UNIVERSITY OF CALIFORNIA
RIVERSIDE

Investigating Transposable Elements for Use in Dipteran Systems

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

Doctor of Philosophy

in

Entomology

by

Jennifer Alicia Wright

December 2011

Dissertation Committee:
Dr. Peter W. Atkinson, Chairperson
Dr. Alexander S. Raihkel
Dr. Shou-wei Ding

The Dissertation of Jennifer Alicia Wright is approved:

Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

I would like to begin by thanking my advisor Dr. Peter Atkinson. He has provided invaluable guidance and mentorship over the past four years. Working in his laboratory has provided me the opportunity to grow as scientist and equipped me with cutting edge molecular biology skills. I would also like to thank my committee members Dr. Alexander Raikhel and Dr. Shou-wei Ding. Both of them have been my instructors and have also provided excellent input which greatly enhanced the quality of my dissertation research. My undergraduate mentor Dr. Robert Kimsey is the person who first encouraged me to go to graduate school. He told me he believed that I would make a great scientist and for that I will always be thankful.

The Atkinson Laboratory members have played a critical role in my development as a scientist. First of all I found the love of my life, and now husband Joshua Knapp working in the lab. He has sat across from me for four years providing me guidance, support and love. Robert Hice has taught me more about being scientist than anyone else in my academic career. He taught me patience, precision and how to think critically when performing molecular biology. He also kept me sane when I wanted to break things out of frustration and taught me not to cry over spilled anti-bodies. I would also like to thank Dr. Peter Arensburger for his patience and guidance. He taught me bioinformatics and was always there to lend an ear when I needed him.

My family has been my foundation and strength. Thank you to my mother who always believed in me and never let me give up even when I felt like a failure. Thank you to Charlie Inman for always making me laugh and never letting me lose my perspective.

My dad and Ann Wright for always reminding me that graduate school was a phase and that no matter how hard it was or how poor I was I would make it through to the other end.

Friends have played a critical role in the completion of my dissertation. Jennifer Neelsen, Kendall Wright, Deane Zahn and Marka Latif have all provided me strength and encouragement that made getting my Ph.D. possible. Steve and Genene Staats provided financial support in my undergraduate career that made my quest to becoming a scientist possible. It takes a village to raise a Ph.D. and I am thankful for everyone involved in my journey.

ABSTRACT OF THE DISSERTATION

Investigating Transposable Elements
for Use in Dipteran Systems

by

Jennifer Alicia Wright

Doctor of Philosophy, Graduate Program in Entomology
University of California, Riverside, December 2011
Dr. Peter Atkinson, Chairperson

This thesis examines the mechanisms of class II DNA transposon activity and the genomic defense system against transposons in two Dipteran systems, *Drosophila melanogaster* and *Aedes aegypti*. The first half of the dissertation focuses on the affects specific amino acid changes have on target site specificity and transposition activity. Specifically it looks at *AeBuster1* and *piggyBac* mutants. None of the *Aebuster1* mutants were found to have increased transposition activity however several mutants had altered target site specificity. One p*iggyBac* mutant *hyPBase*, which has been shown to have increased activity in HeLA cells, mouse somatic cells and yeast, was studied in-depth. My study determined that *hyPBase* does not maintain its increased activity in Dipteran systems and significantly increases sterility rates when compared to wild type *piggyBac.* The second half of the thesis focuses on investigating the piRNA system in the important disease vector *Ae. aegypti*. The piRNA pathway is a Dicer independent small RNA pathway believed to be responsible for transposon control. In this study I examined the expression of 7 PIWI proteins and studied three of them, AGO3, PIWI2 and PIWI7 using co-immunoprecipitation and RNAi knockdown. Together all of these studies provide

insight into transposition mechanisms and transposon control in D. *melanogaster* and *Ae. aegypti*.

Table of Contents

List of Figures

List of Tables

Chapter 1
The Discovery and Investigation of Hyperactive Transposable Elements in Dipteran Systems

1.1 General Introduction to Transposable Elements

Transposons are mobile genetic elements that can integrate themselves as well as transgenic cargo into host genomes. Transposon vector systems are one of the most promising technologies for the non-viral integration of genes into eukaryotic genomes (Wilson *et al.* 2007). Developing and understanding transposon vectors is critical to their development as efficient genetic tools enabling important techniques such as gene and enhancer trapping to be used in order to discover gene function, especially in organisms in which there exist completed whole genome sequences. This review will examine the current state of Type II DNA transposon technology as it relates to use in higher eukaryotes.

1.2 Type II DNA Transposons

Transposons have a significant presence in many eukaryotes for example they represent approximately 50% or the human genome (Lander *et al.* 2001). Two separate classes of transposable elements have been identified to date: class I and class II. Class I elements transpose using reverse transcription of an RNA intermediate; these elements are commonly referred to as retrotransposons, and will not be discussed in this review (Walbot 1992). Class II transposable elements are composed of DNA flanked by terminal inverted repeats (TIRs). The integration of class II transposable elements typically

requires the use of a transposase enzyme which is encoded by the transposon itself. The transposase protein is the enzyme catalyzes transposition and hence is responsible for generating transposition events.

Class II transposases utilize a "cut and paste" mechanism whereby the transposase binds to the TIRs and catalyzes both the excision and integration of the transposon into the target site (Craig *et al.* 2002). The region into which a transposon is integrated is referred to as its target site. Class II elements are characterized by TIR that are recognized by the transposase protein. Autonomous transposons contain the open reading frame of a specific transposase within the two TIRs. TIRs vary in size and sequence between transposons however they are typically composed of sequences that are imperfect palindromes. These class II DNA transposons make up a smaller proportion of eukaryotic genomes compared to class I elements, composing approximately 3% of the human genome (Plasterk and Van Luenen 2002) This is due to the fact that class I elements move via a copy and paste mechanism replicating thus each time they move their copy number in the genome increases.

Increasing target site specificity is a major goal for those seeking to use transposable elements as genetic tools. *Mariner* elements including *Sleeping Beauty* and *Mos1* integrate into target sites with a TA dinucleiotide which is duplicated during insertion. This low specificity results in uncontrollable insertions into areas that may be undesirable to researchers or the *hAT* elements themselves. Random insertions can lead to mutagenesis and position effects on transgene expression (Handler and Harell 1999; Lorenzen *et al.* 2002). "Hotspots" of integration have been found which are regarded as

areas with favorable DNA conformations due to high levels of flexibility which increases the accessibility of the target sequence (Vigdal *et al.* 2002; Guerts *et al.* 2006).

To date the class II *piggyBac* element has proven the most promising as an efficient and precise transposon to be used as vector in gene therapy and insect transgenesis (Ding *et al*. 2005; Wilson *et al.*2007). *PiggyBac* is approximately 2.4 kb bases long with 13 bp terminal inverted repeats and an additional 19 bp internal terminal repeats. The *piggyBac* transposase uses a cut and paste mechanism to integrate into TTAA nucleotide sites (Wang and Frasier 1993). Some of the major advantages of using *piggyBac* is its ability to insert large (9.1-14.3kb) segments of genomic material without a loss of efficiency or alteration of the excision sequence (Li *et al.* 2001; Li *et al*. 2005; Ding *et al.* 2005). *PiggyBac* has proven its usefulness as a genetic tool in a wide variety of insects including but not limited to: *S. fruggiperda, Ae. aegypti, An. stephensi, An. albimanus, An. gambiae* and *T. castaneum* (Olsen *et al.* 1994; Gueiros-Filho and Beverly 1997; Goryshin and Reznikoff 1998; Handler *et al.* 1998; Sundaraajan *et al.* 1999; Grossman *et al.* 2000).

PiggyBac is especially promising for several of its key characteristics. First the element preferentially inserts within genes, with 50% to 67% of insertions occurring within transcriptional units (Ding *et al.* 2005; Wilson *et al.* 2007; Wang *et al*. 2008). Perhaps most importantly to this study, is that the element is not sensitive to over-production inhibition in which elevated levels of the transposase causes decreased transposition (Wilson *et al.* 2007). And lastly, a unique characteristic of *piggyBac* is that it excises cleanly from a donor site, leaving no "footprint" which makes it an excellent

candidate for use in higher order eukaryotes (Carey *et al.* 1989; Fraser *et al.* 1995; Wilson *et al.* 2007).

The Class II element *Hermes* can be used as a viable vector in mosquito germ-line transformation. *Hermes* was used in the transformation of *Anopheles gambiae* to block the transmission of the malaria pathogen through the midgut and salivary glands (Ito *et al.* 2002; Zhao and Eggleston 1998). The *Hermes* element is a 2,749 bp long, has 17 bp terminal inverted repeats and encodes a 70 kDa transposase (Warren 1994). It uses a cut and paste mechanism where the DNA flanking the excised element temporarily forms a hairpin before the gaps are repaired. The *Hermes* crystal structure shows it to contian a functional heterodimer and forms an hexamer that interacts with the host DNA (Hickman *et al.* 2005). *Hermes* is a desirable model transposon for several reasons. First it has proven effective in transforming a variety of organisms providing proof of its ability to be useful across a wide range of organisms (Medhora *et al.* 1991; Sakar *et al.* 1997; Peloquin *et al.* 2000; Allen *et al.* 2001). Second, the crystallized structure has been partially solved resulting in a deeper understanding of the mechanisms involved with its oligermorization and interactions with host DNA (Perez *et al.* 2005; Hickman *et al.* 2005). Despite these advantages *Hermes* has been found in mosquitoes to insert flanking plasmid DNA along with the transgene of interest. Therefore understanding how it recognizes target DNA and how it excises imprecisely in mosquitoes could provide valuable information concerning how transposons interact with their host.

The extinction of active class II DNA transposons in many vertebrate genomes may be considered an advantage when using them to perform mutagenesis in mammals.

This is could be due to the fact that the host may have few defense mechanisms in place to control their mobility. Because natural selection processes have been able to reduce transposon movement in some genomes, it is possible that transposition activity can be restored through mutagenesis. Indeed, there is a precedent for increasing transposition rates in class II transposons through mutagenesis (Baus 2005). An active *Tc1*/*mariner* transposon, synthesized from the alignment of inactive transposons from the *salmonid* subfamily of elements, was resurrected to form the transposon *Sleeping Beauty* (Ivics *et al.* 1997). Endogenous DNA transposons from this family are completely inactive in vertebrate genomes and are thought to have lost functional ability many millions of years ago, with only their inactive remnants found throughout vertebrate genomes. This restoration of activity shows that silencing through natural selection can be reversed through the manipulation of a transposon's sequence (Yant *et al*. 2004). It therefore stands to reason manipulation of currently active transposable elements may lead to higher transformation rates.

1.3 Transposons in Genetic Modification

Elucidating basic knowledge and learning to effectively manipulate transposons will aid in the creation of novel genetic tools used to combat significant human health issues. Transposons are demonstrating great potential in insect transgenesis and disease prevention (O'Brochta & Atkinson 1996). For example, each year, approximately 300 to 500 million human malaria infections lead to over one million deaths (Snow *et al.* 1999). Malaria is most prevalent in sub Saharan Africa in poor regions lacking proper disease

management programs. The World Health Organization has recognized malaria control as one of its top priorities emphasizing the need for a global effort to control the disease (Roll Back Malaria 2000). The ability to inhibit the transfer or hosting of the malarial parasite by the mosquito vector could prove more efficient in regions where mosquito pest management programs are unable to provide effective suppression due to pesticide resistance, pesticide cost, or political instability. Using genetically altered mosquitoes alleviates the issue of pesticide resistance thus offering a long term solution (Robinson *et al.* 2004).

Transposable elements have already been used in germ cell and somatic cell transformation across a large range of eukaryotic organism, including *An. gambiae,* a major vector of malaria in sub Saharan Africa (Ito *et al.* 2002; Zhao and Eggleston 1998). Mosquitoes have been created that are able to block the movement of a *Plasmodium* through the insect midgut inhibiting the transmission of the parasite (Ito *et al.* 2002; Ghosh *et al.* 2001). These studies were performed in *Anopheles stephensi*, a less important vector of malaria, using rodent malaria thus limiting the application of these studies to the realm of proof of principal. *An. gambiae* is difficult to transform do to the size and nature of its chorion and in order to reach full potential an increase in transformation rates must be achieved. Transformation rates of mosquitoes have been reported at 4% and 8% for *mariner* and *Hermes* respectively in *Aedes aegypti* (Coates *et al.* 1998) but these levels are higher than rates many other labs report anecdotally. Often these rates are much lower especially when dealing with other species such as *Anopheles stephensi* and *Anopheles gambiae* anecdotally (personal communication).

Developing and understanding gene vectors is critical to the success of insect transgenesis. However, currently the field is limited by low transformation frequencies and unpredictable integration sites (Glover *et al.* 2005). An increase in transposition rates and predictability of transposon target sites 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 every year. For example an advance in the field of insect transgenics could result in the creation of mosquitoes refractory to a disease.

1.4 Transposable Elements as Genetic Tools

Forward genetics in *D. melanogaster* the class II *P*-element transposon was used to create an insertion library resulting in the creation of numerous strains with a single stable *P* transposon insertion. Mutagenesis was initiated by crossing two strains; one strain contained the *P*-element transposase and the other strain contained the non-autonomous transposon; a transposon that lacks a transposase between the TIRs (Cooley *et al.* 1998. Random *P*-element transpositions were captured in individual stocks that no longer contained the transposase thus the integrations remained stable. This method was used to construct 1,300 single *P*-element insertion stocks. The insertion of a single *P* transposable element into the genome creates a disruption which may impact the function of nearby genes. For example, an insertion in the promoter region of a gene can knock out its function. This library of single-element insertion strains is maintained by a stock center and has allowed the structure and function of *Drosophila* genes to be readily

correlated (Cooley *et al.* 1988). Unfortunately insertion libraries such as the one described for *Drosophila melanogaster* library cannot currently be produced in some other organism such as mosquitoes due to low transformation rates and general inefficiency of transposons in these systems. If there were transposons that could be used efficiently in mosquitoes a great deal of functional genomic data could be elucidated.

Enhancer trapping is another genetic strategy that utilizes transposable elements to gain valuable gene expression and function data in fruit flies. The GAL4/ UAS system is a powerful tool that utilizes a transcription factor and promoter from yeast. The system is comprised of two parts: the GAL4 gene which encodes a yeast transcription factor that is able to specifically bind the second Upstream Activation Sequence (UAS). The GAL4 promoter is driven by endogenous weak promoters and UAS is fused to a reporter gene such as green florescence protein (GFP). Transposable elements are used to integrate the GAL 4 transcription factor and UAS promoter into two separate lines of an organism, the activator and affecter lines respectively. When the lines are crossed the reporter gene is expressed in tissues in which the promoter is expressed endogenously. For example, if the GAL4 transcription factor inserts near the promoter of a gene involved with the nervous system one would expect to see expression of the UAS linked maker (such as GFP) in nervous tissue. The Gal4/ UAS system thus provides valuable expression data in regards to the driver gene. This system is commonly used in *Drosophila melanogaster* and has been used in other organisms such as *Bombyx mori* , *Xenopus* and zebrafish (Scheer and Campos-Ortega 1999; Duffy *et al.* 2002; Hartly *et al.* 2002; Imamura *et al.*

2003; Davison *et al.* 2007). In *Drosophila* the *P*- element is commonly used where as *piggyBac* is used in other organisms such as the silk worm.

More recently a large-scale mutagenesis of *Tribolium castaneum*, the red flour beetle was undertaken to create a similar insertion library as that created by Cooley *et al.* 1988 (Trauner *et al.* 2009). However since *P*- element has only been shown to function in *D. melanogaster,* the class II DNA transposase *piggyBac* was used. The 3xP3 promoter was inserted using EGFP as a marker for expression (Berghammer *et al.* 1999). The 3xP3 promoter is responsive to nearby chromosomal enhancers thereby allowing for the combination of insertional mutagenesis and enhancer trapping. Researchers were able to produce 6,500 *piggyBac* insertions of which 412 were recessive lethal, 75 were semi-lethal and 505 showed enhancer trap patterns. This study demonstrated that classII DNA transposable elements are capable of acting as powerful genetic tools in coleopteran systems.

Transposable elements are also being utilized in RNAi reverse genetics studies (Shin *et al*. 2003). *Ae. aegypti* has been stably transformed using the *piggyBac* transposon with constructs designed to trigger an RNAi response (Bian *et al.* 2005). The constructs were designed to form a long double stranded region that was used to silence AaRel1 (Bian *et al.* 2005). The ability to integrate RNAi constructs into mosquitoes is dependent on appropriate endogenous promoters, genes of interest and transposon integration. Gene promoters are now more accessible with the sequencing of the *An. gambiae* and *Ae. aegypti* genomes however there is little functional annotation available limiting the usefulness of the genome sequences (Holt *et al.* 2002; Nene *et al*. 2007). The

techniques described above are hampered in non-model organisms due to low transformation frequencies and the fact that few tissue specific and developmentally specific promoters have been identified for organisms such as mosquitoes. In conclusion the genetic "tool box" of the mosquito geneticist would be greatly strengthened by higher transformation rates and transposons with increased mobility once they are integrated.

1.5 Mutagenesis of Transposons with the Goal of Hyperactivity

Previous studies have attempted to create hyperactive transposases for use as genetic tools however none of these studies have provided a hyperactive transposases suitable for use in insect systems (Pledger and Coates 2005; Lampe et al. 1996). In 1996 Lampe *et al.* created *Himar1 mariner* transposases that were hyperactive using random mutagenesis and screening for activity in *E. coli.* They discovered that by changing the residues E127 and E244 in the *Himar1* transposase they could increase activity between 4- and 50- fold. However the *Himar1 mariner* transposases that proved hyperactive in *E. coli* assays lost all activity when moved into eukaryotic systems. Pledger and Coates (2005) attempted to create *Mos1* transposases with increased activity in *Aedes aegypti*. The researchers aligned the *Mos1* and *Himar 1 mariner* transposase used previously in Lampe *et al.* (1996) and made the identical changes in *Mos1* at residues E137 and E267. While the mutated *Mos1* transposases that were hyperactive in *E. coli* did not completely lose function in *Ae. aegypti* embryos, they did not show any signs of increased activity. The transposition assays in *Ae. aegypti* did show that the mutated *Mos1* transposases had an increase frequency of canonical or perfect transposition events.

Baus *et al*. (2005) generated a hyperactive *Sleeping Beauty* transposase through targeted mutagenesis. The study used a phylogenetic based mutagenesis approach wherein *Sleeping Beauty (SB)* 10 was aligned with other members related to the transposases in order to determine the most conserved residues. This study was the first to compare some of the previously generated hyperactive *SB* transposases directly in order to optimize activity levels. The results of this study determined that some combinations of mutations can have synergistic affects that increase activity while other combinations had a negative impact (Baus et al 2005). Cui *et al* (2002) also used mutagenesis to alter the activity of *SB*. Through alteration of the direct repeats (DR) and TIR they were also able to demonstrate increased transposition activity. Therby illustrating that mutation of the transposase protein is not the only strategy that can be used to affect transposition activity.

Yant *et al.* 2004 used an alanine scanning approach in *SB* to determine how mutations in specific domains within the transposase affect transposition. The study examined 95 different residues within the transposase using a large-scale mutational analysis and found that amino acid substitutions within either the putative oligomerization domain (L11A, L18A, L25A, and L32A) or the nuclear localization signal (K104A and R105A) severely impede the transposition capacity of the transposon (Yant et al. 2004). Alternatively, each of the 10 alanine substitutions within the DNA-binding domain greatly enhanced *SB*'s transposition activity in mammalian cells. These mutations acted synergistically when combined significantly improving the transposition

rate of *SB*. These studies show that DNA transposons can be altered to enhance their usefulness as genetic tools via mutagenesis in model organisms.

1.6 Modification of Target Site Recognition via DNA Binding Domains

One method used to increased target specificity is to fuse a DNA binding domain (DBD) onto a transposase protein. These DBDs are expected to "direct" the transposase to insert near or in a specific target site. There are several strategies that involve DBD which include: 1) adding a DBD directly to the transposase, 2) using two fused DBD where the first DBD binds to a specific genomic integration site and the other DBD interacts with a sequence inserted within the transposase and 3) an interaction between the transposase and transposase binding protein that is fused to DBD. (Ivics *et al.* 2007) One major drawback to this technology is that transposases tend to lose activity once fused to a DBD. The length and composition of the linker between transposase and DBD can also have affects on activity and add another level of complexity.

Feng *et al.* (2009) provides a recent example of fusing a DBD directly to a transposase. This study used the ISY100 transposase which a member of the *Tc1/mariner* family fused to the Zinc binding domain of Zif268, a mouse transcription factor. *Tc1/ mariner* transposons prefer to integrate at TA dinucleotides, but by pairing the transposase with the DBD of Zif268 the researcher were able to generate a transposase with increased target site specificity (Feng et al. 2009). However the enzymatic activity was reduced between 2-100 fold depending on target site construction and the linker

used. The study was performed in *E. coli* and thus the usefulness of the ISY100-Zif268 in eukaryotic systems has yet to be determined.

Margathavally *et al.* (2006) fused the GAL4 transcription factor DBD to the *Mos1* and *piggyBac* transposases and tested the altered transposase in *Ae. aegypti* embryos. Interplasmid transposition assays were performed using a pGDV1 target plasmid which contained the UAS sequence. The group found that the GAL4-*Mos1* transposed at a frequency 12.7 times greater than controls and inserted in the same location in the target plasmid 96% of the time. The control in this experiment was a pGDV1 target plasmid lacking a UAS target site that is therefore the GAL4UAS interaction was absent. *PiggyBac* had and 11.6 fold increase in activity with 67% of the integrations occurring at one site. The increased transposition frequency along with improved target site specificity makes these transposon systems valuable as genetic tools and demonstrates how DBD can be used to alter transposon target site selection in eukaryotes. However these researchers did not show how target site specificity is affected when integrating into the genome due to the fact that they only performed an interplasmid assays.

Yant *et al.* (2007) took the next step (along with Wilson *et al.* 2005 and Ivics *et al.* 2007) and looked at genomic integrations of fusion transposases. *Sleeping Beauty* was fused to the E2C DBD as well as to GAL4. The fusion proteins were tested in human cells and did demonstrate some levels of increased specificity. The fusion protein with the most promise showed an increase of 8 fold more transposition events in the target region than the control. However the transposition rate was far below what would be needed for use in gene therapy treatments.

The Gal4-UAS system has also been used in *Ae. aegypti* (Kokoza & Raikhel, 2011). This study used the 5' upstream region of the *Ae. aegypti* vitellogenin gene and created a construct with Gal4 and an EGFP reporter gene linked to a 3xP3 promoter. The construct was inserted into the genome using *piggyBac* (Kokoza & Raikhel, 2011*)*. A separate UAS/EGFP line was also engineered in order to evaluate the binary Gal4/UAS system (Kokoza & Raikhel, 2011). This study was able to achieve high levels of tissue-, sex- and stage-specific expression of the reporter gene. Additionally the study used in vitro organ culture experiments as a means of studying hormone regulation of gene expression (Kokoza & Raikhel, 2011). Together these experiments demonstrate the refinement of the UAS/GAL4 system in a mosquito system. 9090c

1.7 Other Genetic Technologies used for Insect Transformation

The *Streptomyces* phage ΦC31 is a site specific recombinase. ΦC31 is able to mediate unidirectional integration of genomic DNA at a specific attP sequence using an integrase enzyme (Novick *et al.* 1967). This integrase allows for the integration of a transgene using a donor plasmid which contains the attB recognition site. The attB recognition site docks to the attP located in the genome of interest and the ΦC31 integrase catalyses the integration (Pattee *et al*.1977; Lee and Iandolo 1986; Lee and Iandolo 1988). Once the transgene has been inserted into the attP site two new attL (left) and attR (right) sites are generated due to the disruption of the attP site which stops any additional integrase action (Patee *et al*. 1977). The major benefit of using an integrase is that the integration site is known removing the need to map where the insertion is located

and eliminating positional affects that plague transposon mediated integration. One major limitation of this technology is that an attP site must be either available in the genome or inserted through transgenesis, thus organism which are hard to transform using traditional transposon mediated transgenesis would be difficult to use (Raush and Lehmann 1991).

ΦC31 has been successfully used in D. melanogaster cell culture (S2 cells) and embryos. In S2 cells that were transfected with two plasmids, a donor (attB) and target (attP) intramolecular recombination events occurred at a frequency of 47%. Once it was established that ΦC31 worked in cell culture two lines of *D. melanogaster* were engineered to contain an attP site using a *P*- element. A pUASTB plasmid, containing a marker, attB recognition site and transgene of interest, was co-injected with ΦC31 integrase mRNA into the embryos. The donor was then integrated into the preplaced attP sequence by ΦC31 integrase into the germ line of *Drosophila melanogaster*. The microinjection resulted in up to 55% of fertile adults producing transgenic offspring with 100% precise integration into the attP site. (Groth *et al.* 1994, Fish *et al.* 2007)

ΦC31 can also be used in conjunction with recombinase-mediated cassette exchange (RMCE) (Wimmer *et al.* 2005; Bateman *et al.* 2006). RCME involves the exchange of two sequences instead of simply an insertion. This is an important distinction because it allows for the insertion of a transgene with a marker since the loss of the first cassette can result in the necessary phenotypic change for selection. Another advantage of using RCME is that it only integrates the sequence within the cassette, unlike ΦC31 which integrates the plasmid backbone and selection marker as well. The target plasmids were constructed in such a manner that there were two attP sites in the genomic target

with a gene in between, such as the mini-white marker (Bateman *et al.* 2006). The donor plasmid had two attB sites with a different gene or marker inside side. When the two sets of attP and attB sites lined up the ΦC31 integrase mediated the transfer (Bateman *et al.* 2006).

ΦC31 integrase has also been used in systems outside of *Drosophila*. Nimmo *et al.* 2006 tested performed site-specific genetic engineering of *Aedes aegypti*. Five transgenic strains containing the attP site were created using *piggyBac* as the vector. Transformation frequency was based on individual males and ranged between 16.7-31.8%. Transformation of females was also observed. Non-specific integration into pseudo-attP sites in the *Ae. aegypti* genome occurred at a frequency of 1.8%. It is interesting to note that this study also investigated the use of CRE/Lox system and found that it failed to produce stable transformants. The ΦC31 system has also efficiently been used in human cell lines (Thyagarajan *et al.* 2001). Studies designed to assess the risk of ΦC31 integration into pseudo-attP sites in the human genome have concluded that the system is safe for use in human gene therapy *(*Chalberg *et al. 2006).*

Rong and Golic 2000 developed a method for the utilization of an alternate site directed recombinase in *D. melanogaster*. Their method was comprised of three parts: a transgene that expresses a site-specific recombinase (FLP), a transgene that expresses a site-specific endonuclease (I-SceI), and a transgenic donor construct that carries recognition sites for both enzymes and DNA from the locus to be targeted. Broken-ended extrachromosomal DNA molecules were used to produce homology-directed changes in a target gene, which in this case was the *white* gene (w+). Once it was determined that the

I-SceI could make double stranded DNA breaks, a transgenic line containing a mutated *yellow*' body mutant gene was made and used to target the recessive *yellow* body gene. The FLP recombinase was able to catalyze the recombination between copies of FLP target, FRT, and the I-SceI recognized and cut its 18 base pair recognition sequence. Double stranded DNA breaks were formed allowing for recombination and insertion of the mutated *yellow*' body gene masking the recessive phenotype. Thus the FLP recombinase in conjunction with I-SceI was able to engage in site specific recombination. However, it is important to note that first the transgenic enzymes (FRT and I-SceI) had to be introduced to the fly genome using *P*-element mediated transformation demonstrating once again the importance of these transposons.

This technology has not been limited to *D. melanogaster*. The *cre-loxP* site-specific recombination system has been used in *Ae. aegypti*. Jasinkiene *et al.* (2003) were able to remove a marker gene from a transgenic strain carrying the *cinnabar* gene. The class II DNA transposon *Hermes* was used to insert a construct containing the FRT and *lox P* recognition sites. The *cre* recombinase was able to successfully remove the *cinnabar* gene at a frequency of 33%. Interestingly, the FLP/ FRT system did not appear to function in *Ae. aegypti* (Jasinkiene *et al.* 2003). It is important to note that while excision did occur the researchers did not accomplish insertion of a novel gene via the recombinase system limiting the usefulness of this technology in mosquitoes.

1.8 Discussion of Transposons as Genetic Tools

There are two common problems in the fields of insect transgenesis and gene therapy: 1) the lack of understanding and control over target site preference and 2) inefficient transformation rates. The creation of transposases that are able to be used in an efficient and predicable manner would accelerate the fields of gene therapy and insect transgenesis in many beneficial ways. For example, an increased understanding of transposons and their underlying mechanisms will result in faster and more effective transfer of genes into new systems in general, thereby increasing the rate of turnover between the discoveries of genes *ab initio* and the understanding of their function in an organism

Zinc finger DBD have shown great potential for increasing target site specificity however the aberrant affects on transposition activity provide yet another hurdle to overcome. The GAL4 *Sleeping Beauty* fusion transposases demonstrate that while it may be the norm to lose activity with fused transposases this is not always the case. Thus there is hope that hyperactive transposons can be tethered to DNA binding domains in order to create transposons suitable for use in gene therapy and with increased transformative capability. Experimentation with linker composition and length will likely contribute to the efficient use of transposases fused to DBD.

The phage integrase ΦC31 has demonstrated its usefulness in a variety of systems including *D. melanogaster,* mosquitoes and human cell lines. The unidirectional nature and precision make it an excellent tool for gene replacement. Strides are also being made in the use of recombinase technology such as the CRE/Lox and FLP/ FRT systems, yet

little progress has been made in increasing their usefulness as tools in insects beyond *Drosophila*. Further investigation will be required before these genetic tools can be used efficiently. Both the phi C31 integrase and the recombinases require an initial insertion of a specific recognition site, thereby still requiring the initial use of class II DNA transposable elements. Hence the need to continue the search for means of making more efficient transposons.

The future of transgenesis is exciting and constantly changing. This review has demonstrated that advances in field of transformative genetics tools are being made, but there is still room for improvement. The search for hyperactive transposases that can be coupled to site-specificity altering DBD may yield exciting tools that will aid in the enhanced efficiency of other tools such as phage 1 integrases.

An increase in transposition rates and predictability of transposon target sites would greatly propel research in both fields allowing for the acceleration of important gene integration studies that could save the lives of millions of people every year. Higher rates of turn over between gene discovery and application of those genes for use in research would save valuable time and money in the current quest for treatment solutions to diseases such as cancer, diabetes, high blood pressure, heart disease, cystic fibrosis, hemophilia and malaria. Therefore focusing on improved transposons should remain a high priority for the scientific community.

Introduction Part II: Investigating the mechanisms that underlie transposable element control in Dipteran systems

1.9 Introduction to Transposon Control Systems

Transposons are often seen as genome parasites. These mobile elements are able to disrupt gene function and alter regulatory cascades via insertions in protein coding regions of the genome. Transposons can also result in large scale damage such as chromosome breakage and genome rearrangement (McClintock 1951). However insertions can have beneficial effects on gene regulation and genome organization as well providing valuable variation as a result of altered gene expression such as silencing or up-regulation (Feschotte 2008). Thus an adaptive control system is necessary to preserve the delicate balance between the beneficial and maladaptive affects transposons have on the genome

1.10 Prevalence of Transposons

A transposon control regime would have to be able to distinguish between host genome transcripts and those of transposable elements. Mobile elements have a variety of structures and mechanisms for moving within the genome. As previously described transposons can generally be placed into two main classes: class I retrotransposons and class II DNA transposons. Within these two classes there are multiple families each with members that posses unique TIR and vary greatly in length (Craig 2002). Endogenous retroviruses, such as *gypsy* and *ZAM*, are another group of mobile genome parasites that have the ability to move in a fashion very similar to retrotransposons (Desset *et al.* 2003).

Thus a genome regulation mechanism for transposons would have to be able to respond to wide range of ever evolving invaders.

Transposons are present in all eukaryotic genomes so far examined to varying degrees. Even within the class Insecta there is a great deal of diversity in transposon loads. For example in *T, castaneum 1*6% of the genome is composed of transposable elements, *D. melanogaster* 5%, *Bombyx mori* 12%, *An.s gambiae 24*% and *Ae. aegypti* 47% (*Tribolium* Genome Sequencing Consortium 2008; Bergmen *et al.* 2006; Mita *et al*. 2004; Aultman *et al.* 2002; Nene *et al*. 2007). The ubiquitous nature of transposons across the Eukaryotic kingdom has lead to the search for a genome wide control system for these elements. The small RNA pathways may represent a mechanism for eukaryotic suppression of transposable element movement.

1.11 Small RNA Biogenesis and Function

Small silencing RNA were first discovered in 1993 and since that time multiple classes of small RNA have been identified including siRNA, microRNA and piRNAs (Wightman *et al.* 1991; Zamore *et al.* 2000;Quintana *et al.* 2001; Aravin *et al.* 2001;Vagin *et al*. 2006). 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. Fire and Mello are credited for the 1998 discovery that double stranded RNA (dsRNA) is able to trigger silencing in *Caenorhabditis elegans* (Fire *et al.* 1998). With this discovery began the race to understanding the underlying mechanisms that direct this silencing. Between 1998 and

2002 it was discovered that in the canonical RNAi pathways the Dicer protein works in a process in which small RNAs are cut from long double stranded RNA. The Dicer produced small RNA then interacts with a second complex composed of Argonaute proteins that together form the RNA-induced silencing complex (RISC). In the end a single stranded small RNA is generated that acts as a template for target specificity (Zamore *et al.* 2000; Hammond *et al.* 2000; Bernstein *et al.* 2002; Elbashir *et al.* 2001a; Elbashir *et al.* 2001b; Hutvágner *et al.* 2001). It was also determined that the microRNA and the siRNA pathways utilize the same proteins and have a similar biogenesis pathway yet have separate regulatory affects on the genome (Hutvágner *et al.* 2001*;* Grishok *et al.* 2001; Ketting *et al.*2001*).* Figure 1.1 is a schematic diagram of the siRNA, microRNA and piRNA pathways in *Drosophila* from Ghildiyal and Zamore 2009 as they are presently understood.

1.12 Small RNA in *Drosophila*

Small RNAs in *Drosophila* were first characterized in 2003 by Aravin *et al.* 2003 and Brennecke 2003). This group identified three separate classes of small RNAs: siRNA, microRNA and repeat associated small interfering RNAs or rasiRNA. The repeat associated RNAs were suggested to play a role in transposon regulation and had been first been identified (Aravin *et al.* 2001). Tissue specific sampling of these rasiRNAs allowed for the discovery that these elements are most active in the germ-line and early embryonic phase. In 2006 the rasiRNAs were reevaluated and renamed piRNAs (Aravin *et al.* 2006).

1.13 piRNAs in *Drosophila*

piRNA biogenesis is distinct from the siRNA and microRNA biogenesis in that it is a dicer-independent pathway. piRNAs are 24-30nt small RNAs that bind PIWI proteins (Aravin *et al.* 2001). The PIWI class proteins were first discovered over ten years ago in *D. melanogaster* mutants and are members of the Argonaute clade (Lin and Spradling 1997). Argonaute proteins are characterized by the presence of a PAZ and PIWI domain which function by forming a channel in which a single-stranded small RNA transcript is held at each end by one of the domains (Song *et al.* 2003). Specifically the proteins Piwi, Aubergine (AUB) and Argonaute3 (AGO3) are actively involved in the piRNA pathway. The lack of production of any of three PIWI proteins results in deregulation of transposons in ovaries (Brennecke et al 2007). Similar studies have not been performed in *Ae. aegypti* due to the difficulty of genetically manipulating this medical significant organism. Another distinctive feature of piRNAs is that they are 2′-*O*-methylated at their 3′ termini, unlike miRNAs but similar to siRNAs in flies (Horwich *et al.* 2007; Vagin *et al.* 2006; Saito *et al.* 2007). Currently piRNAs are believed to be involved in transposon and endogenous retrovirus control in both the germ-line and somatic cells (Chung *et al*. 2008; Brennecke *et al.* 2008; Aravin *et al.* 2007; Malone *et al.* 2009; Li *et al.* 2009).

The biogenesis pathway of piRNAs is the least understood of the current three small RNA pathways. The generation of piRNAs in *D. melanogaster* is believed to be controlled by two separate biogenesis pathways: somatic and germ-line. Somatic suppression of piRNAs is controlled through what is known as primary biogenesis pathway which is described as the "the linear biogenesis of piRNA from transcripts to

complexes with PIWI proteins" (Senti and Brennecke 2010). The primary biogenesis pathway relies on only the PIWI protein (Malone *et al.* 2009; Li *et al*. 2001). The precise location of where somatic piRNA production occurs is unknown. It has been shown that PIWI protein accumulates in the nucleus however an overwhelming proportion of piRNAs originate from 3' UTR regions suggesting the processing begins in the cytoplasm (Saito *et al* 2009, Robine *et al.* 2009). In *D. melanogaster* somatic piRNA appear to originate predominately from loci of densely packed transposon fragments such as *flamenco* (Brennecke *et al*. 2007). It is currently believed that single stranded transcripts from these regions are loaded onto Piwi which is then able to silence transposon transcripts. However Piwi does not act alone in this process. It has been demonstrated that Armitage (ARMI), Zuchinni (ZUC) and Female Sterile 1 (FS (1)) proteins are necessary for primary synthesis of piRNA since they are responsible for the localization of PIWI in the nucleus which is critical to transposon silencing (Saito *et al*. 2010, Olivieri *et al.* 2010, Qi *et al.* 2011).

1.14 An Epigenetic Role of piRNAs

piRNAs have been shown to play a role in hybrid dysgenesis in *D. melanogaster*. Brennecke *et al.* 2008 demonstrated that piRNAs are maternally deposited in eggs and function to protect the embryo from transposon movement during the period before the zygotic genome is transcribed. These piRNAs help mount a silencing response that is able to save the zygote from an invasion of transposons before the embryo's genome is able to produce its own response. Thus naïve embryos that do not have maternally

deposited piRNAs targeted toward highly mobile transposons are susceptible to harmful integrations into the genome.

The role piRNAs play in transposon control is only beginning to be understood. The important proteins involved in the pathway have been defined but the details of how they function together to create a targeted response are still not fully worked out. Many questions have yet to be answered such as the function of piRNAs during meiosis, the exact mechanism of controlling transposons and possible alternative functions of piRNAs in the genome (O'Donnell and Boeke 2007). Most of the studies on this pathway thus far have been performed on model organism and little is known about piRNA biogenesis in other organisms.

1.15 piRNA in *Ae. aegypti*

Bioinformatic studies indicate that the piRNA pathway of *Ae. aegypti* varies from that of *D. melanogaster*. While an ortholog of the AGO3 protein has been discovered no homologous AUB proteins have been described. The most dramatic difference is the gene expansion of the PIWI sub family proteins. Where *D. melanogaster* only has one copy of the Piwi gene there are seven different Piwi genes that have been identified in *Aedes aegypti* (Campbell et al. 2008). However, our studies indicate that there are only six unique Piwi genes because Piwi 1 and Piwi 3 are in fact the same gene (Arensburger et al. submitted). Our lab previously sequenced seven libraries from *Aedes aegypti* and demonstrated that 19% of the sequenced piRNAs mapped to transposons. This is in contrast to *D. melanogaster* libraries where 50% of the

sequenced piRNAs mapped to transposons (Arensburger et al. submitted). This study estimated the total piRNA abundance in *Ae. aegypti* to be $2x10^7$ which is comparable to *D. melanogaster*. piRNA clusters were found to cover 20.6% of the genome and were shown to be able to potentially generate up to 84% of the observed piRNAs. These loci were not as transposon rich and were less dense than those clusters found in *D. melanogaster* (Brennecke et al. 2007, Arensburger et al. submitted)

1.16 Transposon Defense in *Ae. aegypti*

Ae. aegypti is a medically important vector of human diseases and has been the subject of many genetic studies. Researchers have been able to use transposon mediated genetic transformation, RNAi and site specific recombinases as tools to manipulate *Ae. aegypti* (Nimmo et al. 2006, Attardo et al. 2003, Clemons et al. 2010). However the use of transposons as a means of genetically altering these mosquitoes has proven challenging and resulted in relatively few labs perform genetic transformations. *PiggyBac, Hermes* and *Mos1* have all been used to transform *Ae. aegypti* yet their inability to be remobilized has hampered the development of such genetic strategies as gene tagging and gene and enhancer trapping (Smith et al. 2010, O'Brachta et al. 2003, Wilson et al. 2003, Sethuraman et al. 2007). One possible explanation for this phenomenon is that *Ae. aegypti*'s unusually high transposon load (47% of the 1.38GB genome composed of transposons) results in a primed piRNA pathway that is able to quickly silence any transposon activity.

1.17 Discussion of the piRNA Pathway in Diptera

The piRNA system in animals has been studied the most in depth in mice and *D. melanogaster*. Many of the key proteins have been identified as well as the structure and composition of the piRNA populations in these organisms. Two major issues that have not been resolved in these model systems are the enzymatic machinery that process piRNA and the mechanisms the regulate recognition of precursor transcripts. Studies thus far have demonstrated a correlation between piRNAs and transposon control however this has yet to be definitively confirmed. The piRNA pathway is still a black box and there is still much to be discovered in these two model systems.

D. melanogaster is used as a model system that can be used as a guide for studying other insects but there have been instances where *D. melanogaster* is the exception to the rule. For example the *P*-element is excellent for use in transposon mediated transformation yet has not proven useful in other insect systems. As well, the gene "sex-lethal" which works in sex determination in *Drosophila* systems does not have a conserved function even when compared to related Dipteran species (Cline 1983, Traut et al. 2006). It is important as we begin to explore the piRNA pathway in alternate organisms that we explore possible differences and be open to the reality that the system may work in a completely different manner. In our case *D. melanogaster* has a smaller genome with a lower transposon load compared to *Ae. aegypti* which could result in varied.

Studies that link the piRNA pathway to transposon movement in *D. melanogaster* have relied on hybrid dysgenesis and immunostaining of transposase ORFs to correlate

the piRNA pathway to transposon control. These types of studies would be difficult to perform in *Ae. aegypti* due to differences in egg morphology and a lack of known crosses that could result in hybrid dysgenesis. Therefore alternate approaches must be taken when working in mosquito systems to demonstrate that the piRNA pathway is linked to transposon control.

Understanding the piRNA system in *Ae. aegypti* could eventually lead to a better understanding of transposon defense thereby allowing researchers a means of evading these defenses. Currently the piRNA system is the least studied of the small RNA pathways and there is much to be learned about the key proteins involved and the cascade in which they work. In order to better understand the piRNA pathway in *Ae. aegypti* it is important to determine which of the AGO3 and PIWI proteins are expressed in each tissue and systematically study how the knock down of each protein affects transposon transcription. Previous studies of the piRNA pathway in *Ae. aegypti* have performed purely bioinformatic analysis and to the best of our knowledge there has been no functional data presented to date. Determining the relationship of each protein to transposon transcription via knock down experiments paired with mRNA SEQ could provide invaluable data regarding transposon regulation in the yellow fever mosquito.

1.18 Goals of this Dissertation

This thesis explores the mechanisms of DNA transposon movement with the goal of increased transformation frequencies in Dipteran systems. Chapters two and three investigate how mutation of the transposase open reading frame affect target site

preference, transposition frequency and transformation frequency. We were unable to achieve hyperactivity however a great deal of information regarding how specific amino acids contribute to transposase activity. Chapter four takes a first look at the piRNA pathway in *Ae. aegypti* which is believed to play a role in transposon defense. A better understanding of the genomes defenses could enable us to shut down this pathway as a means of increasing transformation efficiency. Together these chapters work to increase the potential of DNA transposon as genetic tools in Dipteran systems.

1.19 References:

Allen, M.L., O'Brachta, D.A., Atkinson, P.W. and Levesque, C.S. (2001) Stable, Germ-line Transformation of Culex quinquefasciatus. *J. Med. Entomol*. **38**:701-710

Aravin, A. A. *et al.* (2001) Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. *Curr. Biol.* **11**:1017–1027

Aravin A.A., Lagos-Quintana, M., Yalcin, A., Zavolan, M., Marks, D., Snyder, B., Terry Gaasterland, Meyer, J and Tuschl, T. (2003) The small RNA profile during *Drosophila melanogaster* development. *Developmental Cell*. **5**: 337–350

Aravin, A.A., Sachidanandam, R., Girard, A., Fejes-Toth, K. and Hannon, G.J. (2006) Developmentally regulated piRNA clusters implicate MILI in Transposon Control. *Science*. **316**:744-747

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

Attardo, G.M., Higgs, S., Klinger, K.A., Vanlandingham, D.L. and Raikhel, A.S. (2003) RNA interference-mediated knockdown of GATA factor reveals a link to anautogeny in the mosquito *Aedes aegypti. Pr.o Natl. Acad. Sci USA*. **100**: 13374-13379

Aultman, K.S., Gottlieb, M., Giovanni, M., and Fauci, A.S. (2002) The *Anopheles gambiae* genome: completing the malaria triad. *Science* **298**: 13

Bateman, J.R., Lee, A.M. and Wu C. (2006) Site-specific Transformation of *Drosophila* via ΦC31 integrase-mediated cassette exchange. *Genetics*. **173**:769-777

Baus, J., Liu, L., Heggestad, A.D., Sanz, S. and Fletcher, B.S. (2005) Hyperactive Transposase mutants of the *Sleeping Beauty* transposon. *Mol. Ther.* **12**: 1148-1156

Berghammer, A.J., Klinger, M., and Wimmer, E.A.(1999) A Universal marker for transgenic insects. *Nature*, **402:** 370–371.

Bergman, C.M., Quesneville, H., Anxolabe´ he` re, D., and Ashburner, M. (2006). Recurrent insertion and duplication generate networks of transposable element sequences in the *Drosophila melanogaster* genome. *Genome Biol.* **7**: R112.

Bian, G., Shin, S.W., Cheon, H.M., Kokoza, V. and Raikhel, A.S. (2005) Transgenic alteration of Toll immune pathway in the female mosquito *Aedes aegypti. PNAS*. **102**: 13568-13573

Brennecke**,**J., Aravin, A.A., Stark, A., Dus, M., Kellie, M., Sachidanandam, R. and Hannon, G.J. (2007) Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila. Cell.* **128***:* 1089-1103

Brennecke, J., Malone, C.D., Aravin, A.A., Sachidanandam, R., Stark, A., Hannon, G.J. (2008) An epigenetic role for maternally inherited piRNAs in transposon silencing. Science. **322**: 1387-1392

Campbell, C.L., Black, W.C., Hess, A. M. and Foy, B.D. (2008) Comparative genomics of small RNA regulatory pathway components in vector mosquitoes. *BMC Genomics***. 9**:425

Chalberg, T.W., Portlock, J.L., Olivares, E.C., Thyagarajan, B., Kirby, P.J., Hillman, R.T., Hoelters, J. and Calos, M.P. (2006) Integration specificity of phage phiC31 integrase in the human genome. *J Mol Biol.* **357**:28-48

Chung, W.J., Okamura, K., Martin, R., and Lai E.C. (2008) Endogenous RNA interference provides a somatic defense against *Drosophila* transposons. Current Biology. **18**: 795-802

Clemons, A., Haugen, M., Severson, D and Dunman-Scheel, M. (2010) Functional analysis of gene in *Aedes aegypti* embryos. Cold Spring Harb Protoc.

Cline, T.W. (1983) The interaction between daughterless and Sex-Lethal in triplois: A sex lethal transforming maternal effect linking sex determination and dosage compensation in Drosophila melanogaster. *Dev. Biol*. **95**: 260-274

Craig, N.L. (2002) Mobile DNA: an introduction. In: Craig, R., Gellert, M., Lambowitz, A.M. (Eds), *Mobile DNA II*. ASM Press, Washinton, DC. 3-11

Coates, C.J., Jasinskiene, N., Miyashiro, L. and James, A.A. (1998) *Mariner* transposition and transformation of the yellow fever mosquito, *Aedes aegypti. PNAS*. **95**: 3748–3751

Cooley, L., Kelley, R. and Spradling A. (1988) Insertional mutagenesis of the *Drosophila* genome with single P elements *Science.* **239**: 4844, 1121-1128

Cui, Z., Guerts, A.M., Liu, G., Kaufman, C.D., and Hackett, P.B. (2002) Structure function analysis of the inverted terminal repeats of the *Sleeping Beauty* Transposon. *J. Mol. Biol*. **318**:1221-1235

Davison J.A., Akitake, C.M., Goll, M.G., , J.M., Gosse, N., Baier, H., Halpern, ME., Leach, S.D. and Parsons, M.J. (2007) Transactivation from Gal4-VP16 transgenic insertions for tissue-specific cell labeling and ablation in zebrafish. *Developmental Biology.* **304**:811-24

Desset, S., Meignin, C., Dastugue, B., and Vaury, C. (2003) COM, a heterochromatic locus governing the control of independent endogenous retroviruses from *Drosophila melanogaster*. *Genetics.* **64**: 501-509

Ding, S., Wu, X., Li, G., Han, M., Zhuang, Y. and Xu, T. (2005) Efficient transposition of the *piggyBac* (PB) transposon in mammalian cells and mice. *Cell.* **122**:473–483

Duffy, J. N. (2002) GAL4 System in *Drosophila*: A Fly Geneticist's Swiss Army Knife. *Genesis.* **34**:1–15

Elbashir, S. M., Lendeckel, W. and Tuschl, T. (2001a) RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* **15**:188–200

Elbashir, S. M., Martinez, J., Patkaniowska, A., Lendeckel, W. and Tuschl, T. (2001b) Functional anatomy of siRNAs for mediating efficient RNAi *in Drosophila melanogaster* embryo lysate. *EMBO J.* **20**: 6877–6888

Feng, X., Denarz, A.M., and Colloms, S.D. (2009) Precise targeted integration by a chimaeric transposase zinc-finger fusion protein. *Nucleic Acids Research* **38**:1204-1216

Feschotte, C. and Pritham, E.J. (2007) DNA transposons and the evolution of eukaryotic genomes. *Annu Rev Genet*. **41**: 331–368

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., Mello, C.C. (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391**: 806–811

Fish, M.P., Groth, A.C., Calos, M.P. and Nusse, R. (2007) Creating transgenic Drosophila by microinjecting the site-specific ΦC31 integrase mRNA and transgene-containing donor plasmid.*Nature Protocols*. **2**:2325-2331

Frechotte, C. (2008) The contribution of transposable elements to the evolution of regulatory networks. *Nat. Rev. Genet*. **5**:397-405

Germon, S., Bouchet, N., Casteret, S., Carpentier, G., Adet. J., Bigot. Y. and Augé-Gouillou, C. (2009) *Mariner Mos1* transposase optimization by rational mutagenesis. *Genetica.***137**:265–276

Geurts, A.M., Hackett, C.S., Bell, J.B., Bergmann, T.L., Collier, L.S., Carlson, C.M. Largaespada, D.A. and Hacket, P.B. (2006) Structure-based prediction of insertion-site preferences of transposons into chromosomes. *Nucleic Acids Res.* **34**:2803-2811

Ghosh, A., Ribolla, P.E. and Jacobs-Lorena, M. (2001) Targeting *Plasmodium* ligands on mosquito salivary glands and midgut with a phage display peptide library. *Proc. Natl. Acad. Sci*. **98:** 13278–13281

Glover, D. J., Lipps, H.J., and Jans, D. A. (2005) Towards safe, non-viral therapeutic gene expression in humans. *Nat. Rev. Genet*. **6**: 299 – 310

Goryshin I.Y and Reznikoff, W.S. (1998) *Tn5 in vitro* transposition. *J. Biol. Chem*. **273**:7367-7374

Grishok, A., Pasquinell, A.E**.,** Conte, D., Li, N., Parrish**, S**., Ha**,** I., Baillie**,** L.D., Fire, A., Ruvkun, G. and Mello, C.C**.** (2001) Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell.* **106**: 23–34

Grossman, G.L., Rafferty, C.S., Fraser, M.J. and Benedict, M.Q. (2000) The *piggyBac* element is capable of precise excision and transposition in cells and embryos of the mosquito, *Anopheles gambiae*. *Insect Biochem. Mol. Biol.* **30**: 909-914

Groth, A.C., Fish, M., Nusse, R. and Carlos, M.P. (1994) Construction of transgenic Drosophila by using the site-specific integrase from phage ΦC31. *Genetics*. **166**:1775-1782

Gueiros-Filho, F.J. and Beverley, S.M. (1997) Trans-kingdom transposition of the Drosophila element Mariner within the protozoan Leishmania. *Science*. **276**:1716-1719

Hammond, S. M., Bernstein, E., Beach, D. and Hannon, G. J. (2000) An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature.* **404**: 293–296

Handler, A.M., McCombs, S.D., Fraser, M.J. and Saul, S.H. (1998) The lepidopteran transposon vector piggyBac, mediates germ-line transformation in the Mediterranean fruit fly. *Proc Natl Acad Sci*. **95**:7520-7525

Handler, A.M. and Harell, R.A. (1999) Germline transposition of *Drosophila melanogaster* with the *piggyBac* transposon vector. *Insect Mol Biol*. **8**: 449-457

Hartley, K.O., Nutt, S.L. and Amaya, E. (2002) Targeted gene expression in transgenic *Xenopus* using the binary Gal4-UAS system. *PNAS*. **99**:1377-82

Hickman, A.B., Perez, Z.N., Zhou, L., Musingarimi, P., Rodolfo,G., Hinshaw, J.E., Craig, N.L. and Dyda, F. (2005) molecular architecture of a eukaryotic DNA transposases. *Nature structural and molecular biology*. **12**: 715-721

Holt *et al.* (2002) The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science.* **298**: 129-149

Horwich, M. D., Li, C., Matranga , C., Vagin ,V., Farley, G., Wang, P. and Zamore, P.B**.** (2007) The *Drosophila* RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Curr. Biol.* **17**: 1265–1272

Hutvágner, G**.,** McLachlan, J., Pasquinelli, A.E., and Balint, E. (2001) cellular function for the RNA interference enzyme Dicer in the maturation of the *let‑7* small temporal RNA. *Science* **293**:834–838

Imamura, M., Nakai, J., Satoshi Inoue, S., Quan, G.X., Kanda, T. and Tamura, T. (2003) targeted gene expression using the *GAL4/UAS* system in the silkworm *Bombyx mori* . *Genetics* **165**: 1329–1340

Ito, J., Gosh, A., Moreira, L., Wimmer, E.A. and Jacobs-Lorena, M., (2002) Transgenic Anopheline mosquitoes impaired in transmission of a malaria parasite. *Nature.* **417**: 452-455

Ivics, Z*.,* Hackett, P.B., Plasterk, R.H., and Izsvák, Z. *(*1997) Molecular reconstruction of *Sleeping Beauty*, a *Tc1*-like transposon from fish, and its transposition in human cells. *Cell.* **91**: 501-510

Ivics, Z., Katzer, A., Stuwe, E.E., Fiedler, D., Knespel, S. and Izsvak, Z. (2007) Targeted *Sleeping Beauty* transposition in human cells. *Mol Ther* **15**:1137-1144

Jasinskiene, N., Coates, C.J., Ashkyan, A. and James, A.A. (2003) High efficiency, site – specific excision of a marker gene by the phage P1 *cre-loxp* system in the yellow fever mosquito*, Aedes aegypti*. *Nucleic Acid Res*. **31**: e147

Ketting, R. F. Fischer, S.E.J., Bernstein, E.J., Sijen, T., Hannon, G.J. and Plasterk, R.H.A**.**
(2001) Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev.* **15**: 2654–2659

Lampe, D.J., Churchill M.E.A. and Robertson, M.A. (1996) A purified mariner transposase is sufficient to mediate transposition *in vitro. EMBO J.* **15**:5470-5479

Lagos-Quintana, M., Rauhut, R., Lendeckel, W. and Tuschl, T. (2001) Identification of novel genes coding for small expressed RNAs. *Science* **294**: 853–858

Lander, E.S. et al. (2001) Initial sequencing and analysis of the human genome. *Nature.* **409**: 860-921

Lee, C.Y. and Landolo, J.J.(1986) Integration of staphylococcal phage L54a occurs by site-specific recombination: structural analysis of the attachment sites. *Proc. Natl. Acad. Sci*. USA **83**:5474-5478

Lee, C. Y., and Iandolo, J.J. (1988) Structural analysis of staphylococcal bacteriophage 411 attachment sites. *J. Bacte- riol*.**170**:2409-2411

Li, C., Vagin, V.V., Lee, S., Xu, j., Ma, S., Xi, H., Seitz, H., Horwich, M.D., Syrzycka, M., Honda, B.M., Kitller, E.L.W., Zapp, M.L., Klattenhoff, C., Sculz, N., Theurkauf, W.E., Weng, Z., and Zamore, P.D. (2009) Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. *Cell*. **137**: 509-521

Li, X., Lobo, N., Bauser, C.A. and Fraser, M.J. Jr. (2001) The minimum internal and external sequence requirements for transposition of the eukaryotic transformation vector *piggyBac*. *Mol Genet Genomics*. **266**: 190-198

Li, X., Harrell, R.A., Handler, A.M., Beam, T., Hennessy, K. and Fraser, M.J. Jr. (2005) *piggyBac* internal sequences are necessary for efficient transformation of target genomes. *Insect Mol Biol.* **14**: 17-30

Lin, H. and Spralding, A.C. (1997) A novel group of pumilio mutations affects the asymmetric dvision of germline stem cells in the *Drosophila* ovary. Development. **124**: 2463-2476

Liu,H. & Visner, G.A. (2005) Applications of *Sleeping Beauty* transposons for nonviral gene therapy. *Life*. **59**: 1 – 6

Lorenzen, M.D., Brwon , S.J., Denell, R.E. and Beeman, R.W. (2002) Transgene expression from the *Tribolium castaneum* polyubiquitin promoter. *Insect Mol Biol*. **11**:133-139

Magathavally, K.J., Kaminski, J.M. and Coates, C.J. (2006) Chimeric *Mos1* and *piggyBac* transposases result in site-directed integration. *FASEB* **20**:E1188-E1195

Malone, C.D., Brennecke, J., Dus, M. Stark, A. McCombie, W.R., Sachidanandam, R. and Hannon, G.J. (2009) Specialized piRNA pathways act in germline and Somatic tissues of the *Drosophila* ovary. Cell. **137**: 522-535

McClintock, B. (1951) Chromosome organization and genic expression. Cold Spring Har. Symp. Quant. Biol. **16**: 13-47

Medhora, M., Maruyama, K. and Hartl, D.L. (1991) Molecular and functional analysis of the mariner mutator element *Mos1* in *Drosophila*. *Genetics*. **128**:311-318

Mita, K., Kasahara, M., Sasaki, S., Nagayasu, Y., Yamada, T., Kanamori, H., Namiki, N., Kitagawa, M., Yamashita, H., Yasukochi, Y., Kadono-Okuda, K., Yamamoto, K., Ajimura, M.,
Ravikumar, G., Shimomura, M., Nagamura, Y., Shin, I.T., Abe, H., Shimada, T., Morishita, S., Sasaki, T. (2004) The genome sequence of silkworm, Bombyx mori. *DNA Res*. **11**: 27–35.

Morris, A.C., Eggleston, P. and Crampton, J.M. (1989). Genetic transformation of the mosquito Aedes aegpti by micro-injection of DNA. *Med Vet Entomol.* **3**:1-7

Nene, V. *et al.* (2007) Genome sequence of *Aedes aegypti*, a major arbovirus vector. *Science* **316:**1718-1723

Nimmo, D.D., Alphey, L., Meredith, J.M. and Eggleston, P. (2006) High efficiency site-specific genetic engineering of the mosquito genome. *Insect Mol Biol*. **15**: 129-136

Novick, R. P. (1967) Properties of a cryptic high-frequency transducing phage in Staphylococcus aureus. *Virology* **33**:155- 166.

Obrochta D.A. and Atkinson, P.W. (1996) Transposable elements and gene transformation in non-drosophilid insects. *Insect Biochem. Molec. Biol.* **26**: 739-753

Obrochta D.A., Sethuramuran, N., Wilson, R., Hice, R.H., Pinkerton, A.C., Levesques, C.S., Bideshi, D.K., Jasinskiene, N., Coates, C.J., James, A.A. *et al.* (2003) Gene vector and transposable element behavior in mosquitoes. *Journal of Experimental Biology.* **3**: 3823-3834

O'Donnell, K.A. and Boeke, J.D**.** (2007) Mighty Piwis defend the germline against genome intruders. *Cell*. **129**: 37-44

Olivieri, D., Sykora, M.M., Sachidanandam, R., Mechtler, K., and Brennecke, J. (2010) An *in vivo* RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in *Drosophila. EMBO J*. **29**: 3301-3317

Olson, K.E., Higgs S., Hahn, C.S., Rice, C.M., Carlson, J.O. and Beaty, B.J. (1994) The expression of chloroamphenicol acetyltransferase in *Aedes albopictus* (C6/36) cells and *Aedes triseriatus* mosquitoes using a double subgenomic recombinant Sindbis virus. *Insect Biochem. Mol. Biol.* **24:**39-48

Pattee,P.A., Thompson,N.E., Haubrich, D. and Novick, R.P. (1977) Chromosomal map locations of integrated plasmids and related elements in Staphylococcus aureus. *Plasmid* **1**:38-51.

Peloquin, J.J. Thibault, S.T. and Miller, T.A. (2000) Genetic transformation of the pink bollworm *Pectinophora gossypiella* with the piggyBac element. *Insect Mol. Biol*. **9**: 323–333

Perez, A.N., Musingarimi, P. Craig, N.L., Dyda, F. and Hickman, A.B. (2005) Purification, crystallization and preliminary crystallographic analysis of the Hermes transposases. *Structural biology and Crystallization communications.* **61**:587-590

Plasterk, R.H.A. and Van Luenen, H.G.A.M. (2002) In Mobile DNA II (ed. Craig, N.L., Craigie, R., M. and Lambowitz, A.M.) ASM Press, Washington, DC, USA. 519-532

Pledger, D.W. and Coates, C.J. (2005) Mutant Mos1 mariner transposons are hyperactive in *Aedes aegypti*. *Insect Biochem Mol Biol.* **35**:1199–1207

Qi, H., (2011) The Yb body, a major site for Piwi-associated RNA biogenesis and a gateway for Piwi expression and transport to the nucleus in somatic cells. *J. Biol. Chem*. **286**: 3789-3797

Raush, H. and Lehmann, M. (1991) Structural analysis of actinophage phi C31 attachment site. Nucleic Acids Res. **19**:5187-5189

Robine, N., Lauc, N.C., Balla, J.S., Jin, Z., Okamura, K., Kuramochi-Miyagawa, S., Blower, M.D. and Lai, E.C. (2009) A Broadly conserved pathway generates 3' UTR directed primary piRNAs. *Curr. Biol.* **19**: 2066-2076

Robinson, A.S., Franz, G., Atkinson, P.W. (2004) Insect transgenesis and its potential role in Agriculture and human health. *Insect Biochem. Molec. Biol.* **34**: 113-120

Roll Back Malaria. Framework for monitoring progress and evaluating outcomes and impact. Geneva, World Health Organization, 2000 (WHO/CDS/RBM/2000.25).

Rong, Y.S. and Golic, K.G. (2000) Gene targeting by homologous recombination in *Drosophila*. *Science.* **288**: 2013-2018

Saito, K., Sakaguchi,Y., Suzuki,T., Suzuki, T, and Siomi, H. (2007) Pimet, the *Drosophila* homolog of HEN1, mediates 2′-*O*-methylation of Piwi- interacting RNAs at their 3′ ends. *Genes Dev.* **21**: 1603–1608

Saito, K., Ingaaki, S., Mituyama, T., Kawamura,Y., Ono, Y., Sakota, E., Kotani, H., Asai, K., Siomi, H. and Siomi, M. (2009) A regulatory circuit for Piwi by the large Maf gene traffic jam in *Drosophila. Nature*. **461**:1296-1299

Saito, K. and Siomi, M.C. (2010) Small RNA-mediated quiescence of transposoable elements in animals. *Dev. Cell*. **19**: 687-697

Sakar, A., Coates, C.J., Whyard, S., Willhoeft, U., Atkinson, P.W. and O'Brochta, D.A. (1997) The *Hermes* element from *Musca domestica* can transpose in four families of Cyclorrhaphan flies. *Genet*. **99**:15-29

Scheer, N. and Campos-Ortega, J.A. (1999) Use of the Gal4-UAS technique for targeted gene expression in the zebrafish. *Mech Dev*. **80**: 153-158.

Senti, K.A. and Brennecke, J. (2010) The piRNA pathway: a fly's perspective on the guardian of the genome. *Cell*. **26**: 499-509

Sethuraman, N, M. J. Fraser Jr., P. Eggleston and D. A. O'Brochta (2007). Post-integration stability of *piggyBac* in *Aedes aegypti*. *Insect Biochem. Mol. Biol*. **37**: 941-951

Shin, S.W., Kokoza V.A. and Raikhel A.S. (2003) Transgenesis and reverse genetics of mosquito innate immunity. *The Journal of Experimental Biology* **206**:3835-3843

Smith, R.C. (2007) Germ line expression of genetic markers and transposases in the yellow fever mosquito *Aedes aegypti*. University of California, Riverside

Smith, R.C. and Atkinson, P.A. (2010) Mobility properties of the *Hermes* transposable element in transgenic lines of *Aedes aegypti*. *Genetica*.**139**: 7-22

Snow, R.W., Craig, M., Deichmann, U. and Marsh, K. (1999) Estimating mortality, morbidity and disability due to malaria among Africa's non-pregnant population. *Bulletin of the World Health Organization*. **8**:624-40

Song, J.J., Liu, J., Tolia, N.H., Schneiderman, J., Smith, S.K., Martienssen, R.A. Hannon, G.J., and Joshua-Tor, L. (2003) The crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **305**:1434-1437

Sundaraajan, P., Atkinson, P.W. and O'Brachata, D.A. (1999) Transposable elements interactions in insects: crossmobilzation of *hobo* and *Hermes*. *Insect Mol. Biol.* **8**:359-368

Thyagarajan, B., Olivares, E.C., Hollis, R.P., Ginsburg, D.S. and Carlos, M.P. (2001) Site-specific genomic integration in mammalian cells mediated by phage phi C31 integrase. *Mol Cell Biol.* **21**:3926-3934

Trauner, J., Schniko, J., Lorenzen, M.D., Shippy, T.D., Wimmer, E.A., Beeman, R.W., Klinger, M., Bucher, G. and Brown, S.J. (2009) Large-scale insertinal mutagenesis of a coleopteran stored grain pest, the red flour beetle *Triblolium castaneum*, identifies embryonic lethal mutations and enhancer traps. *BMC Biology*. **7**:73

Traut, W., Nimi, T.,Ikeo , K. and Sahara, K. (2006) Phylogeny of the sex-determining gene Sex–Lethal in insects. *Genome*. **49**: 254-262

Vagin, V. V., Sigova, A., Li, C., Seitz, H., Gvozdev, V. and Zamore, P.D. (2006) A distinct small RNA pathway silences selfish genetic elements in the germline. *Science.* **313**: 320–324

Vigdal, T.J., Kaufman, C.D., Izsvak, Z., Voytas, D.F. and Ivics, Z. (2002) Common Physical properties of DNA affecting target site selection of sleeping beauty and other *TC1/ mariner* transposable elements. *Nucleic Acids Res* **23**:784-790

Walbot, V. (1992) Strategies for mutagenesis and gene cloning using transposon tagging and T-DNA insertional mutagenesis. *Annu. Rev. Plant Phys. Plant Mol. Biol.* **43**: 49-82

Wang, H.H. and Fraser, M.J. (1993) TTAA serves as the target sire for TFP3 lepidotpteran transposon insertions in both nuclear polyhedrosis virus and *Trichoplusia ni* genomes. *Insect Mol. Biol.* **1**: 109-1996

Warren, W.D., Atkinson, P.W. and O' Brochta, D.A. (1994) The *Hermes* transposable element from the house fly, *Musca domestica*, is a short inverted repeat-type element of the *hobo, Ac*, and *Tam3* (hat) element family. *Genet. Res.* **64**: 87-97

Wightman, B., Burglin, T. R., Gatto, J., Arasu, P. & Ruvkun, G. (1991) Negative regulatory sequences in the lin-14 3′-untranslated region are necessary to generate a temporal switch during Caenorhabditis elegans development. *Genes Dev.* **5**: 1813–1124

Wimmer E.A. (2005) Insect transgenesis by site-specific recombination. *Nature Methods*. **2**: 580-582

Wilson, R., Orsetti, J., Klocko, A.K., Aluvihare, C., Peckham, E., Atkinson P.W., Lehane, M.J. and O'Brachta, D.A. (2003) Post- integration behavior of a *Mos1 Mariner* gene vector in *Aedes aegypti*. *Insect Biochem Mol Biol*. **33**:853-863

Wilson, M.H., Kaminski, J.M. and George, A.L. Jr. (2005) Functional zinc finger/sleeping beauty transposase chimeras exhibit attenuated overproduction inhibition. FEBS. **2**:6205-6209.

Wilson, M.H., Coates, C.J. and George, A.L. (2007) *PiggyBac* transposon –mediated gene transfer in human cells. *Molecular Therapy*. **15**: 139-145

Yant, S.R., Park, J., Huang, Y., Mikkelsen, J.G. and Kay, M.A. (2004) Mutational analysis of the N-Terminal DNA-binding domain of the *Sleeping Beauty* transposase: Critical residues of the DNA binding and hyperactivity in mammalian cells. *Mol. Cell. Biol.* **24**:9239-9247

Yant, S.R. Huang, Y., Akache, B., and Kay, M.A. (2007) Site-directed transposon integration in human cells. *Nucleic Acids Res* **35**:1-13

Zamore, P. D., Tuschl, T., Sharp, P. A. and Bartel, D. P. (2000) RNAi: double-stranded RNA directs the ATP dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell,* **101**: 25–33

Zhao, Y. and Eggleston, P. (1998) Stable transformation of *Anopheles gambiae* cell line mediated by the *Hermes* mobile genetic element. *Insect Biochem. Mol. Biol*. **28**: 213-219

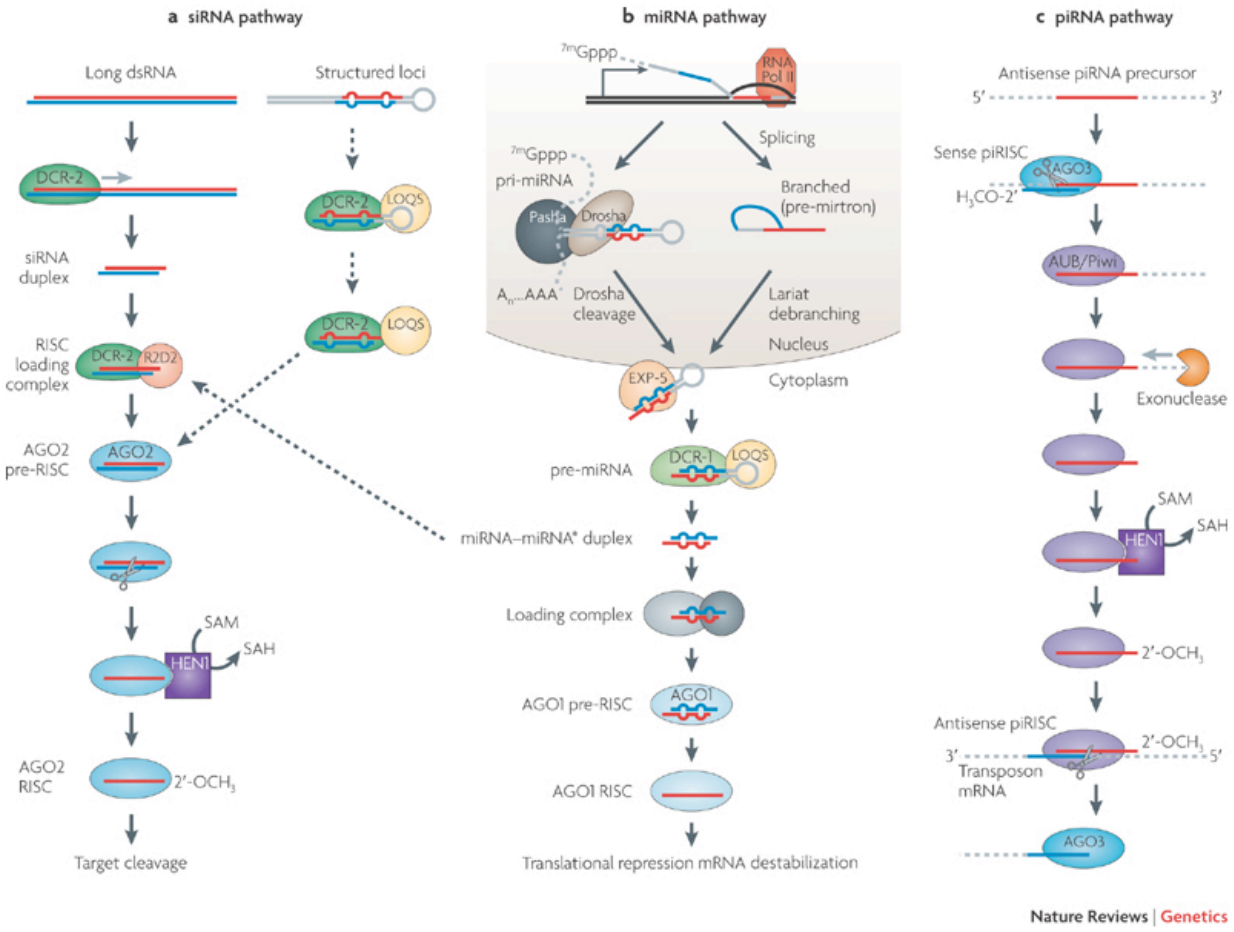

Figure 1.1. Megha Ghildiyal and Phillip D. Zamore, Nature 2009. This figure shows each of the three small RNA pathways in D. melanogaster. The siRNA and microRNA pathways are both Dicer dependent while piRNA pathway involves PIWI proteins.

Chapter 2

AeBuster1 Transposase Mutants Demonstrate Altered Target Site Preference

2.1 Abstract

AeBuster1 is class II DNA transposon belonging to the hAT superfamily (Arensburger *et al*. 2011). Randomly mutated *AeBuster1* transposases' transposition activities were screened in yeast via a high throughput assay. This study took those *AeBuster1* mutants that had the highest increases in transposition rates and tested their activity in *Drosophila melanogaster.* None of the mutants retained their hyperactive activity in *D. melanogaster.* Transposition events for each mutant were sequenced in order to determine the target site preference. The mutants L368P, V597A/ L368P, and V597A all had significantly altered target site preference from the wild-type *AeBuster1*.

2.2 Introduction

Transposons are mobile genetic elements that are able to integrate themselves as well as the genes they carry into host genomes. Transposable elements are able to move within an organism's genome and can therefore act as powerful mutagens and effect the evolution of genomes (Feschotte and Pritham 2007;Craig 2002) By elucidating basic knowledge and learning how to effectively manipulate transposons, novel genetic tools can be created to combat human health issues. Transposons can represent a major proportion of a genome and have been shown to compose up to 35 percent of the human genome (Lander *et al.* 2001).

The *hAT* superfamily is composed of Class II DNA transposons named after three well-studied elements the *hobo* transposon of *D. melanogaster*, *Ac* transposon of *Zea*

mays, and *Tam3* in snapdragons (Atkinson et al. 2003; Calvi *et al.* 1991). The hAT super family is composed of 11 separate super families as described in Freschotte and Pritham (2007).The family is also composed of the *Hermes* transposon of the housefly *Musca domestica* and the *Tol2* transposon of the Japanese Medaka fish, *Oryzias latipes* (Blackman. *et al.* 1989; Kawakami and Shima 1999; McClintok *et al.* 1950; O'Brochta *et al.* 1996). *Buster* transposons such as *TcBuster* from *T. castaneum* and *AeBuster1* from *Ae. aegypti* also belong to the *hAT* superfamily (Arensburger et al. 2011). *hAT* transposons are found in mammalian genomes including humans where they are the most abundant DNA transposons comprising 1.55% (195 Mb) of the total genome. However none of the *hAT* elements in the human genome have been shown to be active during the past 50 million years (Lander *et al*. 2001).

The transposition mechanism of the *hAT* family was previously determined using the *Hermes* element from *M. domestica.* This study found that transposition proceeds through a cut-and paste mechanism involving DNA breakage and rejoining. The transposase first forms a synaptic complex with transposon DNA through binding to the transposon ends and then excises the DNA from the donor site through hairpin intermediates on the flanking DNA. These intermediates are formed at the flanking donor DNA sites. The freed transposon ends then integrate into the target site through the nucleophilic attack of exposed 3'OH at the terminal ends of the transposon (Zhou *et al.* 2004).

AeBuster1 and *TcBuster* were recently discovered using a customized algorithm that searched the *Ae. aegypti* and *T. castaneum* genomes for new active *hAT* transposons (Arensburger *et al.* 2011; Trauner *et al*. 2008; Nene *et al.* 2007) The elements were named *Buster* elements after it was discovered that the *hAT*-like sequences were closely related to the *Buster* genes in vertebrates and yet they contained signature motifs of the *hAT* superfamily. Interplasmid transposition assays demonstrated that these elements are active in both *D. melanogaster* and *Ae. aegypti* embryos. The transposition assays also revealed that both *AeBuster1* and *TcBuster* generate an 8 bp target site duplication however the target site preference is for nnnTAnnn (Arensburger *et al.* 2011).

AeBuster1 is the first active transposase discovered from the dangerous disease vector *Ae. aegypti,* commonly known as the yellow fever mosquito. *AeBuster1* is active in yeast, and as a result mutants of *AeBuster1* were screened for hyperactivity in a high throughput assay in yeast. The resulting mutants were tested with in vivo and in vitro assays.

2.3 Methods and Materials

2.3.1 Fly stocks

A strain of *D. melanogaster* C-s,w 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.

2.3.2 Plasmid Constructs

All of the plasmids made for the Dipteran assays were constructed using a standard protocol as determined in Sambrook *et al.* (1989). Each of the hyperactive transposase constructs were made through the excision of the transposase from pGalsHermes using sticky-ended speI-XhoI and then each fragment was ligated into the vector pkH70new. pKh70new is an altered form of the pkhsp70 (Arensburger *et al.* 2005). The hsp70 5' promoter and 3' terminator were moved so that the restriction enzyme sites are found between the two regions allowing for easy excision and insertion of fragments between pgalsHermes and pKh70new. pGalHermes was derived from the p414GalS plasmid with the addition of *Hermes* transposase between base pairs 3245 and 3246 (Mumberg *et al*. 1994).

2.3.3 *PiggyBac* donor plasmid

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

ATAGGTACCCTATTATTAACCCTAGAAAGATAATCATATTGTG-3') and XhoI-pBac RE F (5'-ATACTCGAGGACTAATAAGTATAATTTGTTTCTATTATGTATAAGTTAAGC-3'). Using restriction sites incorporated into the PCR products, the left and right end samples were digested and inserted into pBluescript II KS+ using the *SacII-XbaI* and *KpnI-XhoI* sites respectively. The resultant plasmid, pBSpBacLR, was then linearized by digestion with *XbaI* to incorporate a *NheI* linearized pGENToriEGFP plasmid to create the pBacGoEGFP plasmid. The pGENToriEGFP plasmid was created by excising the LacZα ORF from the pGENToriα plasmid by digestion with *BamHI* and *NheI*, and replaced by a *BamHI*/*NheI* fragment containing the EGFP ORF. The *BamHI*/*NheI* EGFP fragment was generated through the PCR amplification of the ORF using the primers BamHI EGFP and NheI EGFP and the pBac[3xP3-EGFPafm] plasmid as a template. The PCR product was then digested with *BamHI* and *NheI* and ligated into the pBacGoEGFP (Smith 2007).

2.3.4 *PiggyBac* helper plasmid

The *piggyBac* helper is 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.

2.3.5 *AeBuster1* Donor Plasmid

The transposable element was amplified in sections from *Ae. aegypti* (Orlando). Genomic DNA was purified using the Wizard Genomic DNA Kit (Promega), and 360ng of DNA was used as template in 50ul PCR reactions using the TripleMaster PCR System (Eppendorf). The pGT-AeBusterL2 clone was made by PCR amplification using primers which encompassed the region from the target site duplication of copy of AeBuster1 to the sequence immediately upstream of the ATG of the AeBuster1 ORF. The following primers were used for AeBuster1 L2 (AATGGTACCGCTTATGGCATAGATTCCCAAACTGTG, ITR) and AeBuster L Rev (GATCTCGAGATCTGAAATTATCAAATAATGAATCGCATATTCTG)

The right end PCR product was cloned as above, first into pGEM-T Easy and then into the left end clone following digestion with SacI and XbaI (NEB) to yield the clone pBSAeBusterLR.

Clone pBSAeBuster donor 1 was made by ligation of SmaI digested pBSAeBLR with XmnI-digested pGENToriAlpha, which was derived from pK19 and in which the Kanamycin^R gene was replaced with the gentamycin^R gene from pFastBac HTb (Invitrogen). The donor element thus contained a replication origin, a gentamycin^R gene, and a lacZ-alpha gene.

Cloning the *Aedes Buster* ORF: The primers used for were *AeBuster* ORF For (GATGATATCAGATATGGATAAATGGTTGTTGAAGAAGC) and *AeBuster* ORF Rev (GATGATATCTTAGTGTGATGGATGGCTTGCTTG). DNA was purified as above, ligated into pGEM-T Easy and sequenced. One clone (pGEMT-AeBuster clone 9

ORF) had a sequence which had 2 silent mutations, but no differences in protein sequence from *AeBuster* copies 2 and 3 and was used for the helper plasmid

2.3.6 *AeBuster1* Helper Plasmid

pGEMT-AeBuster clone 9 ORF was digested with *EcoRV* (NEB) and cloned into the *SmaI* site of the pKhsp70 vector (Arensburger *et al*. 2005).

2.3.7 Target Plasmid

pGDV1 will function as the system's target plasmid. pGDV1 is a *Bacillus subtillus* low copy plasmid (Sakar *et al*. 1997) that is unable to replicate within *E. coli* without the addition of an *E. Coli* origin of replication.

2.3.8 Plasmid Mixes

Microinjection of *D. melanogaster* involves several variables that can be difficult to hold constant; therefore it is necessary to create an injection mix that is able to eliminate the background noise inherent to the process. In order to compare relative rates of transposition frequency across trials there must be a means of normalizing the data. This can be achieved through the use of a five plasmid interplasmid transposition assay as developed in Smith (2007). In this assay *AeBuster1* transposon is tested by co-injecting the appropriate helper and donor plasmids along with pGDV1 as a target. This injection mix contains the following: 250ul of *Aebuster1* donor, 250ul of the mutant *AeBuster1* helper-transposase, 250ul of *piggyBac* donor, 250ul of wild-type *piggyBac* helper-transposase and 1000ul of pGDV1 as a donor plasmid. By using *piggyBac* in the assay

along with *AeBuster1* we can compare the number of transposition events of the *AeBuster1* mutated transposase relative to the number of *piggyBac* transposition events, thus transposition rates can be normalized between trials. In each of the trials the transposition events produced from separate transposases can be distinguished by the markers present in the donor plasmids. The *AeBuster1* donor plasmids contains the *LacZ* gene and therefore expresses blue pigment in the presence of IPTG- XGAL where as the *piggyBac* donor plasmid contains the GFP marker.

2.3.9 Microinjection into *Drosophila melanogaster*

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 were 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 then heat shocked at 37°C for one hour. The embryos were then processed using the transposition assay described below.

2.3.10 Dipteran Transposition Assay

The embryos are heat shocked at 37°C for one hour and then allowed to rest at 30° C for thirty minutes before further processing. The embryos were collected from the

glass slides and placed into a proteinase buffer. The embryos were then lysated and incubated for 30°C minutes at 55°C. Next the plasmids were recovered and electroporated into the DH10B *E. coli* strain (Gibco-BRL). Of these cells 1/200th were plated on LB plates containing the appropriate antibiotics for 24 hours at 37° to screen for donor titer. The remaining cells were then placed onto plates containing a mixture of antibiotics for three days at 37°C to screen for transposition events as described in Smith (2007). The transposition events were sequenced at the University of California, Riverside Genomic Core Facility to determine the integration sites of *AeBuster1* mutants into the pGDV1target plasmid.

2.3.11 Sequencing of Transposition Events

AeBuster1 transposition events were sequenced using the primers:

AebusterL Rev-GATCTCGAGATCTGAAATTATCAAATAAGAATCGCATATTCTG

AeBuster R For-GATTCTAGATGCGCATCGAACAACATTTTTAGTGAG.

2.4 Results

2.4.1 Evaluation of Hyperactive Mutants *in vivo*

Utilizing transposition assays we tested the activity of mutant *Aebuster1* transposases in *D. melanogaster*. All hyperactive mutants were cloned into pKh70new which is an altered form of the pkhsp70 (Arensburger *et al.* 2005). A five plasmid mix containing the mutant *AeBuster1* donor and helper plasmids, *piggyBac* donor and helper

plasmids and the target plasmid pGDV1 was injected into 0-2 hour old *D. melanogaster* embryos. The target plasmids were then recovered and sequenced to determine the target site into which the transposons were inserted.

2.4.2 *Aebuster1* Mutants are not Hyperactive in Dipteran Transposition Assays

The *D. melanogaster* transposition assay allows for the analysis of *AeBuster1* mutants in higher eukaryotes. Each *AeBuster1* mutant was co-injected into *D. melanogaster* embryos with the *piggyBac* helper and donor as a means of normalizing the resulting frequencies (Table 2.1.). Mutants L368P and L368P/V597A had the highest transposition rates of 3.49x 10^{-4} and 3.48x10^{-4} respectively, however they were both at least 50% as active as *piggyBac* in the same trial. No significant difference was found between the transposition rates of L368P and L368P/V597A compared to wild type *AeBuster1* in *D. melanogaster*. *AeBuster1* mutants V168A/T324A/N465S and E260G/S485L had similar transposition frequencies ~2.5 X 10^{-4} while mutants L348V/L368P, V597A have frequencies ~5.5x 10^{-5}. No transposition events were found for the mutant E260G/S485L resulting in a transposition frequency of zero.

2.4.3 *AeBuster1* mutants have altered target site preference

Transposition events in *D. melanogaster* were sequenced from the right and left ends outward to determine the 8 bp target site duplications. These sequences were then compared to the target sites sequenced for wild type *AeBuster1* which has an 8 bp

consensus sequence of nnnTAnnn (Arensburger *et al.* 2011). Figures 2.1, 2.2, 2.3 and 2.4 are Weblogos depicting the preference for certain nucleotides at each of the 8 positions for wild-type, L368P, V597A and L368P/V597A (Crooks *et al*. 2004). Table 2.2 shows the p-values for the statistical difference between the mutants and wild-type *AeBuster1* at each of the 8 base positions as determined using the Fisher Exact Test. Mutant L368P illustrates no significant difference between any of the 8 positions, however, Mutant L368P/V597A demonstrated significant differences in positions 3, 4, 6, and 7 from the wild-type *AeBuster1*. Mutant V597A's target site preference is significantly different from wild-type at positions 3, 5 and 6. L368P/V597A has the most divergent target site preference suggesting that while V597A is most likely the mutation leading to the altered activity there is reasonable evidence to suggest synergistic effects involving L368P. These data show that the manipulation of a transposase protein can lead to altered target site preference in *D. melanogaster*.

2.4.4 Integrations patterns of *AeBuster1* wild-type and mutants

In transposition assays some transposases demonstrate a preference for insertion into specific locations in the pGDV1 target plasmid known as "hot spots". Transposition events were sequenced in order to determine where the transposase inserts and the resulting data is depicted in Tables 2.3 and 2.4. *AeBuster1* wild-type does not show strong bias toward any one hot spot but did insert into position 943 four times and position 1993 twice out of 25 sequenced events. All of the *AeBuster1* mutants showed insertion preference to position 1993 in pGDV1. L369P had the strongest preference for

position 1993 with 12 insertions out of 26 sequenced occurring at this hot spot. There was no preference seen overall for insertion into position

2.5 Discussion

2.5.1 *AeBuster1* Mutants in *D. melanogaster*

Transposition assays are subject to a great deal of variation within and between experiments, thus standard deviations arising from this experimental technique tend to be large. Embryo desiccation, needle shape and injection pressure are several variables that cannot be kept perfectly constant between experiments. Additionally, the assay can only be taken as an indicator of somatic activity due to the fact that somatic nuclei comprise the majority of nuclei found in the developing embryos. Transposition assays in *D. melanogaster* demonstrated that the mutant *AeBuster1* transposases did not maintain significantly increased activity in fly embryos as was demonstrated in the yeast assay. One possible explanation for this difference is that chromatin remodeling and transcription activation requires the initiation of complexes which involves some common machinery between yeast and *Drosophila* however these complexes are also composed of several unique components that are able to create specific targets and specific functions (Eisen *et al.* 2001).

Only one *AeBuster1* mutant V168A/T324A/N465S demonstrated a higher transposition rate than *piggyBac* however this difference was not significant. Additional genetic studies would need to be conducted to demonstrate any true increase in the

mutants' transposition rates such as germ line transformation which is a more reliable means of determining the activity levels of a transposon in an organism. Most mutants had an activity that was at least half that of the *piggyBac* within the same assay which is not enough to consider them as hyperactive genetic tools for use in *D. melanogaster*. Another mutant S229F/F249L/C326S appeared to have no activity however it is possible that it maintains a very low activity which could be explored with more injections.

The most interesting observation made from the *D. melanogaster* trials was the discovery of altered target site preference in *AeBuster1* L368P/V597A and V597A. L386P/V597A has a significantly different target site preference at four of the eights base pairs while V597A is significantly different at three positions of the 8bp target positions when compared to wild-type *AeBuster1*. To the best of our knowledge this is the first report of altered target site preference due to amino acid alteration of the transposase sequence. Mutants L369P, L368P/ V597A and V597A were each independently analyzed in order to determine which of the amino acids substitutions were responsible for the change in preference (Figures 2.1-2.4). We have determined that V597A is the mutation primarily responsible for this shift however there may be some synergistic affects from L368P since the double mutant has a more varied preference from wild-type than V597A alone. Also intriguing is the shift towards a T in the second position and an A in the seventh position in mutant V597A. This shift resembles the target site preference of another distantly related *hAT* element *Hermes* (Arensburger *et al.* 2011). The *Hermes* target site consensus sequence is nTnnnnAn and this data suggests that by altering the

protein sequence a transposase can be made to revert back to an ancestral target site preference.

2.5.2 Future Directions

AeBuster1 does not appear to be an ideal candidate for use in Dipteran systems and there is little reason to continue testing the mutants described in this study. However if further studies were to be undertaken I would suggest that *AeBuster1* be codon optimized for *D. melanogaster* and then tested in the *D. melanogaster* cell line S2. This would still allow for high throughput screening but may provide better candidates for use in flies.

2.5 References:

Arensburger, P.A., O'Brachta, D.A., Orsetti, J. and Atkinson, P.W. (2005) *Hermes*, a new active *hAT* transposable element from the African malaria mosquito *Anopheles gambiae*. *Genetics*. **169:** 697-708

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

Blackman, R.K., Macy, M., Koehler, D., Grimaila, R. and Gelbart, W.M. (1989(Identification of a fully functional *hobo* transposable element and its use of germline transformation of *Drosophila. EMBO J*. **8**:211-217

Calvi, B.R., Hong, T.J., Findley, S.D. and Gelbart, W.M. (1991) Evidence for a common evolutionary origin of inverted repeat transposons in Drosophila and plants: *hobo, activator* and *Tam3*. *Cell*. **66**: 465-471

Craig, N.L. (2002) Mobile DNA: an introduction. In: Craig, R., Gellert, M., Lambowitz, A.M. (Eds), *Mobile DNA II*. ASM Press, Washinton, DC. 3-11

Crooks, G.E., Hon, G., Chandonia, J.M., Brenner, S.E. (2004) WebLogo: A sequence logo generator. *Genome Research*, **14**:1188-1190

Feschotte, C. and Pritham, E.J. (2007) DNA transposons and the evolution of eukaryotic genomes. *Annu Rev Genet*. **41**: 331–368

Handler, A.M., McCombs, S.D., Fraser, M.J. and Saul, S.H. (1998) The lepidopteran transposon vector piggyBac, mediates germ-line transformation in the Mediterranean fruit fly. *Proc Natl Acad Sci*. **95**:7520-7525

Kawakamai, K. and Shima, A. (1999) Identification of the *Tol2* transposase of the medaka fish *Oryzias latipes* that catalyzes excisions of a nonautonomous *Tol2* element in zebrafish *Danio rerio*. *Gene*. **240**:239-244

Lander, E.S. et al. (2001) Initial sequencing and analysis of the human genome. *Nature.* **409**: 860-921

Levy, A., Schwartz, S. and Ast, G. (2009) Large-scale discovery of the insertion hotspots and preferential integration sites of human transposed elements. *Nucleic Acids Res.* **38**: 1515-1530

McClintock, B. (1950) The origins and behavior of mutable loci in maize. *Proc. Natl. Acad. Sci. USA*. **36**: 344-355

Nene, V. *et al.* (2007) Genome sequence of *Aedes aegypti*, a major arbovirus vector. *Science* **316:**1718-1723

Obrochta D.A. and Atkinson, P.W. (1996) Transposable elements and gene transformation in non-drosophilid insects. *Insect Biochem. Molec. Biol.* **26**: 739-753

Plasterk, R.H.A. and Van Luenen, H.G.A.M. (2002) In Mobile DNA II (ed. Craig, N.L., Craigie, R., M. and Lambowitz, A.M.) ASM Press, Washington, DC, USA. 519-532

Sakar, A., Coates, C.J., Whyard, S., Willhoeft, U., Atkinson, P.W. and O'Brochta, D.A. (1997) The *Hermes* element from *Musca domestica* can transpose in four families of Cyclorrhaphan flies. *Genet*. **99**:15-29

Smith, R.C. (2007) Germ line expression of genetic markers and transposases in the yellow fever mosquito *Aedes aegypti*. University of California, Riverside

Trauner, J., Schniko, J., Lorenzen, M.D., Shippy, T.D., Wimmer, E.A., Beeman, R.W., Klinger, M., Bucher, G. and Brown, S.J. (2009) Large-scale insertinal mutagenesis of a coleopteran stored grain pest, the red flour beetle *Triblolium castaneum*, identifies embryonic lethal mutations and enhancer traps. *BMC Biology*. **7**:73

Walbot, V. (1992) Strategies for mutagenesis and gene cloning using transposon tagging and T-DNA insertional mutagenesis. *Annu. Rev. Plant Phys. Plant Mol. Biol.* **43**: 49-82

Zhou, L., Mitra, R., Atkinson, P.W., Hickman, A.B., Dyda, F. and Craig, N.L. (2004) Transposition of hat elements links transposable elements and V(D)J recombination. *Nature*. **432**: 955-1001

Mutant	*pBac* DT	*AB1* DT	*pBac* Events	*AB1* Events	*pBac* rate	*AB1* rate	*AB1/PBac*	Stdev *pBac*	Stdev *AB1*
L368P/V597A	190,400	210,000	129	73	$6.78x10^{-4}$	$3.48x10^{-4}$	.513	$3.5x10^{-4}$	$2.51x10^{-4}$
V168A/T324A/N465S	41, 200	44,200	5	7	$1.21\ x10^{-4}$	$1.58x10^{-4}$	1.62	$1.37x10^{-4}$	$1.05x10^{-4}$
S229F/F249L/C326S	63 400	61,000	7	0	$1.10x10^{-4}$	0	0	$1.46x10^{-4}$	0
E260G/S485L	32,200	27,600	19	6	$5.9x10^{-4}$	$2.1x10^{-4}$	.355	$5.18x10^{-4}$	$2.08x10^{-4}$
L368P	177,200	180,400	152	63	$8.58x10^{-4}$	$3.49x10^{-4}$	.407	$5.38x10^{-4}$	$3.98x10^{-4}$
L348V/L368P	39,600	40,600	16	4	$1.9x10^{-4}$	$6.10x10^{-5}$	0.32	$4.04x10^{-4}$	$9.85x10^{-5}$
V597A	160,400	265,000	85	24	$5.30x10^{-4}$	$9.06x10^{-5}$	0.17	$1.22x10^{-3}$	$2.35x10^{-3}$
Wild-Type *AeBuster1*	137,000	144,600	26	9	$1.90x10^{-4}$	$6.22x10^{-5}$	.33	$1.58x10^{-4}$	$3.57x10^{-5}$

Table 2.1. Transposition rates for *AeBuster1* mutants vs Wild-Type *piggyBac.* DT donor titer, pB *piggyBac*, Events represents the number of unique transposition rates. AB11/Pbac is the ratio of *AeBuster1* to *piggyBac* transposition rate. Stdev is the standard deviation within each of injection sites.

bp	L368P	L368P/V597A	V597A
1	0.2095	0.07384	.1797
2	0.6282	.2159	.1953
3	0.07124	$2.561e^{-12}$	$3.949e^{-6}$
4	0.2468	$4.802e^{-5}$	.2674
5	1	1	$2.775e^{-3}$
6	0.07683	$1.913e^{-07}$	$5.907e^{-4}$
7	0.1383	$4.787e^{-3}$	.2645
8	0.1019	0.1920	.5947

Table 2.2. States the p value for relatedness of the base pairs 1-8 of the duplicated target region of each mutant to the wild-type *AeBuster1*. Any value >.01 was considered significant.

AeBuster1 Wild Type	Target site	Number of integrations	Site	V597A	Target site	Number of Integrations	Site
	CATCAAGA	1	54 (+)		CTCTAGAG	5	1993 (+)
	GTTTAAGG	1	83 (-)		GACTACAT	1	509 (+)
	ATTCAAAT	1	144 (+)		GTTTATAT	1	389 (+)
	ATTTAGAC	1	168 (+)		TTATATTA	1	2495 (+)
	ACTTAACC	1	207 (-)		ATTTACAC	1	807 (+)
	ATTTACAC	1	810 (+)		TTGTATTT	1	2059 (-)
	GATTAAAG	3	943 (+)		ATTTACAC	2	810 (-)
	CTTTAATC	1	943 (-)		GTTTATTT	1	869 (-)
	CTCTAGAG	2	1993 (+)		GACTACAT	1	509 (-)
	ACTTATAG	1	2160 (+)		GTTTATTT	1	869 (+)
	CTATAAGT	1	2160 (-)		ATGGTTTA	1	866 (+)
	ACATAGCC	1	2207 (-)		ATGGGATA	2	2072 (-)
	AGCTAACA	2	2214 (-)		TGTTAGCT	1	2214 (-)
	TTATAACA	2	2381 (+)		GGCTACTC	1	191 (+)
	GTTTACCA	1	2425 (+)		GTTTATTT	1	869 (+)
	TACCAGAT	1	2428 (+)		GCTAGAGC	1	1956 (-)
	GCTTAAAT	2	2445 (+)				
	GATTAGAC	1	2470 (+)				
	GTCTAATC	1	2470 (-)				

Table 2.3. Integration sites of *AeBuster1* and V597A in *Drosophila melanogaster* transposition assays. The *AeBuster1* target site information is taken from Smith (2007). The + and –signs refer to the orientation of integration into pGDV1.

L368P	Target site	Number of Integrations	Site	L368P/V597A	Target site	Number of Integrations	Site
	CTCTAGAG	11	1993 (+)		ATCTACCT	1	1828 (-)
	GGCTACTC	3	191 (+)		CTCCAATT	5	958 (+)
	GATTAAAG	3	943 (+)		GATTAGAC	1	2470 (+)
	TGCTATGA	1	256 (+)		GGCTACTC	4	191 (+)
	GGCTACTC	2	191 (+)		TGTTAGCT	6	2213 (-)
	GATTAAAG	1	943 (-)		CTCTAGAG	4	1993 (+)
	ACATAGCC	1	949 (-)		TGTTAGCT	5	2213 (-)
	ATTTACAC	1	810 (+)		TGTTAGCG	1	2226 (+)
	GTTTATTT	1	868 (+)		AGCTAACA	1	2221 (-)
	CCTTAAAC	1	83 (+)		AACCTATG	1	256 (+)
	CTCTAGAG	1	1993 (-)		CTCTAGAG	1	1993 (-)

Table 2.4. Integration sites of L3628P and L368P/ V597A in *Drosophila melanogaster* transposition assays. The + and –signs refer to the orientation of integration into pGDV1.

Wild-Type *AeBuster1*

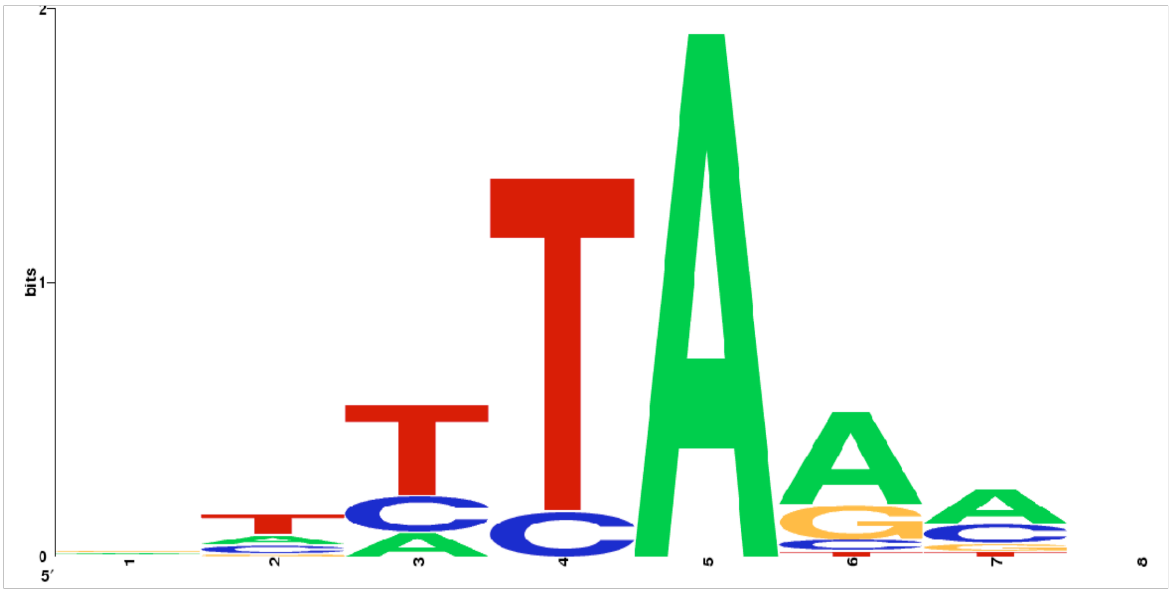

Figure 2.1. Weblogo of the consensus target site preference for Wild-type *Aebuster1*.

L368P

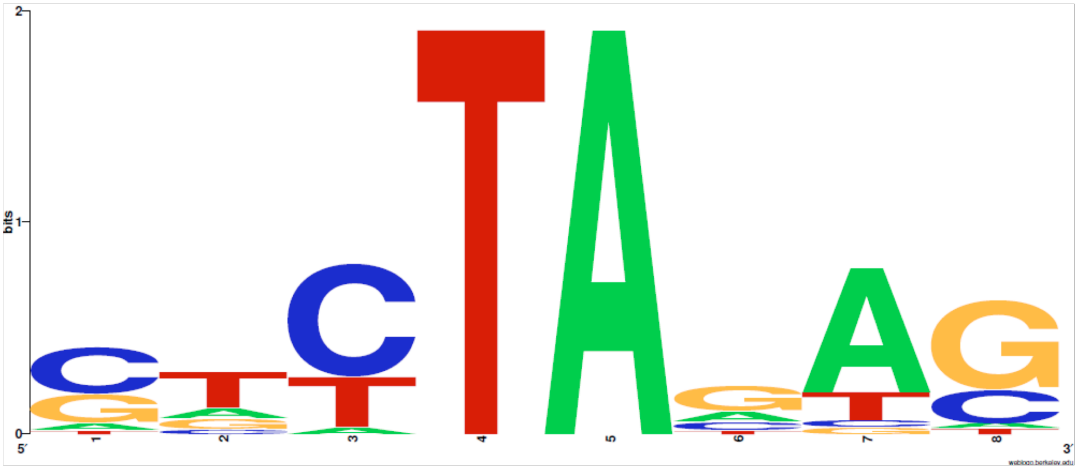

Figure 2.2. Weblogo of consensus target site preference for *AeBuster* 1 with mutation L368P.

L368P/V597A

Figure 2.3. Weblogo of consensus target site preference for *AeBuster* 1 with mutation L368P/V597A.

V597A

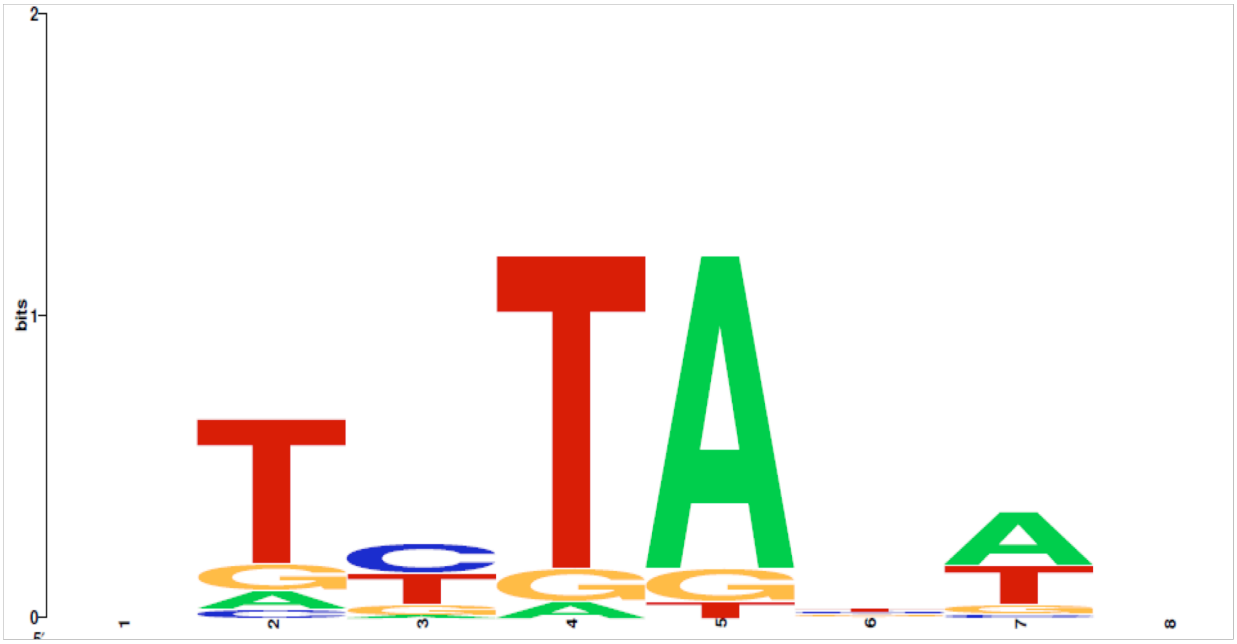

Figure 2.4. Weblogo of consensus target site preference for *AeBuster* 1 with mutation V597A.

Chapter 3

HyPBase Transposase Behavior in *Drosophila melanogaster* and *Aedes aegypti*

3.1 Abstract

Completed whole genome projects have revealed the diversity of this class of transposons in eukaryotes there now being at least 12 superfamilies of them identified. The majority of these are inactive and so relatively few have been developed as genetic tools however their impact on genetics has been profound, leading to sophisticated genetic techniques such as gene tagging, and enhancer and gene trapping that are now widely used to identify the function of genes and regulatory sequences (Wilson et al. 1989; Spradling *et al.* 1995; Clark *et al.* 2004,). Amongst these active class II transposons , *piggyBac* from the cabbage looper moth, *Trichoplusia ni*, has a wide host range extending into protozoans, rodents and humans, and has the unique property of excising from donor DNA leaving a perfectly repaired empty site (Handler *et al.* 1998; Lobo *et al.* 1999; Mitra *et al.* 2008). The mobility properties of this element have led to strategies to further improve its effectiveness as a genetic tool through the selection of a hyperactive mutant *piggyBac*, *hyPBase* (Yusa *et al*.,submitted). *HyPBase* was found to be highly active in mouse embryonic stem cells (ESC) with no apparent deleterious effects. We have tested the somatic and germ-line activity of *hyPBase* in two Dipteran systems: *Drosophila melanogaster* and the yellow fever mosquito, *Aedes aegypti*. Somatic transposition assays revealed that while there was a trend toward increased transposition frequency there was no significant difference between *hyPBase*'s transposition rate and wild-type *piggyBac* from *T. ni* in either *D. melanogaster* or *Ae.*

aegypti. The *hyPBase* germ-line transformation rate in *D. melanogaster* was 0.37% which was significantly lower than the wild type *piggyBac* of 10.5%. Interestingly the *hyPBase* GO flies had an increased sterility rate of 39.3% while the wild type *piggyBac* resulted in a sterility rate of 8.13%. After further investigation we found that the ovaries of sterile *hyPBase* injected flies were completely dysgenic. *HyPBase* transformation experiments in *Ae. aegypti* resulted in zero transgenic progeny produced. This study has demonstrated that while *hyPBase* may be an effective tool in mammalian cell culture studies we did not find it to be an effective tool for Dipteran systems in its current state.

3.2 Introduction

The *piggyBac* transposon from the cabbage looper, *Trichoplusia ni*, has proven to be a highly efficient transposon in invertebrates, vertebrates and protozoa and is currently used in gene therapy and transposon-based strategies to identify mammalian genes (Ding *et al*. 2005; Wilson *et al.*2007). *PiggyBac* integrates into TTAA nucleotide sites and due to its unique staggered cleavage upon excision, leads to the perfect repair of donor sites meaning that it can be introduced into a genome to achieve a desired goal and then precisely removed (Wang and Frasier 1993; Elick *et al*. 2007; Mitra *et al.* 2008). Another advantage of *piggyBac* is its ability to insert large (9.1-14.3kb) segments of genomic material without a loss of efficiency (Li *et al.* 2001; Li *et al*. 2005). *PiggyBac* has proven its utility as a gene vector in a wide variety of insects including but not limited to: *S. fruggiperda, Ae. aegypti, An. stephensi, An. albimanus, An. gambiae* and *T. castaneum*

(Olsen *et al.* 1994; Gueiros-Filho and Beverly 1997; Goryshin and Reznikoff 1998; Handler *et al.* 1998; Sundaraajan *et al.* 1999; Grossman *et al.* 2000).

PiggyBac preferentially inserts within genes, with 50% to 67% of insertions occurring within transcriptional units in human and mouse systems (Ding *et al.* 2005, Wilson *et al*. 2007). Perhaps most important to this study, is that the element is not sensitive to over-production inhibition in which elevated levels of the transposase causes decreased transposition. *PiggyBac* has been used to modify a wide variety of organisms and is one of the most commonly used transposases in insects. While transformation of mosquitoes has been obtained, frequencies remain low, typically less than 4% for *Mos1*, 7% for *piggyBac* and 6% for *Hermes* in *Ae. aegypti*, the mosquito species most amenable to genetic manipulation (Adelman *et al.* 2002). One strategy to increase the transformation rate of mosquitoes, particularly *Ae. aegypti*, is to generate transposases that have hyperactivity.

There is a precedent for increasing transposition rates in class 2 transposons through mutagenesis (Baus *et al.* 2005). The *Tc1*/*mariner* transposon *Sleeping Beauty* was synthesized from the alignment of inactive transposons belonging to the *salmonid* subfamily of elements (Ivics *et al.* 1997). Endogenous DNA transposons from this family are completely inactive in vertebrate genomes and are thought to have lost functional ability many millions of years ago, with only their inactive remnants found throughout vertebrate genomes (Yant et al. 2004). This restoration of activity demonstrates how disabled transposons can be restored through the manipulation of a transposon sequence (Yant *et al*. 2004). It therefore stands to reason that currently active transposable

elements that can be engineered a produce higher transformation. Progress in generating hyperactive transposons for use in insect transgenesis has been at best, marginal. A hyperactive form of the *P* transposase was obtained by selectively altering serines in to alanines testing the hypothesis that inhibition of any phosophorylation may increase activity (Beall *et al*. 2002). One mutant S129A was obtained which had a 172% increase in excision frequency in embryos and a 3 to 5 times increase in transformation frequency (Beall *et al.* 2002). Several hyperactive forms of the *Mos 1* and *Himar* transposase were obtained using a papillation assay in *E. coli*; however none retained any activity when tested in *D. melanogaster* (Lampe *et al.* 1999).

Recently a mutated *piggyBac* element named *hyPBase* was created and found to be hyperactive in mammalian cells (Yusa *et al.* 2011). *HyPBase* is a mammalian codon optimized *piggyBac* which contains seven amino acid substitutions. The combination of five mutations I30V, G165S, S103P, M282V, S509G, N570S and N538K were discovered using random mutagenesis and high throughput yeast assays (Yusa *et al*. 2011). The authors found that *hyPBas*e has a 17-fold increase in excision and 9-fold increase in integration when tested in mouse ESCs (Yusa *et al*. 2011). The study also found that *hyPBase* did not affect the genomic integrity of the host cells and that footprints were found at a frequency of approximately 1% which is comparable to the wild-type transposase. In this study we test to see whether or not *hyPBase* is capable of achieving hyperactive levels of germ-line transformation in *D. melanogaster* and *Ae. aegypti.*

Doherty *et al.* (2011) took the Yusa et al. (2011) experiments to the next step moving from yeast and embryonic mouse cells into human HeLA and HEK293 cells and *in vivo* in somatic cells in mice. This study compared *hyPBase* to wild-type *piggyBac* and a hyperactive *SleepingBeauty* transposon known as *SB100X*. *SB100X* reportedly has a transposition rate 100 times more active than the original *SleepingBeauty* in a variety of human cells (Mates *et al*. 2009). *HyBase* activity was tested in HeLa cells and found to be more active than *SB100X* at every transposon to transposase ratio tested (Doherty *et al*. 2011). *HyPBase* activity in HeLA cells was also shown to be unaffected by the addition of N- terminal tags (Doherty *et al.* 2011). This is important because it affects a transposons usefulness in technology such as gene therapy where it may be necessary to add N-terminal tag. Perhaps most importantly is that hyPBase increases long-term gene expression when introduced into mice via hydrodynamic tail vein injections (Doherty *et al*. 2011). Constructs using hyPBase containing a luciferase gene were injected into mice and it was found that 6 months post injection the hyPBase transposon resulted in approximately 10-fold greater luciferase signal than mice injected with either wild-type *piggyBac* or *SB100X* (Doherty et la. 2011).

3.3 Materials and Methods

3.3.1 Lab strains

D. melanogaster strain C-S;*w* was used in all fly experiments; the Orlando strain of *Ae. aegypti* was used in all mosquito experiments. This strain is maintained at the

University of California, Riverside under standard insectary conditions as described in Munstermann (1997) with the one exception that larvae are fed on a diet of Milkbone Original Dog Biscuits mixed with Red Star Specialty Nutrex 55. Adults were fed on a 10% sucrose solution and females blood fed on mice. The mosquitoes are maintained at 26° C with a photoperiod of 16:8 (L:D).

3.3.2 *PiggyBac* donor plasmid

The pBacGoEGFP plasmid was constructed by amplifying the left and right arms of piggyBac using the [3xp3-EGFPafm] (Horn and Wimmer 2000) . The left end, TSD, a short region of the ORF and a portion of flanking DNA was amplified with the primers SacII-pBac LE F (5'-ATACCGCGGTCTTTTTTAACCCTAGAAAGATAGTCTGCC-3') and XbaI-pBax LE R +ORF (ATATCTAGAGCTCATCGTCTAAAGAACTACCC-3') Next the right end, TSD and flanking plasmid were amplified with the primers KpnI-pBac RE R (5'-ATAGGTACCCTATTATTAACCCTAGAAAGATAATCATATTGTG-3') the left and right portions were digested using restriction sites incorporated into the PCR products and inserted into pBluescript II KS+ using the *SacII-XbaI* and *KpnI-XhoI* sites. The resulting plasmid pBSpBacLR was then linearized through digestion with *XbaI* as a means of incorporating a *NheI* linerarized pGENToriEGFP plasmid thus creating pBacGoEGFP plasmid. The pGENToriEGFP plamid was created through the excision of the lacZα ORF from the pGENToriα plasmid by digestion with *BamHI* and *NheI* and replaced by a *BamHI/NheI* fragment containing the EGFP ORF. The *BamHI/NHeI* EGFP fragment was generated though PCR amplification of the pBac[3xp3-EGFPafm] ORF

using the primers *BamHI* EGFP start (5'-ATAGGATCCACCGGTCGCCACCATGGTG-3') and *NheI* EGFP stop (5' ATAGCTAGCTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGG-3'). Following PCR the product was digested with *BamHI* and *NheI* in preparation for ligation.

3.3.3 *PiggyBac* helper plasmid

The *piggyBac* helper construct was made through the removal of the transposase ORF from pGalsHYPBASE (Yusa *et al*. 2011) using sticky-ended *SpeI-XhoI* ends and ligation into the vector pKH70new. pKH70new is an altered form of plasmid pkhsp70 (Arensburger *et al.* 2005) in which the hsp70 5' promoter and 3' terminal sequences were moved so that the restriction enzyme sites are found between the two regions allowing for easy excision and insertion of fragments between pgalsHYPBASE and pKh70new. pGalHYPBASE was derived from the p414GalS plasmid with the addition hyPBase transposase between base pair 3245 and 3246 (Mumberg *et al*. 1994). The pGalHYBASE plasmid was given to our laboratory from the Craig Laboratory at Johns Hopkins University.

3.3.4 *PiggyBac* reporter plasmid used for transformations.

The pBac[3xP3-EGFP]af donor plasmid was previously described (Horn and Wimmer 2000).

3.3.5 *Hermes* donor plasmids used for transposition assays.

pHDG1 was the *Hermes* donor plasmid used in transposition assays. pHDG1 includes the complete left and right ends of the *Hermes* transposon with none of the *Hermes* open frame included. The left end was amplified from pHERKS3, also called pHermesKSacOα (Sarkar et al 1997) by PCR using the primers:

HL8bpSac (5'-TGAGAGCYCGYCGTATCAGAGAACAACAACAAGTGGCTTATTTG-3')

HLinternalRxba (5'-TGATCTAGACTCTACTACTGCTTAAACTCTTTGAGGATGAACAG-3')

The product was then digested with *SacI* and *XbaI* and cloned into pBLuescript SK+ producing pSHLfull. The right end was amplified from pHERK3 by PCR with the primers HRinternalFxho (5'-TGACTCGAGGACTGGAAAAATTTAAGTTTAAAAGAAGC-3') AND HR8bpRKpn (5'-TGAGGTACCATACAGACAGAGAACTTCSAACAAGTCACAGGC-3') and digested with *XhoI* and *KpnI* and cloned into pBluescript SK+ digest with the same enzymes to produce pBSHRfull. pBSHRfull was digested with *KpnI* and *XhoI* to excise the right end which was cloned into pBSHLfull which was digested with the same enzymes to produce pBSHLRfull. pBSHLRfull was digested with *XbaI* and ligated with pGentOriaα digested by *NheI* to produce the donor plasmid pHDG1

pGentOriα was made by digesting pK19 with *NaeI* and *NheI* to isolate the *E. Coli* replication origin and *LaZα* gene and *NheI* site was filled by Klenow fragment. pFastBac

ORF (Invitrogen) was digested using *MscI* and *SnaBI* to isolate the GentR ORF. The two fragments were then ligated together to create pGentOriα.

3.3.6 *Hermes* Helper Plasmid Used for Transposition Assays

pKHSHH14 was the *Hermes* helper plasmid used and was described in Sakar *et al.* 1997.

3.3.7 Target Plasmid Used for Transposition Assays

pGDV1 has been described previously (Sakar *et al*. 1997) .

3.3.8 Plasmid Mixes

Microinjection of *D. melanogaster* involves several variables that can be difficult to hold constant; therefore it is necessary to create an injection mix that is able to eliminate the background noise inherent to the process. In order to compare relative rates of transposition frequency across trials there must be a means of normalizing the data. This can be achieved through the use of a five plasmid interplasmid transposition assay as developed in Smith (2007). In this assay *hyPBase* transposon was tested by co-injecting the appropriate helper and donor plasmids along with the *Hermes* donor, *Hermes* helper and pGDV1 as a target. This injection mix contains the following: 250ul of *pBacGoEGFP* donor, 250ul of the mutant *hyPBase* helper-transposase, 250ul of *Hermes* donor, 250ul of *Hermes* helper-transposase and 1000ul of pGDV1 as a donor plasmid. By using *Hermes* in the assay along with *hyPBase* we can compare the number of transposition events of the *hyPBase* mutated transposase relative to the number of *Hermes* transposition events, thus transposition rates can be normalized between trials. In each of the trials the transposition events produced from separate transposases can be

distinguished by the markers present in the donor plasmids. The *piggyBac* donor plasmids contains the EGFP gene and therefore expresses blue pigment in the presence of IPTG- XGAL where as the *Hermes* donor plasmid contains the *LacZ* gene.

3.3.9 Microinjection into *D. melanogaster*

Reference back chapter two page section 2.3.9 page 44 for methods.

3.3.10 *D. melanogaster* Transposition Assay

Reference back chapter two page section 2.3.10 page 44 for methods.

3.3.11 Microinjection into *Ae. aegypti* Embryos

Ae. aegypti adults were blood fed on anaesthetized mice for 10 minutes. Females were provided with a small tub of water with a Whatman paper filter in it and allowed to lay eggs for 30 minutes. The embryos were then aligned on double stick tape and mounted on a plastic slide and covered with halocarbon oil 27. The embryos are injected in the same manner as described for *D. melanogaster*. After microinjection the oil was washed away and the slides were placed in a pan of diH_2O. After 16-18 hours post injection the embryos were heat shocked at 42° C for one hour. The embryos were then processed using the transposition assay described above.

3.3.12 Insect Transformation

Transformation procedures for *D. melanogaster* and *Ae. aegypti* were described previously (Rubin and Spradling, 1982, Jasinskiene *et al*., 1998).

3.3.13 Ovary dissections

D. melanogaster ovaries were dissected two weeks after GO adults were crossed. Mosquito ovaries were dissected after the female had imbibed three blood meals. Each *Ae. aegypti* female had the opportunity to lay eggs after the first two blood feedings and if she did not produce viable offspring she was considered sterile. Dissected ovaries were placed in a glycerol based mounting fluid and photographed using a Zeiss Axioskop2.

3.4 Results

3.4.1 *HyPBase* Shows no Significant Increase in Somatic Transposition in *D. melanogaster*

We evaluated the somatic transposition rate of *hyPBas*e in *D. melanogaster* using a normalized interplasmid transposition assay in which the activity of the *hyPBase* transposase was compared with an internal control transposase, each transposase being placed under the control of the *D. melanogaster* hsp70 promoter.

HyPBase had a transposition rate of 1.6×10^{-3} while the transposition frequency of the wild-type *piggyBac* transposase was 1.12×10^{-3} which was consistent with the frequency of wild-type *piggyBac* transposition in this insect observed by others (Table 3.1) (Lobo and Fraser 1999). This difference in normalized transposition activities between *hyPBase* and wild-type *piggyBac* transposase was not statistically significant (Figure 3.1). Since *hyPBase* in mammalian cells showed clear hyperactivity we were surprised by this outcome but also interested to determine whether *hyPBase* had

increased germ-line activity in *D. melanogaster*, which is not measured in the somatic interplasmid transposition assays.

3.4.2 *HyPBase* Produces a High Rate of Sterility in *D. melanogaster*

Germ-line transformation was achieved by injecting the *hyPBase* helper plasmid and the reporter plasmid pBac[3xP3-EGFP]af into *D. melanogaster* embryos. Wild-type *piggyBac* transposase and the same *piggyBac* reporter were injected as controls at the same plasmid concentrations.

A total of 209 embryos injected with wild-type *piggyBac* transposase survived to adulthood (a survival rate of approximately 43%) and of those 192 (92% of G0 crosses) generated progeny upon backcrossing (Table 3.2). Screening the G1 progeny for GFP expression revealed that approximately 10% of the wild-type *piggyBac* transposase crosses produced transgenic offspring. Injections with *hyPBase* resulted in 239 embryos surviving to adulthood (a survival rate of 42%) of which 145 were fertile (67% of G0 crosses). Surprisingly, only 0.68% of the *hyPBase* fertile crosses resulted in transgenic progeny (Table 3.2). These data are in contrast to the activity of *hyPBase* in somatic cells however the decline in transformation rate is matched with an increase in sterility. The overall rate of sterility found with wild-type *piggyBac* transposase was approximately 8% while it increased to approximately 39% with *hyPBase* (Table 3.2). This significant increase in sterility associated with the use the *hyPBase* could be due to the transposase causing higher or sustained rates of insertion of *piggyBac* thereby increasing its mutagenic effect, a direct effect of the transposase on chromosomal DNA such as double

strand breaks that are repaired inefficiently, or a combination of both. To distinguish between these possibilities, we injected the *hyPBase* helper plasmid alone into embryos and measured G0 sterility. These experiments yielded a sterility rate of 23.0% indicating that increased integration of the transposon alone is probably not the only explanation of the increased sterility (Table 3.2).

The increased sterility led us to investigate the physiological effects of *hyPBase* in G0 females. Gonadal dystrophy has been intimately associated with hybrid dysgenesis that arises through the release of suppression of *P* element mobility when *P* element containing sperm fertilize an egg devoid of *P* elements in its genome (Kidwell *et al.* 1977). More recently piRNAs have been proposed to be the actual basis of hybrid dysgenesis with the deficiency of these in the egg allowing *P* elements present in the sperm genome to move in the fertilized zygote resulting in insertional mutagensis, gonadal atrophy and hybrid dysgenesis (Simmons *et al.* 1990; Brennecke *et al.* 2008). To determine if the sterility observed in G0 crosses arising from using *hyPBase* was accompanied by gonadal atrophy we dissected ovaries from sterile female *D. melanogaster* that were either injected with wild-type *piggyBac* transposase or *hyPBase* (Figure 3.2). Most ovaries taken from sterile flies injected with wild-type *piggyBac* were smaller in size yet still maintained the overall structure of the fertile ovary as previously described (Spradling 1993) (Figure 3.2C). However all but one ovary dissected from sterile flies injected with *hyPBase* were completely atrophied with only the peritoneal sheath that surrounds the ovaries remaining (Figure 3.2D). Ovaries from the sterile

females injected with *hyPBase* alone also displayed the same type of atrophy (Figure 3.2E).

3.4.3 The Activity of *hyPBase* in *Ae. aegypti*

Despite the high degree of sterility caused by *hyPBase* in *D. melanogaster* we reasoned that the generally low frequency of transformation obtained by using exogenous transposons in *Ae. aegypti* may result in a similar decrease in the activity of *hyPBase* and perhaps therefore lead to at least a detectable transformation frequency. Interplasmid transposition assays using *hyPBase* revealed no significant difference in its activity in *Ae. aegypti* relative to that obtained using wild-type transposase and so reflected what was observed in *D. melanogaster* (Table 3.3). We attempted germ-line transformations using the same helper and reported plasmids as used for *D. melanogaster* but did not recover any transgenic progeny from 21 fertile crosses (Table 3.4). As was seen in *D. melanogaster* we did observe a high degree of sterility (77%) and so examined the ovaries of these G0 females and compared them to ovaries dissected from adult female mosquitoes of the same age. The ovaries from these G0 females displayed the same type of atrophy seen in corresponding *D. melanogaster* females. While we did not perform injections with helper plasmid alone, we concluded that *hyPBase* had the same type of hyperactive behavior in *Ae. aegypti* as it did in *D. melanogaster*.

3.4.4 Transposase concentration has no affect on transformation rate

Transformation injections were performed with three different concentrations of the *hyPBase* transposase: 250 ng/ul, 50 ng/ul and 25 ng/ul. Transformation rates were not

affected by decreasing concentrations of the transposase in the injection mix. The transformation rate of the 50ng/ul and 25ng/ul experiments remained at 0% with no transgenic offspring produced. However the sterility rate did decrease with reduced transposase concentration. The sterility rate decreased approximately 10% with each dilution (Table 3.5).

3.5 Discussion

3.5.1 *hyPBase's* usefulness in Dipteran systems

The use of transposons as genetic tools has led to interest in generating hyperactive forms of them to further increase their efficiency. These gain of function mutations are rare and their usefulness can be balanced against their hyperactivity being so high that their mutagenic properties exceeds their ability to be used effectively. In insects, gains have been modest. A hyperactive form of the *P* transposase obtained through directed mutagenesis of serine residues potentially targeted by kinases showed a small increase in transposition in somatic assays with no substantial increase in germ-line transformation frequency (Beall *et al.* 2002). *Mos1* and *Himar* hyperactive mutants selected in *E. coli* using papillation assays had no activity at all when tested in *D. melanogaster* (Lampe *et al.* 1999). In contrast hyperactive mutants of the *Sleeping Beauty* transposase from teloest fish have shown repeated high rates of transposition in human cell culture and in live rodents illustrating that such mutants can be obtained and

used to improve existing transposon-based genetic techniques (Ivics *et al.* 1997). *Sleeping Beauty* is however, for reasons unknown, inactive in *D. melanogaster*.

The properties of the *piggyBac* transposon make it an obvious target for hyperactive mutagenesis strategies and, given *piggyBac's* utility as a gene vector across the animal kingdom, it is reasonable to propose that a hyperactive form of its transposase that functions in vertebrates may well preserve this activity in invertebrates, especially since this transposon was originally identified in an insect. *HyPBase* was isolated through screening for hyperactive mutants in a yeast-based transposition assay followed by their integration into a single open reading frame that contained seven amino acid substitutions relative to the wild-type transposase. (Yusa *et al.* 2011). *HyPBase* showed a 17- and 9-fold increase in excision and transposition frequency respectively in mouse embryonic stem cells and, in the vast majority of excisions, preserved the perfect excision of *piggyBac* from donor sites (Yusa *et al.* 2011). This hyperactivity was not accompanied by any obvious deleterious effects on the genomes of these cells as measured by comparative genome hybridization between genomic DNA isolated from two control stem cell lines and ten lines expressing hyPBase in which all *piggyBac* elements had been removed following excision (Yusa et al. 2011). While half of the *hyPBase* lines contained abnormalities, none were linked to any known *piggyBac* integration sites. Furthermore, abnormalities were found in control lines leading to the conclusion that these genomic alterations were most likely not due to action of *hyPBase* (Yusa et al, 2011).

To determine if hyperactivity was retained in insects, we evaluated the activity of *hyPBase* in the model insect *D. melanogaster* and in *Ae. aegypti*, a vector of several human pathogens. We concluded that *hyPBase* had the same activity as wild-type *piggyBac* transposase in interplasmid-based assays performed in somatic cells however its expression in germ-line cells led to high rates of sterility, accompanied by gonadal dystrophy, in both *D. melanogaster* and *Ae. aegypti*. In the former this was also associated with a low rate of genetic transformation while our results in *Ae. aegypti* suggest that *hyPBase* would not be likely to produce a high frequency of genetic transformation.

Indeed, the germ-line transformation experiments in flies revealed that the *hyPBase* transformation rate was dramatically decreased compared to wild-type *piggyBac* transposase. The overall wild-type transformation rate using wild-type *piggyBac* transposase was approximately 10% compared to *hyPBase* which had an overall transformation rate of under 1%. Wild-type *piggyBac* transposase generated a sterility rate of 10.9% while fly embryos injected with both the *hyPBase* helper and the reporter plasmids led to a sterility rate of 40.2%. Perhaps most intriguing is that injecting the *hyPBase* transposase without a donor element available for transposition generated an increased sterility rate of 23% suggesting that the overall sterility rate of approximately 39.3% seen in the presence of the *piggyBac* reporter was due both to the increased mutagenic action of the transposon as well as a direct effect of the transposase on the genomic DNA. The former could be caused either by an increase in stability of *hyPBase* or an increase in its enzymatic function. The later could, in turn, be due to the

mobilization of *piggyBac* end-like sequences incumbent in the genome or due to some direct enzymatic effect of *hyPBase* on chromosomal DNA.

Increased sterility in both species was associated with complete gonadal atrophy in G0 females. Gonadal atrophy is a hallmark of hybrid dysgenesis which has been best characterized for the *P* element of *D. melanogaster* (Kidwell et al. 1997). More recently, Brennecke *et al.* (2008) found that novel transposons introduced into a naive host genome led to increased sterility and complete dysgenesis of the ovaries. The *piggyBac* element has not been reported in *D. melanogaster* C-S, w^+ and only four *piggyBac* family elements with a total of 31 copies have been found in the reference genome (Liverpool) strain of *Ae. aegypti* although it is formally possible that the Orlando strain used in our injections has a different complement of these transposons. To our knowledge, this is the first example of *piggyBac* transposase activity leading to gonadal atrophy (Nene et al. 2007).

The hsp70 promoter used to control *hyPBase* is expressed in both the germ-line and somatic tissues. In order to determine any other mutagenic affects of *hyPBase,* sterile individuals were examined to count total appendages, abdominal, segments, mouthparts and wing structure. Testes were also dissected however the relatively small size of this organ made distinguishing between naturally occurring variation and true mutagenic affects difficult to determine. We could find no other phenotypic manifestations of *hyPBase* activity in these individuals. We conclude that either the effect of *hyPBase* is confined to the gondal tissue despite the fact that expression of the gene should be throughout the insect, or there are affects in the somatic tissue that we were unable to

detect. Consistent with the former explanation is our failure to detect an increase in transposition of *piggyBac* in the presence of *hyPBase* in somatic cells.

The *hyPBase* transposase may also have applications in genetic approaches to insect control since it can clearly act as a sterilization agent through gonadal atrophy, at least in females. This can be more comprehensively explored through the creation of genetic strains in which *hyPBase* is placed under the control of a conditional promoter and activator, such as the GAL4-UAS system. This approach would also allow any somatic effects of *hyPBase* expression through all developmental stages to be better addressed and would also allow any effects on testes development and sperm generation to be determined. In a similar vein, the placement of *hyPBase* under the control of a testes-specific promoter, such as the ß 2-tubulin promoter may help to confine expression to the testes, leading to strains that can mate but are sterile. That this transposase leads to similar affects in at least the ovaries of both *D. melanogaster* and *Ae. aegypti* indicates that this phenotype may appear across several other species.

While hyPBase in its current form is not useful as a genetic tool for the mobilization of piggyBac, its generation from the yeast-based screen illustrates that this approach does result in transposases with increased levels of activity and so should be applicable to other transposon systems. The effects we have observed in both *D. melanogaster* and *Ae. aegypti* are significantly more detrimental to the host organism than what was seen in rodent embryonic stem cell culture (Yusa et al., 2011). One explanation for this difference is that cell culture is not an appropriate platform for determining all the possible effects on phenotype that a hyperactive transposase may have

on a multicellular organism through development, in which case studying the effects on *hyPBase* on living rodents may reveal any pleiotropic effects. Alternatively, but not necessarily exclusively, *hyPBase* protein may be more efficiently produced in insects given that *piggyBac* was first detected in an insect. that the fact that the *hyPBase* ORF has been optimized for mammalian expression does not appear to be any significant impediment to its translation in these two insect species.

We do not know the biochemical basis of *hyPBase* hyperactivity, nor do we know which of the seven point mutations in it contribute most to this phenotype. The hyperactivity could be due to increased catalytic activity of the enzyme, increased stability, or a combination of both. A component of this does require the presence of the *piggyBac* transposon suggesting that increased insertional mutagenesis of the transposon is a significant factor. However it is also clear from our data that *hyPBase* also exerts a detrimental effect in the absence of a supplied *piggyBac* transposon. It remains possible that the analysis of each of the seven mutations of hyPBase may result in a transposase that provides enhanced activity in insects without any accompanying detrimental effects. At the very least the discovery of *hyPBase* validates the use of genetic screens based in yeast that generate transposases with enhanced and interesting properties that may be useful in metazoans. Furthermore, once a crystal structure of the *piggyBac* transposase is obtained, mapping these mutations onto it should provide valuable insights into how this important transposase functions.

3.5.2 Future Directions

The next step in this study would be to tease apart the affects each individual mutation has on transposition and transformation frequencies. It may well be that a decrease in transposition frequency could reduce the sterility rate and increase the transformation rate. I would also recommend that future screening of *piggyBac* mutants be performed in *D. melanogaster* S2 cells with codon optimized transposases. I believe this would yield better results than screening in yeast. It would also be interesting to determine what about the transposase is causing increased sterility. It maybe that the protein is causing double stranded breaks, therefore taking on a DNAse like activity. Isolating the protein and presenting it with plasmid may allow for the visualization of this activity. The *hyPBase* transposase is deserving of further study.

3.6 References:

Adams, M.D. *et al.* (2000). The Genome Sequence of *Drosophila melanogaster. Science*. **286**: 2185-2195

Adelman, Z. N., Jasinskiene, N. and James, A. A. (2002). Development and applications of transgenesis in the yellow fever mosquito, *Aedes aegypti*. *Mol. Biochem. Parasitol.* **121**: 1-10

Arensburger, P., Kim, Y.J., Orsetti, J., Aluvihare, C. O'Brochta, D.A. and Atkinson, P.W. (2005) The transposable element, *Herves*, from the African malaria mosquito *Anopheles gambiae*. *Genetics* **169**: 697–708

Baus, J., Liu, L., Heggestad, A.D., Sanz, S. and Fletcher, B.S. (2005) Hyperactive transposase mutants of the *Sleeping Beauty* transposon. *Mol. Ther.* **12**: 1148-1156

Beall, E.L., Mahoney, M.B. and Rio, D.C. (2002) Identification and analysis of a hyperactive mutant Form of *Drosophila P*-element transposase. *Genetics.* **162**: 217-227

Brennecke, J., Malone, C.D., Aravin, A.A., Sachidanandam, R., Stark, A. and Hannon, G.J. (2008) An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science*. **322**: 1387-1392

Ding, S., Wu, X., Li, G., Han, M., Zhuang, Y. and Xu, T. (2005) Efficient transposition of the *piggyBac* (PB) transposon in mammalian cells and mice. *Cell*. **122**: 473–483

Doherty, J.E., Huye, L.E., Yusa, K., Zhou, L., Craig, N.L. and Wilson, M.H. (2011) Hyperactive *piggyBac* gene transfer in human cells in vivo. *Human Cell Therapy.* At press

Elick, T. A., Bauser, C. A. and Fraser, M. J. (1996) Excision of the *piggyBac* transposable element in vitro is a precise event that is enhanced by the expression of its encoded transposase. *Genetica.* **98**: 33–41

Handler A.M., McCombs S.D., Fraser M.J. and Saul S.H. (1998) The lepidopteran transposon vector *piggyBac*, mediates germ-line transformation in the Mediterranean fruit fly. *Proc Natl Acad Sci* USA. **95**:7520-7525.

Ivics, Z., Hackett, P.B., Plasterk, R.H. and Izsvák, Z. *(*1997) Molecular reconstruction of *Sleeping Beauty*, *a Tc1*-like transposon from fish, and its transposition in human cells. *Cell*. **91**: 501-510.

Jasinskiene, N., Coates, C. J., Benedict, M. Q., Cornel, A. J., Rafferty, C. S. and James, A. A., Collins, F. H. (1998) Stable transformation of the yellow fever mosquito, *Aedes aegypti*, with the *Hermes* element from the housefly. *Proc. Natl. Acad. Sci USA*. **95**: 3743-3747.

Kidwell, M.G., Kidwell, J.F. and Sved, J.A. (1997) Hybrid dysgenesis in *Drosophila melanogaster*: A syndrome of aberrant traits including mutation, sterility and male recombination. *Genetics*. **86**: 813-833

Goryshin, I.Y., and Reznikoff, W.S. (1998) *Tn5 in vitro* transposition. *J. Biol. Chem*. **273**: 7367-7374

Grossman, G.L., Rafferty, C.S., Fraser, M.J. and Benedict, M.Q. (2000) The *piggyBac* element is capable of precise excision and transposition in cells and embryos of the mosquito, *Anopheles gambiae*. *Insect Biochem. Mol. Biol.* **30**: 909-914.

Gueiros-Filho, F.J. and Beverley, S.M. (1997) Trans-kingdom transposition of the *Drosophila* element *Mariner* within the protozoan *Leishmania. Science*. **276**: 1716-1719.

Handler A.M., McCombs S.D., Fraser M.J. and Saul S.H. (1998) The lepidopteran transposon vector *piggyBac*, mediates germ-line transformation in the Mediterranean fruit fly. *Proc Natl Acad Sci US.A* **95**: 7520-7525.

Horn, C. and Wimmer, E. A. (2000) A versatile vector set for animal transgenesis. *Dev Genes Evol*. **210**: 630–637.

Lampe, D.J., Akerley, B.J., Rubin, E.J., Mekalanos, J.J. and Robertson, H.M. (1999) Hyperactive transposase mutants of the *Himar1 mariner* transposon. *Proc. Natl. Acad. Sci. USA*. **96**: 11428-11433.

Li, X., Lobo, N., Bauser, C.A. and Fraser, M.J. Jr. (2001) The minimum internal and external sequence requirements for transposition of the eukaryotic transformation vector *piggyBac*. *Mol Genet Genomics.* **266**: 190-198

Li, X., Harrell, R.A., Handler, A.M., Beam, T., Hennessy, K. and Fraser, M.J. Jr. (2005) *piggyBac* internal sequences are necessary for efficient transformation of target genomes. *Insect Mol Biol.* **14**: 17-30

Lobo, N., Li, X. and Fraser, M. J., Jr. (1999) Transposition of the *piggyBac* element in embryos of *Drosophila melanogaster*, *Aedes aegypti* and *Trichoplusia ni*. *Mol. Gen. Genet.* **261**: 803–810

Mitra, R., Fain-Thornton, J. and Craig, N.L. (2008) *piggyBac* can bypass DNA synthesis during cut and paste transposition. EMBO J. **7:** 1097-109

Mumberg, D., Muller, R., and Funk, M. (1994) Regulatable promoters of *Saccharomyces cerevisiae*: comparison of transcriptional activity and their use for heterologous expression. *Nucleic Acids Res.* **22**: 5767-5768.

Munstermann, L.E. (1997) Care and maintenance of *Aedes aegypti* colonies. The Molecular Biology of Disease Vectors, J.M. Crampton, Beard, C.B. and Louis, C., ed. London Chapman and Hall

Nene, V. *et al.* (2007) Genome Sequence of *Aedes aegypti*, a major arbovirus vector. *Science* **316**:1718-1723

Olson, K.E., Higgs S., Hahn, C.S., Rice, C.M., Carlson, J.O. and Beaty, B.J. (1994) The expression of chloroamphenicol acetyltransferase in *Aedes albopictus* (C6/36) cells and *Aedes triseriatus* mosquitoes using a double subgenomic recombinant Sindbis virus. *Insect Biochem. Mol. Biol.* **24**: 39-48

Rubin, G. M. and Spradling, A. C. (1982) Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**: 341-347

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

Simmons, M.J., Raymond, J.D., Rasmusson, K.E., Miller, L.M., C.F. and Zunt, J.R. (1990) Repression of *P* Element-Mediated Hybrid Dysgenesis in *Drosophila melanogaster. Genetics.* **124**: 669-676

Spradling, A. C. (1993) The development of *Drosophila melanogaster*. Cold Spring Harbor. **1**: 1-70

Spradling, A. C., Stern, D. M., Kiss, I., Roote, J., Laverty, T. and Rubin, G. M. (1995) Gene disruptions using *P* transposable elements: An integral component of the *Drosophila* genome project. *Proc. Natl. Acad. Sci. USA*. **92**: 10824–10830

Sundaraajan, P., Atkinson, P.W. and O'Brachata, D.A. (1999) Transposable elements interactions in insects: crossmobilzation of *hobo* and *Hermes*. *Insect Mol. Biol.* **8**: 359-368

Wang, H.H. and Fraser, M.J. (1993) TTAA serves as the target sire for TFP3 lepidotpteran transposon insertions in both nuclear polyhedrosis virus and *Trichoplusia ni* genomes. *Insect Mol. Biol.* **1**: 109-1996

Wilson, M.H., Coates, C.J. and George, A.L. (2007) *PiggyBac* transposon –mediated gene transfer in human cells. *Molecular Therapy*. **15**: 139-145

Yant, S.R., Park, J., Huang, Y., Mikkelsen, J.G. and Kay, M.A. (2004) Mutational analysis of the N-terminal DNA-binding domain of the *Sleeping Beauty* transposase: critical residues of the DNA binding and hyperactivity in mammalian cells. *Mol. Cell. Biol.* **24**: 9239-9247

Yusa, K., Zhou, L., Li, M.A., Bradley, A. and Craig, N.L. (2011) A hyperactive *piggyBac* transposase for mammalian applications. *PNAS.* **108**: 1531-1536

D. Melanogaster **Transposition Assay**	**Number of Experiments**	**Donor Titer**	**Confirmed events**	**Transposition Rate**	**Standard Deviation**	**Ratio of *PiggyBac: Hermes***
Hermes Control I	12	134,400	7	5.2×10^{-5}	7.85×10^{-3}	
hyPBase	12	60,400	101	1.6×10^{-3}	2.48×10^{-4}	3.94
Hermes Control II	3	21,000	8	3.8×10^{-4}	4.5×10^{-4}	
piggyBac WT	3	25,000	28	1.12×10^{-3}	$1.06 \times^{-3}$	2.94

Table 3.1. Transposition rates in *D. melanogaster* when injected with *hyPBase, Hermes* and *piggyBac.* The standard deviations within each assay provided as well as the *piggyBac* or *hyPBase* to *Hermes* activity ratio. *The published *piggyBac* rate is taken from Smith 2007.

Transposase	Experiment	Total injected	Survived inj.	Total Eclosed	Total Crosses	Fertile	Sterile	Transgenics	Sterility Rate	Transformation Rate
Wild-type *PiggyBac*	I	169	104	59	59	55	4	7	6.77%	12.72%
	II	170	120	82	82	73	9	8	10.97%	10.96%
	III	152	117	68	68	64	4	4	5.88%	6.25%
Totals		491	341	209	209	192	17	19	8.13%	9.89%
Average			69.70%	61.10%	100%	91.20%	8.61%		8.80%	9.97%
Stdv			7.75%	6.33%	0%	2.70%	3.25%		2.97%	2.73%
hyPBase	I	150	75	38	38	13	25	0	65.60%	0%
	II	162	86	76	72	56	16	1	20%	1.78%
	III	91	58	30	30	24	6	0	20%	0%
	IV	157	120	103	99	52	47	0	47.40%	0%
Totals		560	339	247	239	145	94	1	39.30%	0.68%
Average			60.8%	69.15%	99.12%	60.10%	84.10%		38.20%	0.45%
Stdv			11.96%	20.08%	1.94%	20.88%	74.10%		22.30%	0.009
Transposase Only	I	192	160	143	143	110	33	0	23%	0%

Table 3.2. Sterility results from the transformation experiments in *D. melanogaster* using *hyPBase* and *piggyBac* as a control.

Ae. aegytpi **Transposition Assay**	**Number of Experiments**	**Donor Titer**	**Confirmed Events**	**Transposition Rate**	**Standard Deviation**
Hermes Control	3	46,800	1	2.14×10^{-5}	1.92×10^{-4}
hyPBase	3	19,600	3	1.53×10^{-4}	1.11×10^{-4}
Published *piggyBac**	4	611,200	35	5.72×10^{-5}	

Table 3.3. *Aedes aegypti* transposition assays results with *Hermes, hyPBase* and wild- type *piggyBac*. *Wild-type *piggyBac* data was taken from Smith 2007. The standard deviation within trials is given as well.

Transposon	Trials	Fertile	Sterile	Total Crosses	Sterility%	Transformation Rate
hyPBase	3	21	70	91	77%	0

Table 3.4. *Aedes aegypti* transformation rates with sterility when injected with *hyPBase* transposase and donor plasmid.

Transposase Conc.	Experiment	Total injected	Survived inj.	Total Eclosed	Total Crosses	Fertile	Sterile	Transgenic	Sterility Rate	Transformation Rate
50 ng/ul	I	210	128	119	110	74	36	0	32.72%	0
	II	146	87	72	63	40	23	0	36.51%	0
	III	163	99	83	72	46	26	0	36.11%	0
Totals		519	314	274	245	160	95	0		
Average			60.42%	86.52%	88.89%	64.88%	54.22%		35.11%	
Stdv			7.32%	5.61%	3.09%	2.08%	4.85%		2.08%	
25 ng/ ul	I	151	107	63	56	46	10	0	17.8%	0
	II	327	191	163	159	119	40	0	25.16%	0
	III	171	127	83	78	65	13	0	16.67%	0
Total		649	425	309	293	230	63	0		
Average			67.80%	69.86%	93.47%	80.10%	25.12%		19.90%	
Stdev			8.34%	13.79%	4.35%	4.00%	7.41%		4.60%	

Table 3.5. Represents transformation experiments with diluted transposase concentrations. Mosquitoes were injected with 50ng/ul or 25ng/ul and then survivorship and fertility were monitored. The averages and standard deviations for each column are given as well.

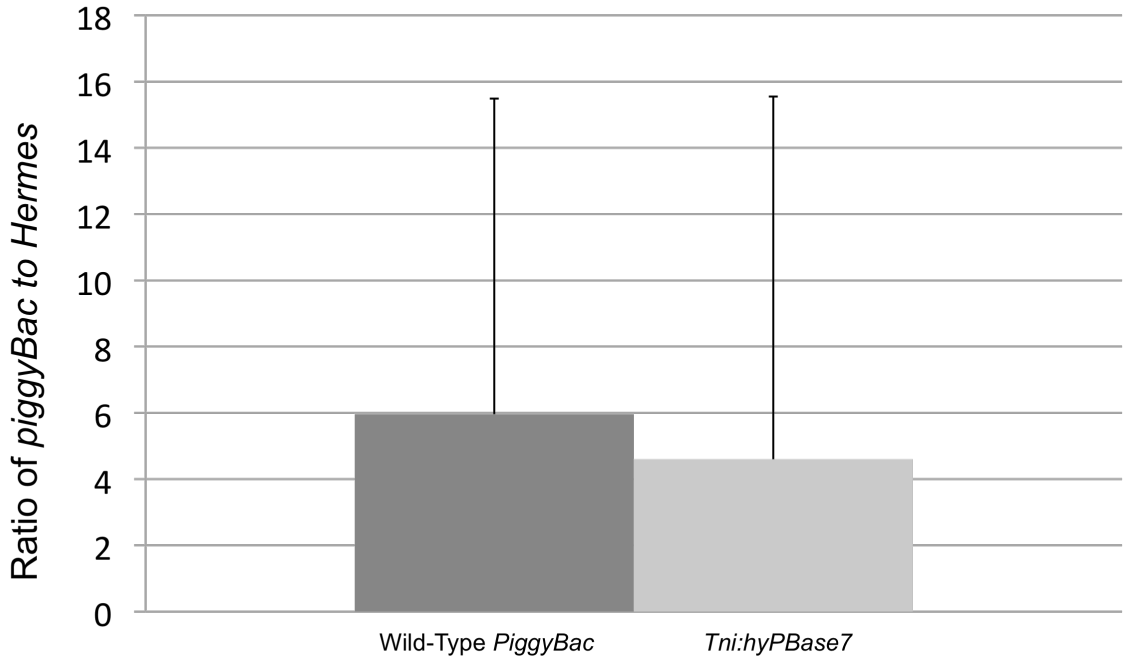

Figure 3.1. The normalized transposition ratios of *piggyBac:Hermes in Drosophila* transposition assays. The results are not statistically significant.

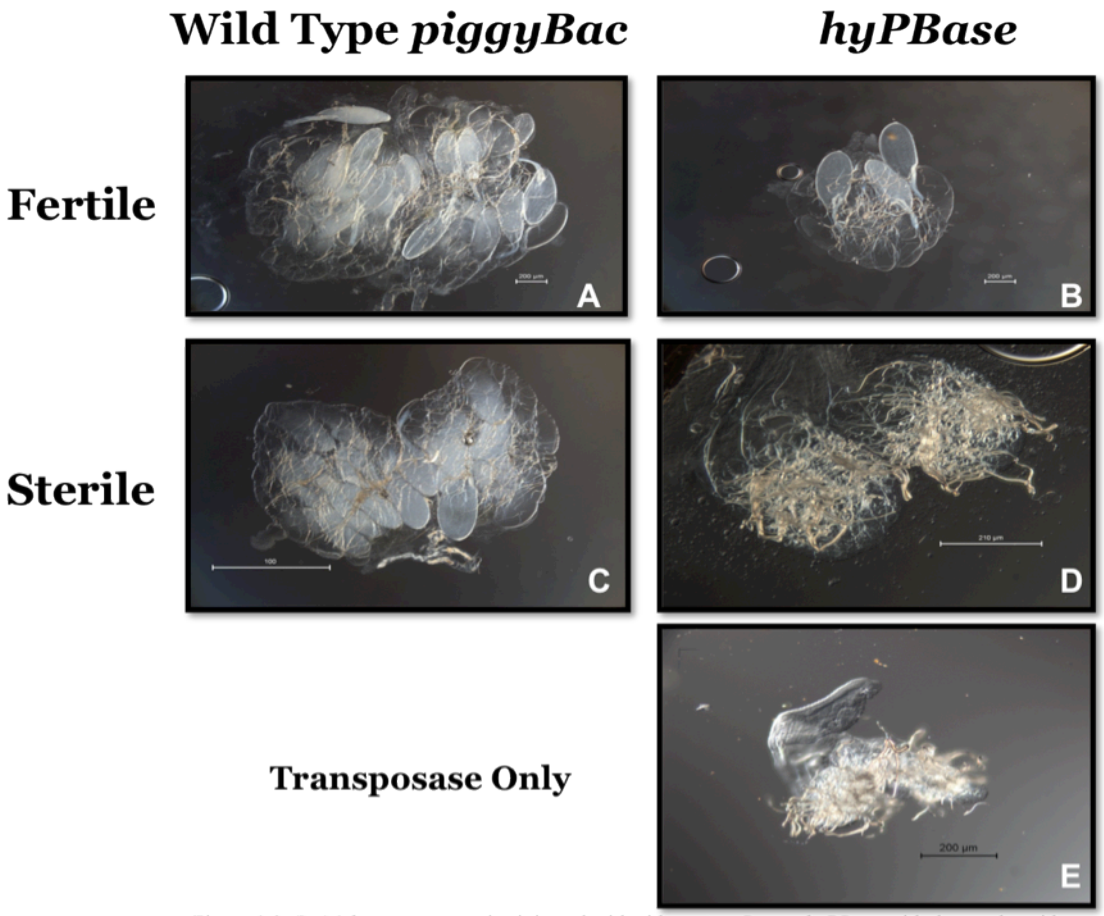

Figure 3.2. *D. Melanogaster* ovaries injected with either *piggyBac* or *hyPBase* with donor plasmid or the *hyPBase* transposase protein only.

Chapter 4

An Investigation into piRNA system in *Aedes aegypti*

4.1 Abstract

piRNA are 24-30nt long small RNA that are generated via a Dicer independent pathway known as the piRNA pathway. The piRNA pathway is least well understood of the small RNA pathways however piRNA have been linked with the control of transposon movement in the genome. The piRNA pathway in *Drosophila melanogaster* has been studied relatively well compared to non-model insects. In *D. melanogaster* the piRNA pathway is composed of Aubergine (AUB), Argonaute3 (AGO3) and PIWI. This study looks at the piRNA pathway in the important human disease vector *Aedes aegypti*. The *Ae. aegypti* piRNA pathway consists of one AGO3 and six PIWI proteins. Piwi 2 is the only germ-line specific PIWI protein. We created RNAseq libraries to determine the expression levels of each of the proteins involved in the piRNA pathway and the proper transcript structure. We performed immunoprecipitation and followed by mass spectrometery for the proteins AGO3, PIWI2 and PIWI7 and then created small RNA libraries from the piRNAs found in each sample. These libraries demonstrated that each of the PIWI proteins tested is associated with piRNA that map to transposons, with PIWI2 being associated with the greatest proportion of piRNAs mapping to transposons.

4.2 Introduction

Understanding genomic mechanisms of transposon regulation is essential to the optimization of transposons as genetic tools. Research into transposon regulation has

revealed that small RNA regulatory pathways play a role in controlling transposon movement. In *D. melanogaster* the PIWI-interacting small RNAs (piRNAs) appear to function primarily in the regulation of transposons in both the germ-line and somatic tissues (Chung et al, 2008; Khurana & Theurkauf, 2010; Lau et al, 2009; Malone et al, 2009; Robine et al, 2009). piRNA *in D. melanogaster* also map to a great deal of other sequences in the genome however the importance of these associations has yet to be fully understood. The transposon load of the *D. melanogaster* genome is 15.8% which is relatively small when compared to that of *Ae. aegypti* which has transposon load of 46% (Kaminker et al. 2002; Nene et al 2007). This difference in genome composition may result in the *Ae. aegypti* piRNA pathway functioning in a different manner (Kaminker et al, 2002; Nene et al, 2007; Smith et al, 2007).

piRNAs are 24-30nt small RNAs that are believed to be involved in transposon control. PiRNA transcripts are generated from PIWI proteins in a dicer-independent pathway. The *D. melanogaster* piRNA system is composed of three proteins in the PIWI subfamily: Piwi, Aubergine (AUB) and Argonaute3 (AGO3). AUB and AGO3 work together in the germ line functioning in a feedback loop known as the ping-pong model (Bernstein et al, 2001; Brennecke et al, 2007; Olivieri et al, 2010; Vagin et al, 2006). The PIWI protein is involved in primary biogenesis of piRNA and functions in somatic tissues. It is currently believed that piRNA generated by PIWI are responsible for priming the germ-line ping pong pathway.

AGO3 and PIWI are characterized by their PAZ and Piwi domains (Hammond et al, 2001). The structure of the PAZ domain has been well defined in *D. melanogaster* Ago1(Yan et al, 2003). Nuclear resonance has demonstrated that this Ago1 is composed of a left handed six-stranded β-barrel core (Yan et al, 2003). The PAZ domain is a conserved domain across species and has been demonstrated to interact with RNA (Yan et al, 2003)Mutations of Piwi in *D. melanogaster* demonstrated that the Piwi gene is required for the maintenance of cell propagation in the male and female germ-line. Mutations in another PIWI protein, AGO2 led to neurological defects (Kataoka et al, 2001). Thus PIWI family members have been shown to be involved in the control of development as well as RNAi.

Prior to our laboratories work little was known about how the piRNA system works in the human disease vector, *Ae. aegypti*. Bioinformatic studies indicate that the piRNA pathway of *Ae. aegypti* may var from that of *D. melanogaster*. Previous studies have shown that while an AGO3 homology can be found in *Ae. aegypti* there has been no AUB proteins found in Ae. *aegypti*. The most notable difference between *D. melanogaster* and *Ae. aegypti* is the gene expansion of the PIWI proteins. Seven PIWI proteins were described in *Ae. aegypti* compared to one described PIWI protein in D. melanogaster (Cambell et al, 2008). One possible explanation for this expansion is the increased transposon load of *Ae. aegypti* compared to *D. melanogaster*.

To the best of our knowledge there have been no functional studies of the individual *Ae. aegypti* PIWI proteins. However there is one stud that investigated the total piRNA populations in *Ae. aegypti* (Arensburger *et al*. 2011). In 2011 seven piRNA

libraries were created from *Ae, aegypti* Orlando and four lines of transgenic *Ae. aegypti* that contained the DNA transpsosons *Hermes*, *Mos1* or *piggyBac*. The study demonstrated that 19% of the sequenced piRNAs from whole tissue libraries mapped to transposons and a majority of these mapped to the antisense strand. In contrast a whole tissue *D. melanogaster* library in the same study reported that 50% of the sequenced piRNAs mapped to transposons. There are an estimated $2x10^7$ piRNA in *Ae. aegypti* which is comparable to *D. melanogaster*. piRNA clusters were found to cover 20.6% of the genome and were shown to potentially generate up to 84% of the observed piRNAs. These loci were not as transposon rich and were less dense than those clusters found in *D. melanogaster* (Arensburger et al, 2011).

Transposon control is not the only function that has been attributed to the piRNA pathway. PiRNA appear to play a role in controlling deadenylation of early maternally deposited mRNA that control development in early embryogenesis. Nanos (nos) is a morphogen that is expressed in a gradient from the posterior to anterior pole of *D. melanogaster* embryos. The expression of this transcript controls abdominal segmentation (Gavis & Lehmann, 1994). It was demonstrated that an interaction between piRNAs and nos mRNA is required for nos mRNA deadenylation and translational repression. The knock down of PIWI proteins resulted in stabilized nos mRNA thus leading to head development defects. The study proposed that in complex with Piwi-type Argonaute proteins Aub and Ago3, target nos maternal mRNAs and recruit or stabilize the CCR4-NOT deadenylation complex together with Smaug (Rouget et al, 2010). Thus

piRNA are involved in developmental processes that control embryo development and may have many other functions that have yet to be discovered.

Aedes aegypti is a medically significant vector of human diseases and has been the object of many genetic studies. Researchers have been able to use transposon mediated genetic transformation, RNAi and site specific recombinases as tools to manipulate *Ae. aegypti* (Attardo et al, 2003; Clemons et al, 2010; Nimmo et al, 2006). *PiggyBac, Hermes* and *Mos1* have all been used to transform *Ae. aegypti* yet their inability to be remobilized has hampered the development of such genetic strategies as gene tagging and gene and enhancer trapping. There are only a few laboratories in the world that are able to transform mosquitoes with any sort of efficiency (O'Brochta et al, 2003; Sethuraman et al, 2007; Smith & Atkinson, 2010; Wilson et al, 2003) One possible explanation for this phenomenon is that *Aedes aegypti*'s unusually high transposable element load (47% of the 1.38GB genome composed of transposons) results in a primed piRNA pathway that is able to quickly silence any transposon activity.

This study investigated the *Aedes* piRNA pathway by determining which of the putative PIWI proteins are expressed and in what tissues. We have also performed immunoprecipitation in order to determine where piRNAs that are associated with the different proteins map to in the genome. RNAi knockdown of the PIWI proteins was performed to determine the downstream affects on the piRNA population.

4.3 Methods and Materials

4.3.1 Mosquito Stocks

Aedes aegypti Liverpool strain was used throughout these experiments. This strain is maintained at the University of California, Riverside under standard insectary conditions as described in (Munstermann, 1997) with the one exception that larvae were fed on a diet of Milkbone Original Dog Biscuits mixed with Red Star Specialty Nutrex 55. Adults were fed on a 10% sucrose solution and blood fed on mice. The colony was kept at 26° with an L:D of 14:10

4.3.2 RT-PCR

Reverse Transcription PCR (RT-PCR) was performed using Invitrogen's SuperScript III Reverse Transcriptase. The primers used for RT-PCR and qPCR are listed in Table 1.

4.3.3 Quantitative PCR

The carcass samples contained equal males and females with their genitalia and heads removed. The heads are removed since the eyes of mosquitoes contain a pigment that may disrupt polymerase activity in subsequent reactions. The ovaries were dissected 48 hours post blood feeding and the eggs were collected less than three hours post-oviposition. RNA was isolated using TRIzol® and cDNA was synthesized using

Superscript® II from Invitrogen. Real time quantitative PCR was performed on a BioRad MyiQ™ thermal cycler using iQ™ SYBR® Green Supermix.

4.3.4 Small RNA Library

Ovaries were harvested from *Ae aegypti* females 48 hours post blood feeding. RNA was extracted using TRIzol®. The libraries were then sequenced using the Illumina Small RNA kit v1.5 revision C. The runs were then sequenced on an illumina Genome Analyzer II DNA Sequencer at the Univeristy of California, Riverside's Genome Core.

RNAseq figures were designed using the Vectorbase website. Once at the site I searched for the gene of interest. Then pushed the following tabs in order: configure this page, custom data and attach remote file. Next I entered the IP address of the server of interest followed by the location of the file on the computer, the library name and the designate the read out format For example: http://xxxxxxxxxxxx/ucr.edu/Tophat-files/RNAlib19_aedes.bam. Lastly choose the format BAM. This command line maps the transcripts in the library to the gene of interest. In order to predict novel gene transcript and test how the transcripts map the annotated gene return to configure this page, custom data and attach remote file. Then enter the following link: http://xxxxxxxxxxx/ucr.edu/cufflinks-files/RNAliv-novel.gtg. Next select the format GFF and press enter. The two figures will be merged as one.

4.3.5 Immunoprecipitation of Small RNA

Ovaries were dissected as described above. 100 ovaries were collected and antibodies were added in a ratio of 1:10 for AGO3, PIWI2, and EGFP while PIWI7 antibody was used in a 1:1 ratio due to the small amount present in the ovaries compared to the other proteins. Affinity purified antibodies were raised against the following peptides: anti-Ago3-GQSVKRNPDEALNDKLFYL, anti-Piwi2-RPTFQHPGAEGRAMTHRDASAGRGAS and anti-Piwi7-EYRPRGGRGGNNQARGNVGGEG by Open Biosystems. The affinity purified anti-EGFP was purchased from Invitrogen. Dissected ovaries were homogenized in 200 ul volume of Buffer A (described below). The lysate was cleared via centrifugation for 40 minutes at 3000g at 4°C. The supernatant was filtered through a 0.45um filter Spin X column. $1/10^{th}$ volume primary of antibody was added and incubated overnight at 4°C on a rocker. 50 ul of Protein G Mag beads were placed into an RNase-free tube and pulled down with a magnet. The beads were washed once with 200 ul of Buffer A. The beads were then pulled down with a magnet to remove all of the supernatant. Next the homogenate/antibody solution was added to the beads and incubated for 1 hour at 4°C on a rocker. Once this was completed the beads were washed for 20 minutes in lysis Buffer A with rocking and then repeated. Following the second wash the beads were washed for 20 minutes in Buffer B (described below). Then the beads were once again pulled down with the magnet. 50 ul of Buffer B was added after the final wash. To re-suspend, 500 ul of Trizol® was added to the buffer-bead suspension. RNA was extracted from the proteins using Trizol®. The small RNAs were then sequenced using the Illumina

TruSeq™ RNA kit Revision A on the Illumina Genome Analyzer II DNA Sequencer at the University of California, Riverside's Genome Core.

Buffer A: 20 mM Tris pH7.5, 150 mM NaCl, 1.5 mM MgCl2, 0.5% Triton-X 100, 0.5% NP-40. To make complete Buffer A add fresh to 10 ml of A:10 ul 1M DTT to give final 1 mM DTT. Add 1 tablet Roche EDTA-free protease inhibitor mix .

In a new tube Add 2.5 ul of RNAaseOUT or RNasin to 200 ul of the above mix.

Buffer B: 20 mM Tris pH7.5, 300 mM NaCl. 1.5 mM MgCl2, 0.5% Triton-X 100, 0.5% NP-40. To 1 ml of B add fresh: of 1M DTT to give 1 mM DTT final concentration.

4.3.6 Mass Spectrometry

Ovaries were dissected as described in the small RNA library section. Immunoprecipitation was then performed as described above with the exception that the beads were not treated with Trizol® after the final wash. Instead the proteins were left bound to the magnetic beads and submitted for mass spectometery University of California, Riverside's Proteomics Core. The proteins were treated with chymotrypsin and then analyzed using a Waters Q-TOF nano-ESI/MS/MS with 2D-nanoAcquity UPLC. Matrix Science's Mascot Daemon software was used to analyze the results.

4.3.7 RNAseq libraries

Ovaries were dissected as above and eggs were collected less than three hours after oviposition. We used the illumina TruSeq™ RNA kit Revision A. All libraries were

sequenced at the UCR Institute for Integrative Genome Biology on an Illuminia GAx2 sequencing machine. Prior to analysis, ribosomal degradation products were identified and removed by matching each small sequence to a database of known ribosomal sequences from the *Aedes* genus using the BLAT alignment program (Kent 2002). 3' adapter sequences were identified on the remaining sequences using custom written PERL and R scripts. Adapter sequences were removed bioinformatically. The remaining library was mapped on to the *Ae. aegypti* assemby (AaegL1) hosted by VectorBase (http://aaegypti.vectorbase.org/) using the program TopHat with default parameters (C. Trapnell, Pachter, and Salzberg 2009). Transcripts were assembled using the Cufflinks program and expression levels (using RPKMs) for both known transcripts (AaegL1.2 gene set from VectorBase) and novel transcripts were determined (Cole Trapnell et al. 2010).

4.3.8 RNAi

Newly emerged females were injected with dsRNA and allowed to recover for 24 hours. DsRNA was synthesized using the MEGAscript® RNAi kit by Ambion with the primers listed below. The females were then blood fed and dissected 72 hours later. Ovaries were dissected as described above. The RNA was extracted with TRIzol® and then RNA seq libraries were made using the Illumina TruSeq™ RNA kit Revision A.

4.3.9 RNAi Survivorship

Unblood fed young mosquitoes were injected with dsRNA and allowed to recover for 24 hours and then blood fed on mice. At 72 hours post blood feeding, ovaries were dissected and mounted onto glass slides to monitor development and determine whether or not the dsRNA has an affect on ovary development. Also at 72 hours post injection 20 of the injected and blood fed females were separated into oviposition cups with three young males. The females were provided with 10% sucrose and cup of water for oviposition and drinking purposes. The cups were monitored once every day to determine when eggs were laid, eggs hatched, pupae formed and adults emerged. A cup was marked as having achieved each stage on the day that at least one individual in the cup reached the next benchmark. This means that the first day that an egg or larvae was seen in the ovipoistion cup the females was scored as having met the benchmark. The number of eggs laid by each female was estimated and placed into one of four categories >10, 10-50, 50-100 or, <100. There were no significant differences found in number of eggs between any of the treatments.

4.4 Results

4.4.1 PIWI Transcription

An initial reverse transcription PCR (RT-PCR) was performed to determine if and where each of the PIWI proteins are transcribed. RT-PCR indicated that AGO3, PIWI3, PIWI4, PIWI5 and PIWI6 are expressed in both the somatic and germ-line tissues. PIWI2 and PIWI7 appeared to be germ line specific with expression only being seen in the

ovaries and eggs (Figure 4.1 and Table 4.2). We decided to pursue the germ-line specific proteins and AGO3 for further study. Relative transcription levels for three PIWI proteins: AGO3, PIWI2 and PIWI7 were determined using quantitative real time PCR (qPCR). The AGO3 results showed that AGO3 is expressed at similar levels in the carcass, egg and ovaries (Figure 4.2). The expression levels of AGO3 in the ovaries and eggs were verified by the mRNA-seq data. This is markedly different from *D. melanogaster* in which AGO3 is only transcribed in the germ-line. PIWI2 transcription was verified to be germ line specific with no transcripts being detected in the carcass samples (Figure 4.3). mRNA-seq data showed that PIWI2 is expressed in both eggs and ovaries in similar levels. Additionally through mRNA-seq were able to detect a novel PIWI2 isoform in the ovaries that is not described in Vectorbase. PIWI7 had originally appeared to be germ-line specific in the initial RT-PCR experiment; however qPCR demonstrated that PIWI7 transcription occurs in both germ-line and somatic tissues (Figure 4.4). The mRNA-seq data revealed that there is a very low transcription of PIWI7 in the ovaries and that transcription is greatly increased in the eggs. No isoforms of PIWI7 were discovered. We do not currently have mRNA-seq data for expression in the carcass but that information could provide valuable insight into what is occurring in the carcass.

4.4.2. Immunoprecipitation Small RNA Libraries

Small RNA libraries were made from immunoprecipitation pull downs (IP pull downs) using antibodies to AGO3, PIWI2, PIWI7 and EGFP in order to determine which piRNA transcripts are associated with each of the individual proteins and where those piRNAs map to in the genome. Over 100 million sequences were generated from each IP small RNA library. The sequences that did not map to the annotated *Ae. aegypti* genome were removed from the data set and then the remaining sequences were analyzed to determine what proportion of the pulled down sequences mapped to transposable elements (Figure 4.5). PIWI2 had the largest proportion of transcripts that matched transposable elements with 27% mapping to retrotransposons, 1.6% to DNA transposons, 0.16% MITES and 0.29% to other for a total of 29%. The remaining portions map to non-transposon regions of the genome. The PIWI7 pull down data set contained half as many sequences mapping to transposable elements with 11.3% mapping to retrotranposons, 0.51% to DNA transposons, 0.11% to MITES and 0.25% mapped to other types for a total of 12.11% of the total data base. The AGO3 pulls suggested had similar percentages of piRNA mapping to transposons as PIWI7 with 13.39% mapping to retrotransposons, 0.75% to DNA transposons, 0.18% to MITES and 0.27% mapping to other.

4.4.3 Mass Spectrometry

Mass spectrometry (MS) was performed using immunoprecipitations of AGO3, PIWI2, PIWI7 and EGFP (Tables 4.3, 4.4, and 4.5). MS allowed us to ensure that we were indeed pulling down the proteins of interest. EGFP was used as a control in order to

determine background proteins that were binding non-specifically to either the beads or the antibodies. Vitellogenin, vitellogenin precursor, myosin and AAEL006138 were found in each of the pull downs suggesting these proteins are abundant in the ovaries. The AGO3 pull down demonstrated that we were in fact pulling down AGO3, and while other proteins were detected by the mass spectrometer they were not specific to the AGO3 pull down. The PIWI2 pull down revealed that PIWI3 was also pulled down suggesting that either the two possibly form a complex or that they cross react to the antibody. PIWI3 is found in both germ line and somatic tissues and thus it is reasonable for the two proteins to be interacting. PIWI7 is found in very low concentrations in the ovaries and so we were unable to detect the protein in the pull down and the other proteins detected did not vary significantly from those found in the EGFP pull down. However, PIWI7 is found in much greater quantities in the eggs so this may be a better tissue to investigate in order to determine which proteins interact with PIWI7.

4.4.4 Knock Down Experiments

Double stranded RNA (dsRNA) was synthesized using cDNA from the Liverpool strain of *Ae. aegypti* and targeted to AGO3, PIWI2, PIWI7 with Luciferase used as a control. Surprisingly we were unable to achieve knock down of any of the proteins in the ovaries. After examining the mRNA-seq libraries we found that transcript levels of each of the targeted proteins was similar to that of our wild-type library. Analysis of our control library in which we injected dsRNA to luciferase revealed that no sequences

mapping to LUC had been amplified. Typically in a knock down experiment it is expected that the library will contain a large number sequences that match the dsRNA that was injected. This suggests that no dsRNA entered the ovary. Furthermore we verified the lack of knock down via qPCR (Figure 4.6.)

This result was unexpected because the several days after blood feeding is the period in which the follicle cells surrounding the developing eggs open and begin to absorb a great deal of proteins and lipids from the hemolymph thus we surmised this would be the most vulnerable period to dsRNA. Previous studies have demonstrated the injection of dsRNA into adult female mosquitoes can result in efficient knock down of proteins in the fat body however we have not seen any studies that report on the ability of dsRNA to enter the ovaries (Attardo et al, 2003). Insects do encounter dsRNA viruses and it makes sense that the insect immune system would wish to protect the ovaries from infection or exogenous RNA.

The knock down survivorship study may provide some insight into this problem. As a control luciferase dsRNA was injected into 20 females resulting in 85% of the injected females producing F1 adults. When AGO3 and PIWI7 dsRNA was injected in females the number of females that produced adult F1 individuals was greatly decreased when compared to the luciferase control. Of the 20 females injected with dsRNA targeted to AGO3 and PIWI7, 45% and 30% respectively, resulted in adult FI progeny. However, those females injected with dsRNA to PIWI2 showed no difference from the control with 80% percent of crosses resulting in F1 adults. This suggests that AGO3 and PIWI7 may have some sort of developmental function in the somatic tissue and that PIWI2 was

protected in the ovaries from the dsRNA thus we saw no affects on survivorship (Table 4.6).

4.4.5 RNAseq Libraries

The RNAseq libraries were made from eggs 0-3 hours old and ovaries that were dissected in the manner described above. These libraries provide valuable data regarding the expression levels and transcripts of each of the PIWI genes (Table 4.7) . PIWI 1 and 3 are almost identical however they are located on separate supercontigs and there are some small nucleotide polymorphisms (snps) that exist between the two genes. However the sequences are close enough that anti-bodies nor primers could be designed to tell the two apart. Expression levels can be determine for each gene by normalizing the data sets between the two libraries, Egg (19) and Ovary (20). The expression level of PIWI1 was 13.07 in Eggs with a confidence range of (0-20.31 LOWDATA) and 34.54 in Ovaries with a confidence range of (0-46.29 LOWDATA). The low data suggests there were not enough transcripts present to determine the expression level with any confidence. PIWI3 has an expression level of 1.71013 in eggs with confidence range of (1.49- 1.92) and 2.93 in ovaries with a confidence range of (2.60-3.27) (Figure 4.7).

PIWI2 has a second unannotated found on another supercontig. The unannotated gene has a sequence similarity of 98.4% and produces an identical protein. It is possible that this second copy is an artifact from misassembly and the sequence differences are from heterlogy found in the *Ae. aegypti* genome. PIWI2 has an expression level of 1.71 with a confidence range of (1.49-1.92) in eggs and 2.93 in ovaries with a confidence range of (2.599-3.27) (Figure 4.6).

PIWI4 transcripts map well to the annotated gene. There is a novel transcript prediction that predicts an additional exon 10kb upstream of the annotated PIWI4 however this is likely not reality since there only a few transcripts mapping to region suggesting it is an artifact. PIWI 4 has an expression level of 44.48 in eggs with a confidence range of (43.31-45.65) and an expression level of 34.16 in ovaries with a confidence range of (33.16-35.15) (Figure 4.8). This high level of expression suggests it could be playing an important role in piRNA pathway.

PIWI5 may a have longer fifth exon than is currently annotated in the *Ae. aegypti* genome. Especially in the ovary library there are a great deal of transcripts that map to the 3' end of the gene beyond predicted transcript. PIWI 5 has an expression level of 0.74 in eggs with a confidence range of (0.57-.90) and an expression level of 1.60 in ovaries with a confidence range of (1.33-1.86) (Figure 4.9). PIWI6 transcripts mapped very closely to the predicted gene sequence. PIWI6 had a high expression levels in eggs of 28.45 with a confidence range of (27.46-29.43) and no transcripts were found in Ovaries. However, when we analyzed the knock down RNA seq libraries we did find expression in the ovaries. In the two ovary knock down libraries we analyzed for PIWI6 expression we found expression levels of 9.88 with a range of (9.31-10.46) in one library and 19.69 with a range of (9.31-10.46) in the other (Figure 4.10). This may mean that PIWI6 is only expressed under stressed physiological conditions or it may be that we just didn't pick up PIWI6 transcripts in the ovary library (20) even though they were indeed present.

PIWI7 transcripts mapped well to the predicted gene sequence in Vectorbase. PIWI7 has an expression level of 7.428 in eggs with a confidence range of (7.00-7.85).

The expression is significantly lower in eggs with an expression level of 0.04 and a confidence level of (0.01-0.08) (Figure 4.11). This is also supported by the qPCR which suggests that expression is significantly greater in eggs than ovaries. And lastly, AGO3 maps to the Vectorbase prediction with each of the introns highly supported and no novel transcripts suggested. AGO3 with expressed at high levels in both the egg and ovary libraries. In eggs AG03 has an expression level of 24.76 with a confidence range of (24.76- 25.43). In ovaries the expression level is 26.123 with a confidence range of (0-36.45 LOWDATA) (4.12). However when we look at the ovary knock-down library we get an expression of 23.273 with a confidence range of (22.62-23.92) supporting the expression level found in the ovary library (20).

4.5 Discussion

4.5.1 The first look at the piRNA pathway in *Ae. aegypti.*

Ae. aegypti contains one AGO protein and six PIWI proteins. Of these proteins AGO3, PIWI3, PIWI4, PIWI5, PIWI6 and possibly PIWI7 are transcribed in both germ-line and somatic tissues. PIWI 2 is the only PIWI protein that is restricted to the germ-line. Expression levels were determined for each protein using RT-PCR. qPCR was used to determine the relative expression rates of AGO3, PIWI2 and PIWI7 in three separate tissues: carcass, ovary and egg. There was no significant difference in transcription levels of AGO3 and PIWI7 between the three tissues while PIWI2 was found to only be expressed in egg and ovary with similar expression found in both tissues.

The IP pull downs showed that PIWI2 has the highest proportion of piRNAs mapping to transposable elements with nearly a third of the data set mapping to TEs. AGO3 and PIWI7 had similar proportions with ~13% of the data sets mapping to TEs. As stated earlier, previous studies have shown that 19% of the total piRNA population in *Ae. aegypti* so the 13% that map to AGO3 and PIWI7 suggest they may not play a critical role in transposon regulation.

The knock down experiments were not successful in the way we had originally intended however valuable information was still discovered about the affects of the knock down of these proteins on survivorship of F1 progeny and may provide insight into a dsRNA defense system in the ovaries. None of the three proteins AGO3, PIWI2 or PIWI7 were able to produce a knock down affect in ovaries as was verified via mRNA-seq and qPCR. This was surprising since previous studies had shown that knock down via RNAi is achievable in *Aedes aegypti*. These studies were performed on the fat bodies which come in constant contact with the hemolymph and may have less reason to defend themselves against dsRNA than the germ line. We could not find any studies where RNAi was attempted in ovaries. Other studies that analyzed the lack of a specific PIWI protein in the ovaries utilized *Drosophila* mutants that were deficient in the protein and thus did not need to deal with the presence of dsRNA in the system. The survivorship study suggests that while there was no detectable knock down in the ovaries the presence of dsRNA that targets somatic PIWI proteins resulted in decreased survivorship of F1 progeny.

It seems that the most likely candidate PIWI protein to be involved with transposon regulation that we have studied thus far is PIWI2. We would expect that if a protein was involved in transposon regulation it would have a great deal of activity in the ovaries where developing embryos are undergoing meiosis and are most vulnerable to transposon movement. PIWI2 is germ line specific and was associated with double the proportion of piRNAs mapping to TEs when compared to PIWI7 and AGO3. The survivorship study suggests that PIWI7 and AGO3 have a somatic function that ultimately affects F1 survivorship however this function remains unclear.

Rouget et al. (2010) recently demonstrated that piRNAs are important to deadenylation and that knocking down components of the piRNA pathway resulted in mutations during embryogenesis. This phenomenon may provide insight as to how the knock down of PIWI 7 and AGO3 resulted in a decrease in FI progeny when the knock down experiments clearly show that dsRNA is not entering the ovary. It may be that the piRNA generated from AGO3 and PIWI 7 are involved in controlling genes important in oogenesis. The fat body is responsible for the production of vitellogenin and other proteins that vital to embryo development. It has already been demonstrated that RNAi machinery is capable of affecting the fat body and so it may be that knocking down AGO3 and PIWI7 affects the fat bodies production of proteins or transcripts during Vitellogenesis. This study cannot address what or how these affects are occurring but may point to other possible roles of the piRNA pathway in *Ae. aegypti*. The next experiment to perform would be inject newly emerged females with dsRNA targeted to the proteins in question and instead of creating RNAseq libraries from the germ-line

ovary tissue dissect the fat bodies and determine if there are altered transcription levels of the PIWI proteins in somatic tissue.

The RNAseq libraries enabled us to determine relative transcription levels for each of the PIWI proteins. PIWI1, PIWI2, PIWI3, and PIWI5 were each expressed in the eggs and ovaries at relatively low levels. PIWI4 had high transcription levels in both the eggs ovaries suggesting that it may play an important role at both developmental stages. PIWI 6 had high expression in the eggs but there was insufficient data to determine the transcription levels in ovaries. However when we analyzed the knock-down RNAseq libraries we found medium to high transcription levels. This may be indicative that PIWI6 is involved in dsRNA defense or some sort of immune response but it may also just be an anomaly in our ovary data set. PIWI7 had very low levels of transcription in the eggs but was highly transcribed in the ovaries. Future immunoprecipitation and knock-down studies on PIWI7 would benefit from focusing on eggs.

Overall this study has provided valuable information on the expression of the proteins involved in the piRNA pathway. It has also investigated which piRNAs are associated with AGO3, PIWI2 and PIWI7. Future studies should look into immunoprecipitation RNAseq libraries for each of the PIWI proteins in order to better understand which piRNAs are associated with the individual proteins. Furthermore RNAi Knock-down experiments should be performed targeted to each of the PIWI proteins and the effects investigated in somatic tissue, ovaries and eggs.

4.5.2 Future Directions

This study has provided a valuable platform to begin further studies into the *Ae. aegypti* piRNA pathway. First I would make an RNAseq library using carcass only since thus far we have only made them for egg and ovaries. Next it would be beneficial to make antibodies to each of the PIWI proteins so that co-immunopreciptation studies can be done for each of the PIWI in order to determine which piRNA are associated with each protein. Repeating the pull down of PIWI7 using eggs 0-3 hours old could provide valuable information since PIWI7 is most highly expressed in eggs and is barely expressed in the ovaries.

I would also repeat the RNAi knock down studies only this time making RNAseq libraries from the fat body tissues. This would allow for us determine if knock down of the PIWI proteins is occurring in somatic tissues and possibly explain the decreased survivorship of those females injected with dsRNA targeting PIWI proteins expressed in somatic tissue. And lastly I would perform in-situ hypbridization in the ovaries and views in order to better understand the expression patterns of each of the PIWI proteins over the development of the embryo. These experiments would provide greater insight in this newly investigated pathway.

4.6 References:

Arensburger PA, Hice RH, Wright JA, Craig NL, Atkinson P (2011) The mosquito Aedes aegypti has a large genome size and high transposable element load but contains a low proportion of transposon-specific piRNAs. *BMC Genomics* **Submitted:** submitted

Attardo GM, Higgs S, Klingler KA, Vanlandingham DL, Raikhel AS (2003) RNA interference-mediated knockdown of a GATA factor reveals a link to anautogeny in the mosquito Aedes aegypti. *PNAS* **100:** 13374-13379

Bernstein E, Caudy AA, Hammond SM, Hannon GJ (2001) Role for a bidentate ribonuclease in the initiation step of RNA interferance. *Nature* **406:** 363-366

Brennecke J, Aravin AA, Stark A, Dus M, Kellie M, Sachidanandam R, Hannon GJ (2007) Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* **128:** 1089-1103

Chung W-J, Okamura K, Martin R, Lai EC (2008) Endogenous RNA interference provdes a somatic defense against *Drosophila* transposons. *Current Biology* **18:** 795-802

Clemons A, Haugen M, Severson D, Dunman-Scheel M (2010) Functional analysis of gene in Aedes aegypti embryos. *Cold Spring Harb Protoc*

Gavis ER, Lehmann R (1994) Translational regulation of nanos by RNA localization. *Nature* **369:** 315–318

Hammond SM, Boettcher S, Caudy AA, Kobayashi R, Hannon GJ (2001) Argonaute2, a Link Between Genetic and Biochemical Analyses of Rnai. *Science* **293:** 1146-1150

Kaminker JS, Bergman CM, Kronmiller B, Carlson J, Svirskas R, Patel S, Frise E, Wheeler DA, Lewis SE, Rubin GM, Ashburner M, Celniker SE (2002) The transposable elements of the Drosophila melanogaster euchromatin: a genomics perspective. *Genome Biology* **3:** 0084.0081-0084.0020

Kataoka Y, Takeichi M, Unemura T (2001) Developmental Roles and Molecular Characterization of a Drosophila homologue of Arabidopsis Argonaute1, the Founder of a novel gene superfamily. *Genes to Cells* **6:** 313-325

Khurana JS, Theurkauf W (2010) piRNAs, transposon silencing, and Drosophila germline development. *J Cell Biol* **191:** 905-913

Kokoza VA, Raikhel AS (2011) Targeted Gene Expression in the Transgenic *Aedes aegypti* using the Binary Gal4-UAS System. *Insect Biochem Mol Biol* **41:** 637-644

Lau NC, Robine N, Martin R, Chung W-J, Niki Y, Berezikov E, Lai EC (2009) Abundant primary piRNAs, endo-siRNAs, and microRNAs in a Drosophila ovary cell line. *Genome Research* **19:** 1776-1785

Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, Sachidanandam R, Hannnon GJ (2009) Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. *Cell* **137:** 522-535

Munstermann LE (1997) Care and maintenance of Ades aegypti colonies. The Molecular Biology of Disease Vectors. *JM Crampton, Beard, CB and Louis, C, ed London Chapman and Halls***:** 13-20

Nene V, Wortman JR, Lawson D, Haas B (2007) Genome sequence of *Aedes aegypti*, a major arbovirus vector. *Science* **316:** 1718-1723

Nimmo DD, Alphey L, Meridith JM, Eggelston P (2006) High efficiency site-specific engineering of the mosquito genome. *Insect Molecular Biology* **15:** 126-136

O'Brochta DA, Sethuramuran N, Wilson R, Hice RH, Pinkerton AC, Levesques CS, Bideshi DK, Jasinskiene N, Coates CJ, James AA, al. e (2003) Gene vector and transposable element behavior in mosquitoes. *Journal of Experimental Biology* **3:** 3823-3834

Olivieri D, Sykora MM, Sachidanandam R, Mechtler K, Brennecke J (2010) An in vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in Drosophila. *EMBO J* **29:** 3301-3317

Robine N, Lau NC, Balla S, Jin Z, Okamura K, Kuramochi-Miyagawa S, Blower MD, Lai EC (2009) A broadly conserved primary pathway generates 3'UTR-direceted primary piRNAs. *Current Biology* **19:** 2066-2076

Rouget C, Papin C, Boureux A, Meunier A, Franco B, Robine N, Lai EC, Pelisson A, Simonelig M (2010) Maternal mRNA deadenylation and decay by the piRNA pathway in the early Drosophila embryo. *Nature* **467:** 1128-1132

Sethuraman N, Fraser MJ, Eggleston P, O'Brochta DA (2007) Post-integration stability of piggyBac in Aedes aegypti. . *Insect Biochem Mol Biol* **37:** 941-951

Smith CD, Shu S, Mungall CJ, Karpen GH (2007) The release 5.1 annotation of *Drosophila melanogaster* heterochromatin. *Science* **316:** 1586-1591

Smith RC, Atkinson PA (2010) Mobility properties of the Hermes transposable element in transgenic lines of Aedes aegypti. *Genetica* **139:** 7-22

Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD (2006) A distinct small RNA pathway silences selfish genetic elements in the germ.line. *Science***:** 320–324

Wilson R, Orsetti J, Klocko AK, Aluvihare C, Peckham E, P.W. A, Lehane MJ, O'Brachta DA (2003) Post- integration behavior of a Mos1 Mariner gene vector in Aedes aegypti. . *Insect Biochem Mol Biol* **33:** 853-863

Yan KS, Yan S, Faroq A, Han A, Zeng L, Zhou MM (2003) Structure and conserved RNA binding of the PAZ domain. *Nature* **426:** 269-474

Primer	Sequence
Lucifase For	TGCTTTTACAGATGCACATATCGAG
Luciferase Rev	CGTATTTGTCAATCAGAGTGCTTTTG
Luciferase ForT7	TAATACGACTCACTATAGGGTACTTGCTTTTACAGATGCACATATCGAG
Luciferase Re T7	TAATACGACTCACTATAGGGTACTCGTATTTGTCAATCAGAGTGCTTTTG
Piwi2 For T7	TAATACGACTCACTATAGGGTACTATGTCCGACCGTCAATCGCA
Piwi2 Rev Kpn	AAAGGTACCTTTTGCAACGCCACGACGG
Piwi2 For KpnI	AAAGGTACCATGTCCGACCGTCAATCGCA
Piwi2 Rev T7	TAATACGACTCACTATAGGGTACTTTTTGCAACGCCACGACGG
Piwi 7 For	GTAATGTCGGAGGAGAAGGTAGC
Piwi 7 Rev	CATTGCCTTGCGATTAATCAG
Piwi 7 For T7	TAATACGACTCACTATAGGGAGAGTAATGTCGGAGGAGAAGGTAGC
Piwi 7 Rev T7	TAATACGACTCACTATAGGGAGACATTGCCTTGCGATTAATCAG
Ago3 For	TATTACAAACAGCAGTACAACATCGAC
Ago3Rev	CCCAGTTTGCAGTTCATCTG
Ago3 For T7	TAATACGACTCACTATAGGGAGATATTACAAACAGCAGTACAACATCGAC
Ago3 Rev T7	TAATACGACTCACTATAGGGAGACCCAGTTTGCAGTTCATCTG

Table 4.1. Primers used for rtPCR and qPCR.

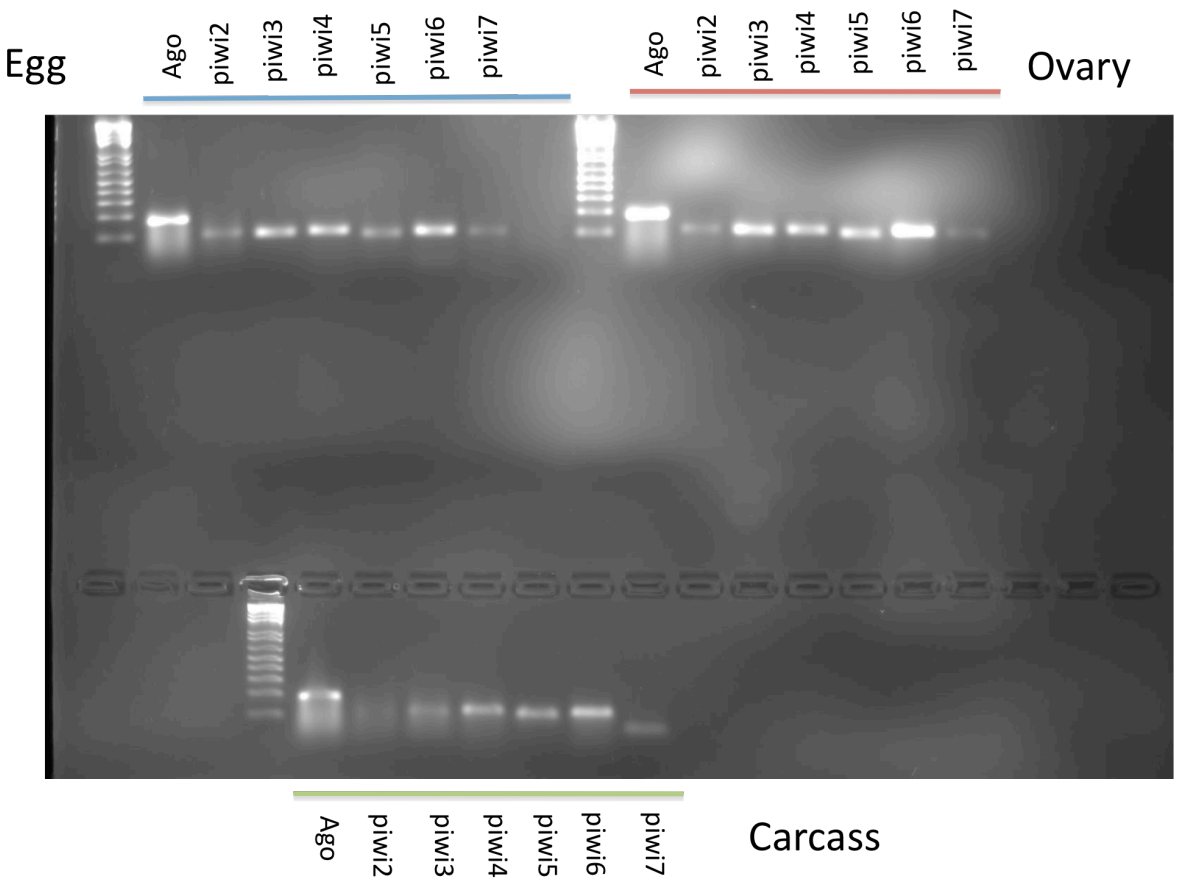

Figure 4.1. Reverse Transcription PCR Results for the PIWI genes. PCR was performed on cDNA from three separate tissues : carcass, ovary and eggs.

Protein	Carcass	Ovary	Egg
AGO3	+++	+++	+++
PIWI2	+ *	+++	+++
PIWI3	+++	+++	+++
PIWI4	+++	+++	+++
PIWI5	++	++	++
PIWI6	+++	+++	+++
PIWI7	+*	+++	+++

Table 4.2. Relative expressions of each PIWI protein as determined using rtPCTR . +++ is the highes level and + is the lowest expression. *Expression only seen at higher cycles.

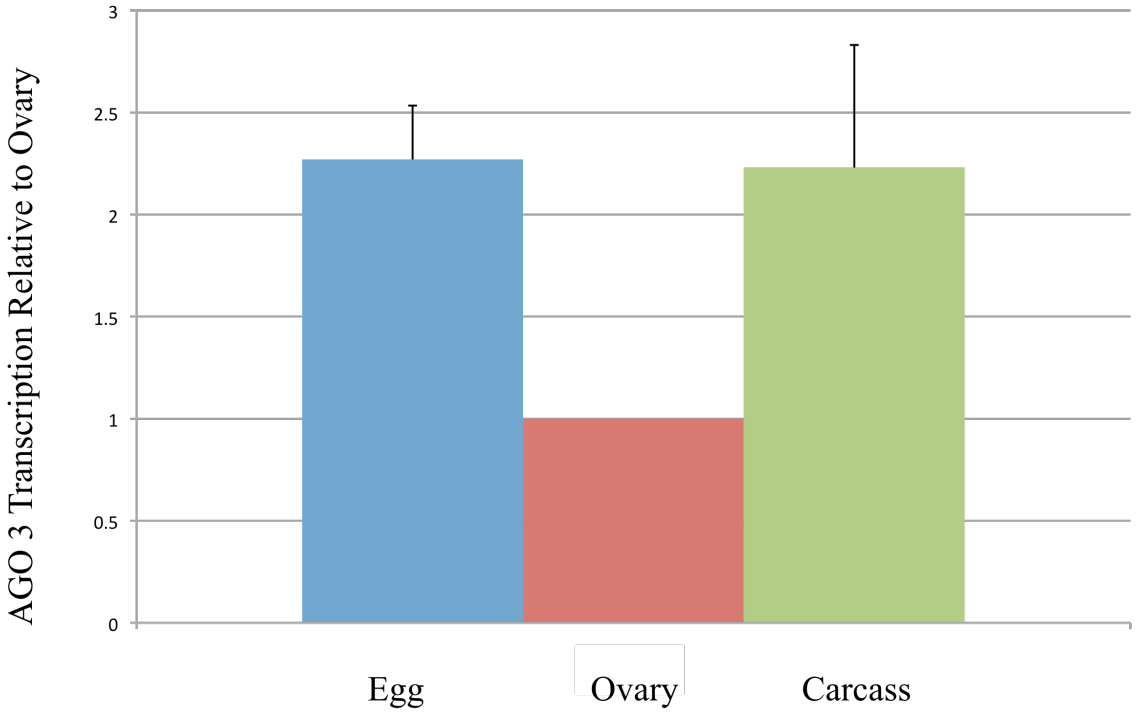

Figure 4.2. The relative transcription of AGO3 in egg, ovary and carcass in relation to ovary transcription. RSP7 was used as a reference gene. E=1.95

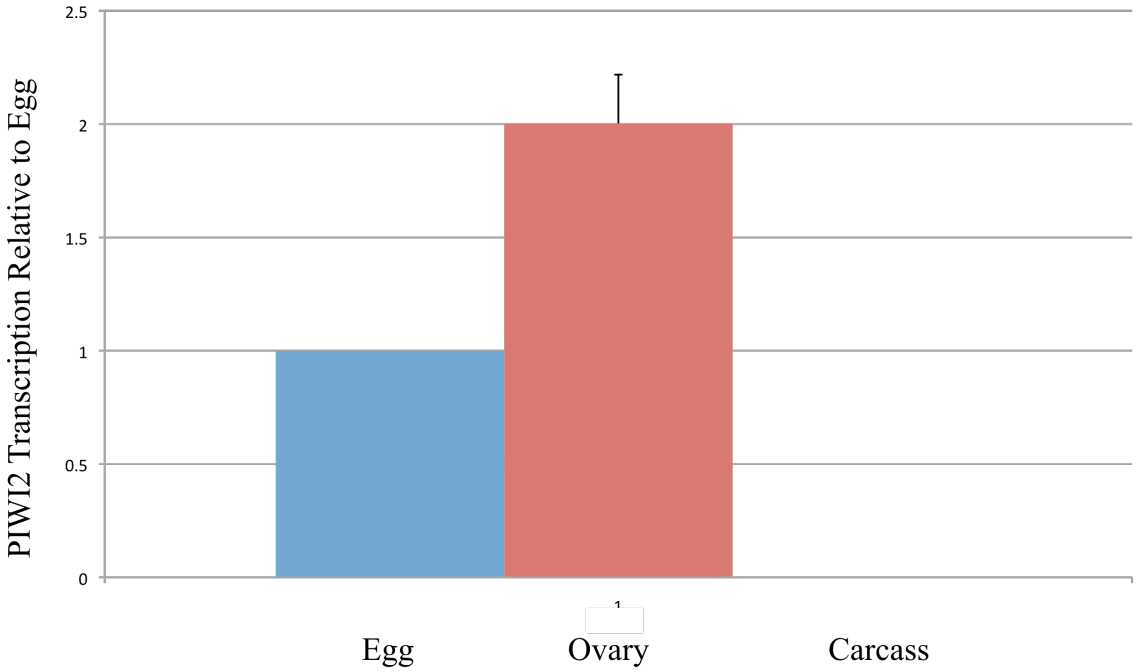

Figure 4.3. The relative transcription of PIWI2 in egg, ovary and carcass in relation to egg transcription. RSP7 was used as a reference gene. E= 2.01

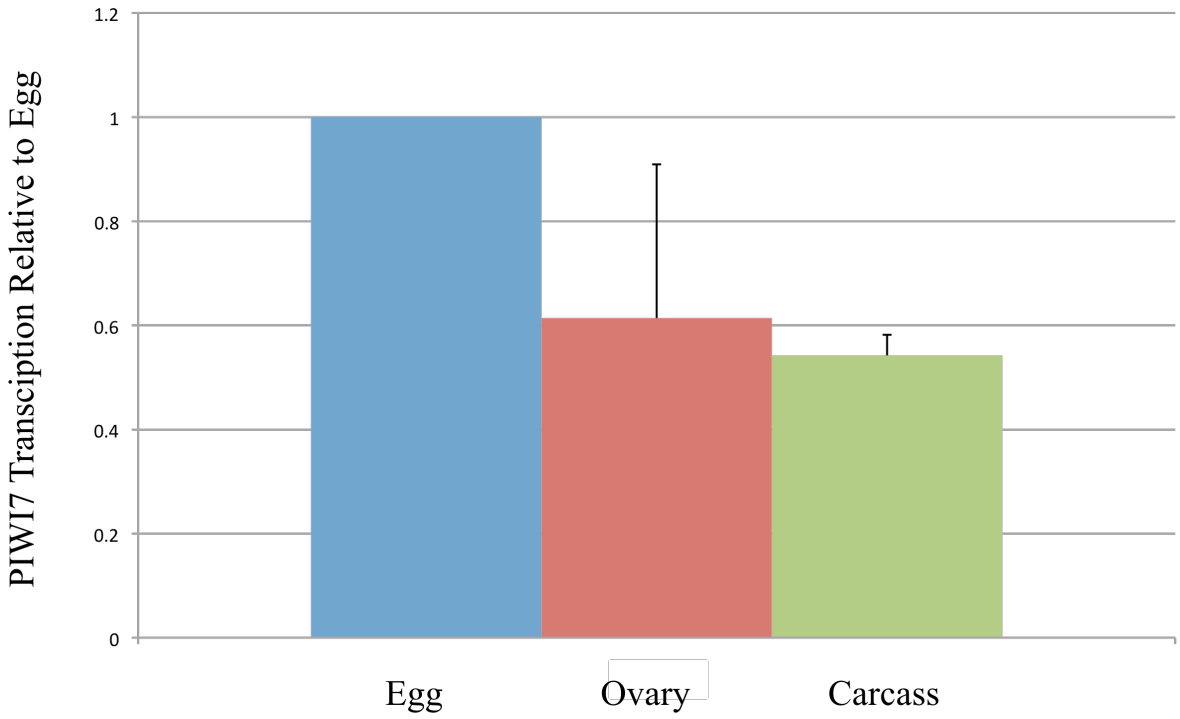

Figure 4.4. The relative transcription of PIWI7 in egg, ovary and carcass in relation to egg transcription. RSP7 was used as a reference gene. E= 2.02

Gene acession number	Gene Description
gi|157118639	myosin heavy chain, nonmuscle or smooth muscle
gi|157118641	myosin heavy chain, nonmuscle or smooth muscle
gi|157110721	myosin heavy chain, nonmuscle or smooth muscle
gi|472308	vitellogenin
gi|157126144	hypothetical protein AaeL_AAEL010434
gi|157112385	hypothetical protein AaeL_AAEL006138
gi|94468486	actin
gi|71383976	actin 6
gi|157116094	myosin light chain 1, putative
gi|157105157	myosin vi
gi|157125380	actin
gi|157131795	dihydrolipoamide succinyltransferase component of 2-oxoglutarate dehydrogenase
gi|157108966	spectrin
gi|157112672	beta chain spectrin
gi|157114119	2-oxoglutarate dehydrogenase
gi|157117037	AGO3
gi|157120145	myosin regulatory light chain 2 smooth muscle
gi|157131825	tropomyosin invertebrate
gi|157131827	tropomyosin invertebrate
gi|157167683	myosin regulatory light chain 2 (mlc-2)

Table 4.3. Above are the top twenty *Ae. aegypti* proteins found in the AGO3 immunoprecipiation.

Gene acession number	Gene Description
gi|157118639	myosin heavy chain, nonmuscle or smooth muscle
gi|472308	vitellogenin
gi|157117882	PIWI2
gi|157126144	hypothetical protein AaeL_AAEL010434
gi|157112385	hypothetical protein AaeL_AAEL006138
gi|157136936	PIWI3
gi|157116094	myosin light chain 1, putative
gi|157112377	hypothetical protein AaeL_AAEL006126
gi|157111095	myosin heavy chain, nonmuscle or smooth muscle

Table 4.4. Above are the nine *Ae. aegypti* proteins found in the PIWI 2 immunoprecipiation.

Gene acession number	Gene Description
gi|157118639	myosin heavy chain, nonmuscle or smooth muscle
gi|472308	vitellogenin
gi|157126144	hypothetical protein AaeL_AAEL010434
gi|37528873	vitellogenin-B
gi|157112377	hypothetical protein AaeL_AAEL006126
gi|157116094	myosin light chain 1, putative

Table 4.5. Above are the six *Ae. aegypti* proteins found in the PIWI 7 immunoprecipiation.

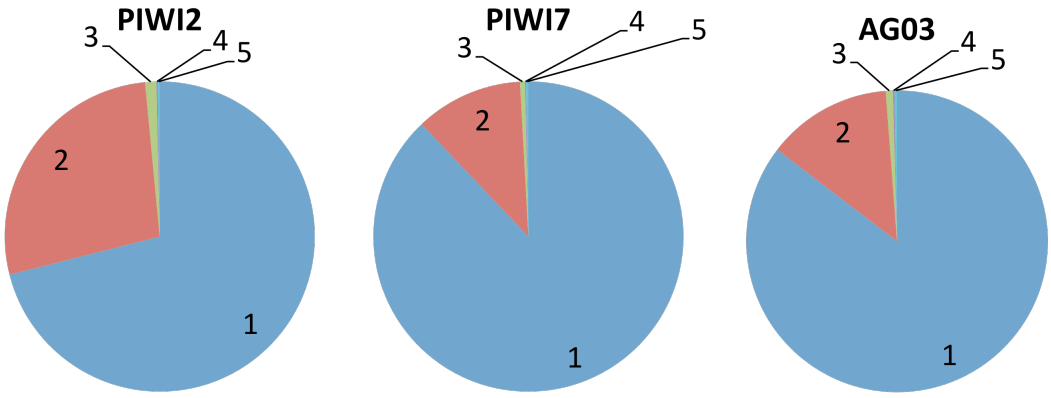

Figure 4. 5. 1.) Percentage of piRNA mapping to non-transposon regions of the genome. 2.) Percentage of piRNA mapping to retrotransposons. 3.) Percentage of piRNA mapping to DNA transposons. 4.) Percentage of piRNA mapping to MITES. 5.) Percentage of piRNA mapping to other transposons.

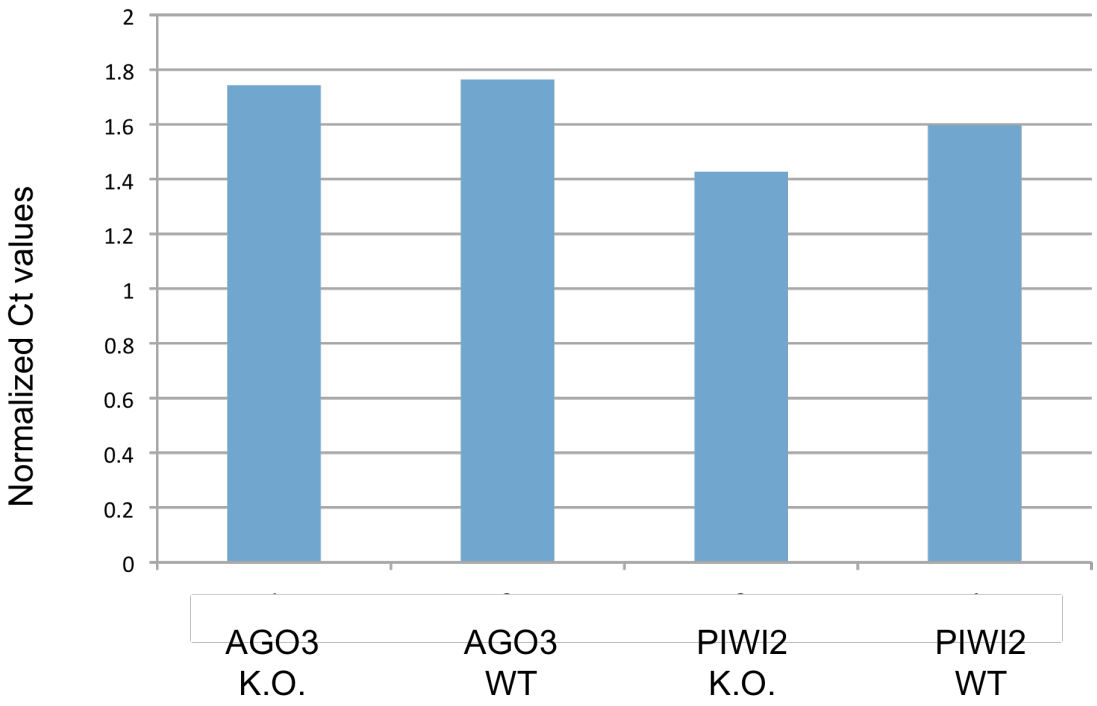

Figure 4.6. qPCR was performed on RNA collected from dissected ovaries of females injected with dsRNA targeted to AGO3 and PIWI2 as well as WT ovaries. The Ct values for each sample were normalized against RSP17.

Protein	Egg	Larvae	Adult
Luciferase	100%	85%	85%
Ago 3	60%	45%	45%
Piwi 2	80%	80%	80%
Piwi 7	70%	30%	30%

Table 4.6. This table shows the percent survivorship of the offspring of 20 females injected with dsRNA targeted to the indicated proteins. Luciferase was used as a control.

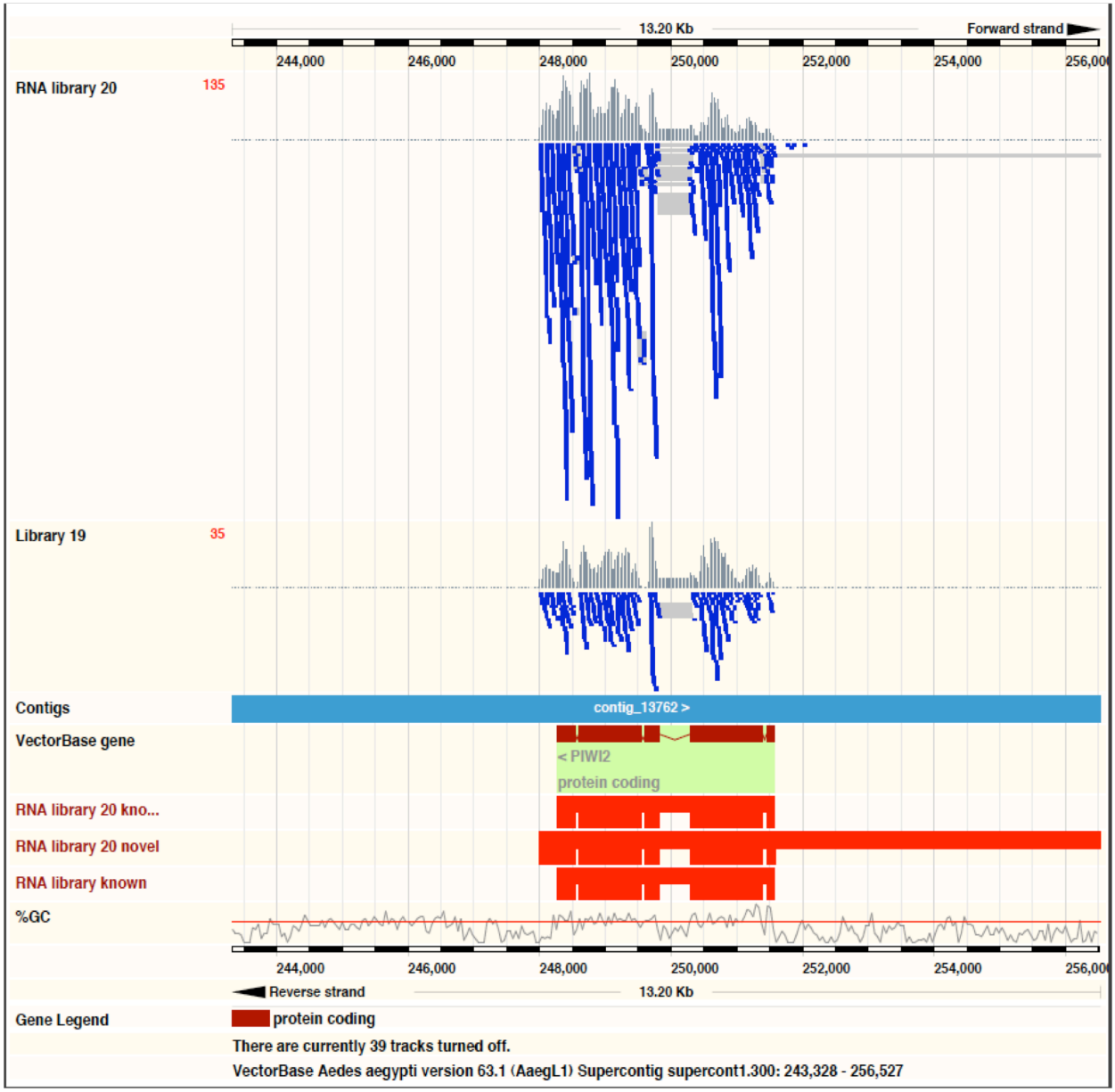

Figure. 4.7. RNAseq data for Piwi2 using libraries 19 (Egg) and 20 (ovary). The red bars are transcript predictions using the program cufflinks.

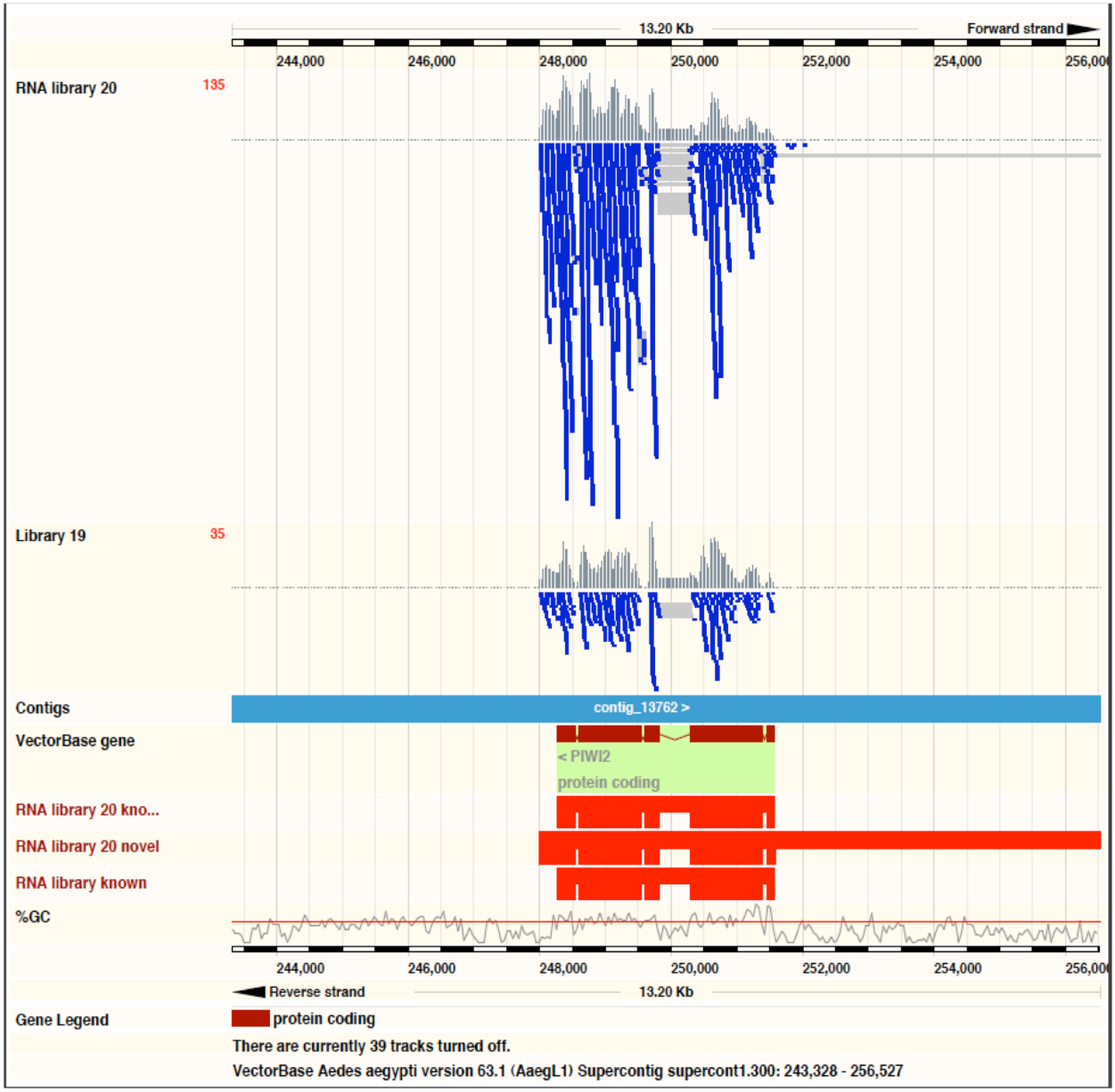

Figure. 4.8. RNAseq data for Piwi3 using libraries 19 (Egg) and 20 (ovary). The red bars are transcript predictions using the program cufflinks.

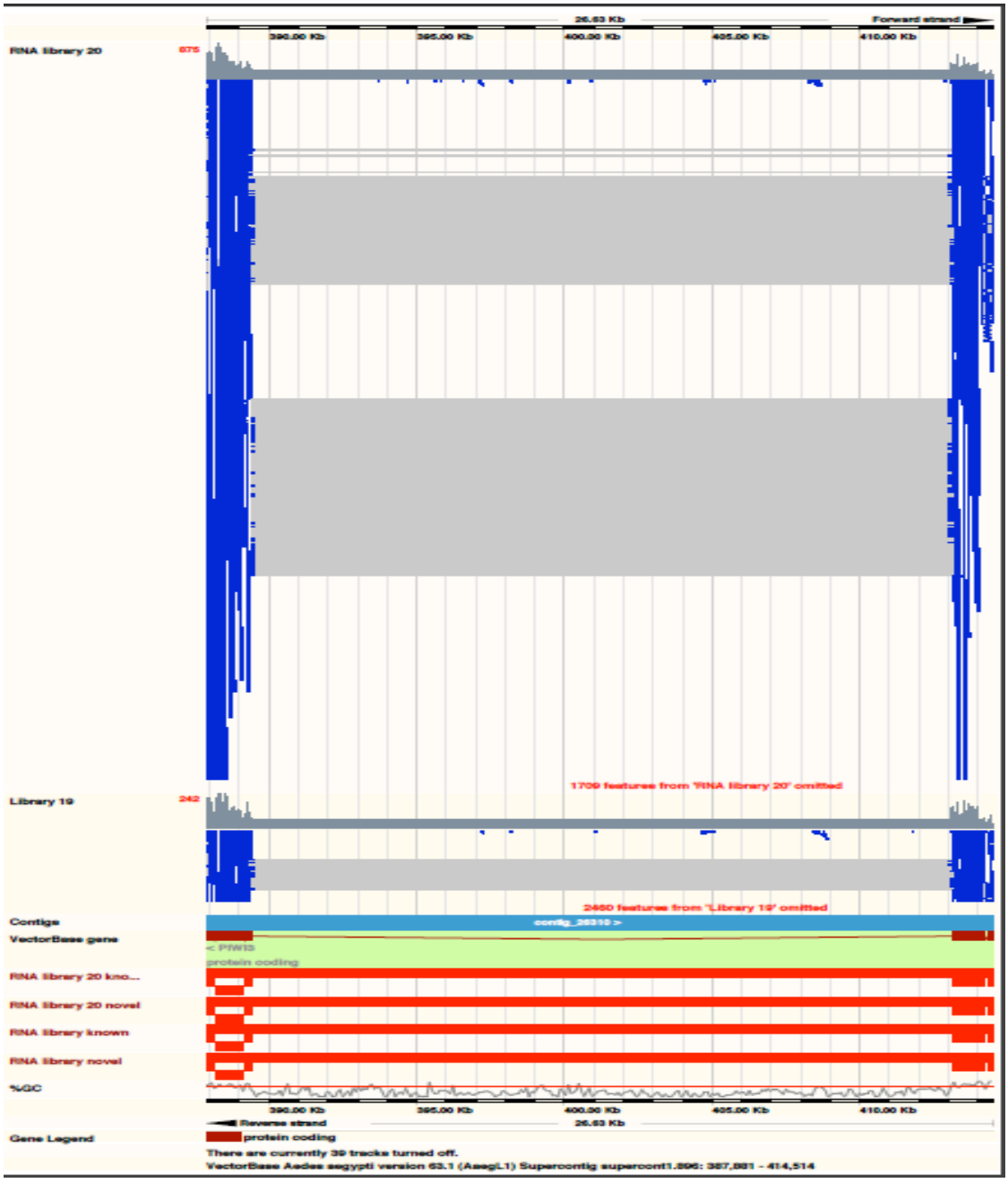

Figure. 4.9. RNAseq data for Piwi4 using libraries 19 (Egg) and 20 (ovary). The red bars are transcript predictions using the program cufflinks.

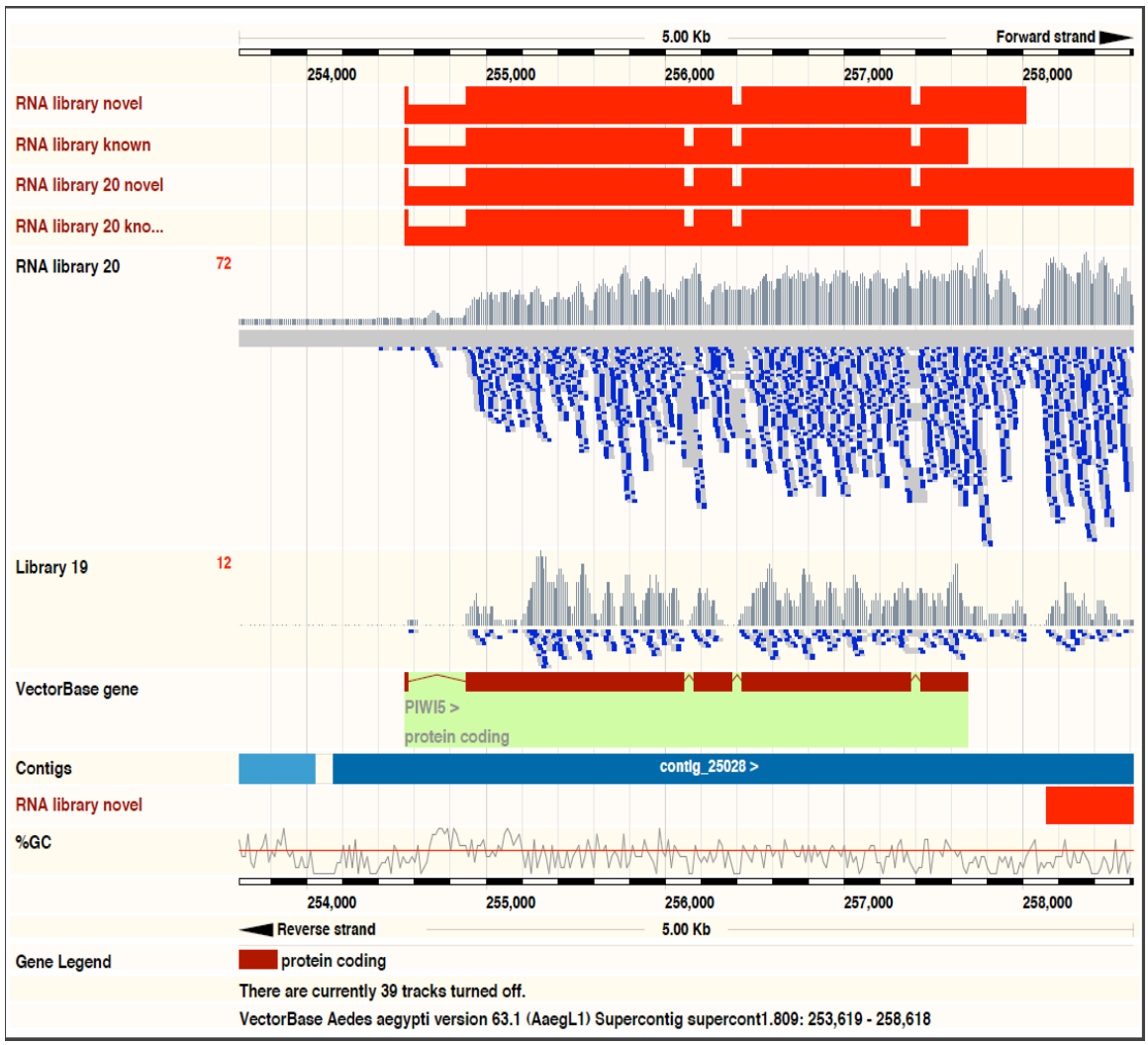

Figure. 4.10. RNAseq data for Piwi5 using libraries 19 (Egg) and 20 (ovary). The red bars are transcript predictions using the program cufflinks.

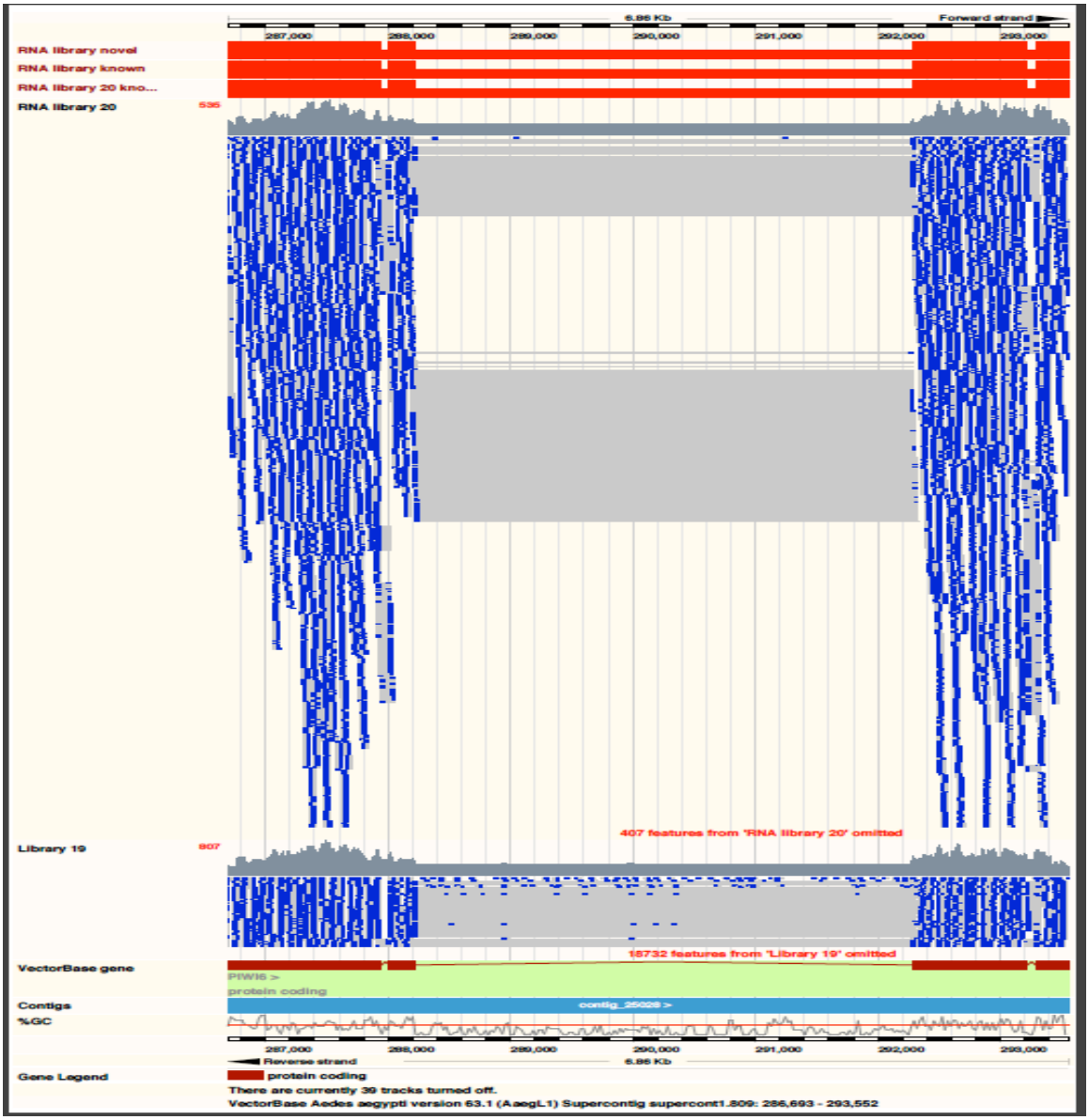

Figure. 4.11. RNAseq data for Piwi6 using libraries 19 (Egg) and 20 (ovary). The red bars are transcript predictions using the program cufflinks.

Figure. 4.12. RNAseq data for Piwi7 using libraries 19 (Egg) and 20 (ovary). The red bars are transcript predictions using the program cufflinks.

Figure. 4.13. RNAseq data for AGO3 using libraries 19 (Egg) and 20 (ovary). The red bars are transcript predictions using the program cufflinks.

Chapter 5

Summary and Conclusions

5.1 Summary

This dissertation examines 1) the affects of mutations on *AeBuster1* transposition activity and integration 2) the usefulness of *hyPBase* in Dipteran systems and 3) the properties of the piRNA pathway in *Aedes aegypti.*

5.2 *Aebuster1* Mutants are not Hyperactive in *Drosophila melanogaster*.

AeBuster1 mutants were screened using a high throughput assay in yeast in the Craig laboratory at Johns Hopkins University. Seven mutants that demonstrated the greatest increases in transposition activity were selected for testing in *Drosophila melanogaster*. Of those mutants tested L368P and L368P/V597A had the highest transposition rates of $3.49x\ 10^{-4}$ and $3.48x10^{-4}$ respectively. Unfortunately they were both at least 50% as active as wild-type *piggyBac* in the same trial. *PiggyBac* is one the most commonly used transposons due to its ability to carry large loads of cargo, high transposition activity and lack of over production inhibition (Wang and Fraser 1993; Li et al. 2001; Li et al. 2005; Mitra et al. 2008). Thus we had hoped to develop transposases that are at least competitive with *piggyBac* if they are to be deemed especially useful. No significant difference was found between the transposition rates of L368P and L368P/V597A compared to wild type *Aebuster1* in *D. melanogaster.* Thus even though

they demonstrated hyperactivity in yeast this activity failed to translate into *D. melanogaster*. *Aebuster1* mutants V168A/T324A/N465S and E260G/S485L had similar transposition frequencies around 2.5 X 10^{-4} while mutants L348V/L368P and V597A have frequencies around 5.5x 10^{-5}. No transposition events were found for the mutant E260G/S485L resulting in a transposition frequency of zero. This demonstrated that even when a transposase greatly outperforms wild-type *AeBuster1* in yeast they may not only be maintainted but greatly decreased in *D. melanogaster*.

None of the mutants tested demonstrated any sort of increase in transposition that could lead to its increased usefulness as a genetic tool. One suggestion that could help with the creation of hyperactive transposon for use in Dipteran systems would be to perform the initial high throughput screening assay in *D. melanogaster* cells lines. This may result in the discovery of transposases that once determined to be hyperactive in cell culture could more easily translate to increased *in vivo* activity.

Integration patterns were monitored in conjunction with transposition acitivity. And while the goal of increased activity was not achieved altered target sites were identified. Wild-type *Aebuster1* has an 8bp consensus sequence of nnnTAnnn (Arensburger *et al.* 2011). Target site preference is an important characteristic of transposases because increased or decreased target specificity can affect its usefulness as a genetic tool depending on the application; for instance in human gene therapy increased specificity can be desirable in order to keep the transposon from integrating into detrimental regions of the genome (Doherty et al. 2011) Thus it is important to determine

which of the amino acids in the protein are responsible for controlling target site preference.

When the transposon is inserted into the pGDVI and 8bp target region if found flanking the left and right ends of the transposon. These 8bp target sites were sequenced from insertions into the pGDV1 target plasmid for each of the mutants. Mutant L368P illustrated no significant difference between any of the 8 positions, however, mutant L368P/V597A demonstrated significant differences in positions 3, 4, 6, and 7 from the wild-type *AeBuster1*. Mutant V597A's target site preference is significantly different from wild-type at positions 3, 5 and 6. This information allowed us to deduce that V597 is an important amino acid for target site preference and that even though the mutation L368P alone had no significant affect on target site preference L368 may have a synergistic affect with L597 since the mutation of the two amino acids resulted in the greatest divergence from the wild-type *AeBuster1* target site preference. These data demonstrate that altered target site preference is achievable via the manipulation of a transposase protein in *D. melanogaster*.

Transposases often demonstrate a preference for insertion into specific locations within the pGDV1 target plasmid in transposition assays. These sites are known as "hot spots". Wild-type *AeBuster1* does not show strong bias toward any one location in the pGDV1 plasmid. All of the *AeBuster1* mutants demonstrated at least a small preference for insertion into position 1993 in pGDV1. L369P had the strongest preference for position 1993 with 46% of the insertions sequenced occurring at this hot spot. The

L368P/V597A and V597A did not demonstrate a dramatic increase in preference for a specific "hot spot". Thus even though a change in target site preference was demonstrated there appears to be no increased in specification associated with these two mutants.

This study did not result in the discovery of any hyperactive *AeBuster1* mutants for use in *D. melanogaster*. The study did however improve our understanding of how specific amino acids govern *AeBuster1* transposition. In-depth studies that investigate how specific mutations affect integration, frequency and transposition rates provide critical information regarding the potential manipulation of transposable elements. Our results demonstrate that the target site preference of transposase can by altered via mutation of specific amino acids. We also demonstrated that transposases with increased in activity in yeast do not necessarily translate into hyperactivity in *D. melanogaster*.

5.3 *hyPBase* does not demonstrate hyperactivity in *D. melanogaster* or *Ae. aegypti*.

HyPBase is a mutated *piggyBac* element that was demonstrated to be hyperactive in yeast and mammalian cell lines (Yusa et al. 2011; Doherty et al. 2011). We tested the somatic transposition rate of *hyPBas*e in *D. melanogaster* using a normalized interplasmid transposition assay in which the activity of the *hyPBase* transposase was compared with an internal control transposase. *HyPBase* had a transposition rate of 1.6 x 10^{-3} while the transposition frequency of the wild-type *piggyBac* transposase was 1.12 x 10^{-3} which was consistent with the frequency of wild-type *piggyBac* transposition in this insect observed by others (Lobo and Fraser 1999). This difference in normalized transposition activities between *hyPBase* and wild-type *piggyBac* transposase was not

statistically significant. While we did not find increased activity in the somatic assays we decided to test the germ-line activity since hyPBase had shown increased activity in other organisms.

A total of 192 embryos injected with wild-type *piggyBac* resulted in fertile crosses. Of these crosses 10% of G1 progeny expressed the GFP marker. Injections with *hyPBase* resulted in 145 fertile crosses with only 0.68% of the *hyPBase* crosses resulting in transgenic progeny. These data are in contrast to the activity of *hyPBase* in somatic cells however the decline in transformation rate corresponds with an increase in sterility. The wild-type *piggyBac* rate of sterility was approximately 8%. In contrast hyPBase had a GO sterility rate of 39%. The significant increase in sterility associated with the use the *hyPBase* may have been caused by several factors: higher rates of insertion of *piggyBac* thereby increasing its mutagenic effect, a direct effect of the transposase on chromosomal DNA, or a combination of these factors. To distinguish between these possibilities, we injected the *hyPBase* helper plasmid alone into embryos and measured G0 sterility. These experiments yielded a sterility rate of 23.0% indicating that increased integration of the transposon alone is likely not the only explanation of the increased sterility.

The increase in sterility rate led us to investigate the effects of *hyPBase* on the ovaries of G0 females. Gonadal dystrophy has been associated with hybrid dysgenesis in *D. melanogaster*. This hybrid dysgenesis results from the release of suppression of *P* element mobility when *P* element containing sperm fertilize an egg devoid of *P* elements in its genome (Kidwell *et al.* 1977). Brennecke *et al*. (2008) found that increased

mobility of transposons in the ovaries was associated with a lack of maternal deposition of piRNAs in the egg mapping to the transposon of interest.

In order to determine if the increased sterility we observed in G0 crosses injected with *hyPBase* was associated with gonadal atrophy we dissected ovaries from sterile female *D. melanogaster* that were injected with either wild-type *piggyBac* transposase or *hyPBase*. Wild-type *piggyBac* is associated with low levels of sterility. The ovaries dissected from sterile GO females injected with wild-type *piggyBac* varied from slightly smaller than un-injected females to normal size and shape as previously described in Spradling (1993). This is in contrast to hyPBase where all but one ovary dissected from sterile flies injected with *hyPBase* were completely atrophied with only the peritoneal sheath that surrounds the ovaries remaining. Additionally we injected flies with the *hyPBase* transposase without the donor plasmid and found that the ovaries from these sterile females also displayed the same type of atrophy.

Despite the high levels of sterility *hyPBase* caused in *D. melanogaster* we decided to test the transposase in *Ae. aegypti*. Generally the use of exogenous transposons has resulted in low transformation frequencies in *Ae. aegypti.* We reasoned that perhaps the decrease in the activity of *hyPBase* may lead to at least a detectable transformation frequency. However the *Ae. aegypti* interplasmid transposition assays demonstrated that there was no significant increase in hyPBases somatic activity when compared to wild-type *piggyBac*. In spite of this we attempted germ-line transformations of *Ae. aegypti*. No transgenic progeny were detected from 21 fertile crosses. Twenty one crosses may seem like a small amount but we observed a high degree of sterility (77%). Next we examined

the ovaries of these G0 females and compared them to ovaries dissected from adult female mosquitoes of the same age. The ovaries from *Ae. aegypti* G0 females displayed similar atrophy as seen in corresponding *D. melanogaster* females. We concluded that *hyPBase* has the same type of hyperactive behavior in *Ae. aegypti* as it does in *D. melanogaster*.

5.4 The piRNA Pathway in *Ae. aegypti*

The first step in understanding the piRNA pathway in *Ae. aegypti* was to perform RT- PCR in order to determine which of the proteins were expressed in either the carcass (whole adult body minus sexual organs), ovaries and eggs. This initial step suggested that PIWI2 and PIWI7 were germ-line specific leading us to chose them for further decision since we would expect germ-line tissues to be where transposon control is most essential. We also chose to further investigate AGO3 since it is involved in germ-line piRNA synthesis in *D. melanogaster*. The qPCR results showed that AGO3 is expressed at similar levels in the carcass, egg and ovaries while PIWI2 transcription was verified to be germ line specific with no transcripts being detected in the carcass samples. PIWI7 had appeared to be germ-line specific in the original RT-PCR but the qPCR data suggests that the PIWI7 is also found in the carcass. Immunopreciptiation of PIWI2 revealed that 29% of the isolated piRNA map to transposons. The PIWI7 pull down data set contained half as many sequences mapping to transposons with 12.11% of the total data base mapping to transposons. The AGO3 pull down library suggested had similar percentages of piRNA mapping to transposons as PIWI 7 with 14.59% mapping to transposons.

Immunoprecipitations pull (IP) downs of AGO3, PIWI2, PIWI7 and EGFP were further analyzed using Mass Spectrometery (MS). The PIWI2 MS data revealed that PIWI3 was pulled down as well as PIWI2 however it was detected at lower levels than PIWI2. This suggests that either PIWI2 is forming a complex with PIWI3 or that antibody may not specific to only PIWI2. PIWI7 is found in very low concentrations in the ovaries and so we were unable to detect the protein in the pull down and the other proteins detected did not vary significantly from those found in the EGFP pull down later the low transcript level was verified using the RNAseq data. For future studies it would be best to perform IP pull downs using egg tissue. The AGO3 IP pull down verified that we were pulling down AGO3 and while the MS detected other proteins they were not specific to the AGO3 IP pull down and thus were likely background.

Unexpectedly we were unable to achieve knock down of any of the proteins in the ovaries. Examination of the RNA SEQ libraries revealed that transcript levels of each of the targeted proteins was similar to that of our wild-type library. Previous studies have demonstrated that the injection of dsRNA into adult female mosquitoes can result in efficient knock down of proteins in the fat body however we have not seen any studies that report on the ability of dsRNA to enter the ovaries (Attardo *et al.* 2003). It is possible that this inability of dsRNA to enter the ovary could be due to some sort of defense system to dsRNA. Insects encounter RNA viruses and it is logical that the insect would harbor a system tasked with the protection of the ovaries from infection or exogenous RNA.

The knock down survivorship study may provide some insight into this problem. Our study showed a dramatic decrease in the production of viable F1 progeny when the FO female was injected with AGO3 or PIWI7. However those females injected with dsRNA targeted to PIWI2 did not vary from the control. luciferase dsRNA was injected into FO females as a mock control. The mock injections resulted in 85% of the injected females producing F1 adults. When AGO3 and PIWI7 dsRNA was injected in females the number of females that produced adult F1 individuals was greatly decreased to 45% and 30% respectively, resulted in adult FI progeny. Those females injected with dsRNA targeted to PIWI2 demonstrated no difference from the control with 80% percent of crosses resulting in F1 adults. These results indicate that AGO3 and PIWI7 may play a role in embryo development in the somatic tissues such as the fat body. It is also possibly PIWI2 mRNA in the embryo was protected in the ovaries from the dsRNA thus we saw no affects on survivorship.

5.5 Conclusions

5.5.1 Mutation of Transposases with the Goal of Hyperactivity

We had hoped that moving high throughput mutation screens from the prokaryotic *E. Coli* system to a Eukaryotic yeast system would result in a more rapid discovery of hyperactive transposases. Our studies demonstrate that hyperactivity in one Eukaryotic organism does not necessarily translate to hyperactivity in Dipteran systems. We tested two DNA transposases from the hAT Super family, *PiggyBac* and *AeBuster1*. *PiggyBac* is a commonly used transposase and has been effective in a diverse range of organisms. *AeBuster1* was discovered in *Aedes aegypti* and is believed to still be active

which is why it was choosen for our study. Mutation of the transposases did not lead to hyperactivity however it did help us to determine which amino acids are important to target site selections. Understanding target site selection is the first step toward the creation of transposases that can be directed to specific sequences.

5.5.2 The first look into the piRNA pathway

After being unable to create hyperactive transposases we turned our attention to the genomic regulation of transposons in *Aedes aegypti*. Research into transposon regulation in model organisms has revealed that small RNA regulatory pathways play a role in controlling transposon movement. In *D. melanogaster* the Piwi-interacting small RNAs (piRNAs) appear to function primarily in the regulation of transposons in both the germ-line and somatic tissues Thus we decided to perform functional studies of the piRNA pathway in *Aedes*. Our investigation demonstrated that there are six PIWI proteins and one AGO3 that are transcribed in *Ae. aegypti*. We choose three of the proteins to study more in depth: AGO3, PIWI2 and PIWI7. AGO3 and PIWI7 are expressed in both somatic and germ-line tissues and 13% of the piRNA associated with them map to transposable elements. PIWI2 is the only germ-line specific PIWI protein and 25% of the piRNA associated with it maps to transposable elements. The knock down of both AGO3 and PIWI7 resulted in a dramatic affect on the viability of F1 progeny. This suggests that these proteins may play a role in the somatic tissues that affect oogenesis. PIWI2 is associated with PIWI3 in the ovaries and appears to be the strongest candidate for playing a role in transposon regulation. These experiments are the

first step towards understanding the complexities of piRNA pathway in *Aedes aegypti*. Future studies should shed more light on how all of the PIWI proteins work in concert to regulate transposon movement and development.

5.6 References:

Attardo, G.M., Higgs, S., Klinger, K.A., Vanlandingham, D.L. and Raikhel, A.S. (2003) RNA interference-mediated knockdown of GATA factor reveals a link to anautogeny in the mosquito *Aedes aegypti. Pr.o Natl. Acad. Sci USA*. **100**: 13374-13379

Brennecke, J., Malone, C.D., Aravin, A.A., Sachidanandam, R., Stark, A., Hannon, G.J. (2008) An epigenetic role for maternally inherited piRNAs in transposon silencing. Science. **322**: 1387-1392

Doherty, J.E., Huye, L.E., Yusa, K., Zhou, L., Craig, N.L. and Wilson, M.H. (2011) Hyperactive *piggyBac* gene transfer in human cells in vivo. *Human Cell* Therapy. At Press

Kidwell, M.G., Kidwell, J.F. and Sved, J.A. (1997) Hybrid dysgenesis in *Drosophila melanogaster*: A syndrome of aberrant traits including mutation, sterility and male recombination. *Genetics*. **86**: 813-833

Li, X., Lobo, N., Bauser, C.A. and Fraser, M.J. Jr. (2001) The minimum internal and external sequence requirements for transposition of the eukaryotic transformation vector *piggyBac*. *Mol Genet Genomics.* **266**: 190-198

Li, X., Harrell, R.A., Handler, A.M., Beam, T., Hennessy, K. and Fraser, M.J. Jr. (2005) *piggyBac* internal sequences are necessary for efficient transformation of target genomes. *Insect Mol Biol.* **14**: 17-30

Lobo, N., Li, X. and Fraser, M. J., Jr. (1999) Transposition of the *piggyBac* element in embryos of *Drosophila melanogaster*, *Aedes aegypti* and *Trichoplusia ni*. *Mol. Gen. Genet.* **261**: 803–810

Mitra, R., Fain-Thornton, J. and Craig, N.L. (2008) *piggyBac* can bypass DNA synthesis during cut and paste transposition. EMBO J. **7:** 1097-109

Spradling, A. C. (1993) The development of *Drosophila melanogaster*. Cold Spring Harbor. **1**: 1-70

Wang, H.H. and Fraser, M.J. (1993) TTAA serves as the target sire for TFP3 lepidotpteran transposon insertions in both nuclear polyhedrosis virus and *Trichoplusia ni* genomes. *Insect Mol. Biol.* **1**: 109-1996

Yusa, K., Zhou, L., Li, M.A., Bradley, A. and Craig, N.L. (2011) A hyperactive *piggyBac* transposase for mammalian applications. *PNAS.* **108**: 1531-1536