CONSERVED NON-CODING ELEMENT m2de3
DIRECTED GENE EXPRESSION DURING DEVELOPMENT

A Thesis
by
LAITON REID STEELE

Submitted to the Graduate School
at Appalachian State University
in partial fulfillment of the requirements for the degree of
MASTERS OF SCIENCE

May 2017
Department of Biology

CONSERVED NON-CODING ELEMENT m2de3
DIRECTED GENE EXPRESSION DURING DEVELOPMENT

A Thesis
by
LAITON REID STEELE
May 2017

APPROVED BY:

_____
Dr. Ted Zerucha
Chairperson, Thesis Committee

_____
Dr. Cortney Bouldin
Member, Thesis Committee

_____
Dr. Andrew Bellemer
Member, Thesis Committee

_____
Dr. Zack Murrell
Chairperson, Department of Biology

_____
Max C. Poole, Ph.D.
Dean, Cratis D. Williams School of Graduate Studies

Abstract

CONSERVED NON-CODING ELEMENT m2de3
DIRECTED GENE EXPRESSION DURING DEVELOPMENT

Laiton Reid Steele
B.S., Wingate University
M.S., Appalachian State University

Chairperson: Dr. Ted Zerucha

Meis genes have important roles in gene regulation and development. The *meis* family encodes homeodomain proteins that are known to work in combination with other transcription factors such as the Hox and Pbx families to regulate development. The primary purpose of this study is to gain a better understanding of the genetic mechanisms that control the expression of *meis* genes during development. The Zerucha lab identified four highly conserved non-coding elements, named *m2de1-m2de4* (for *meis2* downstream element) that we hypothesize regulate expression of the *Meis2* gene. All four elements have been identified in all tetrapod genomes examined however only one element, *m2de1*, has been identified in teleosts. These elements were identified within a linked gene that is found located adjacent to *Meis2* in all vertebrates examined, which has been temporarily named *M2lg* (zebrafish homolog *zgc:154061*). The role of the *M2lg* gene is not currently known, however in recent research it has been linked to Congenital Dyserythropoietic Anemia which suggests it plays a role in primitive hematopoiesis. In my thesis research, I have characterized the *m2de3*

element to determine if this element is able to direct spatial and temporal gene expression that is consistent with *Meis2* expression patterns over different stages of development. Zebrafish (*Danio rerio*) were selected as a model organism in combination with a Tol2 injection cassette to introduce gene expression constructs containing murine *m2de3* through microinjection at the single cell stage. These constructs use *m2de3* to drive the expression of eGFP, through the *cfos* minimal promoter. *m2de3* driven expression can then be visualized under laser confocal microscopy throughout development. In primary transgenic embryos, *m2de3* directed expression was observed in a punctate expression pattern along the notochord, within motor neurons along the notochord and around the heart and in developing muscle fibers along the trunk of zebrafish embryos. This is also consistent with murine *Meis2* expression. In addition, a stable *m2de3* transgenic zebrafish genetic line was generated and confirmed by genotyping. This transgenic line demonstrates an expression pattern along the entire length of the notochord and within the heart directed by *m2de3*.

Acknowledgments

I would first like to thank my chair Dr. Ted Zerucha for his exceptional support, patience and guidance throughout my time at Appalachian, and I am extremely and thoroughly grateful for the opportunity to pursue science and research in his laboratory. I would also like to thank Dr. Cort Bouldin for his time, encouragement and thorough knowledge of zebrafish development that helped me complete this project. I would like to also thank Dr. Andy Bellemer for his time on my committee and for always having excellent feedback and follow-up questions to better my understanding of both my field of study and research project. My grateful thanks and appreciation goes to all the past and present graduate students of Zerucha lab; Kyle Nelson, Zach Williams, Alicia Ramsaran, Tucker Munday, Tyler Ferrara, Hannah Freundlich, Tray Neilson and Megan Tennant for all their assistance and friendship while finishing this project. I would also like to thank several undergraduate students for their assistance in experiments and fish care, specifically Patrick Bearden, Harlie Walkup, Riley Parr and Mackenzie Trapp. Finally I would like to thank our high school lab intern, Lucy Edy for her admirable help in fish care and experiments.

I am tremendously grateful for the support and funding of Appalachian State University, the Cratis D. Williams Graduate School, Appalachian State University Office of Student Research and the Institutional Animal Care and Use Committee of Appalachian State that made this project possible. I am also extremely thankful for the Biology Department at Appalachian State, and I would like to specifically mention a few faculty and students from

other laboratories. I would first like to thank Mrs. Monique Eckerd in the animal facility for great conversation and her help in zebrafish care. I am also extremely thankful to Dr. Gichuan Hou, who was an amazing resource in the microscopy facility and was extremely helpful in both LSCM microscope training and in improving the quality of my research images. Dr. Mary Kinkel was an outstanding help and source of information in zebrafish care, protocols and in In-Situ Hybridization optimization. In addition, I would like to thank a few other members of the faculty and staff for their support and help in my classes, teaching and completion of my project- Dr. Ece Karatan, Dr. Annkatrin Rose, Dr. Ray Williams, Randy Cockerell, and finally Patrick Sullivan and Carra Parker the biology general laboratory managers. In addition, the friendship and comradery of biology graduate students in other laboratories made the process of graduate school enjoyable and I would like to specifically mention Libby Villa, Caitlin Wotanis, Richard Sobe, Laura Ellis, Mara Cloutier, Austin Harbison, Dakota Goad, Meghan Polzin, Brewer Logan, Jamison Slate, Brandon Smith and Brandon Tate.

I would finally like to thank my extended family and friends for all their endless thoughts, encouragement and support during the completion of this thesis. Finally, I would like to thank Dr. Brian Odom, Dr. Edward Mills and Dr. Acchia Albury and other faculty from Wingate University for cultivating my love of science and academic research.

Dedication

Dedicated to Robinette and Wesley Steele, for their endless support and dedication to my pursuits and endeavors.

Table of Contents

Foreword

The format of this manuscript, figures, tables and references are all formatted according to publication submission guidelines for the peer reviewed journal *Developmental Biology* published by Elsevier. This journal is the official publication of the Society of Developmental Biology and is the most pertinent journal for research in the field of developmental genetics.

Introduction

The expression and regulation of an organism's genes is vital for development as well as the maintenance of body systems and functions. The central dogma of molecular biology outlines how DNA is transcribed into RNA which is in turn is translated into protein final products, but gene regulation enables this simple process to produce a wide variety of protein products from the same template DNA sequence. Regulatory mechanisms must accurately affect only the correct gene targets to specific locations in the organism at specific times in development; for example in developmental body patterning where exact gene activation and patterning of multiple genes is necessary to determine body axis and segmentation. The process of gene regulation is also able to respond to and be inducible by internal signaling, feedback loops and external environmental stimuli for development of the organism. The complex level of regulation maintained in organisms requires interaction between multiple regulatory genes within the DNA itself. This task is undertaken with specific regulatory proteins and general transcriptional machinery interacting to achieve complete combinatorial gene control (Bhattacharjee et al., 2013; Venters and Pugh, 2009).

DNA in all organisms exists as a flexible double helix structure. In eukaryotes, DNA is organized into chromosomes made of repeated chromatin complexes which are in turn constructed from nucleosome units. Nucleosomes are 147 nucleotides of DNA material wrapped around eight histones proteins that make up a barrel or spool shape that organizes the DNA, and histone structure and function is highly conserved across eukaryotic species. There are four different types of histone proteins, labeled H1 through H4 and each subunit is found twice in a single histone for a total of 8 per histone. In addition, each histone subunit has a protein structural "tail" domain that is accessible to modification at a selection of sites.

The wide variety of modifications essentially function as a key or code, that cause differential changes in histone structure affecting the availability of the genetic material located on the nucleosome. Some examples of common histone modification include methylation, acetylation, and phosphorylation. Specific combinations of different modifications have been shown to act as signals for function as a repressor or activator. For example, the H1 histone subunit has been demonstrated to serve as a sectional repressor, limiting expression of its target genes when the hormone progesterone is present and also while under selective phosphorylation or methylation. On the other hand, trimethylation of the H3 subunit can act as either a repressor or activator merely by changing the location of the methyl groups on the amino acids within the histone tail. (Crane-Robinson, 2016; Ferraiuolo et al., 2010; Jenuwein, 2001; Khochbin, 2001; Ng et al., 2016; Vicent et al., 2016).

During embryogenesis, many developmental control genes are found to be positioned in clusters within open chromatin regions between histones with exposed promoters for ease of transcription. RNA polymerase II is often present at the exposed promoters of key developmental genes ready for the start of transcription without being active, which contradicts the normal understanding of RNA polymerase recruitment but enables faster gene transcription for time sensitive expression. This “poised” or bivalent state is suspected to be maintained through the combination of both activator and repressor mechanisms that holds the promoter in an inactive but ready state (Muse et al., 2007; Zeitlinger and Stark, 2010).

Once a DNA sequence is open and in an exposed chromatin state, gene transcription can then take place. RNA polymerase II (RNAPII) is the most common transcriber protein of DNA in eukaryotes, but RNAP II requires general transcription factors (GTF) to recruit it to

the template DNA sequence (Venters and Pugh, 2009). General transcription starts with transcription factor II D (TFIID), a large complex containing a TATA binding protein that attaches to TA-rich sequences within the promoter and causes a bend in the DNA strand. TFIID then serves as a stable site for the binding of other general transcription factors. Transcription factor II B (TFIIB) is next recruited and binds directly to TFIID. Once these two interact, transcription factor II F, factor II E and factor II H bind together respectively to form a zinc ribbon complex which increases the stability of the complex. The entire complex then acts to recruit the RNA polymerase, which binds within the transcription starting position alongside TFIIE and TFIIH. Finally, on the opposite side of TFIID, transcription factor A (TFIIA) attaches to stabilize the complex to the DNA strand. This process can be facilitated by gene-specific activator proteins bound to *cis* elements. These proteins that are bound to a *cis* element interact with the general transcriptional machinery via DNA looping and can increase or decrease the rate of transcription. An additional role of the GTF TFIIH is to loosen DNA from tightly wound histones and for weakening DNA strand binding in preparation for helicase activity. (Buratowski et al., 1989; Conaway and Conaway, 1993; Dvir et al., 2001; Liu et al., 2013; Roeder, 1996; Venters and Pugh, 2009).

DNA transcription can then occur moving along the gene sequence generating an unprocessed complimentary mRNA strand, which can then be post-transcriptionally modified before being sent for translation. First, RNA is capped with a modified guanine that prevents RNases from degrading the 5’ end of the mRNA. This cap has also been shown to be necessary for splicing initiation and as a signal tag for direction of mRNA to the ribosome. Second, polyadenylation of the 3’ end of the mRNA strand occurs when a 250 nucleotide section of adenines are added to the tail end of the sequence which slows sequence damage

by units of RNase since the polyadenylation must be removed first and helps with translation initiation. Finally, mRNA must be accurately spliced to remove introns or non-coding regions from the transcript.

Splicing is driven by the looping and production of lariats which are then spliced out (Amrani et al., 2008; Klerk and Hoen, 2015; Topisirovic et al., 2011). Disorders involving incorrect splicing have been linked to a selection of diseases related to neurodegeneration, genetic and developmental disorders and cancer in humans. Evidence has even shown that as a preventive measure of protection against splicing error, a small subsection of vital growth and developmental genes will not contain introns (Grzybowska, 2012). Splicing is further complicated in that splicing order of the exons together can be modified to produce different forms of processed mRNA from the same gene template. This process is termed alternative splicing, meaning a single genes transcript can generate multiple versions of the final sequence product that can alter the final splicing pattern of the processed exon mRNA or even act as splicer enhancers or suppressors. This process was expanded upon through the ENCODE genome sequencing project which examined splicing patterns on a large scale. The final result of this study is known as the spliceosome, which identifies the multiple possible products of the genes in the genome (Blencowe, 2006; Consortium 2007; Klerk and Hoen, 2015; Wang et al., 2008).

Transcription and translation are highly dependent on gene regulation controlled by elements of the DNA sequence which are termed *cis*-acting elements. These elements are usually short sequences associated with the sequence to be transcribed that serve as binding sites that can be located upstream or downstream from the target gene. The most common of these sequences are the previously discussed promoters, which serve as the starting position

for transcription machinery of their specific target gene. Promoters are commonly found located within 40 to 50 base pairs upstream of the transcription start site and are extremely specific to their gene target, but a single gene can have multiple promoters for different gene products. In addition to the proximal promoter, long distance distal promoters have been identified that require DNA looping to interact with the core promoter, but still bind in a similar fashion as the proximal promoters with the initiation complex.

The promoter's role in gene regulation is to serve as a binding site for the pre-initiation complex where transcription of a gene starts. This role is based on a selection of binding sites that are recognized by the general transcriptional machinery and specific transcription factors together to activate transcription. Originally, promoters were subdivided into major categories depending on specific nucleotide sequences found within them (Carnicnci, 2006; Cooper et al., 2005; Li et al., 2015). The first category are TATA box enriched promoters, which occur in around 25% of human core promoters and around 33% of vertebrate promoters. However, the TATA sequence promoter is the most common overall when all organisms are considered (Gagniuc et al., 2012, 2013). The second category is CpG-rich promoters, which are sequences containing cysteine guanine nucleotide pairs, and are prevalent in housekeeping genes. CpG binding sites are of particular interest in gene regulation because the addition or removal of a methyl group to a CpG site affects gene expression, with demethylation of a gene possibly even leading to overexpression. However, further research showed that promoters of the same type could respond differently depending on location, orientation and different gene interactions associated with the new location. This led to comprehensive examination analyzing promoter physical structure, gene sequence and nucleotide count to further differentiate the promoter categories, resulting in a minimum of

ten promoter category groups which are highly conserved even across species (Gagniuc et al., 2012; Gagniuc et al., 2013). Promoter function also results from location, for example promoters on the same chromosome have even been shown in humans to exhibit small chromosome specific structural similarities. These chromosome similarities are present regardless of gene function or cell type, leftover from evolutionary dynamics during the formation of chromosomes and what genetic material is available for transcription and accumulation of mutations. (Bhattacharjee, 2013; Cooper et al., 2005; Gagniuc et al., 2012, 2013; Li et al., 2015).

While promoters serve as the staging grounds for the initiation of transcription, interactions with the correct enhancer are required for optimal transcription. An enhancer's function in the genome is to modify the strength of affinity of promoters, which allows them to directly control regulation of a gene by affecting the accessibility of the target gene to both general and specific transcription machinery. Enhancers were originally discovered within the simian virus genome and function regardless of location or orientation in relation to the promoter. Enhancers positively affect transcription rates through direct binding with promoters in combination with activator proteins that can interact with the promoter. The promoter must be in accessible chromatin regions to the enhancer, but enhancers are commonly found some distance away from their gene promoter target. (Arensbergen et al., 2014; Banjeri et al., 1981; Bulger and Groudine, 2011). For example, the farthest distance an enhancer has been identified from its associated promoter is close to 1 million nucleotides away from the developmentally important gene *Sonic hedgehog* (Lettice et al., 2003).

One example of developmental enhancers are those associated with the gene *Sox10*. *Sox10* is an extremely important developmental gene and is expressed during sex

differentiation as well as within regions of neural, optical and skeletal development. *Sox10* is also shown to be involved with the manifestation of Hirschsprung disease, one of the most common developmental defects in humans (Bondurand and Sham, 2013; Guth and Wegner, 2008). The gene *Sox10* is indirectly controlled by the enhancer Sox10E2 and is activated through binding with Sox8, Pea3 and cMyb genes to positively affect transcription. Expression of the gene is directed to different locations by Sox9, Ets1 and cMyb (Betancur et al., 2011).

Recently, a second specialty class of enhancers has been described. Shadow enhancers currently appear to be specialized secondary enhancers which drive expression in the same expression pattern and are bound by the same transcription factors as the primary enhancer. Originally identified with *brinker* and *sog* genes in *Drosophila*, shadow enhancers are extremely common within developmental genes. 70% of shadow enhancer sequences have more than two shadow enhancers, with the highest number discovered so far being five for *Traf1* (Cannavò et al., 2016). Shadow enhancers are not fully understood but seem to serve as a measure of redundancy and robustness of critical genes in regulation and development against any changes that may occur within the genome. There are multiple hypotheses on shadow enhancers and why they remain so prevalent within different genomes. One theory suggested by Cannavo et al. suggests that multiple enhancers have evolved to increase expression of genes whose products are required in high quantities. (Barolo, 2012; Cannavò et al., 2016).

Repressors and silencer elements serve the opposite of the role of enhancers, acting to downregulate or stop the transcription of a target gene. The well-known canonical example of gene silencing was the discovery of the lac-operon by Jacob and Monod. The enzyme

needed for the breakdown of lactose is only produced when lactose is present as a food source, and this is accomplished through a gene silencer. When no lactose is present, the silencer *I* gene produces a repressor protein that is able to bind to the operator region that enables enzyme production, and prevents the RNA polymerase recruitment and thus transcription of the gene. However, if lactose is present, the repressor protein is able to bind to lactose instead of binding to the repressor and thus enabling production of beta-galactosidase.

Silencer elements commonly target gene promoters or general transcription machinery, while a subset of silencers in specific cases can instead target other transcription factors specific to the gene of question. (Arnosti, 2011; Lewis, 2011; Tajbakhsh et al., 2011). Silencer elements must be accessible within the chromatin and intron exon location of a silencer has been shown to be both a limiting factor to either gene or to interrupting active transcription. Repressors demonstrate a wide variety of different functions, although most silencers are orientation and promoter dependent. Certain silencer elements can even function in both read directions due to specialized inverted repeat sequences such as human Aromatase Silencer S1 (Zhou, 1998) Silencer elements are also specific to application and transcription state; some can only affect the gene when it's not actively being transcribed while others can interrupt expression completely at any time, such as the *sog* gene and the Snail silencer element which is capable of stopping active expression of *sog*. Silencers are also vulnerable to mutation and play a role in disorders such as Huntington's disease and are a target for a variety of common cancers and tumors. Silencer elements are vital for development genes that must quickly stop transcription and also inhibit other genes downstream of primary development genes.

Because genes can be located within close proximity to other genes with different expression patterns and enhancer specificity, a mechanism of isolation and separation is necessary between genes. Insulator regions serve as boundaries between different genes directing vastly different expression patterns and functions, preventing miscommunication. Insulator regions perform this role by mediating chromatin availability and DNA looping through insulator protein interactions that allows them to selectively control the separation of genes (Ali et al., 2016; Burgess-Beusse et al., 2002; Grosveld et al., 1987; Kyrchanova and Georgiev, 2014; West et al., 2002). For example, modification of the *HoxA* HK3K27me3 insulator domain shifted the active chromatin domain, moving a methylated unavailable gene area and enabling the activation of a different more caudal *Hox* gene in the original expression area (Narendra et al., 2015). In addition to preventing access to promoters, in certain cases insulators can also interact directly with promoters and can be necessary to enable gene transcription. The presence of the Wari and 1A2 insulators was shown to be necessary for the target *white* promoter and subsequent *yellow* gene transcription through the formation of a promoter-insulator loop (Erokhin et al., 2011; Kyrchanova and Georgiev, 2014). Insulators can also be positioned between a promoter and enhancer, serving as an enhancer binding blocker. This function enables the insulator to block the ability of the enhancer to interact with the promoter. An interesting point of enhancer binding blockers occurs due to the polarity that these insulators demonstrate for different effects. For example, an insulator can block transcription in one direction and allow transcription in the reverse direction. One example of this is the upstream *digit* and downstream *hernia* insulators and how they regulate access of various *HoxD* genes. In normal orientation, *digit* is able to access *Hoxd10* through *Hoxd13* while *hernia* can only access *Hoxd10* through *Hoxd12* by

physically blocking access across *Hoxd12* in one read direction. Conversely, by inverting the insulators direction within the sequence, *digit* can only access *Hoxd11* through *Hoxd13* while *hernia* can access *Hoxd10* through *Hoxd13* since *digit* is blocked, demonstrating how gene positioning and insulator interactions are important. This is even more important within the context of development, as *Hox* genes are positioned within the genome in the order in which they are expressed within the different body sections, meaning the positional access of *Hox* genes can change the entire body plan. (Burgess-Beusse et al., 2002; Duboule, 1998; West et al., 2002.)

The development of an organism from zygote to adult is directed by an intricate hierarchy of genes that interact together to generate the specific and accurate expression patterns that outline the body's structure and organize precursor structures for correct development. The careful regulation of body plan patterning is entirely dependent on the correct determination of the anterior-posterior and dorsal-ventral axis of the body. This "blueprint" patterning occurs in a distinctive order and in relation to other developmental expression patterns which serves as a template precursor for all forthcoming structures (Delgado and Torres, 2016; Gilbert, 2000; Sander, 1975). Once these patterns have been established, development can advance properly through key developmental milestones such as gastrulation, neurulation, and organogenesis to generate an adult organism. These functions of gene regulation are also tightly regulated temporally and spatially as they develop and maintain an organism in conjunction with a wide variety of transcription factors. It is also well demonstrated in many model organisms that disruptions of these patterns and developmental stages can lead to a variety of disorders and mutant phenotypes. (Arnosti, 2011; Goulding and Gruss, 1989; Zeitlinger and Stark, 2010).

During embryonic development, one of the most important and intensively studied developmental genes are the homeobox gene family. All homeobox genes contain a characteristic homeobox DNA sequence approximately 180 base pairs in length. The homeobox region is a highly conserved sequence encoding a homeodomain binding site region of approximately 60 amino acids that contain a three alpha helix domain including a helix-turn-helix DNA binding motif with an N-terminal arm on one end (Desplan et al. 1988; McGinnis 1984). The homeodomain DNA binding domain allows the regulation of target genes, often containing TAA or TAAT binding sites (Banerjee-Basu and Baxevanis, 2001; Brennan and Matthews et al., 1989; Desplan et al., 1988; Duboule, 1998; Gehring et al., 1994; Goulding and Gruss, 1989; Krumlauf, 1994; Ruddle, 1985). Homeobox genes were discovered in the 1980's during research about developmental patterning and segmentation patterns in *Drosophila melanogaster*. Homeobox genes were found to be highly conserved across multiple organisms and even the non-coding DNA flanking regions share remarkable conservation across species in multiple cases (Goulding and Gruss, 1989; Holland and Hogan, 1988; Jaynes and O'Farrell, 1988; Lawrence and Morata, 1976; McGinnis et al., 1984; McGinnis and Krumlauf, 1992; Ruddle et al., 1985). Abnormal disorders and phenotypes have been demonstrated to be caused by mutations to homeobox genes, and across multiple species. (Lawrence and Morata 1976, 1994).

One subset of Homeobox genes is the aforementioned *Hox* family of genes which are regulators that pattern the position of the anterior-posterior axis and organs through precise gene regulation (Amores et al., 1998; Akam 1989; Goulding and Gruss, 1989; Krumlauf, 1994; McGinnis and Krumlauf, 1992). *Hox* genes or homeotic genes were originally identified due to their ability to generate mutant phenotypes in *Drosophila melanogaster*

where one body segment exhibited the identity or characteristics of another segment. The classical example of these mutants is *Antennapedia*, which causes the wild type antenna located on the head of the fly to form legs (Lewis, 1978; Mallo et al., 2010; McGinnis et al., 1984, McGinnis and Krumlauf 1992; Nüsslein-Volhard and Wieschaus, 1980). *Hox* genes are clustered in the genome and exhibit collinearity; *Hox* genes are found clustered in the genome where the order of expression pattern reflects their genomic organization where genes at the 3' end of the cluster are expressed earlier and more anteriorly than genes located more 5' (Deschamps, 2007; Duboule, 1998, 2007; Krumlauf, 1994; Mann, 2005; McGinnis and Krumlauf, 1992). This genomic arrangement is found in all metazoan species examined and research has found a wide variety of disorders and abnormal phenotypes associated with mutant variations of selected *Hox* genes (Duboule, 1998; Graham et al., 1989; Krumlauf, 1994; Quinonez and Innis, 2014). Hox proteins function as transcription factors to preferentially bind DNA and regulate gene expression, however Hox proteins demonstrate poor binding specificity and stability of binding. Hox proteins overcome the problem of non-specific binding by forming complexes with other transcription factors that increase specificity and affinity enabling accurate gene control (Gehring et al., 1994; Slattery et al., 2011). This combinatorial binding of their targets through multiple factors to overcome non-specific binding has been termed "latent specificity" (Slattery et al., 2011). The interaction with other protein cofactors is key to optimal binding, most notably with members of PBC and *MEIS* proteins, inside of the TALE family. (Chang et al., 1996; Jacobs et al., 1999; Mann and Chan, 1996; Moens and Selleri, 2006).

One group of the homeobox genes are the TALE (**T**hree **A**mino Acid **L**oop **E**xtension) homeodomain superclass family of which feature an additional three amino acids

located within the loop of the helix loop helix binding domain, found between helix 1 and 2 (Bertolino et al., 1995; Bürglin, 1997). This family was categorized by Burglin in 1997 due to highly conserved upstream DNA motifs and currently consists of the *Meis*, PBC, IRO and TGIF families in animals and the KNOX and BEL in plants. Burglin et al. also identified that modifying a single amino acid completely changed the binding interaction potential of the homeodomain, and that a variety of amino acids can interchange within the same spot while preserving functionality. For example, TALE homeodomains will commonly contain a smaller non-polar amino acid residue in position 50 in the homeodomain sequence that other homeodomains do not, which allows TALE to bind DNA-protein interactions that other homeoboxes cannot. This versatility in binding restriction increases stability, particularly by forming dimers and trimers with other TALE members such as the canonical relationship between Hox, PBX and *Meis*. (Bürglin, 1997; Jacobs et al., 1999; Merabet and Lohmann, 2015).

The *Meis* gene family was originally discovered in a murine model during efforts to study the Myeloid Ecotropic Leukemia Virus. Researchers found the virus targeted and integrated itself in a coding region for a then unknown gene, disrupting its expression and leading to its naming: Myeloid Ecotropic Leukemia Insertion Site. Recently, *Meis* genes have been connected with human acute myeloid leukemia and also as targets in in cancer cell proliferation and lifespan as a regulation target through Hox A7/A9 and Pbx1 (Moskow et al., 1995; Nakamura et al., 1996a; Nakamura 2005). *Meis* genes encode multiple separate functional domains, leading to gene control through binding with Meis alone or in combination with one or more factors. The key to this binding ability lies in the structure of the Meis protein product, with three binding site domains; a homeodomain, a C-terminus and

an N-terminus binding sites. The homeodomain serves as the primary DNA binding site and includes the aforementioned helix-turn-helix motif. The C-terminal binding site serves to bind to other Hox protein products and Meis similar transcription factors. Finally, the N-terminus otherwise known as the Meinox domain binds specific transcription factors such as Pbx. This is key for the production of dimers and trimers, as the homeodomain can be securely bound to DNA and the C-terminus is still accessible to other transcription factors (Bürglin, 1997; Geerts et al., 2005; Jacobs et al., 1999; Moskow et al., 1995). *Meis*2 also shows extensive alternate mRNA splicing, enables even further variation while maintaining highly conserved intron and exon regions. For example *Meis*2 has four different alternative splicing variants which all change the C-terminal structure and are identified as *Meis*2A-*Meis2*D. A *Meis*2E splice variant has also been found, but is missing major parts of the homeodomain structure, meaning that it can still interact with proteins through the Meinox domain but cannot interact directly with DNA. *Meis*2E possibly plays a role in direct negative regulation of normal *Meis2* function, but this needs further research to confirm. (Geerts et al., 2005; Moskow et al., 1995).

Once the importance of these genes was understood, subsequent searches for *Meis* genes were performed in other species. DNA-DNA Hybridization of *Meis*1 homeobox sequences and from *Meis*1 untranslated regions was performed to identify additional genes. Hybridization revealed the same sequence similarity in the untranslated regions within a variety of species including *Xenopus*, humans and mice. However, the *Meis*1 homeobox probe bound to the same sequences as before, but also identified *Meis*2 and *Meis*3, which share the highly conserved homeodomain region but do not share the untranslated regions, meaning that they featured different unique flanking regions (Moskow et al., 1995;

Nakamura et al. 1996a; Nakamura et al. 1996b; Steelman et al., 1997). This research lead to the identification and sequencing of these genes, where it was found that the *Meis*1, *Meis*2 and *Meis*3 genes were conserved among vertebrate model organisms including humans, chickens, monkeys, mice, *Xenopus* and zebrafish. (Cecconi et al., 1997; Moskow et al., 1995; Salzberg et al., 1999; Smith et al., 1996; Steelman et al., 1997; Zerucha and Prince, 2000). In addition, an ortholog called Homothorax or *Hth* was discovered in *Drosophila melanogaster*, which serves a similar role with *Hox* in development and segment patterning. The genomes of teleost fishes such as zebrafish feature the normal *meis*1-*meis*3 however their genomes contain 2 *meis1* genes, m*eis1a* (originally labeled *meis4*) and *meis1b* as well as 2 *meis2* genes, *meis*2a and *meis*2b , due to a proposed genome duplication event in the teleost lineage after its branching from tetrapods (Santini, 2009, Slattery et al., 2011).

*Meis*1 body pattern expression is highly characterized in both development and in specific body systems and disorders. *Meis*1 was originally discovered as a gene target for myeloid leukemia. Since that original discovery *Meis*1 has since been further tied to cancer growth and proliferation, particularly in neuroblastomas (Geerts et al., 2005). *Meis*1 is also more importantly essential for embryonic development and survival due to its extensive role in hematopoietic development. Mutations in *Meis1* in mice embryos were found to be embryonic lethal during mid gestation as a consequence of internal hemorrhaging and anemia from system-wide hematopoiesis failure, and also defects in the neural and sensory regions. Thus *Meis1* plays a role in hematopoietic stem cells and also in differentiation of endothelial cells for the formation of the heart and blood vessels (Azcoitia et al., 2005; Minehata et al., 2008). *Meis*1 expression is also found in early development of somites and the spinal cord as well as in the forebrain, midbrain and hindbrain, and this continues throughout late

developmental stages. Meis1 works in conjunction with Pbx and Prep, another vertebrate TALE gene, to regulate neural stem cell regulation and differentiation in these areas (Barber et al., 2013; Salzberg et al., 1999; Waskiewicz et al., 2001; Waskiewicz et al., 2002). Associated with the neural expression and neural stem cells, Meis1 is also key in eye development as well where it interacts in combination with more than 20 other homeodomain proteins and multiple transcription factors to develop the optical nerves and eyes. When *meis*1 and *meis*2 are knocked down in the eye, microphthalmia or abnormally small eyes and reduced retinal ganglion cells are observed in the mutant phenotype (Zagozewski et al., 2014). *Meis*1 is also known to be selectively expressed in the developing limb buds and skeletal system, most notably in the cranial and branchial arches, where *meis1* knockdowns lead to a missing cranial skeleton as well as in selected cartilage (Amin et al., 2015; Waskiewicz et al., 2001; Waskiewicz et al., 2002).

Zebrafish contain two *meis1* genes, *meis1a* and *meis1b*. Splice variants of *meis1a* were originally classified as *meis4* but subsequent genomic analysis confirmed these to be in fact, be a *meis1* gene. (Waskiewicz et al., 2002; ZFIN Unpublished data reference (ZDB-GENE-02122-3)). Since the nomenclature transition from *meis4* to *meis1a* and *meis1b* is fairly recent, most literature still cites *meis4*. *meis4* was the least characterized among all the *meis* genes and only two papers have been published on research projects involving zebrafish *meis4/meis1b,* which is only found naturally occurring in teleost fish. In a zebrafish *meis* survey, *Waskiewicz et al.,* identified and cloned *meis* 4.1 and 4.2, which are splice variants of the same *meis4.1* gene, but they did not perform any specific experimentation on *meis*4 expression (Amin et al., 2015; Waskiewicz et al., 2001; Waskiewicz et al., 2002). The second research paper by Amin et al. examined a full *meis*1-*meis*4 transcript knockout in the

branchial arch structure of developing zebrafish. With all *meis* transcripts knocked down using morpholino oligonucleotides, this resulted in a mutant phenotype with a minimal / non-existent craniofacial skeletal structure and diminished nervous system (Amin et al., 2015; Waskiewicz et al., 2001; Waskiewicz et al., 2002).

*Meis*2 shares an almost identical homeobox sequence with *Meis1* and tends to share the same generalized areas of expression, but does exhibit some differences in expression patterning. *Meis*2 neural expression is found in the same common areas as *meis*1 in the nervous system, but exhibits a different pattern in the forebrain, midbrain and hindbrain, where it is generally expressed more anteriorly in the head and more posteriorly in the neural tube. *Meis*2 also demonstrates expression early on in the neural plate and at later time points within the somites. Fully sectioned somites show solid *meis*2 expression while developing somites start with very weak *meis*2 expression. (Biemar et al., 2001; Cecconi et al., 1997; Toresson et al., 2000; Zerucha and Prince, 2000). The hindbrain in particular exhibits major *meis*2 expression, possibly due to the large amount of retinoic acid synthesis used in signaling in the hindbrain as it is hypothesized that retinoic acid signaling regulates *Meis* gene expression. This control is used to setup specific expression concentration gradients tied with gene regulation for exact patterning of retinoic acid which is necessary for correct hindbrain development and segmentation (Vitobello et al., 2011). *Meis*1 and *Meis*2 share overlapping expression patterns in cardiac hematopoiesis and hematopoietic stem cells, however *Meis*2 expression was found in the myocardium, cardiac cushions and all four cardiac regulatory valves. Weak *Meis2* expression is also found in the nascent heart tube, a precursor structure in heart development. Mice lacking *Meis*2 die similarly to *Meis*1 mutants with systemic hemorrhaging and anemia but additionally demonstrate an abnormal liver and

small body size (Azcoitia et al., 2005; Biemar et al., 2001; Machon et al., 2015). In humans, a de novo microdeletion of a section of *MEIS2* has been found in infants demonstrating Tetralogy of Fallot, a common congenital heart defect pattern characterized by altered heart chamber size and an incomplete ventricular septum allowing flow between the left and right ventricles (Chen et al., 2016). *Meis*2 also plays a role in retinal development like *Meis1*, being expressed in the entire optic cup. Furthermore, *Meis2* is also involved in cell differentiation and connection of the retinal nerve, and in addition when mutated a reduced cornea and eyelid structure (Machon et al., 2015; Zagozewski et al., 2014). As mentioned previously for *meis*1, *meis*2 is expressed in the developing cartilage and bone structures, with severely repressed craniofacial bones and completely missing hyoid bones and shorter jaw structure when observed in a mutant phenotype (Amin et al., 2015; Machon et al., 2015).

*Meis*3 tends to shows expression patterns in similar areas to *Meis*1 and *Meis*2 but tends to fulfill niche specialist roles in development. In early mesodermal development of the neural plate, a nervous system precursor structure, *Meis*3 is a known downstream target of Wnt3. These two together form a regulatory network that helps stem cell proliferation and acts as a major factor in anterior-posterior patterning in the neural plate, particularly the future hindbrain, cerebellum and hippocampus (Yaniv and Frank, 2010). In a congenital heart defect study, *Meis*3 and *Meis*1 were implicated as a possible target of Pbx3 in patients with congenital heart defects and as a possible risk screen allele when bound with PBX3, with a particular affinity for cardiac outflow tract patients (Arrington et al., 2012). *Meis*3 also serves a role in neural connections to both precursor neural crest cells and the developed nerves to organs such as the spleen and pancreas (Uribe and Bronner, 2015). *Meis3* works in these organs in conjunction with Shh and foxa2, where *meis*3 acts as an upstream regulator of

liver and pancreas organs. When *meis*3 is disrupted in this area, the liver and pancreas are abnormally located and lack the full size and functionality of wild type *meis*3 affected organs. This altered functionality includes hyperinsulinemia and hypoglycemia in both young and adult zebrafish (DiIorio et al., 2007). More recently, *Meis3* has been shown to be expressed within developing and adult epithelial and beta lung cells in mice, and that *Meis3* is required for continued survival in these cells through the PDK1 pathway (Liu et al., 2010). In a related study by Toyotome, fungus sourced mycotoxin Deoxynivalenol repressed expression of *Meis3* in these cells, leading to increased toxicity in the lungs possibly followed by death. In this research A549 cells modified to transiently express more *Meis3* were less susceptible to the toxicity of the mycotoxin (Toyotome et al., 2016).

Meis genes have proven to be extremely important in multiple cellular processes and particularly in organogenesis and homeostasis of organs such as the brain and heart. *Meis* genes as a whole have been well characterized and a significant amount of research has been performed on their molecular function and role in body patterning. To date, however, little is known concerning the complex gene regulation required for these genes to correctly work and how their expression is directed to each location within the organism. The Zerucha lab used a comparative genomics approach to identify conserved noncoding sequences of DNA in proximity to *Meis2* and that contained binding sites for proteins that are candidates for regulating *Meis* gene expression. Four regulatory elements were discovered downstream of *Meis2*, named *m2de1-m2de4* (for *Meis2* downstream element) (Figure 1). These regulatory elements are located within a gene that is consistently located downstream from Meis2 and is temporarily being called *Meis2* link gene or *m2lg*. *m2de1* drives expression in the midbrain and hindbrain as well as muscle fibers along the trunk. *m2de2* expression is centralized

within neurons in the brain and along the entire trunk of zebrafish within all muscle fibers (Ferrara, 2015; Freundlich, 2016;). My project characterizes the regulatory element *m2de3,* which I hypothesize is controlling *Meis* expression and thus provides insights into the regulatory mechanisms controlling expression within this family of genes.

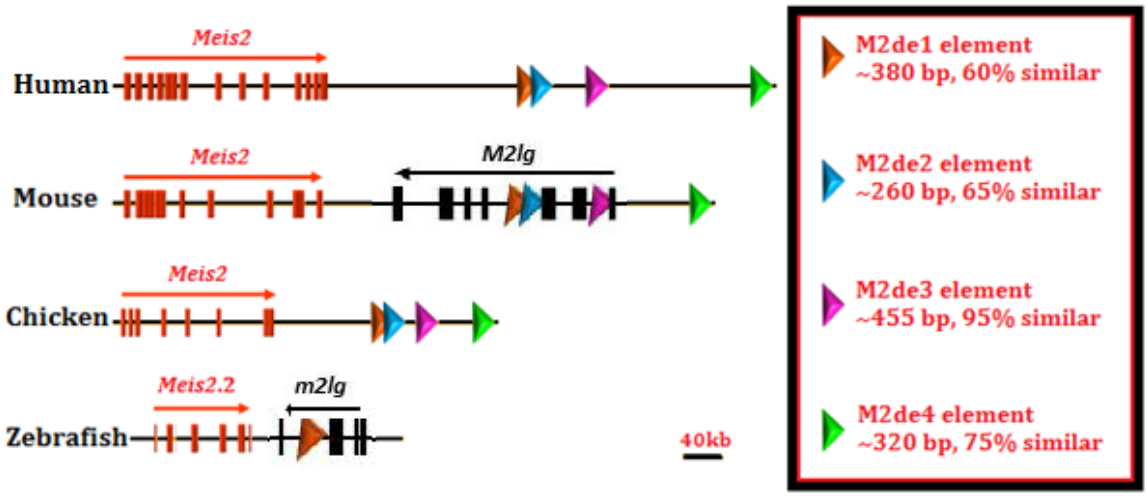


Figure 1- Diagram demonstrating m2de elements size, similarity across species and location in sequence relative to *meis2* and *m2lg*

Materials and Methods

General Zebrafish Husbandry and Care

Adult Zebrafish are housed in an enclosed Marine Biotech Z-Mod system (Aquatic Habitats) that uses a combination of UV and particulate filters to generate a large, constantly running quality water supply to each tank in the system. The fish habitat system is maintained and the wild type / transgenic fish lines are fed and cared for based upon standardized conditions as defined in the Zebrafish Book (Westerfield, 2000) and located on the Zebrafish Information Network or ZFIN (http://zfin.org/). The entire aquatic system is kept on a synchronized day/night light cycle timer for a 14 hour day and a 10 hour night cycle. Water conditions are monitored daily, and kept within optimal conditions with a temperature of 27° C, pH within the range of 7.2-7.5, a water conductivity within 450ppm to 650ppm. The water is also conditioned using crushed coral, which buffers the pH and releases calcium into the water. Feeding consists of a combination of Adult Complete Zebrafish Diet dry food (Zeigler Brothers) and live brine shrimp (Brine Shrimp Direct) fed to adults and a specialized nursery diet (Zebrafish Maintenance Systems Ltd. UK) adjusted based on the size and age of young fish.

Zebrafish pair-wise crosses for experiments are conducted by selecting a breeding tank in the late afternoon around approximately 5:30-9:00pm and then separating the male and female zebrafish in a specialized divider tank with a mesh bottom that fits within the normal tank, which along with a plastic plant is used to simulate the shallow shores where natural spawning would occur. This specialized cross tank is then carefully returned to the system for the night. In the morning when the day cycle starts, the plastic divider is removed

and the fish are allowed to spawn at a controllable time point for around 30 minutes to an hour with the fertilized eggs falling through the mesh bottom of the tank. Embryos are then collected and cleaned from the bottom of the tank for use in experiments. When a specific synchronized time point is not needed, zebrafish crosses for embryos can also be performed by marbling, where two or three layers of clear glass marbles are placed in the bottom of the tank and this can be left for up to two days. Zebrafish will lay eggs normally, and these will fall through the marbles and can be collected. Embryos are collected by straining the contents of the fish tank through a mesh filter from the tank and then rinsing the tank thoroughly two or three times with RO and filtering again to collect all embryos. The mesh is then placed in a glass pyrex dish filled with 0.3x Danieau solution (17 mM NaCl, 2mM KCl, 0.12 mM MgSO4, 1.8 mM Ca(NO3)2, 1.5 mM HEPES, pH 7.6) rinsed with RO water and then the dish is transferred to a 30° C incubator. Optionally, 0.8% Danieau methylene blue solution (10 drops of methylene blue per 500ml of 0.3x Danieau solution) can be added to the dish to hinder growth of any organic material. Embryos are cleaned by hand using a disposable glass pipette and solution water changes are performed daily to keep embryos healthy until experimental time points are reached. After approximately one week, embryos can be transferred to a full size tank in the system nursery.

Generation of HCNE m2de3 Construct

The *m2de3* element was originally isolated from the mouse (*Mus Musculus*) genome using PCR (Nelson, 2011). This element was cloned into the pCR-2.1©-TOPO (Invitrogen) vector plasmid from which the injection plasmid was made. The *mm m2de3* element was subcloned into pGW-cfos-eGFP using Gateway Site-Specific Recombination cloning to

generate the pGW-cfos-mm-*m2de3*-eGFP expression construct, and was stored at -80° C as a glycerol stock in DH5α cells (Ramsaran, unpublished) (Appendix A, Figure A). To produce the injection construct, the glycerol stock is scraped with a P20 pipette tip which is used to transfer bacteria to an Ampicillin (Amp) Luria-Betani Broth (LB) agar plate and cultured overnight at 37° C. The next morning, the plate is examined for growth and single colonies are selected for liquid culture. 40ml glass liquid culture tubes are filled with 3ml of liquid LB broth and 3 µl of 100 ng/ul concentration of Ampicillin antibiotic and then the tube is placed in a heated 37° C gyroshake table overnight. Liquid cultures are removed in the morning and a Wizard Plus SV Miniprep DNA Purification Kit (Promega A1460) is used to isolate the DNA construct. The *m2de3* expression construct is then cleaned using a phenol chloroform extraction to purify the DNA. Phenol chloroform extractions start by raising the reaction volume to 100 µl and then 10 µl of 3M NaAc (pH 5.2) solution is added. 110 µl volume of Phenol-Chloroform-Isoamyl alcohol (25:24:1 v/v) solution is then added to the reaction tube, gently vortexed for one minute to thoroughly mix and then spun in a centrifuge for five minutes at 13,000rpm at 4 °C to separate the aqueous layer. Upon the completion of the spin, two clear layers should be present. If a third white protein interface is present, more phenol-chloroform mixture is added and the tube is centrifuged again. The upper layer which now contains the DNA is transferred to a new eppendorf tube. 100 µl of Chloroform-Isoamyl alcohol is then added to the new tube and vortexed for one minute to mix. The eppendorf tube is then returned to the centrifuge and a three minute spin is performed at 13,000rpm at 4 °C temp. The upper layer from this spin is transferred to a new eppendorf tube and 250 µl of 100% ice cold ethanol is added and then the vial is transferred to the -20° C overnight. To collect the DNA pellet, the tube is spun down for twenty minutes at 13,000rpms at 4° C. The

ethanol is then pipetted off and 500µl of 70% ethanol/DEPC water is added to wash the pellet. The tube is then briefly inverted and then is spun for a final collection of DNA at room temperature for twenty five minutes at 13,000rpm. Once the final centrifuge cycle is complete, around 80% of the ethanol wash is carefully pipetted off as to not disturb the DNA pellet and then placed uncapped sideways on a clean chemwipe to evaporate the remaining ethanol leaving the DNA pellet behind. This process takes anywhere from 5 to 45 minutes and is carefully observed to prevent over drying of the DNA. Once the ethanol has evaporated, the pellet is resuspended in 20 ul of DEPC water and quantified by spectrophotometry (Thermo Scientific ND-1000). Once complete, the DNA construct is diluted to 125ng/ul and aliquoted into eppendorf tubes of 6 µl to minimize freeze/thawing between injections and stored at -20° C.

Generation of transposase mRNA

Tol2 Transposase mRNA used in microinjection experiments is generated from a DNA construct in PCS2-FA originally sourced from the lab of Dr. Chi-Bin Chien (Kwan et al., 2007), stored in glycerol stock at -80C (Appendix A, Figure B). The glycerol stock is plated on Ampicillin (Amp) LB agar plates and cultured overnight at 37° C. Single colonies are selected and then transferred to 3ml liquid cultures which are grown overnight at 37° C. The plasmid DNA is then isolated from the culture using a Wizard Plus SV Miniprep DNA Purification Kit (Promega A1460). 20 µg of Tol2 transposase DNA template is prepared for transcription through a linearization digest using restriction endonuclease Kpn1 (NEB R0142S); [20µg plasmid DNA template, 5.0µl of Kpn1 enzyme (NEB R0142S), 10.0µl NEB buffer 1.1(NEB) and brought to a final volume of 100µl using Gibco distilled water (Invitrogen 15230)]. DNA template is digested overnight at 37° C, and then ran on a 0.5%

TBE gel to ensure complete linearization of the template. Finally, the template DNA is purified using a Phenol-Chloroform extraction as detailed above and then purity and concentration determined by spectrophotometry and stored at -20° C.

Linearized Tol2 transposase DNA template is then used to produce mRNA for microinjection using an SP6 mMESSAGE mMACHINE© RNA Transcription kit (Invitrogen by Thermo-Fisher). The work area and all gloves used during the reaction are thoroughly cleaned with RNaseZap (Ambion© AM9780) and dried. Enzyme Mix and Transposase template are thawed in an ice bath while the 10x reaction buffer and 2x NTP/CAP are thawed at room temperature. All solutions are briefly vortexed and spun down once thawed and then the following is added in order to an autoclaved eppendorf tube at room temperature- [1µg of Tol2 Transposase Template, 10.0µl of NTP/CAP, 2.0µl of 10x Reaction Buffer, 2.0µl of Enzyme mix, and brought to a full volume of 20.0µl using brand new aliquot of DEPC treated nuclease free water (Invitrogen 46-2224)]. The tube is briefly mixed by flicking and briefly spun down and is then incubated at 37° C for two hours. At the one hour forty five minute mark, Turbo DNase is added to the reaction tube, mixed by flicking and spinning down and is then returned to the water bath for the final fifteen minutes. 30ul of DEPC water and 30ul of Lithium Chloride precipitation solution (7.5M lithium chloride, 100mM EDTA) is added to the reaction tube, which is then stored in -20° C for a minimum of two hours or overnight. The mRNA is pelleted by centrifugation at 4° C for 15 minutes (13,000rpm). The supernatant is removed leaving the mRNA pellet, and 1 ml of 70% Ethanol wash solution is added to clean the pellet and then the vial is returned to the 4° C centrifuge for another ten minutes at 13,000rpm. Upon completion of the spin, the supernatant is removed and the pellet is allowed to air dry for five minutes by inverting the eppendorf tube over a clean wipe.

The mRNA pellet is then resuspended in 20µl of DEPC treated nuclease free water. The mRNA is quantified for both purity and concentration by spectrophotometry, and is then separated into single use aliquots of 4 µl 175ng/µl concentration for injections and stored at -20° C. A single aliquot is then selected and run on a 1% TBE agarose gel in RNA loading dye along with an ssRNA ladder (NEB N0362S) to confirm the mRNA is of the correct size.

Microinjections of m2de3 vector construct

Micro-injections of embryos are performed on single cell stage embryos using a Nanoliter 2000 micro injector (World Precision Instruments Model B203XVY). The overall solution injected is a 5 µl solution containing 125 ng/µl DNA, 175 ng/µl transposase mRNA, 2µl Phenol Red with the remaining volume composed of DEPC treated RNase free water. Transposase mRNA is translated by the zebrafish embryo to produce the transposase protein required to insert the *m2de3* expression cassette into the host genome. Phenol red (Sigma Aldrich 0.5% Phenol Red in DBPS solution) is utilized as a red visual marker injected and visualized through the clear embryo to confirm the injection. The injector must then be loaded with an injection needle, which is an RNase free 3.5 nanoliter capillary tube heated and pulled using a Vertical pipette puller (David Kopf Instruments Model 700C) with a heat setting of 85 and a solenoid setting of 0. This needle is then beveled by hand using forceps tweezers and then locked into the micro-injector, back filled with mineral oil filler and then attached to the micro injector needle. Prepared injection solution is then drawn up through the needle point and then the mircoinjectior is set to 4 nl and is used to inject the injection solution into the yolk of each embryo. Embryos are lined up against the edge of a microscope slide which is directly attached to a petri dish stage to align and prevent movement of the

embryos. Embryos are then injected from the side directly into the yolk on the far side clear away from the developing cells in the embryo by pushing the needle through to the yolk itself. Red injected solution should be clearly visible within the yolk material once injected and then the needle is carefully removed to prevent tearing of the chorion. Injected embryos are then removed from the slide, cleaned and then kept in 0.3x Danieau buffer in an incubator at 27° C. Embryos are then cleaned two to three times daily and dechorionated as necessary before the imaging process if they have not hatched on their own accord.

Analysis of Gene Expression through Optical and Laser Confocal Microscopy

Embryos are prepared for laser confocal imaging by first being anesthetized using a 20% tricaine (Tokyo Chemical Industry, Ltd.) Danieau buffer solution and are observed until there is no swimming movement. Embryos are then immobilized in a 0.8% Danieau buffer agarose solution, which is pipetted onto a deep well microscope slide, and the embryos are arranged in imaging positon before the agar sets and a glass cover slip is applied. Transgenic embryos are observed under a Zeiss laser confocal microscope (Zeiss LSM 510) using standardized laser channel gain settings to visualize any eGFP expression that is clearly separate from naturally occurring autoflourescence, using FITC and DAPI filter visual settings. If expression is observed, images are collected and compiled into single and compound Z-Stack images, using a FITC-CY2-GFP Long pass setting using the Zen 2009 computer program, with the exact stack size and speed parameters adjusted to each individual fish. This process is performed at varied time points, inside the range of 10 hpf to 72 hpf, saved as a high quality .tif format with the only image modification being incremental size markers applied by the laser confocal software.

Embryos that are to be screened purely for breeding or transgenic status confirmation purposes can be easily screened without being subjected to immobilization and anesthetic. Embryos are placed in a thin layer of 0.3x Danieau buffer inside a clear glass Pyrex dish, into which around 4-5 drops of tricaine solution (20%) are placed to slow the swimming and twitching of young embryos. At this point, the dish can be placed under the normal optical fluorescent microscope (Olympus IX81.3 using Microsuite Biological Suite 5 v.3.1), usually at 48hpf. Embryos can be quickly viewed through the microscope and transgenic embryos can then be removed via a disposable glass pipette and transferred into another dish.

Genome Analysis of Adult Zebrafish through Fin Clip protocol

Clipping the caudal fin in zebrafish is a common method to isolate genetic material for analysis and further molecular genetic study. Potential transgenic adult fish were selected and moved to a quarantine tank. Surgical forceps and metzenbaum scissors were sterilized and thoroughly cleaned with 90% ethanol and air dried. The transgenic fish were moved one at a time to a 0.8% tricaine solution in system water in a wide mouth glass bowl where the fish is observed for continual gill movement and swimming for a maximum of five minutes. When the fish was properly anesthetized, the tail fin was gently held upright with forceps and a small section of the upper part of the caudal tail was removed, no more than halfway between the tip and centerline. This should occur without blood loss. The fish was then immediately transferred to a bubbler oxygenated recovery tank and observed carefully for return to normal behavior within five minutes of the transfer. From this point on the fish was carefully monitored in a separate recovery tank until the caudal fin section has completely regrown, usually around two weeks. The removed tail section is carefully placed into an

Eppendorf tube and can be used immediately or stored at -20° C until genome extraction was performed. Genomic extraction was performed using an activated Proteinase K genome extraction buffer (10mM Tris pH 8.2, 10mM EDTA, 200mM NaCl, 0.5% SDS, 200ug/mL Proteinase K) to break down the fin tissue and cartilage on a heated gyroshake table (56°C, 100rpm, 3 hours or until completely digested). The product was then purified by phenol chloroform DNA purification and genomic DNA analyzed by polymerase chain reaction (PCR) to determine the presence of the transgenic expression construct using *m2de3* specific primers (Appendix A, Table 1) and Taq PCR Polymerase (ThermoFisher Scientific). PCR was performed under the following conditions; initial melt of 93° C for 1 minute thirty, normal melt cycle at 93° C for thirty seconds, normal annealing cycle at 58.9° C for 30 seconds, normal extension cycle for 68° C for one minute, repeated for 30 cycles followed by a completion hold at 68° C for ten minutes and indefinite final hold at 4° C until the samples are removed from the PCR machine. PCR samples were loaded on a 1% TBE gel and examined for bands of the correct size against a positive *m2de3* control looking for bands around 491bp. (455bp element plus two 18bp primers) (Figure 2).

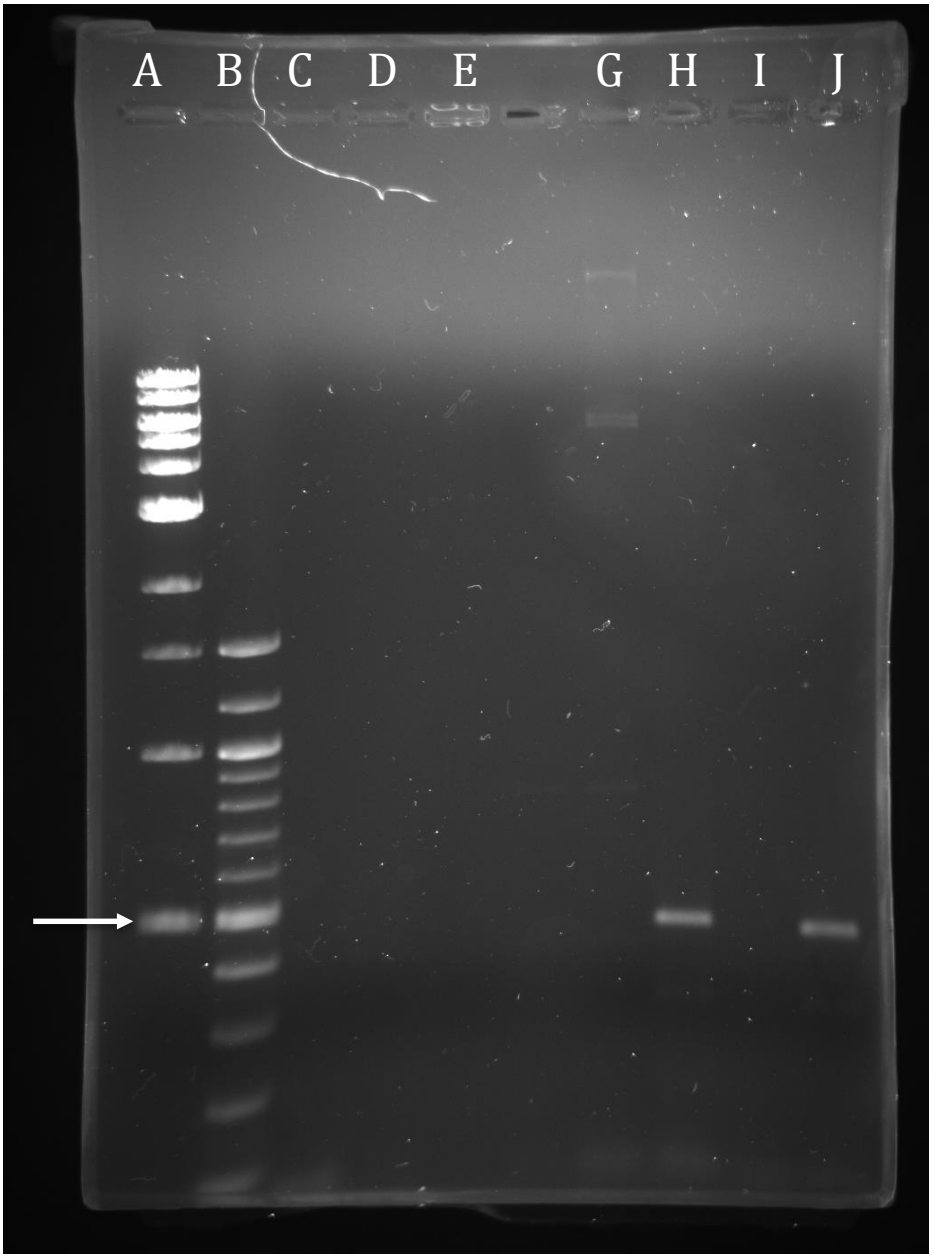


Figure 2- Gel Image Example of Transgenic Fin Clip Screening. Column A is 1 kilo base DNA Ladder with a bottom band of 500bp (white arrow), column B is a 100bp DNA ladder with a 500bp ladder band (white arrow). Column C, D and E are negative result PCR columns. Column F is a negative experimental result with no bands, Column G has incorrect bands size and is a negative result, Column G is a positive transgenic result, and Column H is

a negative experimental result. Finally, Column I is a positive PCR result from *m2de3* topo DNA.

Results and Discussion

Microinjection Characterization of m2de3

Expression of the *m2de3* element was characterized through the eGFP reporter gene construct microinjections and all embryos were imaged using a Zeiss laser confocal microscope. Initial test screens of transgenic zebrafish initially showed limited expression at 24 hours post fertilization (hpf) and clearer expression at 36 and 48 hpf. Injected embryos were screened at 24hpf multiple times and faint GFP expression was observed in a select few embryos, but expression was extremely inconsistent within each clutch of embryos and across multiple clutches. In later experiments, large batches of embryos were split and half were imaged at 24hpf and 48hpf and the earlier embryos demonstrated no GFP expression while the later batch demonstrated notochord expression (Figure 3, Figure 4) In addition, expression time points before at 20hpf and after at 26hpf and 28hpf yielded similar irregularity and no repeatable expression pattern. The irregular expression observed at 24hpf is more than likely a consequence of irregular or altered development timelines or the GFP expression cassette being inserted adjacent to another regulatory element within the zebrafish genome.

Consistent *m2de3* expression begins with faint and gradual expression along the notochord, first observed around 30hpf to 32hpf in a distinctive pattern starting in large sections of the notochord and in dorsal ventral aligned stripes of the hypochord and floorplate structures located along the notochord (Figure 3A, 3B). Around the 34 to 36 hpf mark, the expression pattern brightens in eGFP expression and the pattern becomes visible along the entire length of the notochord, but expression still remains intermittent through the entire notochord (Figure 3C, 3D). At 38 to 42 hpf, the zebrafish notochord pattern is fully visible

and is generally bright and clear along the notochord with large notochord cells exhibiting GFP expression (Figure 3E, 3F). Expression strength peaks after 48 hpf, with bright clear expression within the notochord (Figure 3F). The eGFP expression pattern tends to expand in size and strength of expression over previous time points, often wrapping in a curve within the notochord as the notochord cells fill out (Figure 3F). Expression continues at a high level of consistently bright expression through 56 and 58hpf, and the notochord expression pattern expands to the largest observed size (Figure 3G, 3H). This expression is present along the entire length of the notochord in varying strength and remains through 60hpf (Figure 3I). Motor neuron expression was observed clearly through all time points examined from 30hpf to 60hpf but rarely followed a consistent pattern completely along or around the notochord. The neural pattern was also seen rarely around the heart in around 6 embryos total during microinjections, for example Figure 3E and 3F. The heart expression was originally thought to be autoflourescence associated with the yolk since it was commonly imaged either against or close to the yolk of the developing embryo but upon comparison of all the collected images along with the transgenic line images this expression was confirmed as part of the heart. Muscle fiber expression was seen in embryos at a selection of different time points after 48hpf during microinjection screens, specifically 32hpf, 48hpf, 52hpf, 56hpf, and 58hpf (Figure 7). Observed muscle fibers were only seen in one or two embryos in each clutch of embryos and were always seen in short sections of the embryos either on one side or mirrored along both sides of the notochord region. All three of these secondary patterns were always seen in the presence of to at least partial notochord expression of the usual pattern.

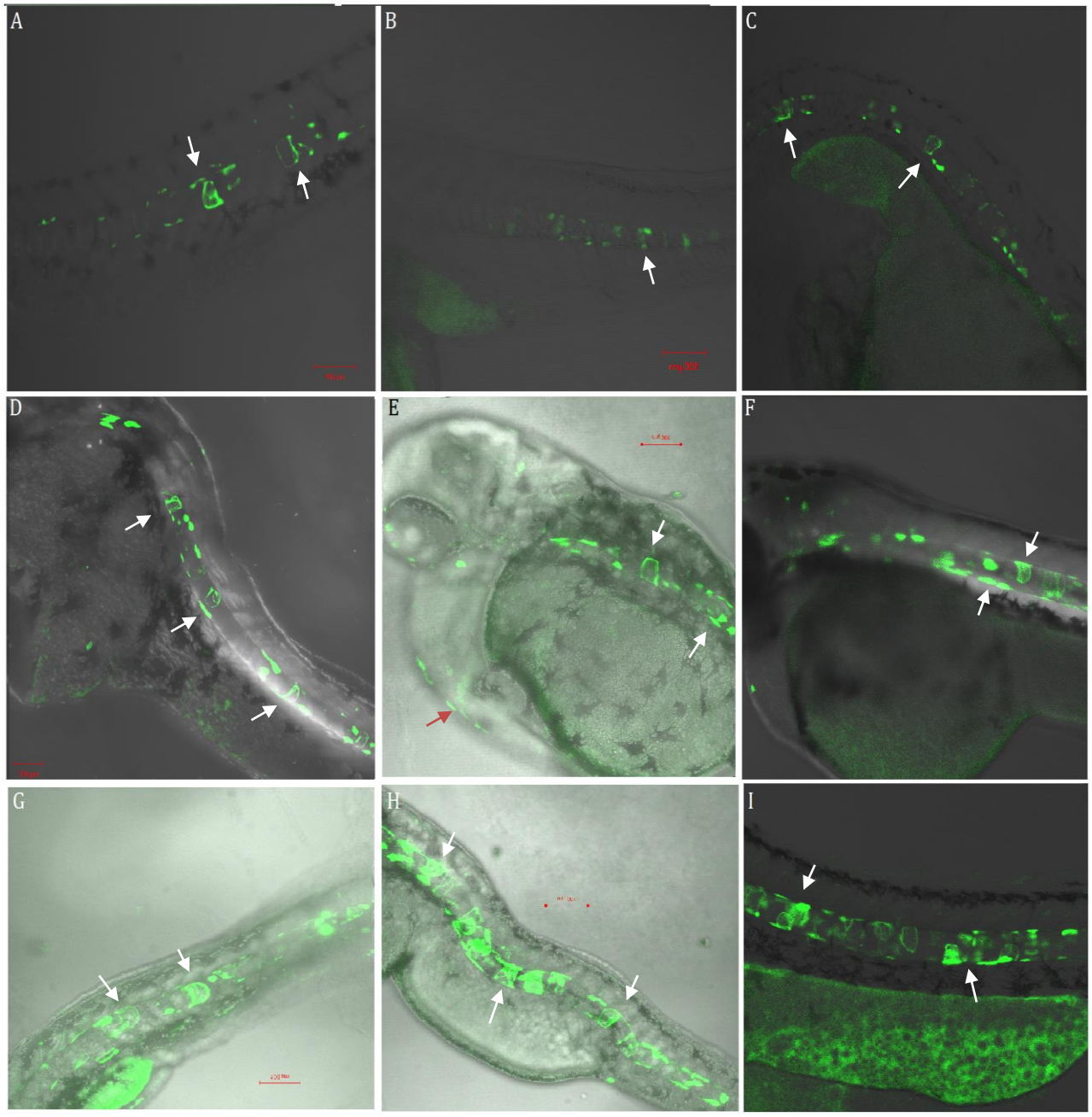


Figure 3- Microinjection *m2de3* transgenic fish. Figure 3A is a 30hpf embryo with arrows pointing to developing notochord cells; Figure 3B is a 34hpf embryo with an arrow pointing to early faint notochord expression; 2C is a 32hpf embryo with arrows showing full notochord cell expression; 2D is a 36hpf embryo shows arrows pointing out large notochord cells and notochord sheath expression; 2E is a 38hpf embryo with heart expression marked

by a red arrow and white arrows marking notochord cell and motor neuron expression; 2F is a 48hpf embryo with notochord cell, motor neuron and notochord sheath expression marked by white arrow; 2G is a 56hpf embryo with large notochord cell expression; 2H is a 58hpf embryo showing full length notochord expression with arrows highlighting larger sections of expression; 2I is a 60 hpf embryo with white arrows showing the sharp notochord sheath boundary and internal notochord cells ; All images are arranged anterior left / posterior right. Scale Bar is 100 μm in length.

Transgenic Line Characterization of m2de3

To generate a stable *m2de3-eGFP* transgenic line, 38 zebrafish embryos were raised from injected embryos that exhibited notochord eGFP expression. Each embryo was raised to adulthood, and a fin clip DNA extraction was performed to test the transgenic status of each animal. 8 zebrafish generated a positive confirmation PCR result as determined by gel electrophoresis and were separated as full F1 Generation transgenics. These zebrafish were then crossed pairwise 1 male and 1 female with both petshop non-transgenics and other *m2de3* transgenics to test for GFP expression patterns in the F2 Generation. Parings between 1 male transgenic zebrafish and 2 female transgenic zebrafish were found to generate a very consistent pattern of GFP expression within the F2 generation. This expression was observed at 32hpf, 38hpf, 48hpf and 56hpf, key time points selected from the microinjection experiments. Expression appears extremely similar in pattern to the original primary injected embryos, but is observed along the entire length of the notochord, with sharply defined boundaries along the outside of the notochord sheath and is present in all notochord cells (Figure 4A, Figure 4B). In addition, heart expression is much clearer and illuminates the entire heart instead of the restricted areas seen in microinjected embryos (Figure 4A, Figure 4B). This heart expression was observed in two patterns, one as a highlighted outline of neuron notochord style expression (Figure 4A) and one of the entire heart expressing eGFP (Figure 4B). The heart expression pattern was consistently observed in all time points while the whole heart image shown was demonstrated by two embryos Figure 4B.

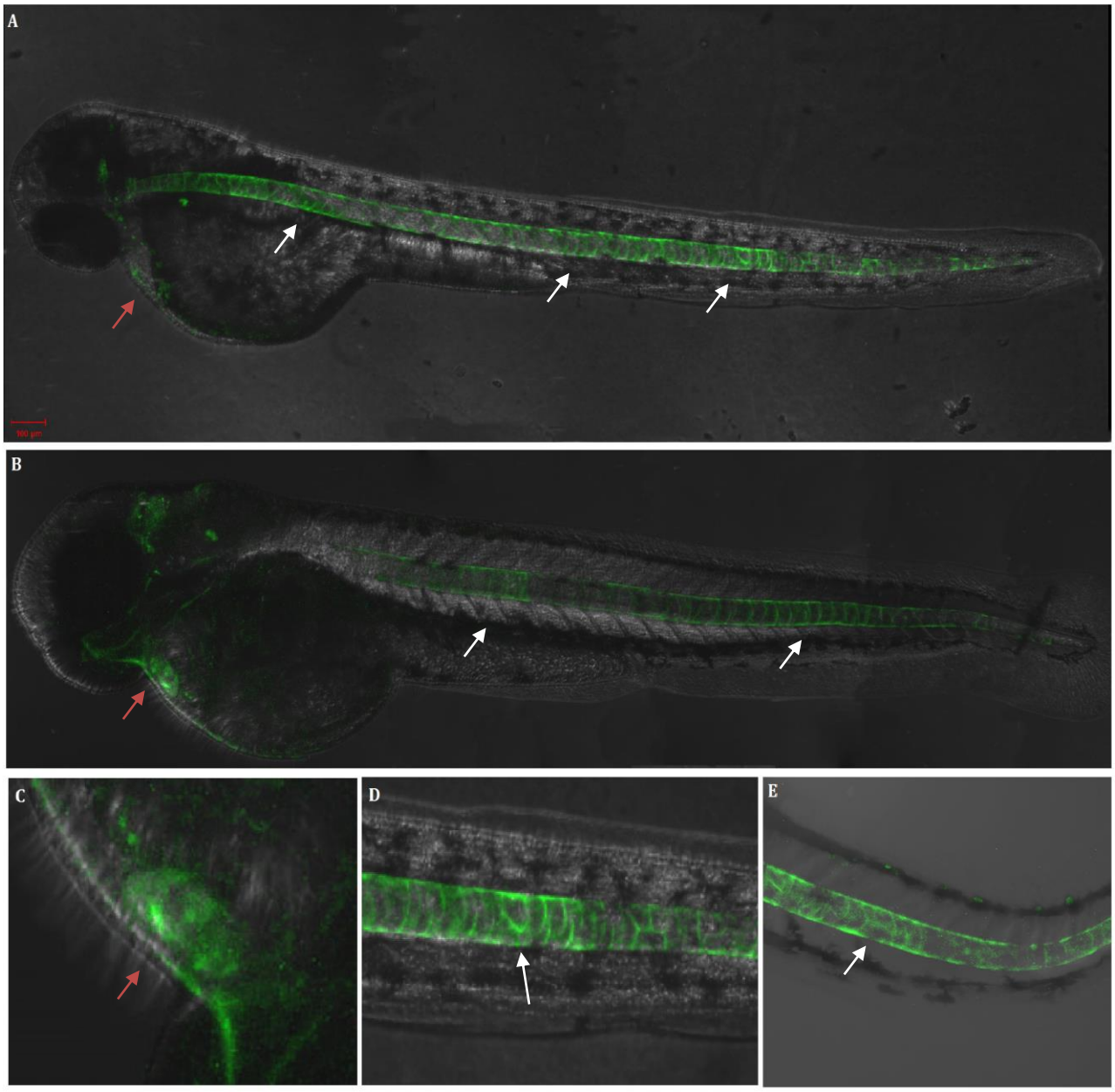


Figure 4- F2 Generation Transgenic Fish. Figure 4A is a whole fish image compilation from a 56hpf embryo with 100 µm Scale Bar. Heart expression is marked with a red arrow and some detailed notochord expression is marked in white; Figure 4B is a whole fish image compilation of a 48hpf embryo. Heart expression is marked with a red arrow and the white arrow points to the notochord sheath boundaries and notochord cells; Figure 4C is an expanded zoom image of the circular heart expression taken from Figure 4B source images,

indicated with a red arrow; Figure 4D is a zoomed image of notochord expression taken from Figure 4A source images with the white arrow pointing to individual notochord cell boundaries; Figure 4E is an image from a 38hpf embryo showing mixed expression in the notochord and boundaries of the notochord sheath. All images are arranged anterior left / posterior right. Scale Bar is 100 μm in length.

Discussion of Expression patterns

The most prominent expression observed to be directed by *m2de3* is in the notochord. In vertebrates, the notochord or *Chorda dorsalis* is a semi rigid rod like structure located in-between the dorsal spinal cord and the anterior gut tube that travels the entire length of the trunk. The notochord is necessary and required for correct development, with mutations of the notochord leading to an abnormally short body plan which directly hinders zebrafish movement particularly in swimming and feeding and in extreme cases are developmentally lethal due to incomplete nervous systems. The notochord serves as a structural strength "backbone" for the skeleton until bone ossification occurs, upon which the notochord is either ossified into the vertebrae structure or in certain organisms remains as a separate structure. The notochord also serves as a developmental organizer splitting the right and left halves of an organism and sending patterning signals to other nearby structures for body positioning and cell-fate differentiation. The most notable role of the notochord in development is as the signaling center for primary neurulation, somitogenesis and neuron pathfinding to connections in the spinal cord. Current research has also tied notochord-based signaling to development of cardiac neural connections, formation of the aorta and other blood vessels as well as the development of the pancreas (Eriksson and Löfberg, 2000; Fouquet et al., 1997; Pourquié et al., 1993; Stemple, 2004, 2005; Yamada et al., 1991; Yamada et al., 1993). This control and organization of development is performed through production and release of a variety of signaling molecules such as *BMP, Wnt, Shh,* and *Sox9a* (DiIorio et al., 2007; Lien et al., 2013; Pourquié et al., 1993; Yamada et al., 1993; Yamada et al., 1991).

The notochord develops from a line of multipotent cells originally designated during gastrulation of an embryo, which then differentiates into patterned chordamesoderm which then develops into the notochord. The final notochord structure must remain flexible from side to side, but also must be able to mechanically support the body skeleton and not travel past physical limits in the vertical up and down axis but remain flexible enough for swimming. This structure is described well in an analogy in Stemple as a "firehose packed full of water balloons with steel cables at top and bottom" (Gilbert, 2000; Stemple, 2004, 2005). Using this analogy, the firehose itself is the flexible but strong notochord sheath which serves as an outer boundary for the notochord that retains the notochord cells and provides high outer wall strength and resistance to interior cells. The notochord cells themselves are highly vacuolated, and press outward against the sheath and resemble the water balloons in this analogy, creating tension through the structure. The steel cables running along the top and bottom are the floorplate and the hypochord respectively, where in this position they provide more rigid structure in the vertical axis thanks to large concentrations of cartilage proteins, but remain flexible in the horizontal axis for locomotion (Stemple, 2004, 2005). In zebrafish, the designated mesoderm is clearly visible around 12 hpf and transitions into a clearly visible notochord around 24 hpf which is then present throughout the adult life of zebrafish. The *m2de3* directed expression in the notochord region is normally observed within single cells of the notochord in primary embryos that have been microinjected with the expression cassette. However, imaging of a stable transgenic *m2de3* line has shown that this expression pattern occurs throughout the entire notochord instead of in single cells, and that the single cell expression pattern is likely a result of mosaicism. Microinjections of an expression cassette into zebrafish embryos does not ensure that the

construct is integrated into every single cell of the organism, thus leading to the mosaic expression patterns from all cells not necessarily containing the injection construct. Comparison of the transgenic *m2de3* directed eGFP pattern and the transgenic F2 expression shows clear irregular vertical lines which are extremely similar to the structural boundaries visible between notochord cells seen in unstained and stained wild type zebrafish images (Figure 6) available through ZFIN (Abdelilah et al., 1996; Stemple, 2005; Thisse et al., 2001). The extremely sharp and clean outer boundaries of the expression pattern also make sense with the tension of the notochord sheath which retains the shape and sharp lines along the edge of the notochord and provides its strength, which provides clear evidence for this type of expression. In addition, some of the longer line sections seen along the top and bottom of the notochord in Figure 3D, Figure 4D and Figure 4E clearly resemble sections of the hypochord and floorplate regions, being located extremely close to the notochord and often demonstrate mirrored expression on both dorsal and ventral sides of the notochord. Finally, different *Hox* genes and *Meis/Pbx* have been independently shown in the literature to be expressed actively in the notochord during time points when *m2de3* directed expression has been observed. This provides even more evidence for *m2de3* expression driven by a highly conserved non-coding element that is known to be associated with binding to these elements in the notochord (DiIorio et al., 2007). This validates my observations and also suggests that *m2de3* does indeed direct expression of *Meis2* expression in the notochord.

The next area of expression we have to examine is the spinal column, particularly the posterior region of the spinal cord and its motor neurons. The spinal cord runs above or dorsal to the notochord along the trunk of the body and serves as the primary relay and connection between the peripheral nervous system of most of the body and limbs and the

central nervous system. The spinal cord is formed from ectodermal cells which make up the precursor neural plate structure during a process called primary neurulation. This plate is then signaled by Sonic hedgehog (Shh) from the notochord to transform in shape and fold downward to the notochord itself, then extend in the elevation phase upward. A hinge is then formed at the dorsolateral hinge point in the middle of the neural plate and the two ends fold together into the middle over the neural tube space creating the neural tube structure, which is connected together by the neural crest cells to form the final neural tube that will become the full spinal cord. This neural tube region is then further differentiated based on gradient strength of signal molecules BMP4, BMP5, BMP7 and Shh to differentiate dorsal ventral axis patterning and to pattern each section of the neural tube for its future role as the spinal cord (Gilbert, 2000; Zelenchuk and Brusés, 2011). Some of the cells from the neural tube will become motor neurons, or the neural connection between muscle fibers and the nervous system. Motor neurons differentiate from the neural tube and form precursor motor neurons by 16 hpf in zebrafish (Myers et. al., 1985; Myers et al., 1986). After this original differentiation, two types of motor neurons will arise from the designated precursor cells, primary and secondary motor neurons.

Primary motor neurons are much larger than their secondary counterparts, and have different branching patterns and size differences. Primary motor neurons will push outward from the spinal cord region around 20 hpf. In zebrafish there are three primary motor neurons always found in each somatic segment, labeled rostral (RoP), Middle (MiP) and Caudal (CaP) and by a variable primary neuron (VaP) that occurs in half the segments, alternating between present and absent in each segment. Primary motor neurons demonstrate larger and longer reaching dendrites, and an axon connector that travels along the interior side of the

Mauthner axons. In comparison, Secondary motor neurons have much smaller cell bodies, around 170 um in overall diameter versus the primary motor neurons which are around 670 um. There are 25 secondary motor neurons formed within each segment with varying patterns of much smaller dendrites and their thinner axon falls on the outside of the Mauthner axon. Finally, secondary axons exit the spinal cord region later around 26 to 28 hpf, and reach full size by 48hpf. Both neuron types direct dendrite and axon connections specific to the surrounding tissues as Hox and LIM homeodomain proteins which drive genes such as Ret and Gfra to establish correct connections with other neurons, muscle cells and limb neural connections (Catela et al., 2016; Mallo et al., 2010; Myers et al., 1985; Myers et al., 1986; Rousso et al., 2008; Westerfield et al., 1996; Zelenchuk and Brusés, 2011).

The *m2de3* directed reporter gene expression also appears to be located in the motor neurons. This expression is observed throughout the embryo within a punctate pattern, which is clearly separate from the notochord expression pattern. In embryos microinjected with the expression construct, I observed small sections of the fish exhibit a small selection of these repeated motor neuron cell bodies expression pattern, with other sections of the zebrafish demonstrating little or no expression in the same pattern (Figure 4A, 4B). This is more than likely also due to mosaicism, as described previously, from the microinjections meaning that the construct is not present in all cells. In the transgenic line imaging, the motor neuron pattern is still visible but is more washed out against the brighter notochord expression (Figure 6). However, many characteristics of the expression pattern fit with motor neuron bodies demonstrating circular or oval structures in the correct area that matches stained expression (Myers et al., 1985; Myers et al., 1986; Zelenchuk and Brusés, 2011). Second, the size of the punctate expression structures does indeed vary in size, with some much larger

and brighter expression sites as primary motor neurons and smaller secondary motor neurons in the flanking sites in the same region (Hutchinson and Eisen, 2006; Lin et al., 2012; Zelenchuk and Brusés, 2011). Probably the most important piece of evidence is why this expression is directed only to the cell body of the motor neuron and not the axon and dendrite regions of the cell that should be visible in later time points. While certain *m2de3* directed expression research images do show small projections from the central cell body, for example in Figure 3C, this expression pattern makes sense for a motor neuron branching out into the trunk muscles but we cannot determine exactly which motor neuron type each expression shape is. This would also coordinate with when *m2de3* directed expression is first observed shortly after these motor neurons have reached their final position and are starting to branch out in association with other factors, and why no expression is seen before 30hpf. (Hutchinson and Eisen, 2006; Jung et al., 2010; Rousso et al., 2008; Zelenchuk and Brusés, 2011). However, since eGFP located within the cytoplasm of cells it should be seen throughout the entire cell instead of just the cell body. One possible explanation is that since axons and dendrites are small thin structures extending from the larger cell body that these structures may not be always visible due to less eGFP located within these extensions.

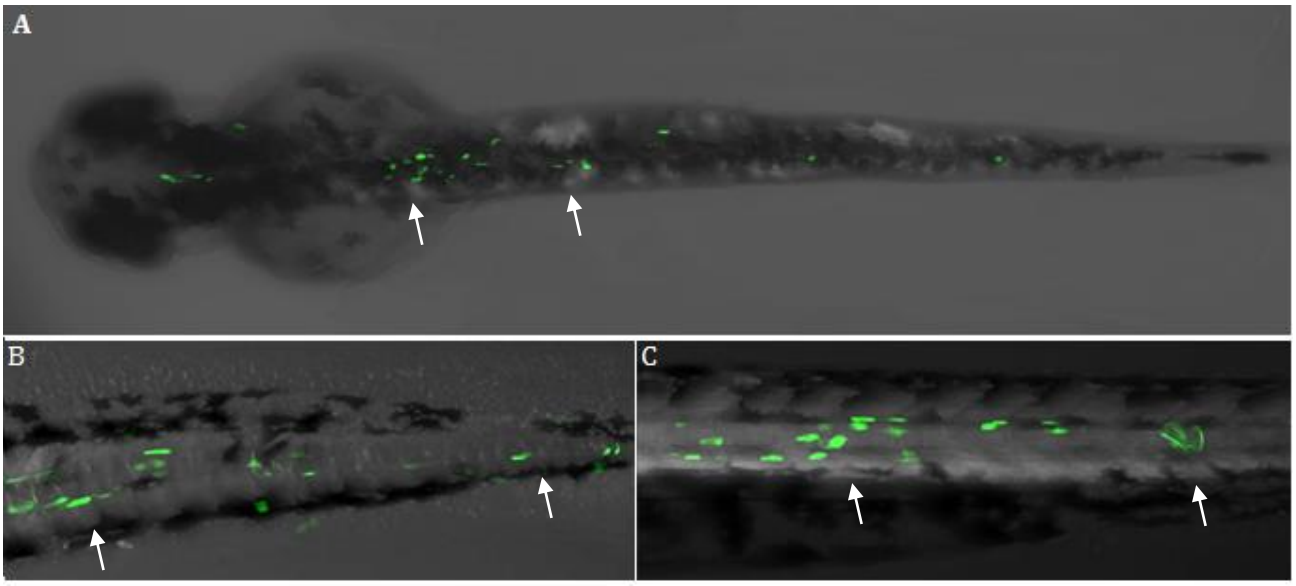


Figure 5- Motor Neuron Examples at Selected Time points of microinjected embryos. Figure 5A is a whole fish image at 48hpf with white arrows pointing to large groups of motor neurons; Figure 5B is a tail section motor neuron example at 36hpf with arrows showing motor neurons against the notochord; Figure 5C is a tail section at 60hpf and has arrows that show expression on all sides of the notochord. All images are arranged anterior left / posterior right.

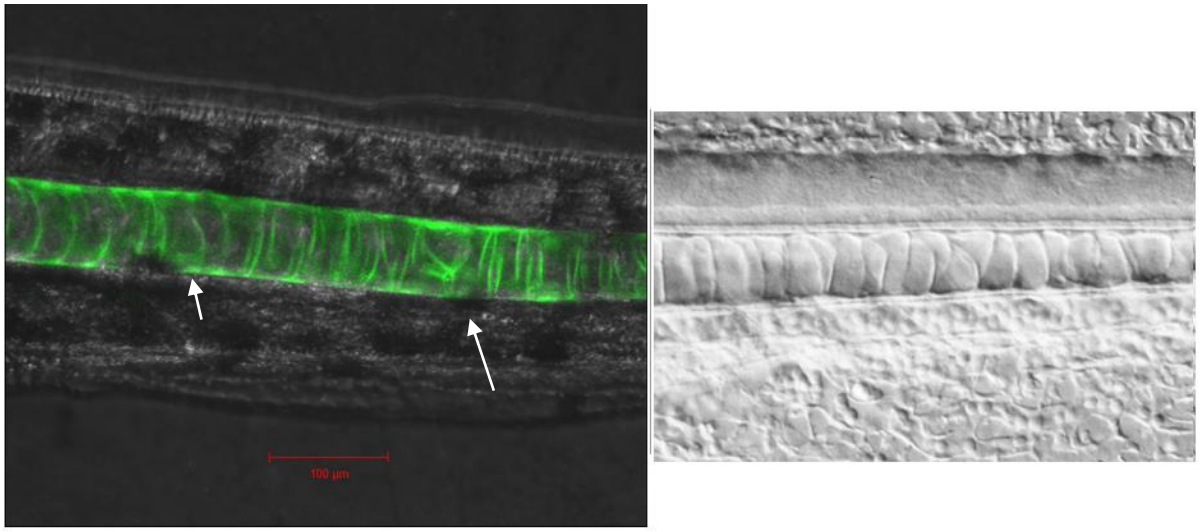


Figure 6- Notochord section containing Motor Neuron Example Selected from a 48hpf Transgenic F2 embryo aligned with a reference image of an unstained notochord from Figure

4 of (Stemple, 2004; Stemple, 2005). White arrows point to notochord cells with clear cell wall boundary detail. Images are arranged anterior left / posterior right. Scale Bar is 100 µm.

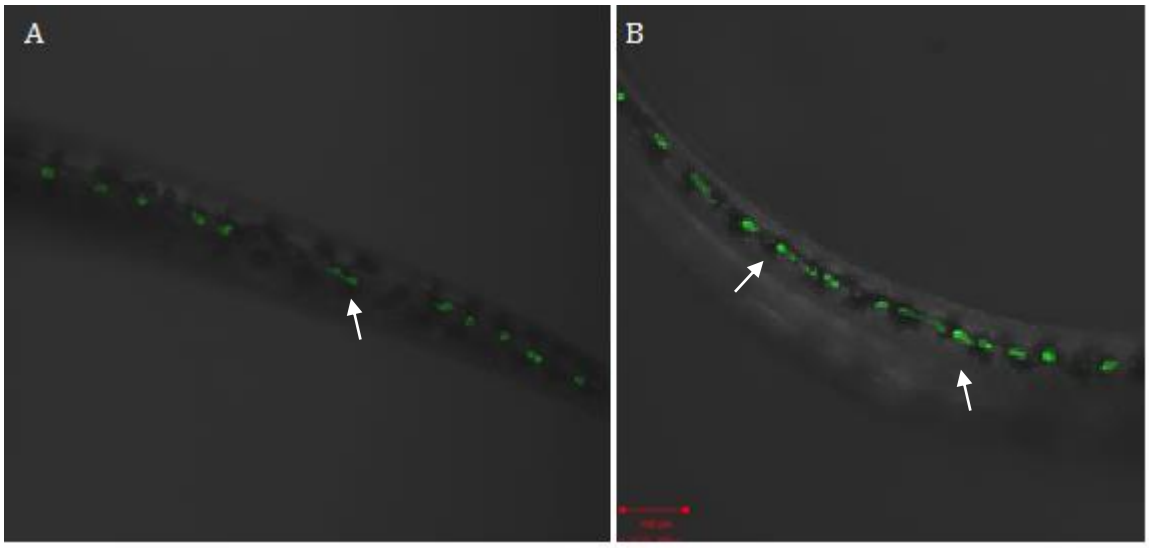


Figure 7- Lateral Line System Examples of F1 generation. Figure 7A is a 50hpf embryo with an arrow showing one pair of lateral line cells; Figure 7B is a 56hpf embryo showing the lateral line pattern along the entire trunk marked by white arrows. Scale Bar is 100 µm. All images are arranged anterior left / posterior right.

M2de3 directed expression could also be involved in the lateral line system in zebrafish. The lateral line system is a unique structure that fish and salamanders have that travels the entire length of their body and has several mechanoreceptors in a sensory hair cell or neuromast that serve to detect environmental conditions and other fish around them (Whitfield et al., 1996) Zebrafish also use this system similar to an internal gyroscope for movement and orientation in the water column, as well as detection of possible lethal contaminants in the water environment. The primary structure consists of 29 neurosensory sites concentrated around the head, eye and along the trunk region of the zebrafish, as seen on ZFIN (Whitfield et al., 1996, Whitfield, unpublished data). The lateral line sites occur along the same area in the trunk of the zebrafish, also similarly in pairs like the ZFIN reference (Figure 7A, 7B). While zebrafish do not have the most sensitive or complex lateral line system, we should still consider it as a possible target to direct expression for *m2de3* due to its direct ties to the neural system, and the lateral line system is also known to share common neural factors (Fuentes-Pérez et al., 2015; Germanà et al.,2009; Wada et al., 2014). Since the *m2de3* gene was originally isolated from an organism that does not naturally have the lateral line system, it could also be a byproduct of directed neural expression associated with the notochord that the lateral line system is responding to. This is further complicated since the lateral line system is located further out from the midline of the zebrafish, making differentiation between expression sites hard in a horizontal view as zebrafish have been imaged in this research. Research of the lateral line system is mostly limited to toxicological and neural research, so the current images have shown expression in the neuromast cell

locations but we do not currently have enough information to completely rule these out as certain expression points in the zebrafish.

The final area of *m2de3* directed expression that we have observed during this project is observed in certain muscle fibers and limited expression that has been observed in previous studies regarding the *m2lg* element transgenics. Muscle fiber expression as directed by *m2de3* is always found located along the trunk and always together with the presence of the expression patterns in the notochord and motor neuron expression type. This expression is commonly significantly weaker in magnitude than the neural and notochord expression and generally requires higher channel gain on the laser confocal microscope to be viewed clearly. In microinjection experiments, visible in Figure 8A are muscle fibers of the same signal strength as other expression patterns while Figure 8B shows a high gain contrast image of weaker strength muscle fibers located along the trunk. Muscle fibers were observed at a variety of time points, specifically 32hpf, 36hpf, 48hpf, 56hpf, 60 hpf and 72hpf. Since this expression is inconsistent in location and strength, further breeding of the *m2de3* transgenic line zebrafish may provide further clarity. In another research project in the Zerucha lab, another element, *m2de2*, was shown to direct transgene reporter expression across the full trunk muscle fibers within a stable transgenic line (Freundlich, 2016). In a bioinformatic analysis of *m2de3* a binding site for MyoD, a muscle fiber and neural gene was identified within the *m2de3* sequence (Appendix A, Figure C). Comparison of expression patterns of *MyoD* in zebrafish at similar time points to the observed zebrafish pattern shows extremely similar expression patterns as demonstrated in Figure 8. Similar muscle fiber expression was also observed in zebrafish injected with the *m2de1* and *m2de2* elements and MyoD binding sites were also identified within those sequences as well. Also, the *m2de2*-eGFP transgenic

line showed complete trunk expression patterns, and further examination of the *m2de3* line could reveal similar results (Ferrara, 2015; Freundlich, 2016;). In context, this would make sense given the presence of the binding site itself. This is also supported by the fact that MyoD has also recently been discovered to interact with neural development factor NeuroD2 and with other neural regulation factors similar to *m2de3* (Fong et al., 2015; Messmer et al., 2012).

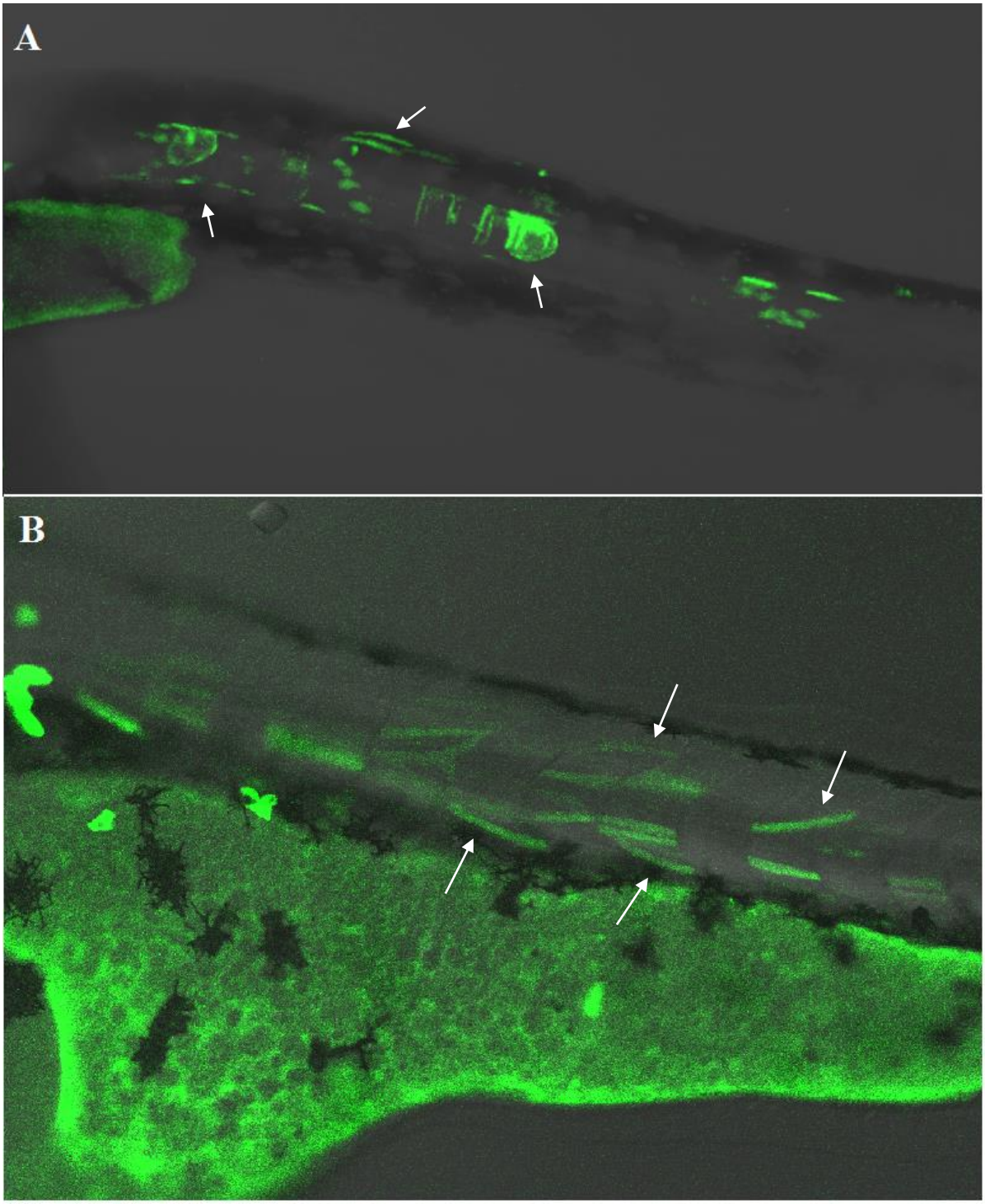


Figure 8- Muscle Fiber Expression Examples of *m2de3*. Figure 8A is from a 52hpf embryo using standard imaging values and white arrows pointing to individual muscle fibers, and the notochord pattern. Figure 8B is an enlarged image of a 56hpf embryo with an increased

channel gain of eGFP and select muscle fibers marked with white arrows. All images are arranged anterior left / posterior right.

Conclusions and Future Directions

In this project, we have undertaken a thorough examination of gene expression targets of the *m2de3* element by introducing the element to a model organism that naturally lacks it. The *mm-m2de3* eGFP expression construct was microinjected into zebrafish and were grown to select time points. These experiments have yielded a select area of definitive expression targets within the notochord, motor neurons, heart and muscle fibers, consistent across direct microinjection and within a stable transgenic line. The notochord is a definitive expression site of *m2de3* with extremely consistent experimental expression along the notochord, with patterns that confirm definitive physical characteristics of the notochord. Motor neurons located along the notochord have also shown expression serving as a link between the notochord and other targets such as the heart and muscle fibers. Muscle fibers have also shown to be an interesting tie in showing a possible connection between neural and muscle fiber expression that will be extremely interesting to further research. Finally, the lateral line system is located within these regions and must also be considered as a possible expression target, but warrants further examination. In addition, these microinjected zebrafish were then screened through fin clip genome extraction and tested for *m2de3* transgenics which generated a transgenic zebrafish line for use in future experiments.

Based off the information learned from this project, m2de3 definitely deserves further research. The first direction of future research would continue to examine the roles of zebrafish *meis2a* and *meis2b* to determine if gene expression of these two genes overlaps where the expression pattern directed by *m2de3* was observed. To this end, I performed preliminary studies into zebrafish *meis2* expression through in-situ hybridization and generation of RNA probes that target *meis2a, meis2b* and *M2lg.* I generated the RNA probes

necessary to complete this in-situ hybridization and performed successful hybridizations, but these were incomplete for most time points and inconclusive at the end of this project. Another direction of future research will be the comparison of *m2de3* expression in zebrafish against *meis2* expression in mouse embryos at comparable developmental time points. In addition, comparison of the *m2lg* gene and *m2de3* in both mice and zebrafish will allow further study related to the possibility that *meis2* and *M2lg* might be sharing the *m2de3* element (Carpenter, 2016)). This would be evidence for selective pressure involving the linkage that is observed between these two genes in all vertebrates examined. Another exciting experiment would be a full knockout of the *m2de3* element from the mouse genome as it would yield a unique opportunity to examine these mice for anatomical and physical changes in development. These changes might include incomplete or missing development of certain structures, changes in performance of organs and also the determination if *m2de3* is necessary and essential for survival and development of the organism. Finally, we identified multiple possible transcription factors that could potentially interact with *m2de3* such as Myod. These deserve further examination to see if their expression patterns occur in the same time points and locations observed in zebrafish and also if they are capable to binding to the element.

This characterization of the *m2de3* element and the generation of the transgenic *m2de3* line will serve as a vital foundation for further research, and the data we have generated on the *m2de3* element shows great promise and potential for future research developments in developmental genetics of the heart and nervous system.

References

Abdelilah, S., Driever, W., Gorn, A., Malicki,J., Neuhauss, S., Pack, M., Rangini, Z., Schier, A., Solnica-Krezel, L., Stanier, D. and Stemple, D. (1996). Anatomy of the 24, 48, 72 and 120 hours zebrafish (Danio rerio) embryo. ZFIN Direct Data Reference Submission.

Akam, M. (1989). Hox and HOM: Homologous gene clusters in insects and vertebrates. Cell. **57**, 347-349.

Ali, T., Renkawitz, R. and Bartkuhn, M. (2016). Insulators and domains of gene expression. Current Opinion in Genetics & Development. **37**, 17-26.

Amin, S., Donaldson, I., Zannino, D., Hensman, J., Rattray, M., Losa, M., Spitz, F., Ladam, F., Sagerstro, C. and Bobola, N. (2015). Hoxa2 selectively enhances meis binding to change a branchial arch ground state. Developmental Cell. **32**, 265-277.

Amores, A., Force, A., Yi-Lin, Y., Lucille, J., Amemiya, C., Fritz, A., Ho, R., Langeland, J., Prince, V., Wang, Y., Westerfield, M., Ekker, M. and Postlethwait, J. (1998). Zebrafish Hox clusters and vertebrate genome evolution. Science. **282**, 1711-1714.

Amrani, N., Ghosh, S., Mangus, D.A. and Jacobson, A. (2008). Translation factors promote the formation of two states of the closed-loop mRNP. Nature. **453**, 1276-1280.

Arensbergen, J., van Steensel, B. and Bussemaker, H.J. (2014). In search of the determinants of enhancer–promoter interaction specificity. Trends in Cell Biology. **24**, 695-702.

Arnosti, D.N. (2011). Transcriptional repressors: Shutting off gene expression at the source affects developmental dynamics. Current Biology. **21**, R859-R860.

Arrington, C., Dowse, B., Bleyl, S. and Bowles, N. (2012). Non-synonymous variants in pre-B cell leukemia homeobox (PBX) genes are associated with congenital heart defects. European Journal of Medical Genetics. **55**, 225-237.

Azcoitia, V., Aracil, M., Martı´nez-A, C. and Torres, M. (2005). The homeodomain protein Meis1 is essential for definitive hematopoiesis and vascular patterning in the mouse embryo. Developmental Biology. **280**, 307-320.

Banerjee-Basu S., S., D. and Baxevanis, A. (2001). The Homeodomain resource: sequences, structures, DNA binding sites and genomic information. Nucleic Acids Research. **29**, 291-293.

Banjeri, J., Rusconi, S. and Schagffner, W. (1981). Expression of a B-globin gene is enhanced by remote SV40 DNA sequences. Cell. **27**, 299-308.

Barber, B., Liyanage, V., Zachariah, R., Olson, C., Bailey, M. and Rastegar, M. (2013). Dynamic expression of MEIS1 homeoprotein in E14.5 forebrain and differentiated forebrain-derived neural stem cells. Annals of Anatomy. **195**, 431-440.

Barolo, S. (2012). Shadow enhancers: Frequently asked questions about distributed cis-regulatory information and enhancer redundancy. BioEssays : news and reviews in molecular, cellular and developmental biology. **34**, 135-141.

Bertolino, E., Reimund, B., Wildt-Perinic, D. and Clerc, R. (1995). A novel homeobox protein which recognizes a TGT core and functionally interferes with a retinoid-responsive Motif. The Journal of Biological Chemistry. **270**, 31178-31188.

Betancur, P., Sauka-Spengler, T. and Bronner, M. (2011). A Sox10 enhancer element common to the otic placode and neural crest is activated by tissue-specific paralogs. Development. **138**, 3689-3698.

Bhattacharjee, S., Renganaath, K., Mehrotra, R. and Mehrotra, S. (2013). Combinatorial contol of gene expression. BioMedical Research International, **3**, 11.

Biemar, F., Devos, N., Martial, J., Driever, W. and Peers, B. (2001). Cloning and expression of the TALE superclass homeobox Meis2 gene during development. Mechanisms of Development. **109**, 427-431.

Blencowe, B.J. (2006). Alternative splicing: New insights from global analyses. Cell. **126**, 37-47.

Bondurand, N. and Sham, M.H. (2013). The role of SOX10 during enteric nervous system development. Developmental Biology. **382**, 330-343.

Brennan, R. and Matthews B. (1989). The helix-turn-helix DNA binding motif. Journal of Biological Chemistry. **264**, 1903-1906.

Bulger, M. and Groudine, M. (2011). Functional and mechanistic diversity of distal transcription enhancers. Cell. **144**, 327-339.

Buratowski, S., Hahn, S., Guarente, L. and Sharp, P.A. (1989). Five intermediate complexes in transcription initiation by RNA polymerase II. Cell. **56**, 549-561.

Burgess-Beusse, B., Farrell, C., Gaszner, M., Litt, M., Mutskov, V., Recillas-Targa, F., Simpson, M., West, A. and Felsenfeld, G. (2002). The insulation of genes from external enhancers and silencing chormatin. PNAS. **99**, 16433-16437.

Bürglin, T. (1997). Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals. Nucleic Acids Research. **25**, 4173-7180.

Cannavò, E., Khoueiry, P., Garfield, David A., Geeleher, P., Zichner, T., Gustafson, E.H., Ciglar, L., Korbel, Jan O. and Furlong, E.M. (2016). Shadow enhancers are pervasive features of developmental regulatory networks. Current Biology. **26**, 38-51.

Carnicnci, P. E. A. (2006). Genome-wide analysis of mamalian promoter architecture and evolution. Nature Genetics. **38**, 626-635.

Carpenter, B.G., B. and Zerucha, T. (2016). Identification and developmental gene expression of the zebrafish zgc; 154061 gene. Eastern Biologist. **5**, 1-17.

Catela, C., Shin, Maggie M., Lee, David H., Liu, J.-P. and Dasen, J.S. (2016). Hox proteins coordinate motor neuron differentiation and connectivity programs through Ret/Gfrα genes. Cell Reports. **14**, 1901-1915.

Cecconi, F., Proetzel, G., Alvarez-Bolado, G., Jay, D. and Gruss, P. (1997). Expression of Meis2, a knotted-related murine homeobox gene, indicates a role in the differentiation of the forebrain and the somitic mesoderm. Developmental Dynamics. **210**, 184-190.

Chang, C., Brocchieri, L., Shen, W., Largman, C. and Cleary, M. (1996). PBx modulation of Hox homeodomain amino-terminal arms establishes different DNA-binding specificities across the Hox locus. American Society of Microbiology. **16**, 1734-1745.

Chen, C., Chen, C., Chern, S., Wu, P., Chen, Y., Chen, S., Chen, L., Yang, C. and Wang, W. (2016). Prenatal diagnosis and molecular cytogenetic characterization of a de novo 4.858-Mb microdeletion in 15q14 associated with ACTC1 and MEIS2 haploinsufficiency and tetralogy of fallot. Taiwanese Journal of Obstetrics & Gynecology. **55**, 270-274.

Conaway, R.C. and Conaway, J.W. (1993). General initaion-factors for RNA polymeras-II. annual review of biochemistry. **62**, 161-190.

Consortium, E.P. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature. **447**, 799-816.

Cooper, S., Trinklein, N., Anton, E., Nguyen, L. and Myers, R. (2005). Comprehensive analysis of transcriptional promoter structure and function in 1% of the human genome. Genome Research. **16**, 1-10.

Crane-Robinson, C. (2016). Linker histones: History and current perspectives. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms. **1859**, 431-435.

Delgado, I. and Torres, M. (2016). Graidents, waves and timers, an overview of limb patterning models. Seimnars in Cell and Devlopmental Biology. **49**, 109-115.

Deschamps, J. (2007). Ancestral and recently recruited global control of the Hox genes in development. Current Opinion in Genetics & Development. **17**, 422-427.

Desplan, C., Theis, J. and O'Farrell, P. (1988). The sequence specificity of homodomain-DNA interaction. Cell. **54**, 1081-1090.

DiIorio, P., Alexa, K., Choe, S., Etheridge, L. and Sagerström, C. (2007). TALE-Family homeodomain proteins regulate endodermal sonic hedgehog expression and pattern the anterior endoderm. Developmental Biology. **304**, 221-231.

Duboule, D. (1998). Vertebrate Hox gene regulation: clustering and/or colinearity? Current Opinion in Genetics and Development. **8**, 514-518.

Duboule, D. (2007). The rise and fall of Hox gene clusters. Development. **134**, 2549-2560.

Dvir, A., Conaway, J.W. and Conaway, R.C. (2001). Mechanism of transcription initiation and promoter escape by RNA polymerase II. Current Opinion in Genetics & Development. **11**, 209-214.

Eriksson, J. and Löfberg, J. (2000). Development of the hypochord and dorsal aorta in the zebrafish embryo (Danio rerio). Journal of Morphology. **244**, 167-176.

Erokhin, M., Davydova, A., Kyrchanova, O., Parshikov, A., Georgiev, P. and Chetverina, D. (2011). Insulators form gene loops by interacting with promoters in Drosophila. Development. **138**, 4097-4106.

Ferraiuolo, M.A., Rousseau, M., Miyamoto, C., Shenker, S., Wang, X.Q.D., Nadler, M., Blanchette, M. and Dostie, J. (2010). The three-dimensional architecture of Hox cluster silencing. Nucleic Acids Research. **38**, 7472-7484.

Ferrara, T. (2015). Characterization of the *Meis2a* downstream regulatory element *DR-M2de1*. *Masters Thesis* - Appalachain State University, Boone NC.

Fong, Abraham P., Yao, Z., Zhong, Jun W., Johnson, Nathan M., Farr Iii, Gist H., Maves, L., and Tapscott, S.J. (2015). Conversion of myoD to a neurogenic factor: binding site specificity determines lineage. Cell Reports. **10**, 1937-1946.

Fouquet, B., Weinstein, B.M., Serluca, F.C. and Fishman, M.C. (1997). Vessel patterning in the embryo of the zebrafish: Guidance by notochord. Developmental Biology. **183**, 37-48.

Freundlich, H. (2016). Identification and characterization of a highly conserved noncoding element associated with the Meis2 gene; *M2de2*. *Masters Thesis* - Appalachain State University, Boone NC.

Fuentes-Pérez, J.F., Tuhtan, J.A., Carbonell-Baeza, R., Musall, M., Toming, G., Muhammad, N. and Kruusmaa, M. (2015). Current velocity estimation using a lateral line probe. Ecological Engineering. **85**, 296-300.

Gagniuc, P. and Ionescu-Tigoviste, C. (2012). Eukaryotic genomes may exhibit up to 10 generic classes of gene promoters. BMC Genomics. **13**, 2-16.

Gagniuc, P. and Ionescu-Tigoviste, C. (2013). Gene promoters show chormosome-speceifcity and reveal chromosome territories in humans. BMC Genomics. **14**, 2-12.

Geerts, D., Revet, R., Jorritsma, G. Schilderink, N. and Versteeg, R. (2005). MEIS homeobox genes in neuroblastoma. Cancer Letters. **228**, 43-50.

Gehring, W., Qian, Y., Billeter, M., Tokunaga, K., Schier, A., Resendez-Perez, D., Affolter, M., Otting, G. and Wuthrich, K. (1994). Homeodomain-DNA regonition. Cell. **78**, 221- 223.

Germanà, A., Montalbano, G., Guerrera, M.C., Laura, R., Levanti, M., Abbate, F., de Carlos, F., Vega, J.A. and Ciriaco, E. (2009). Sox-2 in taste bud and lateral line system of zebrafish during development. Neuroscience Letters. **467**, 36-39.

Gilbert, S. (2000). *Developmental Biology 6th Edition*, 6th ed. Sinauer Associates, Sunderland, Massachusetts.

Goulding, M. and Gruss, P. (1989). The homeobox in vertebrate development. Current Opinion in Cell Biology. **1**, 1088-1093.

Graham, A., Papalopulu, N. and Krumlauf, R. (1989). The Murine and drosophila homeobox genecomplexes have common features of organization and expression. Cell. **57**, 367-378.

Grosveld, F., Blom van Assendelft, G., Greaves, D. and Kollias, G. (1987). Position-independent, high-level expression of the human β-globin gene in transgenic mice. Cell. **51**, 975-985.

Grzybowska, E.A. (2012). Human intronless genes: Functional groups, associated diseases, evolution, and mRNA processing in absence of splicing. Biochemical and Biophysical Research Communications. **424**, 1-6.

Guth, S.I.E. and Wegner, M. (2008). Having it both ways: Sox protein function between conservation and innovation. Cellular and Molecular Life Sciences. **65**, 3000-3018.

Holland, P. and Hogan, B. (1988). Expression of homeo box genes during mouse development. Gene Development. **2**, 773-782.

Hutchinson, S. and Eisen, J. (2006). Islet1 and Islet2 have equivalent abilities to promote motoneuron formation and to specify motoneuron subtype identity. Development. **133**, 2137-2147.

Jacobs, Y., Schnabel, C. and Cleary, M. (1999). Trimeric association of Hox and TALE homeodomain proteins mediates Hoxb2 hindbrain enhancer activity. Mollecular and Cellular Biology. **19**, 5134-5142.

Jaynes, J. and O'Farrell, P. (1988). Acitvation and repression of transcritption by homeodomain-containing proteins that bind a common site. Nature. **336**, 744-749.

Jenuwein, T. (2001). Translatinig the histone code. Science. **293**, 1074-1080.

Jung, H., Lacombe, J., Mazzoni, E.O., Liem Jr, K.F., Grinstein, J., Mahony, S., Mukhopadhyay, D., Gifford, D.K., Young, R.A. and Anderson, et al. (2010). Global control of motor neuron topography mediated by the repressive actions of a single Hox gene. Neuron. **67**, 781-796.

Khochbin, S. (2001). Histone H1 diversity: bridging regulatory signals to linker histone function. Gene. **271**, 1-12.

Klerk, E. and Hoen, P. (2015). Alternative mRNA transcription, processing, and translation: insights from RNA sequencing. Trends in Genetics. **31**, 128-139.

Krumlauf, R. (1994). Hox genes in vertebrate development. Cell. **78**, 191-201.

Kwan, K., Fujimoto, E., Grabber, C., Mangum, B., Hardy, M., Campbell, D., Parant, J., Yost, J., Kanki, J. and Chien, C. (2007). The Tol2kit: A multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. Developmental Dynamics. **236**, 3088-3099.

Kyrchanova, O. and Georgiev, P. (2014). Chromatin insulators and long-distance interactions in Drosophila. FEBS Letters. **588**, 8-14.

Lawrence, P. and Morata, G. (1976). Compartments in the wing of Drosophila: A Study of the engrailed gene. Developmental Biology. **50**, 321-337.

Lawrence, P. and Morata, G. (1994). Homeobox Genes: Their function in Drosophila segmentation and pattern formation. Cell. **78**, 181-169.

Lettice, L., Heaney, S., Purdie, L., Li, L., Beer, P., Oostra, B., Goode, D., Elgar ,G., Hill, R. and Graaff, E. (2003). A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. Human Molecular Genetics. **12**, 1725-1735.

Lewis, E.B. (1978). A gene complex controlling segmentation in Drosophila. Nature. **276**, 565-570.

Lewis, M. (2011). A Tale of two repressors. Journal of Molecular Biology **409**, 14-27.

Li, Y., Chen, C.-y., Kaye, A.M. and Wasserman, W.W. (2015). The identification of cis-regulatory elements: A review from a machine learning perspective. Biosystems. **138**, 6-17.

Lien, H.-W., Yang, C.-H., Cheng, C.-H., Liao, Y.-F., Han, Y.-S. and Huang, C.-J. (2013). Zinc finger protein 219-like (ZNF219L) and Sox9a regulate synuclein-γ2 (sncgb) expression in the developing notochord of zebrafish. Biochemical and Biophysical Research Communications. **442**, 189-194.

Lin, J.-F., Pan, H.-C., Ma, L.-P., Shen, Y.-Q. and Schachner, M. (2012). The Cell neural adhesion molecule contactin-2 (TAG-1) is beneficial for functional recovery after spinal cord injury in adult zebrafish. PLoS ONE. **7**, e52376.

Liu, J., Wang, Y., Birnbaum, M. and Stoffers, D. (2010). Three-amino-acid-loop-extension homeodomain factor Meis3 regulates cell survival via PDK1. PNAS. **107**, 20494-20499.

Liu, X., Bushnell, D.A. and Kornberg, R.D. (2013). RNA Polymerase II transcription: structure and mechanism. Biochimica et Biophysica Acta. **1829**, 2-8.

Machon, O., Masek, K., Machonova, O., Krauss, S. and Kozmik, Z. (2015). Meis2 is essential for cranial and cardiac neural crest development. BMC Developmental Biology. **15**, 86-93.

Mallo, M., Wellik, D. and Deschamps, J. (2010). Hox genes and regional patterning of the vertebrate body plan. Development. **344**, 7-15.

Mann, R. (2005). Why are Hox genes clustered. BioEssays. **19**, 661-664.

Mann, R. and Chan, S. (1996). Extra specificity from extradenticle. the partnership between HOX and PBX/EXD homeodomain proteins Trends in Genetics. **12**, 258-262.

McGinnis, W., Garber, R., Wirx, J., Kuroiwa, A. and Gehring, W. (1984). A homologus protein-coding sequence in drosophila homeotic genes and its conservation in other metazoans. Cell. **37**, 403-408.

McGinnis, W. and Krumlauf, R. (1992). Homeobox genes and axial patterning. Cell. **68**, 283-302.

Merabet, S. and Lohmann, I. (2015). Towards a new twist in Hox and TALE DNA-binding speceficity. Cell. **32**, 259-261.

Messmer, K., Shen, W.-B., Remington, M. and Fishman, P.S. (2012). Induction of neural differentiation by the transcription factor NeuroD2. International Journal of Developmental Neuroscience. **30**, 105-112.

Minehata, K., Kawahara, A. and Suzuki, T. (2008). Meis1 regulates the development of enothelial cells in zebrafish. Biochemical and Biophysical Research Communications. **374**, 647-652.

Moens, C. and Selleri, L. (2006). Hox cofactors in vertebrate development. Developmental Biology. **291**, 193-206.

Moskow, J., Bullrich, F., Huebner, K., Daar, I. and Buchberg, A. (1995). Meis1, a PBx1-related homeobox gene involved in myeloid leukemia in BXH-2 mice. Mollecular and Cellular Biology. **15**, 5434-5443.

Muse, G.W., Gilchrist, D.A., Nechaev, S., Shah, R., Parker, J.S., Grissom, S.F., Zeitlinger, J. and Adelman, K., (2007). RNA polymerase is poised for activation across the genome. Nat Genet. **39**, 1507-1511.

Myers, P., Eisen, J. and Westerfield, M. (1986). Development and axonal outgrowth of identified motoneurons in the zebrafish. Journal of Neuroscience. **5**, 2278-2289.

Myers, P.Z. (1985). Spinal motor neurons of the larval zebrafish. Journal of Comprehensive Neurology. **236**, 555-561.

Nakamura, T. 2005. Meis and Hox: a mighty pair defeats apoptosis. Inside Blood. **105**, 909.

Nakamura, T., Jenkins, N. and Copeland, N. (1996b). Identification of PBx-related homeobox genes. Oncogene. **13**, 2235-2242.

Nakamura, T., Largaesoada, D., Shaughnessy, J., Jenkins, N. and Copeland, N. (1996a). Cooperative activation of Hoxa and PBx1-related genes in murine myeloid leukaemias. Nature Genetics. **12**, 149-153.

Narendra, V., Rocha, P., An, D., Raviram, R., Skok, J., Mazzoni, E. and Reinberg D. (2015). CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. Science. **347**, 1017-1021.

Ng, H.H., Robert, F., Young, R.A. and Struhl, K. (2016). Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. Molecular Cell. **11**, 709-719.

Nüsslein-Volhard, C. and Wieschaus, E. (1980). Mutations affecting segment number and polarity in Drosophila. Nature. **287**, 795-801.

Pourquié, O., Coltey, M., Teillet, M.A., Ordahl, C. and Le Douarin, N.M. (1993). Control of dorsoventral patterning of somitic derivatives by notochord and floor plate. Proceedings of the National Academy of Sciences. **90**, 5242-5246.

Quinonez, S. and Innis, J. (2014). Human Hox gene disorders. Molecular Genetics and Metabolism. **111**, 4-15.

Roeder, R.G. (1996). The role of general initiation factors in transcription by RNA polymerase II. Trends in Biochemical Sciences. **21**, 327-335.

Rousso, D.L., Gaber, Z.B., Wellik, D., Morrisey, E.E. and Novitch, B.G. (2008). Coordinated actions of the forkhead protein Foxp1 and Hox proteins in the columnar organization of spinal motor neurons. Neuron. **59**, 226-240.

Ruddle, F., Hart, C. and McGinnis, W. (1985). Structural and functional aspects of the mamalian homeo-box sequences. Trends in Genetics: Cell. **1**, 48-51.

Salzberg, A., Eliasa, S., Nachaliela, N., Bonsteina, L., Heniga, C. and Frank, D. (1999). A meis family protein caudalizes neural cell fates in xenopus. Mechanisms of Development. **80**, 3-13.

Sander, K. (1975). Pattern specification in the insect embryo. CIBA Foundation Symposium. **1**, 241-263.

Santini, F., Harmon, L., Carnevale, G. and Alfaro, M. (2009). Did genome duplication drive the orign of teleosts? A compartive study of diversification in ray-finned fishes. BMC Evolutionary Biology - BioMed Central. **9**, 194

Slattery, M., Riley, T., Liu, P., Abe, N., Gomez-Alcala, P., Dror, I., Zhou, T., Rohs, R., Honig, B., Bussemaker, H. and Mann, R. (2011). Cofactor binding evokes latent differences in DNA binding specificity between Hox proteins. Cell. **147**, 1270-1282.

Smith, J., Bollekens, J., Inghirami, G. and Takeshita, K. (1996). Cloning and mapping of the MEIS1 gene, the human homolog of a murine leukemogenic gene. Genomics. **43**, 99-103.

Steelman, S., Moskow, J. and Muzynski, K. (1997). Identification of a conserved family of Meis1-related homeobox genes. Genome Research. **7**, 142-156.

Stemple, D.L. (2004). The notochord. Current Biology. **14**, R873-R874.

Stemple, D.L. (2005). Structure and function of the notochord: an essential organ for chordate development. Development. **132**, 2503-2512.

Tajbakhsh, S., Cavalli, G. and Richet, E. (2011). Integrated gene regulatory circuits: Celebrating the 50th anniversary of the operon model. Molecular Cell. **43**, 505-514.

Thisse, B., Pflumio, S., Fürthauer, M., Loppin, B., Heyer, V., Degrave, A., Woehl, R., Lux, A., Steffan, T., Charbonnier, X.Q. and Thisse, C. (2001). Expression of the zebrafish genome during embryogenesis. ZFIN Direct Data Submission

Topisirovic, I., Svitkin, Y.V., Sonenberg, N. and Shatkin, A.J. (2011). Cap and cap-binding proteins in the control of gene expression. Wiley Interdisciplinary Reviews: RNA. **2**, 277-298.

Toresson, H., Parmar, M. and Campbell, K. (2000). Expression of meis and Pbx genes and their protein products in the developing telencephalon: implications for regional differentiation. Mechanisms of Development. **94**, 183-187.

Toyotome, T., Takahasi, H. and Kamei, K. (2016). MEIS3 is repressed in A549 lung epithelial cells by deoxynivalenol and the repression contributes to the deleterious effect. Journal of Toxological Sciences. **41**, 25-31.

Uribe, R. and Bronner, M. (2015). Meis3 is required for neural crest invasion of the gut during zebrafish enteric nervous system development. Mollecular and Cellular Biology. **247**, 3728-3740.

Venters, B.J. and Pugh, B.F. (2009). How eukaryotic genes are transcribed. Critical reviews in Biochemistry and Molecular Biology. **44**, 117-141.

Vicent, G.P., Wright, R.H.G. and Beato, M. (2016). Linker histones in hormonal gene regulation. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms. **1859**, 520-525.

Vitobello, A., Ferretti, E., Lampe, X., Vilain, N., Ducret, S. Ori. M., Spetz, J., Selleri, L. and Filippo M. (2011). Hox and Pbx factors control retinoic acid synthesis during hindbrain segmentation. Developmental Cell. **4**, 480-482.

Wada, H., Iwasaki, M. and Kawakami, K. (2014). Development of the lateral line canal system through a bone remodeling process in zebrafish. Developmental Biology. **392**, 1-14.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P. and Burge, C.B. (2008). Alternative isoform regulation in human tissue transcriptomes. Nature. **456**, 470-476.

Waskiewicz, A., Rikhof, H., Hernandez, R. and Moens, C. (2001). Zebrafish Meis functions to stabilize Pbx proteins and regulate hindbrain patterning. Development. **128**, 4139-4151.

Waskiewicz, A., Rikhof, H., Hernandez, R. and Moens, C. (2002). Eliminating zebrafish Pbx proteins reveals a hindbrain ground state. Developmental Cell. **3**, 723-733.

West, A., Gaszner, M. and Felsenfeld, G. (2002). Insulators: many functions, many mechanisms. Genes and Development. **16**, 271-278.

Westerfield, M. (2000) *The Zebrafish book. A Guide for the laboratory use of zebrafish (danio rerio) 4th Edition.* Eugene, OR: University of Oregon Press.

Westerfield, M., McMurray, J. and Eisen, J. (1996). Identified motorneurons and their innervation of axial muscles in the zebrafish. Journal of Neuroscience. **6**, 2267-2277.

Whitfield, T., Granato, M., van Eeden, F.J., Schach, U., Brand, M., Furutani-Seiki, M., Haffter, P., Hammerschmidt, M., Heisenberg, C.P. and Jiang, et al. (1996). Mutations affecting development of the zebrafish inner ear and lateral line. Development. **123**, 241-254.

Yamada, T., Pfaff, S.L., Edlund, T. and Jessell, T.M. (1993). Control of cell pattern in the neural-tube - motor-neuron induction by diffusable factors from notochord and floor plate. Cell. **73**, 673-686.

Yamada, T., Placzek, M., Tanaka, H., Dodd, J. and Jessell, T.M. (1991). Control of cell pattern in the developing nervous-system - Polarizing activity of the floor plate and notochord. Cell. **64**, 635-647.

Yaniv, E. and Frank, D. (2010). Mesodermal Wnt signaling organizes the neural plate via Meis3. Developmental Biology. **344**, 476.

Zagozewski, J., Zhang, Q., Pinto, V., Wigle and Eisenstat, D. (2014). The role of homeobox genes in retinal development and disease. Developmental Biology. **393**, 195-208.

Zeitlinger, J. and Stark, A. (2010). Developmental gene regulation in the era of genomics. Developmental Biology. **339**, 230-239.

Zelenchuk, T.A. and Brusés, J.L. (2011). In vivo labeling of zebrafish motor neurons using an mnx1 enhancer and Gal4/UAS. Genesis. **49**, 546-554.

Zerucha, T. and Prince, V. (2000). Cloning and devlopmental expression of Zebrafish meis2 homeobox gene. Mechanisms of Development. **102**, 247-250.

Zhou, D. and Chen S. (1998), Characeterization of a silencer element in the human aromatase gene. Archives of biochemistry and biophysics. **353**, 213-220.

Appendix A- Supplementary Data

Thesis Oligo Primers			
Oligo	Primer Sequence	Length in BP	Description
mm-m2de3-5’	GGCTAAGAAGAAGGCATC	18	5’ end of the murine *Meis2* downstream element 3
mm-m2d3-3’	GAGATCTGTCTACTGTCC	18	3’ end of the murine *Meis2* downstream element 3

Table 1- Oligo Primers used in this thesis project.

Figure A- M2de3 Construct figure

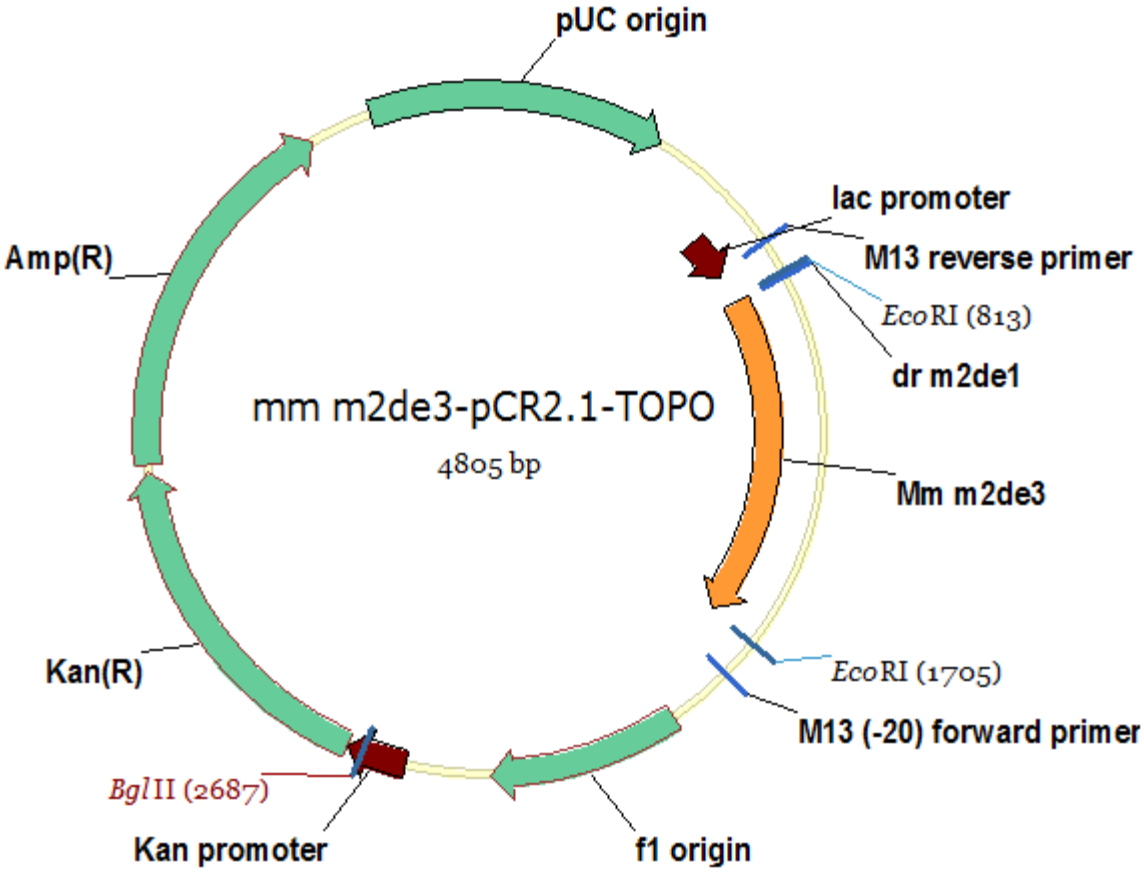
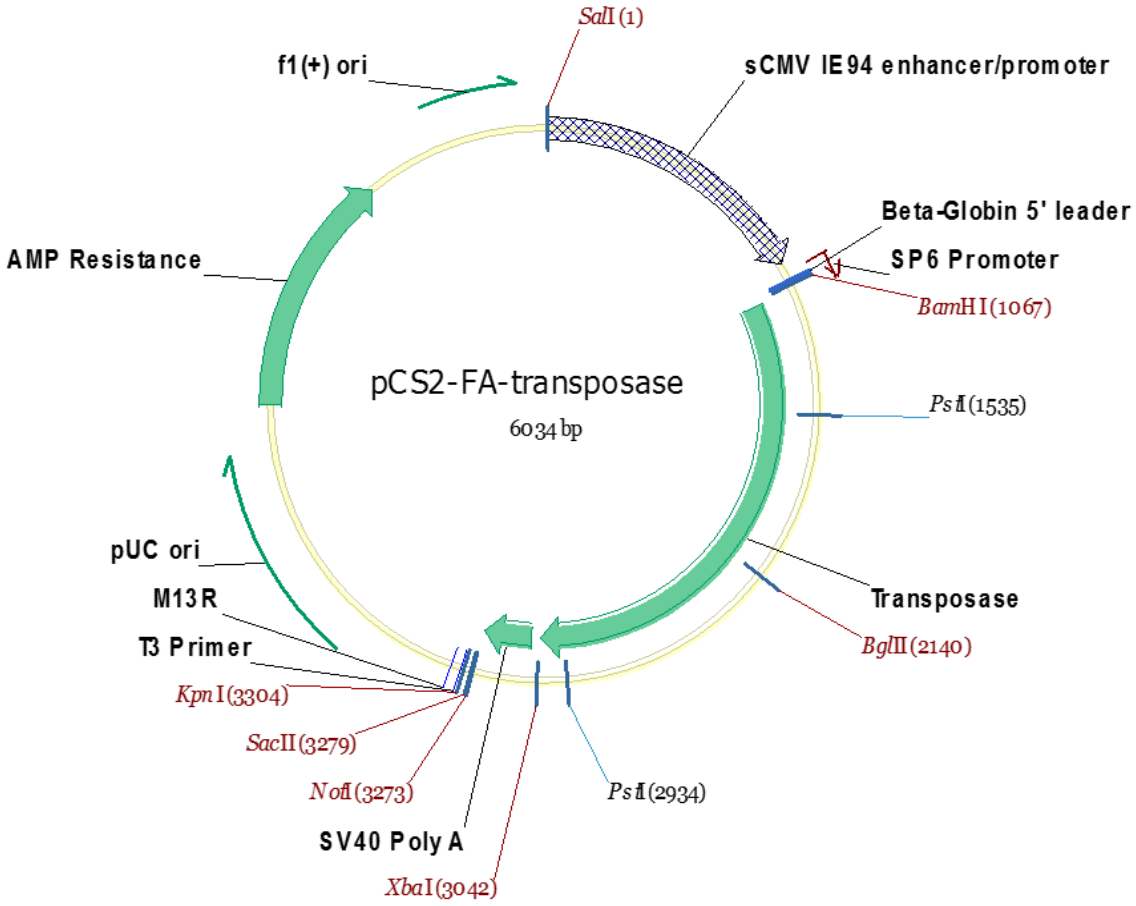


Figure B- Transposase mRNA plasmid Map



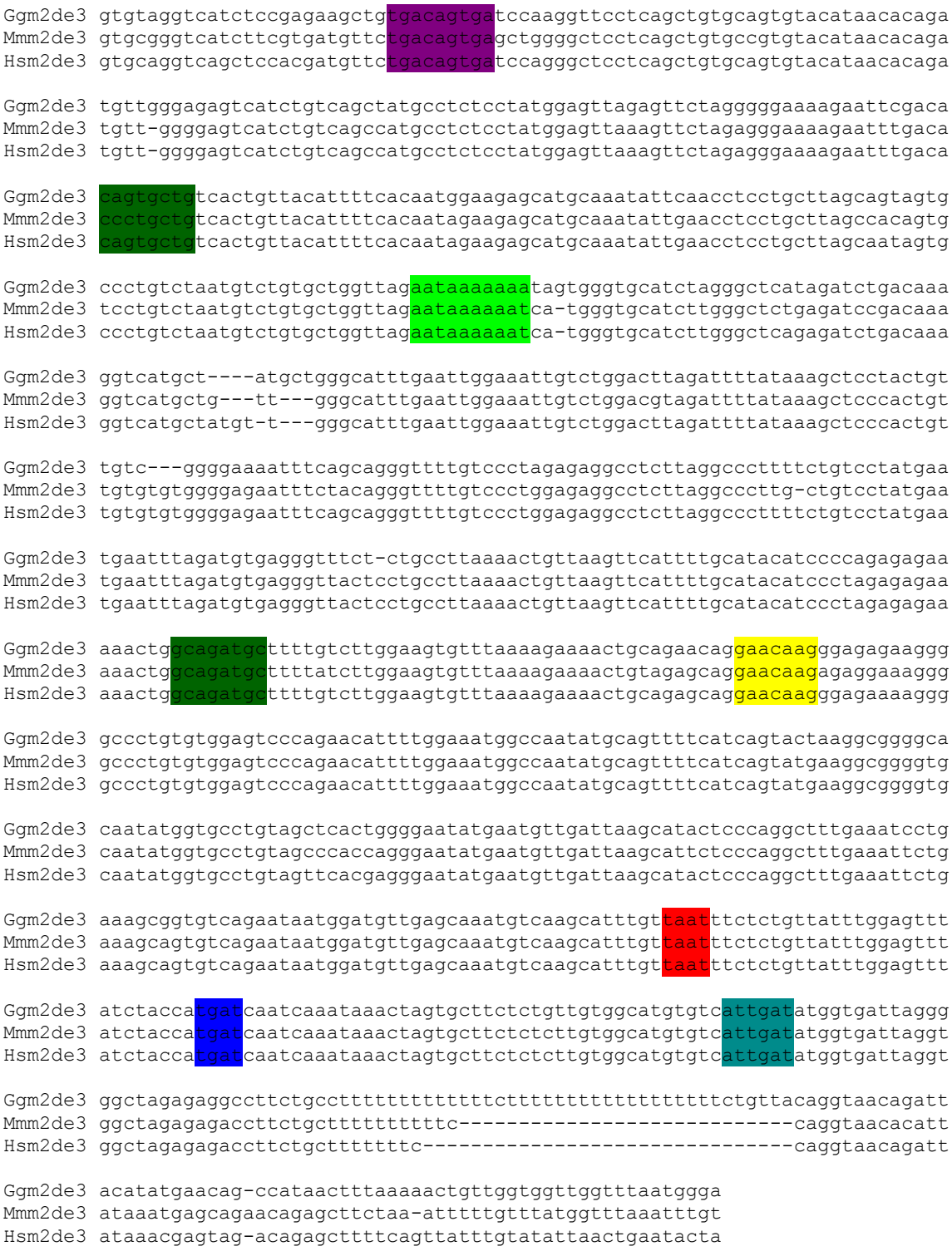


Figure C- PROMO Analysis Multiple sequence alignment of *m2de3*: Gg is *Gallus* (Chicken), Mm is *Mus musculus* (Mouse), Hs is *Homo sapiens* (Human). Species

conservation is unmarked. Binding sites: Blue is a known *Pbx* binding site, Red is a known *Hox* binding site, Yellow is a known *SRY* binding site, Dark Green is a known *Myod* binding site, Teal is a known *GATA-1* binding site, Lime green is a known *HoxD9* binding site, purple is a known *meis1a* binding site.

Appendix B- Abbreviations

Amp– Ampicillin antibiotic
BSA- Bovine Serum Albumin
DepC- DepC treated nuclease free water
DNA- Deoxyribonucleic Acid
Dr– Danieau Rerio or Zebrafish
EDTA- Ethylenediaminetetraacetic acid
GipCo- GipCo brand distilled water
GFP / eGFP– Green Fluorescent Protein
GTP- General Transcription Factors
Hpf– Hours post fertilization of embryos
Hyb. – Hybridization Buffer
LB– Liquid broth
MeOH- Methanol
Mm- Mus Musculus or Mouse Genome
PBS- Phosphate Buffered Saline Solution
PBT- Tween-Phosphate Buffered Saline Solution
PCR- Polymerase Chain Reaction
PFA- Paraformaldehyde
PI Buffer- Pre-Immunization Buffer
Ppm– Parts per million
RNA- Ribonucleic Acid
RO- Reverse osmosis treated water
UV- Ultra violet

Vita

Laiton Steele was born in Statesville, North Carolina, to Wesley and Robinette Steele. He graduated from West Rowan High School in North Carolina in June 2010. Immediately following, he was accepted to Wingate University and began studies for a Bachelor's of Science in the fall of 2010. In the spring of 2014, he was awarded a Bachelor's of Science in Biology from Wingate University. Shortly thereafter, Laiton started his master's research in fall 2014 with Dr. Ted Zerucha at Appalachian State University. Following the completion of this thesis, he plans to pursue further education of a M.D. degree or towards a Ph.D. in Biology.