, F., Packila, M. L., Li, J., Ashlock, D. A., Nikolau, B. J., & Schnable, P. S. (2002). Maize Mu transposons are targeted to the 5' untranslated region of the gl8 gene and sequences flanking Mu target-site duplications exhibit nonrandom nucleotide composition throughout the genome. *Genetics*, *160*(2), 697–716.

Doseff, A., Martienssen, R., & Sundaresan, V. (1991). Somatic excision of the Mu1 transposable element of maize. *Nucleic Acids Research*, *19*(3), 579–84.

Gao, D. (2012). Identification of an active Mutator-like element (MULE) in rice (Oryza sativa). *Molecular Genetics and Genomics : MGG*, *287*(3), 261–71. http://doi.org/10.1007/s00438-012-0676-x

Ghildiyal, M., & Zamore, P. D. (2009). Small silencing RNAs: an expanding universe. *Nature Reviews. Genetics*, *10*(2), 94–108. http://doi.org/10.1038/nrg2504

Hanada, K., Vallejo, V., Nobuta, K., Slotkin, R. K., Lisch, D., Meyers, B. C., … Jiang, N. (2009). The functional role of pack-MULEs in rice inferred from purifying selection and expression profile. *The Plant Cell*, *21*(1), 25–38. http://doi.org/10.1105/tpc.108.063206

Handler, A. M. (2004). Understanding and improving transgene stability and expression in insects for SIT and conditional lethal release programs. *Insect Biochemistry and Molecular Biology*, *34*(2), 121–30. http://doi.org/10.1016/j.ibmb.2003.08.005

Holligan, D., Zhang, X., Jiang, N., Pritham, E. J., & Wessler, S. R. (2006). The transposable element landscape of the model legume Lotus japonicus. *Genetics*, *174*(4), 2215–28. http://doi.org/10.1534/genetics.106.062752

Hua-Van, A., & Capy, P. (2008). Analysis of the DDE motif in the Mutator superfamily. *Journal of Molecular Evolution*, *67*(6), 670–81. http://doi.org/10.1007/s00239-008-9178-1

Jiang, N., Bao, Z., Zhang, X., Eddy, S. R., & Wessler, S. R. (2004). Pack-MULE transposable elements mediate gene evolution in plants. *Nature*, *431*(7008), 569–73. http://doi.org/10.1038/nature02953

Jiang, N., Ferguson, A. A., Slotkin, R. K., & Lisch, D. (2011). Pack-Mutator-like transposable elements (Pack-MULEs) induce directional modification of genes through biased insertion and DNA acquisition. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(4), 1537–42. http://doi.org/10.1073/pnas.1010814108

Li, Y., Harris, L., & Dooner, H. K. (2013). TED, an autonomous and rare maize transposon of the mutator superfamily with a high gametophytic excision frequency. *The Plant Cell*, *25*(9), 3251–65. http://doi.org/10.1105/tpc.113.116517

Lisch, D. (2002). Mutator transposons. *Trends in Plant Science*, *7*(11), 498–504. http://doi.org/10.1016/S1360-1385(02)02347-6

Lisch, D., Chomet, P., & Freeling, M. (1995). Genetic characterization of the Mutator system in maize: behavior and regulation of Mu transposons in a minimal line. *Genetics*, *139*(4), 1777–96

Marquez, C. P., & Pritham, E. J. (2010). Phantom, a new subclass of Mutator DNA transposons found in insect viruses and widely distributed in animals. *Genetics*, *185*(4), 1507–17. http://doi.org/10.1534/genetics.110.116673

Neuvéglise, C., Chalvet, F., Wincker, P., Gaillardin, C., & Casaregola, S. (2005). Mutator-like element in the yeast Yarrowia lipolytica displays multiple alternative splicings. *Eukaryotic Cell*, *4*(3), 615–24. http://doi.org/10.1128/EC.4.3.615-624.2005

Pritham, E. J. (2009). Transposable elements and factors influencing their success in eukaryotes. *The Journal of Heredity*, *100*(5), 648–55. http://doi.org/10.1093/jhered/esp065

Robertson, D. S. (1980). The timing of mu activity in maize. *Genetics*, *94*(4), 969–78.

Sarkar, A., Coates, C. J., Whyard, S., Willhoeft, U., Atkinson, P. W., & O'Brochta, D. A. (1997). The Hermes element from Musca domestica can transpose in four families of cyclorrhaphan flies. *Genetica*, *99*(1), 15–29.

Silber, J., Bazin, C., Lemeunier, F., Aulard, S., & Volovitch, M. (1989). Distribution and conservation of the foldback transposable element in Drosophila. *Journal of Molecular Evolution*, *28*(3), 220–4.

Smith, R. C., & Atkinson, P. W. (2011). Mobility properties of the Hermes transposable element in transgenic lines of Aedes aegypti. *Genetica*, *139*(1), 7–22. http://doi.org/10.1007/s10709-010-9459-7

Sousa, A., Bourgard, C., Wahl, L. M., & Gordo, I. (2013). Rates of transposition in Escherichia coli. *Biology Letters*, *9*(6), 20130838. http://doi.org/10.1098/rsbl.2013.0838

Talbert, L., & Chandler, V. (1988). Characterization of a highly conserved sequence related to mutator transposable elements in maize. *Mol. Biol. Evol.*, *5*(5), 519–529.

Tan, B.-C., Chen, Z., Shen, Y., Zhang, Y., Lai, J., & Sun, S. S. M. (2011). Identification of an active new mutator transposable element in maize. *G3 (Bethesda, Md.)*, *1*(4), 293–302. http://doi.org/10.1534/g3.111.000398

Wang, Q., & Dooner, H. K. (2006). Remarkable variation in maize genome structure inferred from haplotype diversity at the bz locus. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(47), 17644–9. http://doi.org/10.1073/pnas.0603080103

Wright, J. A., Smith, R. C., Li, X., Craig, N. L., & Atkinson, P. W. (2013). IPB7 transposase behavior in Drosophila melanogaster and Aedes aegypti. *Insect Biochemistry and Molecular Biology*, *43*(10), 899–906. http://doi.org/10.1016/j.ibmb.2013.06.009

Xu, Z., Yan, X., Maurais, S., Fu, H., O'Brien, D. G., Mottinger, J., & Dooner, H. K. (2004). Jittery, a Mutator distant relative with a paradoxical mobile behavior: excision without reinsertion. *The Plant Cell*, *16*(5), 1105–14. http://doi.org/10.1105/tpc.019802

Yu, Z., Wright, S. I., & Bureau, T. E. (2000). Mutator-like Elements in Arabidopsis thaliana: Structure, Diversity and Evolution. *Genetics*, *156*(4), 2019–2031.

Table 3.1 Somatic transposition of *Muta1* in *D. melanogaster* and *Ae. aegypti*. The table describes the plasmid used for the transposition assay. The *Muta1* element flanked by an 8 bp TSDs or a 9 bp TSDs, number of experiments performed with each plasmid, the number of embryos that were injected, the donot titer obtained for the *Muta1* element and the *piggyBac* element from the transpostion assays, and the number of transposition events recovered from each set of experiments. The transposition frequency for the *Muta1* and the *piggyBac* elements was determined by dividing total number of transpostion events obtained to the donor titer. The zero control transposition assays demonstrate transposition activity in absense of the *Muta1* transposase, this expreiment used the *piggyBac* element with the *piggyBac* helper plasmid as an internal control. Thus, it shows transposition frequency for the *piggyBac* element.

Insect Species	*Muta1* Transposase	TSD Length	No. of Expts.	No. of Embryos Injected	*Muta1* Donor Plasmid Titer	*piggyBac* Donor Plasmid Titer	No. of *Muta1* Transpos itions	No. of *piggyBac* Transpos itions	*Muta1* Transposition Frequency (std. dev.)	*PiggyBac* Transpositi on Frequency (std. dev.)
D. melanogaster	+	8 bp	4	444	127,200	154,400	51	69	5.7×10^{-4} (2.8×10^{-4})	3.9×10^{-4} (8.2×10^{-5})
"	-	8 bp	2	200	319,000	30,800	0	21	0	6.8×10^{-4} (5.0×10^{-6})
"	+	9 bp	3	300	765,000	709,400	254	217	3.96×10^{-4} (1.67×10^{-4})	3.05 x 10-4 (9.65×10^{-5})
"	-	9 bp	2	200	732,000	993,600	0	215	0	2.18×10^{-4} (2.25×10^{-5})
Ae. aegypti	+	8 bp	2	120	102,000	128,800	1,234	74	8.3×10^{-3} (6.8×10^{-3})	5.2×10^{-4} (2.3×10^{-4})
"	-	8 bp	3	250	779,000	521,800	0	20	0	3.7×10^{-5} (3.5×10^{-5})
"	+	9 bp	3	300	359,800	257,000	3,306	192	9.45×10^{-2} (1.29×10^{-3})	7.19×10^{-4} (1.23 x 10-4)
"	-	9 bp	3	300	841,800	331,200	0	155	0	2.84×10^{-4} (0.7×10^{-4})

Table 3.2 Excision of *Muta1* in developing embryos of *D. melanogaster* and *Ae. aegypti*. The two-plasmid assays were performed with the *Muta1* donor plasmid and the helper plasmid. The donor plasmid contained the *Muta1* element flanked an 8 bp TSDs or a 9 bp TSDs. The table describes total number of experiments performed, the donor titer obtained, and the number of excision events recovered. The excision frequency was calculated by dividing total number of excision events to the donor titer.

Insect Species	Transposase	TSD Length	Number of Experiments	Number of Embryos Injected	Donor Plasmids Titer	Number of Excisions	Excision Frequency (Std. dev.)
D. melanogaster	+	8 bp	5	250	91,105	2,484	1.4×10^{-2} (1.7×10^{-2})
"	+	9 bp	2	200	168,000	3,780	2.24×10^{-2} (1.4×10^{-3})
"	-	8 bp	7	350	94,383	0	0
Ae. aegypti	+	8 bp	4	250	199,516	486	1.9×10^{-3} (6.8×10^{-4})
		9 bp	3	150	3,600	32	8.73×10^{-3} (3.85×10^{-3})
"	-	8 bp	3	150	109,272	0	0

Table 3.3 Sequences of TSDs generated by transposition of *Muta1* into the target plasmid pGDV1.

A. TSD and locations of *Muta1* transpositions events into target plasmid for *Muta1* element in donor plasmid flanked by 8 bp TSD in *D. melanogaster.*

Insertion Site in Target Plasmid (bp)	TSD Sequence	Orientation	TSD Length (bp)
220	GTAAAAAAA	-	9
308	CATAAAATC	-	9
398	GAATATGAC	+	9
409	AGTAAAAG	+	8
592	TGATAAAGC	+	9
676	CTAATAAAT	-	9
725	GTTGAAGTT	+	9
819	GTTTCAAAA	-	9
887	TGAAGAAGG	+	9
895	GCCTTCTTC	-	9
1014	GTTTTTGCA	-	9
1993	CTCTAGAG	-	8
2001	CCGGGGATC	-	9
2032	GTGGCAAAG	-	9
2164	TGTTCTATA	-	9
2184	CGAGAAAAC	+	9
2184	CGAGAAAAC	+	9
2199	GTTGGAATG	+	9
2225	CTAACAAG	-	8
2394	GTACATACT	-	9
2547	GTCGTTCAC	-	9

B. TSD and locations of *Muta1* transpositions events into target plasmid for *Muta1* element in donor plasmid flanked by 8 bp TSD in *Ae. aegypti.*

Insertion Site in Target Plasmid (bp)	TSD Sequence	Orientation	TSD Length (bp)
15	AATCTTGTA	+	9
198	CTAATAGCC	+	9
306	CTGATTTTA	+	9
336	GGCTTTGGG	+	9
349	GTTTGTAAT	-	9
411	TAAAAGCAG	+	9
515	ATGGAAGAT	+	9

554	TTCAAAATC	-	9
580	CTACGCAAT	+	9
676	ATTTATTAG	+	9
699	AACAAGAAC	+	9
788	ATTCCAGTA	-	9
886	CCTTCTTCA	-	9
907	ACTTGAAAG	+	9
920	GCTAAATAT	+	9
979	AATGAAATC	+	9
2296	CCTACAGGG	-	9
2543	GTTTGTGAA	+	9

C. TSD and locations of *Muta1* transpositions events into target plasmid for *Muta1* element in donor plasmid flanked by 9 bp TSD in *D. melanogaster.*

Insertion Site in Target Plasmid (bp)	TSD Sequence	Orientation	TSD Length (bp)
54	CATCAAGAT	+	9
62	ATCTTGATG	-	9
71	CAAATATAT	-	9
223	AAAATTACC	-	9
411	CTGCTTTT	-	8
624	GTCGTAATG	+	9
864	GCTTGTTCT	-	9
1139	GATTTTTTC	-	9
2072	ATGGGATAT	+	9
2223	TTAGCTAAC	-	9

D. TSD and locations of *Muta1* transpositions events into target plasmid for *Muta1* element in donor plasmid flanked by 9 bp TSD in *Ae. aegypti.*

Insertion Site in Target Plasmid (bp)	TSD Sequence	Orientation	TSD Length (bp)
26	TTTAGATAT	+	9
65	ATATTTGGG	+	9
219	ATTTTTTTA	+	9
301	CTTTTTAGC	-	9
623	GGTCGTAAT	+	9
1028	GATTTTTG	+	8
2183	CGAGAAAAC	+	9
2283	GTTGAGTTG	-	9
2389	AGTAAGTAT	+	9
2416	GTTTTAACT	+	9

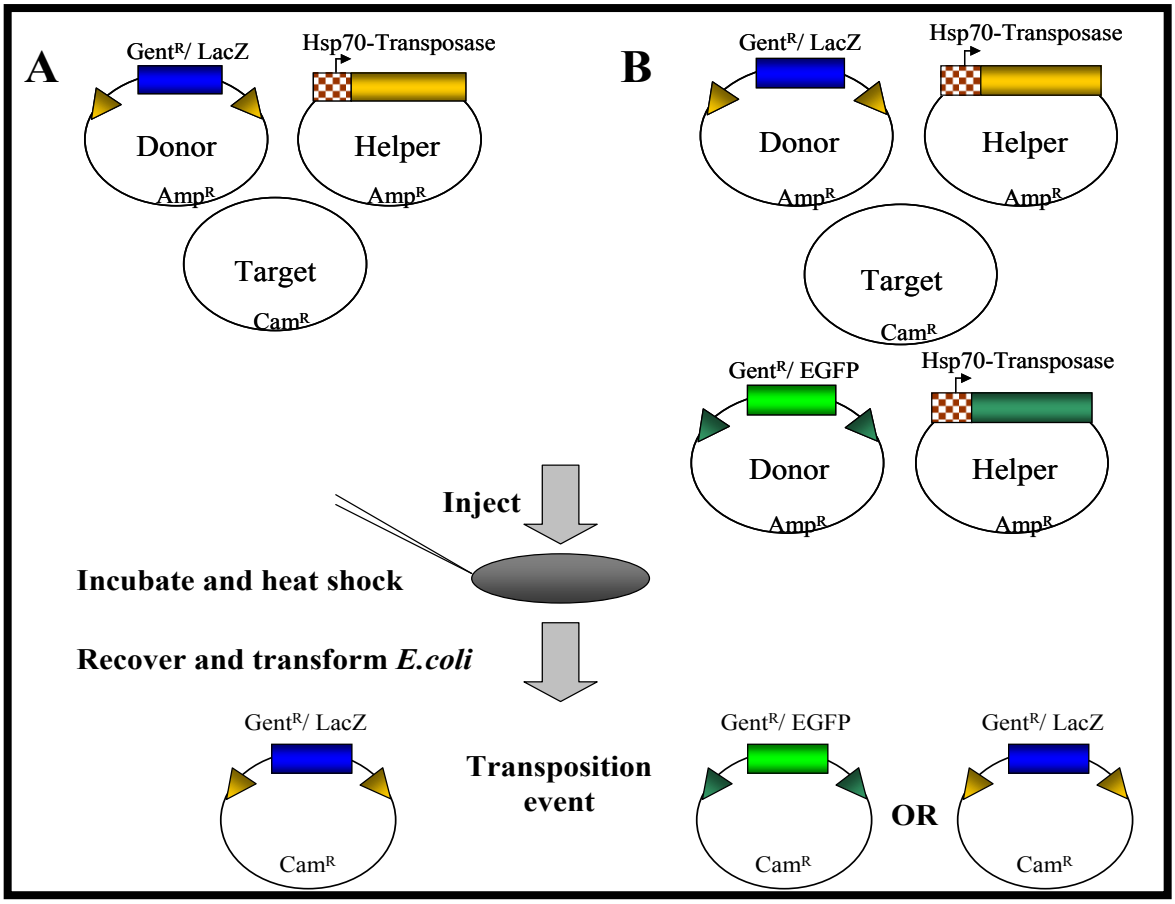
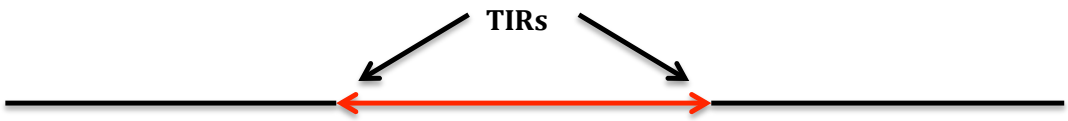


Figure 3.1 Five plasmid normalized transposition assay. (Wright, Smith, Li, Craig, & Atkinson, 2013). The donor plasmid contained the *Muta1* ends with *LacZ* reporter gene, the helper plasmid contained *Muta1* transpose under Hsp70 promoter, the *piggyBac* donor plasmid contained EGFP reporter gene with *piggyBa*c ends, *piggyBac* helper plasmid contained *piggyBac* transposase under Hsp70 promoter, and the pGDV1 is the target plasmid with chloramphenicol resistance gene. The five-plasmid injection mix is injected into embryos, and transformed into *E. coli.* The recombinant plasmids for *Muta1* and *piggyBac* transposition event are screened for *LacZ*/Cam^R and EGFP/Cam^R, respectively



…….GGTACC**GACGCAGT** **GACGCAGT(AG)**GAGCTC……

Kpn1 **TSD** **TSD** *Sac1*

12…. GGTACC**GACGCAGT** **(AG)**GAGCTC…

2…… 40 bp Deletion 40 bp Deletion…

1…. GGTACC**GACGCAGGT** **(AG)**GAGCTC…

1…. GGTA **(G)**GAGCTC…

1…. GGTACC**GA** GAGCTC…

1…. 34 bp Deletion 36 bp Deletion….

Figure 3.2 The figure shows excision pattern of the *Muta1* flanked by an 8 bp TSDs and the respective enzyme sites in *D. melanogaster*. The *Muta1* element in the donor plasmid has AG dinucleotide flanking to the right end TSD. These extra nucleotides do not interfere with the transposition mechanism. 1) Precise excision has a TSD flanked by *Kpn1* and (AG) *Sac1* restriction sites. The other patterns of excision were 2) excision outside the left end and right end in the flanking plasmid DNA, 3) TSD with one additional base flanked by their respective restriction sites, 4) cleavage in the *Kpn1* restriction site, 4) cleavage in the TSD flanked by their respective restriction sites, and 5) excision in the flanking plasmid DNA.

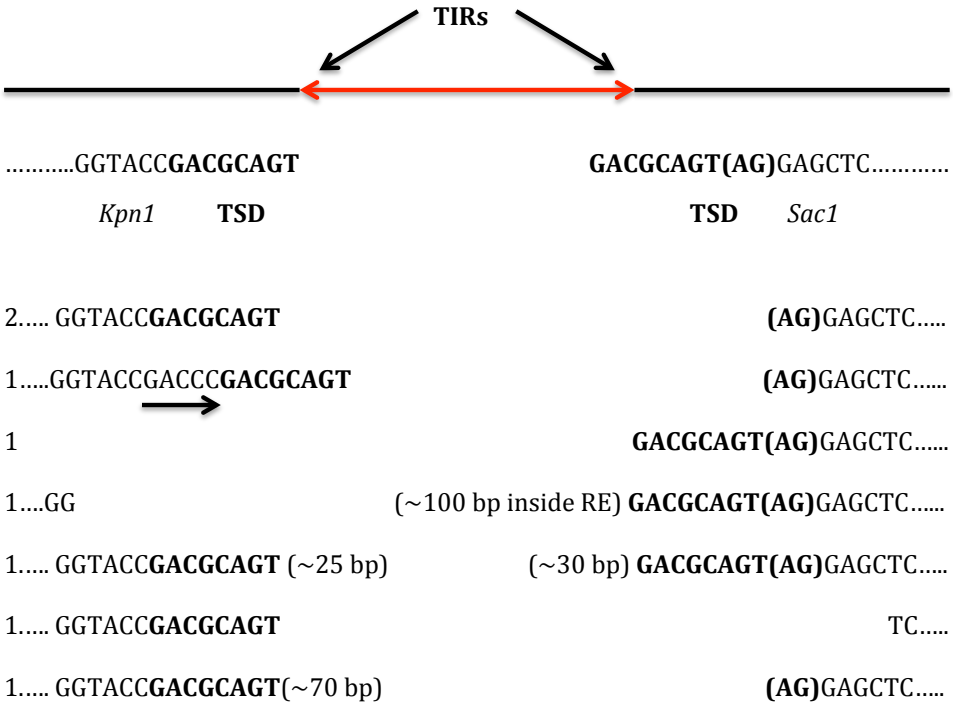


Figure 3.3 The figure shows excision pattern of the *Muta1* flanked by an 8 bp TSDs and the respective enzyme sites in *Ae. aegypti*. The *Muta1* element in the donor plasmid has an extra AG dinucleotide flanking to the right end TSD. These extra nucleotides do not interfere with the transposition mechanism. 1) Precise excision has a TSD flanked by *Kpn1* and (AG) *Sac1* restriction sites. The other patterns of excision were 2) Excision acquiring five extra bases flanked to TSD, 3) Excision outside the left end in the flanking plasmid DNA, 4) Cleavage in the *Kpn1* restriction site flanked by a TSD and *Sac1* site 5) Cleavage in the right and left end flanked by TSDs and their restriction site, 6) Cleavage in the *Sac1* site flanked by a TSD, and 7) cleavage in the left end flanked by a TSD and their restriction site

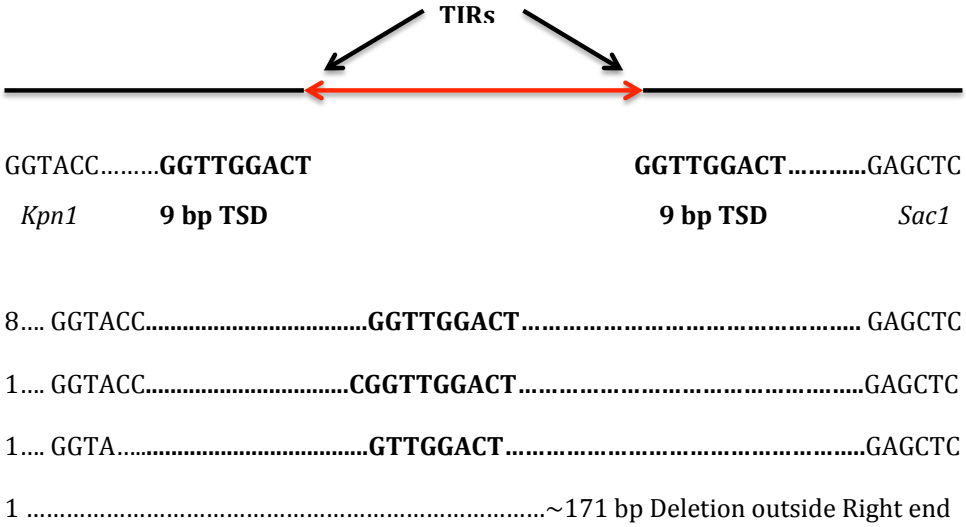


Figure 3.4 The figure shows excision pattern of *Muta1* flanked by a 9 bp TSDs and their respective enzyme sites in *D. melanogaster*. 1) Precise excision has a single TSD flanked by *Kpn1* and *Sac1* restriction sites. Other patterns of excision were 2) TSD flanked by one extra base with respective restriction sites, 3) TSD with cleavage in the *Kpn1* restriction site, and 4) excision outside the right end in the flanking plasmid DNA.

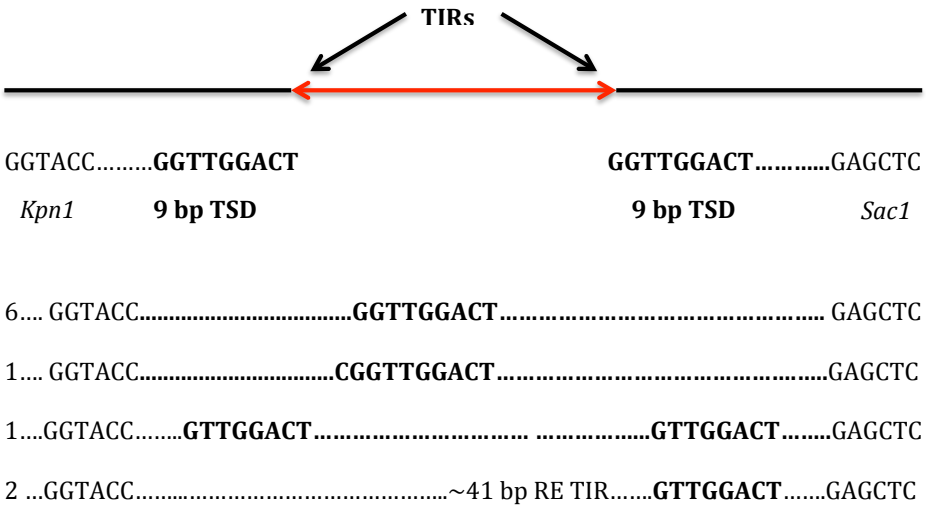


Figure 3.5 The figure shows excision pattern of *Muta1* flanked by a 9 bp TSDs and their respective enzyme sites in *D. melanogaster*. 1) Precise excision has one with a single TSD flanked by *Kpn1* and *Sac1* restriction sites. Other patterns of excision were 2) with a TSD flanked by one extra base, 3) two TSDs flanked by restriction sites, and 4) TSD with 41 bp of the right end.

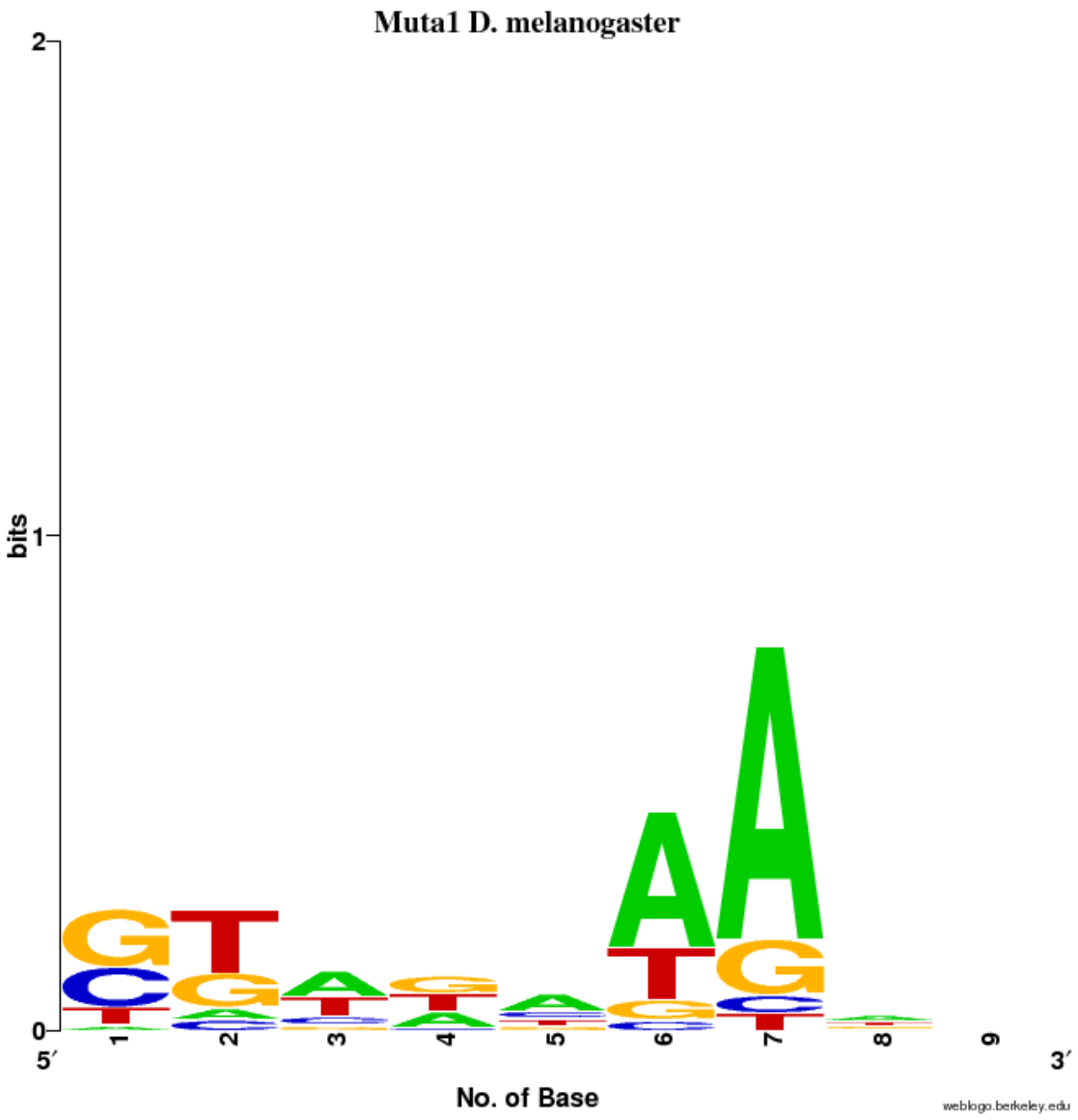


Figure 3.6 WebLogo of consensus target site preference of *Muta1* flanked by an 8 bp TSD in *D. melanogaster*. The figure shows the preference for 'A' at the 7th and 8th position in the TSDs sequence obtained from the transposition events.

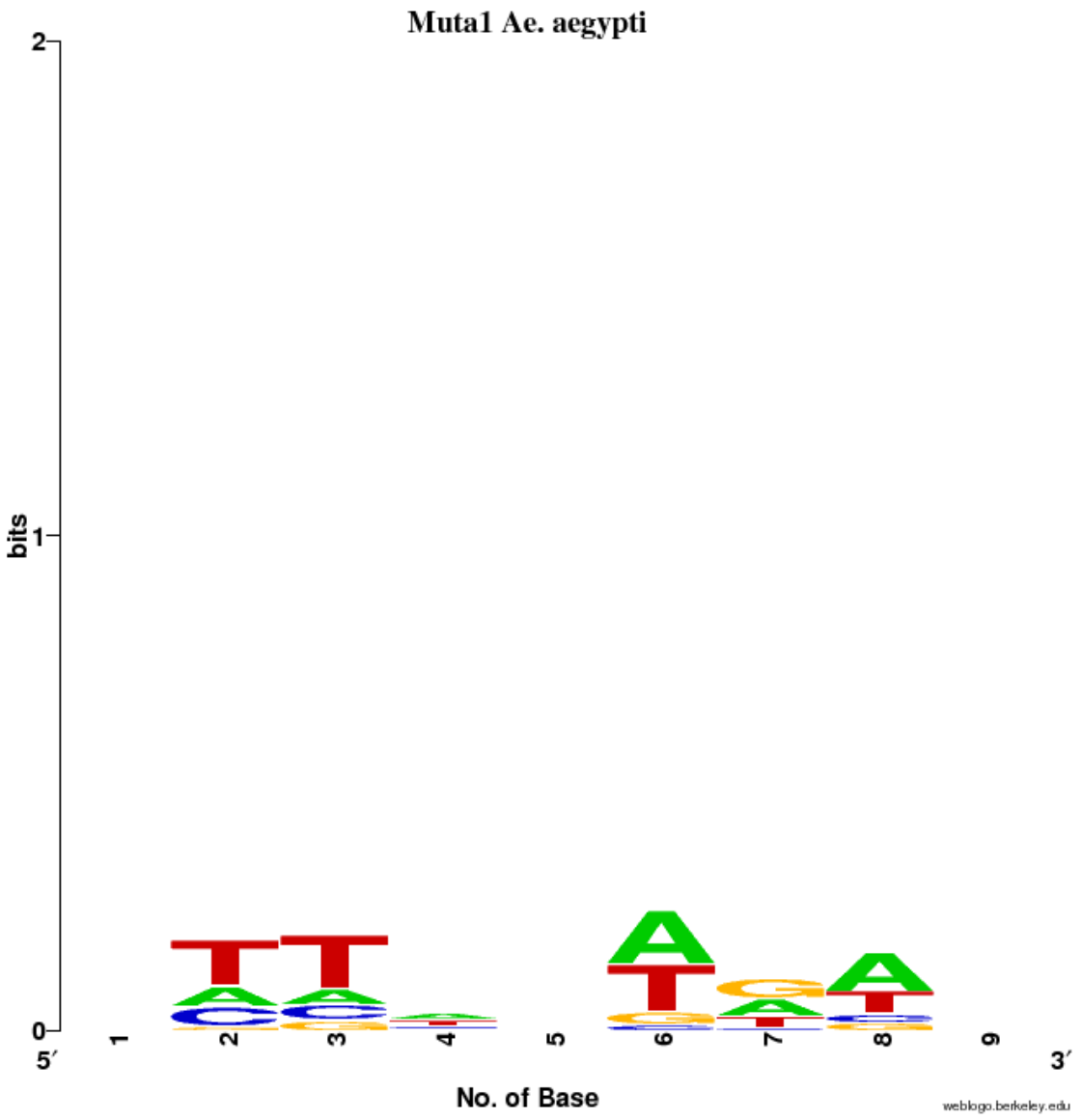


Figure 3.7 WebLogo of consensus target site preference of *Muta1* flanked by an 8 bp TSD in *Ae. aegypti.* The figure shows the preference for 'T' at the 2nd and 3rd position, and 'A' at the 6th and 8th position in the TSDs sequence obtained from the transposition events.

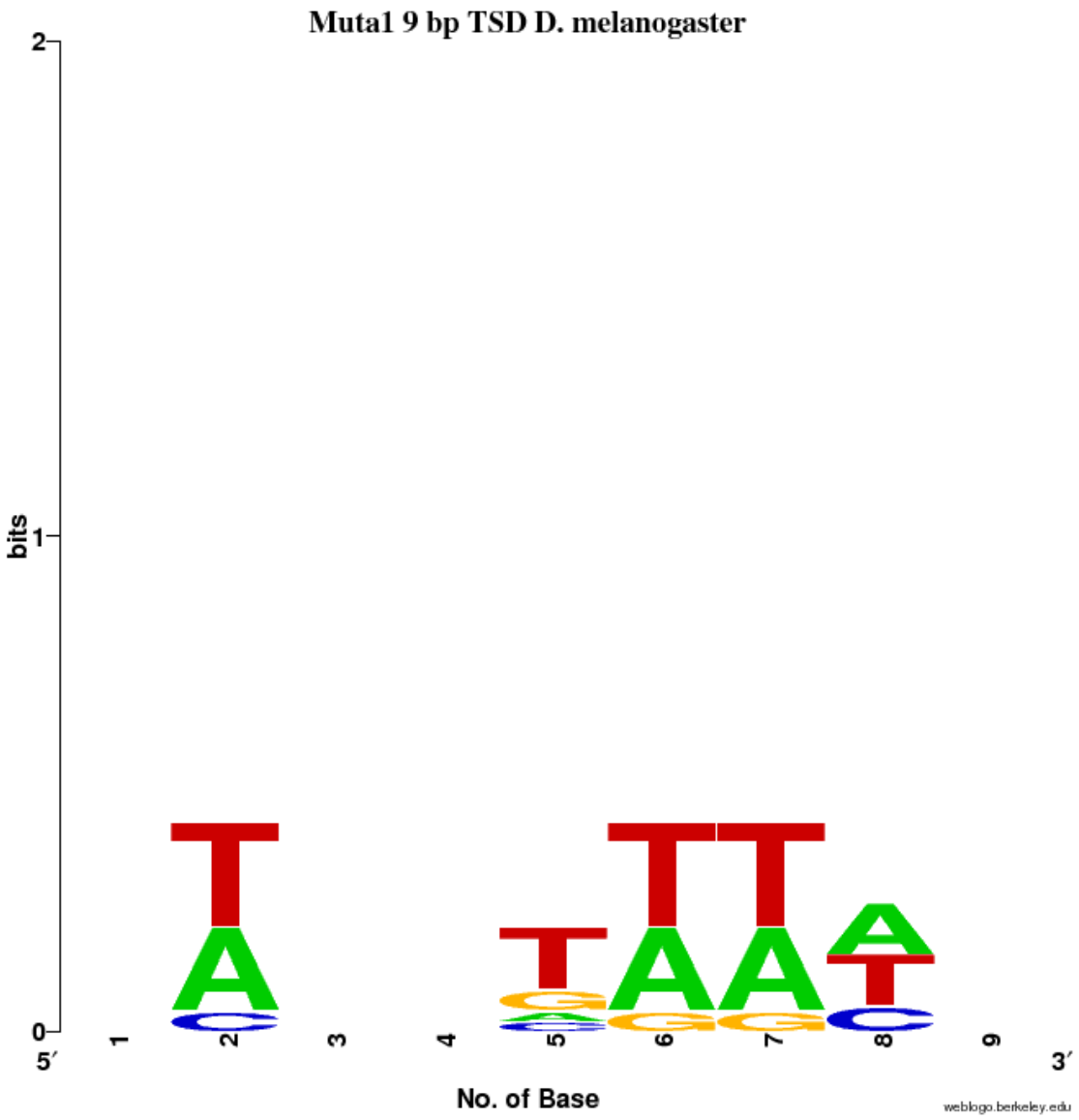


Figure 3.8 WebLogo of consensus target site preference of *Muta1* flanked by a 9 bp TSD in *D. melanogaster*. The figure shows the preference for 'T' at the 2nd, 5th, 6th, 7th and 'A' at the 8th position in the TSDs sequence obtained from the transposition events.

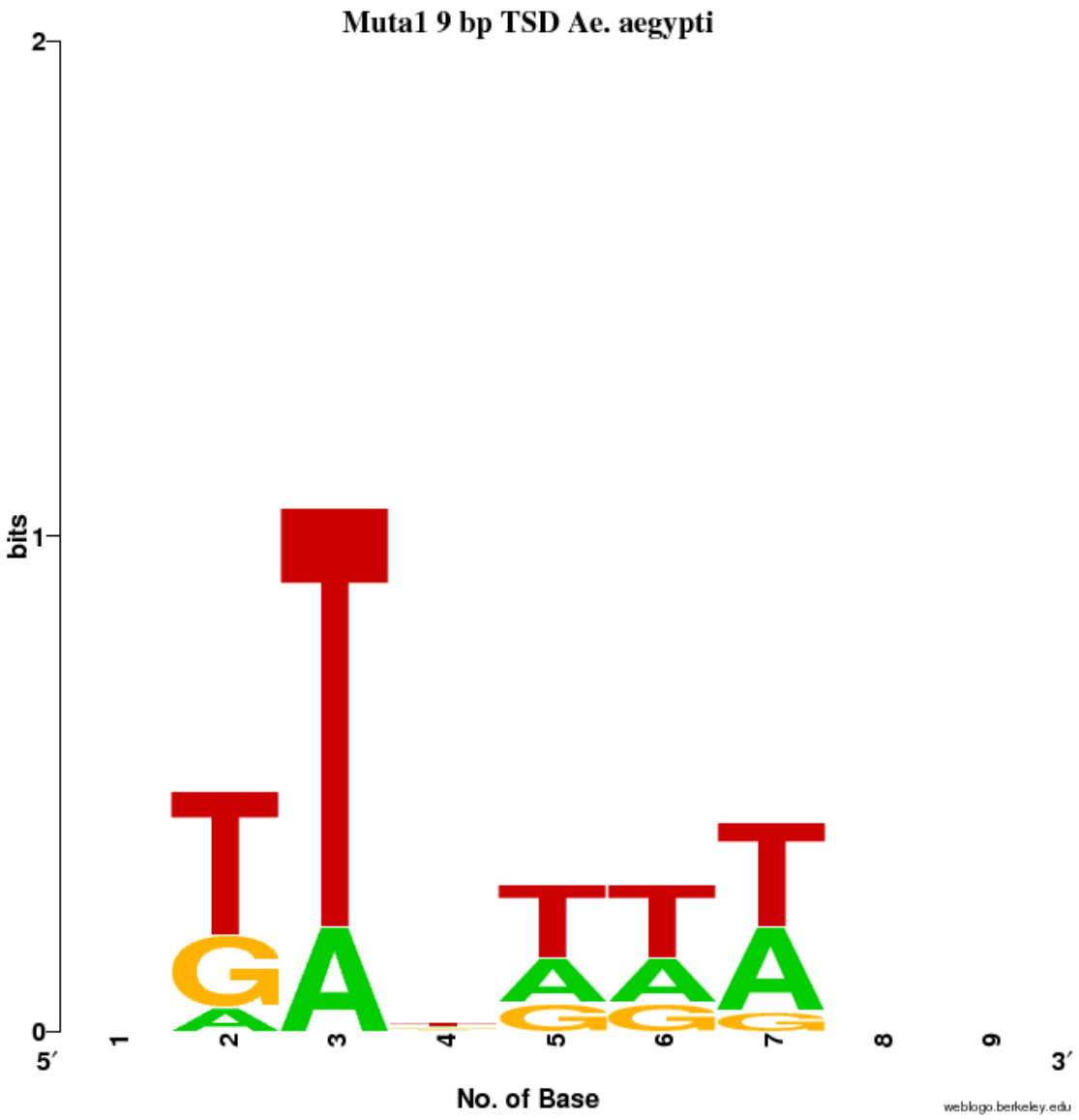


Figure 3.9 WebLogo of consensus target site preference of *Muta1* flanked by a 9 bp TSD in *Ae. aegypti*. The figure shows the preference for 'T' at the 2nd, 3rd, 5th, 6th, and 7th position in the TSDs sequence obtained from the transposition events.

Chapter 4

Muta1 is active in the germline of *D. melanogaster* and *Ae. aegypti*.

4.1 Abstract

The germline activity of the *Muta1* element in *D. melanogaster* and *Ae. aegypti* was assayed using a non-autonomous *Muta1* element containing either the DsRed marker or white marker genes. The *Muta1* germline transformation rate for a non-autonomous *Muta1* in *D. melanogaster* was 5.6% with the *white* gene and 14.8% with the DsRed marker. The remobilization of *Muta1* was determined relative to the rate of *piggyBac* remobilization. The transformation frequency for this second, *Muta1* element was 4.9% and for the *piggyBac* element was 14.8%. New phenotypes arising from the *Muta1* remobilization were observed with DsRed expression in full body, thorax, abdomen, labellum and genitalia. The remobilization rate for the *Muta1* element was 21.2% in *D. melanogaster*. Despite containing endogenous copies of the *Muta1* transposon, an engineered *Muta1* transposon containing the DsRed genetic marker was found to be able to be used successfully genetically transform its host, the transformation rate being 4% in *Ae. aegypti*. Remobilization experiments were also performed in *Ae. aegypti*. Two potential new phenotypes were observed for the *Muta1* element in *Ae. aegypti* with DsRed expression in the legs and in the midgut. Molecular characterization of the phenotype with DsRed expression in legs revealed to be a remobilization event of the *Muta1* element. Thus, this was the first evidence of a transposon being able to remobilize post-integration in *Ae. aegypti.*

4.2 Introduction

Genetic transformation in *Drosophila* was a breakthrough with the use of *P* element for genetic manipulation (Bingham, Kidwell, & Rubin, 1982) (Spradling et al., 1999). The *P*-element was discovered as a result of hybrid dysgenesis, a phenomenon that results in progeny with variety of phenotypes such as sterility, high mutation rate and high frequency of chromosomal aberration and nondisjunction (Rubin & Spradling, 1982). The *P*-element was subsequently found to be restricted to use in *D. melanogaster*; germ-line transformation of the *P* element into non-drosophilids has not proven successful in other insect species (O'Brochta & Handler, 1988). As a consequence a search for new transposons that could function in non-drosophilid species was commenced (Jasinskiene et al., 1998) (David A O'Brochta et al., 2003). Other transposons that were subsequently discovered and showed activity in Drosophilids include *piggyBac*, which has proven to be a highly efficient genetic tool, due to a high transformation efficiency and precise excision (Fraser, Ciszczon, Elick, & Bauser, 1996) (Handler & Harrell, 1999). In addition, *piggyBac* preferentially inserts within genes, with 50 to 67% of insertion occurring within transcriptional units in *D. melanogaster,* which is useful for discovery of new genes and promoters through enhancer trapping (Berg & Spradling, 1991). The *Minos* element, isolated from *D. hydei,* belongs to the *Tc1/mariner* superfamily and has also been shown to transform other insect species (Franz & Savakis, 1991) (Loukeris, Arcà, Livadaras, Dialektaki, & Savakis, 1995) (Catteruccia et al., 2000). *Hermes* is a *hAT* element isolated from *Musca domestica* and is used for genetic transformation of *D. melanogaster* and *Schizosaccharomyces pombe* (Evertts, Plymire,

Craig, & Levin, 2007). It is also able to remobilize within the germline of *D. melanogaster* at a rate of approximately 0.03 jumps per element per generation (Guimond, Bideshi, Pinkerton, Atkinson, & O'Brochta, 2003). The *Mos1* element efficiently transforms *D. mauritiana* but is almost immobile in *D. melanogaster* (Bryan, Jacobson, & Hartl, 1987) (Lidholm, Lohe, & Hartl, 1993) (David A O'Brochta et al., 2003).

The *piggyBac* element has been very successful in *Drosophila* for gene tagging and enhancer trapping, but has failed to retain both germline and somatic activity in transgenic lines of *Ae. aegypti* (Palavesam, Esnault, & O'Brochta, 2013). The transformation rate of *Hermes* in *Ae. aegypti* is less than 10% compared to *D. melanogaster,* where the transformation rate is greater than 50% (D. A. O'Brochta, Warren, Saville, & Atkinson, 1996) (Jasinskiene et al., 1998). The *Mos1* element has been used for transformation in *Ae. aegypti* with a 4% transformation rate (Coates, Jasinskiene, Miyashiro, & James, 1998). The integration pattern of *Mos1* was found to be similar to the *Hermes* element, which integrates along with flanking plasmid DNA (Wilson et al., 2003). The post-integration mobilization of *Mos1* has not been observed in *D. melanogaster* (Lozovsky, Nurminsky, Wimmer, & Hartl, 2002) (A. R. Lohe, Lidholm, & Hartl, 1995). Immobility of *Mos1* was attributed to the requirement of the three regions of sequences spread throughout the element which are located 200 bp from the ends (Allan R. Lohe & Hartl, 2002).

The *Mutator* superfamily is the most widespread family of transposons with elements

discovered in plants (D Lisch, Chomet, & Freeling, 1995), fungi (Chalvet, Grimaldi, Kaper, Langin, & Daboussi, 2003), bacteria (Eisen, Benito, & Walbot, 1994), protozoans (Pritham, Feschotte, & Wessler, 2005), metazoans (Hua-Van & Capy, 2008), and insect viruses (Marquez & Pritham, 2010). The majority of investigation of *Mutator* transposons was performed in maize, with the *MuDR* transposon being the most commonly studied transposon of this superfamily (D Lisch et al., 1995) (Manish N. Raizada & Walbot, 2000) (M N Raizada, Benito, & Walbot, 2001) (Damon Lisch, 2002) (McCarty et al., 2005). Studies of *Mutator* elements in maize have demonstrated that the introduction of exogenous *MuDR* transposase into lines where endogenous *MuDR* elements have been silenced via cytosine methylation in the TIRs, results in a demethylation of endogenous *MuDR* elements followed by their somatic excision in maize (Manish N. Raizada & Walbot, 2000). The means by which this silencing occurs is currently unknown. However, it is believed that methylation and inactivation of *MuDR* elements might be regulated by host genes (Damon Lisch, 2002).

Jittery, a *Mutator* transposon had demonstrated high excision frequency in maize, causing somatic and germinal reversion, but has failed to generate new insertions (Xu et al., 2004). It has been proposed that *Mutator* elements use a cut and paste mechanism in somatic cells, and in germline cells they use copy and insert mechanism (Tan et al., 2011). In somatic cells, the *Mutator* transposon excises itself and reinserts it in a new location in the genome. Moreover, the high frequency of excision of *Mutator* element was restricted to late stage of cells involved in development during organogenesis (Tan et al., 2011). In the germline, element replicates just before meiosis or in the gametophyte

and inserts into a new location in the genome (Tan et al., 2011). Therefore, in the germline, element duplication and insertion may cause increase in copies of the *Mutator* transposons (Tan et al., 2011).

Here, I determined transformation and remobilization activity of the *Muta1* transposon in *D. melanogaster* and *Ae. aegypti*. In *D. melanogaster*, I also characterized the degree of preference for *Muta1* integrations into exons, introns, and intergenic regions.

4.3 Materials and Methods

Plasmid constructions

4.3.1 *Muta1 white* gene transformation donor

pMuta1whiteTFD was constructed through ligation of a fragment from pBSHermesw+ and the vector pBSMuta1LR (described in chapter three page section 3.3.4). The vector pBSMuta1LR was *EcoR1*-digested, treated with FastAP (Thermo-Fisher) and purified by agarose gel electrophoresis. The insert was *EcoR1*-digested *white* gene fragment, along with the 3xP3 promoter from pBSHermesw+ and purified by agarose gel electrophoresis. Vector and insert were ligated with T4 DNA ligase (Thermo-Fisher).

4.3.2 *Muta1* DsRed transformation donor plasmid.

pMuta1DsRedTFD was constructed through ligation of 3xP3-DsRed2-1fragment and the vector pBSMuta1LR. The vector pBSMuta1LR is digested using FastDigest *Xma1* and *Xba1*, treated with FastAP (Thermo-Fisher) and purified by agarose gel electrophoresis. The insert was *Sma1* and *AvrII* digested 3xP3-DsRed2-1 fragment. 3xP3-DsRed2-1

fragment (1.2 kb) was cut from pMos (3xP3-DsRed) ((Smith, Walter, Hice, O'Brochta, & Atkinson, 2007). The vector and insert were ligated with T4 DNA ligase (Thermo-Fisher).

4.3.3 *Muta1* helper plasmid

The plasmid is described in Chapter three section 3.3.5

4.3.4 *piggyBac* transformation donor plasmid

The pBac[3xP3-EGFP]af plasmid was previously described (Horn & Wimmer, 2000)

4.3.5 *piggyBac* helper plasmid

The *piggyBac* helper is phsp70-Bac (formerly pBhsΔSa) as described previously in chapter three page section 3.3.7 was used for transformation experiment (Handler *et al*. 1998).

4.3.6 pMuta13EHpBac *Muta1* donor plasmid for remobilization experiments.

pMuta13EHpBac was constructed through ligation of phsp70-pBac fragment from *piggyBac* helper and the vector pMuta1DsRedTFD. phsp70-pBac is digested with *EcoR1* and *HindIII* to yield 3.5 Kb, fragment were blunted using blunting enzyme (Pjet kit), gel-purified fragment (Zymoclean Gel DNA Recovery Kit, Zymo Research) and quantified on an agarose gel, to obtain *piggyBac* transposase fragment. This fragment is ligated into pMuta1DsRedTFD described above following digestion with *EcoRV* (Thermo-Fisher), treated with FastAP (Thermo-Fisher), gel purified 4.8 Kb fragment on an agarose gel electrophhoresis. Vector and insert were ligated with T4 DNA ligase (Thermo-Fisher). Insert was sequenced using pBacEGFPaf 3462 Rev and For2 for LE.

4.3.7 pBac3EHpMuta1 *Muta1* helper plasmid for remobilization experiments.

pBac3EHspMuta 1 was constructed through ligation of a PCR fragment of hspMuta1 from pMuta1 helper and the vector pBac3E(Afm). The vector was digested with FastDigest *AvrII* (Thermo-Fisher), treated with FastAP (Thermo-Fisher), and purified by agarose gel electrophoresis. The hspMuta1 fragment was amplified with the primers Nhe F and Nhe R using Phusion polymerase (New England Biolabs) and the program: 98° for 30 sec, 5 × (98° for 8 sec, 57° for 20 sec, 72° for 1.5 min), 25 × (98° for 8 sec, 67° for 20 sec, 72° for 1.5 min), 72° for 5 min, 4°. Column purified PCR product was digested with Nhe1 and purified again. Vector and insert were ligated with T4 DNA ligase (Thermo-Fisher).

4.3.8 Embryo microinjection in *D. melanogaster* for transformation experiments with non-autonomous *Muta1*.

The embryo microinjection was carried out as described in chapter three section 3.3.10. The plasmid mix used for microinjection had two plasmids, 250 ng/ul of pMuta1DsRedTFD and 250 ng/ul of pMuta1 helper. In second set of experiment, embryos were microinjected with plasmid mix that contained 250 ng/ul of pMuta1whiteTFD and 250 ng/ul of pMuta1 helper. The two experiments had different donor plasmids, first experiment had a donor plasmid with the DsRed2-1 marker and the second experiment had a donor plasmid with the *white* gene marker.

4.3.9 Embryo microinjection for remobilization experiments in *D. melanogaster*.

As described above embryos were injected with a plasmid mix that contains 250 ng/ul of pMuta13HsppBac and 250 ng/ul of pBac3EHspMuta1.

4.3.10 Embryo microinjection in *Ae. aegytpi* for remobilization experiments.

Muta1 donor (pMuta13HsppBac) and helper plasmid (pBac3EhspMuta1) were sent to the Insect Transformation Facility, University of Maryland, to generate transgenic mosquitoes. Recovered embryos were collected and eclosed embryos were used for pool mating, G1 progeny from this pool mating of 75 adults were then screened for DsRed and EGFP marker expression in larvae and pupae stage.

4.3.11 Establishment of crosses in *D. melanogaster* for remobilization experiments.

Two crosses were established, here 20 males from reporter line M32 with DsRed marker were cross with the 20 females from the helper line P32 and P42 with GFP has the marker, respectively. Similarly, reciprocal crosses were set up, here 20 females from reporter line M32 with DsRed marker were cross with the 20 males from the helper line P32 and P42 with GFP has the marker, respectively. Resulting reporter/helper progeny segregating with both fluorescent markers were collected based on the presence of both transgenic constructs. Reporter/helper hybrids were self-crossed with other reporter/helper siblings. After the crossing, 3 days later they were heat shocked at 37°C for 1 hour everyday till they reach adult stage. Progeny with both markers were screened for several generations to detect any change in phenotypic expression for DsRed marker, which may have indicated remobilization of this tagged *Muta1* transposon.

4.3.12 Establishment of crosses in *Ae. aegypti* for remobilization experiments.

Four crosses were established, 10 males from reporter line M1 with DsRed marker were cross with the 10 females from the helper line P11 and P14 with GFP has the marker; In second set of cross 10 males from line M2 with dsRed marker were cross with the 10 females from the helper line P9 and P11, respectively. Resulting reporter/helper progeny segregating with both fluorescent markers were collected based on the presence of both transgenic constructs. Reporter/helper hybrids were self-crossed. Progeny in embryo, larvae and pupae stage were heat shocked at 37°C for 2 hour everyday until they reached adult stage. Progeny with both markers were stored for characterization of new integration location for *Muta1* element and were screened for several generations to detect any change in phenotypic expression for DsRed marker.

4.3.13 Transgenic line validation.

Genomic DNA of a transgenic fly was purified using a DNeasy Blood and Tissue kit (Qiagen), and 25 ng of DNA was used as template in PCR reaction. Transgenic fly was confirmed by amplifying the region between the LE (left end) and the marker and also through amplification of fragment between marker and the RE (right end) for both *white* gene containing transgenic line and the one with DsRed marker.

pMuta1whiteTFD transgenics PCR was performed using tag polymerase and primers used were Muta1LE For: 5′-GATGGTACCGACGCAGTGGGTCTACCC-3′ and *White*

gene marker Rev: 5'-CGCTGAGTGGAATGTC-3'. *Muta1* RE Rev: 5′-GATGAGCTCCTACTGCGTCGGGTCTACC-3′ and *White* gene For: 5'-GAATACAAGTATTTCCCCTTCGAAC-3' with the following PCR program: 30 x (94° 2', 94° 30", 61° 30", 72° 7′), 4°. pMuta1DsRedTDF transgenic PCR was performed using phusion polymerase and primers used were Muta1LE For: 5′-GATGGTACCGACGCAGTGGGTCTACCC-3′ and DsRed marker Rev : 5'-CCTTGGTCACCTTCAGCTTC-3'. *Muta1* RE Rev: 5′-GATGAGCTCCTACTGCGTCGGGTCTACC-3′ and DsRed For: 5'-TGATGCAGAAGAAGACCATGG-3' with the following PCR program: 98° 30 sec, 30× (98° 08″, 58° 30″, 72° 30″), 72° 7′, 4°.

4.3.14 Molecular characterization of transposition events through inverse PCR

To verify the genomic location of the transposon in transgenic flies inverse PCR technique was used. The method uses the polymerase chain reaction (PCR), it has the primers oriented in the reverse direction of the usual orientation. The template for the PCR reaction was a restriction fragment that has been self ligated to form a circle. DNA was extracted using Promega kit and was digested with Taq1 enzyme at 65°C (Inverse PCR and sequencing protocol on 5 fly preps. Exelxis, Inc.). Digested DNA was heat killed at 80°C for 10 minutes and is then ligated using T4 DNA ligase (Thermo Fisher). 1 ul of ligated reaction was used as a template for first round of nested PCR using *Taq* DNA polymerase, with the following PCR program: 94° for 3 min, 30 × (94° for 20 sec, 60° for 20 sec 72° for 1 min), 72° for 5 min, 4°. Second round of inverse PCR was

performed using 1 ul template from first round with the following PCR program: 94° for 3 min, 30 × (94° for 20 sec, 60° for 20 sec 72° for 1 min), 72° for 5 min, 4°.

Primers for first round for left and right end:

Rev1 for LE: 5'-CATAACAGTGTGAGAAGCGTACG-3'
For1 for LE: 5'-TGTCACGATCCATTAGTCACTGT-3'
REV1 for RE: 5'-GTATACATAGAAATGTAATGAAAAACCTCTCTC-3'
For1 for RE: 5'-TGACTCATGTGAACAACGGTAAC-3'

Primers for second round for left and right end:

Rev2 for LE: 5'-GTACGACTAGATAAAGATGTTCATCATGT-3'
For2 for LE: 5'-GCCTTATACTAGTTTATTTGATATTTGTACTACG-3'
Rev2 for RE: 5'-TCCTAAGTGAGCTGCAATTGC-3'
For2 for RE: 5-GTTTTAAAAATACGATTTCTGGTTATGG-3'

4.3.14 Molecular characterization of transposition events through genome walking.

For each library that was made, DNA was digested with blunt cutters *SnaB1* (Thermo Fisher) and *EcoRV* (Thermo Fisher). 250 ng/ul of DNA was digested in fast digest buffer for 3 hours at 37°C. Digested DNA was then column purified (Qiagen kit), DNA was eluted in 30 ul of 1mM Tris at pH8.5. DNA should be at 250 ng, at concentration of 8.33 ng/ul. To total amount of eluted DNA, add 1 ul of freshly annealed adapters at concentration of 25 uM and ligate reaction with T4 DNA ligase (Thermo Fisher). Ligation was carried out for 3 hours and 1ul of ligation reaction was used as template for first round of genome walking. Adapter annealing was carried out for 3 minutes at 95°C in NEB buffer 2 and was allowed to cool down for 45 minutes on the heat block at RT using adapter Top

5’GTAATACGACTCACTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGT-3’ and adapter bottom 5’-ACTATAGGGCACGCGTGGT-3.

The first round of genome walking was performed using Q5-hot DNA polymerase and primers used were GWAP1: 5′-GTAATACGACTCACTATAGGGC-3′ and *Muta1* RE GW1: 5’-CTTGAGGCAATTGCAGCTCACTTAGG-3’ with the following PCR program: 98° for 30 sec, 6 × (98° for 25 sec, 72° for 3 min), 31 × (98° for 25 sec, 67° for 20 sec, 72° for 90 sec), 72° for 5 min, 4°. 1ul template from first round was used for second round of genome walking with Q5-hot DNA polymerase and primers used were GWAP2: 5′-ACTATAGGGCACGCGTGGT-3′ and *Muta1* RE GW2: 5’-GGTTATGGTTATGCCAAACGACTATTATGCCAAATG-3’ with the following PCR program: 98° for 30 sec, 4 × (98° for 25 sec, 72° for 3 min), 19 × (98° for 25 sec, 67° for 20 sec, 72° for 3 min), 72° for 5 min, 4°.

4.4 Results

4.4.1 *Muta1* is active in germline of *D. melanogaster*

Germ-line transformation of *D. melanogaster* was achieved using pMuta1DsRedTFD as the reporter plasmid and the *Muta1* helper plasmid. Transformation was also obtained using pMuta1whiteTFD as the reporter plasmid, which has the *white* gene as the reporter.

A total of 45 embryos injected with pMuta1DsRedTFD survived to adulthood (a survival rate of approximately 18%) and of those, 27 individuals (60% of G_0 crosses) generated progeny upon backcrossing (Table 4.1). Out of 27 fertile crosses, 4 crosses produced transgenic offspring, which resulted in transformation efficiency of 14.8% for the *Muta1* element with DsRed marker. Injections with p*Muta1*whiteTFD along with wildtype *Muta1* transposase cloned into p*Muta1* helper resulted in 21 embryos surviving to adulthood (a survival rate of approximately 15%) and 18 of those were fertile (86% of the G_0 crosses). Out of 18 fertile crosses, only one cross-produced transgenic offspring upon backcrossing, which resulted in a transformation efficiency of 5.6% with the *white* gene as the reporter (Table 4.1).

Microinjection with pMuta13EHspBac and pBac3EHspMuta1 plasmids resulted in 104 embryos that survived to adulthood (a survival rate of approximately 35%) and of those, 41 individuals (85% of G_0 crosses) generated progeny upon backcrossing (Table 4.1). Out of 41 fertile crosses, 2 crosses produce transgenic offspring for *Muta1* element (Table 4.1). Screening G_1 progeny showed that 4.9% of crosses produced transgenic offspring for *Muta1*, having DsRed expression, and 14.8% of the crosses produced transgenic offspring for *piggyBac*, GFP expression.

PCR analysis confirmed presence of the DsRed gene and the *white* gene was due to integration of the respective *Muta1* element in *D. melanogaster* (Figure 4.6). Molecular characterizations of the actual integrations in the transgenic lines were carried out using

inverse PCR. One line with DsRed expression had integration into the exon of gene CG30643 on chromosome 2R with 9 bp TSDs and another line had integration into the exon of gene CG10663 on chromosome 3L with 9 bp TSDs (Table 4.3). Two lines showed integration along with flanking plasmid DNA. The transgenic line obtained with *white* gene showed integration into chromosome 2R into a repeat region and had 9 bp TSDs (Table 4.3).

For the remobilization experiment, I used transgenic lines M32, P32 and P42, which were generated with pMuta13EHspBac and pBac3EHspMuta1 plasmids. Molecular characterization of the integration in these parental lines was carried out using genome-walking protocol. Because all the primers in inverse PCR are required to be in the ends of *Muta1* element, which are filled with, direct repeats and that makes characterization of integration very difficult. The M32 line, which is a *Muta1* transgenic line, had three integrations: one on chromosome 2L in an intron of the *Sickie* gene forming a 9 bp TSD, a second on chromosome X in an intron of the *Regucalcin* gene with 8 bp TSDs and a third with the flanking plasmid DNA whose exact integration location was not determined. Transgenic line P32, which is a *piggyBac* transgenic line, had integration into an intron of the Acetylcholine receptor gene on chromosome 2L and formed the expected TTAA TSDs upon integration. Line P42 contained integration into an exon of the *Arginase* gene with TTAA TSDs (Table 4.3).

4.4.2 Remobilization activity of *Muta1* in *D. melanogaster.*

Four crosses were established. In cross RA2, 20 M32 males were crossed with 20 P32 females. In cross RA8, 20 M32 males were crossed with 20 P42 females. The reciprocal crosses RC2 and RC8 were established for RA2 and RA8, respectively. The parental line M32 had the *Muta1* element with DsRed expression in eyes; P32 and P42 lines had the *piggyBac* element with GFP expression in eyes.

Five new phenotypes were observed in the offspring for these crosses. DsRed expression in thorax, full body, abdomen, labellum and genitalia. Molecular characterization of remobilization in these flies was performed using genome walking and splinkeret PCR (Table 4.4). The fly from cross RC8 G23 with DsRed expression in full body (Figure 4.3.A) contained a new integration into an intergenic region between gene CG43248 and CG7450 on chromosome 3L, forming a 9 bp TSD. The fly from a cross RA2 G6 with DsRed expression in the dorsal side of abdomen (Figure 4.3.B) contained integration into the *piggyBac* transposase polyA in the same plasmid. The fly from a cross RA2 G22 with DsRed expression in the thorax (Figure 4.3.C) contained two new integrations, one in the intron of gene CG32269/Mrtf, on chromosome 3L, forming 9 bp TSD and the second in an intron of CG11546/Kermit gene on chromosome 2R. The fly from a cross RA8 G18 with expression of DsRed in the abdomen and genitalia (Figure 4.3. D), contained a new integration into an intergenic region between genes CG4717/Knirps and CG44684 on chromosome 3L, forming a 9 bp TSD. The fly from a cross RA8 G14 with

expression of DsRed in the labellum and genitalia (Figure 4.3.E) contained three new integrations, two of which were novel and one was parental integration on X chromosome in intron of CG1803/Regucalcin gene. The two novel insertions were on chromosome 3R in a repeat region, forming a 9 bp TSD and on chromosome 3L into an intron of the CG34418/sif gene, forming a 9bp TSD, respectively.

Flies in which new phenotypes were not observed were look for new integration locations for *Muta1* element, as the *Muta1* element might have remobilized in these flies but does not produced new phenotypes (Table 4.5). The new integrations characterized were (1) a fly from the cross RA2 G6 in an exon of CG3552 on chromosome 3L forming a 9 bp TSD; (2) a fly from the cross RA2 G6 in an intron of CG17664 on chromosome 2R forming a 9 bp TSD; (3) a fly from the cross RA2 G7 in a repeat region, which was found in chromosome 3R, 2R, X, and chromosome Y, forming a 9 bp TSD; (4) a fly from the cross RA2 G10 in an untranslated exon of CG3394/Mir gene on chromosome 2R forming an unusual 5 bp TSD; (5) a fly from the cross RA2 G23 in an exon of CG0721/Arrowhead gene on chromosome 3L forming a 9 bp TSD; (6) a fly from the cross RC8 G15 in an exon of CG16784/Purple gene on chromosome 2L forming a 9 bp TSD; (7) a fly from the cross RA2 G7 in an exon of CG9660/Toucan gene on chromosome 2L forming a 9 bp TSD; (8) a fly from the cross RA8 G6 in an intron of CG7337, forming a 9 bp TSD, (9) a fly from the cross RC2 G4 in an exon of CG1803/Regucalcin gene on X chromosome forming a 9 bp TSD, (10) integration into *piggyBac* polyA was characterized in flies from crosses RA2 G5, RA2 G6, RA2 G7, RA2

G10, RA8 G6, RA8 G7, and RC8 G6, (11) integration with the flanking plasmid DNA were observed in flies from crosses RA2 G6, RA G7, RA2 G12, RA2 G21, RA8 G6, RA8 G7, and RA8 G23 (Table 4.4).

4.4.3 *Muta1* is active in germ-line of *Ae. aegypti.*

Germline transformation of *Ae. aegypti* was performed using the reporter plasmid pMuta13HSPpBac and the helper plasmid pBac3HSPpMuta1 at Insect Transformation facility, University of Maryland. 75 adults were used in a pool mating from injected embryos , three transgenic progeny were obtained for *Muta1*, and 13 transgenic progeny were obtained for *piggyBac* (Figure 4.4). Transformation efficiency of 4% was observed for *Muta1* element and 17.3% for *piggyBac* element (Table 4.6). The *Muta1* transgenic lines were confirmed through PCR for the integration of the *Muta1* element with DsRed marker (Figure 4.7).

Molecular characterization of the three *Muta1* lines revealed integration of *Muta1* in Supercontig 1.21 for line M3 with a 9 bp TSD, line M1 and M2 showed integration along with the flanking plasmid DNA (Table 4.7).

Four crosses were established. In cross 1, 10 males from *Muta1* line 1 were crossed with 10 females from *piggyBac* line 11. In cross 2, 10 males from *Muta1* line 1 were crossed with 10 females from *piggyBac* line 14. In cross 3, 10 females from *Muta1* line 2 were crossed with *piggyBac* line 9. In cross 4, 10 females from *Muta1* line 2 were cross with

females from *piggyBac* line 14. New phenotypes were not observed in the larvae or pupae stages. In the adult stage, two potentially new phenotypes were observed after screening 1,454 adults; one female with expression in the legs and one female with expression in the abdomen (Figure 4.5).

As the DNA extraction produced a low yield from these mosquitoes, the DNA sample was first amplified with genomic amplification protocol (Qiagen Repli-G minikit) before molecular characterization of the region by genome-walking and inverse PCR protocols. The molecular characterization of the mosquito with DsRed expression in legs contained a new integration into an intergenic region on supercontig 1.70, the insertion formed a 9 bp TSD. The parent of the mosquito with DsRed expression in legs had an integration of the *Muta1* element with flanking plasmid DNA (Figure4.12). This suggests that the *Muta1* element has excised precisely and integrated into a new location. This integration location has not been characterized previously in the parental lines. Thus, confirms that it is a *Muta1* remobilization event. The mosquito with DsRed expression in midgut contained integration with the flanking plasmid DNA. This pattern was similar to the parental integration and further molecular verification needs to be performed to determine if this is a new transposition event.

The mosquito with DsRed expression in the legs died before it could be mated, so I was unable to determine whether it was a germline event or arose through somatic transposition of the *Muta1* element during development. The mosquito with DsRed

expression in the midgut was backcrossed to wild-type Orlando and progeny from this cross did not show DsRed expression in the midgut, leading to the possibility that this phenotype may have reason though of somatic remobilization.

4.5 Discussion

Muta1 is active in the germline of *D. melanogaster* and *Ae. aegytpi*. The transformation rate obtained for the non-autonomous *Muta1* element containing the white gene marker was low compared to the same transposon containing the DsRed marker. The white gene is relatively larger in size, 3,681 bp in length compared to the size of the DsRed2-1 gene, which is 657 bp, which may be the explanation. If so it may indicate that cargo capacity is a factor in the efficiency of the *Muta1* activity on a vector. This possibility would need to be more fully explored with further experiments using different size inserts. Studies have shown that the transformation efficiency is also affected by the cargo capacity, which is amount of the internal DNA inserted, position of this DNA in the vector, and the amount of sub-terminal DNA remaining in the vector (Balciunas et al., 2006).

A remobilization rate of 21.3% was observed in *D. melanogaster* with characterization of 17 new integration locations for the *Muta1* element from molecular analysis of 80 flies. Molecular characterization of flies with new integrations of *Muta1* element revealed that these flies retain parental integration of *Muta1* element in X chromosome with multiple new integrations elsewhere. The new transposition events that were recovered from molecular characterization showed integration into exons in 41.2% of events, into intron

in 29.4% of the cases, while 23.5% were intergenic or into repeat regions (Figure 4.9). The *Muta1* element was found to have integrated into the polyA of *piggyBac* transposase in 50% of the characterized flies; the *piggyBac* transposase was encoded in between the ends of *Muta1* element. The integration into *polyA* was obtained from characterization of the right end, as left the end amplification has been difficult due to direct repeats in the end. Moreover, there is a possibility that these flies with integration into polyA region might have integrated into a new location, which I was unable to characterize because of the bias towards amplification of the right end integration into polyA (Figure 4.10). More detail molecular verification needs to be performed to determine exact mechanism of *Muta1* transposition. These events might be result of the *Muta1* local hopping, where a transposon prefers to integrate into sequence immediately adjacent to initial integration Two integrations on chromosome X at location 12016142 and 12016483 were observed in an exon of CG1803/Regucalcin gene, which was in near vicinity of the previous parental integration into an intron of CG1803/Regucalcin gene (Figure 4.11).

The local hopping of *Muta1* might be a result of the mechanism of *Muta1* transposition, in which the transposase requires internal terminal sequences for excision and integration. Moreover, there is a possibility that the *Muta1* element might have a preference for polyA target sites for integration, which is easily available in the plasmid sequence of the vector that contains *piggyBac* transposase polyA flanked by the TIRs of the *Muta1* element. This pattern of integration has not been observed in other *Mutator* transposons. The fly characterized with unusual 5 bp TSDs, might be due to alternative mechanism of

the *Muta1* element insertion or it could be an artifact. Further analysis needs to be done to understand the mechanism of transposition of the *Muta1* transposase.

The phenomenon of hybrid dysgenesis in relation to the introduction of *Muta1* element in *D. melanogaster*, which is naïve to *Muta1* element, was determined in the crosses RA2, and RA8; and their reciprocal crosses RC2 and RC8. In the crosses RA2 and RA8, males carrying *Muta1* element were crossed with the females from *piggyBac* lines that do not contain *Muta1* element. In the reciprocal crosses RC2 and RC8, females carrying *Muta1* element were crossed with the males from *piggyBac* lines. No sterility or mortality was observed in these crosses, and the screened progeny were determined to be healthy and fertile Thus, the *Muta1* element does not cause hybrid dysgenesis which has been observed with the *P*-element (Bingham et al., 1982) and the *Penelope* element (Evgen'ev et al., 1997) when introduced into *D. melanogaster* genome, that was originally naïve to these elements.

The transformation efficiency observed for the *Muta1* element in *Ae. aegypti* was low as compared to *piggyBac* element, which was used as an internal control *(*Table 4.6). The decrease in transposition efficiency of *Muta1* element in germline compared to the somatic cells in *Ae. aegypti* could be due to requirement of internal DNA sequence of the *Muta1* element or the germline host factor that might be influencing transformation frequency. In a study of *piggyBac* element with minimal sequence cartridge, which is capable of efficient embryo interplasmid transposition assays, failed to produced

transformants at significant frequency in *D. melanogaster* compared to full length or less extensive internal deletion constructs (Li et al., 2005). This study also demonstrates that the internal DNA sequence adjacent to 5' and 3' terminal repeat domains are crucial for germline transformation of *piggyBac* but not for excision and somatic transposition in *D. melanogaster*. Therefore, analysis of internal DNA sequence and their influence on transformation efficiency of the *Muta1* element can be very useful.

The *Muta1* element was observed to have a precise integration forming a 9 bp TSDs in one of the transgenic line. Two other lines showed integration with flanking plasmid DNA. The RT-PCR experiments performed on third instar larvae showed expression of *Muta1* transposase in absence of heat shock, which indicates that the Hsp70 promoter is leaky in the transgenic lines (Figure 4.8). Previously it has been reported that the heat shock promoter Hsp70 from *D. melanogaster* was able to induce transcription without heat shock (D. A. O'Brochta, Pilitt, Harrell, Aluvihare, & Alford, 2012). The Orlando larvae used as a control in the RT-PCR experiment showed no expression of the *Muta1* transposase. Thus, the new potential phenotypes observed in the transgenic lines were most likely a result of *Muta1* transposase expressed in the helper line and not due to the eight endogenous wild-type copies of the *Muta1* element in *Ae. aegypti.* Further, it is possible that endogenous Muta1 transposase is expressed at another life stage or tissue. Comprehensive time course tissue study might help in determining endogenous *Muta1* transposase expression in *Ae. aegypti.*

The new potential phenotypes were observed in two individuals from screening 1,454 mosquitoes, that is, 0.13% of the progeny screened. The mosquito with the expression in the midgut was backcrossed with wild-type *Ae. aegypti* (Orlando strain), but the *Muta1* integration was not inherited in the offspring from this cross. This could be due to a somatic transposition event of the *Muta1* element early on during development of the progeny. Similar somatic transposition events have been reported with the *piggyBac* element in *An. Stephensi* (D. A. O'Brochta et al., 2012). Whereas in *Ae. aegypti, piggyBac* elements did not show germline or somatic activity post-integration (D. A. O'Brochta, 2003) (David A O'Brochta, Alford, Pilitt, Aluvihare, & Harrell, 2011). My work is the first example of somatic transposition events in *Ae. aegypti*. Further studies are needed to determine the post integration mobility of the *Muta1* in *Ae. aegypti*.

The ability of the *Muta1* transposon to efficiently transform *D. melanogaster* and *Ae. aegypti*, will make it an important genetic tool that can be used for wide variety of applications. Despite the few progeny analyzed, my experiments found that the *Muta1* element was able to remobilize in *D. melanogaster*. Molecular characterization of new phenotypes with DsRed expression in legs showed that the *Muta1* element has integrated into a new location in *Ae. aegypti*. This integration location has not been observed in the parental lines, which confirms that it is a new jump of the *Muta1* element. It integrated into an intergenic region between protein coding genes AAEL002860 and AAEL002864. The analysis of the proteins of these two genes revealed that AAEL002864 has a ZINC finger domain and four FLYWCH domains, which are DNA binding domains. The gene

AAEL002860 has MPC domain whose function is unknown. There is a possibility that these two genes might have a promoter that lead to the expression of DsRed marker in the legs. To date, the transposons used for *Ae. aegypti* transformation have been exogenous to this insect. My research demonstrated the first endogenous *Mutator* element to be active in germline of *Ae. aegypti* and *D. melanogaster*, with indication that it may also be somatically active. Thus, the *Muta1* element might serve as an efficient genetic tool for wide variety of insects and in particular for *Ae. aegypti*.

4.6 Reference

Balciunas, D., Wangensteen, K. J., Wilber, A., Bell, J., Geurts, A., Sivasubbu, S., … Ekker, S. C. (2006). Harnessing a high cargo-capacity transposon for genetic applications in vertebrates. *PLoS Genetics*, *2*(11), e169. http://doi.org/10.1371/journal.pgen.0020169

Berg, C. A., & Spradling, A. C. (1991). Studies on the rate and site-specificity of P element transposition. *Genetics*, *127*(3), 515–524.

Bingham, P. M., Kidwell, M. G., & Rubin, G. M. (1982). The molecular basis of P-M hybrid dysgenesis: The role of the P element, a P-strain-specific transposon family. *Cell*, *29*(3), 995–1004. http://doi.org/10.1016/0092-8674(82)90463-9

Bryan, G., Jacobson, J., & Hartl, D. (1987). Heritable somatic excision of a Drosophila transposon. *Science*, *235*(4796), 1636–1638. http://doi.org/10.1126/science.3029874

Catteruccia, F., Nolan, T., Loukeris, T. G., Blass, C., Savakis, C., Kafatos, F. C., & Crisanti, A. (2000). Stable germline transformation of the malaria mosquito Anopheles stephensi. *Nature*, *405*(6789), 959–62. http://doi.org/10.1038/35016096

Chalvet, F., Grimaldi, C., Kaper, F., Langin, T., & Daboussi, M.-J. (2003). Hop, an active Mutator-like element in the genome of the fungus Fusarium oxysporum. *Molecular Biology and Evolution*, *20*(8), 1362–75. http://doi.org/10.1093/molbev/msg155

Coates, C. J., Jasinskiene, N., Miyashiro, L., & James, A. A. (1998). Mariner transposition and transformation of the yellow fever mosquito, Aedes aegypti. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(7), 3748–51.

Eisen, J. A., Benito, M.-I., & Walbot, V. (1994). Sequence similarity of putative transposases links the maize Mutator autonomous element and a group of bacterial insertion sequences. *Nucleic Acids Research*, *22*(13), 2634–2636. http://doi.org/10.1093/nar/22.13.2634

Evertts, A. G., Plymire, C., Craig, N. L., & Levin, H. L. (2007). The hermes transposon of Musca domestica is an efficient tool for the mutagenesis of Schizosaccharomyces pombe. *Genetics*, *177*(4), 2519–23. http://doi.org/10.1534/genetics.107.081075

Evgen'ev, M. B., Zelentsova, H., Shostak, N., Kozitsina, M., Barskyi, V., Lankenau, D.-H., & Corces, V. G. (1997). Penelope, a new family of transposable elements and its possible role in hybrid dysgenesis in Drosophila virilis. *Proceedings of the National Academy of Sciences*, *94*(1), 196–201. http://doi.org/10.1073/pnas.94.1.196

Franz, G., & Savakis, C. (1991). Minos, a new transposable element from Drosophila hydei, is a member of the Tc1-like family of transposons. *Nucleic Acids Research*, *19*(23), 6646.

Fraser, M. J., Ciszczon, T., Elick, T., & Bauser, C. (1996). Precise excision of TTAA-specific lepidopteran transposons piggyBac (IFP2) and tagalong (TFP3) from the baculovirus genome in cell lines from two species of Lepidoptera. *Insect Molecular Biology*, *5*(2), 141–51.

Guimond, N., Bideshi, D. K., Pinkerton, A. C., Atkinson, P. W., & O'Brochta, D. A. (2003). Patterns of Hermes transposition in Drosophila melanogaster. *Molecular Genetics and Genomics : MGG*, *268*(6), 779–90. http://doi.org/10.1007/s00438-002-0800-4

Handler, A. M., & Harrell, R. A. (1999). Germline transformation of Drosophila melanogaster with the piggyBac transposon vector. *Insect Molecular Biology*, *8*(4), 449–57.

Horn, C., & Wimmer, E. A. (2000). A versatile vector set for animal transgenesis. *Development Genes and Evolution*, *210*(12), 630–7.

Hua-Van, A., & Capy, P. (2008). Analysis of the DDE motif in the Mutator superfamily. *Journal of Molecular Evolution*, *67*(6), 670–81. http://doi.org/10.1007/s00239-008-9178-1

Jasinskiene, N., Coates, C. J., Benedict, M. Q., Cornel, A. J., Rafferty, C. S., James, A. A., & Collins, F. H. (1998). Stable transformation of the yellow fever mosquito, Aedes aegypti, with the Hermes element from the housefly. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(7), 3743–7.

Li, X., Harrell, R. A., Handler, A. M., Beam, T., Hennessy, K., & Fraser, M. J. (2005). piggyBac internal sequences are necessary for efficient transformation of target genomes. *Insect Molecular Biology*, *14*(1), 17–30. http://doi.org/10.1111/j.1365-2583.2004.00525.x

Lidholm, D. A., Lohe, A. R., & Hartl, D. L. (1993). The transposable element mariner mediates germline transformation in Drosophila melanogaster. *Genetics*, *134*(3), 859–68.

Lisch, D. (2002). Mutator transposons. *Trends in Plant Science*, *7*(11), 498–504. http://doi.org/10.1016/S1360-1385(02)02347-6

Lisch, D., Chomet, P., & Freeling, M. (1995). Genetic characterization of the Mutator system in maize: behavior and regulation of Mu transposons in a minimal line. *Genetics*, *139*(4), 1777–96.

Lohe, A. R., & Hartl, D. L. (2002). Efficient Mobilization of mariner in Vivo Requires Multiple Internal Sequences. *Genetics*, *160*(2), 519–526.

Lohe, A. R., Lidholm, D. A., & Hartl, D. L. (1995). Genotypic Effects, Maternal Effects and Grand-Maternal Effects of Immobilized Derivatives of the Transposable Element mariner. *Genetics*, *140*(1), 183–192.

Lozovsky, E. R., Nurminsky, D., Wimmer, E. A., & Hartl, D. L. (2002). Unexpected stability of mariner transgenes in Drosophila. *Genetics*, *160*(2), 527–35.

Marquez, C. P., & Pritham, E. J. (2010). Phantom, a new subclass of Mutator DNA transposons found in insect viruses and widely distributed in animals. *Genetics*, *185*(4), 1507–17. http://doi.org/10.1534/genetics.110.116673

McCarty, D. R., Mark Settles, A., Suzuki, M., Tan, B. C., Latshaw, S., Porch, T., … Curtis Hannah, L. (2005). Steady-state transposon mutagenesis in inbred maize. *The Plant Journal*, *44*(1), 52–61. http://doi.org/10.1111/j.1365-313X.2005.02509.x

O'Brochta, D. A. (2003). Gene vector and transposable element behavior in mosquitoes. *Journal of Experimental Biology*, *206*(21), 3823–3834. http://doi.org/10.1242/jeb.00638

O'Brochta, D. A., Alford, R. T., Pilitt, K. L., Aluvihare, C. U., & Harrell, R. A. (2011). piggyBac transposon remobilization and enhancer detection in Anopheles mosquitoes. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(39), 16339–44. http://doi.org/10.1073/pnas.1110628108

O'brochta, D. A., & Handler, A. M. (1988). Mobility of P elements in drosophilids and nondrosophilids. *Proceedings of the National Academy of Sciences of the United States of America*, *85*(16), 6052–6.

O'Brochta, D. A., Pilitt, K. L., Harrell, R. A., Aluvihare, C., & Alford, R. T. (2012). Gal4-based Enhancer-Trapping in the Malaria Mosquito Anopheles stephensi. *G3: Genes|Genomes|Genetics*, *2*(11), 1305–1315. http://doi.org/10.1534/g3.112.003582

O'Brochta, D. A., Sethuraman, N., Wilson, R., Hice, R. H., Pinkerton, A. C., Levesque, C. S., … Atkinson, P. W. (2003). Gene vector and transposable element behavior in mosquitoes. *The Journal of Experimental Biology*, *206*(Pt 21), 3823–34.

O'Brochta, D. A., Warren, W. D., Saville, K. J., & Atkinson, P. W. (1996). Hermes, a Functional Non-Drosophilid Insect Gene Vector From Musca domestica. *Genetics*, *142*(3), 907–914.

Palavesam, A., Esnault, C., & O'Brochta, D. A. (2013). Post-integration silencing of piggyBac transposable elements in Aedes aegypti. *PloS One*, *8*(7), e68454. http://doi.org/10.1371/journal.pone.0068454

Pritham, E. J., Feschotte, C., & Wessler, S. R. (2005). Unexpected diversity and differential success of DNA transposons in four species of entamoeba protozoans. *Molecular Biology and Evolution*, *22*(9), 1751–63. http://doi.org/10.1093/molbev/msi169

Raizada, M. N., Benito, M. I., & Walbot, V. (2001). The MuDR transposon terminal inverted repeat contains a complex plant promoter directing distinct somatic and germinal programs. *The Plant Journal : For Cell and Molecular Biology*, *25*(1), 79–91.

Raizada, M. N., & Walbot, V. (2000). The Late Developmental Pattern of Mu Transposon Excision Is Conferred by a Cauliflower Mosaic Virus 35S -Driven MURA cDNA in Transgenic Maize. *PLANT CELL*, *12*(1), 5–22.

Rubin, G. M., & Spradling, A. C. (1982). Genetic transformation of Drosophila with transposable element vectors. *Science (New York, N.Y.)*, *218*(4570), 348–53.

Smith, R. C., Walter, M. F., Hice, R. H., O'Brochta, D. A., & Atkinson, P. W. (2007). Testis-specific expression of the beta2 tubulin promoter of Aedes aegypti and its application as a genetic sex-separation marker. *Insect Molecular Biology*, *16*(1), 61–71. http://doi.org/10.1111/j.1365-2583.2006.00701.x

Spradling, A. C., Stern, D., Beaton, A., Rhem, E. J., Laverty, T., Mozden, N., … Rubin, G. M. (1999). The Berkeley Drosophila Genome Project gene disruption project: Single P-element insertions mutating 25% of vital Drosophila genes. *Genetics*, *153*(1), 135–77.

Tan, B.-C., Chen, Z., Shen, Y., Zhang, Y., Lai, J., & Sun, S. S. M. (2011). Identification of an active new mutator transposable element in maize. *G3 (Bethesda, Md.)*, *1*(4), 293–302. http://doi.org/10.1534/g3.111.000398

Wilson, R., Orsetti, J., Klocko, A. D., Aluvihare, C., Peckham, E., Atkinson, P. W., … O'Brochta, D. A. (2003). Post-integration behavior of a Mos1 mariner gene vector in Aedes aegypti. *Insect Biochemistry and Molecular Biology*, *33*(9), 853–63.

Xu, Z., Yan, X., Maurais, S., Fu, H., O'Brien, D. G., Mottinger, J., & Dooner, H. K. (2004). Jittery, a Mutator distant relative with a paradoxical mobile behavior: excision without reinsertion. *The Plant Cell*, *16*(5), 1105–14. http://doi.org/10.1105/tpc.019802

Table 4.1 Transformation rate from transformation and remobilization experiment for *Muta1* in *D. melanogaster*. The table describes number of embryos injected, survived embryos, number of crosses established, number of fertile crosses obtained, number of transgenic lines obtained from the fertile crosses for the *Muta1* element and the transformation rate. The remobilization experiment shows transgenic lines obtained for *piggybac* element and the transformation rate. The *piggyBac* transgenic lines served as helper lines in this experiment and were used as an internal control to compare transformation rate of the *Muta1* element to the *piggyBac* element.

Plasmid Injected	Total embryos injected	Embryos Recovered	Total crosses	G0 Fertile Crosses	*Muta1* Transgenic line	*pBac* Transgenic line	*Muta1* Transgenic Rate	*pBac* Transgenic Rate
pMuta1DsRedTFD+ pMuta1 helper	250	60	45	27	4	NA	**14.8%**	NA
pMuta1whiteTDF +pMuta1 helper	140	40	21	18	1	NA	**5.6%**	NA
pMuta13EHsppBac + pBac3EHsp Muta1	300	104	48	41	2	6	**4.9%**	14.8%

Table 4.2 Remobilization experiment results. Table shows number of flies screened in each cross and new types of phenotype observed in each cross. Total 5 new types of phenotypes were observed from crosses.

Lines	Generations Screened	Number of flies screened	Flies with Different Phenotype
RA2 (M32-males X P32-females)	23	4991	Full body, Thorax, second last segment on the dorsal side of abdomen
RA8 (M32-males X P42-females)	23	6091	Full body
RC2 (M32-females X P32-males)	9	1954	Thorax
RC8 (M32-females X P42-males)	15	3424	Full Body, Abdomen
Total		16,460	5 new types of phenotypes for *Muta1* element (0.03%)

Table 4.3 Molecular characterizations of *Muta1* and *piggyBac* parental lines in *D. melanogaster.*

Marker	Chr	Break Point	TSD	Gene
Muta1 white gene line	2R	3859229	GTTTTTGCG (9bp)	Repeat region
Muta1 DsRed line 1	2R	16676932	CTTAGCAGA (9bp)	Exon of CG30463-Glycosylation activity
Muta1 DsRed line 2	Flanking plasmid DNA	--	--	--
Muta1 DsRed line 3	3L	12481711	GAGAGGCTG (9bp)	Exon of CG10663-Serine type endopeptidase activity
Muta1 DsRed line 4	Flanking plasmid DNA	--	--	--
Muta1 line M4	2R	9104602	ACATGTTTG (8 bp)	Exon of CG8075/Van Gogh
Muta1 line M32	2L, X, Flanking plasmid DNA	19847516 12015935	AAAAAAAAT (9 bp) ATATTTAGG (9 bp) ---	Intron of CG43720/Sickie gene Intron of CG1803/Regucalcin gene --
piggyBac line P32	2L	14073174	TTAA	Intron of CG32975/Acetylcholine receptor
piggyBac line P42	X	514785	TTAA	Exon of CG18104/Arginase or CG4262/elav gene

Table 4.4 Molecular characterizations of new phenotypes in *D. melanogaster* for *Muta1* element integration. The table shows the phenotypic expression observed in the fly, break point of integration of the *Muta1* element, chromosome, TSD sequence, and the gene.

New phenotype flies	Break point	New integration locations	TSDs	Gene
Fly with full body expression (Figure A)	1553430	Chr 3L	TGTGTTGGA (9 bp)	Intergenic between CG43248 and CG7450
DsRed expression in dorsal 2nd last segment (Figure B)	-----	pBac polyA	TTTTTTTAT	------
Thorax (Figure C)	2741335 8151822	Chr 3L Chr 2R	GGTCAAACC GGTCAATTC	Chr 3L Intron of CG32296/ Mrtf RNA polymerase transcription coactivator, Chr 2R Intron CG11546/ Kermit - Regulation of signal transduction
DsRed expression in abdomen (Figure D)	20621644	Chr 3L	CTCACAGGG(9 bp)	Intergenic between Knirps gene and CG44684
Fly with full body, bright expression in labellum and genetalia (Figure E)	12015935 1276726 5702591	Chr X (Parental integration) Chr 3R,3L,2R,X & Y Chr 3L	GATTTATAG (9 bp) TTAAAATAT (9bp) TCCGACCG (8 bp)	Intron of Regucalcin gene, Chr 3R in repeat region, Chr 3L integration is in intron of CG34418/sif - guanyl receptor exchange factor. Involved in developmental stages in embryo, larvae, pupae and adults

Table 4.5 Molecular characterizations of new transposition events in *D. melanogaster* flies with no new phenotypes. The table shows chromosome, break point of integration of the *Muta1* element, TSD sequence, number of flies characterized for a particular integration location, and the gene.

New integration location	Break point	TSD	No. of flies characterized	Gene
pBac PolyA	---	TTTTTTTTT (9 bp)	40	---
Plasmid DNA	----	GACGCAGT (8 bp)	19	----
Chr 3L	9458689	CGATGATAA (9 bp)	1	Exon of CG3552-Phosphorylase activity
Chr 2R	23613641	ATGTATATA (9 bp)	1	Intron of CG17664-Aquaporin protein
Chr 2R,3R, X and Y Chr	------	GTGTTTAAC (9 bp)	3	Repeat region
Chr 2R	24294161	ATTTTT (5 bp)	3	Untranslated exon of CG3394/Mir-AMP binding enzyme
Chr 3L	3860980	GAGCTCGTG (9 bp)	1	Exon of CG072l/ Arrowhead- Zinc ion binding, DNA binding
Chr 2L	20075421	GTCTACGTT (9 bp)	1	Exon of CG16784/purple-phenotype of this allele manifest in pigment cell
Chr 2L	3071477	AATAAATCC (9 bp)	1	Exon of CG9660/ Toucan-Involved in Mitotic spindle formation
RE-pBac polyA--X Chr	12016142	TTTTTTTAT	1	Exon of CG1803/ Regucalcin gene
Chr 2L	1925256	CAAATGATT (9 bp)	1	Intron of CG7337-dehydrogenase
X Chr	12016483	GGACTGGCA (9 bp)	1	Exon of CG1803/ Regucalcin gene

Table 4.6 Transformation rate for *Muta1* and *piggyBac* lines in *Ae. aegypti.* The table shows number mosquito used in the transformation experiment, number of transgenic lines obtained with the *Muta1* element and the *piggyBac* element, and the transformation rate obtained for the *Muta1* element and the *piggyBac* element.

No. of mosquito adults in pool mating	*Muta1* transgenic lines	*piggyBac* transgenic lines	*Muta1* transformation rate	*piggyBac* transformation rate
75	3	13	4%	17.3%

Table 4.7 Molecular characterizations of *Muta1* parental lines in *Ae. aegypti.* The table shows the transgenic line, break point of integration of the *Muta1* element, supercontig location, TSD sequence, and the gene.

Aedes Muta1 transgenic lines	Break point	Location	TSD	Gene
Muta1 line 1	-	Flanking plasmid DNA	--	--
Muta1 line 2	-	Flanking plasmid DNA	--	--
Muta1 line 3 and 4	2698504	Supercontig 1.21	GATGCGCCT (9 bp)	AAGEO2001492.1

Table 4.8 Remobilization experiment results. Table shows number of mosquitoes screened in each cross for 10 generations and new types of phenotypes observed in each cross. Total two new types of phenotypes were observed from crosses.

Crosses	No. of screened mosquitoes	New phenotypes
M1-males X P11-females	846	DsRed expression in legs, Midgut
M1-males X P14-females	130	0
M2-females X P9-males	178	Midgut
M2-males x P11-females	300	0
Total	**1454**	**Two new phenotypes**

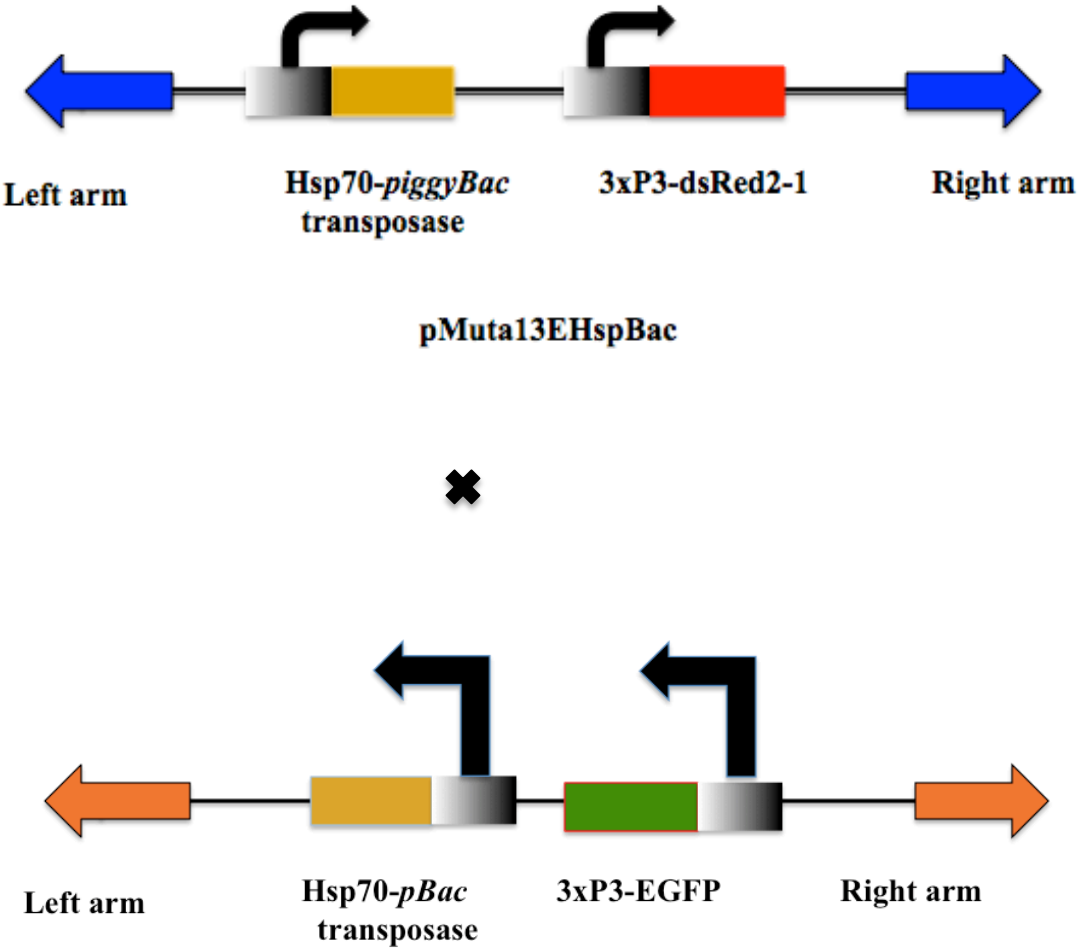


Figure 4.1 Schematic diagram of the reporter and helper line used to check *Muta1* mobility. The reporter line pMuta13EHspBac contains ends of *Muta1* element with *piggyBac* transposase, and the DsRed marker under the control of 3xP3 promoter. The helper line pBac3EHspBac contains the *piggyBac* element with *Muta1* transposase, and the EGFP marker under the control of 3xP3 promoter. The transposase is under the regulation of heat shock promoter 70. Helper/reporter heterozygotes were created by crossing helper and reporter lines. Helper/reporter jumpstarter lines were self-crossed among siblings and there progeny were heat shock to activate expression of respective transposase. DsRed progeny were then examined for new phenotypes and for molecular characterization of new transposition events

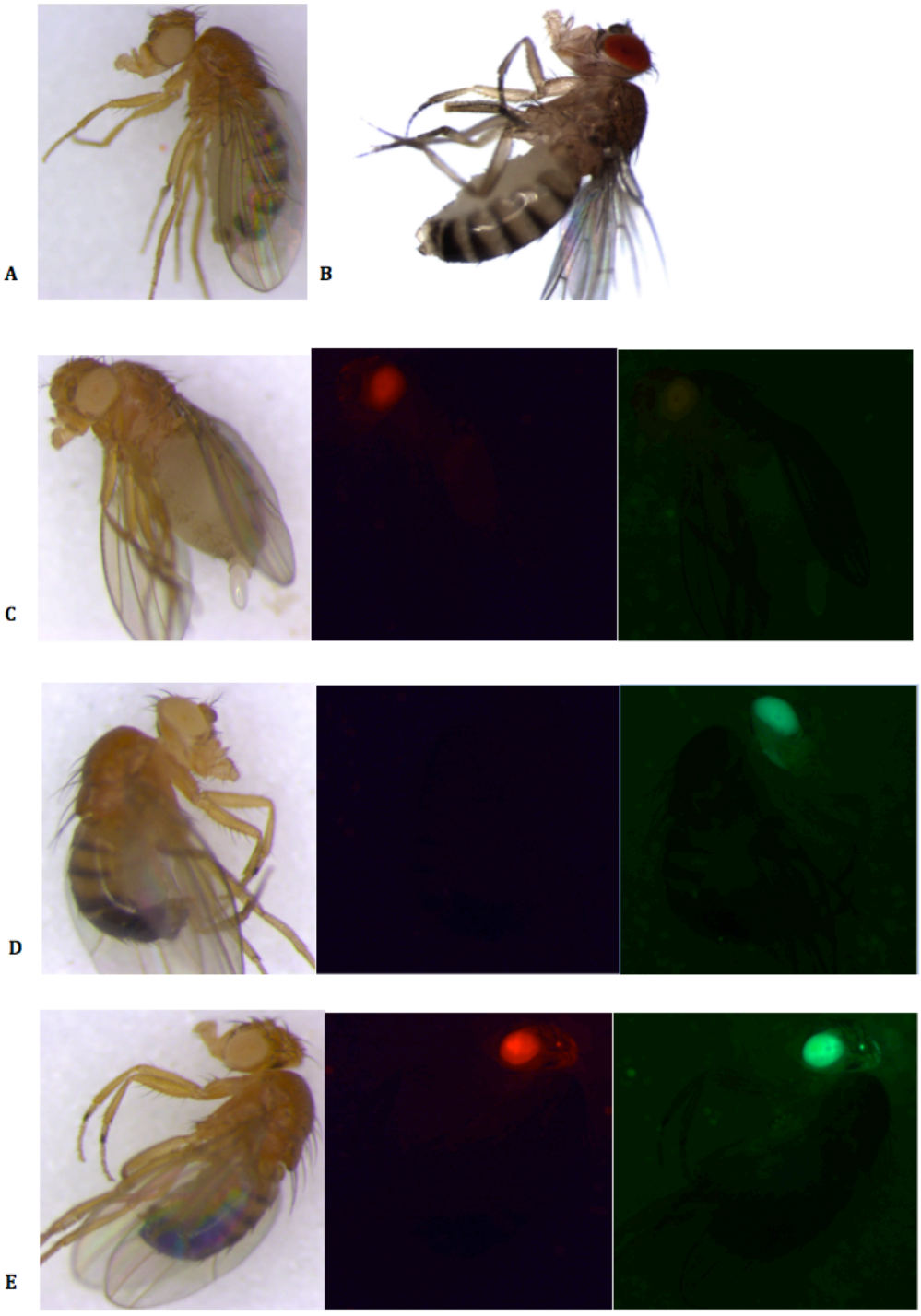
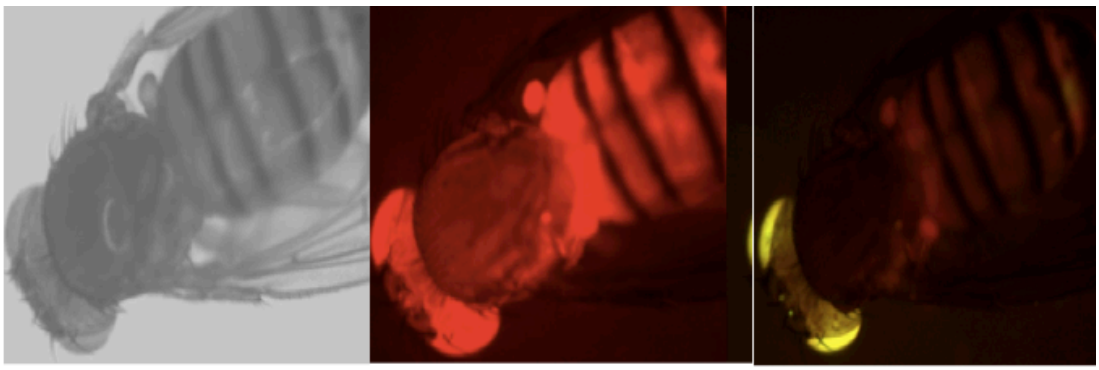
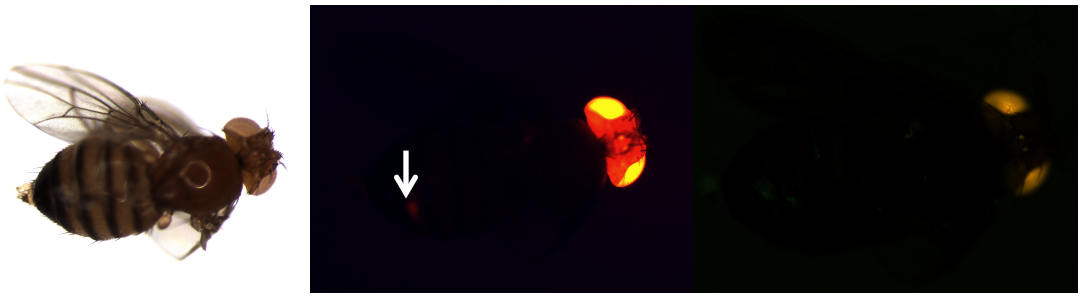


Figure 4.2 Transgenic flies with marker expression in eyes. A. CSw+ wildtype fly, B. *Muta1* transgenic fly with White gene marker, C. *Muta1* transgenic with DsRed marker, D. *piggyBac* transgenic fly with EGFP marker and E. Transgenic fly with both marker expressions

A



B



C

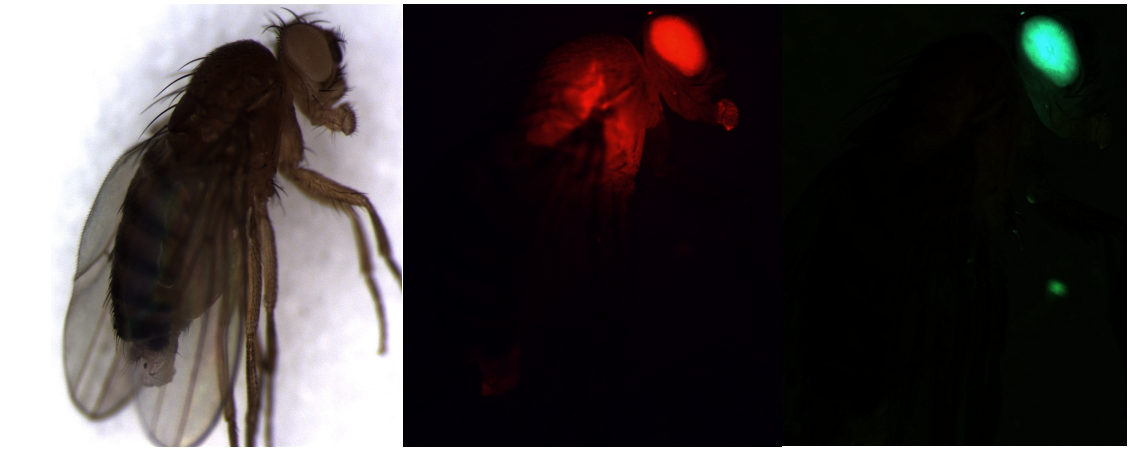
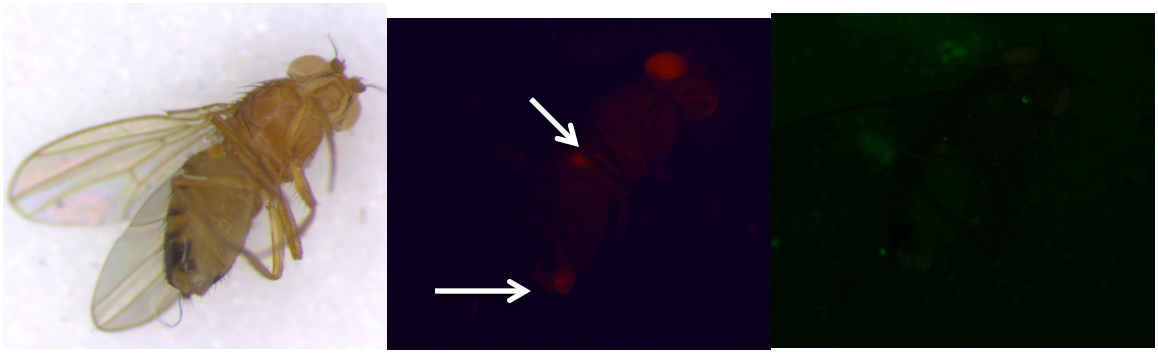


Figure 4.3 New phenotypes observed in crosses for *D. melanogaster* remobilization experiment for *Muta1* element, pictures are taken without filter: DsRed filter and EGFP filter respectively. (A) Adult with new phenotypes expression in full, (B) Dorsal view of a fly with expression in eyes and in 2nd last segment of abdomen marked with an arrow, (C) DsRed expression in thorax.

D



E

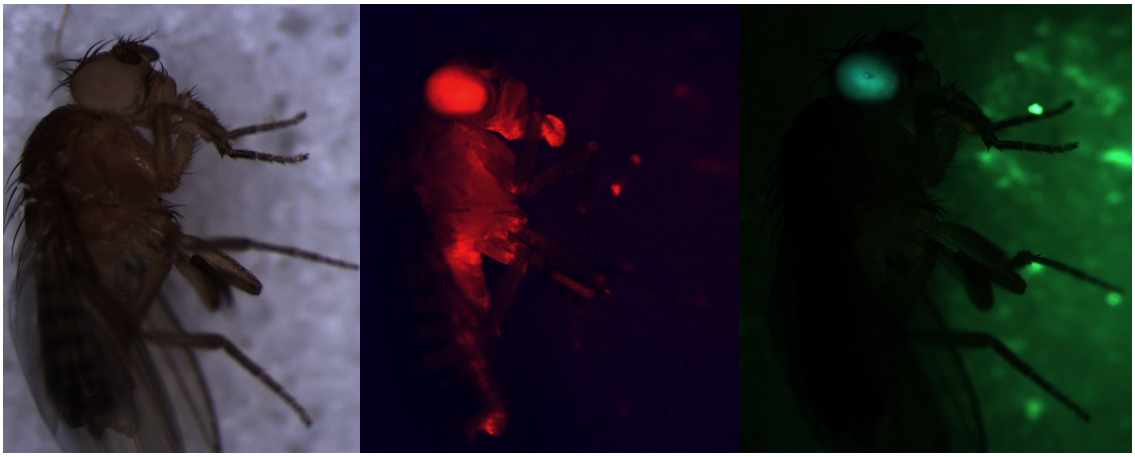


Figure 4.3 New phenotypes observed in crosses for *D. melanogaster* remobilization experiment for *Muta1* element, pictures are taken without filter: DsRed filter and EGFP filter respectively, (D) DsRed expression in abdomen and genitalia (E) DsRed expression in patches all over the body.

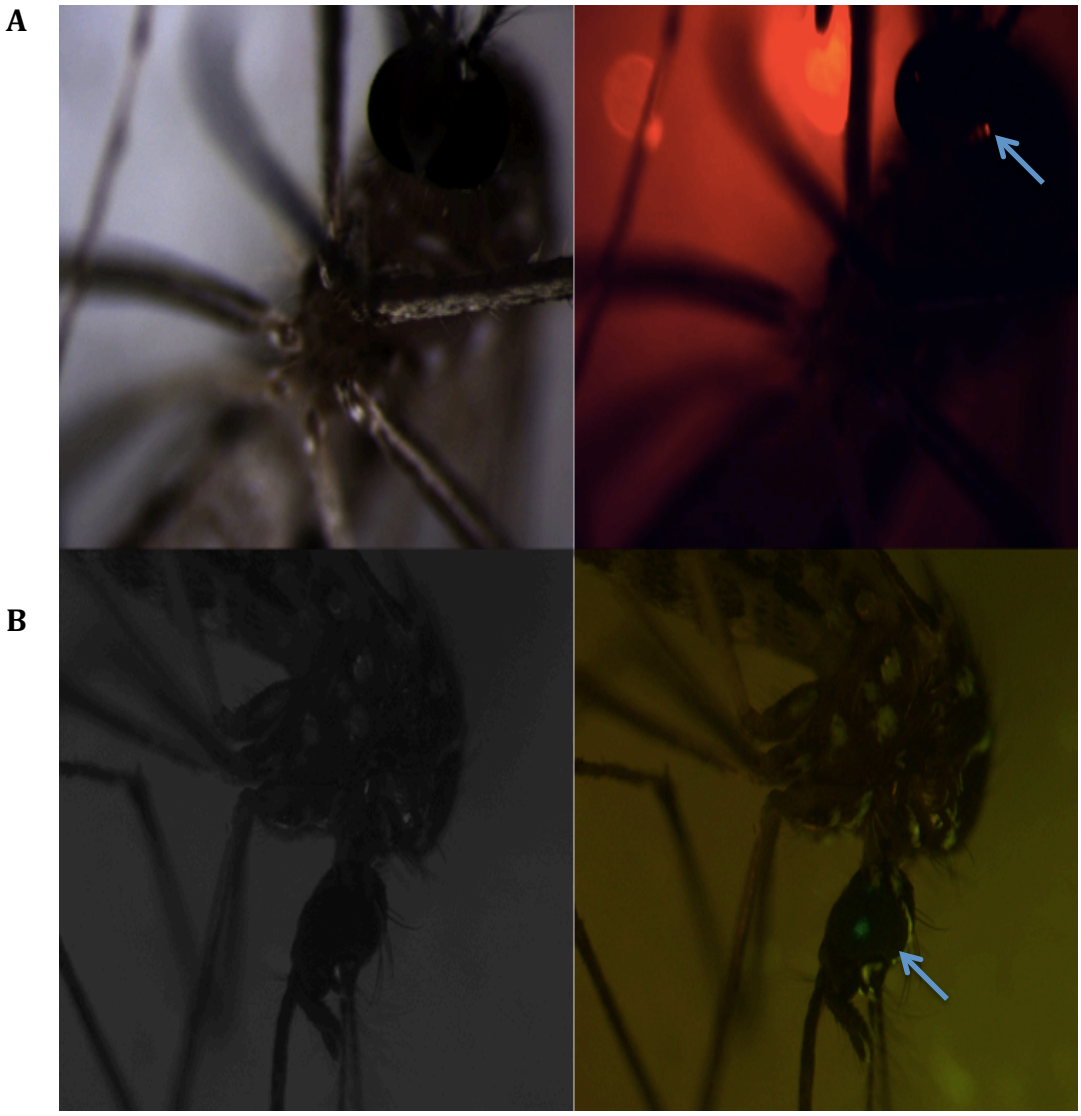


Figure 4.4 *Ae. aegypti* transgenic mosquitoes without and with filter A. Expression of DsRed marker in eyes in adult for *Muta1* element B. Expression of EGFP marker in eyes in adult for *piggyBac* element.

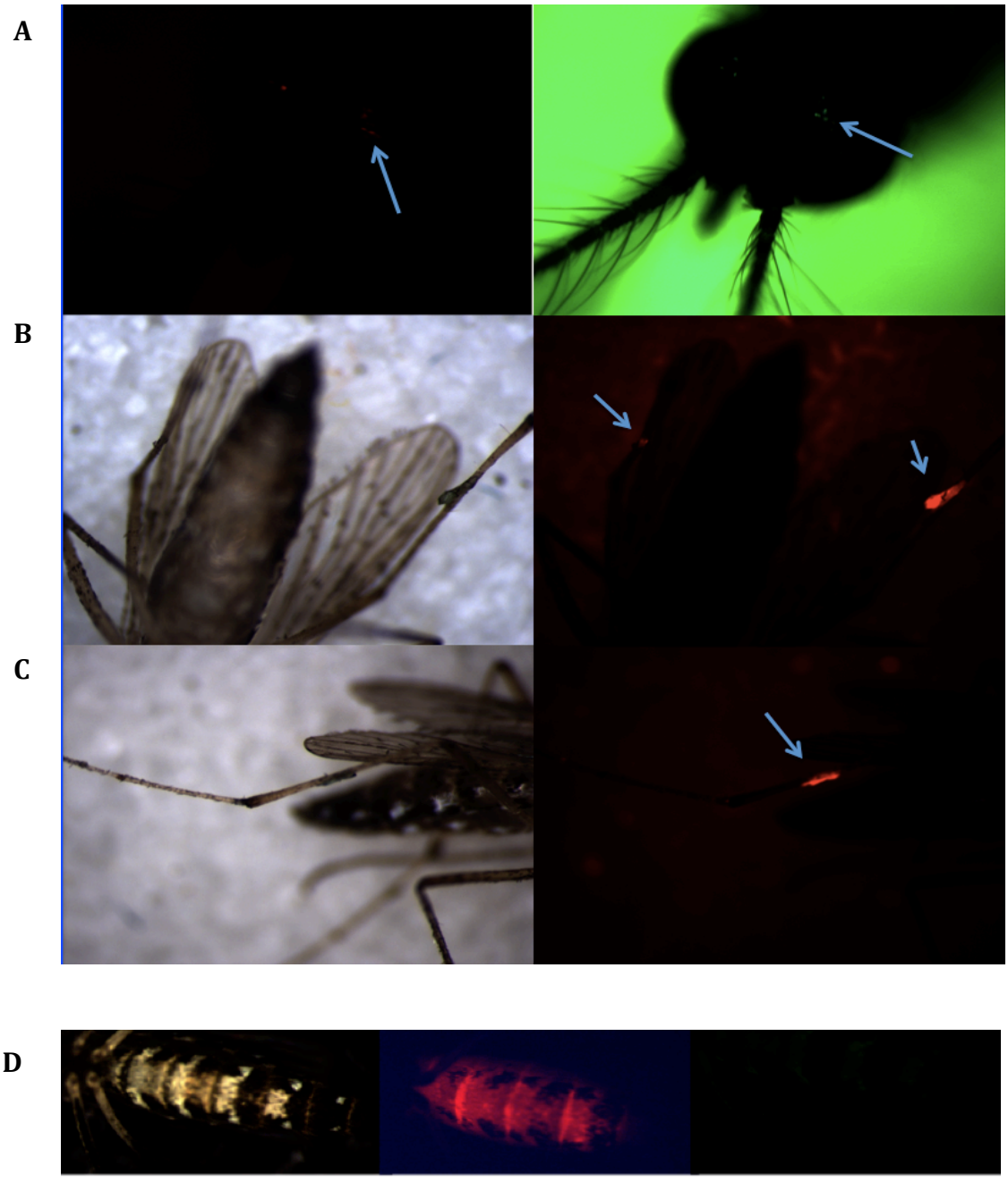


Figure 4.5 Potential new phenotypes observed in crosses for *Ae. aegypti* remobilization experiment for *Muta1* element. A. Adult with new phenotypes expression in legs still retains DsRed expression in eyes, B. Lateral view of expression of DsRed marker in legs, C. Dorsal view of expression in legs and D. DsRed expression in the midgut**.**

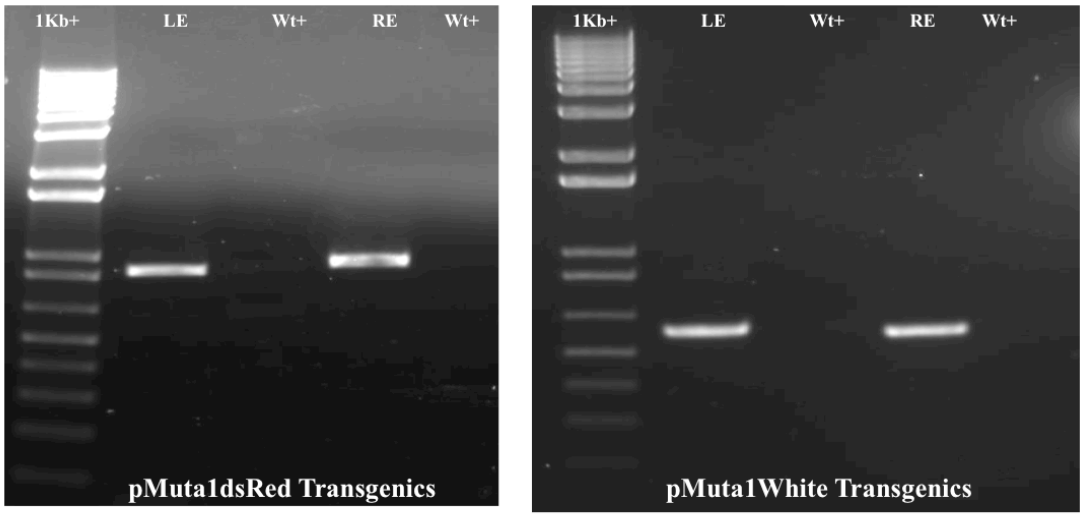


Figure 4.6 Transgenic flies with DsRed marker and White gene marker were confirmed through PCR, where primers were designed for left end and right end of the *Muta1* element and control was CSw+ wild-type flies.

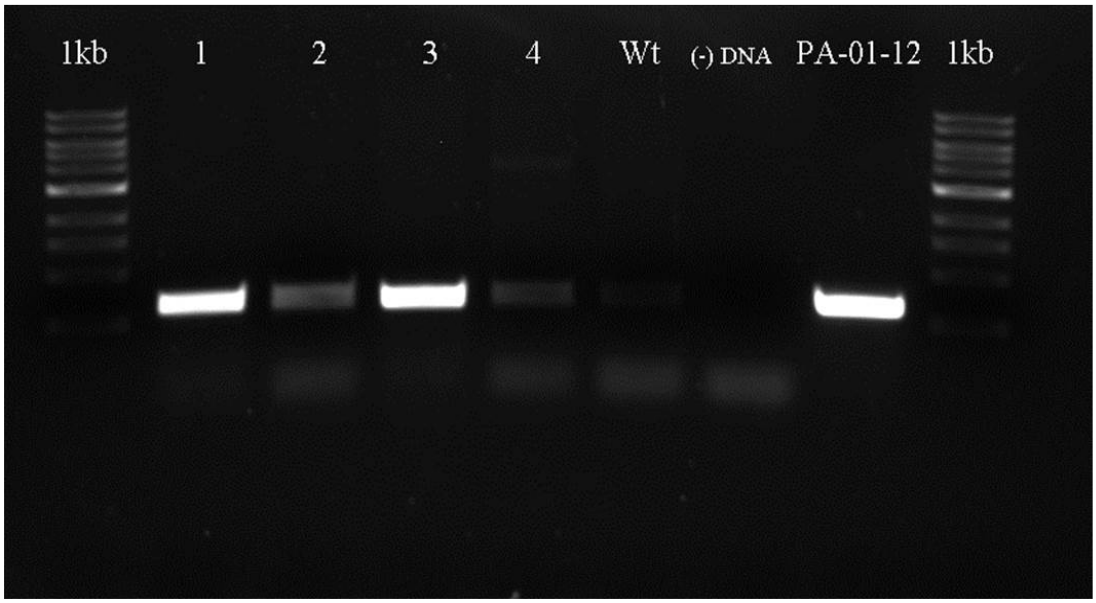


Figure 4.7 Transgenic mosquitoes with pMuta13EHspBac were confirmed through PCR. Lane 1: Muta1 line 1 (PA0112-F1.RG), Lane 2: Muta1 line 2 (PA0112-F2.R), Lane 3: Muta1 line 3 (PA0112-F7.R), Lane 3: Muta1 line 4 (PA0212-M1.R), Lane 5: wild-type Orlando, Lane 6: zero DNA control and Lane 8: positive control pMuta13EhspBac.

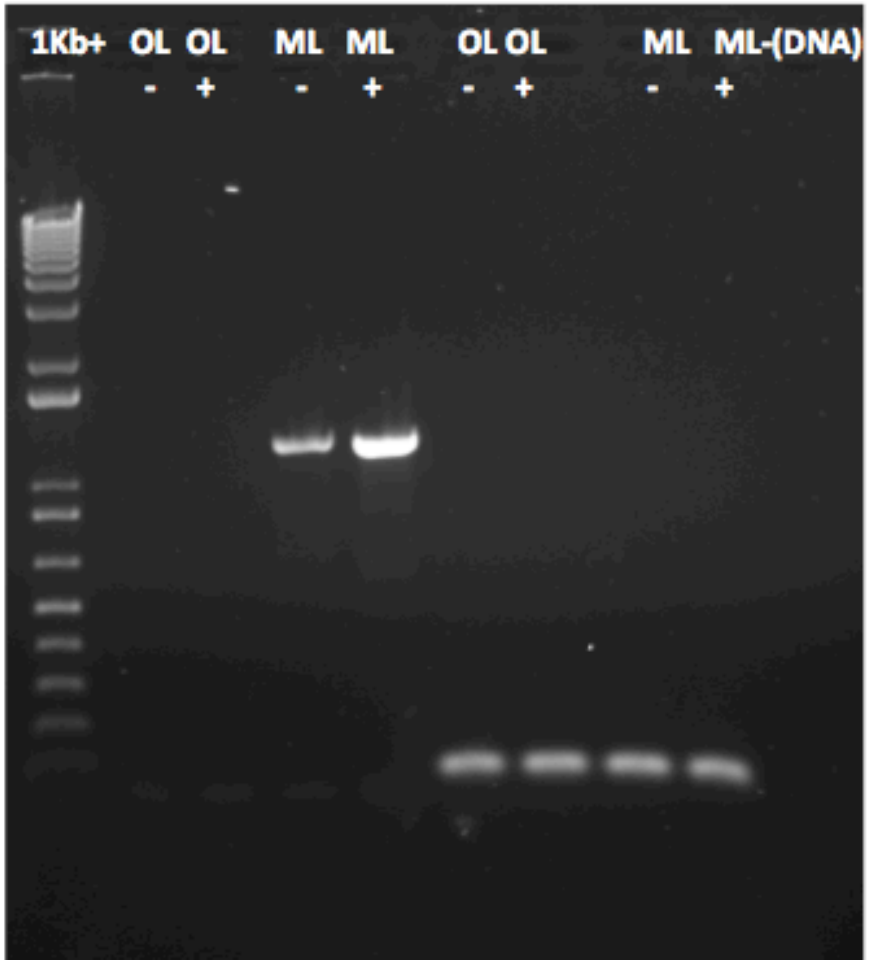


Figure 4.8 RT-PCR is performed on *Muta1* transgenic lines in to check for heat shocked promoter activity in third instar larvae. Lane 1: non-heat shocked Orlando larvae, Lane 2: heat shocked Orlando larvae, Lane 3: non-heat shocked *Muta1* larvae Lane 4: heat shocked *Muta1* larvae. Lane 5-8 has RPS7 gene as a control form Orlando non-heat shocked, heat shocked, *Muta1* non-heat shocked and heat shocked larvae and Lane 9: zero DNA control.

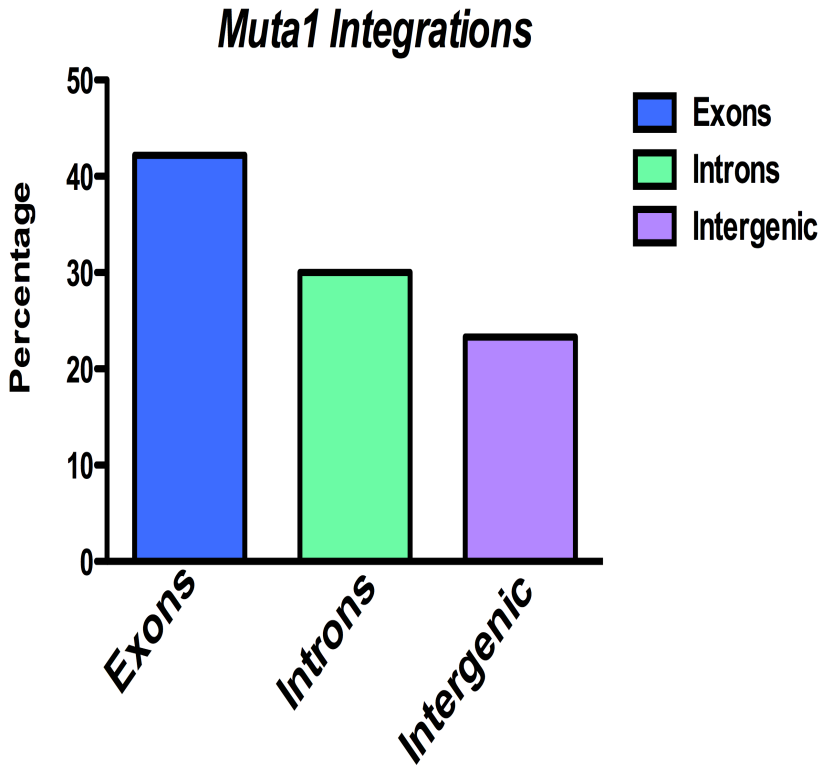


Figure 4.9 Percentage of new *Muta1* integrations, in exons, introns and in intergenic region of *D. melanogaster*.

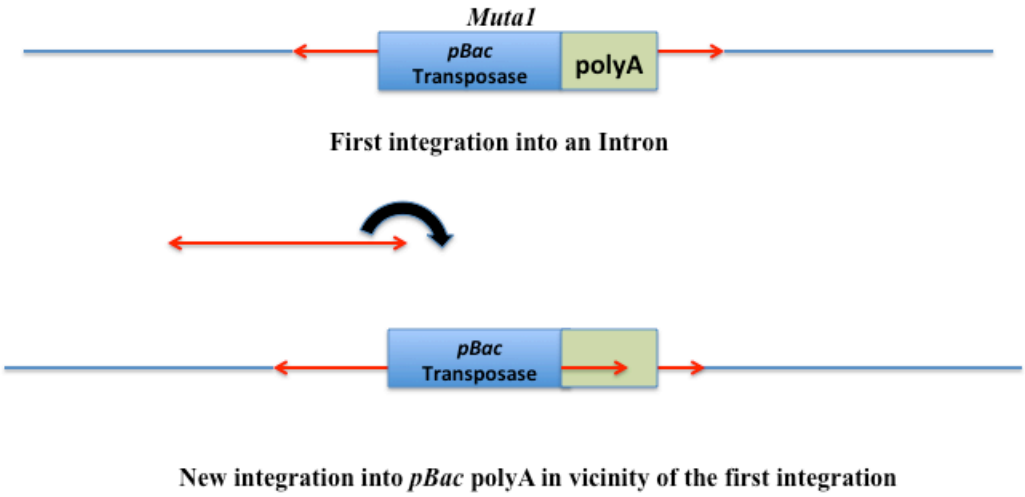


Figure 4.10 Local hopping observed with the *Muta1* element in *D. melanogaster*. The *Muta1* element integrates into itself into a *piggyBac* polyA region. *piggyBac* polyA was cloned in between the ends of the *Muta1* element in the donor plasmid used for the transformation experiment.

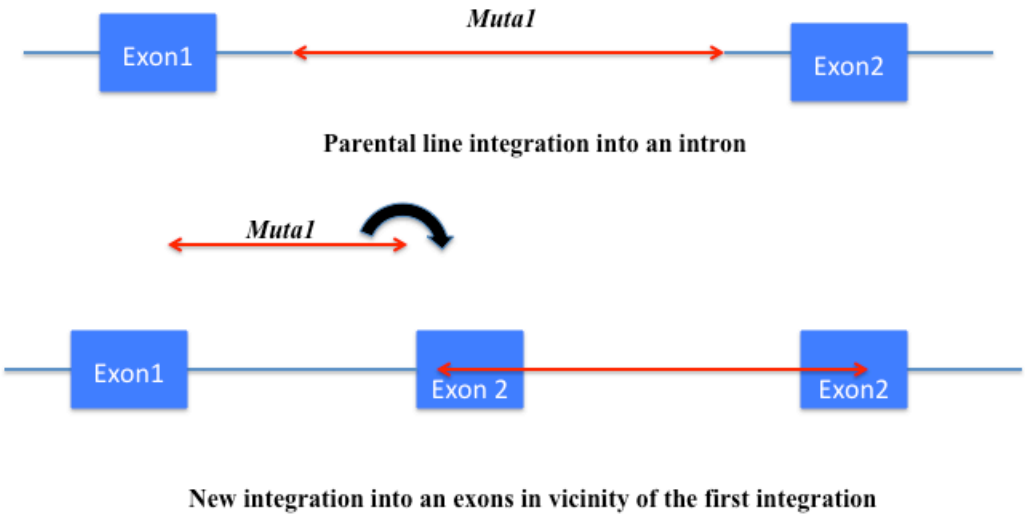


Figure 4.11 Local hopping patterns observed with the *Muta1* element in *D. melanogaster*. The *Muta1* element inserts into an exon, which is in near vicinity of the parental integration into an intron.

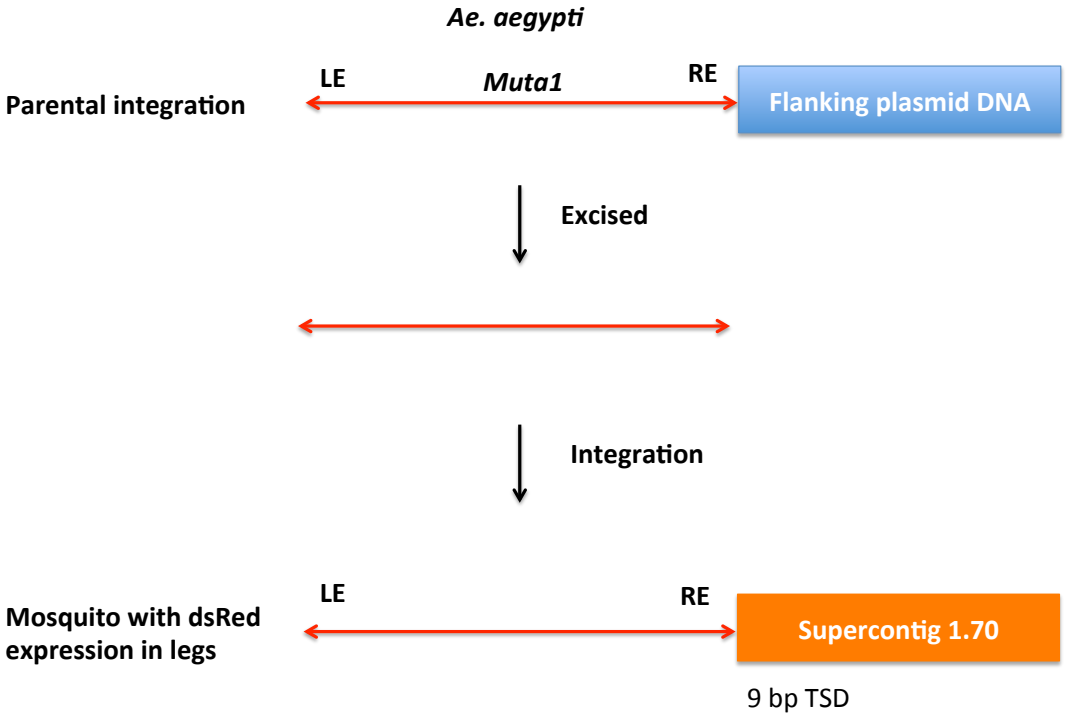


Figure 4.12 The new integration location obtained for the *Muta1* element with DsRed expression in legs in *Ae. aegypti*. The proposed pattern of the *Muta1* element excision and integration into a new location.

Chapter 5. Relationship between piRNAs abundance and *Muta1* transposon in *D. melanogaster*.

5.1 Abstract

Transposable elements impose a threat to the genome because of their propensity to insert into or near genes, creating mutations that disable genes function. The piRNA pathway in *D. melanogaster* is known to participate in regulation of transposable elements (Castañeda, Genzor, & Bortvin, 2011). The small RNA piRNAs are 24-30 nucleotides in length and are generated from long precursor transcripts that are processed in a Dicer independent pathway. This study explores the role of piRNAs in the regulation of the *Muta1* element in *D. melanogaster.* No piRNAs were detected against the *Muta1* element in the wild-type *D. melanogaster,* meaning that this model insect is a perfect platform for studying this element. An autonomous *Muta1* element was constructed into a plasmid and introduced into wild-type embryos of *D. melanogaster* to generate transgenic lines. Progeny were heat shocked to activate the transposase expression. New phenotypes were recorded that were indicative of the movement of *Muta1* element in the genome of *D. melanogaster*. Small RNA libraries were made from generations 11 and 21 of the *Muta1* transgenic lines to detect whether production of piRNAs to the newly introduced *Muta1* element occurred. These libraries revealed that piRNAs were produced that mapped along the length of *Muta1* element, to *Muta1* transposase, the Hsp70 and DsRed promoters, and their polyA regions. Observation of a U1-A10 overlap would be indicative of silencing of a transposon through the secondary piRNA pathway by ping-pong amplification, U1-A10 overlap observed in these lines were not significant, which suggested that the *Muta1* element was not silenced in these transgenic lines by this

secondary pathway.

5.2 Introduction

Germline cells have a particular need to protect their genome, since they must faithfully transmit genetic information to offspring. In animals transposons are suppressed by the piRNA pathway, in which piRNAs binds to member of a specific PIWI clade of Argonaute proteins. Piwi, Aubergine (AUB) and Argoanute3 (AGO3) (Brennecke et al., 2007a). piRNAs were first identified through studies on the *Drosophila Stellate* locus (Aravin et al., 2003), The stellate protein itself has no function, but mutations in the *suppresser of Stellate [Su(ste)]* locus lead to over production of Stellate protein which leads to Stellate crystal formation and reduced fertility. Later, through small RNA cloning and sequencing studies it was shown that *Su(ste)* is required for production of piRNAs directed towards *Stellate* locus (Livak, 1990). These small RNAs that bind to Piwi proteins were initially termed as rasiRNAs for repeat associated small interfering RNAs, since they showed homology to repeat elements (Aravin et al., 2003).

The piRNAs are believed to be generated from transcripts from piRNA clusters, which are filled with TE sequences (Brennecke et al., 2007b). The piRNA clusters are transcribed in the sense and antisense direction from long single stranded RNAs, which serve as the basis for piRNA production. Transposon rich regions are packaged into heterochromatin by specific modification of histones, in which the major modification is methylation of H3 lysine 9 (H3K9) (Soppe et al., 2002) (Grewal & Elgin, 2007). These modifications recruit heterochromatin protein 1 (HP1), which then promotes

transcriptional silencing of transposon transcripts (Brower-Toland et al., 2007a) (Moshkovich & Lei, 2010). In *Drosophila*, these heterochromatic regions serve as a source of piRNA production (Brower-Toland et al., 2007b) (Rangan et al., 2011). There are two-biogenesis pathways that are important for piRNAs production. In the primary piRNA pathway, the long transposon transcript is initially cleaved by the nuclease zucchini (Ipsaro, Haase, Knott, Joshua-Tor, & Hannon, 2012), which has been hypothesized to generate the 5' end of the primary piRNAs. The primary antisense transcripts are transcribed from transposons and/or piRNA clusters, which are processed to piRNAs by an unknown mechanism. The primary piRNAs are then loaded onto PIWI in somatic cells or Aubergine (AUB) in the germline (Siomi, Sato, Pezic, & Aravin, 2011a). The PIWI protein is expressed only in somatic cells, whereas germline cells contains AUB and AGO3 (Malone & Hannon, 2009). The primary piRNAs are found to be antisense and have 5'-Uridine and it is believed that primary piRNAs are responsible for priming germline ping-pong amplification. In secondary piRNA biogenesis, also referred to as the ping-pong mechanism, it has been hypothesized to involve the slicer activity of AUB and AGO3 (Gunawardane et al., 2007). AUB-associated antisense piRNAs guide the RISC complex via sequence complementarity, which cleaves the piRNA precursor from the sense strand at a point 10 nucleotide downstream from the 5'end of the primary antisense piRNAs. The 5' end of the sense piRNA is generated in this manner, which is then loaded onto AGO3. AGO3 associates with sense piRNAs, which can then cleave the piRNA precursor deriving from the antisense strand, generating the 5' end of antisense piRNAs that are subsequently loaded onto AUB. The

3’ end of these piRNAs is further processed by an unknown mechanism, which is then followed by methylation by HEN1 protein (Siomi, Sato, Pezic, & Aravin, 2011b) (Brennecke et al., 2007c).

The absence of piRNAs has severe outcomes, which was demonstrated by the study of *P*-element in *D. melanogaster*. Flies from a P strain have *P*-element in their genomes, but flies of the M strain do not have *P*-element. When *P*-element carrying females were crossed with laboratory strain males, piRNAs were deposited maternally into the embryos and the resulting offspring had wild-type phenotypes (Khurana et al., 2011a). Whereas in the reciprocal cross, when *P-element* carrying males were crossed with laboratory females, the resulting offspring were dysgenic due to lack of maternally derived piRNAs of the *P*-element, which leads to activation of *P-element* and their transposition in the genome causing sterility and gonadal dystrophy. This phenomenon is called P-M hybrid dysgenesis (Simmons et al., 2014a), which is now believed to be due to absence of piRNAs in M-strain (Jensen, Stuart, Goodpaster, Goodman, & Simmons, 2008) (Brennecke et al., 2008).

The hybrid dysgenesis phenomenon has also been observed for *Penelope* retroelement in *Drosophila virilis* (Evgen’ev et al., 1997) (Pyatkov et al., 2002). In one study, the *Penelope* element was introduced into the genome of *D. melanogaster*, which was naïve to the *Penelope* element and transgenic lines were produced (Pyatkov et al., 2002). (Kapitonov & Jurka, 2003). These transgenic lines were then used to study the expression and localization of the *Penelope* element and the biogenesis of *Penelope*-derived small

RNAs in *D. melanogaster* (Rozhkov et al., 2010). A decade later, *Penelope* derived siRNAs and piRNAs were identified in testis and ovaries of the transformed strain (Rozhkov et al., 2010). piRNAs were discovered in 2006 (Aravin et al., 2006) (Girard, Sachidanandam, Hannon, & Carmell, 2006) and their role in transposon regulation was revealed in 2007 (Brennecke et al., 2007b). Therefore, there is a possibility that piRNAs might have been generated to *Penelope* element in previous generations. It was also proposed that appearance of the piRNAs might have arose from accidental transposition into one of the piRNAs cluster, as *Penelope* insertions from two strains within the major *D. melanogaster* germline specific piRNA cluster 42AB were detected (Brennecke et al., 2007a). Progeny from a cross, in which males carrying the *Penelope* element were crossed with females that lacked *Penelope* element, exhibited high levels of *Penelope* transcription was detected in the hybrid (Rozhkov et al., 2013). In the reciprocal cross *Penelope* transcription was not observed. Thus, the authors concluded that the transcription of *Penelope* was suppressed by piRNAs that were maternally deposited. (Rozhkov et al., 2013).

Transposon silencing is not the only role played by the piRNA pathway. In *D. melanogaster* abdominal segmentation is controlled by *Nanos*. *Nanos* (*nos*) is a morphogen that is expressed in a gradient from the posterior to anterior pole of *D. melanogaster* embryos (Gavis & Lehmann, 1994). It was demonstrated that an interaction between piRNAs and *nanos* mRNA is required for *nanos* mRNA deadenylation and translational repression (Rouget et al., 2010). Mutation or knock down of PIWI proteins resulted in stabilized *nanos* mRNA leading to head developmental

defects (Temme, Simonelig, & Wahle, 2014). HSP90α is a major chaperon protein that has been shown to regulate piRNA biogenesis, knockout mutants of *HSP90α* resulted in a large reduction in the expression of primary and secondary piRNAs, as well as mislocalization of MIWI2, a PIWI homolog in mammals (Ichiyanagi et al., 2014). A recent finding has shown that piRNAs play role in long-term changes in neurons for the persistence of memory. The piRNA/PIWI complex facilitates serotonin-dependent methylation of conserved CpG island in the CREB2 promoter, which is a major inhibitory constraints of memory in Aplysia, leading to long-term synaptic facilitation (Rajasethupathy et al., 2012)

Understanding the mechanisms of transposon regulation is essential to the optimization of transposons as genetic tools in insects. Thus, the study of piRNAs involved in regulation of *Muta1* in *D. melanogaster*, where the genome is naïve will help us to understand its expression, localization and biogenesis of *Muta1*-derived piRNAs. Moreover, it will also enable us to determine the targeted region in the *Muta1* element for its suppression, which can then be used to manipulate the structure of transposon in a way that it can evade it detection for silencing by piRNAs.

5.3 Material and Methods

5.3.1 Plasmid construction

pMuta1Auto was constructed by ligation of a fragment that contains Hsp70-*Muta1* transposase-polyA amplified through PCR from pMuta1 helper, and the vector

pMuta1DsRedTFD. The vector was digested with FastDigest *EcoRV* (Thermo-Fisher), treated with FastAP (Thermo-Fisher), and purified by agarose gel electrophoresis. The hspMuta1 fragment was amplified with the primers Hsp70 FP and PolyA RP using Phusion polymerase (New England Biolabs) and the program: 98° for 30 sec, 5 × (98° for 8 sec, 58° for 20 sec, 72° for 1.5 min), 25 × (98° for 8 sec, 67° for 20 sec, 72° for 1.5 min), 72° for 5 min, 4°. Column purified PCR product was digested with FastDigest *EcoRV,* column purified and digested with *Xma1* (New England Biolabs) using NEB buffer 4 and column purified. Vector and insert were ligated with T4 DNA ligase (Thermo-Fisher), transformed into DH10B competent cells, colonies checked with *Sac1* (Thermo-Fisher) and *BglII* (Thermo-Fisher) digestion for correct orientation of *Muta1* transposase fragment, and were sequenced to verify the sequence and the orientation of the transposase.

5.3.2 Embryo microinjections.

Microinjection of *D. melanogaster* embryos was performed as described in chapter three section 3.3.10, with a 250 ng/ul of pMuta1Auto plasmid.

5.3.3 Establishment of crosses in *D. melanogaster* for autonomous *Muta1* experiment

In this experiment, transgenic lines were self-crossed. pMuta1Auto transgenics expressed DsRed gene as the genetic marker and encoded *Muta1* transposase under heat shock promoter Hsp70. Four crosses were established, 20 males from line A19 with DsRed marker were self-crossed with 20 females. Similarly, line A50, A56, and A85 crosses

were also established. Resulting progeny were allowed to be homozygous and screened based on the presence of transgenic constructs. The progeny from these crosses were heat shocked at 37°C for 1 hour everyday until they reach adult stage. Progeny were screened for several generations to detect phenotypic change in expression for DsRed marker. These flies were then used for small RNA library preparation.

5.3.4 Small RNA library construction.

The transgenic flies from lines A19 G11, A19 G21, A50 G11, A50 G21, A56 G11, A56 G21, A85 G11, and A85 G21 were used for small RNA library construction. Five males were crossed with 5 virgin females from each line. Five days later both sexes were used for RNA extraction using TRIzol® and small RNA library was prepared using NEBNext® Mutliplex Small RNA Library Prep set for Illumina® (Set 2). The runs were sequenced on an Illumina Genome Analyzer II DNA Sequencer at the University of California, Riverside's Genomic Core. Total eight libraries were prepared, two libraries for each line from generations 11 and 21. Generation 11 was selected based on new phenotypes observed in generation 7 for lines A56 and A85, which suggests *Muta1* element remobilization in the genome. Generation 21 was selected to determine change in piRNAs expression profile compared to generation 11.

5.3.5 Analysis of small RNA libraries.

Small RNA library were processed in order to remove adapter sequence and were normalized based on size using R-Bioinductor-DESeq package (Anders et al., 2013). Size

selection for piRNA was carried out using R- Bioconductor-Shortread package (Morgan et al., 2009), where 23-30 nt size selection was done and greater than 23 nt for the most. It also performs removal of sequences with Phred quality score of less than 20; which were nucleotides with low confident base call. Small RNA sequences with more than 3 N's were removed and sequences with greater than 20 nt for poly –A, T, C, or G were also removed.

R script (Script 5.1) was used to remove obvious contaminants within the library by mapping the small RNA to known ribosomal RNAs with the SILVA database using Bowtie2 default parameters and eliminating these hits from the library.

The small RNA libraries were generated from two generations; G11 and G21 from *Muta1* transgenic lines A19, A50, A56, and A85, which were then mapped to the autonomous *Muta1* plasmid using Bowtie2 parameters. The control small RNA library was generated from wild-type *D. melanogaster* (RNAlib14), which was analyzed to determine presence of piRNAs to *Muta1* element in *D. melanogaster*. piRNAs were not detected to *Muta1* element and thus, *D. melanogaster* genome was naïve to *Muta1* element (Figure 5.4).

All libraries were normalized based on size with DEseq parameters (Anders et al., 2013). An average was calculated from eight libraries, relative size factor was obtained by dividing size of the library by the average. Normalized counts for each library were obtained by dividing number of hits mapped to the *Muta1* plasmid by the relative size factor (Table 5.5).

Small RNAseq coverage plots were created using custom R script (Script 5.2). It was a simple line plot representing the stranded coverage of piRNAs across a particular reference, which was pMuta1Auto plasmid for piRNA alignment in this case. In this script, essential feature of the *Muta1* element were distinguished based on the color such as; DsRed promoter was color coded red, DsRed polyA was pink, Hsp70 promoter was yellow, *Muta1* transposase was orange, Hsp70 polyA was cyan, left end and right end were color coded light gray. The flanking plasmid DNA was white in color, as three of the transgenic lines had integration with flanking plasmid DNA, it was important to determine if piRNAs were generated to the plasmid DNA. The piRNA hits obtained to the negative strand were color-coded to black and they represent antisense piRNAs and hits to the positive strand were color coded blue, which represent sense piRNAs in the coverage plot. The X- axis represents the length of the pMuta1Auto plasmid and the Y-axis represents number of piRNA hits to a particular region in the map.

Both of these R scripts were written by Patrick Schreiner from Atkinson laboratory, University of California, Riverside.

5.4 Results

5.4.1 An autonomous *Muta1* transposon can transform *D. melanogaster.*

Germ-line transformation was achieved using pMuta1Auto as the reporter plasmid (Figure 5.1). A total of 203 embryos injected with pMuta1Auto survived to adulthood (a survival rate of approximately 68%) and of those, 100 (88% of G0 crosses) generated progeny upon backcrossing (Table 5.1). Screening the G1 progeny for DsRed expression

showed that 16% of the autonomous *Muta1* element crosses produced transgenic offspring (Table 5.1). Out of the total 16 transgenic lines, four of the lines were randomly selected for further experiment. These lines were self-crossed and homozygouzed, their progeny were heat shock to activate the heat shock promoter Hsp70 controlling *Muta1* transposase expression. The genomic site of integration of each *Muta1* transposon in these lines was determined (Table 5.2). Transgenic line A19 showed perfect integration into left arm of chromosome 3, where as the other three lines had integration with the flanking plasmid DNA with intact *Muta1* element. Nonetheless these three lines could serve as parental lines for subsequent remobilization of the *Muta1* transposon.

5.4.2 Remobilization of autonomous *Muta1*.

New phenotypes were detected in transgenic lines, A56 and A85 (Table 5.3). Three new phenotypes were observed from screening of 11,096 flies (Table 5.3). These new phenotypes were DsRed expression in the full body, DsRed expression in the thorax and DsRed expression in the abdomen (Figure 5.2).

Molecular characterization of flies with DsRed expression in eyes, showed two new integrations for the autonomous *Muta1* element (Table 5.4). The parental lines had integration with the flanking plasmid DNA, which made molecular characterization of exact integration location in these lines difficult. Further molecular verification is required to determine exact location of the *Muta1* element integration in the parental lines. 15 flies from line A56 G9 showed integration into an intergenic region between

Lpin/CG8709 and Kermit/CG11546 gene on chromosome 2R forming a 9 bp TSD. Single fly from line A85 G16 was characterized to have integration into chromosome 3R with a 9 bp TSD, into a repeat region (Table 5.4).

5.4.3 Small RNA library analysis.

Coverage plots generated for CSw^{+} control library (RNAlib14) indicated that 12 piRNAs mapped to autonomous *Muta1* plasmid (Figure 5.3.1), 4 of these piRNAs mapped to Hsp70 promoter, 4 piRNAs mapped to the Hsp70 polyA region, a piRNA mapped to the DsRed promoter and 3 piRNAs to the DsRed polyA region. This analysis did not detect any piRNAs that mapped to *Muta1* transposase or to the ends of *Muta1* element, confirming that the D. melanogaster genome was naïve to the *Muta1* element

In the library analysis of library A19 G11 (Figure 5.3.2), 98 piRNAs mapped to the autonomous *Muta1* plasmid. Out of these 98 piRNAs, 70.4% were antisense and 29.6% were sense piRNAs. Furthermore, 62% of these piRNAs mapped to the *Muta1* element, of the piRNAs that mapped to the *Muta1* element, 62.3% mapped to Hsp70 promoter, 9.8% mapped to the *Muta1* transposase, 3.3% mapped to the DsRed gene and 23% mapped to Hsp70 polyA. The piRNAs that mapped to Hsp70 promoter, 97% were antisense piRNAs. The piRNAs that mapped to *Muta1* transposase, all were sense piRNAs. The piRNAs that mapped to DsRed gene, all were sense piRNAs and the piRNAs that mapped to Hsp70 polyA, 86% were antisense piRNAs.

In the library analysis for A19 G21 (Figure 5.3.3), 257 piRNAs mapped to the autonomous *Muta1* plasmid. Out of these 257 piRNAs, 69% were antisense piRNAs and 31% were sense piRNAs. Further, 79.7% of these piRNAs mapped to the *Muta1* element. Of these piRNAs that mapped to the *Muta1* element, 50% mapped to Hsp70 promoter, 10.4% mapped to *Muta1* transposase, and 31% mapped to Hsp70 polyA. The piRNAs that mapped to Hsp70 promoter, 92% of the piRNAs were antisense. The piRNAs that mapped to *Muta1* transposase, 70% were sense piRNAs, and the piRNAs that mapped to Hsp70 polyA, 73.4% of the piRNAs were antisense.

In the library analysis of A50 G11 (Figure 5.3.4), 174 piRNAs mapped to the autonomous *Muta1* plasmid. Out of these 174 piRNAs, 50% were antisense piRNAs and 50% were sense piRNAs. Further, only 13 piRNAs mapped along the length of *Muta1* element. From the piRNAs that mapped to *Muta1* element, 1.2% mapped to Hsp70 promoter and 4.6% mapped to *Muta1* transposase. The piRNAs that mapped to Hsp70 promoter, all of the piRNAs were antisense. The piRNAs that mapped to Muta1 transposase, 90% of the piRNAs were antisense.

In the library analysis of A50 G21 (Figure 5.3.5), a total of 198 piRNAs mapped to autonomous *Muta1* plasmid. Out of these 198 piRNAs, 73% were antisense piRNAs and 28% were sense piRNAs. Further, 88% of these piRNAs mapped to the *Muta1* element. From the piRNAs that mapped to *Muta1* element, 40% mapped to Hsp70 promoter, 18.9% mapped to *Muta1* transposase and 29.3% mapped to Hsp70 polyA. Of the piRNAs that mapped to Hsp70 promoter, 90% of the piRNAs were antisense, the piRNAs that

mapped to *Muta1* transposase, 61% of the piRNAs were antisense, and the piRNAs that mapped to Hsp70 polyA, 82% were antisense.

In the library analysis of A56 G11 (Figure 5.3.6), a total of 220 piRNAs mapped to the autonomous *Muta1* plasmid. Out of these 220 piRNAs, 73% were antisense piRNAs and 28% were sense piRNAs. Further, 92% of the piRNAs mapped to the *Muta1* element. From the piRNAs that mapped to *Muta1* element, approximately 36% mapped to Hsp70 promoter, 35% mapped to *Muta1* transposase, 11.4% mapped to Hsp70 polyA and 12.3% mapped to DsRed promoter. The piRNAs that mapped to Hsp70 promoter, 97% of the piRNAs were antisense piRNAs, the piRNAs that mapped to *Muta1* transposase, 94% of the piRNAs were sense, the piRNAs that mapped to Hsp70 polyA, 61% were sense, and the piRNAs that mapped to DsRed promoter, all piRNAs were sense.

In the library analysis of A56 G21 (Figure 5.3.7), 87 piRNAs mapped to the autonomous *Muta1* plasmid. Out of these 87 piRNAs, 61% were antisense piRNAs and 39% were sense piRNAs. Further, 68% of the piRNAs mapped along the length of the *Muta1* element. From the piRNAs that mapped to *Muta1* element, 60% mapped to Hsp70 promoter and approximately 27% mapped to *Muta1* transposase. The piRNAs that mapped to Hsp70 promoter, 86% of the piRNAs were antisense. The piRNAs that mapped to *Muta1* transposase, approximately 69% of the piRNAs were antisense.

In the library analysis of A85 G11 (Figure 5.3.8), a total of 1066 piRNAs mapped to autonomous *Muta1* plasmid. Out of these 1066 piRNAs, 45% were antisense piRNAs and 54.7% were sense piRNAs. Further, 64% of the piRNAs mapped along the length of the

Muta1 element. From the piRNAs that mapped to *Muta1* element, approximately 32% mapped to Hsp70 promoter, 25% mapped to *Muta1* transposase, 23% mapped to Hsp70 polyA and 16.5% mapped to DsRed promoter. The piRNAs that mapped to Hsp70 promoter, 86% of the piRNAs were antisense piRNAs, the piRNAs that mapped to *Muta1* transposase, 88.9% of the piRNAs were sense, the piRNAs that mapped to Hsp70 polyA, 53.5% were antisense, and the piRNAs that mapped to DsRed promoter, all piRNAs were sense.

In the library analysis of A85 G21 (Figure 5.3.9), 119 piRNAs mapped to autonomous *Muta1* plasmid. Out of these 119 piRNAs, 55% were antisense piRNAs and 45% were sense piRNAs. 99% of the piRNAs mapped to the *Muta1* element, approximately 41% mapped to Hsp70 promoter, 22% mapped to *Muta1* transposase, and 17% mapped to Hsp70 polyA. The piRNAs that mapped to Hsp70 promoter, 76% of the piRNAs were antisense piRNAs, the piRNAs that mapped to *Muta1* transposase, 73% of the piRNAs were sense, and the piRNAs that mapped to Hsp70 polyA, 75% were antisense.

I also analyzed, the preference for Uracil at position 1 (U1) from the 5' end in antisense piRNAs and Adenine at 10^{th} position (A10) of sense piRNAs, which is a signature of ping-pong amplification in piRNA biogenesis (Brennecke et al., 2007b) (Gunawardane et al., 2007). 15-27% antisense piRNAs had U1 and 7-30% sense piRNAs had A10 from 5'end (Table 5.5). Further analysis revealed few overlaps in these libraries. In the library A19 G21, a U1-A10 overlap was observed that mapped to the Hsp70 promoter. In the library A50 G21, two overlaps were observed that mapped to the Hsp70 promoter, one

overlap was observed that mapped to the *Muta1* transposase, three overlaps were observed for the DsRed promoter and one overlap was observed for the DsRed polyA (Table 5.8). In the library analysis for A56 G11, two overlaps were observed that mapped to the *Muta1* transposase and the DsRed promoter, respectively. In the library analysis for A56 G21, two overlaps were observed that mapped to the left end, and a overlap to the DsRed promoter (Table 5.8). In the library A85 G11, one overlap was observed that mapped to the DsRed promoter (Table 5.8). In the libraries A19 G11, A50 G11, and A85 G21 no U1A10 overlap was detected.

The abundance of the antisense and the sense piRNAs were analyzed for four transgenic lines A19, A50, A56, and A85 from generations G11 and G21. In the transgenic line A19, the piRNAs that mapped to Hsp70 promoter, 50% to 80% of the piRNAs were antisense, and 1.5-fold increase in antisense piRNAs was observed in G21 compared to G11. The piRNAs that mapped to the *Muta1* transposase, 70% to 100% of the piRNAs were sense piRNAs, 1.8-fold increase in sense piRNAs was observed in G21 compared to G11. The piRNAs that mapped to Hsp70 polyA region, 73% to 86% of the piRNAs were antisense piRNAs, and 2.3-fold increase in antisense piRNAs was observed in G21 compared to G11. The piRNAs that mapped to DsRed gene, 100% of the piRNAs were sense piRNAs, and 2-fold increase in sense piRNAs was observed in G21 compared to G11 (Table 5.6) (Figure 5.4).

In the transgenic line A50, the piRNAs that mapped to Hsp70 promoter, 98% to 100% of the piRNAs were antisense, and 9-fold increase in pool of antisense piRNAs was

observed in G21 compared to G11. The piRNAs that mapped to the *Muta1* transposase, 40% to 87% of the piRNAs were sense piRNAs, 5-fold increase in antisense piRNAs was observed in G21 compared to G11. The piRNAs that mapped to Hsp70 polyA region, 82% to 100% of the piRNAs were antisense piRNAs, and 10.7-fold increase in antisense piRNAs was observed in G21 compared to G11. The piRNAs that mapped to DsRed gene, 100% of the piRNAs were sense piRNAs, and 9.3-fold increase in sense piRNAs was observed in G21 compared to G11 (Table 5.6) (Figure 5.5).

In the transgenic line A56, the piRNAs that mapped to Hsp70 promoter, 86% to 97% of the piRNAs were antisense, and 2.4-fold decrease in pool of antisense piRNAs was observed in G21 compared to G11. The piRNAs that mapped to the *Muta1* transposase, 31% to 93.5% of the piRNAs were sense piRNAs, 15-fold decrease in antisense piRNAs was observed in G21 compared to G11. The piRNAs that mapped to Hsp70 polyA region, 60% of the piRNAs were antisense piRNAs, and no piRNAs were observed in G21 compared to G11. The piRNAs that mapped to DsRed gene, 100% piRNAs were sense piRNAs, and 8.7-fold decrease in sense piRNAs was observed in G21 compared to G11 (Table 5.6) (Figure 5.6).

In the transgenic line A85, the piRNAs that mapped to Hsp70 promoter, 77% to 86% of the piRNAs were antisense, and 1-fold decrease in pool of antisense piRNAs was observed in G21 compared to G11. The piRNAs that mapped to the *Muta1* transposase, 73% to 88% of the piRNAs were sense piRNAs, 1.6-fold decrease in antisense piRNAs was observed in G21 compared to G11. The piRNAs that mapped to Hsp70 polyA

region, 53%-75% of the piRNAs were antisense piRNAs, and 1.1-fold decrease in sense piRNAs was observed in G21 compared to G11. The piRNAs that mapped to DsRed gene, 100% of the piRNAs were sense piRNAs, and 2.9-fold decrease in sense piRNAs was observed in G21 compared to G11 (Table 5.6) (Figure 5.7).

5.5 Discussion

Microinjection of autonomous *Muta1* plasmid into *D. melanogaster* embryos resulted in transformation efficiency of 16%. The transformation efficiency obtained is relatively high compared to the previous experiment performed with non-autonomous *Muta1* element in remobilization experiment (4.9%). It is important to note, though, that the *Muta1* transposase was supplied in *trans* in the latter case. This suggests that transposition frequency might increase with the *Muta1* transposase acting in *cis*, as the transposase protein might have an easy access to the TIRs of the *Muta1* element that might play an important role in DNA binding. Studies with Tn10 transposon was showed to have decrease in transposition efficiency with the increase in the distance between the transposase gene and its binding site (ends of the transposon), (Kleckner, Chalmers, Kwon, Sakai, & Bolland, 1996),

New integration locations for the *Muta1* element were characterized in flies from generation 9 and 16. The parental lines had integration with the flanking plasmid DNA, which made molecular characterization of the exact integration location in these lines difficult. Further molecular verification is required to determine exact location of the *Muta1* element integration in the parental lines. The new phenotypes were observed

starting from generation 7 (Figure 5.2), which were suggestive of autonomous *Muta1* element remobilization in *D. melanogaster.* This provides an evidence for *Muta1* element transposition into new locations in the genome, which suggest that host-silencing machinery might be able to detect the *Muta1* element for suppression. Thus, small RNA libraries were made from generation 11 and 21 to detect piRNAs response to the *Muta1* element. The small RNA libraries analysis, showed *de novo* piRNAs that mapped to the autonomous *Muta1* plasmid, which suggest that the *Muta1* element might be remobilizing at a higher frequency in the germline of *D. melanogaster*.

The phenomenon of hybrid dysgenesis was analyzed in chapter four in relation to introduction of the *Muta1* element in *D. melanogaster,* which was determined to be naïve to the *Muta1* element. No dysgenesis was observed in the progenies from those crosses.

Previous studies with the *P*-element, it was shown that piRNAs were produced for suppressing transcription of the *P*-element, and integration of the *P*-elements in these flies were characterized into telomeric region of the X chromosome, which is one of the piRNA cluster that produces piRNAs (Simmons et al., 2014b). This study also shows the *P*-element piRNAs from opposite strand have significant bias toward a 10 nt overlap, which is the hallmark of ping-pong amplification (Khurana et al., 2011b). The *Muta1* elements were not characterized to have integration into a piRNA clusters. The analyses of all eight libraries showed 50-80% of the piRNAs produced were antisense piRNAs, and they were mapped to Hsp70 promoter, *Muta1* transposase, and Hsp70 polyA. In transgenic line A19, 4-fold increase in piRNAs was observed for generation 21, more

piRNAs were observed that mapped to the *Muta1* transposase. In the transgenic line A50, 26-fold increase in piRNAs was observed for generation 21, with abundance of piRNAs that mapped to Hsp70 promoter, *Muta1* transposase and Hsp70 polyA. In the transgenic line A85, 2.1-fold increase in piRNAs was observed for transgenic line A85 in generation 21. The transgenic line A56 was an exception with 3.4-fold decrease in piRNAs observed in generation 21.

Analysis of the piRNAs from all four transgenics lines and both generations showed that the piRNAs mapped to the *Muta1* transposase, 70 to 100% of the piRNAs were sense piRNAs. The sense piRNAs are produced from secondary piRNA pathway and are bound to AGO3 protein (Brennecke et al., 2007b). The AGO3 protein is restricted to germline, which suggest that the regulation of the *Muta1* element might be germline specific. This also suggests that the *Muta1* element might have integrated into a germline specific piRNAs cluster eventually leading to production of sense piRNAs. Increase in molecular characterization of the *Muta1* element integrations in *D. melanogaster* might help us in understanding the mechanism of the *Muta1* transposon regulation. In addition, these libraries were made from whole adults, which contains larger proportion of somatic tissue. The piRNAs libraries should me made from ovaries that might give us better insight into the piRNAs specificity to the *Muta1* element in the germline.

Moreover, the U1-A10 overlap observed for Hsp70 promoter, *Muta1* transposase, DsRed promoter and DsRed polyA were not significant (Table 5.6). It is possible as more generations are analyzed, the ping-pong amplification might take place, which might

silence *Muta1* element in *D. melanogaster* through amplification of piRNAs. Furthermore, the small RNA libraries were made from flies with expression in eyes for the *Muta1* element. Flies that have lost expression of DsRed in eyes were observed in lines A50, A56 and A85, which suggest excision of the *Muta1* element or suppression of the *Muta1* element that could result in loss of gene expression of DsRed marker. Analysis of the piRNAs from these flies might give us better insight into regulation of the *Muta1* element in *D. melanogaster*.

Thus, the *Muta1* element is still active in these transgenic lines despite production of piRNAs to the *Muta1* transposase. This suggests that the *Muta1* element is somehow able to evade its detection by host silencing machinery and avoid suppression. The *Muta1* element may serve as efficient genetic tool for insect transgenesis.

5.6 References

Anders, S., McCarthy, D. J., Chen, Y., Okoniewski, M., Smyth, G. K., Huber, W., & Robinson, M. D. (2013). Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nature Protocols*, *8*(9), 1765–86. http://doi.org/10.1038/nprot.2013.099

Aravin, A. A., Lagos-Quintana, M., Yalcin, A., Zavolan, M., Marks, D., Snyder, B., … Tuschl, T. (2003). The Small RNA Profile during Drosophila melanogaster Development. *Developmental Cell*, *5*(2), 337–350. http://doi.org/10.1016/S1534-5807(03)00228-4

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., … Tuschl, T. (2006). A novel class of small RNAs bind to MILI protein in mouse testes. *Nature*, *442*(7099), 203–7. http://doi.org/10.1038/nature04916

Brennecke, J., Aravin, A. A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., & Hannon, G. J. (2007a). Discrete Small RNA-Generating Loci as Master Regulators of Transposon Activity in Drosophila. *Cell*, *128*(6), 1089–1103.

Brennecke, J., Aravin, A. A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., & Hannon, G. J. (2007b). Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell*, *128*(6), 1089–103. http://doi.org/10.1016/j.cell.2007.01.043

Brennecke, J., Aravin, A. A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., & Hannon, G. J. (2007c). Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell*, *128*(6), 1089–103. http://doi.org/10.1016/j.cell.2007.01.043

Brennecke, J., Malone, C. D., Aravin, A. A., Sachidanandam, R., Stark, A., & Hannon, G. J. (2008). An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science (New York, N.Y.)*, *322*(5906), 1387–92. http://doi.org/10.1126/science.1165171

Brower-Toland, B., Findley, S. D., Jiang, L., Liu, L., Yin, H., Dus, M., … Lin, H. (2007a). Drosophila PIWI associates with chromatin and interacts directly with HP1a. *Genes & Development*, *21*(18), 2300–11. http://doi.org/10.1101/gad.1564307

Brower-Toland, B., Findley, S. D., Jiang, L., Liu, L., Yin, H., Dus, M., … Lin, H. (2007b). Drosophila PIWI associates with chromatin and interacts directly with HP1a. *Genes & Development*, *21*(18), 2300–11. http://doi.org/10.1101/gad.1564307

Castañeda, J., Genzor, P., & Bortvin, A. (2011). piRNAs, transposon silencing, and germline genome integrity. *Mutation Research*, *714*(1-2), 95–104. http://doi.org/10.1016/j.mrfmmm.2011.05.002

Evgen'ev, M. B., Zelentsova, H., Shostak, N., Kozitsina, M., Barskyi, V., Lankenau, D.-H., & Corces, V. G. (1997). Penelope, a new family of transposable elements and its possible role in hybrid dysgenesis in Drosophila virilis. *Proceedings of the National Academy of Sciences*, *94*(1), 196–201. http://doi.org/10.1073/pnas.94.1.196

Gavis, E. R., & Lehmann, R. (1994). Translational regulation of nanos by RNA localization. *Nature*, *369*(6478), 315–8. http://doi.org/10.1038/369315a0

Girard, A., Sachidanandam, R., Hannon, G. J., & Carmell, M. A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature*, *442*(7099), 199–202. http://doi.org/10.1038/nature04917

Grewal, S. I. S., & Elgin, S. C. R. (2007). Transcription and RNA interference in the formation of heterochromatin. *Nature*, *447*(7143), 399–406. http://doi.org/10.1038/nature05914

Gunawardane, L. S., Saito, K., Nishida, K. M., Miyoshi, K., Kawamura, Y., Nagami, T., … Siomi, M. C. (2007). A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. *Science (New York, N.Y.)*, *315*(5818), 1587–90. http://doi.org/10.1126/science.1140494

Ichiyanagi, T., Ichiyanagi, K., Ogawa, A., Kuramochi-Miyagawa, S., Nakano, T., Chuma, S., … Udono, H. (2014). HSP90α plays an important role in piRNA biogenesis and retrotransposon repression in mouse. *Nucleic Acids Research*, *42*(19), 11903–11. http://doi.org/10.1093/nar/gku881

Ipsaro, J. J., Haase, A. D., Knott, S. R., Joshua-Tor, L., & Hannon, G. J. (2012). The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. *Nature*, *491*(7423), 279–83. http://doi.org/10.1038/nature11502

Jensen, P. A., Stuart, J. R., Goodpaster, M. P., Goodman, J. W., & Simmons, M. J. (2008). Cytotype regulation of P transposable elements in Drosophila melanogaster: repressor polypeptides or piRNAs? *Genetics*, *179*(4), 1785–93. http://doi.org/10.1534/genetics.108.087072

Kapitonov, V. V, & Jurka, J. (2003). Molecular paleontology of transposable elements in the Drosophila melanogaster genome. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(11), 6569–74. http://doi.org/10.1073/pnas.0732024100

Khurana, J. S., Wang, J., Xu, J., Koppetsch, B. S., Thomson, T. C., Nowosielska, A., … Theurkauf, W. E. (2011a). Adaptation to P element transposon invasion in Drosophila melanogaster. *Cell*, *147*(7), 1551–63. http://doi.org/10.1016/j.cell.2011.11.042

Khurana, J. S., Wang, J., Xu, J., Koppetsch, B. S., Thomson, T. C., Nowosielska, A., … Theurkauf, W. E. (2011b). Adaptation to P element transposon invasion in Drosophila melanogaster. *Cell*, *147*(7), 1551–63. http://doi.org/10.1016/j.cell.2011.11.042

Kleckner, N., Chalmers, R. M., Kwon, D., Sakai, J., & Bolland, S. (1996). Tn10 and IS10 transposition and chromosome rearrangements: mechanism and regulation in vivo and in vitro. *Current Topics in Microbiology and Immunology*, *204*, 49–82.

Livak, K. J. (1990). Detailed structure of the Drosophila melanogaster stellate genes and their transcripts. *Genetics*, *124*(2), 303–16.

Malone, C. D., & Hannon, G. J. (2009). Molecular evolution of piRNA and transposon control pathways in Drosophila. *Cold Spring Harbor Symposia on Quantitative Biology*, *74*, 225–34. http://doi.org/10.1101/sqb.2009.74.052

Morgan, M., Anders, S., Lawrence, M., Aboyoun, P., Pagès, H., & Gentleman, R. (2009). ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. *Bioinformatics (Oxford, England)*, *25*(19), 2607–8. http://doi.org/10.1093/bioinformatics/btp450

Moshkovich, N., & Lei, E. P. (2010). HP1 recruitment in the absence of argonaute proteins in Drosophila. *PLoS Genetics*, *6*(3), e1000880. http://doi.org/10.1371/journal.pgen.1000880

Pyatkov, K. I., Shostak, N. G., Zelentsova, E. S., Lyozin, G. T., Melekhin, M. I., Finnegan, D. J., … Evgen'ev, M. B. (2002). Penelope retroelements from Drosophila virilis are active after transformation of Drosophila melanogaster. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(25), 16150–5. http://doi.org/10.1073/pnas.252641799

Rajasethupathy, P., Antonov, I., Sheridan, R., Frey, S., Sander, C., Tuschl, T., & Kandel, E. R. (2012). A role for neuronal piRNAs in the epigenetic control of memory-related synaptic plasticity. *Cell*, *149*(3), 693–707. http://doi.org/10.1016/j.cell.2012.02.057

Rangan, P., Malone, C. D., Navarro, C., Newbold, S. P., Hayes, P. S., Sachidanandam, R., … Lehmann, R. (2011). piRNA production requires heterochromatin formation in Drosophila. *Current Biology : CB*, *21*(16), 1373–9. http://doi.org/10.1016/j.cub.2011.06.057

Rouget, C., Papin, C., Boureux, A., Meunier, A.-C., Franco, B., Robine, N., … Simonelig, M. (2010). Maternal mRNA deadenylation and decay by the piRNA pathway in the early Drosophila embryo. *Nature*, *467*(7319), 1128–32. http://doi.org/10.1038/nature09465

Rozhkov, N. V, Aravin, A. A., Zelentsova, E. S., Schostak, N. G., Sachidanandam, R., McCombie, W. R., … Evgen'ev, M. B. (2010). Small RNA-based silencing strategies for transposons in the process of invading Drosophila species. *RNA (New York, N.Y.)*, *16*(8), 1634–45. http://doi.org/10.1261/rna.2217810

Rozhkov, N. V, Schostak, N. G., Zelentsova, E. S., Yushenova, I. A., Zatsepina, O. G., & Evgen'ev, M. B. (2013). Evolution and dynamics of small RNA response to a retroelement invasion in Drosophila. *Molecular Biology and Evolution*, *30*(2), 397–408. http://doi.org/10.1093/molbev/mss241

Simmons, M. J., Meeks, M. W., Jessen, E., Becker, J. R., Buschette, J. T., & Thorp, M. W. (2014a). Genetic interactions between P elements involved in piRNA-mediated repression of hybrid dysgenesis in Drosophila melanogaster. *G3 (Bethesda, Md.)*, *4*(8), 1417–27. http://doi.org/10.1534/g3.114.011221

Simmons, M. J., Meeks, M. W., Jessen, E., Becker, J. R., Buschette, J. T., & Thorp, M. W. (2014b). Genetic interactions between P elements involved in piRNA-mediated repression of hybrid dysgenesis in Drosophila melanogaster. *G3 (Bethesda, Md.)*, *4*(8), 1417–27. http://doi.org/10.1534/g3.114.011221

Siomi, M. C., Sato, K., Pezic, D., & Aravin, A. A. (2011a). PIWI-interacting small RNAs: the vanguard of genome defence. *Nature Reviews. Molecular Cell Biology*, *12*(4), 246–58. http://doi.org/10.1038/nrm3089

Siomi, M. C., Sato, K., Pezic, D., & Aravin, A. A. (2011b). PIWI-interacting small RNAs: the vanguard of genome defence. *Nature Reviews. Molecular Cell Biology*, *12*(4), 246–58. http://doi.org/10.1038/nrm3089

Soppe, W. J. J., Jasencakova, Z., Houben, A., Kakutani, T., Meister, A., Huang, M. S., … Fransz, P. F. (2002). DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in Arabidopsis. *The EMBO Journal*, *21*(23), 6549–59.

Temme, C., Simonelig, M., & Wahle, E. (2014). Deadenylation of mRNA by the CCR4-NOT complex in Drosophila: molecular and developmental aspects. *Frontiers in Genetics*, *5*, 143. http://doi.org/10.3389/fgene.2014.00143

Table 5.1 Transformation rate for the autonomous *Muta1* element in *D. melanogaster*. The table describes number of embryos injected, survived injection from the injected embryos, total number of crosses established, total number of fertile crosses from all the crosses that were established and total number transgenic lines obtained from these fertile crosses. The transformation rate was calculated by dividing total number of transgenic lines obtained from total number of fertile crosses.

Plasmid injected	Total embryos injected	Survived Inj.	Total Crosses	Fertile Crosses	Transgenic with Muta1	Transformation Rate
pMuta1Auto	560	203	113	100	16	16%

Table 5.2 Molecular characterization of the integration locations of parental lines in *D. melanogaster*. The transgenic lines obtained for the autonomous *Muta1* element was characterized by inverse PCR method. The table describes the integration location, the break point of the integration in the genome, the TSDs sequence it contained, and the gene.

Lines	Location	Break point	TSD	
pMuta1Auto line A19	Chr 3L	12440359	GACCTAGG (8 bp)	Intergenic between Egy and Ego gene
pMuta1Auto line A50	Flanking plasmid DNA	--	-	
pMuta1Auto line A56	Flanking plasmid DNA	--	--	
pMuta1Auto line A85	Flanking plasmid DNA	--	--	

Table 5.3 New phenotypes observed in the autonomous *Muta1* transformation experiment. Table shows number of flies screened in each cross and new phenotypes observed in each cross.

Muta1 lines	No. of screened flies	New phenotypes
A19	3570	-
A50	2589	-
A56	2548	Full Body, Thorax, Full body with a dot on side in abdomen
A85	2389	Full body
Total	11,096	**Three new phenotypes**

Table 5.4 Molecular characterizations of the new transposition events for the autonomous *Muta1* element. The table describes the integration location, number of flies characterized, the break point of the integration in the genome, the TSDs sequence, and the gene.

New integration locations	Number of flies characterized	Break point	TSDs	Gene
Chromosome 2R	15	4033608	CTTGTTGCC	Intergenic between chromosome Lpin/CG8709 and Kermit/CG11546
Chromosome 3R	1	1447411	TCTGTCCTA	Repeat region

Table 5.5 Number of small RNA sequence obtained from each library and relative number of piRNAs that mapped to the autonomous *Muta1* element. Table shows the normalized counts in each library. Normalization is done based on the size of each library.

Library	Index	Number of small RNA sequences	Raw counts to Auto *Muta1* plasmid	Normalized count
A19 G11	ATCACG	8716353	98	150.7
A19 G21	ACAGTG	14708257	257	233.6
A50 G11	CGATGT	2939490	174	790
A50 G21	GCCAAT	11407368	198	230.2
A56 G11	TTAGGC	8805899	220	333.3
A56 G21	TAGCTT	9181645	87	126
A85 G11	TGACCA	41790391	1066	339.5
A85 G21	GATCAG	8910791	119	177.6
RNAlib 14-control *D. melanogaster*	---	13389112	12	---

Table 5.6 Normalized counts for the piRNAs mapped to the Hsp70 promoter, *Muta1* transposase, Hsp70 polyA region and the DsRed promoter from eight piRNAs libraries from the autonomous *Muta1* transgenic lines. Counts were normalized based on the size of each library.

Library	Region	Normalized counts	Antisense	Sense
A19 G11	Hsp70 promoter	58.46	56.92 (97.3%)	1.54 (2.7%)
	Muta1 transposase	9.23	0	9.23 (100%)
	Hsp70 polyA region	21.53	18.46 (85.7%)	3.07 (4.3%)
	DsRed gene	3	0	3 (100%)
A19 G21	Hsp70 promoter	93.6	86.4 (92.3%)	7.27 (7.7%)
	Muta1 transposase	24.5	7.27 (29.6%)	17.27 (70.4%)
	Hsp70 polyA region	58.2	42.7 (73.4%)	15.45 (26.6%)
	DsRed gene	6.3	0	6.3 (100%)
A50 G11	Hsp70 promoter	9.09	9.09 (100%)	0
	Muta1 transposase	36.36	4.54 (12.5%)	31.8 (87.45%)
	Hsp70 polyA region	4.54	4.54 (100%)	0
	DsRed gene	0	0	0
A50 G21	Hsp70 promoter	81.39	80.23 (98.6%)	1.16 (1.42%)
	Muta1 transposase	38.37	23.25 (60.6%)	15.11 (39.4%)
	Hsp70 polyA region	59.3	48.8 (82.3%)	10.46 (17.6%)
	DsRed gene	9.3	0	9.3 (100%)
A56 G11	Hsp70 promoter	109	106 (97.2%)	3 (2.8%)
	Muta1 transposase	116.7	7.57 (6.5%)	109.09 (93.5%)
	Hsp70 polyA region	34.8	13.6 (39%)	21.2 (61%)
	DsRed gene	37.8	0	37.8 (100%)

A56 G21	Hsp70 promoter	52.17	44.9 (86%)	7.24 (14%)
	Muta1 transposase	23.18	15.9 (68.6%)	7.24 (31.4%)
	Hsp70 polyA region	0	0	0
	DsRed gene	4.34	0	4.34 (100%)
A85 G11	Hsp70 promoter	70.38	60.82 (86.4%)	9.55 (13.6%)
	Muta1 transposase	54.45	6.05 (11.1%)	48.4 (88.9%)
	Hsp70 polyA region	50	26.75 (53.5%)	23.24 (46.48%)
	DsRed gene	35.9	0	35.9 (100%)
A85 G21	Hsp70 promoter	75.38	58.46 (77.6%)	16.9 (22.4%)
	Muta1 transposase	40	10.76 (26.9%)	29.23 (73.1%)
	Hsp70 polyA region	30.76	23.07 (75%)	10.76 (25%)
	DsRed gene	12.3	0	12.3 (100%)

Table 5.7 Frequency of U1 (uracil) in antisense and A10 (adenine) in sense piRNAs from the eight libraries. The table describes number of antisense piRNAs with U1 and number of sense piRNAs with A10 from individual libraries.

Lines	U1 in antisense piRNAs	A10 in sense piRNAs
A19-G11	11/69 (15.9%)	3/29 (10.4%)
A19-G21	31/178 (17.4%)	28/79 (35%)
A50-G11	25/87 (28.7%)	6/87 (6.9%)
A50-G21	29/144 (20.1%)	14/54 (29.9%)
A56-G11	17/95 (17.9%)	34/125 (27.2%)
A56-G21	14/53 (26.4%)	6/34 (17.6%)
A85-G11	106/482 (21.99%)	90/584 (15.4%)
A85-G21	14/66 (21.2%)	15/53 (28.3%)

Table 5.8 U1-A10 overlaps for the piRNA libraries from the autonomous *Muta1* transgenic lines. The piRNAs were examined for a 10 nucleotides overlap between antisense piRNAs with U1 and sense piRNAs with A10. The table describes the transgenic lines with overlaps, location of the overlaps based on the plasmid that was used for experiment, length of the antisense piRNAs, sequence of the antisense piRNAs, length of the sense piRNAs, and the sequence of the sense piRNAs.

Library	Sense RNA location (A10)	length	Sequence	Antisense RNA loc (U1)	length	Sequence
A19-G21	521 (Hsp70)	24	AACAAGCAAAGTGAACACGTCGCT	530	26	TTTGCTTGTTTGAATTGAATTGTCGT
A50-G21	472 (Hsp70)	23	CCGGAGTATAAATAGAGGCGCTT	482	23	TTATACTCCGGCGCTCTTTTCGC
	584 (Hsp70)	23	GCTAAACAATCTGCAGACTAGTA	594	26	GATTGTTTAGCTTGTTCAGCTGCGCC
	1657 (Muta1 transposase)	22	CCGCGTTTGAGAATTTGAAAAA	1666	26	TCAAACGCGGCTGGTATACGTTCAGT
	2462 (Hsp70 polyA)	23	CCGGAGTATAAATAGAGGCGCTT	2472	27	TTATACTCCGGCGGTCGAGGGTTCGAA
	2576 (Hsp70 polyA)	23	TAAACAATCGGGGTACCGCTAGA	2586	25	CAGATTGTTTAGCTTGTTCAGCTGC
	2580 (Hsp70 polyA)	27	CAATCGGGGTACCGCTAGAGTCGACG G	2590	26	TAACCCGATTGTTTAGCTTGTTCAGC
	2706 (DsRed gene)	25	TTCGAGATCGAGGGCGAGGGCGAGG	2715	21	CGATCTCGAACTCGTGGCCGC
A56-G11	864 (Muta1 transposase)	24	TGTTCGGACCATAACCACGAAAGC	875	28	ATGGTCCGAACATGACAGAACGCGGT GT
	2348 (Hsp70 polyA)	24	TTTGCTTAAAAACTCGTTTAGATC	2359	23	TTTTTAAGCAAACTCACTCCCTG
A56-G21	577 (Hsp70)	24	TGAACAAGCTAAACAATCTGCAGT	589	24	TTTAGCTTGTTCAGCTGCGCTTGT
	589 (Hsp70)	21	ACAATCTGCAGACTAGTATGG	599	28	TTGCAGATTGTTTAGCTTGTTCAGCTG C
	2524 (Hsp70 polyA)	22	AACACGTCGCTAAGCGAAAGCT	2535	29	TAGCGACGTGTTCACTTTGCTTGTTTG AA
A85-G11	2576 (Hsp70 polyA)	25	TAAACAATCGGGGTACCGCTAGAGT	2586	25	CAGATTGTTTAGCTTGTTCAGCTGC

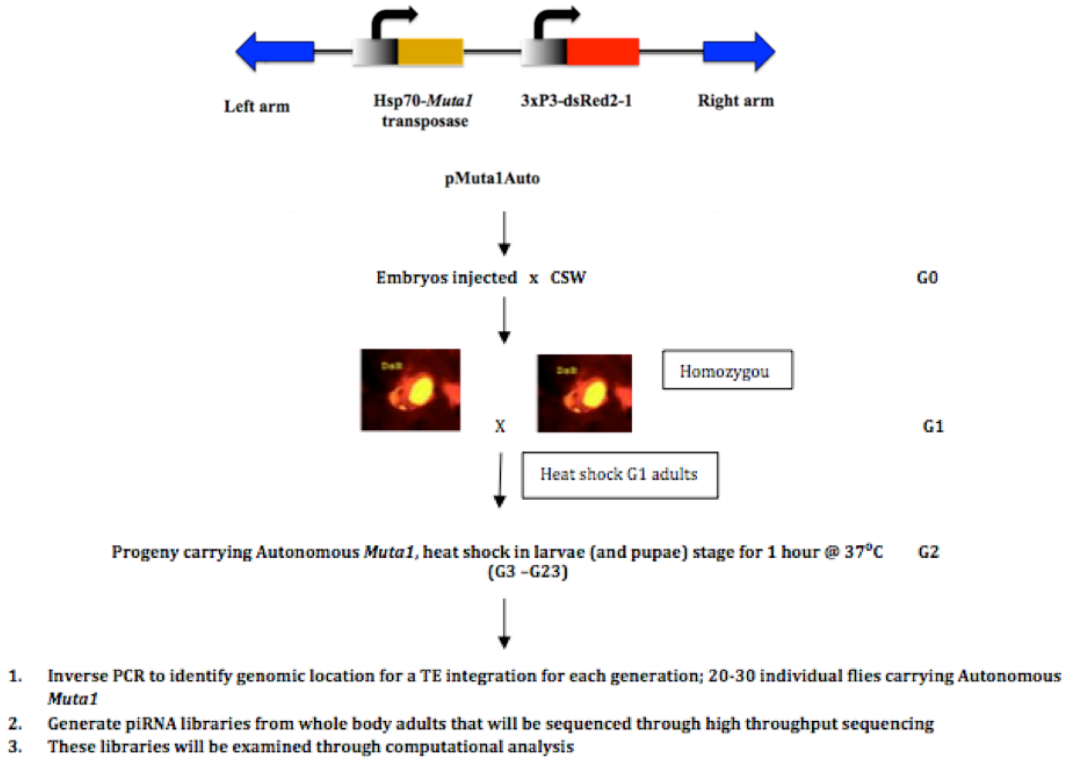
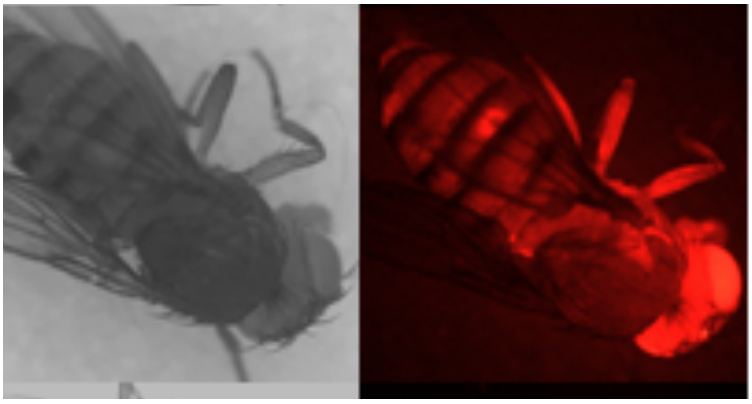
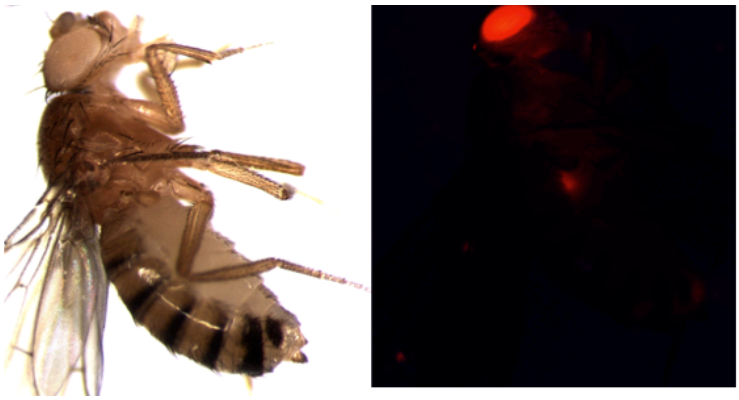


Figure 5.1 Schematic diagram of the autonomous *Muta1* experiment. The flow chart represents the map of the plasmid that was used for the transformation experiments, establishing the homozygous lines, screening of the progeny for new phenotypes, molecular characterization of the new integration location for the autonomous *Muta1* element, and small RNAs library preparation.

A



B



C

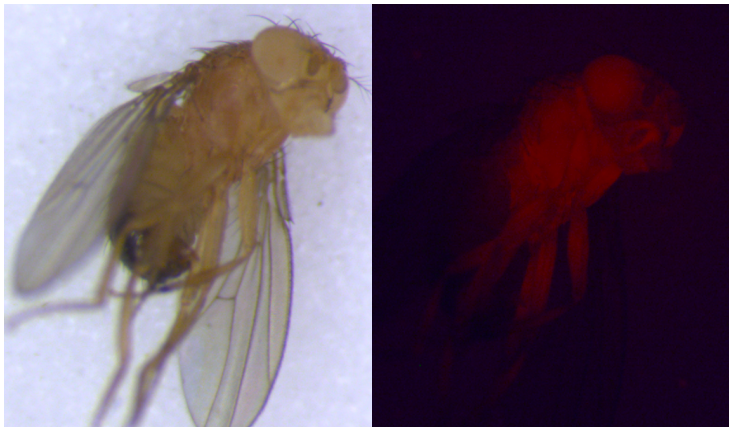


Figure 5.2 New phenotypes observed in the crosses, pictures were taken without filter and with DsRed2-1 filter. A. Adult with DsRed expression in full body, B. DsRed expression in the abdomen, C. DsRed expression in thorax

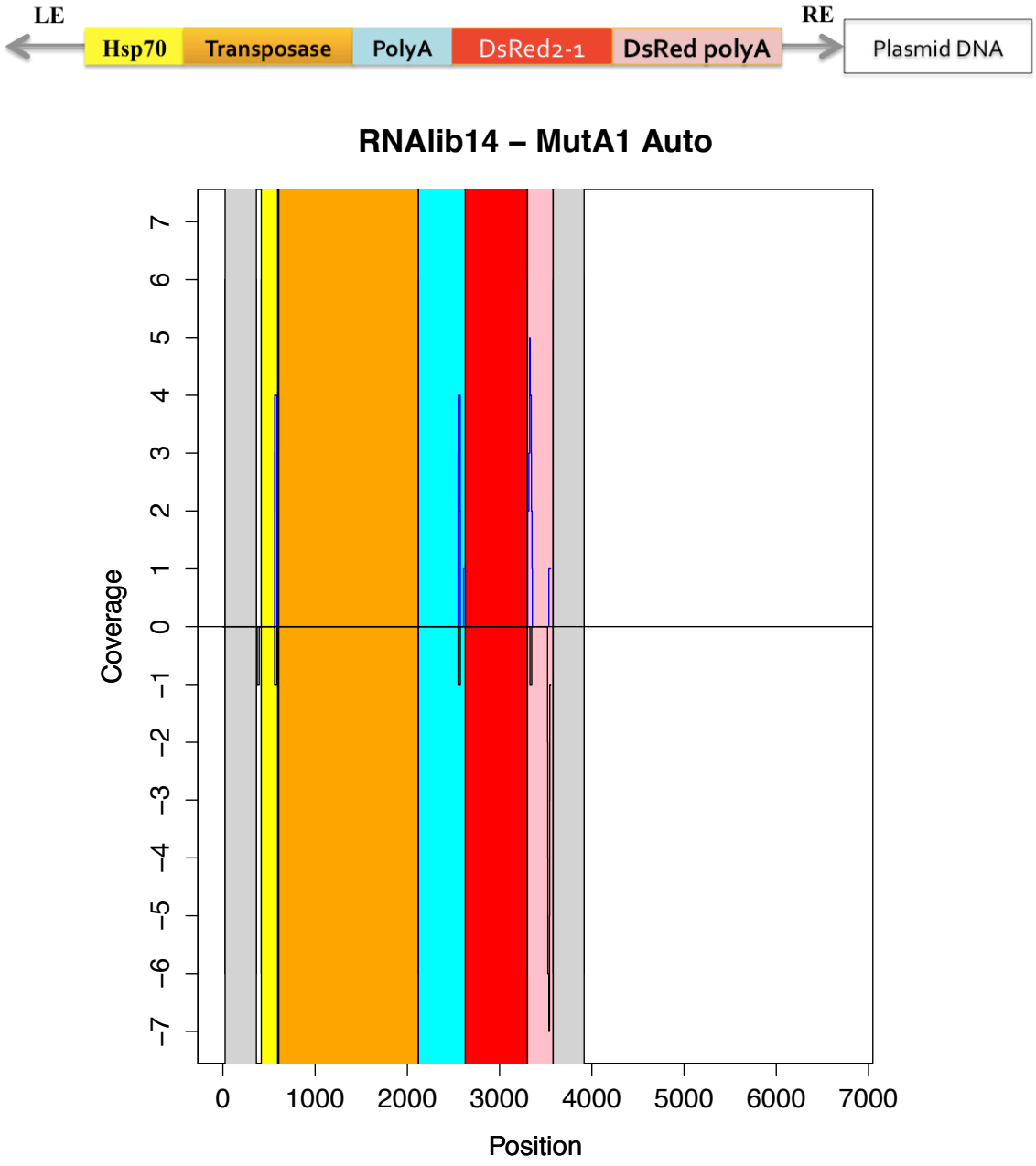


Figure 5.3.1 Coverage plots representing control library RNAlib14 from *D. melanogaster* mapped to autonomous *Muta1* element. The autonomous *Muta1* plasmid contained *Muta1* left end (grey), the Hsp70 promoter (yellow), *Muta1* transposase (orange), Hsp70 polyA (blue), 3xP3-DsRed gene (red), DsRed polyA region (pink), *Muta1* right end (grey), and the plasmid backbone (white). The plots indicate no hits to the *Muta1* transposase or to the ends. Thus, *D. melanogaster* is naïve to *Muta1 element.*

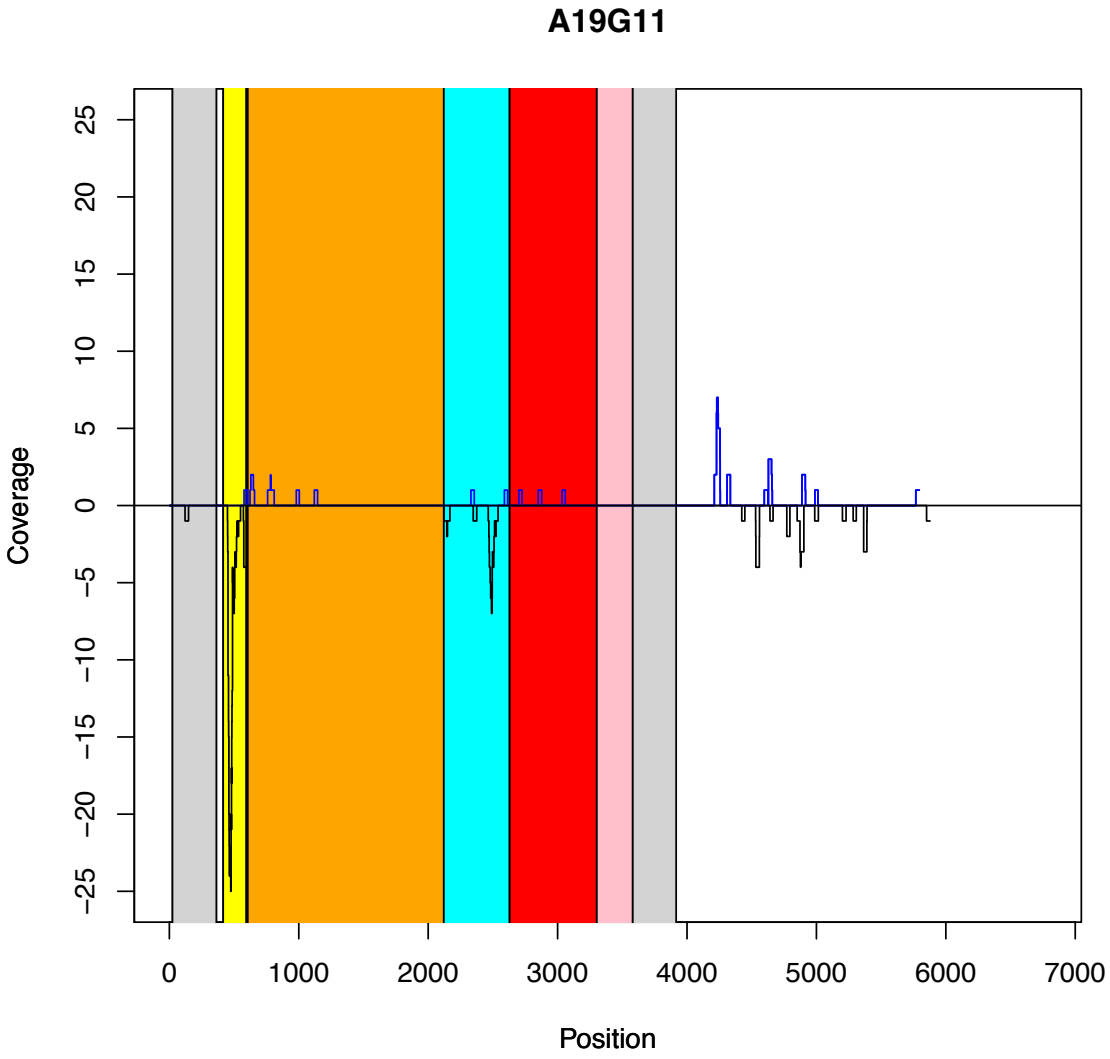


Figure 5.3.2 Coverage plot for library A19 G11, piRNAs mapped to autonomous *Muta1* plasmid. The autonomous *Muta1* plasmid contained *Muta1* left end (grey), the Hsp70 promoter (yellow), *Muta1* transposase (orange), Hsp70 polyA (blue), 3xP3-DsRed gene (red), DsRed polyA region (pink), *Muta1* right end (grey), and the plasmid backbone (white).

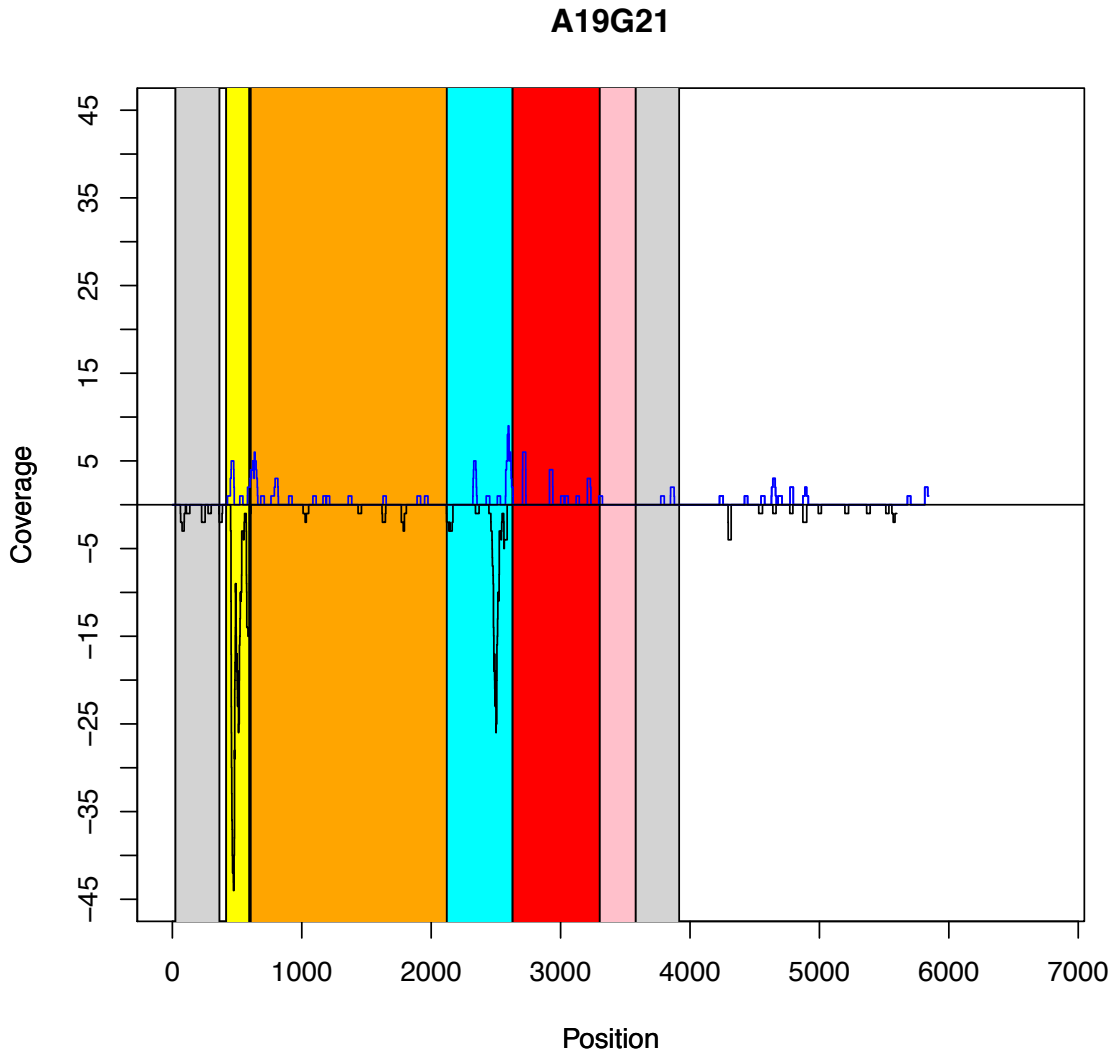


Figure 5.3.3 Coverage plot for library A19 G21, piRNAs mapped to autonomous *Muta1* plasmid**.** The autonomous *Muta1* plasmid contained *Muta1* left end (grey), the Hsp70 promoter (yellow), *Muta1* transposase (orange), Hsp70 polyA (blue), 3xP3-DsRed gene (red), DsRed polyA region (pink), *Muta1* right end (grey), and the plasmid backbone (white).

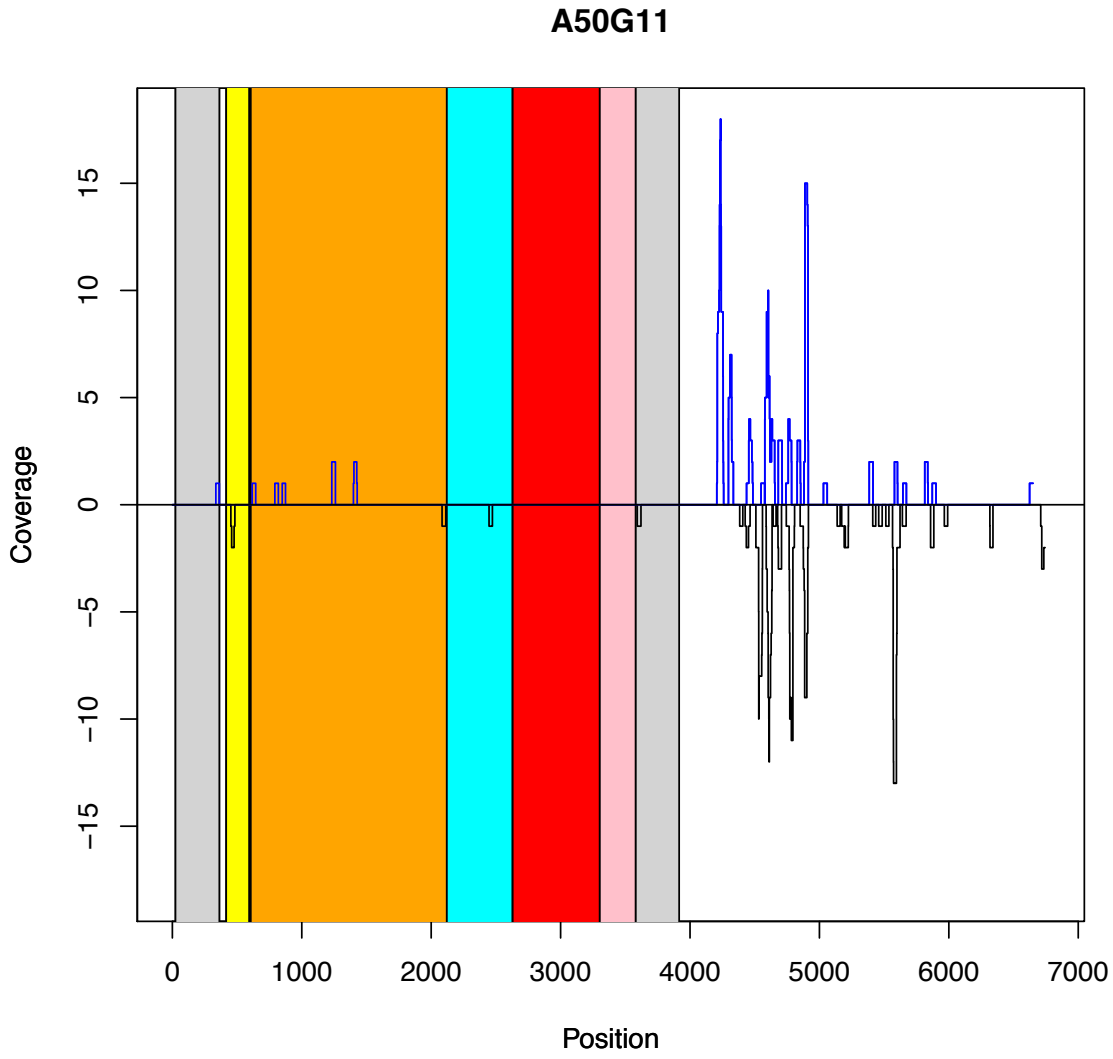


Figure 5.3.4 Coverage plot for library A50 G11. piRNAs mapped to autonomous *Muta1* plasmid. The autonomous *Muta1* plasmid contained *Muta1* left end (grey), the Hsp70 promoter (yellow), *Muta1* transposase (orange), Hsp70 polyA (blue), 3xP3-DsRed gene (red), DsRed polyA region (pink), *Muta1* right end (grey), and the plasmid backbone (white).

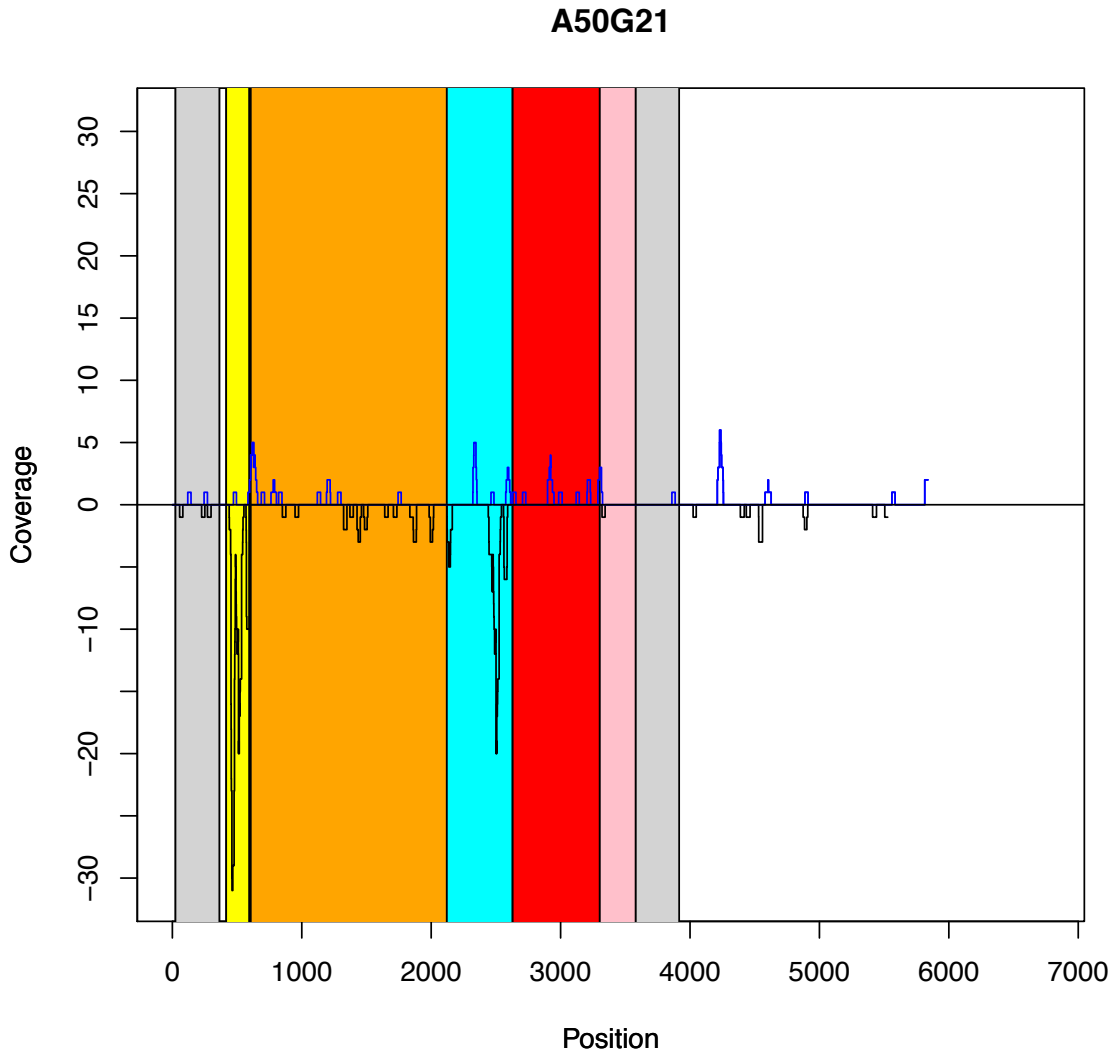


Figure 5.3.5 Coverage plot for library A50 G21, piRNAs mapped to autonomous *Muta1* plasmid. The autonomous *Muta1* plasmid contained *Muta1* left end (grey), the Hsp70 promoter (yellow), *Muta1* transposase (orange), Hsp70 polyA (blue), 3xP3-DsRed gene (red), DsRed polyA region (pink), *Muta1* right end (grey), and the plasmid backbone (white).

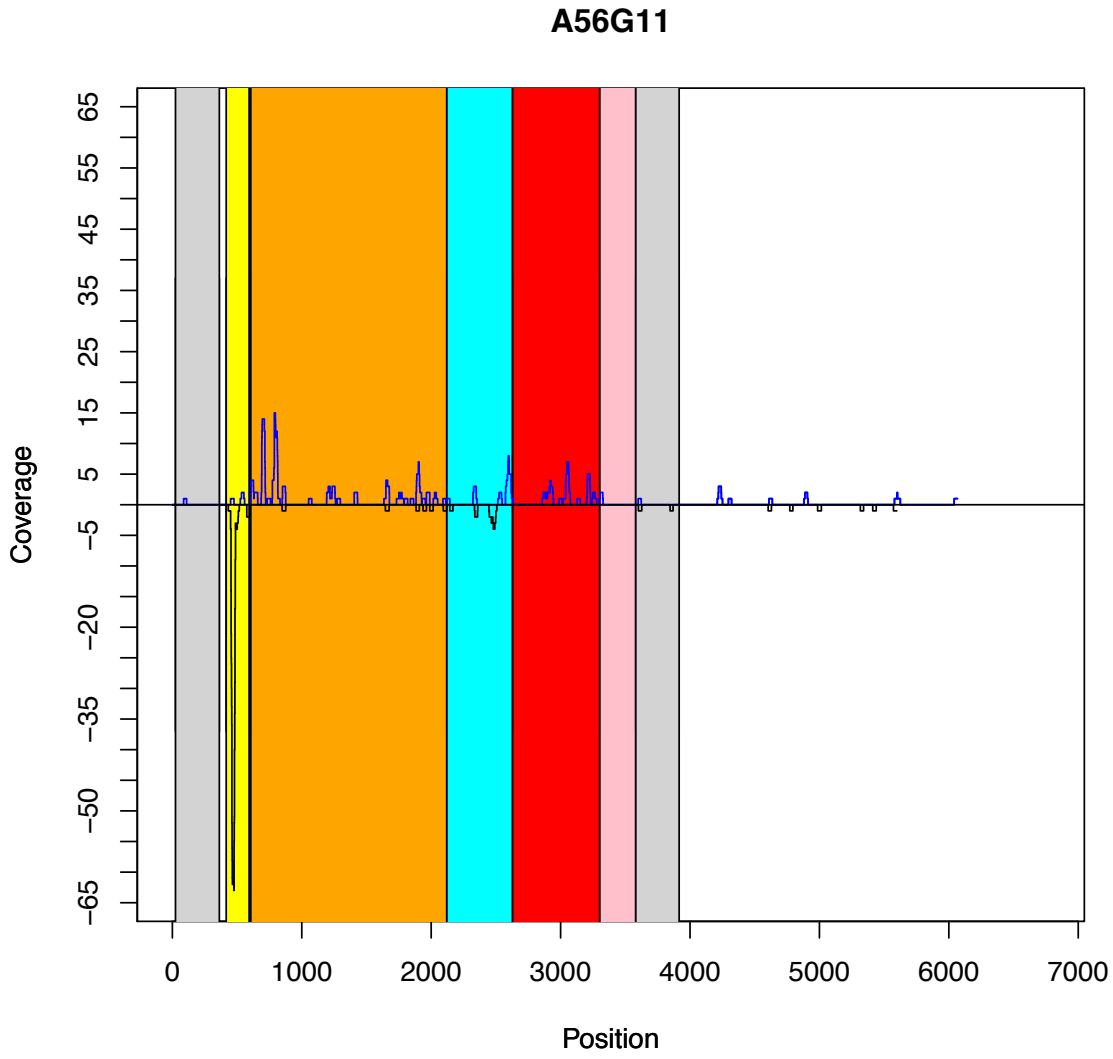


Figure 5.3.6 Coverage plot for library A56 G11, piRNAs mapped to autonomous *Muta1* plasmid. The autonomous *Muta1* plasmid contained *Muta1* left end (grey), the Hsp70 promoter (yellow), *Muta1* transposase (orange), Hsp70 polyA (blue), 3xP3-DsRed gene (red), DsRed polyA region (pink), *Muta1* right end (grey), and the plasmid backbone (white).

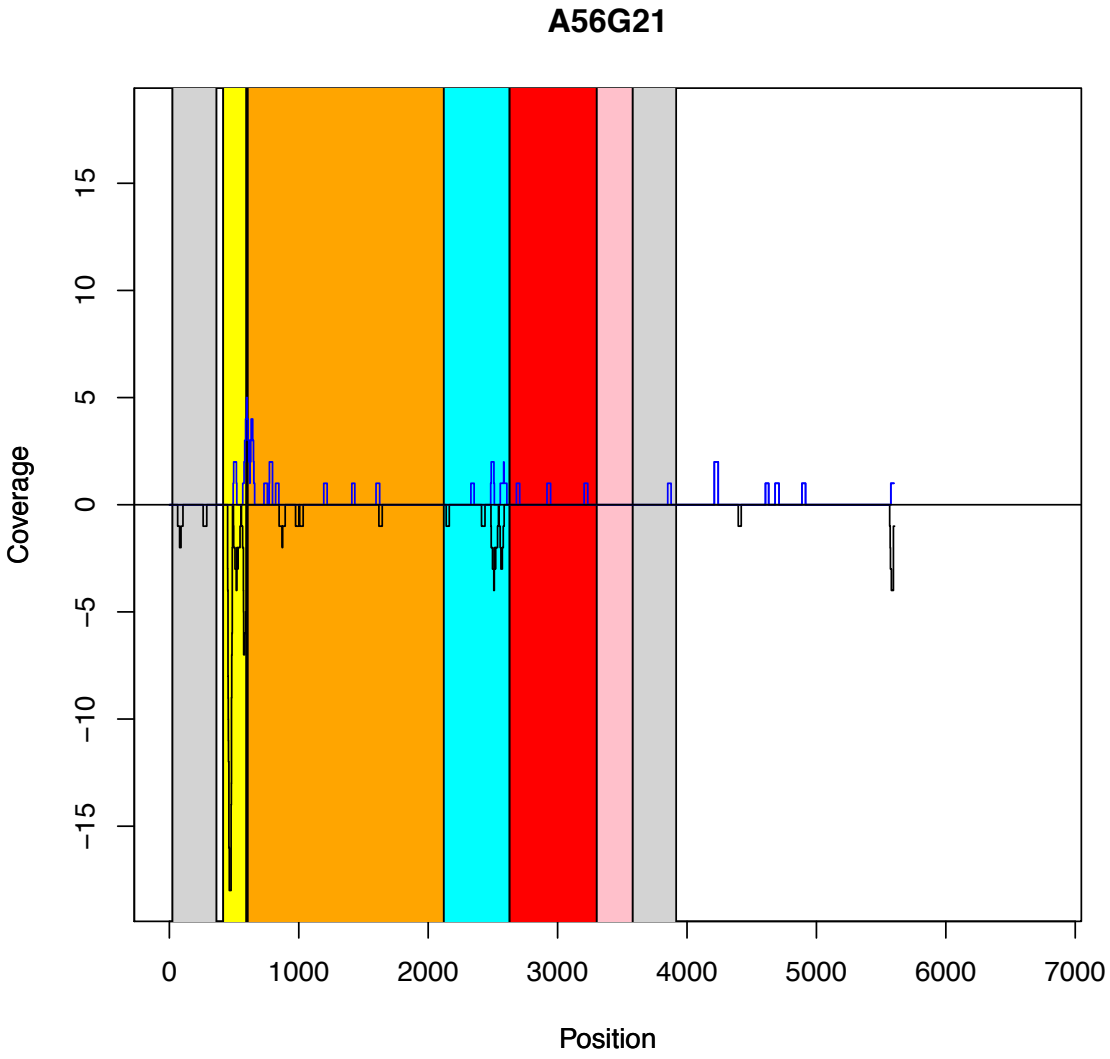


Figure 5.3.7 Coverage plot for library A56 G21, piRNAs mapped to autonomous *Muta1* plasmid. The autonomous *Muta1* plasmid contained *Muta1* left end (grey), the Hsp70 promoter (yellow), *Muta1* transposase (orange), Hsp70 polyA (blue), 3xP3-DsRed gene (red), DsRed polyA region (pink), *Muta1* right end (grey), and the plasmid backbone (white).

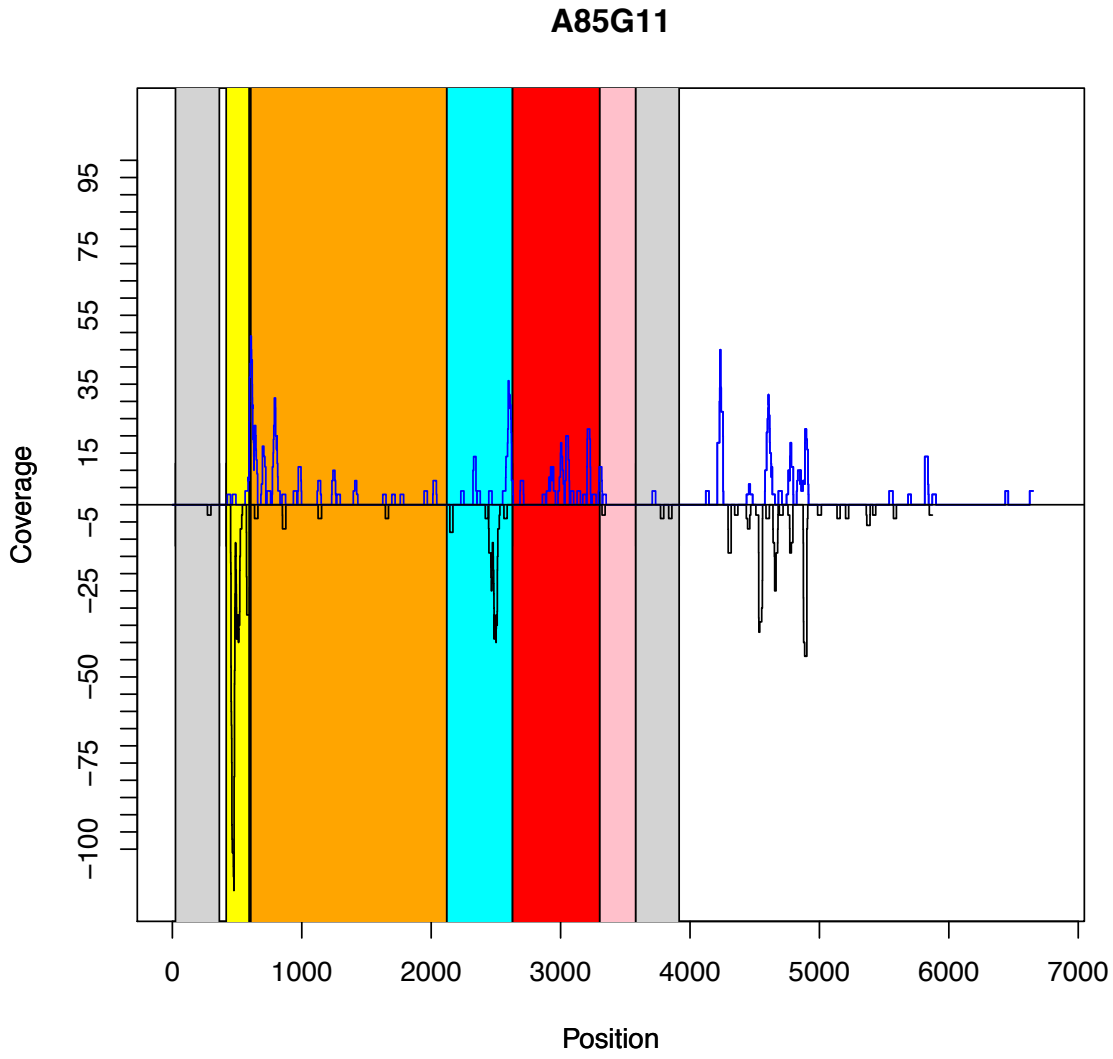


Figure 5.3.8 Coverage plot for library A85 G11, piRNAs mapped to autonomous *Muta1* plasmid. The autonomous *Muta1* plasmid contained *Muta1* left end (grey), the Hsp70 promoter (yellow), *Muta1* transposase (orange), Hsp70 polyA (blue), 3xP3-DsRed gene (red), DsRed polyA region (pink), *Muta1* right end (grey), and the plasmid backbone (white).

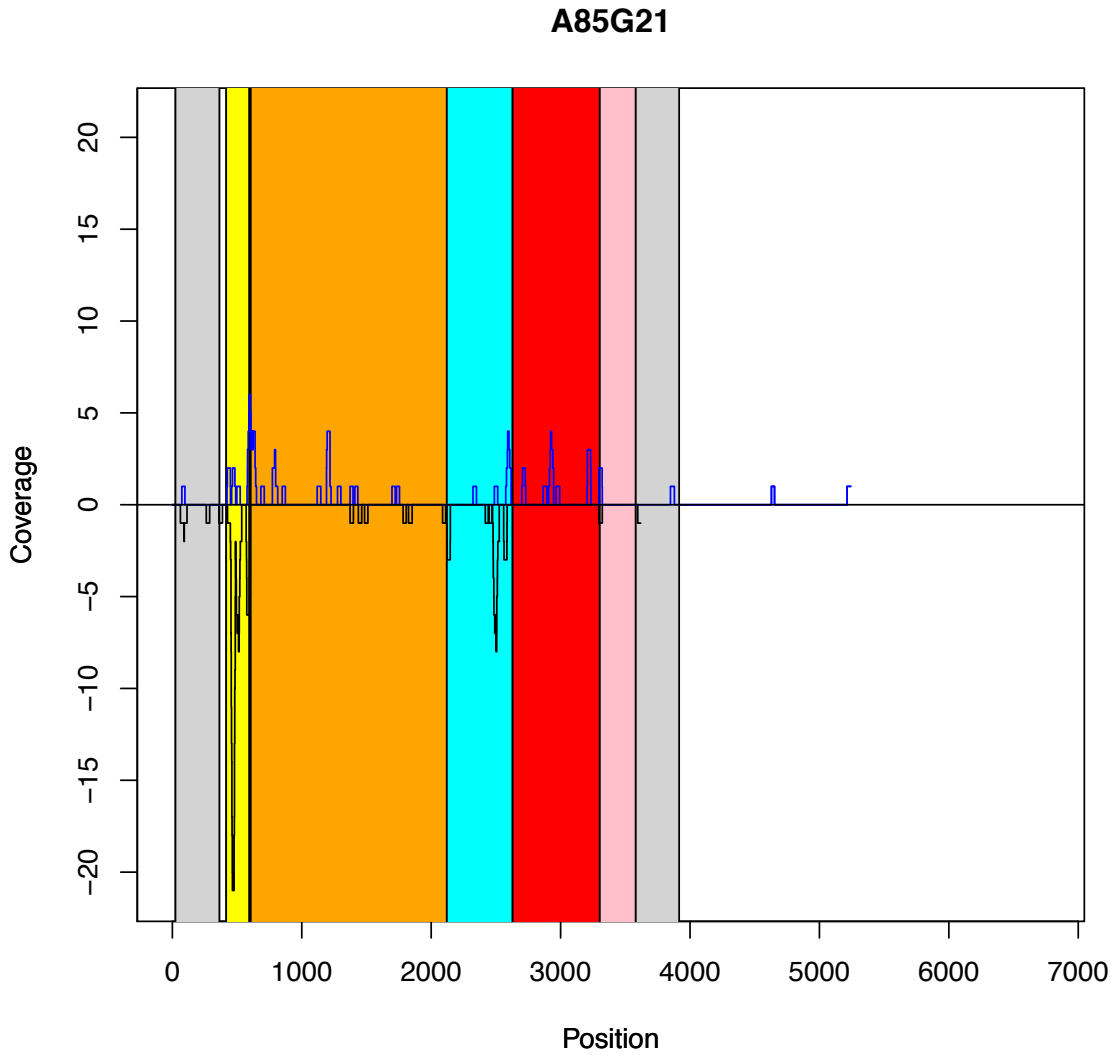


Figure 5.3.9 Coverage plot for library A85 G21, piRNAs mapped to autonomous *Muta1* plasmid. The autonomous *Muta1* plasmid contained *Muta1* left end (grey), the Hsp70 promoter (yellow), *Muta1* transposase (orange), Hsp70 polyA (blue), 3xP3-DsRed gene (red), DsRed polyA region (pink), *Muta1* right end (grey), and the plasmid backbone (white).

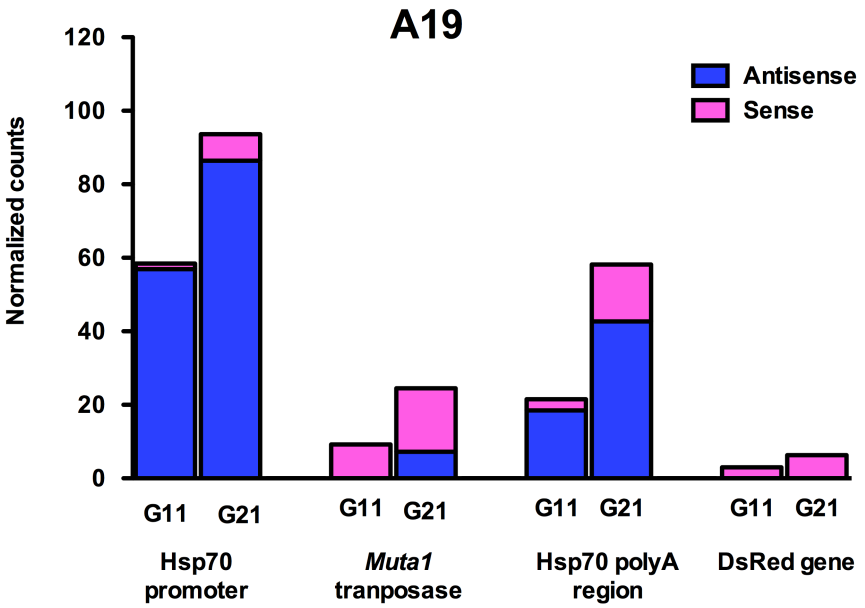


Figure 5.4 Graphs comparing generation 11 and 21 for transgenic line A19, for piRNAs mapped to the Hsp70 promoter, *Muta1* transposase, Hsp70 polyA region, and DsRed gene.

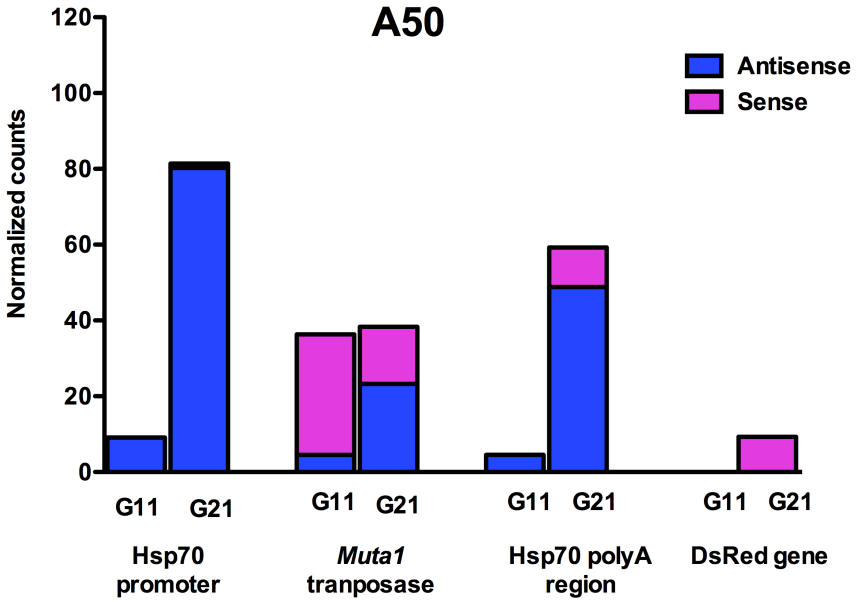


Figure 5.5 Graphs comparing generation 11 and 21 for transgenic line A50, for piRNAs mapped to the Hsp70 promoter, *Muta1* transposase, Hsp70 polyA region, and DsRed gene.

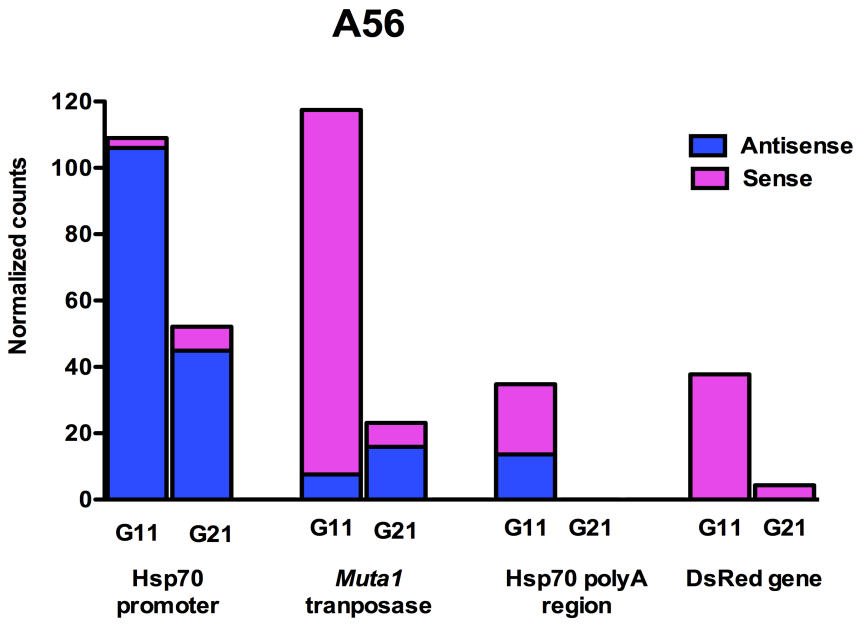


Figure 5.6 Graphs comparing generation 11 and 21 for transgenic line A56, for piRNAs mapped to the Hsp70 promoter, *Muta1* transposase, Hsp70 polyA region, and DsRed gene.

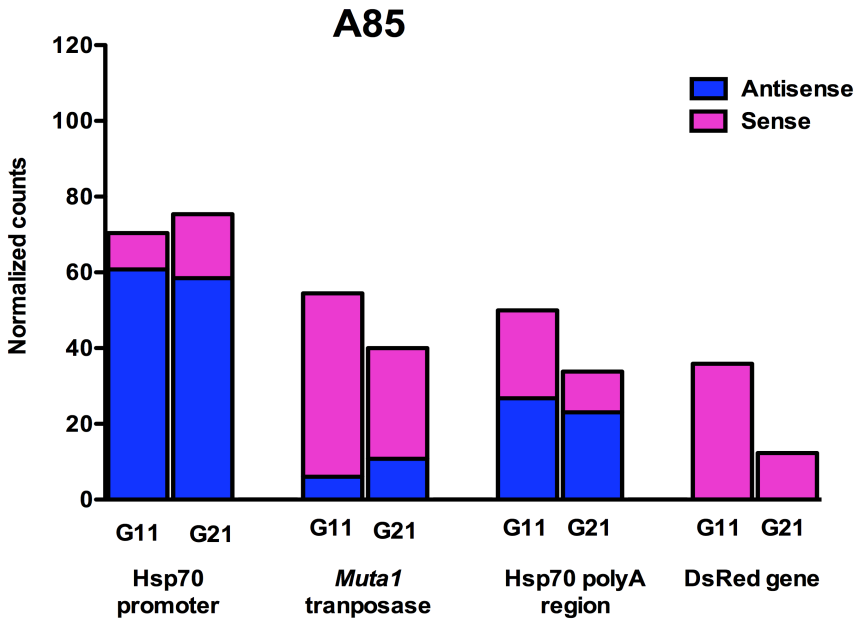


Figure 5.7 Graphs comparing generation 11 and 21 for transgenic line A85, for piRNAs mapped to the Hsp70 promoter, *Muta1* transposase, Hsp70 polyA region, and DsRed gene.

Script 5.1 R script for removal of contaminants from the small RNA library.

```
#####################################
##    Read Preprocessing   #####
#####################################
setwd("/shared/gen240b/VARseq1/")
source("analysis_Fct.R")
library(ShortRead)
fq = FastqStreamer("NAME_OF_FILE.fq")
apply2FQ(fqfiles=fq, myfct=trimReads, batchsize=100000, silent=FALSE, quality=20, Ns=3, polyn=20, polyntype=c("A", "C", "T", "G"), minwidth=32)
### Run Bowtie 2 with default parameters against SILVA rRNA Database
# bowtie2 --no-hd --no-unal --no-sq -p 8 -q -x rRNA_REFERENCE_SEQS.fa -U NAME_OF_FILE.fq.trim -S MAPPED_FILE.sam
### Parse out only read_id (col 1)
# awk '{print $1}' MAPPED_FILE.sam > rRNA_reads.xls
rr = "rRNA_reads.xls"
while(length(rr<-yield(f)) ) { writeFastq(rr[ !(id(rr) %in% rRNA_mapped) ], "RNAlibXX_riboremoved.fastq.trim", mode="a") }
close(f)
rm(f)
#################################
########## Functions ##########
#################################
trimReads <- function(reads, quality=20, Ns=3, polyn=20, polyntype=c("A", "C", "T", "G"), minwidth=32, ...) {
### Trim off low quality tails
if(is.numeric(quality) & length(unique(width(reads))) == 1) { # Note: works only if all reads are of same length
## Inject Ns at low quality positions
qualityCutoff <- quality # remove read tails with quality lower than this
seqs <- sread(reads) # get sequence list
qual <- PhredQuality(quality(quality(reads))) # get quality score list as PhredQuality
myqual_mat <- matrix(charToRaw(as.character(unlist(qual))), nrow=length(qual), byrow=TRUE) # convert quality score to matrix
at <- myqual_mat < charToRaw(as.character(PhredQuality(as.integer(qualityCutoff)))) # find positions of low quality
letter_subject <- DNAString(paste(rep.int("N", width(seqs)[1]), collapse="")) # create a matrix of Ns
letter <- as(Views(letter_subject, start=1, end=rowSums(at)), "DNAStringSet") # trim to length needed for each read
injectedseqs <- replaceLetterAt(seqs, at, letter) # inject Ns at low quality positions
## Get coordinates of polyN tails
adapter <- paste(rep("N", max(width(injectedseqs))), sep="", collapse="")
```

```
mismatchVector <- c(rep(0,width(adapter))) # allow no mismatches at each adapter offset
trimCoords <- trimLRPatterns(Rpattern=adapter, subject=injectedseqs, max.Rmismatch=mismatchVector, ranges=T)
## Apply trimming coordinates from injected reads to non-injected reads
seqs <- DNAStringSet(seqs, start=start(trimCoords), end=end(trimCoords))
qual <- BStringSet(qual, start=start(trimCoords), end=end(trimCoords))
# Use IRanges coordinates to trim sequences and quality scores
qual <- SFastqQuality(qual) # reapply quality score type
reads <- ShortReadQ(sread=seqs, quality=qual, id=id(reads)) # Rebuild reads object
}
## Filter reads
## Maximum number of Ns
if(is.numeric(Ns)) {
filter1 <- nFilter(threshold=Ns) # keep only reads with fewer than 3 Ns
reads <- reads[filter1(reads)]
}
## Homopolymer filter
if(is.numeric(polyn)) {
filter2 <- polynFilter(threshold=polyn, nuc=polyntype) # remove reads with 20 or more of the same letter
reads <- reads[filter2(reads)]
}
## Minimum length filter
if(is.numeric(minwidth)) {
reads <- reads[width(reads) >= minwidth]
}
return(reads)
}
## Usage:
# trimReads(reads, quality=20, Ns=3, polyn=20, polyntype=c("A", "C", "T", "G"), minwidth=32)

## Run trimReads with FastqStreamer on many input files
apply2FQ <- function(fqfiles, myfct, batchsize, silent=FALSE, ...) {
for(i in fqfiles) {
f <- FastqStreamer(i, batchsize) # FastqStreamer usage: fq <- yield(f); sread(fq); quality(fq); id(fq)
while(length(fq <- yield(f))) {
reads <- myfct(reads=fq, ...)
if(length(reads)>0) {
writeFastq(reads, paste(i, ".trim", sep=""), mode="a")
}
if(silent==FALSE) {
cat("N reads from", i, "trimmed and written to file:", length(reads), "\n")
```

```
}
}
close(f)
}
}
```

Script 5.1 R scripts used to make coverage plots.

```
plotCoverage <-
function(x,xlab="Position",ylab="Coverage",xlim=xlim,ylim=ylim)
{
plot(c(start(x),length(x)),c(runValue(x),tail(runValue(x),1)),type="s",col="blue",xlim=xli
 m,ylim=ylim,xlab=xlab,ylab=ylab,yaxt='n')
axis(side=2,at=seq(-10,10,1))
}
plotStrandedCoverage <-
function(positive,negative,name,xlab="Position",ylab="Coverage")
{
xlim <- c(0,6776)
ylim <- max(max(positive),max(negative)) * c(-1,1)
plotCoverage(positive,xlim=xlim,ylim=ylim)
title(main=name,xlab="Position",ylab="Coverage")
rect(2628,-ylim-1,3302,ylim+1,col="red")  # DsRed2-1 marker
rect(3303,-ylim-1,3580,ylim+1,col="pink")# DsRed2-1 polyA
rect(3581,-ylim-1,3916,ylim+1,col="lightgray")  # right end
rect(24,-ylim-1,363,ylim+1,col="lightgray")  # left end
rect(416,-ylim-1,593,ylim+1,col="yellow")  # HSP70
rect(606,-ylim-1,2120,ylim+1,col="orange")  # MutA1 tpase
rect(2121,-ylim-1,2627,ylim+1,col="cyan")# MutA1_tpase polyA
#rect(z+2849,-ylim-1,z+2935,ylim+1,col="green")      # 3' UTR = green
lines(c(start(negative),length(negative)),
- c(runValue(negative),tail(runValue(negative),1)),
type="s",col="black")
lines(c(start(positive),length(positive)),
+ c(runValue(positive),tail(runValue(positive),1)),
type="s",col="blue")
abline(h=0,col="black")
}
```

Chapter 6

Summary and Conclusions

6.1 Summary

The work presented in the thesis demonstrates the following 1) Identification of endogenous active Class II DNA transposon *Muta1* from *Ae. aegypti* 2) somatic activity of *Muta1* in *D. melanogaster* and *Ae. aegypti* 3) germline activity of the *Muta1* in *D. melanogaster* and *Ae. aegypti* 4) the piRNA response to *Muta1* element in *D. melanogaster.*

6.2 Discovery of the *Muta1* element in *Ae. aegypti.*

Mosquito transgenesis has heavily relied on the availability of active DNA transposons that can be used for discovery and analysis of new genes for insect control. Very few transposons have been used for germline transformation of *Ae. aegypti* and all of these transposons have been discovered in other species. *Ae. aegypti* transformation with exogenous DNA transposons have failed to retain their activity such as post-integration mobility which is essential for a transposon to be used as a genetic tool (Wilson et al., 2003) (Smith & Atkinson, 2011) (Palavesam, Esnault, & O'Brochta, 2013). The identification of endogenous DNA transposons in the genome of *Ae. aegypti* can be very useful as these transposons might have been able to evade host silencing and could prove to be an effective tool for *Ae. aegypti* germline transformation.

The research in chapter two describes identification of an endogenous class II DNA transposon *Muta1*, which belongs to the *Mutator* superfamily. *Muta1* is the first member of the *Mutator* superfamily that has demonstrated activity in insects. The *Mutator* transposons have been shown to play a major role in genome evolution in plants. They are involved in the domestication of genes, increases in mutation rate, (Talbert & Chandler, 1988) and have also proved to be a valuable tool for discovery of gene function through insertional mutagenesis (Lisch, 2013). A *Mutator* superfamily transposon *Phantom* has been identified in two insect viruses, which indicates the role of viruses in horizontal transfer of transposon, and *Phantom*-like proteins have been identified in *Ae. aegypti* (Marquez & Pritham, 2010). Preliminary screening for transposons was performed using the MITE-Hunter pipeline, which searches for MITEs in the genome. MITEs have a very high copy number in the genome and can be employed to search for a related autonomous transposons, as transposase from an autonomous transposon is required for transposition of the MITEs in the genome (Han & Wessler, 2010) (Feschotte, Swamy, & Wessler, 2003). Consensus sequences were identified with this MITE-Hunter pipeline (Han & Wessler, 2010), and were then used as a template in the TARGeT pipeline (Han, Burnette, & Wessler, 2009) to search for endogenous transposons in *Ae. aegypti*. Eight copies of *Muta1* transposon were discovered in *Ae. aegypti.* These copies have greater than 99% similarity to each other. Subsequently, target site preference for the wild-type copies of *Muta1* was determined. The *Muta1* element prefers to integrate forming eight or nine base pair target site duplication upon integration. Of these, six out

of eight wild-type copies have 9 bp TSDs and two have 8 bp TSDs in the genome of *Ae. aegypti*.

The *Muta1* element is 3.2 kb in length with 146 bp ITRs. There are seven 11 bp repeats in the left TIR and six 11 bp repeats in the right TIR. The element encodes a single ORF with two exons and a 60 bp intron. The ORF of *Muta1* transposase has 504 amino acids. It contains FLY-WCH domain and a MULEs transposase domain. The FLY-WCH domain is a DNA binding domain (DBDs) classified under the WRKY-GCMI superfamily of DBDs. The WRKY-GCMI DBDs are common feature of some MULEs and plant *MuDR* transposase. (Babu, Iyer, Balaji, & Aravind, 2006). The size, structure and organization of the *Muta1* element and the TSDs were consistent with other *Mutator* superfamily transposons (Feschotte & Pritham, 2007). The structural variation observed in the ends of the *Muta1* element is not unique. Previous study have shown that repeats present in the ends of *Phantom* elements were related to *Foldback* elements like *Galileo* (Marquez & Pritham, 2010). The variation in the TIRs might be to facilitate increased binding specificity of the *Muta1* transposase, which might have a role in the transposition mechanism of the *Muta1* element. Previously, it has been proposed for *Phantom* elements that increasing the number of transposase biding sites within the TIRs might increase transposition frequency of a element (Marquez & Pritham, 2010). Thus, there is a possibility that the tandem repeats in the ends of *Muta1* element might increase transposition frequency.

6.3 Somatic activity of the *Muta1* element in *D. melanogaster* and *Ae. aegypti.*

Studying the transposition and excision mechanisms of a transposon is crucial to determining the behavior of the transposon in an organism. Somatic transposition assay were performed using the *lacZ* reporter and antibiotic selection to determine transposition rate of *Muta1* into a target plasmid, *pGDV1*. *Muta1* has high somatic activity in *D. melanogaster* and *Ae. aegypti*. Zero-control transposition assays performed in absence of *Muta1* transposase showed no integration of the *Muta1* element ends in *D. melanogaster* and *Ae. aegypti.* Thus, the *Muta1* element can be used as genetic tool where stability of the integrated transposon is required such as in human gene therapy and insect transgenesis (Yant et al., 2000) (Franz et al., 2011) (Malcolm J Fraser, 2012).

I determined transposition frequency for the *Muta1* element flanked by an 8 bp as well as a 9 bp TSDs. The transposition frequency obtained for *Muta1* with 8 bp TSD and 9 bp TSD were approximately 1.3 times higher in *D. melanogaster* compared to *piggyBac* element, which served as an internal control. Whereas, in *Ae. aegypti* transposition frequency with 8 bp TSD was approximately 16 times higher than *piggyBac* and approximately 13 times higher than *piggyBac* element with a 9 bp TSDs. Thus, the *Muta1* element has high somatic transposition frequency with an 8 or a 9 bp TSDs. In addition, the transposition events recovered showed preference for integration with 9 bp TSDs, which is also supported by other *Mutator* elements which prefer to integrates producing 9 bp TSD flanking to the site of integration (Marquez & Pritham, 2010).

Limited sequence bias is useful for a transposon used in a mutagenesis project. Most transposons show some sequence preference, which might limit their use for gene tagging, enhancer trapping and mutagenesis (Craig NL: *Mobile DNA II*. Washington, DC: ASM Press; 2002). The analysis of the sequence of TSDs formed by the *Muta1* element integrations in the transposition assays revealed that it does not have a bias towards insertion into specific sequences unlike the *piggyBac* element (M J Fraser, Ciszczon, Elick, & Bauser, 1996) and *Mariner* elements (Plasterk, Izsvák, & Ivics, 1999). Thus, *Muta1* elements can serve as a genetic tool for mutagenesis and cloning with broad target range.

Studies have shown that TEs transposing via DNA intermediates often leave footprints upon excision from their original site. These excision events vary in size and sequence, which are repaired by DNA repair machinery of the host organism (Bryan, Jacobson, & Hartl, 1987) (Xu et al., 2004). The flanking DNA sequence affects the DNA repair process at the site of excision (Scott, LaFoe, & Weil, 1996). Two major models have been proposed for the excision pattern, one of which is 5'-exonuclease mediated and the other one is through hairpin loop formation and an endonuclease cleavage (Scott et al., 1996). Germinal and somatic excision products of *Mu1*, a *Mutator* superfamily transposon has shown footprints with target site deletions suggesting that exonucleolytic degradation occurs upon excision of *Mu1* element in *Zea mays* (Britt & Walbot, 1991). *Mutator* elements have played major role in plant evolution, due to their property of

forming recombination events upon excision leading to intra-chromosomal crossovers and conversions near the site of integration (Doseff, Martienssen, & Sundaresan, 1991).

Excision of *Muta1* was studied through the two-plasmid assay as described in chapter three in *D. melanogaster* and *Ae. aegypti*. In *D. melanogaster,* I observed precise excision of *Muta1* in 67% of the events from excision assays with the *Muta1* element flanked by an 8 bp TSDs and in 73% of the events with the *Muta1* element flanked by a 9 bp TSDs. In *Ae. aegypti*, precise excision was observed in 25% of the events with the *Muta1* element flanked by an 8 bp TSDs and in 60% of the events with the *Muta1* element flanked by a 9 bp TSDs. Few events were found to have *Muta1* excision with nucleotide deletions in TSDs, excision leaving two TSDs, excision cleaving nucleotides in the flanking restriction sites, and an event with acquisition of base pairs flanking to the TSDs in *D. melanogaster* and *Ae. aegypti*. One possible explanation for this behavior could be host DNA repair machinery and the flanking sequence at the site of integration that could affect the excision mechanism of *Muta1* in *Ae. aegypti*. Therefore, it will be very useful to study *Muta1* transposase binding and cleavage specificity that might influence the transposition behavior of the element across different species and specifically studying host factors in *Ae. aegypti*.

6.4 Germline activity of *Muta1* in *D. melanogaster* and *Ae. aegypti*.

TEs have been used for transformation of Drosophilid species; discovery of TEs in non-Drosophilid insects holds great promise for their use as genetic tool for various molecular techniques within these target species. Moreover, identification of new

endogenous TEs in *Ae. aegypti* may help to overcome the problem of lack of mobility and integration properties observed with the use of transposons like *piggyBac*, *Hermes*, *Mos1* and *Minos* (O'Brochta et al., 2003).

Germline transformation was first carried out using a non-autonomous *Muta1* element; a total of 45 embryos injected with along with helper plasmid resulted in 27 fertile crosses. Of these crosses, 14.8% of G1 progeny expressed DsRed2-1 marker. In a second set of experiments, had two transposons injected into *D. melanogaster* embryos, one carrying *Muta1* element with DsRed2-1 marker and a second plasmid contained *piggyBac* element with GFP marker and *Muta1* transposase. A total of 41 fertile crosses were obtained through injection of 104 embryos, which resulted in 4.9% transgenic offspring with DsRed2-1 marker. The difference in transformation rate for both experiments could be a result of the size of the element being integrated or due to increase in load of transposons in the injection mix. Similar transformation experiment with two plasmids was carried out in *Ae. aegypti*, 4% transformation rate for the *Muta1* element was obtained.

Transformation rates observed with *Muta1* element were low in comparison to *piggyBac* element; the reason could be the cargo capacity of the *Muta1* element, the genetic marker being used for this transformation and with ease the fertile crosses obtained through microinjection of embryos. Remobilization experiment was carried out to determine remobilization rate for the *Muta1* element, new phenotypes were observed with expression of DsRed marker in full body, abdomen, thorax, labellum and genitalia. The

Muta1 plasmid used in this experiment was not designed for enhancer trapping, still I were able to score new phenotypes, which is suggests that the *Muta1* element might have high transposition rate eventually integrating into regions that produced new phenotypes. Molecular characterization of flies that did not produced new phenotypes revealed that the *Muta1* element has remobilized in these flies. New integrations were characterized in chromosome 2R, 2L, 3R, 3L and X chromosome. Remobilization rate for the *Muta1* element was 21.2% in *D. melanogaster*.

In *Ae. aegypti*, germline integration of *Muta1* in *Ae. aegypti* resulted in a 4% transformation rate. Two potential new phenotypes were observed with DsRed expression in the legs and the midgut. Molecular characterization of these new phenotypes revealed that the *Muta1* element has remobilized in the mosquito with dsRed expression in legs. This represents the first instance of a endogenous transposon remobilization post-integration in *Ae. aegypti*. With increase in the number of progeny being examined it is possible that the remobilization frequency of *Muta1* in *Ae. aegypti* might increase and molecular characterization of these new events can help us to study the post-integration behavior of the *Muta1* element in the genome.

6.5 The piRNAs response to the *Muta1* element in *D. melanogaster.*

Based on the studies in *Drosophila* small RNA pathways regulate transposon movement in the genome. The genome of *D. melanogaster* protects itself from mutational burden via the Piwi interacting small RNAs (piRNAs), which play a major role in

regulation of transposons in somatic as well as germline cells through production of primary piRNAs, which are present in both types of cells, and secondary piRNAs specific to the germline. Secondary piRNAs are involved in feed forward loop for ping-pong amplification. Beyond transposons regulation piRNAs are found to be involved in other process such as mRNA degradation, germline differentiation, and chromosome segregation (Pek, Patil, & Kai, 2012).

Based on studies of *P*-element in *D. melanogaster.* It was discovered that piRNAs were responsible for protecting the genome against invasion of transposons. Flies that lack maternally deposited piRNAs showed to have abnormal gonads and cause sterility. This phenomenon was called hybrid dysgenesis (Jensen, Stuart, Goodpaster, Goodman, & Simmons, 2008) (Simmons et al., 2014). Moreover, as the dysgenic hybrids age, fertility is restored due to de novo piRNAs production, these de novo piRNAs are produced as a result of resident elements insertion into a piRNA clusters (Khurana et al., 2011).

In chapter five, I determined the relationship between the *Muta1* element and the piRNAs. The *D. melanogaster* genome was determined to be naïve to the *Muta1* element. In this experiment, the autonomous *Muta1* element was introduced into embryos of *D. melanogaster*. Transformation rate of 16% was obtained with the autonomous *Muta1* transposon. New phenotypes were observed with DsRed expression in thorax, abdomen and full body. The autonomous *Muta1* element is able to transform *D. melanogaster* with higher frequency then the plasmid construct used in previous experiments, and thus it is

possible that post-integration mobility of this autonomous element might be high and cause production of piRNAs against the *Muta1* element. The phenomenon of hybrid dysgenesis could not relate to the *Muta1* element introduction in *D. melanogaster,* as observed through the crosses established in the remobilization experiments.

The small RNA libraries generated from generation 11 and generation 21 determined the piRNA profile to the *Muta1* element. These small RNA libraries were analyzed for 23-30 nt small RNAs that were mapped to the autonomous *Muta1* plasmid. The coverage plot showed an abundance of antisense piRNA, found in both somatic and germline cells to the promoter and polyA regions of Hsp70 and 3xP3-DsRed genes. The *Muta1* transposase showed abundance of sense piRNAs in all four transgenic lines from G11 and G21. This suggest that the *Muta1* transposase might be regulated by the AGO3 protein, which is a germline specific Piwi protein (Brennecke et al., 2007) (Khurana & Theurkauf, 2010). Identification of the integration of the *Muta1* element into a germline piRNA cluster might support the hypothesis that the *Muta1* element regulation is restricted to germline. These small RNA libraries were made from whole adults, which contained larger proportion of somatic tissue. Preparation of libraries from ovaries might help in determining the *Muta1* element regulation by germline specific factors such as AGO3.

I did not observe a U1-A10 overlap bias in the piRNAs that mapped to the *Muta1* element. Only two overlaps were observed to the *Muta1* transposase in transgenic line

A50 G21 and A56 G11. In analysis of the piRNAs biogenesis for the *Penelope* element, a U1-A10 bias was observed in the piRNAs that mapped to the element (Pyatkov et al., 2002). It is possible as more progeny are examined there might be increase in piRNAs for suppression of the *Muta1* element or this element is capable to evade its detection by host silencing machinery. The crosses were examined until generation 23 and the *Muta1* element was found to be active in transgenic flies with DsRed expression in eyes. This suggests that the *Muta1* element is active and somehow it is able to evade host suppression machinery.

6.6 Future Direction

Despite a lower germline transformation rate than *piggyBac* in *D. melanogaster* and *Ae. aegypti*. The *Muta1* element can serve as an important genetic tool due to less bias for integration as observed through screening of transposition events in somatic assays and germline integration events. The data presented here shows strong evidence of post-integration movement of the *Muta1* element in *D. melanogaster* and preliminary remobilization of an endogenous element in *Ae. aegypti*. The post integration behaviors of the *Muta1* element can be applied for enhancer trapping and mutagenesis studying in different insect species and foremost for mosquito transgenesis. Further studies with *Muta1* transposase will help in understanding the mechanism of *Muta1* transposition and DNA binding motifs for the same.

6.7 References

Babu, M. M., Iyer, L. M., Balaji, S., & Aravind, L. (2006). The natural history of the WRKY-GCM1 zinc fingers and the relationship between transcription factors and transposons. *Nucleic Acids Research*, *34*(22), 6505–20. http://doi.org/10.1093/nar/gkl888

Brennecke, J., Aravin, A. A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., & Hannon, G. J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell*, *128*(6), 1089–103. http://doi.org/10.1016/j.cell.2007.01.043

Britt, A. B., & Walbot, V. (1991). Germinal and somatic products of Mu1 excision from the Bronze-1 gene of Zea mays. *Molecular & General Genetics : MGG*, *227*(2), 267–76.

Bryan, G., Jacobson, J., & Hartl, D. (1987). Heritable somatic excision of a Drosophila transposon. *Science*, *235*(4796), 1636–1638. http://doi.org/10.1126/science.3029874

Doseff, A., Martienssen, R., & Sundaresan, V. (1991). Somatic excision of the Mu1 transposable element of maize. *Nucleic Acids Research*, *19*(3), 579–84.

Feschotte, C., & Pritham, E. J. (2007). DNA transposons and the evolution of eukaryotic genomes. *Annual Review of Genetics*, *41*, 331–68. http://doi.org/10.1146/annurev.genet.40.110405.090448

Feschotte, C., Swamy, L., & Wessler, S. R. (2003). Genome-wide analysis of mariner-like transposable elements in rice reveals complex relationships with stowaway miniature inverted repeat transposable elements (MITEs). *Genetics*, *163*(2), 747–58.

Franz, A. W. E., Jasinskiene, N., Sanchez-Vargas, I., Isaacs, A. T., Smith, M. R., Khoo, C. C. H., … Olson, K. E. (2011). Comparison of transgene expression in Aedes aegypti generated by mariner Mos1 transposition and ΦC31 site-directed recombination. *Insect Molecular Biology*, *20*(5), 587–98. http://doi.org/10.1111/j.1365-2583.2011.01089.x

Fraser, M. J. (2012). Insect transgenesis: current applications and future prospects. *Annual Review of Entomology*, *57*, 267–89. http://doi.org/10.1146/annurev.ento.54.110807.090545

Fraser, M. J., Ciszczon, T., Elick, T., & Bauser, C. (1996). Precise excision of TTAA-specific lepidopteran transposons piggyBac (IFP2) and tagalong (TFP3) from the baculovirus genome in cell lines from two species of Lepidoptera. *Insect Molecular Biology*, *5*(2), 141–51.

Han, Y., Burnette, J. M., & Wessler, S. R. (2009). TARGeT: a web-based pipeline for retrieving and characterizing gene and transposable element families from genomic sequences. *Nucleic Acids Research*, *37*(11), e78. http://doi.org/10.1093/nar/gkp295

Han, Y., & Wessler, S. R. (2010). MITE-Hunter: A program for discovering miniature inverted-repeat transposable elements from genomic sequences. *Nucleic Acids Research*, *38*(22).

Jensen, P. A., Stuart, J. R., Goodpaster, M. P., Goodman, J. W., & Simmons, M. J. (2008). Cytotype regulation of P transposable elements in Drosophila melanogaster: repressor polypeptides or piRNAs? *Genetics*, *179*(4), 1785–93. http://doi.org/10.1534/genetics.108.087072

Khurana, J. S., & Theurkauf, W. (2010). piRNAs, transposon silencing, and Drosophila germline development. *The Journal of Cell Biology*, *191*(5), 905–13. http://doi.org/10.1083/jcb.201006034

Khurana, J. S., Wang, J., Xu, J., Koppetsch, B. S., Thomson, T. C., Nowosielska, A., … Theurkauf, W. E. (2011). Adaptation to P element transposon invasion in Drosophila melanogaster. *Cell*, *147*(7), 1551–63. http://doi.org/10.1016/j.cell.2011.11.042

Lisch, D. (2013). Regulation of the Mutator system of transposons in maize. *Methods in Molecular Biology (Clifton, N.J.)*, *1057*, 123–42. http://doi.org/10.1007/978-1-62703-568-2_9

Marquez, C. P., & Pritham, E. J. (2010). Phantom, a new subclass of Mutator DNA transposons found in insect viruses and widely distributed in animals. *Genetics*, *185*(4), 1507–17. http://doi.org/10.1534/genetics.110.116673

O'Brochta, D. A., Sethuraman, N., Wilson, R., Hice, R. H., Pinkerton, A. C., Levesque, C. S., … Atkinson, P. W. (2003). Gene vector and transposable element behavior in mosquitoes. *The Journal of Experimental Biology*, *206*(Pt 21), 3823–34.

Palavesam, A., Esnault, C., & O'Brochta, D. A. (2013). Post-integration silencing of piggyBac transposable elements in Aedes aegypti. *PloS One*, *8*(7), e68454. http://doi.org/10.1371/journal.pone.0068454

Pek, J. W., Patil, V. S., & Kai, T. (2012). piRNA pathway and the potential processing site, the nuage, in the Drosophila germline. *Development, Growth & Differentiation*, *54*(1), 66–77.

Plasterk, R. H., Izsvák, Z., & Ivics, Z. (1999). Resident aliens: the Tc1/mariner superfamily of transposable elements. *Trends in Genetics : TIG*, *15*(8), 326–32.

Pyatkov, K. I., Shostak, N. G., Zelentsova, E. S., Lyozin, G. T., Melekhin, M. I., Finnegan, D. J., … Evgen'ev, M. B. (2002). Penelope retroelements from Drosophila virilis are active after transformation of Drosophila melanogaster. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(25), 16150–5. http://doi.org/10.1073/pnas.252641799

Rommens, C. M., van Haaren, M. J., Nijkamp, H. J., & Hille, J. (1993). Differential repair of excision gaps generated by transposable elements of the "Ac family". *BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology*, *15*(8), 507–12. http://doi.org/10.1002/bies.950150803

Scott, L., LaFoe, D., & Weil, C. F. (1996). Adjacent sequences influence DNA repair accompanying transposon excision in maize. *Genetics*, *142*(1), 237–46.

Simmons, M. J., Meeks, M. W., Jessen, E., Becker, J. R., Buschette, J. T., & Thorp, M. W. (2014). Genetic interactions between P elements involved in piRNA-mediated repression of hybrid dysgenesis in Drosophila melanogaster. *G3 (Bethesda, Md.)*, *4*(8), 1417–27. http://doi.org/10.1534/g3.114.011221

Smith, R. C., & Atkinson, P. W. (2011). Mobility properties of the Hermes transposable element in transgenic lines of Aedes aegypti. *Genetica*, *139*(1), 7–22. http://doi.org/10.1007/s10709-010-9459-7

Talbert, L., & Chandler, V. (1988). Characterization of a highly conserved sequence related to mutator transposable elements in maize. *Mol. Biol. Evol.*, *5*(5), 519–529.

Wilson, R., Orsetti, J., Klocko, A. D., Aluvihare, C., Peckham, E., Atkinson, P. W., … O'Brochta, D. A. (2003). Post-integration behavior of a Mos1 mariner gene vector in Aedes aegypti. *Insect Biochemistry and Molecular Biology*, *33*(9), 853–63.

Xu, Z., Yan, X., Maurais, S., Fu, H., O'Brien, D. G., Mottinger, J., & Dooner, H. K. (2004). Jittery, a Mutator distant relative with a paradoxical mobile behavior: excision without reinsertion. *The Plant Cell*, *16*(5), 1105–14. http://doi.org/10.1105/tpc.019802

Yant, S. R., Meuse, L., Chiu, W., Ivics, Z., Izsvak, Z., & Kay, M. A. (2000). Somatic integration and long-term transgene expression in normal and haemophilic mice using a DNA transposon system. *Nature Genetics*, *25*(1), 35–41. http://doi.org/10.1038/75568