IOWA STATE UNIVERSITY
Digital Repository

Graduate Theses and Dissertations

Iowa State University Capstones, Theses and Dissertations

2016

The development of tools for transposon somatic mutagenesis in zebrafish

Joshua Mauldin
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd

 Part of the Biology Commons

Recommended Citation
Mauldin, Joshua, "The development of tools for transposon somatic mutagenesis in zebrafish" (2016). *Graduate Theses and Dissertations*. 15768.
https://lib.dr.iastate.edu/etd/15768

This Thesis is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

The development of tools for transposon somatic mutagenesis in zebrafish

by

Joshua Mauldin

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Genetics, Developmental, and Cellular Biology

Program of Study Committee:
Maura McGrail, Major Professor
Jeffrey Essner
Kristen Johansen

Iowa State University

Ames, Iowa

2016

Copyright © Joshua Mauldin, 2016. All rights reserved.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Maura McGrail, and my committee members, Dr. Jeff Essner and Dr. Kristen Johansen for their guidance and support. I would also like to thank all the members of the McGrail and Essner labs.

ABSTRACT

Transposon somatic mutagenesis is a forward genetics approach that can be used to identify novel genetic drivers of cancer. This research has developed tools required for performing these transposon mutagenesis screens in zebrafish. Transgenic zebrafish lines were created that allow for tissue-specific and inducible expression of the transposase required for the *Sleeping Beauty* (*SB*) transposon system. Mutagenesis screens will be performed in a pRB deficient background, which gives rise to highly proliferative and undifferentiated tumors resembling primitive neuroectodermal tumors (PNETs). This could reveal cooperating genes that interact with *rb1* to affect brain cancer. To this end, neural progenitors were targeted using the *Tg(krt5:SB11)* transgenic line. The ability of this transgenic to mobilize transposons was demonstrated in whole larvae and adult brains. Also, an inducible transposase source in the *Tg(ubi:floxed-SB11)* transgenic line was able to increase transposon mobilization in response to *Cre* injection, despite a low level of background expression. Other tissue-specific promoters were explored, including *gfap* and *nestin*, which may result in increased tumorigenesis.

Several mutagenesis screens are currently being monitored in both wild type and tumor susceptible backgrounds. Common integration sites in the zebrafish genome are also being mapped to aid in the analysis of data from these screens by removing background that does not contribute to tumorigenesis. A small scale mutagenesis screen in a wild type background has resulted in a single fish out of ten that developed a brain tumor after 8 months of age. Importantly, analysis showed that transposon mobilization was specific to tumor tissue, demonstrating that the *Tg(krt5:SB11)* transgenic line can

target mutagenesis to specific cell types. The results of these transposon mutagenesis screens could lead to a better understanding of the genetics behind PNETs and open possibilities for new treatments.

CHAPTER I

INTRODUCTION

History of Transposons

Since the discovery of transposable elements by Barbara McClintock in the 1940's, transposons have revolutionized the field of genetics. Her research into the *Activator/Dissociation* (*Ac/Ds*) transposon system in maize was the first description of mutagenesis through transposition (McClintock, 1950 & 1953).Transposons are defined as segments of DNA that are capable of changing their location in the genome, leading to the popular name jumping genes. It is now known that transposons are present in almost every organism from prokaryotes to humans, and in many cases, make up a significant portion of the host genomes. Their discovery has not only led to a better understanding of our evolution and genetic past, but also has provided important genetic tools for research.

Types of Transposons

Transposable elements can be divided into two main groups corresponding to their mechanism of action (Pray, 2008). Class I transposons are known as retrotransposons, named for the requirement of reverse transcriptase for mobilization. Transposition is achieved by transcribing the transposon into an RNA intermediate. Reverse transcriptase, encoded by the transposon, then produces a DNA copy of the transposon which can be integrated back into the host genome at the target site. The result is the duplication of the transposon from one site to the other, while leaving the original transposon behind. As a result, this is referred to a copy and paste mechanism.

The other group of transposon fall into class II, or DNA transposons. These do not require a RNA intermediate, but instead rely on a transposase that recognizes and binds to the transposon, excises it from the donor DNA, and then inserts it at the new target site. In contrast to the retrotransposons, DNA transposons exhibit a cut and paste mechanism. Due to the sequence specificity that each transposase recognizes, different transposons require a certain transposase to be mobilized.

In addition to these broad classifications, transposon can be further divided based on their ability to catalyze their own transposition. A transposon that encodes for the transposase or reverse transcriptase required for its own function is said to be autonomous. A non-autonomous transposon lacks this ability, and therefore requires other transposable elements for its mobilization.

Transposons in Research

The development of transposons has led to many research applications. In particular, the *Tol2* and *Sleeping Beauty* transposon systems have had significant influence in research and are two transposons that will be used in this research. The *Tol2* transposase system is a well characterized method for introducing transgenes into organisms (Kawakami, 2007 & Ni et al, 2016). It is a class II, DNA transposon that requires the *Tol2* transposase to catalyze transposon integration in the host genome. Transgenes are constructed in a *Tol2* transposon flanked by inverted repeats. For zebrafish, this DNA construct and transposase mRNA are injected into zebrafish at one-cell stage. Although this results in mosaic integration of the transgene in injected fish, transgenic lines can be recovered from their offspring.

Sleeping Beauty Transposon

A good example of a transposable element being used for research purposes is the *Sleeping Beauty* transposon system. As the name implies, *Sleeping Beauty* (*SB*) was an inactive transposon found in the genome of a salmonid species that was engineered to restore its original function (Ivics et al., 1997). A putative functional transposon sequence was generated based on phylogenetic analysis to obtain a consensus sequence with conserved domains. The revived *SB* transposon was then synthesized by eliminating the deleterious mutations to match this consensus sequence. This involved restoring the open reading frame of the transposase and repairing mutations to the DNA binding and catalytic domains. The restored transposase was named *SB10* and was successfully used to drive transposition in human cells (Ivics et al., 1997).

The *SB* transposon falls into the *Tc1/mariner* superfamily. This family includes DNA transposons that are present in a wide variety of organisms. This makes *SB* advantageous for use in vertebrate research as it has a wider host range (Plasterk et al., 1999). Structurally, the transposon contains two terminal inverted repeats flanking the *SB* transposase gene (Ivics et al, 2015). For research purposes, however, the transposase gene can be replaced with a gene of interest or other genetic cargo, and the transposase supplied through another method. This allows the *SB* transposon system to be used for a wide range of research applications (Izsvák et al., 2004).

As a class II transposon, *SB* is mobilized by the cut and paste method (Walisko et al., 2008). Transposition begins when *SB* transposase binds to the inverted repeats on the transposon, with two units of the transposase at each end. The synaptic complex forms as the transposase subunits bind, bringing the two ends together, resulting in DNA cleavage.

Finally, the transposon is excised from the donor site and is introduced at a new target site. The resulting double stranded break in the donor DNA is repaired by non-homologous end joining. One of the advantages of the *SB* system is that the target site requirements for transposition have few restrictions, only requiring a TA dinucleotide. These qualities of have made the *SB* transposon an important tool in research.

CHAPTER II

TRANSPOSON SOMATIC MUTAGENESIS

Background

The use of transposon insertional mutagenesis is valuable for forward genetic screens. This kind of approach is important as it reveals new targets for cancer research to pursue (Mann et al., 2014). Several studies have been done in mice using the *SB11* transposon system to identify cancer genes through the disruption of random genes throughout the genome (Collier et al, 2005, Dupuy et al, 2005, Moriarity et al., 2015), and the system has been used in zebrafish by our lab (McGrail et al., 2011). Zebrafish offer several advantages as a vertebrate model organism for forward genetic screens. They have a short generation time and high fecundity, resulting in great numbers of individuals in a relatively short amount of time. Their genome has been sequenced and a number of genetic tools have been developed for their use, including transposons for generating transgenics and genome editing techniques. The ability to study specific genes and their role in cancer has improved dramatically since the utilization of TALENs and CRISPRs. However, forward genetic approaches are still required to better understand tumorigenesis by identifying novel genes that contribute to the cancer phenotype.

Transposon mutagenesis screens are often carried out in a tumor-susceptible background. This allows for the discovery of cooperating genes that interact to contribute to cancer. In the pRB model, in which *rb1* is targeted using TALENs, adult mosaic mutants develop tumors that share characteristics of a type of cancer called PNETs (Solin et al., 2015). Primitive neural ectodermal tumors (PNETs) are a form of highly malignant

brain cancer that usually affects children and has a low survivability (Smoll, 2012). A somatic insertional mutagenesis screen in the pRB model should identify cooperating genes leading to the PNET phenotype.

The somatic insertional mutagenesis screen using the *SB11* system involves two main components. First, a gene break transposon array is required to disrupt gene express after integration. Transposons such as the *T2/OncZ* enable the identification of both tumor suppressors and oncogenes through either over expression or premature termination. The second component is the transposase source. When *SB11* transposase is introduced, transposon mobilization results in the random integration of the transposon throughout the genome. This transposase source can either be transient, such as injected mRNA for the protein, or expressed by the cell through transgenics. The use of promoters can determine if *SB11* expression is constitutive or limited to specific tissues.

In this research, a tissue-specific and inducible transposase source will be used to drive somatic transposon mutagenesis to identify novel genes that contribute to tumorigenesis. The ability to direct mutagenesis to specific cell types in an inducible manner is a very valuable technique in studying tumorigenesis. It would allow mutagenesis in the cell types of interest and their future lineages, as well as after critical stages of development. Neural progenitors are of particular interest for our PNET model. Several reports have shown that neural progenitors can be targeted through the use of the *keratin5* (*krt5*) promoter. Tumorigenesis was promoted when using the *krt5* promoter as a driver to activate the Sonic hedgehog signaling pathway (Ju et al., 2014). It was also used to drive overexpression of human KRAS in zebrafish, resulting in malignant brain tumors

(Ju et al., 2015). Therefore, this promoter will be used to drive *SB11* transposase expression for the mutagenesis screens.

In addition to targeting the transposon mutagenesis screen in specific tissues, it can also be controlled using an inducible transposase source. This will be accomplished using the *Cre-Lox* system and a tamoxifen-inducible *Cre*. Expression of *SB11* transposase will be prevented until the tamoxifen is introduced. Driving the inducible *Cre* from a tissue specific promoter would allow for both the spatial and temporal control of the mutagenesis screen. Taking this forward genetics approach, the transposon mutagenesis screen should reveal novel genetic drivers of tumorigenesis and provide candidate genes for future study. This will help gain a better understanding of the mechanisms of tumorigenesis and could provide insights into possible treatments.

Methods

Cloning constructs for the expression of a transgene

All PCR reactions described were performed with KOD master mix, following the manufacturer's protocol. Primer sequences are provided in Table 1. The *krt5* promoter was cloned from zebrafish genomic DNA. Two pairs of primers were designed for a nested PCR. In the first round of PCR, the target sequence was amplified with the surrounding 30-40 bps. The second round of PCR used the resulting product as template to amplify up the target sequence to obtain the *krt5* promoter flanked by restriction sites. The restriction sites NotI and NheI were used to insert the amplified promoter into a vector to create *pTol2<krt5:SB11, cmcl2:GFP>*. The construct *pTol2<gfap:SB11, cmlc2:GFP>* was similarly made by amplifying the *gfap* promoter from a plasmid

obtained from Addgene (#39761). The *pTol2<krt5:SB11, cmcl2:GFP>* plasmid was modified to *pTol2<krt5:GFP>* by digesting with NheI and AgeI to remove the *SB11* coding region to the *cmcl2* promoter, followed by a blunt end ligation.

The *ubiquitin* promoter used in the *Tg(ubi:floxed-SB11)* construct was obtained from Addgene (#27320). The promoter region was excised using the restriction sites NheI and BamHI, and inserted into a vector to create *pTol2<ubi:floxed-SB11, cmlc2:GFP>*. The termination signal consisting of two SV40 sequences surrounded by *LoxP* sites was amplified from a plasmid template. The BamHI restriction sites were used to insert this transcription terminator sequence between the *ubiquitin* promoter and *SB11* coding region to yield *pTol2<ubi:floxed-SB11, cmlc2:GFP>*. The *pTol2<ubi:GFP>* plasmid was obtained by digesting *pTol2<ubi:floxed-SB11, cmlc2:GFP>* with BamHI and AgeI to remove the *floxed-SB11* and *cmlc2* promoter, followed by a blunt end ligation. The *nestin* promoter was amplified from genomic DNA and cloned directly into the pCR-BluntII-Topo vector. This allowed the promoter to be cloned into the *p5'E:MCS* vector using the restriction sites KpnI and SacII to yield *p5'E:nestin.*

Table 1. Primer sequences for cloning constructs.

Primer Name	Sequence (5' to 3')
krt5-nested-forward	GCTCATGCTTCCACTACTGG
krt5-nested-reverse	AGAACTGGGAAACTGAGACTG
krt5-promoter-forward	ATTAGCGGCCGCGAAAGCGACTCCACCC
krt5-promoter-reverse	ATTAGCTAGCGAGGGGGTGAGGATCAGA
loxP-SV40-forward	ATTAGGATCCTCCGGAATTCATAACTTCG
loxP-SV40-reverse	ATTAGGATCCTGCAGGATATCGATATAACTTC
gfap-promoter-forward	ATTAGCTAGCGAGGTAAGGACTGAGGTG
gfap-promoter-reverse	ATTAGCATGCGGTGGAGGAGAATGAG
nestin-promoter-forward	AGTCGGTACCAGGAAGTTTCCAGTGTTGTG
nestin-promoter-reverse	AGCTCCGCGGTCACTAGAGGCTGTGAAGAA

Generating and characterizing transgenic zebrafish lines

Transgenics were created by co-injecting the vector containing the transgene in the *Tol2* backbone with *Tol2* transposase. For each transgenic line, 50 pg of the construct and 125 pg of the *Tol2* mRNA was injected. Injected embryos were screened for a reporter gene and founders were identified.

SB11 expression in the *Tg(krt5:SB11)* was analyzed using whose mount *in situ* hybridization on 2 day old larvae. The probe was the entire antisense *SB11* coding region. The control included was the *leptin b* probe, which has been used previously. The ability of these transgenic lines to mobilize the transposon was determined using a transposon excision assay, in which PCR was used to amplify specific sequences along the transposon. Primers were previously designed to amplify across the excision site and the left arm of the transposon. These sequences are provided in Table 2. The *Tg(ubi:floxed-SB11)* system was tested by injecting of 50 pg *Cre* mRNA into progeny from a cross with *T2/OncZ*. Injected and uninjected larvae were raised to 5 days of age, separated according to genotype, and used to extract genomic DNA. The transposon excision assay was used to determine if *SB11* was being expressed.

Table 2. Primer sequences for the transposon excision assays.

Primer Name	Sequence (5' to 3')
T2/OncZ-forward	ATGTGCTGCAAGGCGATTAAGTTG
T2/OncZ-arm-reverse	ATCAAGCTTCTAAAGCCATGACATC
T2/OncZ-across-reverse	TGAGCGGATAACAATTTCACACAGG
T2/OncZgb-forward	TGTAAAACGACGGCCAGT
T2/OncZgb-arm-reverse	GACCCACTGGGAATGTGATGAAAG
T2/OncZgb-across-reverse	GTCGACGGTATCGATAAGCTTG

Preparing sample libraries for common integration site analysis

A cross was setup between the *Tg(β-actin:SB11)* transgenic and the *T2/OncZgb1* transposon line. Progeny were separated based on genotype and raised to adulthood. At 2 months of age, the fish were sacrificed and the entire trunk of the body used for genomic DNA extraction. Linker mediated PCR was used to amplify insertion sites (McGrail et al., 2011). The libraries were constructed using the primers listed in Table 3. A total of 48 samples were prepared for each arm of the transposon. Samples were barcoded with one of 24 barcoded primers, listed in Table 4. The first set of samples containing 24 libraries from the left transposon arm were sent for sequencing on the MiSeq with 150-cycles.

In order to ensure that the common integration site libraries contained sequences of the transposon flanking genomic DNA, random clones were isolated and sequenced. Each sample was run on a gel and excised, purified, and used for TA cloning. This allowed random fragments from each sample to be cloned into the pCR4-Topo vector. Clones were selected and individual colonies were grown up. The plasmid DNA was sent for standard Sanger sequencing.

Table 3. Primer sequences for common integration site library construction.

Primer Name	Sequence (5' to 3')
NlaIII_Linker_+	GTAATACGACTCACTATAGGGCTCCGCTTAAGGGACCATG
NlaIII_Linker_-	Phos-GTCCCTTAAGCGGAGC-3'spacer
BfaI_Linker_+	GTAATACGACTCACTATAGGGCTCCGCTTAAGGGAC
BfaI_Linker_-	Phos-TAGTCCCTTAAGCGGAGC-3'spacer
Primary_IRR1	GCTTGTGGAAGGCTACTCGAAATGTTTGACCC
Primary_IRL1	CTGGAATTTTCCAAGCTGTTTAAAGGCACAGTCAAC
Primary_Linker	GTAATACGACTCACTATAGGGC
Secondary_Linker	CAAGCAGAAGACGGCATACGAGCTCTTCCGATCTAGG GCTCCGCTTAAGGGAC

Table 4. Barcoded primers for common integration site library construction. Barcodes are shown in red.

Primer Name	Sequence (5' to 3')
Secondary_IR-BC1	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTAGGAGTTGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC2	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTGCGAGTTGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC3	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTCTGAGTTGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC4	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTAACAGTTGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC5	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTGGCAGTTGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC6	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTTCCAGTTGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC7	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTGATAGTTGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC8	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTCGTAGTTGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC9	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTACTAGTTGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC10	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTGAAGGTTGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC11	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTAGAGGTTGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC12	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTCCAGGTTGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC13	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTATCACGTGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC14	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTCGATGTTGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC15	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTTTAGGCTGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC16	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTTGACCATGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC17	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTACAGTGTGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC18	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTGCCAATTGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC19	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTCAGATCTGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC20	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTACTTGATGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC21	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTGATCAGTGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC22	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTTAGCTTTGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC23	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTGGCTACTGTATGTAAACTTCCGACTTCAACTG
Secondary_IR-BC24	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTCTTGTATGTATGTAAACTTCCGACTTCAACTG

Crosses performed for transposon mutagenesis screens

A transposon mutagenesis screen was performed using the *Tg(krt5:SB11)* transgenic and *T2/OncZ* transposon line. Progeny of the cross were separated according to genotype. At five days of age, several larvae were used to extract genomic DNA for a transposon excision assay. The remaining fish were raised to adulthood. Of the double transgenics, one developed a tumor originating from the head. The fish was sacrificed and portion of the tumor and muscle was removed for extracting genomic DNA using the Qiagen Blood and Tissue Extraction kit. A transposon excision assay was used to show transposon mobilization in the tumor tissue.

Two large scale screens were performed with the *Tg(krt5:SB11)* and *T2/OncZgb1* transposon line. The first was in a wild type background. Transposon mobilization in double transgenics was verified by dissecting adult brains and extracting genomic DNA that was used for the excision assay. An additional mutagenesis screen was carried out in the pRB deficient background by injecting TALENs for *rb1* at the one-cell stage. The fish from both screens are still being monitored for tumorigenesis.

Results

Isolation of transgenic lines for tissue specific and inducible transposase sources

In this research, transposon somatic mutagenesis will be used to identify genetic drivers that contribute to the development of the PNET phenotype. To accomplish this, tissue specific and inducible transposase sources were explored. The design of the constructs used to create these transgenic zebrafish lines is illustrated in Figure 1. The number of independent lines of each transgenic is summarized in Table 5. The targeting

of neural progenitors is of particular interest, as these cells are thought to be the origin of over proliferation in our model. Studies have shown that the *krt5* promoter could be used to drive tumorigenesis in the brain by targeting neural progenitors (Ju et al., 2014). This promoter was therefore chosen to drive transposase expression for the mutagenesis screens. In total, three independent transgenic lines of *Tg(krt5:SB11)* were isolated.

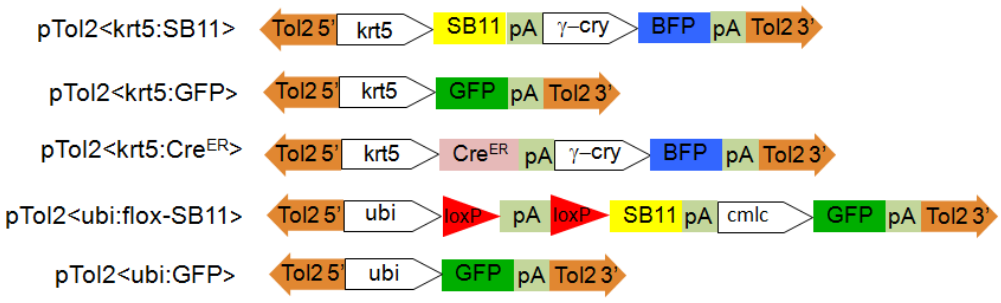

Figure 1. Transposon construction design to generate transgenics, including the tissue-specific and inducible expression of a transposase.

Table 5. Number of independent lines of each transgenic that has been isolated.

Transgenic Line	Independent Lines
Tg(krt5:GFP)	2
Tg(krt5:SB11)	3
Tg(ubi:flox-SB11)	2
Tg(ubi:GFP)	1

The ability of the *Tg(krt5:SB11)* transgenic line to drive transposon mobilization was tested with a small scale mutagenesis screen by crossing with the *T2/OncZ* line. Transposon mobilization was measured using an excision assay, which is shown in Figure 2. In this assay, amplification across the excision site is achieved only when the transposon has been mobilized. The 300 bp band observed in the double transgenics, individuals with both the transposon and transposase, indicates transposon excision.

Controls for the presence of the transposon are included, showing amplification of a 420 bp product on the transposon arm. This confirms that the *Tg(krt5:SB11)* can drive transposon mobilization as expected.

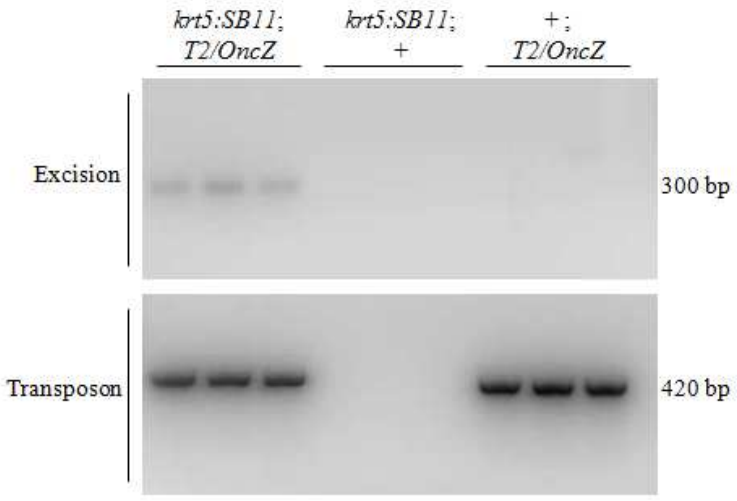

Figure 2. Excision assay of whole larvae showing mobilization of the *T2/OncZ* transposon when crossed to *Tg(krt5:SB11)*.

The large scale screens will be using the *T2/OncZgb1* transposon. The function of the *Tg(krt5:SB11)* transgenic line to catalyze mobilization of the *T2/OncZgb1* transposon was confirmed using an excision assay of dissected brains. The results of this assay are shown in Figure 3. As with the previous excision assay, the amplification across the excision site can only be after transposon mobilization, indicated by the red arrow. The results show that *Tg(krt5:SB11)* can be used for the large scale mutagenesis screens.

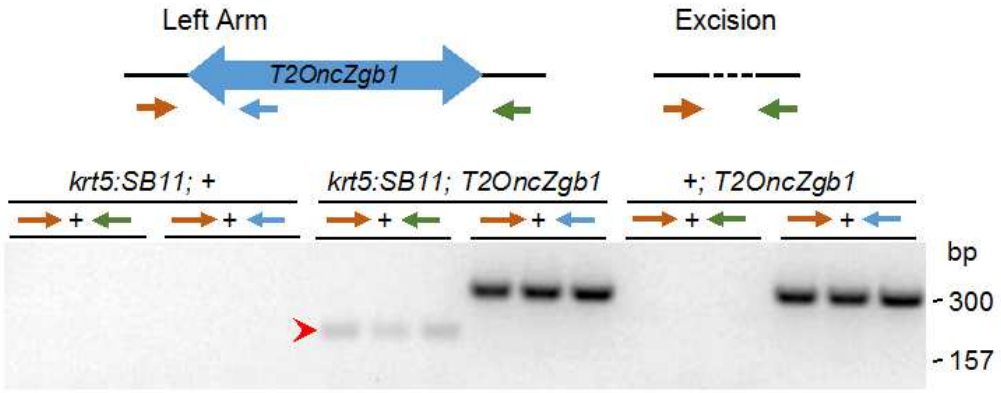

Figure 3. Excision assay of dissected brains showing mobilization of the *T2/OncZgb1* transposon when crossed to *Tg(krt5:SB11)*

The expression pattern of the *Tg(krt5:SB11)* transgenic was analyzed in 2 day old larvae using whole mount *in situ* hybridization. The results are shown in Figure 4. The results were inconclusive, as there was not an appreciable difference between the transgenics expressing *SB11* and the wild type. The probe designed for the *in situ* consisted of the entire antisense sequence of *SB11*. Analyzing this sequence against the zebrafish genome resulted in a high level of homology. This could explain the presence of the probe in the wild type. In future efforts to characterize *SB11* expression in the *Tg(krt5:SB11)* transgenic, an alternative probe will be used that does not map to the zebrafish genome.

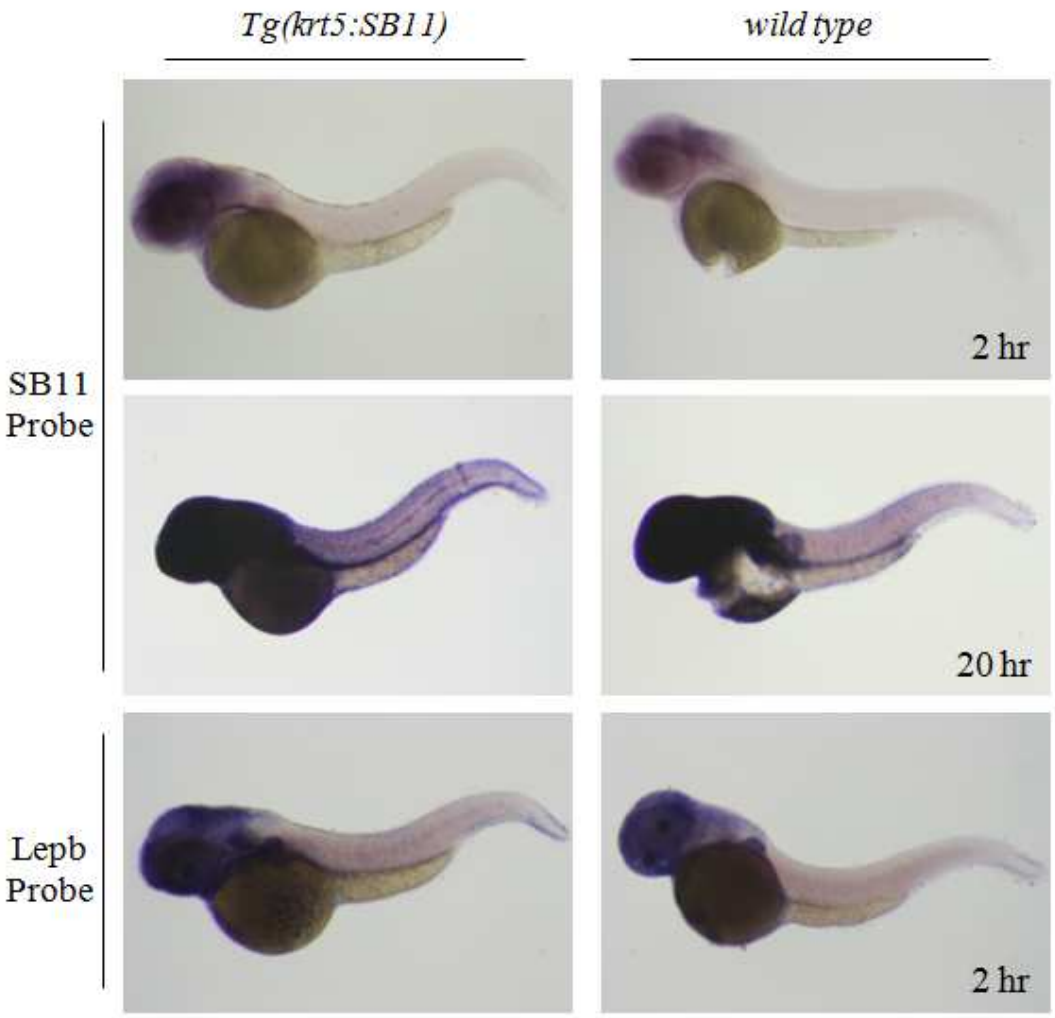

Figure 4. Whole mount in situ hybridization to characterize the expression of *SB11* transposase in *Tg(krt5:SB11)*.

Another method was explored to visualize the activity of the *krt5* promoter. The *Tg(krt5:GFP)* transgenic line was isolated to observe the expression of GFP, in order to have some insight into the *SB11* expression pattern in the *Tg(krt5:SB11)* line. The confocal image shown in Figure 5 shows intense GFP expression throughout the skin, as expected, as well as significant expression in the brain. Sectioning of these fish would provide a more detailed visualization of GFP throughout the brain.

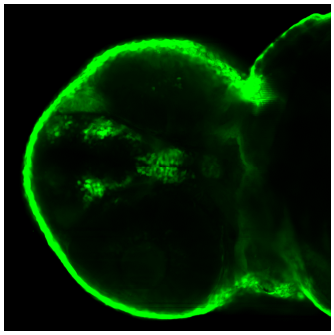

Figure 5. GFP expression in the skin and brain in *Tg(krt5:GFP)* at 2 dpf.

In addition to tissue-specificity, and inducible transposase source would be advantageous for initiating transposon mutagenesis after critical stages of development. A common method for achieving inducible control of a transgene is the *Cre-Lox* system, which was to be used in this research. One part of this system has been developed so far. Three independent lines of the *Tg(ubi:floxed-SB11)* transgenic have been isolated. This transgenic line was generated in order to achieve inducible control of transposase expression with *Cre* recombinase. To test its ability to induce transposon mobilization in the presence of *Cre*, this line was crossed with the *T2/OncZ* line and *Cre* mRNA was injected. A transposon excision assay revealed a relatively low amount of transposon excision in non-injected individuals, as shown in Figure 6. In individuals injected with *Cre* mRNA, the band corresponding to transposon excision is much more intense. These results suggest that the transcription terminator that was meant to prevent transcription of the *SB11* transgene is not completely effective. Therefore, a low level of *SB11* expression is ubiquitously expressed even in the absence of *Cre* recombinase. Despite this leakiness, the presence of *Cre* does increase transposon mobilization. With refinement, this system could be used in conjunction with an inducible *Cre* for temporal control over the transposon mutagenesis screens.

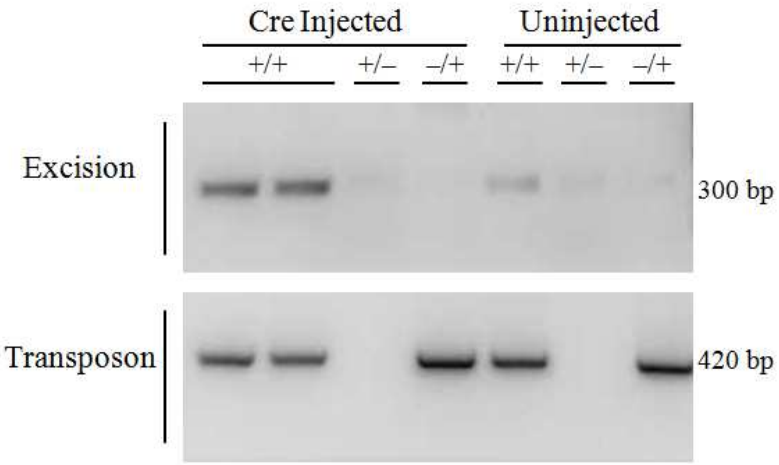

Figure 6. Excision assay demonstrating inducible transposon mobilization with injected *Cre* mRNA. Genotypes refer to the presence of transposase/transposon.

Other tissue specific promoters expressed in the brain that were explored include *gfap* and *nestin* promoters. These promoters have been used as reporters for radial glia and stem cells in the brain (Lam et al., 2009). Progress has been made in cloning these promoters into several constructs, but transgenic lines have not yet been produced. The *gfap* promoter has been cloned into the *pTol2<gfap:SB11, cmlc2:GFP>* vector to drive *SB11* expression. However, there are alternative translation start sites upstream of the *SB11* coding region that may prevent or hinder *SB11* from being successfully translated. Before this construct is used to create a transgenic line, these alternative start sites should be removed. The *nestin* promoter was cloned successfully from genomic DNA and inserted into a multiple cloning site vector.

The construction of these plasmids opened up the ability to create another transgenic line that would be beneficial in the lab. This construct was modified to drive the ubiquitous expression of GFP, which lead to the establishment of the *Tg(ubi:GFP)* line. This transgenic has been used for several experiments involving transplantation.

The transgenics expressing transposase sources were designed and constructed in this research to catalyze the mobilization of a transposon in two other lines, shown in Figure 7. The first is *T2/OncZ* that has been used in the lab previously (McGrail et al., 2011). It contains a termination signal to stop transcription if integrated within a gene, and a *β-actin* promoter to drive overexpression of a downstream gene. The other transposon that will be used is *T2/OncZgb1*. It has the additional advantage of tagging disrupted proteins with RFP when inserted in the correct reading frame. Both of the lines were previously created in the lab by injecting linear DNA into zebrafish embryos. The linear DNA formed a concatemer consisting of many copies of the transposon, and this transposon array was inserted randomly into the genome. In this research, the *T2/OncZ* transposon line will primarily be used at a small scale to test transposase sources, while the *T2/OncZgb1* transposon will be used for the large scale mutagenesis screens.

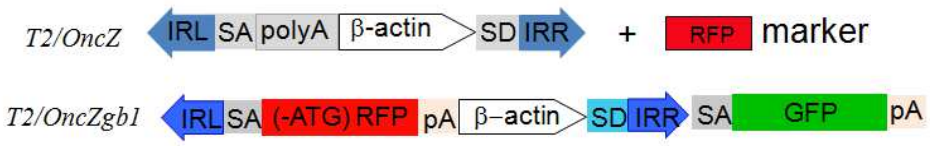

Figure 7. Schematic of the two transposon lines that will be used in this study.

Mapping common integration sites of the T2OncZgb1 transposon following mobilization

Common integration sites are being mapped following the mobilization of the *T2OncZgb1* transposon by the *Tg(β-actin:SB11)* transgenic line. This will identify integration hotspots in the genome, which will allow these sites to be removed as background when analyzing the data from somatic mutagenesis screens. At 2 months of age, the double transgenics were sacrificed and genomic DNA was extracted for library

preparation. In total, 48 individual fish produced libraries for the right and left transposon arms. Each library was prepared with a barcoded primer that would allow multiplexing during sequencing.

In order to verify that the constructed libraries contained fragments of genomic DNA surrounding the transposon insertion site, several random amplified products were cloned and sequenced using standard sequencing. The results revealed that most of the samples contained transposon concatemer sequence, instead of mapping to the genome. However, these segments were flanked by the sequencing and linker primers used to construct the libraries. In addition, the barcode sequence corresponding to the library sample was recovered. This suggested that the library construction was successful and ready for next generation sequencing.

The MiSeq has been performed for the first 24 samples of the left arm of the transposon. Sequencing returned a total of 18.3 million reads. 17.3 million reads remained after removing sequences that did not have the TA dinucleotide requirement for integration. A large percentage of the read mapped to the transposon itself, and only 2% of the read mapped to the genome. This indicates that the libraries contain a significant amount of intact transposon, pointing to an incomplete secondary digestion. Analysis is still ongoing, but a few common integration plots have been constructed for one barcoded sample, which are shown in Figure 8. In particular, there appears to be a common integration site on chromosome 3 at the *β-actin* promoter. This was observed previously as well (McGrail et al., 2011). The process of mapping the reads to the zebrafish genome will identify common integration sites, and is currently underway.

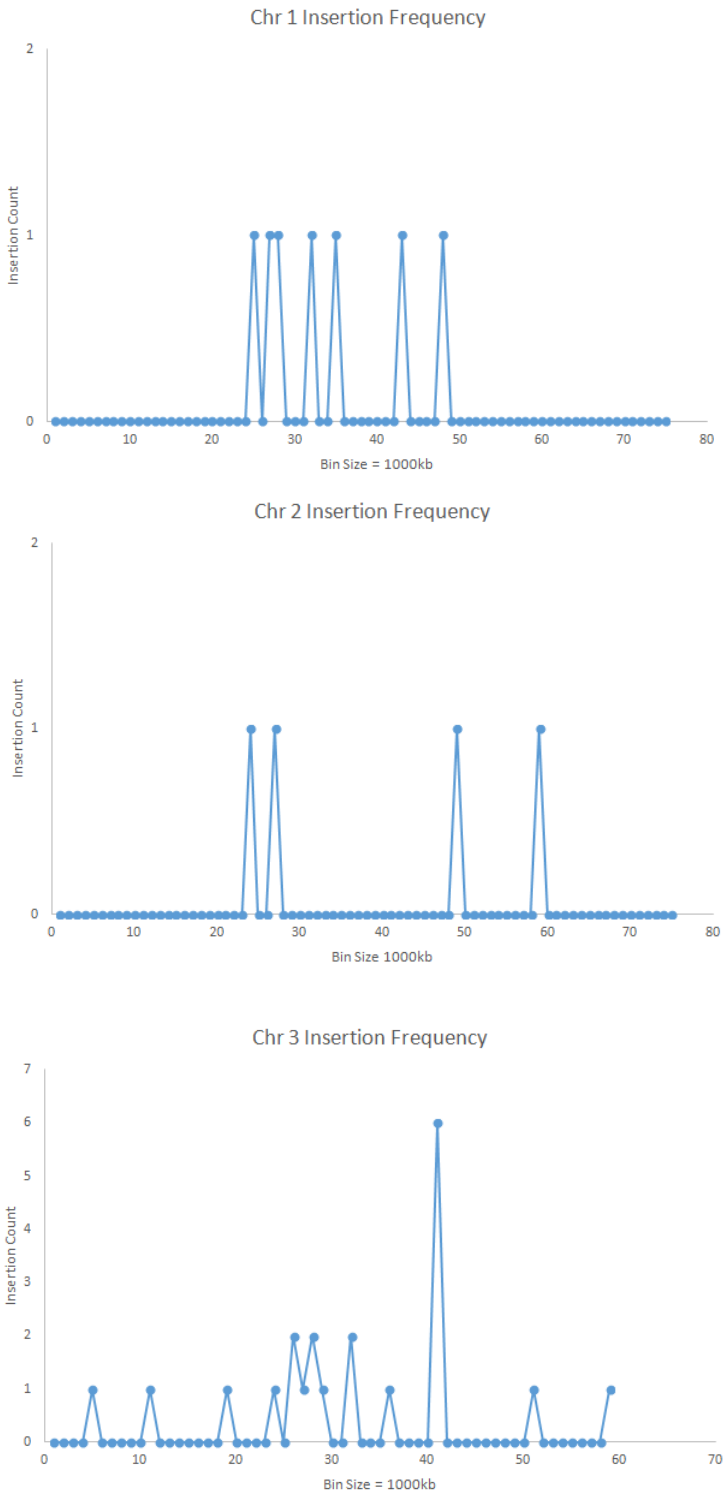

Figure 8. Preliminary analysis of common integration sites showing increased integration at the *β-actin* promoter on chromosome 3.

Genome wide mutagenesis in to screen for tumor suppressors and oncogenes

Transposon mutagenesis screens were targeted to neural progenitors using the described transgenic lines. Four separate screens are currently being monitored in the lab, including two screens using the *T2/OncZ* transposon line and two with the *T2/OncZgb1* transposon line. The first screen using the *Tg(krt5:SB11)* transgenic was with the *T2/OncZ* transposon line. A second screen using this cross was setup to examine changes in tumorigenesis in the pRB deficient background by injecting CRISPRs into single-cell stage embryos. However, analysis showed that there was no mutagenesis at the *rb1* locus. As a result, these fish were raised as a wild type background. These screens are summarized in Table 6. Of this second screen, one individual out of 10 showed tumor formation. Performing the excision assay in Figure 9, it was shown that transposon mobilization was detected in the tumor tissue, but not in the control muscle tissue. This indicates that the *Tg(krt5:SB11)* transgenic can induce tumorigenesis by mobilization of the transposon in specific tissues. Due to the limited numbers of fish involved in this screen, the significance of these results can not be established.

Table 6. The number of adult fish of each genotype from crosses between *Tg(krt5:SB11)* and *T2/OncZ*.

	Number of Adult Fish of Genotype			
Age	**+ ; +**	**krt5:SB11; +**	**+; T2OncZ**	**krt5:SB11; T2OncZ**
15 mo.	12	8	10	6
9 mo.	16	11	11	10

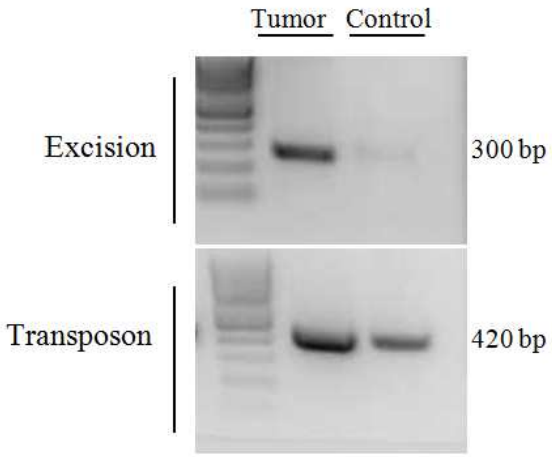

Figure 9. Excision assay showing transposon mobilization specifically in tumorous tissue.

A similar large scale screen is also being monitored using the *T2/OncZgb1* transposon line. Transposon mobilization was again confirmed in double transgenics using an excision assay of larvae. In addition, an excision assay performed on dissected brains of adult individuals showed transposon mobilization in the presence of the transposase in the *Tg(krt5:SB11)* transgenic. Lastly, the other screen in progress was done in the pRB deficient background. This was achieved by targeting *rb1* with TALENs to create mosaic mutants that are known to give raise to tumorigenesis. These screens are summarized in Table 7.

Table 7. The number of adult fish of each genotype from crosses between *Tg(krt5:SB11)* and *T2/OncZ* in wild type and pRB deficient backgrounds.

		Number of Adult Fish of Genotype			
***rb1* targeting?**	**Age**	**+ ; +**	**krt5:SB11; +**	**+; T2OncZgb1**	**krt5:SB11; T2OncZgb1**
No	11 mo.	102	105	127	149
Yes	8 mo.	28	27	35	24

Discussion

The purpose of this research was to utilize transposon somatic mutagenesis to gain a better understanding of brain tumorigenesis. Significant progress has been made in the development of transgenic zebrafish lines to drive transposon mutagenesis in a tissue-specific and inducible manner. The most important has been the *Tg(krt5:SB11)* transgenic line, which targets mutagenesis to neural progenitors. The functionality of this line has been demonstrated through excision assays showing this transposase source can mobilize the *T2/OncZ* and *T2/OncZgb1* transposons. Several mutagenesis screens are currently being monitored in the lab that uses this transgenic line. One of these screens resulted in tumor induction in the brain, with further analysis indicating transposon mobilization was achieved in the tumor, but not in control tissue. This is an important proof of concept that justifies the large scale screens. A screen is also being conducted in the pRB deficient mutant, a tumor susceptible background. To date, the number of fish exhibiting tumor induction in the mutagenesis screens has not been sufficient to establish statistical meaning. If a significant number of fish do not develop tumors, it may be necessary to make modifications to increase the rate of tumorigenesis. One option would be to repeat the screen in a pRB deficient background in combination with *p53* mutants.

One of the most important pieces of data still needed is the characterization of the expression pattern of the *Tg(krt5:SB11)* transgenic. Since the results of the *in situ* hybridization were inconclusive, the experiment needs to be repeated with a redesigned probe. A shorter probe consisting of several hundred base pairs would not have homology with the zebrafish genome, and would likely yield better results.

The inducible *Tg(ubi:floxed-SB11)* transposase source has also been generated and tested. It was shown that the transgenic line could increase transposon mobilization with *Cre* recombinase. However, there was a small amount of *SB11* expression that is ubiquitous, suggesting the 2xSV40 transcription terminator that was used is leaky. Refinements to this system could include replacing the 2xSV40 with a stronger terminator. This modification would be required to make this a viable method for an inducible transposon mutagenesis system.

The construction of the libraries for the common integration site analysis is completed, and the data is currently being analyzed. The results should show hotspots for transposon integration, as well as local hopping around the location of the transposon concatemer. This data will be beneficial when analyzing the mapping results from the mutagenesis screens.

In the course of this research, several molecular tools and transgenic zebrafish lines have been generated for transposon mutagenesis. It is known that cancer results from a complex interaction of several factors. These transposon mutagenesis screens will therefore be important to identify cooperative genes that affect tumorigenesis in the brain. Discovering the genetic mechanisms that drive proliferation of neural progenitors will provide insights into possible treatment of brain cancers, including PNETs. A more complete understanding of the genetic landscape of PNETs will lead to new treatments that will have a positive impact on human lives.

APPENDIX

QUANTIFYING DIFFERENTIAL GENE EXPRESSION IN AN OPTIC PATHWAY TUMOR LINE

In a previous study, the lab characterized the transgenic zebrafish line *Tg(flk1:RFP)is18*, which developed tumors in the optic pathway (Solin et al., 2014). These tumors exhibited properties of retinoblastoma and fibrous glioma. It was determined that the tumors were caused by the disruption of a long, intergenic non-coding RNA gene *lincRNAis18*. A part of the characterization of this mutant line was transcriptome analysis of pre-tumor and tumor retina using RNAseq. The purpose of my experiments in this study was to validate the RNAseq data for several key genes using quantitative PCR in order to measure differences in gene expression. These genes included *ajap1*, *ascl1a*, *atf3*, *bysl*, *hbegfa*, and *insm1a*.

Primers were designed for each gene of interest to amplify 100-110 bp amplicon. Table 8 provides the primer sequences. A standard curve for each primer pair was constructed to test for primer efficiency and determine any off target amplification. This was achieved through a 1 to 10000 dilution of control cDNA. A two-step qRT-PCR was performed using SYBR green master mix. Each reaction was performed in triplicate. Template consisted of 50 ng of control, pre-tumor, and tumor cDNA samples. cDNA was reverse transcribed from tissue samples using oligo-dT primers and the Invitrogen SuperScript III kit. qRT-PCR was performed on the LightCycler instrument. Fold change was calculated using *β-actin* as an endogenous control.

The differential gene expression measured using qRT-PCR was similar to the RNAseq results, as shown in Figure 10. In most cases, a greater increase in gene expression was observed. In particular, *ascl1a* and *hbegfa* showed a much higher fold change than was detected using RNAseq. The results show that the tumor suppressor *ajap1* was significantly decreased in the tumor. The increase expression of *atf3*, which encodes a cAMP-response element binding protein, suggests that cAMP signaling plays a role in tumor growth. The constant expression of *bysl* throughout the samples ruled out optic nerve injury as the cause for this upregulation of *atf3*. In addition, the tumors showed a significant increase in gene expression for several activators of *wnt*-β-cantenin signaling in glial progenitors, including *hbegfa*, *ascl1a*, and *insm1a*. This analysis supports the hypothesis that *Tg(flk1:RFP)is18* retinal tumors result from Müller glia that dedifferentiate to give raise to transformed neuroglial progenitors.

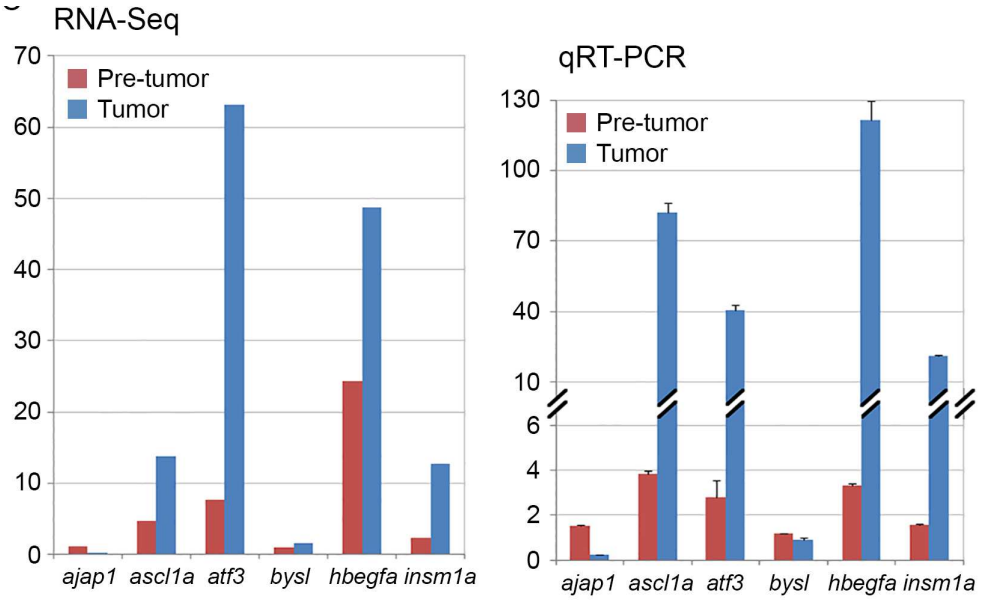

Figure 10. Differential gene expression measured by qRT-PCR and compared to RNAseq.

Table 8. List of primers used in the qRT-PCR analysis.

Primer Name	Sequence (5' to 3')
RT_ajap1_F	GATCATCACCATCACTGTGTCC
RT_ajap1_R	TTTCTCTGATGGCTGCTGTG
RT_ascl1a_F	ATGAACTCTATGGCCGGTTC
RT_ascl1a_R	TAAGTTTCCTTTTACGAACGCTC
RT_atf3_F	CAGCAGCAAAATGTCGGAAC
RT_atf3_R	GCTGCTTTTGGTTCTTCAGC
RT_bactin_F	CGAGCTGTCTTCCCATCCA
RT_bactin_R	TCACCAACGTAGCTGTCTTTCTG
RT_bysl_F	CTTCTACAATCTGGTGCTGCTG
RT_bysl_R	GAGCAGGATCCCTTTGAACC
RT_hbegfa_F	ATTTCTCTCTGCCGGTGAAG
RT_hbegfa_R	TGCTTGTGAAACCTGAGTGC
RT_insm1a_F	AGGTCTACCCGTGCAAATACTG
RT_insm1a_R	GGCATCTGAAGCAGGATCAC

REFERENCES

Collier, L. S., Carlson, C. M., Ravimohan, S., Dupuy, A. J., & Largaespada, D. A. (2005). Cancer gene discovery in solid tumours using transposon-based somatic mutagenesis in the mouse. *Nature*, *436*(7048), 272-276.

Dupuy, A. J., Akagi, K., Largaespada, D. A., Copeland, N. G., & Jenkins, N. A. (2005). Mammalian mutagenesis using a highly mobile somatic Sleeping Beauty transposon system. *Nature*, *436*(7048), 221-226.

Ivics, Z., & Izsvák, Z. (2015). Sleeping Beauty Transposition. *Microbiology spectrum*, *3*(2).

Ivics, Z., Hackett, P. B., Plasterk, R. H., & Izsvák, Z. (1997). Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell*, *91*(4), 501-510.

Izsvák, Z., & Ivics, Z. (2004). Sleeping beauty transposition: biology and applications for molecular therapy. *Molecular Therapy*, *9*(2), 147-156.

Janik, C. L., & Starr, T. K. (2013). Identification of Sleeping Beauty Transposon Insertions in Solid Tumors using Linker-mediated PCR. Journal of visualized experiments: JoVE, (72).

Ju, B., Chen, W., Orr, B. A., Spitsbergen, J. M., Jia, S., Eden, C. J., ... & Taylor, M. R. (2015). Oncogenic KRAS promotes malignant brain tumors in zebrafish. *Molecular cancer*, *14*(1), 1.

Ju, B., Chen, W., Spitsbergen, J. M., Lu, J., Vogel, P., Peters, J. L., ... & Jia, S. (2014). Activation of Sonic hedgehog signaling in neural progenitor cells promotes glioma development in the zebrafish optic pathway. *Oncogenesis*, *3*(3), e96.

Kawakami, K. (2007). Tol2: a versatile gene transfer vector in vertebrates. *Genome Biol*, *8*(Suppl 1), S7.

Lam, C. S., März, M., & Strähle, U. (2009). GFAP and nestin reporter lines reveal characteristics of neural progenitors in the adult zebrafish brain. *Developmental Dynamics*, *238*(2), 475-486.

Mann, M. B., Jenkins, N. A., Copeland, N. G., & Mann, K. M. (2014). Sleeping Beauty mutagenesis: exploiting forward genetic screens for cancer gene discovery. *Current opinion in genetics & development*, *24*, 16-22.

McClintock, B. (1950). The origin and behavior of mutable loci in maize. *Proceedings of the National Academy of Sciences*, *36*(6), 344-355.

McClintock, B. (1953). Mutation in maize. *Carnegie Inst. Wash. Year Book*, *52*, 227-237.

McGrail, M., Hatler, J. M., Kuang, X., Liao, H. K., Nannapaneni, K., Watt, K. E. N., ... & Dupuy, A. J. (2011). Somatic mutagenesis with a Sleeping Beauty transposon system leads to solid tumor formation in zebrafish. PLoS One, 6(4), e18826.

Moriarity, B. S., & Largaespada, D. A. (2015). Sleeping Beauty transposon insertional mutagenesis based mouse models for cancer gene discovery. *Current opinion in genetics & development*, *30*, 66-72.

Ni, J., Wangensteen, K. J., Nelsen, D., Balciunas, D., Skuster, K. J., Urban, M. D., & Ekker, S. C. (2016). Active recombinant Tol2 transposase for gene transfer and gene discovery applications. *Mobile DNA*, *7*(1), 1.

Plasterk, R. H., Izsvák, Z., & Ivics, Z. (1999). Resident aliens: the Tc1/mariner superfamily of transposable elements. *Trends in genetics*, *15*(8), 326-332.

Pray, L. A. (2008). Transposons: The jumping genes. *Nature Education*, *1*(1), 204.

Smoll, N. R. (2012). Relative survival of childhood and adult medulloblastomas and primitive neuroectodermal tumors (PNETs). Cancer, 118(5), 1313-1322.

Solin, S. L., Wang, Y., Mauldin, J., Schultz, L. E., Lincow, D. E., Brodskiy, P. A., ... & Hostetter, J. M. (2014). Molecular and cellular characterization of a zebrafish optic pathway tumor line implicates glia-derived progenitors in tumorigenesis. *PloS one*, *9*(12), e114888.

Solin, S. L., Shive, H. R., Woolard, K. D., Essner, J. J., & McGrail, M. (2015). Rapid tumor induction in zebrafish by TALEN-mediated somatic inactivation of the retinoblastoma1 tumor suppressor rb1. Scientific reports, 5.

Walisko, O., Schorn, A., Rolfs, F., Devaraj, A., Miskey, C., Izsvák, Z., & Ivics, Z. (2008). Transcriptional activities of the Sleeping Beauty transposon and shielding its genetic cargo with insulators. *Molecular Therapy*, *16*(2), 359-369.