Insertional vs Targeted Mutagenesis in the Development of Zebrafish as an *In Vivo* Model for Cardiomyopathy

by

Matiyo Ojehomon

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements for the degree of
Master of Science
in
Molecular and Cellular Biology

Guelph, Ontario, Canada
© Matiyo Ojehomon, April 2020
ABSTRACT

INSERTIONAL VS TARGETED MUTAGENESIS IN THE DEVELOPMENT OF ZEBRAFISH AS AN IN VIVO MODEL FOR CARDIOMYOPATHY

Matiyo Ojehomon
University of Guelph, 2020

Advisor(s):
Dr. John Dawson
Dr. Terry Van Raay

Heart failure is a global economic burden and can be caused by cardiomyopathy, a disease of the myocardium. Mutations in genes encoding sarcomere proteins have been known to cause cardiomyopathy. One of these sarcomere proteins is α-cardiac actin (ACTC), which is needed for proper contraction of the heart. To better understand how mutations in the ACTC protein cause cardiomyopathy, an in vivo model like zebrafish can be used to understand the molecular mechanism that occurs. Using transposons (insertional) or Clustered Regularly Interspaced Palindromic Repeats (CRISPRs) (targeted), zebrafish can be engineered to carry these mutations and the impact can be studied. The transposon system proved to be unstable as the inserted transposon underwent transcriptional repression. To use the CRISPR system, identifying which zfactc gene to target is necessary. Performing WMISH and RT-qPCR, zfacta1b seems to be the best candidate for targeted mutagenesis due to its early expression in the heart.
ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. John Dawson for everything; providing me the opportunity to take on this project, supporting, mentoring, challenging and allowing me to contribute not just in the lab but also within the cardiovascular field. I am forever grateful for the assistance and support from Dr. Terry Van Raay, in his expertise in the area of zebrafish research, allowing me to work in his lab and always being available to render assistance and answer questions during challenging situations and regarding new experiments.

Thank you to all the past and current members of the Dawson lab especially Maria Anillo, who helped edit my thesis and constantly reminding me that I have a social life, Evan Despond (editor), Haidun Liu, Zeeshan Shaikh, Michael Jones and Love Sandhu, for the laughs, Clea Heathfield and Suchandrima Dutta, for the memories and help in collecting data, and Amy Woosely and Bersabhe Solomon, for their help with crossing the fish. A special thank you to Dr. Navneet Sidhu for being a sounding board and Grace (Zi) Teng for forever being graceful. Additionally, I would like to thank past and present members of the Van Raay Lab especially Victoria Rea, for helping me with my graphics and showing me how to use Illustrator as well as Ian Bell, Haider Khan and Dan Pasula. A special thanks to Yuning Su, for her help in maintaining the zebrafish facility. I would also like to thank past and present members of office 2246 especially Dr. Maria Rosales Gerpes for being a friend and constantly cheering me on and Andrea Brumwell for providing me with reagents and equipment to help improve my experiments.

Furthermore, I would like to acknowledge Dr. Andreas Heyland for granting me access to his lab and to use the in situ robot, as well as, Jacquie Matsumoto, for showing me how to
properly dissect the zebrafish, Sarah Alderman, on her expertise in performing and analyzing RT-qPCR and allowing me to work on her bench, Kathy Jacyniak for helping me with histology and Michaela Strueder Kypke for helping me with microscopy. I would also like to thank Matthew Cornish and Mike Davies for their continued assistance in the Hagen Aqualab.

Finally, I would like to thank my mum, my sister, Arinola Ojehomon and Chay Governor for always being in my business and wanting the best for me.
Table of Contents

Abstract............................................................................................................................... ii

Acknowledgements ........................................................................................................ iii

List of Tables ....................................................................................................................... ix

List of Figures .................................................................................................................... x

List of AbbreviationS ......................................................................................................... xi

1 Introduction ....................................................................................................................... 1

1.1 Cardiomyopathy as a Cause of Heart Failure ................................................................. 1

1.2 Characterization of Hypertrophic and Dilated Cardiomyopathy ................................. 2

1.3 Mutations in Cardiac Actin Lead to Cardiomyopathy .................................................. 3

1.4 Actin is a Fundamental Protein for Eukaryote Survival .............................................. 5

1.5 In Vivo and In Vitro Studies Aid in Developing Therapeutic Drugs .............................. 6

1.6 Zebrafish as an In Vivo Model for Cardiovascular Research ......................................... 8

1.7 Comparison of the Heart and Cardiovascular System in Zebrafish and Humans .......... 8

1.8 Heart Development in Zebrafish .................................................................................. 11

1.9 Research Goal and Question ....................................................................................... 14

1.10 Tol2 Transposon System as an Insertional Mutagenesis Tool ..................................... 16

1.11 CRISPR as a Genome Editing Tool for Targeted Mutagenesis in Zebrafish ............... 17
1.11.1 Actin Isoforms in Zebrafish ................................................................. 18

2 Materials and Methods ........................................................................... 21

2.1 Reagents Used ....................................................................................... 21

2.2 Animal Husbandry and Egg Collection .................................................... 21

2.3 Needle Set Up ........................................................................................ 21

2.4 Embryo Injection .................................................................................... 21

2.5 Transposon Plasmid .............................................................................. 22

2.6 Subcloning ACTC Mutants into the Transposon Vector ............................ 22

2.7 Transformation into E. coli DH5α .............................................................. 23

2.8 Screening for Positive Colonies: Successful Ligation and Proper Orientation by Colony PCR 23

2.9 Generating Transposase mRNA ............................................................... 24

2.10 Germ Line Transmission ....................................................................... 24

2.11 Embryo Heart Rate Analysis ................................................................. 25

2.12 Phenotype Trajectories ....................................................................... 25

2.13 Extraction of Genomic DNA from Zebrafish Embryos ............................ 26

2.14 Thermal Asymmetric Interlace Polymerase Chain Reaction (TAIL-PCR) ........ 26

2.15 Designing of CRISPR Targets ............................................................... 26

2.16 gRNA Template Assembly and Generation of sgRNA .......................... 27
2.17 High Resolution Melting Curve Analysis (HRM) .........................................................27

2.18 Sequencing, Tide and Indigo Gear analysis .................................................................28

2.19 Isolation of Hearts from Zebrafish Embryos ...............................................................28

2.20 Extraction of RNA, cDNA Synthesis and Reverse Transcriptase Quantitative PCR (RT-qPCR)29

2.21 Whole Mount In Situ Hybridization (WMISH) .............................................................30

2.22 Statistical Analysis ........................................................................................................31

3 Results ..................................................................................................................................32

3.1 Aim 1: Express Human ACTC cDNA Variants in the Zebrafish genome using Tol2 Transposons

3.1.1 Generation of Transgenic Zebrafish Expressing a Human ACTC Variant using Transposons........32

3.1.2 Functional Analysis of Transgenic Zebrafish ....................................................................35

3.1.3 Heart Rate Analysis .......................................................................................................37

3.1.4 TAIL-PCR ....................................................................................................................37

3.2 Aim 2: Determining the best zfactc gene to target for gene editing ........................................38

3.2.1 WMISH Displays Isoactin Switching in a Chamber-Specific Manner .................................38

3.2.2 Expression Profile of the zfACTC Isoforms Using RT-qPCR ...........................................40

3.3 Mild Phenotype Observed in zfactc1a CRISPR Injected Embryos .......................................42

4 Discussion .............................................................................................................................46

4.1 Heart Specific Expression of Human ACTC Variants Using Tol2 Transposons is not Stable ...46
4.2 Severity of the Cardiac Phenotype Observed is due to the Integration of the Transposon in the Zebrafish Genome ................................................................. 47

4.3 In Situ Hybridization Supports a Cardiac Gene Switch Model ........................................ 49

4.4 Mild Phenotype Observed in CRISPR Injected Embryos May be due to Genetic Compensation ........................................................................................................ 56

5 Conclusion ......................................................................................................................... 58

6 Future Directions ............................................................................................................... 59

7 References .......................................................................................................................... 61
LIST OF TABLES

Table 1: N-terminal sequences of human and zebrafish (zf) actin isoforms. .......................... 6
Table 2 Primers used in this study and not reported in previous literature................................. 31
Table 3: Number of founders discovered per transposon injection. ........................................ 33
Table 4: Phenotype classification of TREE. ............................................................................... 35
Table 5: Analysis of Tol2 integration sites and the possible gene and location of integration .... 38
LIST OF FIGURES

Figure 1 Mutations in ACTC gene cause cardiomyopathy............................................................. 4
Figure 2: Heart development of zebrafish from 24 hpf.. ............................................................... 13
Figure 3: Resulting transgenesis using Tol2 transposon............................................................... 34
Figure 4: Effect of expressing a transgene in the zebrafish genome. .............................................. 36
Figure 5: Localization of zfactc genes using WMISH...................................................................... 39
Figure 6: Relative Expression levels of zfactc genes in the heart using RT-qPCR ...................... 41
Figure 7: Heart rate data in beats per min of zfactc1a CRISPR injected embryos....................... 44
Figure 8: Phenotype to genotype of zfactc1a-CRISPR injected embryo........................................ 45
Figure 9: Schematic illustration of the proposed expression of zfactc genes in the developing heart of an embryo. ................................................................................................................... 54
Figure 10: Schematic illustration showing the results obtained from the RT-qPCR compared to the results obtained from WMISH.......................................................................................................... 55
LIST OF ABBREVIATIONS

ACTA1 - α–skeletal actin
ACTA2 - α-smooth actin
ACTB - β–cytoplasmic actin
ACTC1 - α–cardiac actin
ACTG1 – g–cytoplasmic actin
ACTG2 – γ–smooth actin
bp – Base pair
bpm – Beats per minute
Cmlc – cardiac myosin light chain
CRISPR – Clustered Regularly Interspaced Short Palindromic Repeats
crRNA- CRISPR- RNA
DCM – Dilated Cardiomyopathy
dpf – days post fertilization
GFP – Green fluorescent Protein
gRNA – guide RNA
HCM – Hypertrophic Cardiomyopathy
HDR – Homologous Directed Repair
hpf – hours post fertilization
HRM – High resolution Melting
ITR – Inverted Terminal Repeat
mRNA – messenger RNA
NHEJ – Non homologous end joining

PCR – Polymerase Chain Reaction

SCID – Severe combined immunodeficiency

sgRNA – Single guide RNA

SNV – Single Nucleotide variations

RT-qPCR – Reverse Transcriptase quantitative PCR

TAIL-PCR – Thermal Asymmetric Interlaced PCR

tracrRNA – Trans activating crRNA

TREE – Transposition reporter expressing embryos

Tü – Tübingen

UTR – Untranslated region

WMISH – Whole mount in situ hybridization

WT – Wild type
1 Introduction

1.1 Cardiomyopathy as a Cause of Heart Failure

Heart failure is the outcome that occurs after the heart becomes damaged and its pumping action is too weak to meet the needs of the body. According to the Heart and Stroke Foundation, it is estimated that 600,000 Canadians are living with heart failure; this number is expected to increase due to an increase in the elderly population and increase in conditions like hypertension and diabetes that can generate the development of heart failure (Heart failure | Heart and Stroke Foundation; Lesyuk et al., 2018). The current cost of heart failure on the Canadian economy is more than 22 billion dollars yearly, thereby putting a strain on the Canadian healthcare system (Smith, 2009).

Cardiomyopathy is a disease of the heart muscle, myocardium, that can cause heart failure. The crucial function of the myocardium is to execute contractile activity produced from the interactions of the sarcomeres and its transmission to the extracellular matrix (Liew and Dzau, 2004; Morita et al., 2005). When sarcomere interactions and extracellular matrix activities are reduced, remodeling of the cardiomyocytes occurs to compensate for the changes (Morita et al., 2005; Seidman and Seidman, 2001). These changes can lead to thickening of the myocardium (hypertrophic cardiomyopathy) or thinning of the myocardium (dilated cardiomyopathy). These two compensatory effects of the heart can eventually result in heart failure.
1.2 Characterization of Hypertrophic and Dilated Cardiomyopathy

Hypertrophic cardiomyopathy (HCM) is characterized by thickening of the ventricular walls, thereby decreasing ventricular chamber size (Bakkers, 2011) (Figure 1A). This decrease leads to the inability of the heart to hold a sufficient amount of blood to support the needs of the body, forcing the heart to work harder (Spudich, 2014). The diastolic function of the heart is also impaired as the muscles become stiff due to the thickening of the ventricular walls (Morimoto, 2008). Increased protein synthesis in cardiomyocytes leading to cardiomyocyte enlargement and disarray along with an increase in myofibrils are assumed to lead to the enlargement of the left ventricle (Morimoto, 2008; Nonaka and Morimoto, 2014). It has been estimated that HCM affects 1 in 200 individuals and it is the most common cause of cardiac arrest in young people, particularly trained athletes (Wasfy et al., 2016; Nonaka and Morimoto, 2014). Current therapies for HCM patients improve diastolic dysfunction by indirectly reducing systolic function using β-blockers, Ca^{2+} or Na^{+} channel blockers (Nonaka and Morimoto, 2014).

Dilated cardiomyopathy (DCM) is characterized by the enlargement of one or both ventricles (Bakkers, 2011), where the ventricular walls are thin and stretched, leading to a larger chamber size and impaired systolic performance (Spudich, 2014; Morimoto, 2008) (Figure 1A). DCM accounts for 60% of all cardiomyopathy cases (Kärkkäinen and Peuhkurinen, 2007); it is age dependent and a heart transplant is the only treatment for this type of cardiomyopathy (Olson et al., 1998). This condition is usually accompanied with cardiomyocyte enlargement and interstitial fibrosis but not cardiomyocyte disarray (Seidman and Seidman, 2001; Nonaka and Morimoto, 2014). Patients with DCM present symptoms such as arrhythmia and sudden death that could be due to severe fibrosis and remodeling of the myocardial cells (Nonaka and
Morimoto, 2014). Current therapy for DCM patients is to use cardioprotective drugs like β-adrenergic receptor blockers, vasodilators and diuretics, which help to inhibit the progression of myocardial remodeling and systolic dysfunction (Nonaka and Morimoto, 2014).

1.3 Mutations in Cardiac Actin Lead to Cardiomyopathy

Mutations in genes encoding sarcomeric proteins have been recognized as the major cause of cardiomyopathy (Spudich, 2014; Morimoto, 2008; Jacoby and McKenna, 2012). Over 600 rare mutations have been found in genes encoding sarcomeric proteins that can cause cardiomyopathy (Jacoby and McKenna, 2012). The amount of mutations found in one gene generally correlates with the size of the gene. Some large sarcomere proteins that can be affected are myosin heavy chain, myosin binding protein C and lamin A/C which accounts for 37%, 33% and 28% of mutations found, respectively (Morimoto, 2008). α-cardiac actin (ACTC) is a smaller thin filament sarcomeric protein that accounts for 1.5% of cardiomyopathy cases (Morimoto, 2008). Due to the highly conserved nature of the actin protein, cardiomyopathy-linked ACTC mutations are rare but the phenotypes associated with these mutations can be severe (Monserrat et al., 2007). For example, patients who carry the E99K ACTC mutation not only had symptoms associated with HCM such as septal defects and angina but also cardiac trabeculation (Monserrat et al., 2007; Arad et al., 2005).
Figure 1: Mutations in ACTC gene cause cardiomyopathy. A) Dilated cardiomyopathy (DCM) is characterized with thinner muscular wall leading to an increased volume size and impaired systolic function. Hypertrophic cardiomyopathy (HCM) is characterized with thicker ventricular walls with a decreased volume space and impaired diastolic function (Mcnally et al., 2013). B) Ribbon structure of actin molecule showing the location of the 18 known ACTC gene mutations causing cardiomyopathy. Red indicates the myosin class, green represents the tropomyosin class and, blue represents the tropomyosin and myosin class (Despond and Dawson, 2018). Image generated with PyMOL using PDB 5N0J.
1.4 Actin is a Fundamental Protein for Eukaryote Survival

Actin is the most abundant protein expressed in most eukaryotic cells and aids in many cellular functions (Dominguez and Holmes, 2011). There are six known actin isoforms found in higher vertebrates: α-skeletal (ACTA1) and α-cardiac actin (ACTC1) classified under striated muscle, α-smooth (ACTA2) and γ-smooth actin (ACTG2) classified under smooth muscle and lastly, β-cytoplasmic (ACTB) and γ-cytoplasmic (ACTG1) actin classified under cytoplasmic isoforms. These isoforms have different functions such as aiding in the contractile apparatus of the cell, maintenance of the cytoskeleton and cell motility (Rubenstein, 1990; Dominguez and Holmes, 2011). These isoforms have very similar amino acid sequences, with each sharing at least 93% sequence identity to each other (Perrin and Ervasti, 2010). Variations of the amino acids between actin isoforms are seen in the N-terminus, and studies have shown that these variations influence how actin interacts with other proteins within the cell to function properly (Table 1) (Rubenstein, 1990; Dominguez and Holmes, 2011). Therefore, any slight differences from the conserved amino acid sequence needed in the function of an actin protein could be detrimental to the health of an organism.

Although there are six main isoforms, ACTC is primarily expressed in the heart. To date, there are 18 known ACTC mutations causing cardiomyopathy, 13 associated with HCM and 5 associated with DCM (Kaski et al., 2009; Lakdawala et al., 2012; Mogensen et al., 2004, 1999; Morita et al., 2008; Olivotto et al., 2008; Olson et al., 2000, 1998; Van Driest et al., 2003). These ACTC mutations have been grouped into 3 different class; myosin (M) class, tropomyosin (TP) class and tropomyosin and myosin (TM) class (Despond and Dawson, 2018) (Figure 1B). The functional consequences and exact details of the signalling pathways that bring about the
remodeling of cardiomyocytes which leads to cardiomyopathy are still unknown. Individuals who are at risk can adopt various lifestyle changes or take medication to manage symptoms, but the latter may cause side effects. Therefore, the main goal of this research is to examine how ACTC mutations bring about the disease state seen in humans. Knowing this information, researchers can develop better therapeutic drugs that target the mechanisms bringing about cardiomyopathy instead of treating the symptoms.

**Table 1: N-terminal sequences of human and zebrafish (zf) actin isoforms.** The N-termini of the human actin isoforms differ after the MC amino acid for the muscle isoforms and M for the non-muscle isoforms. N-termini of the zf actin isoforms are the same but difference in the sequence differ in other sections of the protein.

<table>
<thead>
<tr>
<th>Actin Isoform</th>
<th>N-terminal Sequence</th>
<th>UniProt Database Enteries</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTA</td>
<td>MDDEBTTAL</td>
<td>P68133</td>
</tr>
<tr>
<td>ACTC</td>
<td>MDDEBTTAL</td>
<td>P68032</td>
</tr>
<tr>
<td>ACTA2</td>
<td>MCEDSTAL</td>
<td>P62736</td>
</tr>
<tr>
<td>ACTG2</td>
<td>MC-EETTAL</td>
<td>P63267</td>
</tr>
<tr>
<td>ACTB1</td>
<td>M--DDIATAL</td>
<td>P60709</td>
</tr>
<tr>
<td>ACTG1</td>
<td>M--EETIATAL</td>
<td>P63261</td>
</tr>
<tr>
<td>ZFACTC1a</td>
<td>MDDEBTTAL</td>
<td>Q6IQR3</td>
</tr>
<tr>
<td>ZFACTA1b</td>
<td>MDDEBTTAL</td>
<td>Q6XNL8</td>
</tr>
<tr>
<td>ZFACTC1C</td>
<td>MDDEBTTAL</td>
<td>Q6IQL9</td>
</tr>
</tbody>
</table>

**1.5 In Vivo and In Vitro Studies Aid in Developing Therapeutic Drugs**

Variant proteins resulting from ACTC mutations have been studied extensively *in vitro*, revealing the functional consequences that these mutations cause. It has been demonstrated that some ACTC variants exhibit changes in protein stability and as a result, the variant protein might have a different interaction with binding proteins like myosin and tropomyosin (Mundia et al., 2012). Mundia et al. (2012) discovered that some ACTC variants had intrinsic stability defects where the melting temperature was significantly lower than the wildtype (WT). In addition, the
ACTC variants also exhibited short aggregated filaments rather than regular, long filaments (Mundia et al., 2012). Moreover, studies have shown that some ACTC variants, especially E99K, have altered actomyosin interactions (Dahari and Dawson, 2015; Liu et al., 2018). Compared to the WT ACTC, unregulated E99K had a significantly higher activity with myosin (Liu et al., 2018).

In vivo studies are used to understand the physiological consequences of mutations. These studies connect the functional changes from the in vitro studies to the physiological state of the heart. Song et al. (2011) created a transgenic mouse that had 50% E99K variant actin in the heart. They compared the physical state of the mutant mouse heart to patients carrying the same mutation and the same symptoms were observed such as apical hypertrophy and ventricular trabeculations (Song et al., 2011). In addition, the E99K ACTC protein was purified from the transgenic mouse and an in vitro motility assay was conducted. It was observed that the E99K ACTC variant showed a higher Ca\(^{2+}\) sensitivity compared to WT (Song et al., 2011). Many animal models have also been used to discover the pathogenesis of cardiomyopathies caused by mutations in other genes. For example, transgenic mouse models having a mutation in tropomyosin 1 had normal ventricular function despite having cardiomyocytes that were enlarged with fibrosis and cardiac myofilaments that were more sensitive to Ca\(^{2+}\) (Muthuchamy et al., 1999).

Crossing transgenic mice carrying HCM-linked sarcomeric mutations with transgenic mice carrying DCM-linked sarcomeric mutation lead to a decrease in the Ca\(^{2+}\) sensitivity of myofilaments, which prevented worsening of the cardiac tissues (Nonaka and Morimoto, 2014). Results like this have led to the development of drugs such as, W7 – N-(6-Aminohexyl)-5-
chloro-1-naphthalenesulfonamide hydrochloride, that decreases Ca\textsuperscript{2+} sensitivity in HCM myofilaments to bringing it to basal level. Having an animal model to test these drugs would be a critical advancement in cardiomyopathy therapy in patients. In addition, developing animal models that will bring more insight into how cardiomyopathy brings about the disease state is very critical to produce more therapeutic drugs.

1.6 **Zebrafish as an In Vivo Model for Cardiovascular Research**

Different ACTC variants have been studied using various systems, with each system offering its own advantages and limitations. To date, there is no established or efficient *in vivo* model to study *ACTC* mutants. Zebrafish (*Danio rerio*) is a widely used non-mammalian vertebrate model for studying human disease (Genge et al., 2016). Due to their optical transparency at the larval stage, zebrafish are a useful model for studying cardiovascular disease (Genge et al., 2016). The heart can be observed directly with a light microscope, which is a non-invasive technique (Bakkers, 2011). With their transparency and small size, zebrafish do not need a fully functional cardiovascular system to survive early in life, as they can acquire oxygen through passive diffusion from their environment (Bakkers, 2011; Poon and Brand, 2013). This characteristic allows embryos with serious cardiovascular defects such as a non-beating heart or intermittent contraction to continue living, while assessments of their cardiovascular system can be performed (Poon and Brand, 2013).

1.7 **Comparison of the Heart and Cardiovascular System in Zebrafish and Humans**

The heart and cardiovascular system of zebrafish is similar to humans in many respects. The human heart possesses four chambers and undergoes septation, while the zebrafish heart only has two chambers without septation (Chico et al., 2008). In the zebrafish, the deoxygenated
blood from the body reaches the atrium, followed by the ventricle and is then pumped into the gills where it is oxygenated before going to the rest of the body (Rayani et al., 2018). In humans, deoxygenated blood reaches the right atrium, followed by the right ventricle, which then pumps the blood to the lungs where it is oxygenated. The oxygenated blood then returns to the left atrium of the heart, followed by the left ventricle, which produces enough pressure to reach the entire body.

The pressure generated on the right side of the heart is less than the left side, which is beneficial to the lungs since it has a thinner muscular wall. In contrast, the zebrafish heart has to generate enough pressure for both the gills and the body. When the blood leaves the ventricle of the zebrafish heart, the elastic nature of the bulbous arteriosus reduces the pressure from the ventricle so as not to damage the gills (Hu et al., 2001). The pressure of the blood is further reduced by 40% from the resistance of the branchial arteries (Hu et al., 2001). The blood pressure that remains is what drives the blood to the rest of the body.

The two-chambered heart of the zebrafish has an average heart rate of about 150 beats per minute (bpm) in 72-hours post-fertilization (hpf) embryos and adults and a myocardium surrounded by an epicardium and endocardium (Yu et al., 2012; Hu et al., 2001; Nemtsas et al., 2009). Histological studies show that the ventricle of adult zebrafish hearts is composed of more trabeculated myocardium (spongy) and less compact myocardium, while the human heart has a more compact myocardium than a trabeculated myocardium (Sedmera et al., 2003; Yu et al., 2012; Hu et al., 2001). The atrial and ventricular myocytes of the adult zebrafish, although shorter and not as wide compared to human cardiomyocytes, show clear striation (Nemtsas et al., 2009). The human cardiomyocyte has an extensive transverse (T) tubular system, while the adult
zebrafish ventricular myocyte lacks this system (Nemtsas et al., 2009). The T-tubules in the human heart allow impulses to penetrate deeper into the cell of the muscle fibers activating the sarcoplasmic reticulum, which in turn releases sufficient Ca\(^{2+}\) to the sarcoplasm (Nemtsas et al., 2009). Since the zebrafish heart has a small diameter, the heart may not need T-tubules to permit the Ca\(^{2+}\) impulses to reach the center of the muscle fibers (Nemtsas et al., 2009). The missing structure does not imply that the function is missing from the zebrafish.

Zebrfish myocytes have a cardiac action potential similar to humans. The cardiac action potential informs us about the change in the electrical potential that occurs when an impulse passes the cell membrane of the heart cells while contracting and relaxing (Rosen and Pham, 2004). The zebrafish cardiac action potential has a rapid upstroke, which is dominated by Na\(^{+}\) channels, a long-lasting plateau phase contributed by the L-type Ca\(^{2+}\) channels, and a rapid final repolarization phase from the rapidly activating K\(^{+}\) channels (Nemtsas et al., 2009). Due to the long-lasting plateau phase of the zebrafish myocyte, the ventricular zebrafish action potential is more similar in shape and duration to the human ventricular action potential. In contrast, the atria of the zebrafish adult and embryonic hearts has a shorter action potential duration (Nemtsas et al., 2009). The zebrafish atrial myocytes are also sensitive to carbachol, which indicates that the atrial cells contain acetylcholine-activated inward rectifier K\(^{+}\) current that is similar to humans (Nemtsas et al., 2009).

Despite the similarities in the current channels and the shape of the action potential, zebrafish display a strong T-type Ca\(^{2+}\) current, that also contributes to the plateau phase, in both the atrial and ventricular cardiomyocytes. T-type Ca\(^{2+}\) current has only been seen in developing hearts of mammals or when the mammals’ cardiomyocytes are under pathophysiological
conditions (Nemtsas et al., 2009). This difference suggests that the adult zebrafish cardiomyocytes display a phenotype more similar to immature mammals. Lastly, the zebrafish heart displays similar contractile properties to humans as similar signals were obtained from the electrical cardiogram (Yu et al., 2012). The readings from the electrical cardiogram show that the PR interval, which informs us of the impulse conduction from the atria to the ventricle, is shorter in zebrafish compared to humans. This difference might be due to the smaller heart size and the higher mean heart rate of the zebrafish heart (Yu et al., 2012).

1.8 Heart Development in Zebrafish

Heart development in zebrafish is very similar to humans. The heart is composed of 3 cell types; ventricular cardiomyocytes, atrial cardiomyocytes and endocardial cells, which can be traced back to cardiac progenitor cells (Brown et al., 2016; Bakkers, 2011). During heart development in zebrafish, there are two heart fields, which is where the heart is formed. The first heart field refers to the first wave of cardiac progenitor cells that differentiate to form the initial heart tube and expresses muscle specific proteins (de Pater et al., 2009). It is proposed to start at 5 hpf and ends at 24 hpf. At 24 hpf, the disk is elongated into a linear heart tube and separate expression profiles for atrial and ventricular cardiomyocytes and endocardial cells can be seen (Figure 2A) (Brown et al., 2016). The second heart field starts between 28 to 36 hpf and refers to the second wave of differentiating cardiac progenitor cells that are added on the arterial and venous poles to extend the linear heart tube (Figure 2B). The linear heart tube moves leftward, and the heart begins to loop (Brown et al., 2016; Bakkers, 2011). At 48 hpf, this looped heart is located in the pericardial cavity and the atrium and ventricle are clearly distinct with an atrioventricular canal between them (Figure 2C) (Brown et al., 2016). At this stage, the heart is
fully formed but not yet matured and additional structures are being created as the organism develops. The bulbus arteriosus, similar to the mammalian conotruncus, is fully developed at 3 to 5 days post fertilization (dpf) (Brown et al., 2016) (Figure 2D).

The cardiac trabeculae which helps to increase myocardial surface area for blood oxygenation forms at 72 hpf as well as the epicardium (Brown et al., 2016). At 105 hpf, the endocardial cushions thicken and mature into valve leaflets (Brown et al., 2016; Bakkers, 2011). The heart continues to mature through the juvenile and early adult life stage. During early juvenile stage, the ventricle remodels to form a more rectangular morphology while in late juvenile and early adult stage, the ventricle becomes more rounded and the coronary arteries that feed the ventricle are formed (Figure 2 E-F) (Brown et al., 2016).
Figure 2: Heart development of zebrafish from 24 hpf. A) At 24 hpf the cardiac progenitor cells elongates into a linear heart tube forming separate expression profiles of the ventricle, atrium and endocardial cell. B) Between 28 to 36 hpf, the heart begins to move leftward initiating looping. C) At 48 hpf, looping is complete and atrium and ventricle are located in the pericardial cavity. D) At 3 – 5 dpf, the bulbus arteriosus is fully formed as the outflow tract of the heart E) At the juvenile stage, the ventricle has more of a rectangular morphology and the atrium lies dorsally to the ventricle while at the adult stage (F), the ventricle has a more rounded shape. Image adapted from Brown et al., (2016).
1.9 Research Goal and Question

Given the advantages of zebrafish for the study of the heart, can zebrafish be an efficient model for studying ACTC mutations causing cardiomyopathy? Animal models, even with their limitations, have been excellent tools in understanding human physiology. A humanized animal model is a model that carries functional human genes, cells, tissues and/or organs. Such models can be used to understand human disease physiology and help to develop therapeutics to treat the disease. For example, researchers colonized fetal sheep bone marrow with human hematopoietic stem cells, which retained their ability to differentiate into multiple cell lineages for years (Walsh et al., 2017).

Humanized mice have begun to fill the gap of expensive animal models (Walsh et al., 2017). For more than 25 years, researchers have developed humanized mouse models able to produce human T and B cells (Macchiarini et al., 2005; Walsh et al., 2017). The key breakthrough was when a Severe Combined Immunodeficiency (SCID) mutated mouse that lacked both T and B cells was created into which human hematolymphoid cells were transplanted (Walsh et al., 2017). Although the mice produced a number of human cells, the engraftment was very low due to the rejection of the cells by the mouse innate immune system (Walsh et al., 2017). This challenge was overcome by breeding the SCID mutant mice with strains of inbred mice with known defects in innate immunity (Macchiarini et al., 2005; Walsh et al., 2017). This humanized mouse model has led to better engraftment of human cells and has facilitated studies in infectious disease and hematopoietic stem cell development (Walsh et al., 2017). The humanized mouse model has also led to the development of vaccines and other beneficial treatment for humans (Macchiarini et al., 2005; Walsh et al., 2017).
Researchers have humanized the zebrafish liver, expressing human CYP3A4 enzyme, the major detoxifying CYP enzyme (Poon et al., 2017). Using the Tol2 transposon system, the human CYBP3A4 cDNA tagged with mCherry fluorescence was expressed, driven by a highly active zebrafish liver-specific promoter, liver fatty acid binding protein (LFABP) (Poon et al., 2017). The transgenic line significantly increased the metabolism of CYP3A4 substrate drugs producing metabolites similar to humans (Poon et al., 2017).

Several research groups have used zebrafish to study the effect of mutations that causes cardiomyopathy. Sehnert et al., (2002) identified a gene known as the silent heart gene (sih). Embryos with the sih gene developed normally but exhibited a heart with poor contraction (Sehnert et al., 2002). The sih gene was found to encode the sarcomere protein troponin T. Mutations in human cardiac troponin T that are known to cause cardiomyopathy are responsible for cases of sudden cardiac death in young athletes (Morimoto, 2008). Using an inexpensive in vivo model to study mutations causing cardiomyopathy would help examine how mutations in the human ACTC could result in HCM and DCM in humans.

There are different tools that can be used to edit the genome of the zebrafish. I tested using transposons (insertional mutagenesis) and CRISPRs (targeted mutagenesis) to edit the genome of the zebrafish. Therefore my thesis has 2 aims:

**Aim 1** – Express human ACTC cDNA variants in the zebrafish genome using Tol2 transposon

**Aim 2** – Determine the best zfactc gene to target for CRISPR/Homology Directed Repair (HDR) gene editing.

Before describing the methods used in this work, it is important to understand the mechanisms of the different gene editing tools used.
1.10 **Tol2 Transposon System as an Insertional Mutagenesis Tool**

The Tol2 transposon system is an efficient chromosomal engineering vehicle for making transgenic zebrafish (Clark et al., 2011). The Tol2 system brings about more than 80% transgenesis to the next generation; better than the 5% transgenesis obtained from microinjecting naked DNA into the cytoplasm of the zebrafish embryo (Suster et al., 2009; Kawakami, 2007). The Tol2 system consists of two components: a transposon donor plasmid that possesses a Tol2 construct with a promoter and the gene of interest, and a synthetic transposase mRNA to produce the transposase protein catalyzing the movement of the Tol2 construct from its carrier plasmid to the host genome (Clark et al., 2011). In the Tol2 construct, an inverted terminal repeat (ITR) sequence is recognized by transposase as the break site (Kawakami, 2007). The transposon plasmid is cleaved at the ITR site and inserted into the genome of the zebrafish (Kawakami, 2007). After a few hours, the transposase mRNA and protein are degraded (Clark et al., 2011). The integration of the transgene usually occurs during the early stages of embryonic development and can give rise to germ cells containing the transgene (Clark et al., 2011).

In my work, human ACTC variant cDNAs were cloned into the Tol2 transposon vector, which had two promoters expressing two different genes: the cardiac myosin light chain (cmlc) promoter and the $\beta$-actin promoter (Figure 3A). The cmlc promoter is a cardiac-specific zebrafish promoter that drove the expression of the human ACTC cDNA in the heart of the zebrafish. The $\beta$-actin promoter is a cytoplasmic promoter that drove expression of eGFP throughout the cells of the zebrafish. eGFP acts as a reporter gene to track successful transposition, identifying which zebrafish carry the human ACTC cDNA.
1.11 CRISPR as a Genome Editing Tool for Targeted Mutagenesis in Zebrafish

The use of CRISPR/Cas9 as a genome editing tool in zebrafish has improved the outcome of targeted mutagenesis. The CRISPR/Cas9 system consists of a single polypeptide endonuclease, Cas9, that produces a targeted double stranded break in DNA, and a single guide RNA (sgRNA) that recruits the Cas9 endonuclease to the target site. The sgRNA is a complex of two small RNA molecule that researchers have fused by forming a hairpin loop at the ends of both sides (Hwang et al., 2013; Gratz et al., 2014). The sgRNA consist of a CRISPR-RNA (crRNA), that binds to the complementary side of the target site, and trans-activating crRNA (tracrRNA) that helps in the maturation of crRNA and acts as a guide for the nuclease (Hwang et al., 2013; Gratz et al., 2014). When the Cas9 creates a double stranded break in the specific gene of interest, the break triggers the DNA repair mechanism of the organism to try and fix the break site either by non-homologous end joining (NHEJ) or homologous dependent repair (HDR) (Gratz et al., 2014). Implementing the HDR technique, we can supply a repair template, which has sequences flanking the cleaved site, with the CRISPR/Cas9 system. The HDR mechanism can use the homologous DNA sequences as a template for repair. Using this approach, it is possible to introduce point mutations in the zebrafish genome. In my work, I used the classic CRISPR/Cas9 system to examine the impact of knocking out zebrafish cardiac actin (zfactc) genes.

The zebrafish have a partially duplicated genome and they express multiple zfactc genes (Ojehomon et al., 2018). Therefore, the first step toward mutating a zfactc is to identify and know the role of each zfactc in the zebrafish genome.
1.11.1 Actin Isoforms in Zebrafish

Humans only have one ACTC gene and the ACTC protein is the main actin isoform in the adult heart; however, during early muscle development, studies have shown that the two forms of striated and smooth muscle actin pairs are co-expressed (Bertola et al., 2008). This co-expression might be due to the high level of conservation in the regulatory sequences of the actin genes and the need to accelerate the expression of actin proteins during development (Bertola et al., 2008). These isoforms are downregulated during the later stage of fetal development and the main isoform becomes the dominant gene in the adult tissue. For example, ACTC is downregulated in skeletal muscle and upregulated in cardiac tissue and ACTA is upregulated in skeletal muscle and downregulated in cardiac muscle (Bertola et al., 2008; Sartorelli et al.; Laing et al., 2009). From this concept, it is crucial to know which zfACTC is the dominant gene in the development of zebrafish.

In my first work, I identified four zfACTC genes: zfACTC1a which was duplicated in chromosome 19 and 20, zfACTa1b found in chromosome 13 and zfACTC1c found in chromosome 20 (Ojehomon et al., 2018). Protein sequence alignment of the zfACTC variants showed 99% sequence similarity to human ACTC and most of the differences are at the N-terminus. All the recognized zfACTC variants have the same N-terminus (Table 1). Performing in situ hybridization and syntenic analysis, the previously undesignated zfACTC1c (zgc:86709) was localized in the heart at 36 hpf, making it an ACTC gene in zebrafish (Ojehomon et al., 2018). zfACTC1c is predicted to be involved in mesenchyme migration, which is a population of cells that make up the mesoderm; in this case the heart (zgc:86709 | Danio rerio gene | Alliance of
Genome Resources). CRISPR injections against *zfactc1c* performed by Love Sandhu, a previous graduate student in the Dawson lab, displayed reduced cardiac function at 2 dpf (Sandhu, 2019).

As mentioned, *zfactc1a* is a duplicated gene located on chromosome 19 and 20. Previous studies have shown that a point mutation in *zfactc1a*, Y169S, also called the *s434* mutant, causes blood regurgitation at the atrioventricular canal. This mutant was an autosomal recessive gene as only a quarter of its progeny were affected and viable until 7 dpf. The *s434* mutant had fragile actin filaments and developed pericardial edema by 48 hpf and lack of endocardial cushion by 5 dpf (Glenn et al., 2012). The *s434* mutant was mapped to the *zfactc1a* gene on chromosome 20 and not on chromosome 19. We performed syntenic analysis to determine if the location of *zfactc1a* was a genome assembly error or if they are actually two separate identical genes (Ojehomon et al., 2018). Although more experiments are needed to determine their expression patterns, the syntenic analysis proved that the common regions of the two *zfactc1a* genes are not identical and supports the notion that they are two separate genes (Ojehomon et al., 2018). Furthermore, *in situ* hybridization of *zfactc1a* shows expression in the heart and somatic muscle (Glenn et al., 2012). Morpholinos targeted against the *zfactc1a* gene also displayed heart edema, enlarged heart and lack of endocardial cushion by 5 dpf (Glenn et al., 2012).

A mutation in *zfactc1b*, cardiofunk (*cfk*) mutant, causes endocardial cushion defects leading to improper valve formation. Just like *zfactc1a*, *in situ* hybridization shows expression of the *zfactc1b* gene in the heart and somatic muscle. Glenn et al., (2012) demonstrated that there was a partial redundancy effect between *zfactc1b* and *zfactc1a*. Embryos heterozygous for mutations in both *zfactc1a* and *zfactc1b* display a phenotype similar to homozygous *zfactc1a* mutants. Love Sandhu also performed CRISPR injections against *zfacta1b* (Sandhu, 2019). Prior
to 3 dpf, injected embryos had reduced cardiac function and altered cardiac phenotypes, but after this point their morphologies and cardiac function became more similar to mock injected embryos. Sandhu’s work proposed there could be compensation for the loss of zfacta1b occurring from the other zfactc genes.

To identify the best target for knock-in mutagenesis, knowledge of the function of each zfactc gene is crucial. I performed whole mount in situ hybridization (WMISH) to localize the expression of each zfactc gene during zebrafish development. I also performed reverse transcriptase quantitative PCR (RT-qPCR) on zebrafish hearts to determine the level of expression of the zfactc genes at different stage of development. Complementing previous work, CRISPR injections were done against zfactc1a to observe if a detrimental or a compensation effect occurs.
2 Materials and Methods

2.1 Reagents Used

1 kb plus ladder was obtained from Thermo Scientific (Burlington ON). EZ Vision dye was purchased from Fisher Scientific. All restriction enzymes were obtained from New England Biolab (NEB, Whitby ON). Mini prep, gel extraction and PCR purification kits were obtained from Qiagen (Germany).

2.2 Animal Husbandry and Egg Collection

Wild type (WT) zebrafish were originally obtained from Aquality Tropical Fish Wholesale, Mississauga, ON. Tübingen (Tü) zebrafish were obtained from Sick Kids. These lines were further maintained by the Hagen Aquality Facility on a 14-hour light schedule at approximately 28°C.

2.3 Needle Set Up

Injection needles were prepared using a Flaming Brown micropipette puller (Sutter Instrument, Novato CA) to pull a 1.0 mm OD glass capillary. The formed needle was backloaded with at least 1.5 µL of the injection material. #5 student Dumont Donostar forceps (Fine Science Tools, Vancouver BC) were used to clip the needle to obtain the right bolus size. A micrometer slide was also used to calibrate the needle to the right diameter of the bolus.

2.4 Embryo Injection

Injections were performed using a Pneumatic Pico Pump Injector (H.Saur Laborbedarf, Germany). Injections were done at or before the one-cell stage. The CRISPR components were injected into WT eggs while transposon components were injected into Tü eggs. The injection material was injected into the cell if the cell was present. A bolus size of 0.12 mm (~1 nL) was
used to inject the CRISPR while a bolus size of 0.10 mm (~0.5 nL) was used to inject the transposon. After injection, embryos were kept at 28°C in E3 embryo medium (5mM NaCl, 0.17 mM KCl, 0.33 mM CaCl₂ and 0.33 mM MgSO₄) with 0.1% methylene blue.

2.5 Transposon Plasmid

The pT3Ts plasmid (transposase), Tol2 transposon vector containing the cardiac myosin light chain (cmlc) promoter behind the GFP gene (pK Tol2cmlc-GFP) and Tol2 transposon vector containing the cmlc promoter and also the β-actin promoter behind the GFP gene (pK Tol2cmlcBA-GFP) were obtained as a gift from the Clark Lab, Mayo Clinic. WT Human ACTC cDNA and 4 ACTC variant cDNA (E99K, F90del, H88Y and R95C) were subcloned into the pK Tol2cmlcBA-GFP vector in front of the cmlc promoter by Mutagenex, a sub-cloning company (Ohio). The other ACTC variants were subcloned in-house by myself and Clea Heathfield, a 4th year undergraduate student, into the pK Tol2cmlcBA-GFP vector in front of the cmlc promoter.

2.6 Subcloning ACTC Mutants into the Transposon Vector

The ACTC variants originally cloned into pTopo2 plasmids were mini-prepped and quantified using a Nanodrop spectrophotometer (Thermo Scientific, Burlington ON). 5 µg of the purified plasmid was digested with 3 µL of EcoRI for 3 hrs in a 50 µl reaction volume and gel purified. The pK Tol2cmlc-WTACTC-BA-GFP was used as the receiving vector and digested with 3 µL of Mfel enzyme using 5 µg of the plasmid in a 50 µL reaction volume. The reaction was gel purified and treated with shrimp alkaline phosphatase (NEB) for 1 hr at 37°C. The shrimp alkaline phosphatase was then heat inactivated for 65°C for 10 mins. An insert to vector ratio of 3:1 (250 ng of vector) was used with T4 DNA ligase enzyme (NEB) and incubated
overnight at 16°C using an Eppendorf master cycler. The gels were not imaged prior to cutting to limit the amount of UV exposure. Ligated products were transformed into *E. coli* DH5α.

### 2.7 Transformation into *E. coli* DH5α

*E. coli* DH5α competent cells (a gift of the Khursigara lab) were prepared with calcium chloride treatment. 10 μL of the ligation reaction was transformed into *E. coli* DH5α competent cells by heat shock transformation. This reaction was left to incubate on ice for 30 mins, followed by a 30 sec incubation at 42°C and lastly, an incubation on ice for 1.5 mins. Afterwards, the heat shocked cells were diluted with 250 μL of SOC media and left to recover at 37°C for 1 hr at 200 RPM. The transformation reaction was spun down with a tabletop centrifuge at 3000 RPM for 4 min after which 200 μL of the supernatant was removed. The cell pellet was resuspended with the remaining supernatant and spread on a LB-agar plate containing kanamycin at a concentration of 50 μg/mL. The plates were grown overnight at 37°C.

### 2.8 Screening for Positive Colonies: Successful Ligation and Proper Orientation by Colony PCR

Approximately 10 colonies were chosen from the LB agar plate. Each colony was resuspended in 12 μL of sterile water. Two PCR reactions were performed. One confirmed a successful ligation while the second checked for proper orientation. In a PCR tube, 5 μL of the resuspended cells were required for each reaction. A PCR master mix cocktail was made and 20 μL was aliquoted into each PCR tube containing the resuspended cell. Each tube contained 0.2 μM of reverse and forward primer each, 0.4 μM dNTP, 1X Taq Thermo poly buffer and 0.2 μL of Thermo Taq polymerase. Thermocycler conditions for the PCR reaction consisted of an initial
denaturation at 95°C for 5 mins, followed by 35 cycles of denaturation at 95°C for 30 sec, annealing at 42°C (for successful ligation) or 46°C (for proper orientation) for 30 sec, extension at 72°C for 1 min followed by a final extension at 72°C for 3 min and an indefinite hold at 4°C. Primers were purchased from Lab Services (Guelph, Ontario). The remaining 2 μL of the resuspended cell was kept at 4°C until it was verified which cell had the positive colonies. After verification, cultures were grown overnight at 37°C and 200 RPM in LB media containing 50 μg/mL of kanamycin.

2.9 Generating Transposase mRNA

The quantified PT3TS plasmid (approximately 5 μg) was linearized using XbaI restriction enzyme for 2 hours at 37°C. The reaction was PCR purified and the state of the linearized plasmid was verified using a 0.6% agarose gel. After linearization, the DNA was transcribed using Ambion’s mMessage Machine T3 Kit (Life Technologies, Burlington ON) and precipitated using lithium chloride.

2.10 Germ Line Transmission

Embryos were sterilized at 1 dpf using a 0.07% sodium hypochlorite solution and dechorionated manually using #5 Dumont Donostar forceps or chemically using pronase at a concentration of 2 mg/mL.

Between 24 and 48 hpf, embryos were screened for GFP expression using a dissecting microscope with a GFP filter. Positive embryos expressing GFP were introduced into the Hagen Aqua Lab at 5 dpf. The embryos were fed with Gemma 75 (Skretting, North America St. Andrews) from 5 dpf to 1 month and then switched to Gemma 300 from 1 month to adulthood. Embryos were fed 3 times a day to increase their rate of development. After the fish reached
sexual maturity at 3 months, the males and females were crossed to identify germ-line founders (P₀ generation). The P₀ that did not produce offspring with GFP expression were not used for further experiments. Parents that produced offspring with GFP expression were out-crossed with the Tü line to create a stable line that possessed a Mendelian inheritance (F₁ generation). The F₁ generation were then outcrossed with the Tü line or in-crossed to search for Mendelian inheritance of the transgene and F₂ embryos were tested.

2.11 Embryo Heart Rate Analysis

A total of 24 embryos each from the transposon, CRISPR injected, or the mock control groups were randomly selected. The embryos were immobilized with 1.2% low melting point agarose on a 24-well plate lid and arranged in a way that the heart was visible. Heart rate videos were taken using a Samsung S7 cell phone camera at 60 frames per second and a video size of 1920 X 1080 pixels. Embryos were carefully removed from the agarose after videos were taken and kept in a 24-well plate. Heart rate videos were taken from 2 dpf to 7 dpf. Genomic DNA (gDNA) was individually extracted from the embryos at the end of the 7th day.

Heart rate analyses for the transposon fish were performed as previously described (Avey et al., 2018), while heart rate analyses for the CRISPR injected fish were analyzed using Danioscope software (Noldus, Leesburg). When the software gave a heart rate that was too high due to the 2nd harmonic peak being read, the heart rate output was divided by two to acquire the peak that would have been read in the first peak.

2.12 Phenotype Trajectories

The phenotypes of the zebrafish embryos were characterized from 2 dpf to 7 dpf at the same time the heart rate videos were taken. The developing embryos were scored based on their
viability, severity of cardiac edema, blood regurgitation, blood circulation, body and tail curvature, tail size, yolk necrosis and heart morphology. Based on the observation, the phenotype of the transposon embryo was categorized as normal (0 defects), moderate (1-3 defects), or severe (>3 defects) based on the observation.

2.13 Extraction of Genomic DNA from Zebrafish Embryos

gDNA was individually extracted from the embryos according to the protocol from Meeker et al., (2007).

2.14 Thermal Asymmetric Interlace Polymerase Chain Reaction (TAIL-PCR)

Tail PCR and primers used were performed and designed as previously described (Parinov et al., 2004). Amplicons produced in both secondary and tertiary reactions were separated on a 1.2% agarose gel, and then gel purified. These bands were sequenced and searched on the nucleotide BLAST database to determine the location of the inserted gene.

2.15 Designing of CRISPR Targets

CRISPR targets were identified using CHOPCHOP (https://chopchop.cbu.uib.no/), an online webtool that has an in-built zebrafish genome reference along with other disease models. With its algorithm and the template we chose, we detected off target sites (if any), the location of where the Cas 9 endonuclease would cut and identified the primers to amplify the cut site. The gRNA was designed to target the exon 1 of the \textit{zfactc1a} gene. An earlier exon was chosen to increase the possibility of a complete loss of function of the \textit{ACTC} gene. The gRNA and primers were also designed to target both genes in chromosome 19 and 20. Lastly, the gRNA was designed to recognize a 20-nucleotide target sequence next to a trinucleotide NGG protospacer adjacent motif (PAM). Primers were ordered from Lab Services, University of Guelph (Table 2)
2.16 gRNA Template Assembly and Generation of sgRNA

The protocol for template assembly was adapted from Gagnon et al., (2014). This technique is a cloning-free method where a constant oligonucleotide that encodes the reverse complement of the tracrRNA was used. The constant oligonucleotide sequence was 5’AAAAGCACCAGACTCGGTGCCA CTTTTTCAAGTTGATAACCGACTGCTTTATTTAACTTGCTATTTCTAGCTCTAAAAC’3. The gRNA oligonucleotides were then designed to contain the SP6 promoter sequence at one end (for transcription), the 20-nucleotide base target site without the PAM in the middle and a complementary region to anneal to the constant oligonucleotide. The formula for ordering the gRNA oligonucleotide was 5’ATTTAGGTGACACTATA-N_{20}-GTTTTAGAGCTAGAAATAGCAAG’3.

After the gRNA template was prepared and both transcription and purification were successful, the pellet was resuspended in 12 µL of fresh milliQ water and quantified using the Nanodrop spectrophotometer (ThermoFisher Scientific, Burlington ON). 1 mg/mL of Cas 9 protein (PNA BIO Inc, California) and 1 µg of sgRNA were co-injected into the yolk or cell of the embryo.

2.17 High Resolution Melting Curve Analysis (HRM)

HRM analysis was performed to individually validate CRISPR activity in the injected embryos. The PCR reaction was done in triplicate using the Type-it HRM PCR Kit (Qiagen, Germany) using a 10 µL reaction volume from the manufacturer’s protocol. The gDNA was diluted 1:100 and 4.3 µL of the diluted gDNA was used to top off the volume of the reaction. The reaction was then amplified using the Step One Plus Real Time PCR system v2.3 from Applied Biosystem (ThermoFisher Scientific, Burlington ON) with the following conditions; initial denaturation at 95°C for 5 mins, followed by 35 cycles of denaturation at 95°C for 30 sec, annealing at 55°C for 30 sec, extension at 72°C for 1 min followed by a final extension at 72°C
for 3 min and an indefinite hold at 4°C. The amplified reactions were further analysed using HRM software.

2.18 Sequencing, Tide and Indigo Gear analysis

After HRM analysis, the amplicons from samples with CRISPR activity were PCR purified and sequenced. TIDE (https://tide.deskgen.com/) analysis was performed to deconstruct the chromatogram that was received, thereby comparing the mock injected embryo to the CRISPR injected embryo. The chromatogram was also screened through the INDIGO web-based tool (https://www.gear-genomics.com/indigo/), which helped to identify insertions and deletions by performing an alignment of the CRISPR injected embryo to the mock injected embryo.

2.19 Isolation of Hearts from Zebrafish Embryos

The method used to isolate hearts from zebrafish embryos was adapted from Yang and Xu, 2012. Ten hearts were used to make up a sample for each stage of development and 3 samples were made to represent 3 different biological replicates. The embryos were anesthetized in 1xPBS containing 1:20 dilution of 0.4% tricaine. The anesthetized embryos were transferred onto a cover slip that was taped on a dissecting microscope. Two different needles attached to a BD syringe were used to pull the hearts out from the embryo. A 27.5-gauge needle was used to hold the embryo in place while a 26.5-gauge needle was used to pull the heart from the embryo. The pulled heart was transferred into a 1.5 mL microcentrifuge tube using a siliconized glass pipette. This tube was placed on ice until all 10 hearts were pooled into the tube. The tube containing the pooled hearts was centrifuged at 1000 RPM for 5 mins at 4°C. Excess liquid was removed using the 27.5-gauge needle and the tube was placed on dry ice. The excess liquid that was removed was examined to determine if any hearts were also removed. If a heart was
accidentally removed from the tube, then the whole tube was discarded, and a new sample was prepared.

2.20 Extraction of RNA, cDNA Synthesis and Reverse Transcriptase Quantitative PCR (RT-qPCR)

Extraction of RNA, cDNA synthesis and RT-qPCR and analysis were performed as previously described (Ojehomon et al., 2018). Following isolation of the embryonic hearts, the SingleShot Cell Lysis Kit (Bio-Rad Laboratories, Hercules, CA) was used to extract the RNA following the manufacturer’s protocol at half-reaction volumes. The extraction of RNA from adult hearts was performed as previously described (Peterson and Freeman, 2009). The extracted RNA from the adult hearts were further processed by adding 1 ul of Turbo DNase from the Ambion’s mRNAge kit (Life Technologies, Burlington ON) for every 10 ul of total RNA and incubated for 30 mins at 37°C. 1/10th of 3 M sodium acetate was added to the reaction, vortexed and spun down at 14000 RPM for 10 mins. 1 µl of glycol-blue co-precipitant (Invitrogen Life Technologies, Burlington ON) was added to the supernatant with 2x volume of 100% ethanol. The reaction was left to incubate for a minimum of 1 hr and then spun down at 14000 RPM for 15 mins. The supernatant was removed, and the pellet was washed with 1 ml of 70% ethanol. The pellet was left to dry and later resuspended in 12 ul of water.

Synthesis of the cDNA was performed using the iScript Advance cDNA synthesis Kit (Bio-Rad Laboratories, Hercules, CA) following the manufacturer’s instructions such that the final reaction volume contained 40% of the extracted RNA. After cDNA synthesis, the reaction was further diluted by 4X to reduce the reaction components from the Kits from interfering with the RT-qPCR reactions. Gene-specific primers for *zfactc1a, zfacta1b, zfactc1c*, and Ribosomal
Protein L8 (RPL8) were designed and ordered from Lab Services (Table 2). Rpl8 was chosen as the housekeeping gene as a reference to normalize the zfactc gene due to prior research (Ojehomon et al., 2018). Primer efficiencies were measured for each primer, and the generated standard curve was used to analyze the data obtained from the RT-qPCR reactions. A volume of 12 ul was performed for the RT-qPCR reactions for each gene with 1X Power SYBR Green (Life Technologies, Burlington ON), 0.2 µm each of forward and reverse primers, and 3 ul of the diluted cDNA. Two negative controls were also performed; absence of RT and water, and a positive control to act as an internal calibrator was also performed. Triplicates were done for each reaction. The Cq values were averaged, and the standard curve method was used to analyze the expression values. An antilog of the values obtained from the slope of the standard curve was performed and further divided by the rpl8 values to normalize the expression. A standard t-test was performed on the Cq values of rpl8 to measure stability.

2.21 Whole Mount In Situ Hybridization (WMISH)

WMISH was performed on embryos as previously described (Thisse and Thisse, 2008). Antisense RNA probes for zfactc1a, zfacta1b or zfactc1c genes were synthesized as previously described (Ojehomon et al., 2018). The embryos were incubated in the probes for 10 hrs. Embryos were imaged with an Axio Imager-M2 compound microscope (Zeiss, Germany) with a 20x/0.5 numerical aperture objective. Primers to amplify the zfactc1a gene were ordered from Lab Services (Table 2).
2.22 Statistical Analysis

Statistical analysis was performed using SPSS ® Statistics version 16.0 (IBM Corporation. Armonk, NY, USA). Significance was determined using two – way Anova followed by Bonferroni test. Differences were considered significant when $p<0.05$.

### Table 2 Primers used in this study and not reported in previous literature

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zfactc1a gDNA oligo</td>
<td>5’ATTTAGGTGACACTATA ACCATTGTCGCACACGAGTG GTTTTAGAGCTAGAAATAGCAAG’3</td>
<td>sgRNA synthesis</td>
</tr>
<tr>
<td>Zfactc1a HRM F</td>
<td>5’GAACCATCCTAAAAAACAGC’3</td>
<td>HRM</td>
</tr>
<tr>
<td>Zfactc1a HRM R</td>
<td>5’AAACAATAGAGGGAAAGACAG’3</td>
<td>HRM</td>
</tr>
<tr>
<td>Zfactc1a F</td>
<td>5’CCATCGTCCACAGAAAAGTGC’3</td>
<td>WMISH</td>
</tr>
<tr>
<td>Zfactc1a R</td>
<td>5’GGACACATCAGAACTTTTATTAC’3</td>
<td>WMISH</td>
</tr>
<tr>
<td>Zfactc1a F</td>
<td>5’ATCTTCAATCCTTCTCTGTCAGTC’3</td>
<td>RT-qPCR</td>
</tr>
<tr>
<td>Zfactc1a R</td>
<td>5’TATACAGCACAAGGCACAGTACAA’3</td>
<td>RT-qPCR</td>
</tr>
</tbody>
</table>
3 Results

3.1 Aim 1: Express Human ACTC cDNA Variants in the Zebrafish genome using Tol2 Transposons

3.1.1 Generation of Transgenic Zebrafish Expressing a Human ACTC Variant using Transposons

My goal was to develop a stable zebrafish transgenic line that expresses human ACTC WT cDNA and human ACTC variant cDNA related to cardiomyopathy. Human ACTC variant cDNAs were cloned into the Tol2 transposon vector behind the cmlc promoter (Figure 3A), driving the expression of the ACTC cDNA in the heart. A β-actin promoter was also present in the transposon vector and was used to drive eGFP expression throughout the cells of the zebrafish as a reporter (Figure 3A). This reporter gene tracks successful transposition, identifying which zebrafish carry the human ACTC cDNA. A founder in my study is defined as an adult zebrafish that was able to produce embryos with a transgene.

Injecting pKTo12cmlc-GFP (cmlc-GFP, transposon vector without human ACTC cDNA and β-actin promoter) was performed first to determine the efficiency of the transposon system. An efficiency of 75% was obtained where 15 out of 20 adult zebrafish produced embryos with eGFP expression in the heart (glowing heart) (Table 3). Unlike the cmlc-GFP that had a high transgenesis rate, embryos injected with pKTo12cmlc-ACTC-BA-GFP (cmlc-ACTC) had a low transgenesis rate. It was difficult to find an adult transgenic zebrafish that produced embryos expressing eGFP throughout the body (glowing body). Either no founder or only one founder was seen in the number of adult transgenic fish expressing human ACTC variants (Table 3).
Table 3: Number of founders discovered per transposon injection. * indicates a second trial

<table>
<thead>
<tr>
<th>Plasmid injected</th>
<th># of adults survived</th>
<th># of founders</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>cmlc</em>-GFP</td>
<td>20</td>
<td>15</td>
</tr>
<tr>
<td><em>cmlc</em>-WT</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td><em>cmlc</em>-WT*</td>
<td>17</td>
<td>1</td>
</tr>
<tr>
<td><em>cmlc</em>-E99K</td>
<td>43</td>
<td>1</td>
</tr>
<tr>
<td><em>cmlc</em>-R95C</td>
<td>49</td>
<td>0</td>
</tr>
<tr>
<td><em>cmlc</em>-H88Y</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td><em>cmlc</em>-F90del</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td><em>cmlc</em>-A295S</td>
<td>9</td>
<td>0</td>
</tr>
</tbody>
</table>

Injecting *cmlc*-ACTC plasmid (Figure 3A) with transposase mRNA produced embryos expressing GFP in a mosaic manner (Figure 3B I). These embryos were grown to give rise to a founder producing embryos with the ACTC variant in its germline (Figure 3B II). Only one founder from the injection of E99K human ACTC cDNA transposons (E99K-P0) produced embryos that could be grown into adulthood. Two sets of embryos were grown from this E99K-P0: set A and B (Figure 3C).

The F1 generation (4 in total) in set A produced embryos that expressed the transposition reporter in its germline had glowing bodies, but the F2 generation were unable to continue the trend. The F2 generation from an incross of the F1 resulted in embryos that had glowing hearts instead. When these glowing heart embryos reached adulthood, they gave rise to embryos that had glowing hearts and glowing bodies (Figure 3C). In the set B F1 generation, 4 out of 22 transgenic adult fish gave rise to embryos expressing the transposition reporter in their germline. This F2 generation was able to give rise to embryos with varying degree of GFP expression, with some embryos fluorescing more than others. Conversely, not all of the F3 generation were able to give rise to embryos that expressed the GFP reporter (Figure 3C).
Figure 3: Resulting transgenesis using Tol2 transposon. A) Plasmid DNA containing the Tol2 transposon vector co-injected with the transposase mRNA. B) The injected embryo (I) showing mosaicism raised and mated with WT to give rise to F₁ embryos (II). Embryos expressing GFP acted as a control to indicate which embryo is carrying the plasmid. C) Generations of crosses performed on the transgenic zebrafish showing loss of expression after F₂ generation.
3.1.2 Functional Analysis of Transgenic Zebrafish

Functional analysis was performed on embryos obtained from an incross of the F$_1$ generation from set B. Since, it was not determined if the embryos expressed recombinant ACTC protein, the embryos from the incross of the F$_1$ generation were called transposition reporter expressing embryos (TREE) A total of 81 TREE were characterized from 2 dpf to 7 dpf displaying the following disease phenotypes: slow development, blood regurgitation, cardiac edema, enlarged hearts, bent tails, and bent bodies (Figure 4D). Based on the severity of defects observed in this sample, 12% of fish were classified as having a normal phenotype, 47% had moderately severe phenotype, and 41% had severe phenotypes (Table 4). 100% of fish classified as having severe phenotypes died over the 7 dpf period (Figure 4C).

**Table 4: Phenotype classification of TREE.** Embryos were obtained from 4 pairs of transgenic adult zebrafish. Embryos were grouped into 3 categories depending on the severity of the phenotypes observed. The proportion of the sample (N=81) in each category

<table>
<thead>
<tr>
<th></th>
<th>Normal</th>
<th>Moderate</th>
<th>Severe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair 1</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Pair 2</td>
<td>1</td>
<td>15</td>
<td>8</td>
</tr>
<tr>
<td>Pair 3</td>
<td>1</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Pair 4</td>
<td>5</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>Proportion</td>
<td>12.34%</td>
<td>46.91%</td>
<td>40.74%</td>
</tr>
</tbody>
</table>
Figure 4: Effect of expressing a transgene in the zebrafish genome. Heart rate variability plot displaying the average heart rate in bpm of control (A) and TREE (B). The statistics were compared between the heart rate of the control to the heart rate of the TREE within each day. Their significance was determined using two – way ANOVA followed by Bonferroni test, and the error bars represent the standard deviation of the mean. The presence of 3 asterisks above an entry represents $p<0.01$ while 4 asterisks represents $p<0.005$. (C) Kaplan Meier survival plot showing the time of death of E99K ACTC expressing embryos over 7 days. D) Phenotype characterization performed at 2 dpf. Arrows indicate observations that differ from WT phenotype (V) such as blood pooling (I), necrosis of the yolk and body (II), pericardial edema (III), bent tail (IV) and slow development (II, IV, VI). Heart rate analyses and survival graph data were generated by Clea Heathfield under my supervision.
3.1.3 **Heart Rate Analysis**

A sample of 24 zebrafish embryos expressing the human ACTC variant were observed to have a reduced heart rate when compared to WT zebrafish (Figure 4A-B). The heart rate of the TREE increased for the first 4 dpf before consistently decreasing for the rest of the 7 dpf period. The average resting heart rate of the TREE at 7 dpf was 108 bpm, 17% lower than that of the WT. The difference in heart rates of embryos at 2, 3, 5, 6, and 7 dpf was found to be statistically significant (P < 0.01) when compared to the WT. There was no significant difference in the heart rates between the TREE and WT at 4 dpf.

3.1.4 **TAIL-PCR**

TAIL-PCR was performed to determine if the phenotype observed was due to the location or the number of transpositions in the zebrafish genome. An average of 2-3 PCR products were amplified per embryo and the sizes of the amplified fragments varied between 300-400 base pairs (bp) from the secondary round of TAIL-PCR and 200-300 bp from the tertiary round. The shift in the amplicon size in the tertiary reaction when compared to the secondary reaction is indicative of an integration site. Sequencing revealed that *Tol2* insertions were flanked by zebrafish genomic DNA on chromosomes 2, 5, 6, or 14 (Table 5). All of the embryos had approximately 2 to 3 integrations of the *Tol2* construct in their genome, regardless of defect severity. Embryos with insertions that were found in chromosome 6 and 14 had a more severe phenotype irrespective of the number of integrations. Taking a closer look at chromosome 6 and 14, the *Tol2* was inserted into the Transcription Factor Activating Protein 2 (TFAP2) Gamma gene and an uncharacterized region, respectively (Table 5).
Table 5: Analysis of Tol2 integration sites and the possible gene and location of integration. Integration sites were based on the danRer11 genome assembly.

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Chromosome locus</th>
<th>Nearest transcribed region</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2:36055884-36128278</td>
<td>Near transcript CU693370.1</td>
</tr>
<tr>
<td>5</td>
<td>5:40255810-40256328</td>
<td>Near transcript SNX18a</td>
</tr>
<tr>
<td>6</td>
<td>6:56057816-56238415</td>
<td>Within TFAP2 gamma</td>
</tr>
<tr>
<td>10</td>
<td>10:5597096-5597126</td>
<td>Near transcript DIRAS2</td>
</tr>
<tr>
<td>14</td>
<td>14:44266619-44267007</td>
<td>Near transcript RF00001</td>
</tr>
</tbody>
</table>

Ultimately, I was not able to maintain a transgenic zebrafish line expressing human ACTC protein using Tol2 transposons. Therefore, I moved on to test a different approach to examining cardiac actin mutations using zebrafish: editing the cardiac actin genes of the zebrafish itself.

### 3.2 Aim 2: Determining the best zfactc gene to target for gene editing

Recent studies have shown that zebrafish possess multiple ACTC genes and it is unclear which genes are specific to the heart (Ojehomon et al., 2018; Glenn et al., 2012; Bartman et al., 2004). There are 3 zfactc genes that need to be studied; zfactc1a, zfacta1b and zfactc1c. Knowing where these genes are being expressed and the function each gene plays at different stage of development is necessary to identify which gene is the best target for targeted mutagenesis.

#### 3.2.1 WMISH Displays Isoactin Switching in a Chamber-Specific Manner

At 24 hpf and 36 hpf, zfactc1a, zfacta1b and zfactc1c transcripts are shown to be expressed in the heart tube and the somatic muscle of the developing embryo (Figure 5). At 48 hpf, zfactc1a is localized mostly to the ventricle, zfacta1b is expressed in both the ventricle and atrium while zfactc1c is not expressed. At 72 hpf, zfactc1a is expressed only in the heart and not in the somatic muscle. It is expressed in the ventricle and in the outer edges of the atrium.
Figure 5: Localization of zfactc genes using WMISH. Lateral and ventral view of zfactc genes in the heart tube and somatic muscle at different stage of development. A small picture diagram in each figure is displayed to represent what is seen at each stage. The dorsal view of embryos at 24 hpf was taken instead of the ventral view to show the clarity of the expression in the heart.
Zfactc1c ceases to be expressed at 72 hpf and zfacta1b is expressed at a low level in the outer edges of the ventricle and atrium but very high in the somatic muscle. Zfacta1b shows no signal in the heart in embryos obtained at the end of the day at 3 dpf while zfactc1a continues to be expressed.

3.2.2 Expression Profile of the zfACTC Isoforms Using RT-qPCR

10 embryonic hearts at different stage of development were extracted from zebrafish embryos and pooled together to create one sample. Three samples in total were made to make up three different biological replicates. Hearts from adult zebrafish were dissected ensuring that the ventricle, atrium and the bulbous arteriosus were present. RT-qPCR analysis confirmed that all zfactc genes were expressed at different stages, with zfactc1a and zfacta1b following similar levels of expression at each stage (Figure 6). Zfactc1a level of expression at 60 hpf was significantly higher when compared to zfacta1b and zfactc1c. There was no significant difference between the Cq values of rpl8 at each stage.
Figure 6: Relative Expression levels of zfactc genes in the heart using RT-qPCR. mRNA expression of zfactc1a, zfactc1b, zfactc1c of 10 embryonic hearts at each embryonic stage and 1 adult heart. Experiments were done in triplicate. Expression levels of the of the zfactc genes relative to each other and normalized using the house keeping gene, rpl8. Error bars represent the average relative output of 3 biological replicate in standard error measure (SEM). Their significance was determined using two – way ANOVA followed by Bonferroni test, and statistics were compared between the relative expression of the genes within each day. ** indicates $p<0.05$ and *** indicates $p <0.01$
3.3 Mild Phenotype Observed in *zfactc1a* CRISPR Injected Embryos

sgRNA were designed to target *zfactc1a* in both chromosome 19 and 20. CRISPR and mock injections were performed on three different broods of eggs to make up three biological replicates. The replicates were completed to analyse the heart rate data and observe phenotypic changes. A total of 72 CRISPR injected embryos and 72 mock injected embryos were observed for their heart rate and the presence of a cardiac phenotype.

Phenotypes observed in the CRISPR injected embryos showed mild cardiac phenotypes like slow development, blood pooling, edema; of which the embryos looked phenotypically normal at 7 dpf (Figure 7A). In all three replicates 2 to 3 embryos had an intermittent contraction at 4 dpf (Figure 7B). In contrast, only one control embryo in experiment 3 experienced intermittent contraction.

An observation that was consistently observed was that the *zfactc1a* CRISPR injected embryos displayed faster movement than the mock injected embryos because it was harder to catch them in the well or petri plate with a Pasteur pipette during analysis.

Between each biological replicate, the heart rates were not statistically similar; therefore, the data could not be combined. In trial 1, a significant difference between the CRISPR and the mock injected embryos was seen at 5 dpf (*p*=0.043), 6 dpf (*p*=0.004) and 7 dpf (*p*=0.027) (Figure 7A). In trial 2, a significant difference in heart rate was seen between the mock and CRISPR injected embryos at 2 dpf (*p*=0.008), 5 dpf (*p*=0.026), and 6 dpf (*p*=0.049) while in trial 3, a significant difference at 2 dpf (*p*=0.009) and 3 dpf (*p*=0.034) was observed (Figure 7B-C).

HRM analysis on the individual embryos, displayed two melt curves from the mock injected (Figure 8B). These melt curves were used to analyse the CRISPR injected embryos. If a CRISPR injected embryo displayed the same melt curve as one of the mock injected curves, it
was considered to have no CRISPR activity. If a CRISPR injected embryo displayed a different melt curve than both mock injected curves, it was considered to have a high CRISPR activity (Figure 8C). If a CRISPR injected embryo displayed a melting curve different than one mock injected curve but a melting curve similar to the other mock injected curve, the embryo was considered to have a low CRISPR activity.

HRM analysis was performed in trial 2 and 3 and the embryos that showed no CRISPR activity were removed from the heart rate data. The HRM analysis in trial 2 revealed that within the 24 CRISPR injected embryos, 11 embryos had a low CRISPR activity and 5 embryos had a high CRISPR activity. For trial 3, 16 embryos displayed CRISPR activity with 14 considered to have low CRISPR activity and 2 considered to have high CRISPR activity. The sequencing results were done for a few embryos that had high and low CRISPR activity in HRM and showed cardiac phenotypes. The sequencing result displayed a chromatogram similar to the mock injected with the sequence targeted by the sgRNA still intact. This was run through TIDE and the efficiency of a nucleotide change or deletion was not more than 2% for each sequence analyzed. An INDIGO analysis reveals that the chromatograms possess single nucleotide variants (SNV).
Figure 7: Heart rate data in beats per min of zfactca CRISPR injected embryos. Data was obtained from Danioscope software. A) Heart rate of embryos in trial 1, Mock injected (24) Actc1a CRISPR (24) B) Heart rate of embryos in trial 2, Mock injected (24) Actc1a CRISPR (16) C) Heart rate of embryos in trial 3, Mock injected (24) Actc1a CRISPR (16). Data are expressed as means ± standard error means. Significance was determined using a two-way ANOVA followed by a Bonferroni test. The statistics were compared between the heart rate of the mock injected embryos to the heart rate of the CRISPR injected embryos within each day * p<0.05, ** p<0.01.
Figure 8: Phenotype to genotype of zfact1a-CRISPR injected embryo. A) Phenotype observation of mock and CRISPR injected embryos. CRISPR injected embryo showing slow development compared to mock injected. In addition, at 3 dpf the heart of the CRISPR injected was much smaller than the mock injected. B) HRM analysis showing the two mock injected curve that was present in both trial 2 and 3. C) HRM analysis showing the two mock injected curves aligned to a CRISPR injected curve considered as having high CRISPR activity. D) Standard Sanger sequencing chromatogram showing the sequence output of the mock injected and CRISPR injected embryo. Grey area displays the target site for the Cas9 endonuclease. Arrow shows the proposed cut site and where indels should be present.
4 Discussion

4.1 Heart Specific Expression of Human ACTC Variants Using Tol2 Transposons is not Stable

The transposon vector used in my work was designed to produce embryos containing the transgene to have glowing bodies for easier identification. Although I was able to generate transgenic zebrafish with glowing bodies, suggesting successful transposition of the human ACTC cDNA, the transgenic lines were unstable. Additionally, comparing the transgenesis rate of the zebrafish carrying cmlc-GFP to zebrafish carrying cmlc-ACTC, the transgenesis rate for the cmlc-ACTC was relatively low as it was extremely difficult to find a founder. It is possible that the transgene cassette underwent transcriptional repression as an epigenetic mechanism, as this is a potential problem when using transposon vectors (Sharma et al., 2012; Davis et al., 2012; Garrison et al., 2007). Under non-selective conditions in F9 and HeLa cells, transcriptional silencing of Sleeping Beauty transposon has been detected after 7 weeks of passage (Sharma et al., 2012; Garrison et al., 2007). In addition, adding deacetylase inhibitor to cells containing piggyBac transposon vectors, cells were able to regenerate their transgene vectors (Sharma et al., 2012).

Some factors that could cause the repression of the transgene are, the strength of the promoter, the amount of ACTC being expressed, the position that the cDNA was inserted, and the epigenetic mechanism of the zebrafish. It has been revealed that silencing of transposon-based vectors can be diminished by careful transgene design (Garrison et al., 2007; Sharma et al., 2012). Although Tol2 is known to carry at least 10-kb sequences, the adult zebrafish expressing GFP throughout its body (glowing body) gave rise to embryos expressing GFP only in hearts.
(glowing heart). These glowing heart embryos gave rise to glowing body and glowing heart embryos, indicative of epigenetic regulation of the genome (Figure 3C).

Testing different cardiac specific promoters or reducing the size of the plasmid to avoid the potential for reconfiguration could help limit the potential of transcription repression. Lastly, adding more insulators to the transposon vector can block enhanced activation of a promoter and also protect transcribed regions from chromosomal position effects to reduce the likelihood of epigenetic regulation.

4.2 Severity of the Cardiac Phenotype Observed is due to the Integration of the Transposon in the Zebrafish Genome

The embryos used to determine the functional analysis of successful transpositions were obtained from a F1 generation incross. The heart function of the GFP-expressing embryos, suggesting ACTC variant expression, was impaired as the embryos exhibited cardiac phenotypes such as edema, blood pooling, bent or short tails, and a decrease in the heart rate compared to WT embryos. These are common cardiac phenotypes as shown when researchers expose embryos with cardioactive compounds known to affect the function of the heart (Jeanray et al., 2015). Acetaminophen, a cardioactive compound is known to cause malformations in the heart, yolk sac and tail (Jeanray et al., 2015). Although the blood flow of the TREE were not measured, a reduction in the contractility of the heart could reduce the amount of nutrient available for the embryo from the yolk sac which can lead to necrosis and result in a shorter tail to compensate for the loss of nutrients (Fraysse et al., 2006).

While we observed many cardiac phenotypes in the embryos expressing the GFP reporter of ACTC variant transposition, we have to take into account the technique used to edit the
genome. During our observations, we realized that there were more embryos categorized as severe and moderate than normal. This increase in severity could be due to the embryos being obtained from an incross of two parent fish expressing a transgene. Therefore, it was important to validate if the phenotype was due to insertion mutagenesis, the number of transposon insertions, or overexpression of the transgene.

I collaborated with Avey et al.,(2018) to perform a swim test study with 3 different zebrafish lines: WT, tg(cmlc:actc,ba:gfp), and tg(cmlc:gfp). The experiment was done to test if there was a difference in the cardiac and the aerobic function of zebrafish with a transgene. We showed that although there was not a significant difference in the metabolic rate of the three groups of fish, tg(cmlc:actc,ba:gfp) had a 1.8 higher fold difference in its metabolic rate than the other group of fish used, which suggests that the swim efficiency was affected. In addition, the cardiac function of the tg(cmlc:actc,ba:gfp) embryos was also affected, where the heart rate at 4 dpf was 20% lower than WT and 14% lower than tg(cmlc:gfp) (Avey et al., 2018).

Performing TAIL-PCR on the embryos expressing the transgene we showed that the transposon was integrated at multiples sites within the genome. On average, each embryo was observed to have the same number of integrations regardless of phenotype severity. Embryos with more severe phenotypes had integrations in either chromosome 6 or chromosome 14. This result suggest that the Tol2 transposon may have disrupted an essential gene in the zebrafish genome. The transposon integration in chromosome 6 could have disrupted the gene, TFAP2 gamma which is needed in the activation of several developmental genes during early development. The integration pattern of Tol2 transposons in zebrafish are usually integrated in a random pattern but Tol2 vectors do show a preference for transcriptional units (Sharma et al.,
2012). Therefore, I conclude that the number of integrations is not the cause of the severe phenotypes seen but may be due to either overexpression of the transgene or positional effects due to insertional site.

Using a transgenic line proposedly carrying the WT human ACTC cDNA would be a proper control as we can validate if the phenotypes observed are due to overexpression of the transgene. There are limitations when using TAIL-PCR to determine the number of integrations in the zebrafish genome. TAIL-PCR is not a quantitative method; therefore, the number of integrations can be underestimated (Parinov et al., 2004). TAIL-PCR should be used to locate where your transposon is located. Using Southern blot hybridization is a more quantitative experiment and could be done to support the result seen.

In summary, the results obtained from Aim 1 looked promising but since the expression of human ACTC cDNA with the Tol2 transposon was not as stable as I hoped, I decided to work on mutating an endogenous cardiac actin gene of the zebrafish using the CRISPR/Cas9 system.

4.3 In Situ Hybridization Supports a Cardiac Gene Switch Model

The zebrafish genome includes many gene duplications, resulting in gene sub-functionalization and neofunctionalization. My first published work described three zfactc genes (Ojehomon et al., 2018). WMISH was performed to determine where the zfactc genes are being expressed and the function each gene plays at different stages of development. This information is needed to identify which gene is the best target for mutagenesis. In situ hybridization results confirm that zfactc1a is expressed after 48 hpf while zfacta1b and zfactc1c cease to be expressed after 36 hpf and 48 hpf, respectively (Figure 4A). Zfactc1a is expressed in the heart from 24 hpf to 3 dpf, where it is localized in both chambers at 3 dpf, while zfacta1b is expressed in both
chambers at 24 hpf to 48 hpf. Zfactc1c is expressed in the heart at 24 hpf to 36 hpf. It could be hypothesized that zfactc1c is needed during the first heart field during the development of the zebrafish. In addition, previous studies have shown that zfactc1a and zfacta1b are localized in the heart further supporting my results (Glenn et al., 2012; Bartman et al., 2004; Sztal et al., 2018; Bertola et al., 2008). Bertola et al., (2008) using WMISH further demonstrated that at 4 dpf, zfacta1b is only present in the somatic muscle.

On the other hand, qPCR analysis confirms that all zfactc genes continue to be expressed at different stages, with zfactc1a and zfacta1b following similar levels of expression at each stage. Additionally, at each stage, there was no significant difference between the level of expression between the zfactc gene except from 60 hpf, where zfactc1a was significantly higher than zfacta1b and zfactc1a. These results are in line with parallel research done in Bryson Richardson’s lab where they looked at the expression of zfactc1a and zfacta1b, along with other actin genes except for zfactc1c. Their qPCR data showed that zfactc1a is the predominant isoform in the heart in the adult stage (Sztal et al., 2018).

The WMISH and the RT-qPCR data showed two different result outputs. The WMISH suggest that all the zfactc genes are present at the beginning of embryo development but slowly shut down at certain stage of development (Figure 9 & 10). The qPCR data shows that they are all expressed at similar levels at all stages of development (Figure 10). The reason for these two different results obtained from two different analyses is likely due to the technique used to obtain the results. Using in situ hybridization, researchers use synthesized RNA probes to localize at the 3-untranslated region (3’UTR). The hybridization of the probe to the complementary RNA is usually performed in a one to one manner where the RNA probe is bound to one transcript. This
could reduce the sensitivity of the in situ hybridization result since other mRNA transcript can be formed due to splicing. On the other hand, to increase the efficiency of the primers binding to the cDNA during RT-qPCR, primers are designed at any section of the gene but specific to the gene of interest. Due to how highly conserved the zfactc genes are, it is possible for the zfactc primers like zfactc1a primers to mis-prime and amplify the other zfactc gene. This might not be the case for my data as only one melting peak was produced during amplification. Nonetheless, there are different ways in which proteins can be regulated; one such way is regulation by the 3’UTR. The 3’UTR contributes to stability and intracellular localization of the transcript that helps in regulating the expressions of protein (Perrin and Ervasti, 2010). WMISH experiments are performed in fixed embryos where the proteins are crosslinked to preserve the DNA and RNA of interest. Using proteinase K to digest the protein reveals the RNA, making it accessible for the RNA probes to interact. It is possible that transcription factors like microRNA (miRNA) can bind to the 3’UTR at a certain stage, making it inaccessible for the RNA probes to bind during hybridization. 3’UTR in genes usually contain miRNA response elements that miRNAs can bind to thereby repressing the translation of the mRNA (Hu and Coller, 2012). This, however, does not mean that the mRNA undergoes degradation immediately after miRNA interaction.

My results support a cardiac gene switch model where all the zfactc genes are present at the beginning of embryo development but shut down as development continues. The expression of zfactc1c occurs in the heart before 2 dpf, zfactalb occurs before 3 dpf, zfactcla continues to be expressed after 3 dpf. In eukaryotes, cells need synthesized proteins to function correctly. Gene regulation occurs at any stage of gene expression, and it determines how much protein needs to be made and when to stop making the proteins. Cells express different genes in response to
external and internal factors (Cooper, 2000). It is not efficient for a cell to express all genes at the same time since the expression process requires energy and space. The regulation of genes usually takes place at the transcription level, but genes can be regulated past the transcription stage (Cooper, 2000). From my experiment, it could be inferred that the expression of actin is regulated at the RNA level (post-transcription), and since there are similarities in the amino acid sequence of the zfactc gene, an overlap in function is not unexpected. What triggers the cell not to express certain actin proteins at the RNA level is yet to be determined. It has been hypothesized that the co-expression of different actin genes may be necessary to accelerate the expression of actin protein at the early stage of development as cells can rapidly produce a large amount of protein when necessary (Bertola et al., 2008; Alberts et al., 2002). Mechanical and biological stress, cell density, and cell-cell interactions are some external factors that may result in the regulation of the actin genes. Also, internal factors like cis-acting and trans-acting regulatory elements of the promoter regions may be involved (Bertola et al., 2008).

In humans, although ACTC is the main isoform in the adult human heart, half of the total striated muscle actin mRNA content consists of α-skeletal actin and ACTC mRNA makes up 5% of the total striated muscle actin in the adult muscle tissue (Gunning et al., 1983). Lastly, expression profiles of cardiac-specific genes were examined from the different chambers in the adult zebrafish heart using RNA Seq (Singh et al., 2016). Scanning through the list, all three zfactc genes in this study were present in the ventricle, atrium and bulbus arteriosus. These findings support the expression data from the RT-qPCR that was obtained in this project.

The large variability in my qPCR data may stem from loss of sample during the extraction process, dilution error and lack of stability of the housekeeping gene used during analysis. For
example, when analysing the qPCR data for my adult hearts, one of the samples had a low Cq value for my housekeeping gene (rpl8) while the Cq value in the zfactc gene in that sample was similar to the other samples. The difference in the Cq value in my one sample might be because the heart that was extracted from the adult zebrafish could have been from a different gender from the other two samples. Although rpl8 was not tested, McCurley and Callard, (2008) showed that some housekeeping gene’s level of expression differ depending on the tissue type and sex of the fish. Using a second housekeeping gene along-side the rpl8 like elongation factor 1a (eflal1), which has been shown to be stable in both development and tissue type regardless of gender, to analyse the expression levels of the zfactc gene in a more robust fashion. In addition, increasing the sample size so as to remove outliers may also reduce error.

To conclude, although the RT-qPCR does not support the cardiac gene switch model, the in situ hybridization results support the model and depicts an accurate picture of what is occurring in the developing zebrafish. Understanding the expression of the zfactc genes in their morphological context is needed to understand their function. A transcription factor like miRNA may be bound to the 3’UTR making it inaccessible to the RNA probes thereby repressing the translation of the mRNA in the embryo in that stage.
Figure 9: Schematic illustration of the proposed expression of zfactc genes in the developing heart of an embryo. zfactc genes are expressed in a spatial and temporal manner. Following the expression of zfactc1a shown in red, in the initial heart tube (I), all zfactc genes are expressed. At II, all zfactc genes are suggested to be expressed in this manner as revealed from the in situ hybridization of zfact1c. At III, zfactc1a (red) and zfacta1b (orange) are enriched in the ventricle but zfactc1c is no longer expressed. At IV, zfactc1a continues to be expressed filling up the outer edges of the atrium, and at this stage zfacta1b is expressed in both chambers. At V, zfactc1a continues to fill up in the atrium, while zfacta1b slowly diminishes from both the chambers. At VI, zfactc1a expression fills up the atrium and zfacta1b ceases to be expressed in the heart.
Figure 10: Schematic illustration showing the results obtained from the RT-qPCR compared to the results obtained from WMISH. RT-qPCR shows that all the zfactc genes are still being expressed after 3 dpf while WMISH shows that there is little to no expression.
4.4 Mild Phenotype Observed in CRISPR Injected Embryos May be due to Genetic Compensation

To analyze the function of zfactc1a in zebrafish, I injected the CRISPR sgRNA targeting the zfactc1a gene with the Cas9 protein to examine the outcome. Mild phenotypes such as slow development, edema and blood pooling were observed along with rapid movement of the embryo, as it was difficult to catch them in the Petri dish and wells. These cardiac phenotypes improved during the 7 days of analysis. In addition, three different heart rate trials were performed to observe the cardiac function of the CRISPR injected embryos. Although the three different heart rate trials of the CRISPR injected embryos showed a significant difference in some days, the heart rate trials followed the same pattern as the mock injected embryos. These data suggest that compensation from other zfactc genes may be occurring. The heart rate of the zebrafish embryo continues to increase as the embryo develops (Bakkers, 2011). This takes place to make sure that every organ or tissues have enough blood supply for its proper function. It has been hypothesized that the reason for the co-expression of actin isoforms in the heart and muscle during embryonic development is to accelerate the expression of actin (Bertola et al., 2008). It is possible that if the expression of zfactc1a dropped in the CRISPR injected embryo, the other zfactc genes could compensate for this reduction.

The mechanism of genetic compensation is not a new idea. It is, however, unclear how genes compensate for the loss of other genes. An upregulation of ACTA1 is observed in the hearts of patients with HCM compared to patients with normal hearts, even though ACTA1 is downregulated in the late development of the human heart (Lim et al., 2001). It is possible that zfactc1b, an α-skeletal gene in zebrafish, could compensate for the insufficient amount of
Sztal et al., (2018) demonstrated that injecting zfactc1a morpholinos into the embryos of actc1b−/− produced a much more severe phenotype such as loss of striation in the muscle fibre and reduced motility than just a homozygous knockout of actc1b−/−.

In addition, it is likely that zfacta1b can compensate for the loss of zfactc1a. Previous graduate student Love Sandhu performed CRISPR injections targeting zfacta1b and identified severe cardiac phenotypes before 3 dpf, but after 3 dpf, there was an improvement in the cardiac phenotype (Sandhu, 2019). It is possible that in Sandhu’s experiment and my zfactc1a CRISPR experiment, genetic compensation could occur between the two genes.

Upon further examination of the CRISPR injected embryos, HRM analysis suggest a genotype mosaic phenotype where WT DNA and mutant DNA were present. The sgRNA and HRM primers were designed to target both chromosome 19 and 20. It is possible that not all the cells were mutated in both chromosomes and the cells that are considered WT could be sufficient enough to make the functional protein (Ablain et al., 2015). This could be a reason why there was an improvement in the cardiac function of the CRISPR injected embryos. Another feasible reason for the improvement of the cardiac phenotype may be because the cells that were mutated may be dying off due to competition from the WT cells, and the WT cells are duplicating producing more functional protein to help in the development of the embryo. Cell competition has been shown to happen in Drosophila where depending on the type of gene, either the WT cell or the mutant cell can outcompete each other (Baker, 2017). For example, WT cells for the troponin gene in Drosophila can outcompete the troponin mutant cells (Baker, 2017).

HRM is a highly sensitive and specific analysis used to identify nucleotide variations as low as one nucleotide even in a mix of WT DNA (Li et al., 2010). I hypothesize that the Sanger
sequencing did not detect any mutations because of the amount of WT DNA was much higher than mutant DNA. Sanger sequencing has its limitations as it is unable to detect mosaic mutations below a threshold of 15% to 20% (Rohlin et al., 2009). The INDIGO analysis suggests there are SNV in each analyzed embryo but the signal on the chromatogram was low compared to the WT.

5 Conclusion

Overall, Tol2 transposons could not generate transgenic zebrafish expressing human ACTC variants that cause cardiomyopathy. Although Tol2 transposons are a great tool for insertional mutagenesis, due to transcription repression after several generations, transposons may not be an ideal technique for ectopic gene expression. A stable zebrafish line is needed that will continue to express the protein for several years. If we want to use this tool in the future, however, cryopreserving the gametes would be necessary so that the line would not be lost.

Another way of editing the zebrafish genome is through gene targeting by using CRISPR. We can perform site directed mutagenesis on the endogenous zfactc gene and observe the impact the point mutation has on the health of the zebrafish. This would overcome issues with ectopic overexpression of an exogenous gene and actin gene compensation. Therefore, knowing the function of each zfactc gene and when it is expressed is crucial so that we can target the right gene. Although the RT-qPCR results showed a continual expression of all zfactc mRNA in the developing zebrafish, the in situ hybridization data contradicts this finding. The in situ hybridization data gives an accurate picture of what is occurring in the cells of the developing embryo. The embryos are fixed at certain stages and the proteins are crosslinked to preserve the
cells containing the RNA and DNA. \textit{zfactc1c} ceases expression after 36 hpf while expression of \textit{zfacta1b} terminates at 3 dpf. The expression of \textit{zfactc1a} seems to continue expressing after 3 dpf. Hence, the results from the \textit{in situ} data supports a cardiac gene switch model (Figure 9).

We can take advantage of understanding the expression pattern of the \textit{zfactc} genes to use CRISPR to target the earlier gene, \textit{zfacta1b}, for site directed mutagenesis and perform different experimental analyses to see the effect the point mutations have on the embryo. \textit{Zfactc1a}, although expressed at later stages of development, is located on two different chromosomes. Therefore, one would have to make sure that both chromosomes are affected. Therefore, \textit{zfacta1b} would be the best candidate for site directed mutagenesis using CRISPR.

Lastly, compensation was believed to occur in the CRISPR injected embryos targeting \textit{zfactc1a} but it is unclear if the compensation is from the other \textit{zfactc} genes or the CRISPR caused a knockdown of the \textit{zfactc1a} gene specifically in the cardiac muscle cells.

6 Future Directions

To verify if \textit{zfactc1a} is still expressed in the zebrafish heart after 3 dpf, future experiments can look at \textit{in situ} hybridization of later stages using tissue sections, since it is harder to treat the tissues with proteinase K at 5 dpf. Tissues sections can also be done with the adult heart and localization of \textit{zfactc1a} can be identified.

I was not able to determine which \textit{zfactc} compensates for the loss of the other; therefore, future work will have to elucidate this. Performing \textit{in situ} hybridization on the CRISPR-treated embryos targeting the different \textit{zfactc} gene should be performed since it gives an accurate picture of what is occurring to the cells as the embryo develops. Designing \textit{actc1b} probes would also be
necessary since there is a possibility of the gene taking part in the compensation mechanism. Furthermore, since standard Sanger sequence could not pick up any signal from the mutant DNA, MiSeq or parallel sequencing can be performed to see if there are any mutations in the CRISPR region of interest.

Lastly, *zfacta1b* appears to be the best candidate for gene target mutagenesis since it is only located in one chromosome. Using CRISPR to create point mutations in the gene of the zebrafish sgRNA will need to be carefully designed to target the right section in *zfacta1b*, along with a repair template to activate the HDR mechanism in the zebrafish genome. Performing this, we can analyse the cardiac function and the impact of the point mutation causing cardiomyopathy for the first 3 days of zebrafish development. Targeting the *zfacta1b* gene will give us an early indication of problems that could impact the health of the embryos.

Developing zebrafish as an *in vivo* model for cardiomyopathy will be beneficial for the cardiovascular field. Therapeutic drugs created to treat patients with cardiomyopathy can be tested on zebrafish to see the impact it generates on the health of the organism. In addition, creating zebrafish expressing mutations that cause cardiomyopathy can provide more insight on how cardiomyopathy brings about the disease state, thereby manufacturing better therapeutic drugs to treat the cause instead of the symptoms.
7 References


Cooper, G.M. 2000. Regulation of transcription in eukaryotes - the cell - NCBI bookshelf.


Heart failure | Heart and Stroke Foundation.


Sartorelli, V., M. Kurabayashi, and L. Kedes. Mini review muscle-specific gene expression a comparison of cardiac and skeletal muscle transcription strategies.


zgc:86709 | *Danio rerio* gene | Alliance of Genome Resources.