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

Permalink
https://escholarship.org/uc/item/1s91s60r

Author
SHAH, PRESHA VIJAYKUMAR

Publication Date
2015

Peer reviewed|Thesis/dissertation
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, Prithviraj 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 Aedes 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 Aedes 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 Aedes aegypti through excision and transposition assays. The somatic assays revealed that the Muta1 element was active in D. melanogaster and Aedes aegypti. I also determined the germline and remobilization activity of the Muta1 element post integration in D. melanogaster and Aedes 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* superfamly transposon activity in insects specifically *D. melanogaster* and *Ae. aegypti*. 
# Table of Contents

**Chapter One: Introduction.**

1.1 General Introduction to Transposable Elements...........................................1

1.2 Type of Transposons.........................................................................................2

1.3 Transposons in Genetic Modification of Mosquitoes......................................8

1.4 Transposable Elements as Genetic Tools........................................................12

1.5 Genetic Technologies used for Insect Transformation....................................13

1.6 Transposon Control System.............................................................................16

1.7 Thesis Objectives and Aims............................................................................23

1.8 References.......................................................................................................24

**Chapter Two: Identification of new class II DNA transposon in *Ae aegypti.***

2.1 Abstract...........................................................................................................39

2.2 Introduction......................................................................................................40

2.3 Methods and Materials..................................................................................42

2.4 Results.............................................................................................................44

2.5 Discussion.......................................................................................................47

2.6 References.......................................................................................................51

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

3.1 Abstract...........................................................................................................66

3.2 Introduction......................................................................................................66

3.3 Methods and Materials..................................................................................69

3.4 Results.............................................................................................................76

3.5 Discussion.......................................................................................................80

3.6 References.......................................................................................................85
Chapter Four. Muta1 is active in germline of D. melanogaster and Ae. aegypti.

4.1 Abstract ............................................................................. 105
4.2 Introduction ........................................................................... 106
4.3 Methods and Materials ..................................................... 109
4.4 Results .............................................................................. 116
4.5 Discussion ......................................................................... 123
4.6 References ......................................................................... 129

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

5.1 Abstract ............................................................................. 152
5.2 Introduction ........................................................................... 153
5.3 Methods and Materials ..................................................... 157
5.4 Results .............................................................................. 161
5.5 Discussion ......................................................................... 169
5.6 References ......................................................................... 173

Chapter six Summary and conclusions

6.1 Summary ............................................................................. 200
6.2 Discovery of the Muta1 element in Ae. aegypti.................. 200
6.3 Somatic activity of the Muta1 element in D. melanogaster and Ae. aegypti ....... 203
6.4 Germline activity of Muta1 in D. melanogaster and Ae. aegypti .......... 205
6.5 The piRNAs response to the Muta1 element in D. melanogaster .......... 207
6.6 Future Direction ................................................................. 210
6.7 References ......................................................................... 211
List of Tables

2.1 Transposons discovered using TARGeT pipeline..................................................55

3.1 Somatic transposition of Mutal in D. melanogaster and Ae. aegypti...............92

3.2 Excision of Mutal in D. melanogaster and Ae. Aegypti.............................93

3.3 Sequences of TSDs generated by transposition of Mutal..............................94

4.1 Transformation rate from transformation and remobilization experiment for Mutal in
D. melanogaster....................................................................................................133

4.2 Remobilization experiment results.................................................................134

4.3 Molecular characterizations of Mutal and piggyBac parental lines in D.
melanogaster..........................................................................................................135

4.4 Molecular characterizations of new phenotypes in D. melanogaster for Mutal
element integration..............................................................................................136

4.5 Molecular characterizations of new transposition events in D. melanogaster...137

4.6 Transformation rate for Mutal and piggyBac in Ae. aegypti.........................138

4.7 Molecular characterizations of Mutal parental lines in Ae. aegypti..............138

4.8 Remobilization experiment results.................................................................139

5.1 Transformation rate for the autonomous Mutal element in D. melanogaster.....177

5.2 Molecular characterization of the integration locations of parental lines in D.
melanogaster..........................................................................................................177

5.3 New phenotypes observed in the autonomous Mutal transformation experiment...178

5.4 Molecular characterizations of the new transposition events for the autonomous
Mutal element......................................................................................................178

5.5 Number of small RNA sequence obtained from each library and relative number of
piRNAs that mapped to the autonomous Mutal element....................................179

5.6 Normalized counts for the piRNAs mapped the autonomous Mutal plasmid...180

5.7 Frequency of U1 (uracil) in antisense and A10 (adenine) in sense piRNAs from the
eight libraries........................................................................................................182

5.8 U1-A10 overlaps for the piRNA libraries from the autonomous Mutal transgenic
line....................................................................................................................................183
List of Figures

1.1 Types of transposons and their mechanism of transposition.................................37
1.2 piRNA biogenesis and transposon silencing in somatic and germline cells........38
2.1 Steps involved in discovery of consensus sequences from a MITE-Hunter pipeline..56
2.2 Steps involved in discovery of consensus sequences from a TARGeT pipeline......57
2.3 Structure of the Mutal1 element discovered using the TARGeT pipeline.........58
2.4 Supercontig locations of eight wild-type copies of Mutal1 in Ae. aegypti........59
2.5 The analysis of BED domain and the catalytic domain of the two new hAT transposons.................................................................60
2.6 The TIRs and TSDs length, zinc finger DNA binding domain and catalytic DDE motif identified in hAT-3_SM transposon..................................................62
2.7 The TIRs and TSDs length, zinc finger DNA binding domain and catalytic DDE motif identified in hAT-1_AA transposon......................................................63
2.8 The TIRs and TSDs length, BED zinc finger DNA binding domain and catalytic DDE motif identified in Mutal2 transposon.......................................................64
2.9 The TIRs and TSDs length, BED zinc finger DNA binding domain and catalytic DDE motif identified in Mutal3 transposon.......................................................65
3.1 Five plasmid normalized transposition assay.......................................................96
3.2 The excision pattern of the Mutal1 flanked by an 8 bp TSDs in D. melanogaster.97
3.3 The excision pattern of the Mutal1 flanked by an 8 bp TSDs in Ae. aegypti........98
3.4 The excision pattern of the Mutal1 flanked by a 9 bp TSDs in D. melanogaster...99
3.5 The excision pattern of the Mutal1 flanked by a 9 bp TSDs in Ae. aegypti .......100
3.6 WebLogo of consensus target site preference of Mutal1 flanked by an 8 bp TSD in D. melanogaster.................................................................101
3.7 WebLogo of consensus target site preference of Mutal1 flanked by an 8 bp TSD in Ae. aegypti.................................................................102
3.8 WebLogo of consensus target site preference of Mutal1 flanked by a 9 bp TSD in D.
melanogaster .................................................................103

3.9 WebLogo of consensus target site preference of Muta1 flanked by a 9 bp TSD in Ae.
aegypti ........................................................................104

4.1 Schematic diagram of the reporter and helper line for Muta1 remobilization
experiment .................................................................140

4.2 Transgenic flies with marker expression in eyes.................................141

4.3 New phenotypes observed in crosses for D. melanogaster remobilization experiment
for the Muta1 element ..................................................142

4.4 Ae. aegypti transgenic mosquitoes ........................................144

4.5 Potential new phenotypes observed in crosses for Ae. aegypti ..............145

4.6 PCR on transgenic flies with DsRed marker and White gene marker ..........146

4.7 PCR on transgenic mosquitoes with pMuta13EHspBac ..........................146

4.8 RT-PCR on Muta1 transgenic helper lines ........................................147

4.9 Percentage of new Muta1 integrations in D. melanogaster ..................148

4.10 Local hopping observed with the Muta1 element in D. melanogaster .......149

4.11 Local hopping observed with the Muta1 element in D. melanogaster .......150

4.12 The new integration location obtained for the Muta1 element with DsRed expression
in legs in Ae. aegypti ......................................................151

5.1 Schematic diagram of the autonomous Muta1 experiment ....................184

5.2 New phenotypes observed in the crosses for autonomous Muta1 experiment ....185

5.3.1 Coverage plots representing control library RNAlib14 from D. melanogaster
mapped to autonomous Muta1 element ................................186

5.3.2 Coverage plot for library A19 G11 ........................................187

5.3.3 Coverage plot for library A19 G21 ........................................188
5.3.4 Coverage plot for library A50 G11.................................................................189
5.3.5 Coverage plot for library A50 G21.................................................................190
5.3.6 Coverage plot for library A56 G11.................................................................191
5.3.7 Coverage plot for library A56 G21.................................................................192
5.3.8 Coverage plot for library A85 G11.................................................................193
5.3.9 Coverage plot for library A85 G21.................................................................194

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.....195

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.....195

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.....196

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.....196

Script 5.1 R script for removal of contaminants from the small RNA library.........197

Script 5.1 R scripts used to make coverage plots..................................................198
Chapter 1 Introduction

Mobilization of Newly Identified Transposon Mutal 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 (Egger & 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 viridis* 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 produce 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.
The *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 mosquitoes leaves flanking plasmid sequence along with the gene of interest and thus moves via non-canonical cut and paste mechanism (N Jasinskie, 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*.
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, 1993). This system utilizes enhancer trapping with the construct carrying the \textit{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- \textit{FRT} from the \textit{Saccharomyces cerevisiae} (Park, Masison, Eisenberg, & Greene, 2011) and CRE-\textit{lox} from the Bacteriophage P1 (Sauer & Henderson, 1988) are used for gene knockout and
transgene expression. In the Cre-\textit{loxP} or Flp-FRT system, Cre or Flp recombinase recognizes the 34-bp nucleotide sequence named \textit{loxP} or FRT and precisely catalyzes the homologous exchange between the two \textit{loxP} or the two \textit{FRT} sites, respectively. The Cre-\textit{loxP} and Flp-FRT systems (Hoess, Abremski, & Sternberg, 1984) enable the specific manipulation of DNA based on the direction and location of the two \textit{loxPs} or \textit{FRTs} sites. Cre catalyzes the deletion of the DNA between the two \textit{loxP}s sites, when the two \textit{loxP}s are in the same direction on one DNA molecule. When one \textit{loxP} is on a linear DNA molecule and another \textit{loxP} site is on a circular DNA molecule, the circular DNA integrates into the linear DNA at the target. If two \textit{loxP}s sites are oriented in opposite directions, the fragment between them inverts. When one \textit{loxP} site is on a linear strand and a second \textit{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 \textit{D. melanogaster} as a tool kit to inactivate genes (Frickenhaus, Wagner, Mallik, Catinozzi, & Storkebaum, 2015). The \textit{cre-loxP} site-specific recombination system has also been used in \textit{Ae. aegypti} (Nijole Jasinskiene, Coates, Ashikyan, & James, 2003). The \textit{cre-loxP} system was used to remove a marker gene from transgenic mosquitoes. The \textit{cre} recombinase precisely recognizes \textit{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 \textit{Streptomyces} Bacteriophage is used for stable integration of foreign DNA (Thorpe & Smith, 1998). Integrase from phage phiC31 catalyzes the homologous recombination between the \textit{Streptomyces} attachment site \textit{attB}
and the phage attachment site \textit{attP}. The advantage of phiC31 integrase is that recombination is unidirectional, with interaction between \textit{attB} and \textit{attP} sites creating \textit{attL} and \textit{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 \textit{D. melanogaster} S2 cell culture, where intermolecular recombination occurred at a 47\% rate. Transgenic lines were created using \textit{P}-element to integrate \textit{attP} sites and phiC31 integrase mRNA was injected into embryos to promote integration of plasmid with the \textit{attB} site; 55\% transgenic offspring were produced in \textit{D. melanogaster} with precise integration of the \textit{attP} site (Groth, Fish, Nusse, & Calos, 2004). In order for precise targeting of transgenic constructs to a predetermined position in the genome of \textit{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 \textit{Ae. albopictus}, the Asian tiger mosquito, was achieved by microinjection of \textit{piggyBac} with 3xp3-ECFP marker and an \textit{attP} site combined with \textit{piggyBac} transposase mRNA. Transformation frequency obtained with this injection was 2-3\%. The transgenic lines were injected with a second plasmid containing an \textit{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 \textit{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 \textit{D. melanogaster} where the enzyme FRT had to be introduced into the \textit{D. melanogaster} using \textit{P}-element mediated transformation (Rong and Golic, 2000). Furthermore, transformation is not always achieved using this recombination system in mosquitoes, which was observed in \textit{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 interferingRNAs (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 10th and 11th 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 10th 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 (Rajaseethupathy 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,
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., 2011).
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 Mutal transposon. Chapter three and four investigates transposition, excision, target site preference, and transformation efficiency of newly discovered transposon Mutal in D. melanogaster and Ae. aegypti. Chapter five demonstrates relationship between piRNAs and the Mutal element in D. melanogaster. Overall, the research provides valuable information about the activity of the Mutal element, which can have a wide variety of application for insect transgenesis and human gene therapy.
1.8 Reference


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.


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


Figure 1.1 Types of transposons and their mechanism of transposition; Class I-\textit{Retro}elements, and Class II-\textit{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 dTstu1 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 Hd1 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 \textit{hAT} superfamily, I was able to discover autonomous elements that contained characteristics of a \textit{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 \textit{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 \textit{Ae. aegypti} (Table 2.1). Another \textit{hAT} element that was discovered and determined to be present in RepBase database was \textit{AeBuster1}, which had been tested for its transposition activity in \textit{D. melanogaster} and \textit{Ae. aegypti} (Figure 2.5) (Arensburger et al., 2011). Two other \textit{hAT} elements that were discovered were identified from the consensus sequences of the hATx3_SM and hAT-1_AA (Table 2.1). The \textit{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 \textit{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 \( hAT-I_{AA} \) 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 \( hAT-I_{AA} \) 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 \textit{Mule} superfamily elements were discovered. The \textit{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 \textit{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 \textit{Mutator} transposases (Babu, Iyer, Balaji, & Aravind, 2006) and functions in DNA binding for transposition (Figure 2.3). \textit{Muta12} was discovered using the consensus sequences of MuDR7x\_AP, it has 59 bp TIRs and forms 9 bp TSDs (Figure 2.8), while \textit{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 \textit{Muta12} and \textit{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


Table 2.1 Transposons discovered using the TARGeT pipeline.

<table>
<thead>
<tr>
<th>Superfamily</th>
<th>Consensus sequence</th>
<th>TSDs</th>
<th>TIRs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>hAT</strong></td>
<td>TAHAT7</td>
<td>GTTGAGTA</td>
<td>TAGAGATGGGCAA</td>
</tr>
<tr>
<td></td>
<td>(AeHerves2)</td>
<td>8 bp</td>
<td></td>
</tr>
<tr>
<td></td>
<td>hAT_29_HM</td>
<td>GCTTATGG</td>
<td>CATAGATTCCTCA</td>
</tr>
<tr>
<td></td>
<td>(AeBuster1)</td>
<td>8 bp</td>
<td></td>
</tr>
<tr>
<td></td>
<td>hATx-3_SM</td>
<td>CCCGGGAA</td>
<td>AAATCCCGGAATCC</td>
</tr>
<tr>
<td></td>
<td>8 bp</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>hAT-1_AA</td>
<td>CATCGAAT</td>
<td>TAGAGTGTCATGGA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mule/Mutator</strong></td>
<td>MuDr_1_TV</td>
<td>GACGCAGT</td>
<td>GGGTCTACCCCGT</td>
</tr>
<tr>
<td></td>
<td>(Muta1)</td>
<td>8/9 bp</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MuDr7x_AP</td>
<td>TTCCCCCAT</td>
<td>TAGGGCGGTCACCAA</td>
</tr>
<tr>
<td></td>
<td>(Muta12)</td>
<td>9 bp</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MuDR4x_SM</td>
<td>TCTAGAAG</td>
<td>TAAGGGATGCGAAAT</td>
</tr>
<tr>
<td></td>
<td>(Muta3)</td>
<td>8 bp</td>
<td></td>
</tr>
</tbody>
</table>
**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 Mutal element discovered using the TARGeT pipeline. The figure shows the TSDs and TIRs for the Mutal 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 Mutal transposase.
<table>
<thead>
<tr>
<th>Supercontig</th>
<th>Flanking sequence</th>
<th>TSD</th>
<th>TIR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supercontig 1.223</td>
<td>ATATACACTAATAATGTTCTAATAATGACGCAGTGGGTTCTACCCCGTTTGGCATATAATGCCG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supercontig 1.869</td>
<td>TTTTTACATTTTACAGGACACATGTGCCGTGGGTTCTACCCCGTTTGGCATATAATGCCG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supercontig 1.873</td>
<td>TTACTACAAAGCTACTGTAGTTAGCTCACAGGGGTCTACCCCGTTTGGCATATAATGCCG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supercontig 1.801</td>
<td>TGCCGAGTAACCCTACACCTAACACTGATTTGAGGTTCTACCCCGTTTGGCATATAATGCCG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supercontig 1.509</td>
<td>TGGTAAGGTATAAGGTATTTAATTACGTTTTTTGTTACCCCGTTTGGCATATAATGCCG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supercontig 1.304</td>
<td>GCATTCGAGAATGGAAATCGGTGTAATGGTACATGGGTTCTACCCCGTTTGGCATATAATGCCG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supercontig 1.3</td>
<td>TGGTTTGAAATCCCTACTCGGTACCTTTATATTGGGTTCTACCCCGTTTGGCATACAGTGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supercontig 1.922</td>
<td>ACCCTAATCAAAAATCCAAAAATCATGATCCGGGTGCTACCCCGTTTGGCATATAATGCCG</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**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**

<table>
<thead>
<tr>
<th>Domain</th>
<th>BED</th>
<th>DNA binding</th>
<th>CATALYTIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>hATx-3_SM</td>
<td>19 CYK—C (10) HMKN—KH</td>
<td>ARLAADKGFSL 105 TRS 105 SDPR</td>
<td>126 VMKLVS 131 TDF</td>
</tr>
<tr>
<td>hAT-1_AA</td>
<td>23 CEQI—C (40) HMK—KH</td>
<td>ARLAADKGFSL 108 TRS 165 LDEWT</td>
<td>211 RMVQL 229 TDG</td>
</tr>
<tr>
<td>Tol2</td>
<td>47 CVL—C (19) HIER—MH</td>
<td>127 YIQGLKPSST 163 RSK 191 TDCWT</td>
<td>234 EYLVS 258 TDS</td>
</tr>
<tr>
<td>Herves</td>
<td>26 CLY—C (16) HLNL—VH</td>
<td>115 LICKELFPNL 151 SNA 180 SDGW</td>
<td>222 RNMYQ 246 TDN</td>
</tr>
<tr>
<td>Hermes</td>
<td>51 CRK—C (16) HKCC—AS</td>
<td>101 WYVDRCPFSA 147 SRE 179 IDLWT</td>
<td>222 ENYIK 247 TRD</td>
</tr>
<tr>
<td>TcBuster1</td>
<td>77 CVI—C (16) HLDT—KH</td>
<td>147 RIAQKEAYTI 193 SRR 222 MEDST</td>
<td>263 EEIIN 288 TDG</td>
</tr>
<tr>
<td>AeBuster1</td>
<td>75 CVI—C (16) HLET—KH</td>
<td>145 KIAKSGKAHTI 191 ARR 219 MEDST</td>
<td>259 DEIFD 284 TDG</td>
</tr>
</tbody>
</table>

---Block A--

<table>
<thead>
<tr>
<th>Domain</th>
<th>--Block B--</th>
<th>--Block C--</th>
</tr>
</thead>
<tbody>
<tr>
<td>hATx-3_SM</td>
<td>148 ICTHI</td>
<td>243 KLSQDLHRLWVDEMLL 273 SCMQRNP</td>
</tr>
<tr>
<td>hAT-1_AA</td>
<td>249 QYNHA</td>
<td>287 DLLASQGVEYLRHRKDDKVL 325 KKFQKTLQ</td>
</tr>
<tr>
<td>Tol2</td>
<td>325 KCAHL</td>
<td>382 LGQLRPNSTPMAY 449 DILQAEYN</td>
</tr>
<tr>
<td>Herves</td>
<td>265 PCFAJIT</td>
<td>319 LKMTQEVSTFWNNSGYDLML 382 NIVSAQKY</td>
</tr>
<tr>
<td>Hermes</td>
<td>264 NCSSHL</td>
<td>309 SSLKSECRTRWNTYTMRL 372 KELQTCSS</td>
</tr>
<tr>
<td>TcBuster1</td>
<td>316 HCCHVR</td>
<td>369 KNLLHTETVRWLSRGKVL 433 LSLQGPNS</td>
</tr>
<tr>
<td>AeBuster1</td>
<td>312 HCSLHR</td>
<td>365 TSSLHHTETVRWLSRGKVL 420 TYLQGSTS</td>
</tr>
</tbody>
</table>

---Insertion---

<table>
<thead>
<tr>
<th>Domain</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>hATx-3_SM</td>
<td>303 HIFQQLRTPSLLRSSLLRSEDFRIIFRTTEDPWIMLNLNRPVYYVSNRLSKSFRQY</td>
</tr>
<tr>
<td>hAT-1_AA</td>
<td>338 KTRWNS --TEYMLERFQFQTCCLREALEEIGLPDLVLTNDLEVAKLHCLIQPVYLRASEALGRRSDAT</td>
</tr>
<tr>
<td>Tol2</td>
<td>583 --TER-----5LLTYFAICSLAIKTNTPLPASAAACERLQSTAGLLFSKPPARLDTN 84FQILLLKNLRRFY</td>
</tr>
<tr>
<td>Herves</td>
<td>541 LWKE--HQVLKPSLYTLAMSTLCPGTSVPCERLFSKAGQQYSEKRSRAPPKLQELFIQKN---</td>
</tr>
</tbody>
</table>

---Block A--

---Block B--

---Block C--
The analysis of BED domain and the catalytic domain of two new \textit{hAT} transposons \textit{hATx3\_SM} and \textit{hAT\_1\_AA}. The domains identified in these two transposons were based on alignment to other known \textit{hAT} transposons. (Atkinson PW (2015) \textit{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 Mutator transposons. 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 Mutal 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 Mutal transposon in D. melanogaster and Ae. aegypti, this being the first time a member of this family has been shown to function in insects. Mutal 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 Mutal element. Mutal precisely excised in 67% to 73% of the events in D. melanogaster and in 25% to 60% of the events in Ae. aegypti. Mutal 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 Mutator 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′-AATGGTACCGCTTATGGCATAGATTCCCCAAACTGTG-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 KpnI and XhoI (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'-GATTCTAGATGCGCATCGAACAACATTTTAGTGAG-3') and Muta1 RE Reverse (5'-AATGAGCTCCCCATAAGCCATAGGTCCAAACTTTTC-3'), which encompassed the region 350 bp 3' of the stop codon through the target-site duplication. The PCR program was: 94° 2', 4 x (94° 20", 60° 20", 72° 30"), 25 x (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 SacI and XbaI (New England Biolabs, NEB) to yield the clone pBSMuta1LR. XbaI-digested pBSMuta1LR was ligated with NheI-digested pGENToriAlpha, which was derived from pK19 and in which the kanamycin\textsuperscript{R} gene was replaced with the gentamycin\textsuperscript{R} gene from pFastBac HTb (Invitrogen) to generate a donor element, which has a replication origin, a gentamycin\textsuperscript{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 Mutal ORF Rev (5’-GATGATATCTTTAATGACTTTTTTCGCTTGCTTATTATTGCAC-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 SpeI and XhoI, 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’-ARACCGCGGTCTTTTTAACCCTYAGAAAGATAGTCGCC-3’) and XbaI-pBac LE R + ORF (5’-ATATCTAGGCTCCTGCTAAGAAACTACC-3’). The right end, flanking plasmid, and TSD were amplified with the restriction site primers KpnI-pBac RE F 5’-ATAGGTACCCTATTATTAACCCTAGAAAGATACATATTGTG-3’ and XhoI-pBac RE F 5’-ATACTCGAGGACTAATAAATTTTCTATTATGTATAAGTG-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 pGENTor iα plasmid by digestion with BamHI and Nhel, and replaced by a BamHI/Nhel fragment containing the EGFP ORF. The BamHI/Nhel EGFP fragment was generated through the PCR amplification of the ORF using the primers BamHI EGFP and Nhel EGFP and the pBac[3xP3-EGFPafm] plasmid as a template. The PCR product was then digested with BamHI and Nhel 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 Mutal donor, 250 ng/ul of the Mutal 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 Mutal, we can compare the number of transposition events of the Mutal 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 Mutal 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, \( \frac{1}{200} \) 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 Mutal 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 diH₂O. 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 *EcoRV* 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 *EcoRV* cuts only within the *Muta1* donor element, plasmids arising as a result of excision are resistant to *EcoRV* 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\(^{\prime}\)-CGTCCCATTCCCGATTCAAG-3\(^{\prime}\).

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 $5.7 \times 10^{-4}$ in *D. melanogaster* and $8.3 \times 10^{-3}$ 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 $3.96 \times 10^{-4}$ in *D. melanogaster* and $9.45 \times 10^{-2}$ in *Ae. aegypti* (Table 3.1). *Muta1* was therefore approximately 1.3 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. aegypti* 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 $1.4 \times 10^{-2}$ in *D. melanogaster* and $1.9 \times 10^{-3}$ 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 footprint 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 $2.24 \times 10^{-2}$ in *D. melanogaster* and $8.73 \times 10^{-3}$ 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 *KpnI* 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).

Mutal 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 Mutal and Mutal-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 Mutal activity in absence of the helper plasmid. Thus, the D. melanogaster genome is naïve to the Mutal 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 2\(^{nd}\) position, A/T at the 6\(^{th}\) and ‘T’ at the 7\(^{th}\) 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 Mutal 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 Mutal 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 Mutal 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 Mutal 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 Mutal transposase to the ends of transposon for the excision. The Mutal 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


85


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 donor titer obtained for the *Muta1* element and the *piggyBac* element from the transposition 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 transposition events obtained to the donor titer. The zero control transposition assays demonstrate transposition activity in absence of the *Muta1* transposase, this experiment used the *piggyBac* element with the *piggyBac* helper plasmid as an internal control. Thus, it shows transposition frequency for the *piggyBac* element.

<table>
<thead>
<tr>
<th>Insect Species</th>
<th><em>Muta1</em> Transposase</th>
<th>TSD Length</th>
<th>No. of Expts.</th>
<th>No. of Embryos Injected</th>
<th><em>Muta1</em> Donor Plasmid Titer</th>
<th><em>piggyBac</em> Donor Plasmid Titer</th>
<th>No. of <em>Muta1</em> Transpositions</th>
<th>No. of <em>piggyBac</em> Transpositions</th>
<th><em>Muta1</em> Transposition Frequency (std. dev.)</th>
<th><em>piggyBac</em> Transposition Frequency (std. dev.)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>D. melanogaster</em></td>
<td>+</td>
<td>8 bp</td>
<td>4</td>
<td>444</td>
<td>127,200</td>
<td>154,400</td>
<td>51</td>
<td>69</td>
<td>5.7 x 10^4 (2.8 x 10^4)</td>
<td>3.9 x 10^4 (8.2 x 10^4)</td>
</tr>
<tr>
<td>&quot;</td>
<td>-</td>
<td>8 bp</td>
<td>2</td>
<td>200</td>
<td>319,000</td>
<td>30,800</td>
<td>0</td>
<td>21</td>
<td>6.8 x 10^4 (5.0 x 10^4)</td>
<td>0</td>
</tr>
<tr>
<td>&quot;</td>
<td>+</td>
<td>9 bp</td>
<td>3</td>
<td>300</td>
<td>765,000</td>
<td>709,400</td>
<td>254</td>
<td>217</td>
<td>3.96 x 10^4 (1.67 x 10^4)</td>
<td>3.05 x 10^-4 (9.65 x 10^-4)</td>
</tr>
<tr>
<td>&quot;</td>
<td>-</td>
<td>9 bp</td>
<td>2</td>
<td>200</td>
<td>732,000</td>
<td>993,600</td>
<td>0</td>
<td>215</td>
<td>2.18 x 10^4 (2.25 x 10^4)</td>
<td>0</td>
</tr>
<tr>
<td><em>Ae. aegypti</em></td>
<td>+</td>
<td>8 bp</td>
<td>2</td>
<td>120</td>
<td>102,000</td>
<td>128,800</td>
<td>1,234</td>
<td>74</td>
<td>8.3 x 10^4 (6.8 x 10^3)</td>
<td>5.3 x 10^4 (2.3 x 10^4)</td>
</tr>
<tr>
<td>&quot;</td>
<td>-</td>
<td>8 bp</td>
<td>3</td>
<td>250</td>
<td>779,000</td>
<td>521,800</td>
<td>0</td>
<td>20</td>
<td>3.7 x 10^4 (3.5 x 10^4)</td>
<td>0</td>
</tr>
<tr>
<td>&quot;</td>
<td>+</td>
<td>9 bp</td>
<td>3</td>
<td>300</td>
<td>359,800</td>
<td>257,000</td>
<td>3,306</td>
<td>192</td>
<td>9.45 x 10^4 (1.29 x 10^4)</td>
<td>7.19 x 10^-1 (1.23 x 10^-1)</td>
</tr>
<tr>
<td>&quot;</td>
<td>-</td>
<td>9 bp</td>
<td>3</td>
<td>300</td>
<td>841,800</td>
<td>331,200</td>
<td>0</td>
<td>155</td>
<td>2.84 x 10^4 (0.7 x 10^-4)</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 3.2 Excision of MutAl in developing embryos of D. melanogaster and Ae. aegypti. The two-plasmid assays were performed with the MutAl donor plasmid and the helper plasmid. The donor plasmid contained the MutAl 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.

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

<table>
<thead>
<tr>
<th>Insertion Site in Target Plasmid (bp)</th>
<th>TSD Sequence</th>
<th>Orientation</th>
<th>TSD Length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>220</td>
<td>GTAAAAAAA</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>308</td>
<td>CATAAAATC</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>398</td>
<td>GAATATGAC</td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>409</td>
<td>AGTAAAGG</td>
<td>+</td>
<td>8</td>
</tr>
<tr>
<td>592</td>
<td>TGATAAAGGC</td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>676</td>
<td>CTAATAAAT</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>725</td>
<td>GTTGAAGTT</td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>819</td>
<td>GTTTCAAATA</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>887</td>
<td>TGAAGAAGG</td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>895</td>
<td>GCCTTCTTCT</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>1014</td>
<td>GTTTTTGCA</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>1993</td>
<td>CTCTAGAG</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>2001</td>
<td>CCGGGGATC</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>2032</td>
<td>GTGGCACAAG</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>2164</td>
<td>TGTTCTATA</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>2184</td>
<td>CGAGAAAC</td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>2184</td>
<td>CGAGAAAC</td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>2199</td>
<td>GTTGGGAATG</td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>2225</td>
<td>CTAACAAAG</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>2394</td>
<td>GTACACACT</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>2547</td>
<td>GTCGTTTCAC</td>
<td>-</td>
<td>9</td>
</tr>
</tbody>
</table>

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*.

<table>
<thead>
<tr>
<th>Insertion Site in Target Plasmid (bp)</th>
<th>TSD Sequence</th>
<th>Orientation</th>
<th>TSD Length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>AATCTTGTA</td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>198</td>
<td>CTAATAGCC</td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>306</td>
<td>CTGATTTTA</td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>336</td>
<td>GGCCTTGGG</td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>349</td>
<td>GTTTGTAAT</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>411</td>
<td>TAAAAGCAG</td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>515</td>
<td>ATGGAAGAT</td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>Insertion Site in Target Plasmid (bp)</td>
<td>TSD Sequence</td>
<td>Orientation</td>
<td>TSD Length (bp)</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>----------------</td>
<td>-------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>54</td>
<td>CATCAAGAT</td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>62</td>
<td>ATCTTGGATG</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>71</td>
<td>CAAATATAT</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>223</td>
<td>AAAATTACC</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>411</td>
<td>CTGCTTTTT</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>624</td>
<td>GTCGTAATG</td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>864</td>
<td>GCTTGTCTCT</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>1139</td>
<td>GATTTTTTC</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>2072</td>
<td>ATGGGATAT</td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>2223</td>
<td>TTAGCTAACC</td>
<td>-</td>
<td>9</td>
</tr>
</tbody>
</table>

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*.

<table>
<thead>
<tr>
<th>Insertion Site in Target Plasmid (bp)</th>
<th>TSD Sequence</th>
<th>Orientation</th>
<th>TSD Length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>TTTAGATAT</td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>65</td>
<td>ATATTTGGG</td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>219</td>
<td>ATTTTTTTA</td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>301</td>
<td>CTTTGTAGC</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>623</td>
<td>GGTGTAAT</td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>1028</td>
<td>GATTTTTG</td>
<td>+</td>
<td>8</td>
</tr>
<tr>
<td>2183</td>
<td>CGAGAAAC</td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>2283</td>
<td>GTTGAGTTG</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>2389</td>
<td>AGTAAGTAT</td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>2416</td>
<td>GTTTTAACT</td>
<td>+</td>
<td>9</td>
</tr>
</tbody>
</table>

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*.
Figure 3.1 Five plasmid normalized transposition assay. (Wright, Smith, Li, Craig, & Atkinson, 2013). The donor plasmid contained the Mutal ends with LacZ reporter gene, the helper plasmid contained Mutal transposase under Hsp70 promoter, the piggyBac donor plasmid contained EGFP reporter gene with piggyBac 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 Mutal and piggyBac transposition event are screened for LacZ/Cam$^R$ and EGFP/Cam$^R$, respectively.
Figure 3.2 The figure shows excision pattern of the Mutal flanked by an 8 bp TSDs and the respective enzyme sites in D. melanogaster. The Mutal 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.

\[ \ldots \ldots \text{GGTACCGACGCAGT} \quad \text{TIRs} \quad \text{GACGCAGT(AG)GAGCTC} \ldots \ldots \]

\[ \quad \text{Kpn1} \quad \text{TSD} \quad \text{TSD} \quad \text{Sac1} \]

12... GGTACCGACGCAGT
2...... 40 bp Deletion
1.... GGTACCGACGCAGGT
1.... GGTA
1.... GGTACC GA
1.... 34 bp Deletion

(AG)GAGCTC...
40 bp Deletion...
(AG)GAGCTC...
(G)GAGCTC...
GAGCTC...
36 bp Deletion....
Figure 3.3 The figure shows excision pattern of the Mutal flanked by an 8 bp TSDs and the respective enzyme sites in Ae. aegypti. The Mutal 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 Mutal 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.
The figure shows excision pattern of Mutα1 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 3.5**
Figure 3.6 WebLogo of consensus target site preference of Mutal1 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 Mutal 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 2\(^{nd}\), 5\(^{th}\), 6\(^{th}\), 7\(^{th}\) and ‘A’ at the 8\(^{th}\) 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 2\textsuperscript{nd}, 3\textsuperscript{rd}, 5\textsuperscript{th}, 6\textsuperscript{th}, and 7\textsuperscript{th} 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 (Jasinskieene 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, Nurinsky, 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-DsRed2-1 fragment and the vector pBSMuta1LR. The vector pBSMuta1LR is digested using FastDigest Xmal and Xbal, treated with FastAP (Thermo-Fisher) and purified by agarose gel electrophoresis. The insert was SmaI 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 EcoRI 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 electrophoresis. 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 pBac3EHspMuta1 Muta1 helper plasmid for remobilization experiments.

pBac3EHspMuta 1 was constructed through ligation of a PCR fragment of hspMut1 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 NheI 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. aegypti* 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′-GATGGTACCGACGAGCTACCC-3′ and *White*
gene marker Rev: 5’-CGCTGAGTGGAATGTC-3’ and White gene For: 5’-GAATACAAGTATTTCCCCTCGAAC-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’-GATGGTACCGACGCTGGGTCTACCC-3’ and DsRed marker Rev: 5’-CCTTGTCACCTTCAGCTTC-3’. Muta1 RE Rev: 5’-GATGAGCTCTACTGCCCAGGTCTACC-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 TaqI enzyme at 65°C (Inverse PCR and sequencing protocol on 5 fly preps. Exelixis, 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’-CATAACAGTGAGAAGCGTACG-3’
For1 for LE: 5’-TGTCAGATCCATTAGTCACGT-3’
REV1 for RE: 5’-GTATACATAGAAATGTAATGAAAACCTCTCTC-3’
For1 for RE: 5’-TGACTCATGTGAACACGGTAAC-3’

Primers for second round for left and right end:

Rev2 for LE: 5’-GTACGACTAGATAAGATGTTCATCATGT-3’
For2 for LE: 5’-GCTTTATACTAGTTTATTTGGATATTGACTACG-3’
Rev2 for RE: 5’-TCCTAAGTGAAGCTGCAATTGC-3’
For2 for RE: 5’-GTTTTAAAAATACGATTTTCTGGTTATGG-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’-GTAATACGACTCATAGGGCAGCGTGGTACGACGGCGGCCGCTGGT-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’-GTAATACGACTCATATAGGGC-3’ and Mutal RE GW1: 5’-CTTGAGGCATTGACGCTTATAGG-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 Mutal RE GW2: 5’-GGTTATGGTTATGCCAAACGACTATTATGCCAATG-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 Mutal is active in germline of D. melanogaster

Germ-line transformation of D. melanogaster was achieved using pMutalDsRedTFD as the reporter plasmid and the Mutal helper plasmid. Transformation was also obtained using pMutalwhiteTFD 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 G₀ 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 pMuta1whiteTFD along with wildtype Muta1 transposase cloned into pMuta1 helper resulted in 21 embryos surviving to adulthood (a survival rate of approximately 15%) and 18 of those were fertile (86% of the G₀ 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 G₀ 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 G₁ 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/Kninrps 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 Mutal element, as the Mutal 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 (Figure 4.12). 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

*Mutal1* is active in the germline of *D. melanogaster* and *Ae. aegypti*. The transformation rate obtained for the non-autonomous *Mutal1* 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 *Mutal1* 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 *Mutal1* element from molecular analysis of 80 flies. Molecular characterization of flies with new integrations of *Mutal1* element revealed that these flies retain parental integration of *Mutal1* 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 Mutal 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 Mutal element. The integration into polyA was obtained from characterization of the right end, as left 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 Mutal transposition. These events might be result of the Mutal 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 Mutal might be a result of the mechanism of Mutal transposition, in which the transposase requires internal terminal sequences for excision and integration. Moreover, there is a possibility that the Mutal 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 Mutal 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 Mutal element insertion or it could be an artifact. Further analysis needs to be done to understand the mechanism of transposition of the Mutal transposase.

The phenomenon of hybrid dysgenesis in relation to the introduction of Mutal element in *D. melanogaster*, which is naïve to Mutal element, was determined in the crosses RA2, and RA8; and their reciprocal crosses RC2 and RC8. In the crosses RA2 and RA8, males carrying Mutal element were crossed with the females from piggyBac lines that do not contain Mutal element. In the reciprocal crosses RC2 and RC8, females carrying Mutal 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 Mutal 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 Mutal 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 Mutal element in germline compared to the somatic cells in *Ae. aegypti* could be due to requirement of internal DNA sequence of the Mutal 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


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.


Table 4.1 Transformation rate from transformation and remobilization experiment for Mutal1 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 Mutal1 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 Mutal1 element to the piggyBac element.

<table>
<thead>
<tr>
<th>Plasmid Injected</th>
<th>Total embryos injected</th>
<th>Embryos Recovered</th>
<th>Total crosses</th>
<th>G0 Fertile Crosses</th>
<th>Mutal1 Transgenic line</th>
<th>pBac Transgenic line</th>
<th>Mutal1 Transgenic Rate</th>
<th>pBac Transgenic Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>pMuta1DsRedTFD+ pMuta1 helper</td>
<td>250</td>
<td>60</td>
<td>45</td>
<td>27</td>
<td>4</td>
<td>NA</td>
<td><strong>14.8%</strong></td>
<td>NA</td>
</tr>
<tr>
<td>pMuta1whiteTDF +pMuta1 helper</td>
<td>140</td>
<td>40</td>
<td>21</td>
<td>18</td>
<td>1</td>
<td>NA</td>
<td><strong>5.6%</strong></td>
<td>NA</td>
</tr>
<tr>
<td>pMuta13EHspBac + pBac3EHspMutal1</td>
<td>300</td>
<td>104</td>
<td>48</td>
<td>41</td>
<td>2</td>
<td>6</td>
<td><strong>4.9%</strong></td>
<td>14.8%</td>
</tr>
</tbody>
</table>


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.

<table>
<thead>
<tr>
<th>Lines</th>
<th>Generations Screened</th>
<th>Number of flies screened</th>
<th>Flies with Different Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>RA2 (M32-males X P32-females)</td>
<td>23</td>
<td>4991</td>
<td>Full body, Thorax, second last segment on the dorsal side of abdomen</td>
</tr>
<tr>
<td>RA8 (M32-males X P42-females)</td>
<td>23</td>
<td>6091</td>
<td>Full body</td>
</tr>
<tr>
<td>RC2 (M32-females X P32-males)</td>
<td>9</td>
<td>1954</td>
<td>Thorax</td>
</tr>
<tr>
<td>RC8 (M32-females X P42-males)</td>
<td>15</td>
<td>3424</td>
<td>Full Body, Abdomen</td>
</tr>
<tr>
<td>Total</td>
<td>16,460</td>
<td></td>
<td>5 new types of phenotypes for Mutal1 element (0.03%)</td>
</tr>
</tbody>
</table>
Table 4.3 Molecular characterizations of *Muta1* and *piggyBac* parental lines in *D. melanogaster*.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Chr</th>
<th>Break Point</th>
<th>TSD</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Muta1</em> white gene line</td>
<td>2R</td>
<td>3859229</td>
<td>GTTTTTGCAG</td>
<td>Repeat region</td>
</tr>
<tr>
<td><em>Muta1</em> DsRed line 1</td>
<td>2R</td>
<td>16676932</td>
<td>CTTAGCAGAG</td>
<td>Exon of CG30463-Glycosylation activity</td>
</tr>
<tr>
<td><em>Muta1</em> DsRed line 2</td>
<td>Flanking plasmid DNA</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td><em>Muta1</em> DsRed line 3</td>
<td>3L</td>
<td>12481711</td>
<td>GAGAGGCTG</td>
<td>Exon of CG10663-Serine type endopeptidase activity</td>
</tr>
<tr>
<td><em>Muta1</em> DsRed line 4</td>
<td>Flanking plasmid DNA</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td><em>Muta1</em> line M4</td>
<td>2R</td>
<td>9104602</td>
<td>ACATGTTTTG</td>
<td>Exon of CG8075/Van Gogh</td>
</tr>
<tr>
<td><em>Muta1</em> line M32</td>
<td>2L, X</td>
<td>19847516</td>
<td>AAAAAAAAAT</td>
<td>Intron of CG43720/Sickie gene</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12015935</td>
<td>ATATTTAGG</td>
<td>Intron of CG1803/Regucalcin gene</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>--</td>
</tr>
<tr>
<td><em>piggyBac</em> line P32</td>
<td>2L</td>
<td>14073174</td>
<td>TTAA</td>
<td>Intron of CG32975/Acetylcholine receptor</td>
</tr>
<tr>
<td><em>piggyBac</em> line P42</td>
<td>X</td>
<td>514785</td>
<td>TTAA</td>
<td>Exon of CG18104/Arginase or CG4262/elav gene</td>
</tr>
</tbody>
</table>
**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.

<table>
<thead>
<tr>
<th>New phenotype flies</th>
<th>Break point</th>
<th>New integration locations</th>
<th>TSDs</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fly with full body</td>
<td>1553430</td>
<td>Chr 3L</td>
<td>TGTGTTGGA (9 bp)</td>
<td>Intergenic between CG43248 and CG7450</td>
</tr>
<tr>
<td>expression (Figure A)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DsRed expression in</td>
<td>-----</td>
<td>pBac polyA</td>
<td>TTTTTTTTAT</td>
<td>-----</td>
</tr>
<tr>
<td>dorsal 2^nd last</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>segment (Figure B)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thorax (Figure C)</td>
<td>2741335</td>
<td>Chr 3L</td>
<td>GGTCAAAACC</td>
<td>Chr 3L Intron of CG32296/Mrf RNA polymerase transcription coactivator,</td>
</tr>
<tr>
<td></td>
<td>8151822</td>
<td>Chr 2R</td>
<td>GGTCAAATTC</td>
<td>Chr 2R Intron CG11546/Kermit – Regulation of signal transduction</td>
</tr>
<tr>
<td>DsRed expression in</td>
<td>20621644</td>
<td>Chr 3L</td>
<td>CTCACAGGG(9 bp)</td>
<td>Intergenic between Knirps gene and CG44684</td>
</tr>
<tr>
<td>abdomen (Figure D)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fly with full body,</td>
<td>12015935</td>
<td>Chr X (Parental integration)</td>
<td>GATTTTATAG (9 bp)</td>
<td>Intron of Regucalcin gene, Chr 3R in repeat region, Chr 3L integration is in intron of CG34418/sif - guanylyl receptor exchange factor. Involved in developmental stages in embryo, larvae, pupae and adults</td>
</tr>
<tr>
<td>bright expression in</td>
<td>1276726</td>
<td>Chr 3R,3L,2R,X &amp; Y</td>
<td>TTAAATAT (9bp)</td>
<td></td>
</tr>
<tr>
<td>labellum and genitalia (Figure E)</td>
<td></td>
<td>Chr 3L</td>
<td>TCCGACCG (8 bp)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5702591</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>New integration location</th>
<th>Break point</th>
<th>TSD</th>
<th>No. of flies characterized</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBac PolyA</td>
<td>---</td>
<td>TTTTTTTTTT (9 bp)</td>
<td>40</td>
<td>---</td>
</tr>
<tr>
<td>Plasmid DNA</td>
<td>----</td>
<td>GACGCAGT (8 bp)</td>
<td>19</td>
<td>----</td>
</tr>
<tr>
<td>Chr 3L</td>
<td>9458689</td>
<td>CGATGATAA (9 bp)</td>
<td>1</td>
<td>Exon of CG3552-Phosphorylase activity</td>
</tr>
<tr>
<td>Chr 2R</td>
<td>23613641</td>
<td>ATGTATAA (9 bp)</td>
<td>1</td>
<td>Intron of CG17664-Aquaporin protein</td>
</tr>
<tr>
<td>Chr 2R, 3R, X and Y Chr</td>
<td>------</td>
<td>GTGTTTAAC (9 bp)</td>
<td>3</td>
<td>Repeat region</td>
</tr>
<tr>
<td>Chr 2R</td>
<td>24294161</td>
<td>ATTTTT (5 bp)</td>
<td>3</td>
<td>Untranslated exon of CG3394/Mir-AMP binding enzyme</td>
</tr>
<tr>
<td>Chr 3L</td>
<td>3860980</td>
<td>GAGCTCGTG (9 bp)</td>
<td>1</td>
<td>Exon of CG0721/Arrowhead-Zinc ion binding, DNA binding</td>
</tr>
<tr>
<td>Chr 2L</td>
<td>20075421</td>
<td>GTCTACGTGTT (9 bp)</td>
<td>1</td>
<td>Exon of CG16784/purple-phenotype of this allele manifest in pigment cell</td>
</tr>
<tr>
<td>Chr 2L</td>
<td>3071477</td>
<td>AATAAATCC (9 bp)</td>
<td>1</td>
<td>Exon of CG0660/Toucan-Involved in Mitotic spindle formation</td>
</tr>
<tr>
<td>RE-pBac polyA--X Chr</td>
<td>12016142</td>
<td>TTTTTTTAT</td>
<td>1</td>
<td>Exon of CG1803/Regucalcin gene</td>
</tr>
<tr>
<td>Chr 2L</td>
<td>1925256</td>
<td>CAAATGATT (9 bp)</td>
<td>1</td>
<td>Intron of CG7337-dehydrogenase</td>
</tr>
<tr>
<td>X Chr</td>
<td>12016483</td>
<td>GGACTGGCA (9 bp)</td>
<td>1</td>
<td>Exon of CG1803/Regucalcin gene</td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>No. of mosquito adults in pool mating</th>
<th><em>Muta1</em> transgenic lines</th>
<th><em>piggyBac</em> transgenic lines</th>
<th><em>Muta1</em> transformation rate</th>
<th><em>piggyBac</em> transformation rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>75</td>
<td>3</td>
<td>13</td>
<td>4%</td>
<td>17.3%</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th><em>Aedes Muta1</em> transgenic lines</th>
<th>Break point</th>
<th>Location</th>
<th>TSD</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Muta1</em> line 1</td>
<td>-</td>
<td>Flanking plasmid DNA</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td><em>Muta1</em> line 2</td>
<td>-</td>
<td>Flanking plasmid DNA</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td><em>Muta1</em> line 3 and 4</td>
<td>269850 4</td>
<td>Supercontig 1.21</td>
<td>GATGCGCCT (9 bp)</td>
<td>AAGEO2001492.1</td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>Crosses</th>
<th>No. of screened mosquitoes</th>
<th>New phenotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1-MALES X P11-FEMALES</td>
<td>846</td>
<td>DsRed expression in legs, Midgut</td>
</tr>
<tr>
<td>M1-MALES X P14-FEMALES</td>
<td>130</td>
<td>0</td>
</tr>
<tr>
<td>M2-FEMALES X P9-MALES</td>
<td>178</td>
<td>Midgut</td>
</tr>
<tr>
<td>M2-MALES X P11-FEMALES</td>
<td>300</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1454</strong></td>
<td><strong>Two new phenotypes</strong></td>
</tr>
</tbody>
</table>
Figure 4.1 Schematic diagram of the reporter and helper line used to check Mutal mobility. The reporter line pMuta13EHspBac contains ends of Mutal element with piggyBac transposase, and the DsRed marker under the control of 3xP3 promoter. The helper line pBac3EHspBac contains the piggyBac element with Mutal 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
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.
**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 *MutaI* 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 Mutal1 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 Mutal 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 Mutal larvae Lane 4: heat shocked Mutal larvae. Lane 5-8 has RPS7 gene as a control form Orlando non-heat shocked, heat shocked, Mutal 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 Mutal$^{1}$ element in D. melanogaster. The Mutal$^{1}$ element integrates into itself into a piggyBac polyA region. piggyBac polyA was cloned in between the ends of the Mutal$^{1}$ 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 \textit{Muta1} element with DsRed expression in legs in \textit{Ae. aegypti}. The proposed pattern of the \textit{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 *XmaI* (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 *SacI* (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® Multiplex 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 Mutal1 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 Mutal transgenic lines A19, A50, A56, and A85, which were then mapped to the autonomous Mutal 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 Mutal element in D. melanogaster. piRNAs were not detected to Mutal element and thus, D. melanogaster genome was naïve to Mutal 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 Mutal 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 \textit{Mutal} 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 \textit{Mutal} transposase expression. The genomic site of integration of each \textit{Mutal} 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 \textit{Mutal} element. Nonetheless these three lines could serve as parental lines for subsequent remobilization of the \textit{Mutal} transposon.

\textbf{5.4.2 Remobilization of autonomous \textit{Mutal}.}

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 \textit{Mutal} 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 \textit{Mutal} 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 Mutal 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 Mutal transposase or to the ends of Mutal element, confirming that the D. melanogaster genome was naïve to the Mutal element.

In the library analysis of library A19 G11 (Figure 5.3.2), 98 piRNAs mapped to the autonomous Mutal plasmid. Out of these 98 piRNAs, 70.4% were antisense and 29.6% were sense piRNAs. Furthermore, 62% of these piRNAs mapped to the Mutal element, of the piRNAs that mapped to the Mutal element, 62.3% mapped to Hsp70 promoter, 9.8% mapped to the Mutal 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 Mutal 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 *Mutal* 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 *Mutal* plasmid. Out of these 220 piRNAs, 73% were antisense piRNAs and 28% were sense piRNAs. Further, 92% of the piRNAs mapped to the *Mutal* element. From the piRNAs that mapped to *Mutal* element, approximately 36% mapped to Hsp70 promoter, 35% mapped to *Mutal* 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 *Mutal* 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 *Mutal* 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 *Mutal* element. From the piRNAs that mapped to *Mutal* element, 60% mapped to Hsp70 promoter and approximately 27% mapped to *Mutal* transposase. The piRNAs that mapped to Hsp70 promoter, 86% of the piRNAs were antisense. The piRNAs that mapped to *Mutal* 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 *Mutal* 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 10th 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,
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 \textit{Muta1} element remobilization in \textit{D. melanogaster}. This provides an evidence for \textit{Muta1} element transposition into new locations in the genome, which suggest that host-silencing machinery might be able to detect the \textit{Muta1} element for suppression. Thus, small RNA libraries were made from generation 11 and 21 to detect piRNAs response to the \textit{Muta1} element. The small RNA libraries analysis, showed \textit{de novo} piRNAs that mapped to the autonomous \textit{Muta1} plasmid, which suggest that the \textit{Muta1} element might be remobilizing at a higher frequency in the germline of \textit{D. melanogaster}.

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

Previous studies with the \textit{P}-element, it was shown that piRNAs were produced for suppressing transcription of the \textit{P}-element, and integration of the \textit{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 \textit{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 \textit{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, \textit{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 be 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 \textit{Muta1} element in \textit{D. melanogaster} through amplification of piRNAs. Furthermore, the small RNA libraries were made from flies with expression in eyes for the \textit{Muta1} element. Flies that have lost expression of DsRed in eyes were observed in lines A50, A56 and A85, which suggest excision of the \textit{Muta1} element or suppression of the \textit{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 \textit{Muta1} element in \textit{D. melanogaster}.

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


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.

<table>
<thead>
<tr>
<th>Plasmid injected</th>
<th>Total embryos injected</th>
<th>Survived Inj.</th>
<th>Total Crosses</th>
<th>Fertile Crosses</th>
<th>Transgenic with Mutal</th>
<th>Transformation Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>pMuta1Auto</td>
<td>560</td>
<td>203</td>
<td>113</td>
<td>100</td>
<td>16</td>
<td>16%</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Lines</th>
<th>Location</th>
<th>Break point</th>
<th>TSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>pMuta1Auto</td>
<td>Chr 3L</td>
<td>12440359</td>
<td>GACCTAGG (8 bp)</td>
</tr>
<tr>
<td>line A19</td>
<td></td>
<td></td>
<td>Intergenic between Egy and Ego gene</td>
</tr>
<tr>
<td>pMuta1Auto</td>
<td>Flanking plasmid DNA</td>
<td>--</td>
<td>-</td>
</tr>
<tr>
<td>line A50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pMuta1Auto</td>
<td>Flanking plasmid DNA</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>line A56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pMuta1Auto</td>
<td>Flanking plasmid DNA</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>line A85</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5.3  New phenotypes observed in the autonomous \textit{Muta1} transformation experiment. Table shows number of flies screened in each cross and new phenotypes observed in each cross.

<table>
<thead>
<tr>
<th>\textit{Muta1} lines</th>
<th>No. of screened flies</th>
<th>New phenotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>A19</td>
<td>3570</td>
<td>-</td>
</tr>
<tr>
<td>A50</td>
<td>2589</td>
<td>-</td>
</tr>
<tr>
<td>A56</td>
<td>2548</td>
<td>Full Body, Thorax, Full body with a dot on side in abdomen</td>
</tr>
<tr>
<td>A85</td>
<td>2389</td>
<td>Full body</td>
</tr>
<tr>
<td>\textbf{Total}</td>
<td>\textbf{11,096}</td>
<td>\textbf{Three new phenotypes}</td>
</tr>
</tbody>
</table>

Table 5.4  Molecular characterizations of the new transposition events for the autonomous \textit{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.

<table>
<thead>
<tr>
<th>New integration locations</th>
<th>Number of flies characterized</th>
<th>Break point</th>
<th>TSDs</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromosome 2R</td>
<td>15</td>
<td>4033608</td>
<td>CTTGTTGCC</td>
<td>Intergenic between chromosome Lpin/CG8709 and Kermit/CG11546</td>
</tr>
<tr>
<td>Chromosome 3R</td>
<td>1</td>
<td>1447411</td>
<td>TCTGTCCTA</td>
<td>Repeat region</td>
</tr>
</tbody>
</table>
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.

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

<table>
<thead>
<tr>
<th>Library</th>
<th>Region</th>
<th>Normalized counts</th>
<th>Antisense</th>
<th>Sense</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A19 G11</td>
<td>Hsp70 promoter</td>
<td>58.46</td>
<td>56.92 (97.3%)</td>
<td>1.54 (2.7%)</td>
</tr>
<tr>
<td></td>
<td><em>Muta1</em> transposase</td>
<td>9.23</td>
<td>0</td>
<td>9.23 (100%)</td>
</tr>
<tr>
<td></td>
<td>Hsp70 polyA region</td>
<td>21.53</td>
<td>18.46 (85.7%)</td>
<td>3.07 (4.3%)</td>
</tr>
<tr>
<td></td>
<td>DsRed gene</td>
<td>3</td>
<td>0</td>
<td>3 (100%)</td>
</tr>
<tr>
<td>A19 G21</td>
<td>Hsp70 promoter</td>
<td>93.6</td>
<td>86.4 (92.3%)</td>
<td>7.27 (7.7%)</td>
</tr>
<tr>
<td></td>
<td><em>Muta1</em> transposase</td>
<td>24.5</td>
<td>7.27 (29.6%)</td>
<td>17.27 (70.4%)</td>
</tr>
<tr>
<td></td>
<td>Hsp70 polyA region</td>
<td>58.2</td>
<td>42.7 (73.4%)</td>
<td>15.45 (26.6%)</td>
</tr>
<tr>
<td></td>
<td>DsRed gene</td>
<td>6.3</td>
<td>0</td>
<td>6.3 (100%)</td>
</tr>
<tr>
<td>A50 G11</td>
<td>Hsp70 promoter</td>
<td>9.09</td>
<td>9.09 (100%)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td><em>Muta1</em> transposase</td>
<td>36.36</td>
<td>4.54 (12.5%)</td>
<td>31.8 (87.45%)</td>
</tr>
<tr>
<td></td>
<td>Hsp70 polyA region</td>
<td>4.54</td>
<td>4.54 (100%)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>DsRed gene</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A50 G21</td>
<td>Hsp70 promoter</td>
<td>81.39</td>
<td>80.23 (98.6%)</td>
<td>1.16 (1.42%)</td>
</tr>
<tr>
<td></td>
<td><em>Muta1</em> transposase</td>
<td>38.37</td>
<td>23.25 (60.6%)</td>
<td>15.11 (39.4%)</td>
</tr>
<tr>
<td></td>
<td>Hsp70 polyA region</td>
<td>59.3</td>
<td>48.8 (82.3%)</td>
<td>10.46 (17.6%)</td>
</tr>
<tr>
<td></td>
<td>DsRed gene</td>
<td>9.3</td>
<td>0</td>
<td>9.3 (100%)</td>
</tr>
<tr>
<td>A56 G11</td>
<td>Hsp70 promoter</td>
<td>109</td>
<td>106 (97.2%)</td>
<td>3 (2.8%)</td>
</tr>
<tr>
<td></td>
<td><em>Muta1</em> transposase</td>
<td>116.7</td>
<td>7.57 (6.5%)</td>
<td>109.09 (93.5%)</td>
</tr>
<tr>
<td></td>
<td>Hsp70 polyA region</td>
<td>34.8</td>
<td>13.6 (39%)</td>
<td>21.2 (61%)</td>
</tr>
<tr>
<td></td>
<td>DsRed gene</td>
<td>37.8</td>
<td>0</td>
<td>37.8 (100%)</td>
</tr>
<tr>
<td></td>
<td>Hsp70 promoter</td>
<td>Muta1 transposase</td>
<td>Hsp70 polyA region</td>
<td>DsRed gene</td>
</tr>
<tr>
<td>----------</td>
<td>----------------</td>
<td>-------------------</td>
<td>--------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>A56 G21</td>
<td>52.17</td>
<td>23.18</td>
<td>0</td>
<td>4.34</td>
</tr>
<tr>
<td></td>
<td>(86%)</td>
<td>(68.6%)</td>
<td>0</td>
<td>(100%)</td>
</tr>
<tr>
<td></td>
<td>7.24 (14%)</td>
<td>7.24 (31.4%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A85 G11</td>
<td>70.38</td>
<td>54.45</td>
<td>50</td>
<td>35.9</td>
</tr>
<tr>
<td></td>
<td>(86.4%)</td>
<td>(11.1%)</td>
<td>(53.5%)</td>
<td>(100%)</td>
</tr>
<tr>
<td></td>
<td>9.55 (13.6%)</td>
<td>48.4 (88.9%)</td>
<td>23.24 (46.48%)</td>
<td></td>
</tr>
<tr>
<td>A56 G21</td>
<td>75.38</td>
<td>40</td>
<td>30.76</td>
<td>12.3</td>
</tr>
<tr>
<td></td>
<td>(77.6%)</td>
<td>(26.9%)</td>
<td>(75%)</td>
<td>(100%)</td>
</tr>
<tr>
<td></td>
<td>16.9 (22.4%)</td>
<td>29.23 (73.1%)</td>
<td>10.76 (25%)</td>
<td></td>
</tr>
<tr>
<td>A85 G21</td>
<td>58.46</td>
<td>10.76</td>
<td>23.07</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(77.6%)</td>
<td>(26.9%)</td>
<td>(75%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>16.9 (22.4%)</td>
<td>29.23 (73.1%)</td>
<td>10.76 (25%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12.3</td>
<td>0</td>
<td>12.3 (100%)</td>
<td></td>
</tr>
</tbody>
</table>
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.

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

<table>
<thead>
<tr>
<th>Library</th>
<th>Sense RNA location (A10)</th>
<th>Length</th>
<th>Sequence</th>
<th>Antisense RNA loc (U1)</th>
<th>Length</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>A19-G21</td>
<td>521 (Hsp70)</td>
<td>24</td>
<td>AACAAGCAAAGGAGAAGCACGTGCT</td>
<td>530</td>
<td>26</td>
<td>TTGCGTGTGTGAATTGAGTGTGTGCT</td>
</tr>
<tr>
<td>A50-G21</td>
<td>472 (Hsp70)</td>
<td>23</td>
<td>CCGAGTAGATAAAATAGAGGGGCTT</td>
<td>482</td>
<td>23</td>
<td>TTATACTCCGGCGCTCTTTTCG</td>
</tr>
<tr>
<td>584</td>
<td>23 (Hsp70)</td>
<td>23</td>
<td>GCTAAACAATCTGCAGACTAGTA</td>
<td>594</td>
<td>26</td>
<td>GATTIGTTAGCTTGTCAGCTGCAGCC</td>
</tr>
<tr>
<td>1657</td>
<td>22 (Muta1 transposase)</td>
<td>25</td>
<td>CCGGTTTGGAGAATTGAAAAA</td>
<td>1666</td>
<td>26</td>
<td>TCAAACGCGCTGCTGATACTCGTACG</td>
</tr>
<tr>
<td>2462</td>
<td>23 (Hsp70 polA)</td>
<td>23</td>
<td>CCGGATGATAATAGAGGGGCTT</td>
<td>2472</td>
<td>27</td>
<td>TTATACCGCGCTGCAGGAGTTCGAG</td>
</tr>
<tr>
<td>2576</td>
<td>23 (Hsp70 polA)</td>
<td>27</td>
<td>TAAACACTCGGGGTACCCTAGA</td>
<td>2586</td>
<td>25</td>
<td>CAGATTGTTGATCCGACTGCTG</td>
</tr>
<tr>
<td>2580</td>
<td>27 (Hsp70 polA)</td>
<td>25</td>
<td>CAATGCGGTAGTACCCGAGCTGACG</td>
<td>2590</td>
<td>26</td>
<td>TAACCGATGTTGAATGACTGTCGACTG</td>
</tr>
<tr>
<td>2706</td>
<td>25 (DsRed gene)</td>
<td>25</td>
<td>TCGAGATCGAGGCGAGGGCCAGGGG</td>
<td>2715</td>
<td>21</td>
<td>CGATCTCGAGACTCTGGCGG</td>
</tr>
<tr>
<td>A56-G11</td>
<td>864 (Muta1 transposase)</td>
<td>24</td>
<td>TTGCGGACCTGACAACGAGGACG</td>
<td>875</td>
<td>28</td>
<td>ATGTGCAGAACATGACAGAAGCGGAGG</td>
</tr>
<tr>
<td>2348</td>
<td>24 (Hsp70 polA)</td>
<td>24</td>
<td>TTTGCTTTAAAAACACTGTTAGAC</td>
<td>2359</td>
<td>23</td>
<td>TTTTTAGCAAAACTCTACTCCCGG</td>
</tr>
<tr>
<td>A56-G21</td>
<td>577 (Hsp70)</td>
<td>24</td>
<td>TGAACAGCTAAACACCTTCAGT</td>
<td>589</td>
<td>24</td>
<td>TTACGTGTTGCGCTGCGGCTTGTG</td>
</tr>
<tr>
<td>599</td>
<td>21 (Hsp70)</td>
<td>21</td>
<td>ACGATCCTGCAGACTGATGAG</td>
<td>599</td>
<td>28</td>
<td>TGACGATTGTTGACGTGTCAGC</td>
</tr>
<tr>
<td>2524</td>
<td>22 (Hsp70 polA)</td>
<td>22</td>
<td>AACAGGTGCTAAGCGAGGACT</td>
<td>2535</td>
<td>29</td>
<td>TAGCGAGTCAGCGTTACTTTTGGTGGT</td>
</tr>
<tr>
<td>A85-G11</td>
<td>2576 (Hsp70 polA)</td>
<td>25</td>
<td>TAAACATCGGGTACCGCATGAG</td>
<td>2586</td>
<td>25</td>
<td>CAGATTGTTGAGCTGTTACGTCG</td>
</tr>
</tbody>
</table>
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.
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 *MutaI* plasmid. The autonomous *MutaI* plasmid contained *MutaI* left end (grey), the Hsp70 promoter (yellow), *MutaI* transposase (orange), Hsp70 polyA (blue), 3xP3-DsRed gene (red), DsRed polyA region (pink), *MutaI* right end (grey), and the plasmid backbone (white).
Figure 5.3.6 Coverage plot for library A56 G11, piRNAs mapped to autonomous Mutal plasmid. The autonomous Mutal plasmid contained Mutal left end (grey), the Hsp70 promoter (yellow), Mutal transposase (orange), Hsp70 polyA (blue), 3xP3-DsRed gene (red), DsRed polyA region (pink), Mutal right end (grey), and the plasmid backbone (white).
Figure 5.3.7 Coverage plot for library A56 G21, piRNAs mapped to autonomous Mutal plasmid. The autonomous Mutal plasmid contained Mutal left end (grey), the Hsp70 promoter (yellow), Mutal transposase (orange), Hsp70 polyA (blue), 3xP3-DsRed gene (red), DsRed polyA region (pink), Mutal right end (grey), and the plasmid backbone (white).
Figure 5.3.8 Coverage plot for library A85 G11, piRNAs mapped to autonomous Mutal plasmid. The autonomous Mutal plasmid contained Mutal left end (grey), the Hsp70 promoter (yellow), Mutal transposase (orange), Hsp70 polyA (blue), 3xP3-DsRed gene (red), DsRed polyA region (pink), Mutal right end (grey), and the plasmid backbone (white).
Figure 5.3.9 Coverage plot for library A85 G21, piRNAs mapped to autonomous Mutal plasmid. The autonomous Mutal plasmid contained Mutal left end (grey), the Hsp70 promoter (yellow), Mutal transposase (orange), Hsp70 polyA (blue), 3xP3-DsRed gene (red), DsRed polyA region (pink), Mutal 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, Mutal 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, Mutal 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, polytype=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, polytype=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
    letter_subject <- DNAString(paste(rep.int("N", width(reads)[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
  }
  # 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 <- read(seqs) # 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
segs <- 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=segs, 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")
    }
    
  }
}
Script 5.1 R scripts used to make coverage plots.

```r
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=xlim, 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-13302, ylim+1, col="red") # DsRed2 marker
  rect(3303, ylim-13580, ylim+1, col="pink") # DsRed2 polyA
  rect(3581, ylim-13916, ylim+1, col="lightgray") # right end
  rect(24, ylim-1363, ylim+1, col="lightgray") # left end
  rect(416, ylim-1593, ylim+1, col="yellow") # HSP70
  rect(606, ylim-12120, ylim+1, col="orange") # MutA1 tpase
  rect(2121, ylim-12627, ylim+1, col="cyan") # MutA1 tpase polyA
  #rect(z+2849, ylim-1, z+2935, ylim+1, col="green") # 3' UTR = green
  lines(c(runValue(negative), tail(runValue(negative), 1)), type="s", col="black")
  lines(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 Mutal1 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 Mutal1 into a target plasmid, pGDV1. Mutal1 has high somatic activity in D. melanogaster and Ae. aegypti. Zero-control transposition assays performed in absence of Mutal1 transposase showed no integration of the Mutal1 element ends in D. melanogaster and Ae. aegypti. Thus, the Mutal1 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 Mutal1 element flanked by an 8 bp as well as a 9 bp TSDs. The transposition frequency obtained for Mutal1 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 Mutal1 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 Mutal1 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, Mutal1 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 MutaI element. The phenomenon of hybrid dysgenesis could not relate to the MutaI 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 MutaI element. These small RNA libraries were analyzed for 23-30 nt small RNAs that were mapped to the autonomous MutaI 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 MutaI transposase showed abundance of sense piRNAs in all four transgenic lines from G11 and G21. This suggest that the MutaI 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 MutaI element into a germline piRNA cluster might support the hypothesis that the MutaI 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 MutaI 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 MutaI element. Only two overlaps were observed to the MutaI 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


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.


