CONSERVED NON-CODING ELEMENT DERIVED REGULATION OF THE *MEIS2.2* HOMEOBOX GENE DURING EMBRYONIC DEVELOPMENT

A Thesis
by
CODY EVAN BARRETT

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

August 2013
Department of Biology

CONSERVED NON-CODING ELEMENT DERIVED REGULATION OF THE *MEIS2.2* HOMEOBOX GENE DURING EMBRYONIC DEVELOPMENT

A Thesis
by
CODY EVAN BARRETT
August 2013

APPROVED BY:

Dr. Ted Zerucha
Chairperson, Thesis Committee

Dr. Ece Karatan
Member, Thesis Committee

Dr. Mary U. Connell
Member, Thesis Committee

Dr. Susan L. Edwards
Chairperson, Department of Biology

Dr. Edelma D. Huntley
Dean, Research and Graduate Studies

ABSTRACT

CONSERVED NON-CODING ELEMENT DERIVED REGULATION OF THE MEIS2A HOMEOBOX GENE DURING EMBRYONIC DEVELOPMENT

Cody Evan Barrett, B.S., University of North Carolina at Chapel Hill

M.S., Appalachian State University

Chairperson: Dr. Ted Zerucha

Homologues of the *Meis* genes family have been identified in all animals studied. *Meis* genes are a member of the TALE (three amino acid loop extension) superclass of atypical homeobox genes, whose products are characterized by an additional three amino acids between helix 1 and helix 2 of their homeodomain. The vertebrate homeobox-containing *Meis* gene family includes at least four members that are expressed in spatially and temporally conserved fashion throughout development in all vertebrates examined thus far. Products of the *Meis* genes appear to function as cofactors, interacting with both other transcription factors and DNA to assist in the regulation of transcription. Most importantly, they appear to work with the Hox proteins as well as other various homeobox genes' products including the Pbx proteins. At this time, little is known about the regulation of the *Meis* genes. Using phylogenetic footprinting to search for regulatory elements in association with the *Meis* family of homeobox-containing genes, we identified a highly conserved element located downstream of the *Meis2* gene that we have called Meis2 Downstream Element 1 (m2de1). This putative enhancer is conserved in sequence and position across the genomes of all vertebrates examined, including human, mouse, chicken, and zebrafish.

Furthermore, the m2de1 element contains several putative transcription factor binding sites for proteins that could be regulating *Meis2* expression. In this study, we have demonstrated the ability of m2de1 to drive reporter gene expression through microinjection derived transgenic analysis. Reporter gene expression was observed in the developing brain of zebrafish embryos in a manner consistent with endogenous *Meis2* expression, potentially implicating the m2de1 element with *cis*-regulatory function.

ACKNOWLEDGMENTS

There are so many people who have helped me during my years at Appalachian State University. First and foremost, I would like to thank my project advisor Dr. Ted Zerucha for his friendship and encouragement throughout the last three years. His guidance and help were instrumental to the completion of my research project and I feel truly grateful to have been his graduate student. I would also like to thank past and present Zerucha lab members, most notably Kyle Nelson, Caroline Cochrane, and Zach Williams for always lending a hand and ultimately making life in the lab that much more fun.

My utmost gratitude is extended to my committee members, Dr. Ece Karatan and Dr. Mary Connell, for agreeing to work with me even though their own individual workloads were beyond ridiculous. Thank you both for your assistance. I will forever be grateful to Dr. Gary Walker for actually admitting me to graduate school as well as Monique Eckerd for all of her management and supervision in the animal facility.

Special thanks to the Appalachian State University Cratis D. Williams Graduate School, the Appalachian State University Office of Student Research, and the Appalachian State University Graduate Student Association Senate for all of the research and travel funding I have been awarded. Without these monetary awards, the completion of my project would not have been possible.

And last but definitely not least, I would like to thank my family. I have been blessed with the most amazing, loving, supporting parents and sister anybody could ask for. My utmost gratitude goes to them.

DEDICATION

To my family for always being there for me, supporting me in every decision I have ever made, and for loving me unconditionally. Without them, I would not be where I am today. I am forever indebted to them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

FOREWORD

This thesis manuscript and the references, figures, and tables within were arranged in accordance with the instructions detailed for manuscript submission to the peer-reviewed journal *Developmental Biology*, the official journal of the Society for Developmental Biology, published by Elsevier.

INTRODUCTION

The process of fertilization initiates embryonic development which then proceeds through a series of complex, yet coordinated, cellular interactions, resulting in rapid growth, cell division, and ultimately cellular specialization. The underlying commonality behind all organismal development is deoxyribonucleic acid, or DNA. In a process known as transcription, the genetic information stored in DNA is copied into an intermediate structure known as ribonucleic acid, or RNA. The mRNA is then decoded into protein in a procedure called translation. These processes occur at all times in cells for homeostasis and repair and are especially important during development. To insure that these genetic processes proceed in a correct and efficient manner, a series of regulatory steps exist in the gene expression mechanisms of eukaryotes and prokaryotes. This regulation includes: transcriptional control, mRNA processing, mRNA transport control, translation control, and protein control (Alberts, 2008). All these steps taken together help to insure that gene expression, and therefore, embryonic development proceeds properly, resulting in a final, functional adult organism.

The first level of gene regulation is that of transcriptional control. Arguably the most important and diverse level of regulation, transcriptional control consists of a series of steps working together to regulate when and how often a gene is transcribed (Alberts, 2008; Griffiths, 2000; Orphanides et al., 1996; Zawel et al., 1995). Transcription in eukaryotes is a complex process involving numerous proteins and their subsequent interactions with each other as well as the DNA to be utilized (Orphanides et al., 1996; Zawel et al., 1995). The

process of transcription depends on the General Transcription Factors (GTFs); called general because while other transcriptional proteins necessary may vary from gene to gene, the GTFs must be present for any eukaryotic gene expression event (Alberts, 2008; Orphanides et al., 1996). The GTFs work to recruit RNA Polymerase II (RNAPII) to the promoter, help to pull apart the two strands of DNA to allow RNAPII access to the DNA template, and to release the RNAPII from the promoter once transcription has begun (Orphanides et al., 1996). The first GTF to bind to the DNA is Transcription Factor IID (TFIID) (Nikolov et al., 1996; Orphanides et al., 1996; Sawadogo and Sentenac, 1990). One of TFIID's numerous protein subunits is TATA Binding Protein (TBP), which binds specifically to the DNA sequence known as the TATA Box (Li et al., 1999; Nikolov et al., 1996; Orphanides et al., 1996). The TATA Box, although not present in all eukaryotic gene promoters, is one of the most significant DNA sequences utilized by transcriptional proteins to begin transcription (Mathis and Chambon, 1981). Usually located 25 base pairs upstream of the transcriptional start site, it is believed that the TATA site also facilitates subsequent DNA unwinding due to the abundance of the weaker bonded adenines and thymines in these areas (Juo et al., 1996; Kim et al., 1993; Nakatani et al., 1990). Once TFIID has bound to the promoter through its TBP subunit, additional GTFs such as Transcription Factor IIB (TFIIB), Transcription Factor IIE (TFIIE), and Transcription Factor IIH (TFIIH), as well as RNAPII are recruited to the site and bind to each other forming what is called the transcription initiation complex (Buratowski et al., 1989; Roeder, 1991). The GTF TFIIH has a functional helicase that uses ATP to unwind nucleosomes and dislodge histones resulting in the actual unwinding of the DNA helix (Drapkin et al., 1994; Svejstrup et al., 1996). RNAPII, previously recruited to the promoter site, is phosphorylated by a protein kinase subunit of TFIIH, and is released from

the transcription initiation complex to begin elongation (Akoulitchev et al., 1995; Svejstrup et al., 1996). In elongation, the gene of interest is copied into a complimentary, antiparallel mRNA strand through the action of the RNAPII. One of the most important facets in association with the regulation of gene expression is the inherent ability of RNAPII to proofread its own mistakes in mRNA production (Sydow and Cramer, 2009; Thomas et al., 1998). Such a procedure insures that the mRNA is true to its DNA parent strand and that subsequent expression steps produce a viable, functional protein. Another important gene regulation step associated with the RNAPII is that if for some reason it becomes dislodged from the DNA template, it cannot rejoin where it left off (O'Shea-Greenfield and Smale, 1992; Pankotai and Soutoglou, 2013). In fact, the RNAPII must start all over from the promoter (O'Shea-Greenfield and Smale, 1992). This safeguard makes it impossible for the RNAPII to produce an abundance of mRNA fragments; it must transcribe the whole DNA sequence. Upon completion of transcription, the single stranded mRNA that results serves to deliver the genomic information to the protein producing structures located in the cell's cytosol.

Actual eukaryotic transcription events require the implementation of many more proteins and sequences found on the template DNA aside from the promoter and the GTFs to insure successful production of mRNA (Davidson, 2006; Frith et al., 2001). These additional parameters serve to guard against unnecessary and unwanted mRNA production, making sure that transcription occurs only when spatially and temporally necessary (Bustos et al., 1991; Fisher et al., 2006; Kwan et al., 2001). *Cis*-regulatory elements are regions of DNA that regulate the expression of a gene located on the same molecule of DNA as itself and have been found to be especially clustered near genes important for embryonic development

(Alberts, 2008; Kikuta et al., 2007b; Navratilova and Becker, 2009). Unlike the promoter regions which are always found upstream of their associated gene, *cis*-regulatory elements can be found upstream, downstream, within introns of genes, or up to 15000 base pairs away from the initiation complex (Davidson, 2006; Dutton et al., 2008; Echelard et al., 1994; Kikuta et al., 2007b; Valverde-Garduno et al., 2004). These *cis*-regulatory elements are not expressed (Kikuta et al., 2007b). Instead, they are highly conserved sequences ranging in size but usually under 1000 nucleotides long that serve as binding sites for additional transcriptional regulatory proteins (Woolfe et al., 2005). *Cis*-regulatory elements show an extreme amount of sequence conservation due to their importance in gene regulation, and consequent resistance to evolutionary strain (Engstrom et al., 2007; Navratilova and Becker, 2009). Because they are highly conserved, phylogenetic footprinting can be used to identify and locate potential *cis*-regulatory elements (Fisher et al., 2006; Kikuta et al., 2007a; Kikuta et al., 2007b; Woolfe et al., 2005). The process of phylogenetic footprinting involves the comparison of specific regions of DNA believed to contain Highly Conserved Noncoding Elements (HCNEs) across multiple species (Navratilova and Becker, 2009; Woolfe et al., 2005). Previous research has shown that HCNEs often have *cis*-regulatory function, although such assumptions must still be experimentally tested (Fisher et al., 2006; Woolfe et al., 2005).

Currently, it is not known if new *cis*-regulatory elements can evolve from random sequences (Ludwig, 2002). In fact, it is believed that the existing *cis*-regulatory elements are modified, resulting in elements with novel function (Blader et al., 2004; Ludwig, 2002). Exactly how *cis*-regulatory elements do so is currently unknown, but many promising hypotheses exist. The first of these hypotheses states that the new *cis*-regulatory sequences

could be formed by rearrangement and sequence modification; specifically the addition or subtraction of protein binding sequences from existing elements, so that they gain a new function (Blader et al., 2004; Ludwig, 2002). Another reasonable hypothesis proposes that a previously existing element simply develops an association with a previously unassociated gene resulting in a new regulatory function (Blader et al., 2004; Ludwig, 2002).

As mentioned previously every eukaryotic transcriptional event utilizes the GTFs as well as RNAPII meaning differential gene expression and diversity cannot stem from such properties. Instead, it is believed that the *cis*-regulatory elements are the factors that help differentiate when and where transcription of a certain gene occurs (Navratilova and Becker, 2009). Each gene has its own subset of specific *cis*-regulatory elements (Navratilova and Becker, 2009; Tumpel et al., 2006). Although conserved in their sequences, the compliment of *cis*-regulatory elements each gene employs varies from gene to gene, and only when those *cis*-regulatory elements are bound by their transcriptional proteins can transcription of that gene occur (Kikuta et al., 2007b; Navratilova and Becker, 2009; Tumpel et al., 2006). Even though each gene has its own specific subset of *cis*-regulatory elements, it appears as if the *cis*-regulatory elements are not relegated to act upon just one specific gene, but can work to direct the expression of several different genes, thus a level of *cis*-regulatory element sharing is believed to be rather common (Gould et al., 1997; Zerucha et al., 2000). The proteins bound to the *cis*-regulatory sequences work together to regulate the expression of a gene in numerous ways, but one of the most unique and influential is that they can cause the DNA helix to loop back upon itself, bringing the bound proteins into contact with the previously described transcription initiation complex (Gibcus and Dekker, 2013; Tolhuis et al., 2002). Thus, they can directly interact with the GTFs, other *cis*-regulatory bound proteins, or even

the RNAPII, directing the general machinery to proceed with transcription (Calo and Wysocka, 2013).

Owing to these numerous characteristics, *cis*-regulatory elements have been implemented in helping to drive the important developmental processes of differentiation, morphological diversification, and body plan formation (Duboule, 1998; Gompel et al., 2005; Pearson et al., 2005; Whiting et al., 1991). As is the case for most genes, most developmentally important genes are under the influence of a variety of multiple *cis*-regulatory elements as well (Woolfe et al., 2005). The method in which *cis*-regulatory elements cause differentiation and diversification stems from the differential utilization of the elements (Gompel et al., 2005). The spatiotemporal expression patterns of a single gene can be driven in vastly different ways depending upon which *cis*-regulatory elements in that gene's compliment of *cis*-regulatory elements are actively bound by transcriptional proteins (Blader et al., 2004; Gompel et al., 2005; Navratilova et al., 2010; Tumpel et al., 2006). Thus, *cis*-regulatory elements can provide a vast source of expression differentiation and organization depending on the subset of elements a particular gene utilizes in any given transcription event (Blader et al., 2004; Gompel et al., 2005; Navratilova et al., 2010; Tumpel et al., 2006). *Cis*-regulatory elements are also thought to play important roles in evolution as diversification and differential gene expression can be altered through the random, slight mutation of their DNA sequences (Stone and Wray, 2001; Wray, 2007). Such mutations could cause novel transcription factor binding resulting in differential gene expression (Stone and Wray, 2001; Wray, 2007). Prior to the advent of phylogenetic studies, it was believed that species were differentiated from each other solely based on the distinctiveness of their genes (Carroll et al., 2008). Through the influx of sequencing and advanced phylogenetics, it

is commonly known now that many genes are conserved across widely divergent organisms, that evolutionary pressures have been unable to act upon said genes, and therefore organismal differentiation and body plan morphologies must be the result of discrepancies and variation in gene regulation (Castillo-Davis et al., 2004; Davidson, 2001). As stated previously, *cis*-regulatory elements can in fact drive such variation and therefore could be a driving force behind organismal evolution and branching (Castillo-Davis et al., 2004; Davidson, 2001).

There exist three different types of *cis*-regulatory elements: enhancers, silencers, and insulators (Bell et al., 2001; Woolfe et al., 2005). Enhancers theoretically work to increase the probability of transcriptional events, driving expression of a gene in an efficient, yet regulated fashion (Fiering et al., 2000; Walters et al., 1995). As with all *cis*-regulatory elements, the transcriptional proteins that bind enhancers, sometimes referred to as activator proteins, subsequently interact and influence other chromatin bound proteins or even the actual protein complex assembled at the promoter to commence transcription (Popham et al., 1989; Szutorisz et al., 2005). The next classification of *cis*-regulatory elements, silencers, is the antithesis of enhancer elements, working to suppress gene expression (Alberts, 2008; Brand et al., 1985; Clark and Docherty, 1993). Silencers are bound by suppressor proteins, which like activator proteins interact with a various assortment of transcriptional proteins, but unlike activators work to decrease or halt gene expression altogether (Clark and Docherty, 1993). The last of the three *cis*-regulatory elements are insulators. Insulators are regions of DNA located between enhancer elements and the promoters under their influence or between silencers and the promoters they direct (Gerasimova and Corces, 2001). Insulators are important in areas where two genes having vastly different roles and

transcriptional expression patterns are located adjacent to one another (Gerasimova and Corces, 2001). In other genetic regions, the insulators serve to inhibit the enhancers and silencers in association with one gene from interfering with the expression of the adjacent gene (Burgess-Beusse et al., 2002; Geyer, 1997).

Enhancers and silencers can also influence the regulation of gene expression events at the transcriptional level by altering the composition and subsequent conformation of the DNA itself. DNA in eukaryotes is most commonly found in the form of chromatin, which consists of the actual DNA molecule as well as associated proteins known as histones working to protect, package, and strengthen the DNA (Alberts, 2008; Bonner et al., 1968). The interaction of the DNA and its histone proteins makes it physically difficult for the transcriptional machinery to bind to and transcribe genes (Grunstein, 1997; Rombauts et al., 2003; Struhl, 1999). Often times, the composition of the DNA can prevent enhancers and silencers from carrying out their roles during transcription (Lee et al., 1993; Rombauts et al., 2003). DNA methylation, which is the addition of methyl groups to the cytosine nucleotides within a *cis*-regulatory element, can prevent proteins from binding to these sequences, inhibiting their influence (Razin and Riggs, 1980; Rombauts et al., 2003; Weber et al., 2007). However, the activators and suppressors that bind to functional enhancers and silencers can regulate the expression of a gene by recruiting enzymes that alter the degree of packing a gene and its histones employ (Struhl, 1999). Generally, the recruited enzymes will work to add or remove acetyl groups from the histones, resulting in looser or tighter chromatin respectively (Struhl, 1999). Some suppressor proteins have even been found to successfully recruit methylation and deacetylation enzymes to the region of DNA undergoing potential transcription in an effort to prevent the RNAPII from gaining access to the DNA strands,

halting gene expression, while some activators have been shown to do the opposite (Struhl, 1999).

Therefore, it is apparent that transcriptional control is a very powerful and diverse realm through which gene expression is controlled and regulated. The utilization of transcriptional control, most notably the actions of *cis*-regulatory elements, is important in the process of embryonic development; an extremely complex procedure that involves strict adherence to regulated progression. As mentioned previously, developing embryos utilize DNA as a blueprint to direct processes that increase the number of cells present, differentiate the exponentially produced cells, and make sure that the functioning of those cells is regulated in an organized manner (Wolpert, 2007). During these processes, body plan and specific morphological characteristics begin to develop as well, facilitated by cell-cell interactions (Wolpert, 2007). Crucial to such developmental progression is the formation of the body axes so that the anterior of the embryo is differentiated from the posterior (anterior/posterior axis), the right from the left (right/left axis), and the back from the underside (dorsal/ventral axis) (Wolpert, 2007). Such axes formation makes it possible for the later steps of gastrulation, neurulation, and organogenesis to occur in an acceptable fashion (Wolpert, 2007). The formation of these axes and development overall is often associated with a specific set of genes known as the homeobox genes.

Homeobox genes were discovered via studies of *Drosophila melanogaster* that displayed homeotic mutations, where one body segment has the identity of another body segment instead of its own (Morata and Lawrence, 1977; Wolpert, 2007). Eventually, it was determined that such malformations are the result of mutations to a certain set of genes, and that those genes play major roles in axis formation and subsequent segment specification

(Wolpert, 2007). Extremely similar genes were subsequently found in numerous other organisms. These genes, eventually grouped together and named Homeobox genes, all contain a highly conserved 180 base pair sequence known as the homeobox that codes for a well conserved 60 amino acid protein domain named the homeodomain (Gehring, 1987; Gehring, 1993; McGinnis et al., 1984). The homeodomain consists of a helix-loop-helix protein domain commonly seen in other proteins that have the capability of binding DNA (Otting et al., 1988).

Of special importance for the formation and maintenance of the anterior/posterior (A/P) axis are the *Hox* genes (Lemons and McGinnis, 2006). The *Hox* genes are a subset of the Homeobox genes and have been found in all animal species tested so far (Holland and Garcia-Fernandez, 1996; Lemons and McGinnis, 2006). Although they are important in the development of organisms displaying a distinct A/P axis, *Hox* genes have been found in *Cnidarians*, but it is believed that they have no apparent relative in plants, protozoa, or sponges (Balavoine et al., 2002; Lemons and McGinnis, 2006). The *Hox* genes are normally found grouped as clusters in the same arrangement in relation to each other on the chromosome, and this particular organization is utilized to help regulate them spatially and temporally during development (Lemons and McGinnis, 2006). The *Hox* genes positioned more toward the 3' end of the grouping are expressed more towards anterior regions and earlier on in development, while the *Hox* genes located nearer the 5' end are expressed in the embryos posterior areas and at later times during development (Amores et al., 1998; Duboule, 1998; McGinnis and Krumlauf, 1992). This unique function associated with the *Hox* genes is called colinearity (Amores et al., 1998; Duboule, 1998; McGinnis and Krumlauf, 1992; Prince et al., 1998).

Although most metazoans possess *Hox* gene orthologs and these genes are normally found in the same position in relation to each other in all species, the number of clusters present varies from species to species (Amores et al., 1998). This consequently leads to a wide variation in total *Hox* gene number for each species (Amores et al., 1998; Lemons and McGinnis, 2006). In most invertebrates, there exists only one cluster of the *Hox* genes with varying numbers of actual *Hox* genes, although eight are usually attributed to insects (Amores et al., 1998; Holland and Garcia-Fernandez, 1996). In contrast, most tetrapod vertebrates have been found to contain four *Hox* gene clusters with as many as 13 genes existing in a cluster (Greer et al., 2000; Lemons and McGinnis, 2006). These extra clusters have subsequently been designated as *HoxA*, *HoxB*, *HoxC*, and *HoxD*, and each individual gene within those clusters has been assigned a paralog number for organizational purposes (Greer et al., 2000; Lemons and McGinnis, 2006). The number of clusters presented previously only serves as a baseline for predicting how many *Hox* gene clusters actually exist in a species, as variations from the norm often times are found. Some invertebrates have experienced a severe degradation in the conservation and organization of their *Hox* genes, although they still technically possess them (Wagner et al., 2003). With the vertebrates and other chordates, these variations are often times the result of whole genome duplication events that occurred during their evolutionary history (Amores et al., 1998; Kuraku and Meyer, 2009). In fact, genome duplication was the driving force behind the formation of three extra *Hox* gene clusters found in the tetrapod lineage that are not seen in invertebrate genomes (Amores et al., 1998). Another perfect example of such variation is seen in the teleosts, which although vertebrates, possess seven *Hox* clusters and 49 total *Hox* genes (Amores et al., 1998; Brunet et al., 2006; Prince et al., 1998; Prohaska and Stadler, 2004). It

is believed that the teleost fishes, the lineage which includes zebrafish (*Danio rerio*), underwent an additional teleost specific genome duplication event after the evolutionary divergence of ray-finned and lobe-finned fishes, just before teleost expansion, resulting in a total of eight *Hox* gene clusters (Amores et al., 1998; Brunet et al., 2006; Prohaska and Stadler, 2004). Somewhere within that evolutionary history, one of the clusters was lost, leaving the seven clusters present today (Brunet et al., 2006; Prohaska and Stadler, 2004). Research has shown that the genome duplication events may have coincided with the increasing complexity and expansion of various body plans seen during the course of evolutionary history (Venkatesh, 2003). Such duplications would have presented numerous new genes and allowed for novel developmental control mechanisms, resulting in increased axial diversity (Holland and Garcia-Fernandez, 1996).

The methodology behind how and why the *Hox* genes play such a major role in embryonic development stems from the fact that as mentioned previously, the homeodomain present within their protein structures have the ability to bind DNA (Dorn et al., 1994; Gehring, 1993; Gehring et al., 1994). Therefore, the *Hox* genes code for transcription factors that bind to *cis*-regulatory elements associated with developmentally important genes, helping direct those genes' regulation and activation (Dorn et al., 1994; Gehring et al., 1994). Each *Hox* protein, regulated in a highly stringent spatiotemporal manner, serves to direct the identity and morphology of those regions in which they are expressed (Moens and Selleri, 2006). However, such Hox-DNA binding has been shown to be unspecific and inefficient by itself; a major problem considering the implications the *Hox* genes have concerning the developmental processes (Ekker et al., 1994; Moens and Selleri, 2006; Waskiewicz et al., 2001). Residue mapping studies concerning the Hox proteins revealed that individual Hox

proteins could be distinguished from each other by differences in amino acids away from their areas that bind DNA; more specifically in the areas in which they would interact with other various proteins (Mann and Affolter, 1998; Sharkey et al., 1997). Thus, in addition to binding DNA at the *cis*-regulatory sequences, it would appear that the Hox transcription factors have the ability to bind other proteins (Mann and Affolter, 1998). By binding other transcriptional proteins, known as cofactors, binding specificity is increased simply by increasing the DNA surface area to which proteins are bound, resulting in a much more regulated gene transcription event (Mann and Affolter, 1998; Moens and Selleri, 2006).

Many different groups of cofactors exist, one of them being the Three Amino Acid Loop Extension (TALE) class of homeobox genes, named after the presence of an additional three amino acids in the region between helix one and helix two of the homeodomain (Burglin, 1997). The TALE class includes, but is not limited to, the MEIS, PBC, KNOX, IRO, and recently identified MKX subclasses of genes (Burglin, 1997; Mukherjee and Burglin, 2007). The *Meis* genes, short for Myeloid Ecotropic Leukemia Virus Integration Site, were first identified in murine models when the Myeloid Ecotropic Leukemia Virus integrated itself into the previously unidentified *Meis* gene's open reading frame (Moskow et al., 1995). Ensuing analysis lead to the discovery of *Meis* gene paralogs in a wide variety of divergent organisms and the summation of such research led ultimately to the categorization of the *Meis* gene family. In vertebrates, three separate *Meis* genes were found although it appears as if teleosts have a total of four: *Meis1*, *Meis2*, *Meis3*, and *Meis4* (Geerts et al., 2005; Moskow et al., 1995; Nakamura et al., 1996; Waskiewicz et al., 2001). A *Meis* gene family ortholog was found in invertebrates as well as evidenced by the *homothorax* gene in the fruit fly, *Drosophila melanogaster* (Kurant et al., 1998; Pai et al., 1998; Rieckhof et al.,

1997; Ryoo et al., 1999). Interestingly, similarities between the *Meis* genes and *Knox* genes in plants have been found, and all indications are that they diverged from a common ancestor (Becker et al., 2002; Burglin, 1997; Burglin, 1998). The vertebrate genomes examined and observed to contain *Meis* gene orthologs include humans (Geerts et al., 2005; Lawrence et al., 1999; Smith et al., 1997; Steelman et al., 1997), chicken *Gallus gallus* (Coy and Borycki, 2010; Mercader et al., 2000; Sanchez-Guardado et al., 2011), mouse *Mus musculus* (Bomgardner et al., 2003; Cecconi et al., 1997; Nakamura et al., 1996), and the zebrafish *Danio rerio* (Chong et al., 2009; Sagerstrom et al., 2001; Waskiewicz et al., 2001; Zerucha and Prince, 2001;). To add to the complexity, it appears as if *Meis* genes have the capability to be alternatively spliced as well, which suggests that Meis proteins can have various functions (Burglin, 1997; Geerts et al., 2005).

As stated previously, the Meis proteins act as cofactors, binding to *cis*-regulatory sequences found within DNA as well other various transcriptional proteins. In order to carry out its function, the Meis cofactor has two domains set aside for such roles: the homeodomain and the flexible N-terminal domain (Burglin, 1997; Chang et al., 1997; Jacobs et al., 1999; Shanmugam et al., 1999). The homeodomain is the region the Meis cofactor utilizes to bind to the actual DNA (Chang et al., 1997). The Meis homeodomain to DNA interaction, like all other TALE class cofactor and DNA interactions, is generally not believed to be strong and may vary depending on the proteins bound to other areas of the cofactor (Burglin, 1997). The N-terminal domain, commonly called the Meinox domain, is the region which enables and facilitates the binding of a specific transcriptional protein; the Pbx cofactor (Berthelsen et al., 1998; Chang et al., 1997; Choe et al., 2002; Jacobs et al., 1999; Shanmugam et al., 1999). There exists a third region, known as the C-terminal

domain. This domain is responsible for the Meis paralog functional variation, with binding the actual Hox proteins themselves, and in helping to distinguish the Meis cofactors from the other similarly related proteins (Burglin, 1997; Huang et al., 2005; Irimia et al., 2011; Moskow et al., 1995; Shanmugam et al., 1999; Williams et al., 2005).

Transcription factor proteins have the ability to interact with a wide variety of cofactors in order to increase their binding affinity and specificity when binding DNA as a monomer (Chang et al., 1997; Moens and Selleri, 2006; Shanmugam et al., 1999). The transcription factors execute such interactions through the utilization of adjacent DNA-binding half sites and aforementioned conserved motifs flanking their homeodomain (Chang et al., 1996; Chang et al., 1997; Phelan and Featherstone, 1997). The Hox proteins are no exception to these rules. One of the protein-protein interactions seen associated with Hox proteins is the formation of a Hox-Pbx dimer in which the Pbx cofactor uses its C-terminal region to interact with the N-terminal region of the Hox protein (Chang et al., 1997; Shanmugam et al., 1999). In fact, the majority of the Hox proteins (Paralogs 1-10) can interact with Pbx (Shanmugam et al., 1999; Shen et al., 1997;). In addition, Hox transcription factors can utilize their N-terminal structure to form dimers with the Meis C-terminal domain as indicated earlier, although such binding is not that common as only Hox paralogs 9-13 can act with Meis (Shanmugam et al., 1999; Shen et al., 1997;). Such dimer formation occurs on the adjacent DNA binding half sites with the sequence TGACAG-TTA(T/C), but can occur in the absence of DNA (Shen et al., 1997). The Meis cofactor is able to form dimers with Pbx cofactors through the interactions of N-terminal ends and a DNA bipartite sequence consisting of 5’ Pbx and 3’ Meis half sites (Chang et al., 1997; Shanmugam et al., 1999; Waskiewicz et al., 2001). Co-immunoprecipitation studies have

shown that the Meis-Pbx dimers even form in the absence of DNA indicating that the Meis protein is the favored binding partner for Pbx (Chang et al., 1997). Alternatively, Meis will bind Pbx in situations where Pbx is bound to DNA and the Meis binding motif is absent, further indicating the affinity the two proteins have for one another (Shanmugam et al., 1999). The Meis cofactors, similar to the TGIF cofactors, seem to bind the atypical TGACAG(G/C)T sequence with the first six nucleotides acting as the binding core (Chang et al., 1997; Fujino et al., 2001; Zhang et al., 2006). This preferred binding site is unlike that recognized by Hox and other TALE proteins that favor the sequence TAAT (Popperl and Featherstone, 1992). The Meis-Pbx cofactors form dimers on the sequence TGATTGACAG where the 5' TGAT is the preferred binding sequence for the Pbx protein (Chang et al., 1997; Fujino et al., 2001; Huang et al., 2005). The ability for each protein to bind together as well as bind to their complete preferred DNA binding sites indicates that Pbx does not alter the binding specificity of Meis which varies from that seen in Pbx-Hox interactions (Chang et al., 1997). Aside from their obvious transcriptional importance, Meis-Pbx dimers have been implicated in nuclear localization of cofactors as well (Berthelsen et al., 1999; Rieckhof et al., 1997; Saleh et al., 2000; Waskiewicz et al., 2001).

In addition to the dimers listed above, Meis cofactors have been shown to form trimeric complexes with Pbx and Hox proteins (Jacobs et al., 1999; Shanmugam et al., 1999). Such complex interactions serve to increase the binding specificity. The trimeric complex consists of a Pbx-Hox dimer with each protein bound to its respective DNA binding site, and a Meis cofactor attached to the Pbx as well as its own unique DNA binding site (Chang et al., 1997; Jacobs et al., 1999; Mann and Affolter, 1998; Shanmugam et al., 1999; Shen et al., 1999). These trimeric complexes, like the dimers previously mentioned, have been found

formed without the presence of DNA, indicating a potential preference for trimer formation when possible (Shanmugam et al., 1999; Shen et al., 1999). Moreover, the incorporation of the Meis cofactor within such a complex has been proven to increase the chances and stability of trimer formation and function (Jacobs et al., 1999; Schnabel et al., 2000).

As stated previously, there are four members in the *Meis* gene family. The first of the *Meis* genes is *Meis1* which, at early stages of development, is present in posterior regions of the embryo, most notably in the primitive streak, somites, and mesoderm, although no expression is present in the anterior regions (Coy and Borycki, 2010; Maeda et al., 2002). At later stages of development, *Meis1* expression is found in highly restricted areas within the midbrain, hindbrain, and the developing eye and the retina, but can still be found in posterior regions, most notably in the lateral mesoderm, intermediate mesoderm, and neural tube (Bessa et al., 2008; Coy and Borycki, 2010; Maeda et al., 2002; Waskiewicz et al., 2001). *Meis1* expression is also observed throughout the proximal limb buds as well as the branchial arches (Coy and Borycki, 2010; Maeda et al., 2002; Waskiewicz et al., 2001).

As stated earlier, the zebrafish have undergone an additional teleost specific genome duplication event resulting in duplicate copies of the *meis* genes (Amores et al., 1998). Due to the evolutionary loss of coding exons, zebrafish have lost the second *Meis* gene (*meis1.2)* and only possesses the *meis1.1*, or *meis1* gene as it is commonly referred (Irimia et al., 2011). The *Meis1* gene is believed to play a major role in the patterning of hindbrain segmentation (Waskiewicz et al., 2001). In zebrafish studies, mutating the *meis1.1* gene so that its protein was non-functional was shown to result in disorganization in the compartmentalization of the hindbrain (Waskiewicz et al., 2001). Such deformation is believed to be a direct result of Hox protein inhibition caused by insufficient meis1 cofactor functionality (Waskiewicz et al.,

2001). The *Meis1* gene is also believed to help in the embryonic development of other regions of the central nervous system by directing the regulation of *Sox3* and in the positioning of neurons in the dorsal horn of the developing spinal cord (Mojsin and Stevanovic, 2010; Rottkamp et al., 2008). Aside from neuronal roles, *Meis1* is involved in the construction and regulation of limb development (Mercader et al., 1999). In chicken studies, the localization of *Meis1* under the direction of retinoic acid to vertebrate limb proximal areas was found to play a role in cell and regional differentiation along the proximodistal axis of the limb (Mercader et al., 1999; Mercader et al., 2000; Mercader et al., 2005). The *Meis1* gene has also been implicated in the proliferation, positioning, and maintenance of retinal progenitor cells throughout the period of retinal neurogenesis by influencing the expression of cell cycle control proteins (Bessa et al., 2008; French et al., 2007; Heine et al., 2008). *Meis1* genes have been determined to play roles in the hematopoietic processes as *Meis1* deficiencies result in blood vessel abnormalities, arterial deficiencies, and hemorrhaging (Azcoitia et al., 2005; Hisa et al., 2004). Lastly, research has shown that the *Meis1* gene helps direct the expression of *Pax6*, in turn helping drive pancreatic islet cell differentiation (Carbe et al., 2012; Zhang et al., 2006).

The second of the *Meis* genes, *Meis2*, has been revealed to show expression in a wide variety of areas during embryonic development including the forebrain, midbrain, hindbrain, branchial arches, somites, mesoderm, limb buds, neural tube, spinal cord, and developing eye as well as the female genital tract (Cecconi et al., 1997; Coy and Borycki, 2010; Heine et al., 2008; Mercader et al., 2005; Oulad-Abdelghani et al., 1997; Waskiewicz et al., 2001). The teleost specific genome duplication event resulted in two separate *Meis2* genes seen within the genome of zebrafish: *meis2a* (*meis2.2*) and *meis2b* (*meis2.1*) (Biemar et al., 2001;

Waskiewicz et al., 2001; Zerucha and Prince, 2001). In zebrafish, *meis2b* is seen expressed early throughout the embryo, but expression becomes more localized into distinct domains within the forebrain, midbrain, hindbrain, spinal cord, and retina as development continues (Biemar et al., 2001; Zerucha and Prince, 2001). The expression of *meis2a* is similar to that for *meis2b* with expression seen in a few additional areas such as the limb buds and branchial arches (Coy and Borycki, 2010; Mercader et al., 1999; Mercader et al., 2000; Waskiewicz et al., 2001).

Similar to the *Meis1* gene, the *Meis2* gene has been implicated in development of the eye through its interaction with *Pax6* expression, and is believed to help with the proliferation of the undifferentiated cells found in the early retina (Heine et al., 2008; Zhang et al., 2002). Research has also shown that the removal of microRNA miR-204 expression leads to a malformed eye displaying abnormal lens formation and an unusual dorsal-ventral patterning of the retina (Conte et al., 2010). *Meis2* is believed to serve as a target for proper miR-204 function which is elevated in situations where miR-204 is removed leading to the dysfunctional phenotype (Conte et al., 2010). Further implicating the *Meis2* gene with neuron development is the abundance of *Meis2* expression seen in the lateral ganglionic eminence where it marks striatal progenitor neurons and subsets of amacrine cells for differentiation (Bumsted-O'Brien et al., 2007; Toresson et al., 1999). Conversely, *Meis2* expression has been shown to be down regulated prior to actual retinal differentiation by sonic hedgehog signaling indicating that retinal differentiation may vary from region to region (Heine et al., 2009). As with the *Meis1* genes, the *Meis2* genes serve as targets for retinoic acid, help to specify cellular differentiation in the limb, and construct the proximodistal axis of limbs (Mercader et al., 2005). Successful limb development depends

on the restriction of *Meis2* to proximal regions of the limb and such localization is enforced by orthologs of *Drosophila dpp* genes and *Hoxd* genes (Capdevila et al., 1999). The Meis2 proteins also play important roles in the pancreas but in a very different way than that seen by the Meis1 proteins (Liu et al., 2001). Within the developing pancreas, the Meis2b cofactor forms a trimer complex with PDX1 and PBX1b which subsequently participates in the activation of the *ELA1* enhancer in pancreatic acinar cells, in turn helping in the formation of the pancreas's exocrine functions and eventual excretion of digestive enzymes (Liu et al., 2001). Aside from these major roles, the *Meis2* gene has been implicated with the development of the cardiac septum, somitic differentiation leading to subsequent segmentation, and the inhibition of myeloid differentiation (Cecconi et al., 1997; Crowley et al., 2010; Fujino et al., 2001). Of special interest is the discovery that Meis2 cofactors were found to be expressed in the islands of Calleja within the striatum of primate brains (Takahashi et al., 2008). The islands of Calleja are believed to aid in reward processing and the drug addiction pathways (Takahashi et al., 2008).

The third member of the Meis gene family is *Meis3*, whose expression is first seen during the gastrula stage and has its most notable expression within the caudal hindbrain where it eventually fades out completely (Sagerstrom et al., 2001; Salzberg et al., 1999; Waskiewicz et al., 2001). In time, during later steps of development including segmentation, *Meis3* expression is seen in the somites, the budding pectoral fin, the neural tube, and cells immediately adjacent to embryonic pancreatic cells (diIorio et al., 2007; Sagerstrom et al., 2001; Waskiewicz et al., 2001). As is the case with the *meis1.2* gene, it appears as if the second *Meis3* gene (*meis3.2)* has been lost from the zebrafish genome since the teleost specific genome duplication event (Waskiewicz et al., 2001).

The Meis3 cofactors are extremely important in the formation of the posterior hindbrain (Choe et al., 2002; Salzberg et al., 1999). Ectopic *Meis3* expression in the anterior regions of embryos can lead to a loss of forebrain, midbrain, and eye structures and an increased expression of posterior neural markers in those same areas (Salzberg et al., 1999; Vlachakis et al., 2001). Research has shown that such truncations and malformations are the result of the proteins Hoxb1b, Pbx4, and Meis3 forming a trimer complex (Vlachakis et al., 2000; Vlachakis et al., 2001). This trimer is found endogenously during development in the hindbrain to aid in specification and subsequent development, but can also influence the expression profiles and development in other anatomical areas (Vlachakis et al., 2001). In *Meis3* antisense morpholino oligonucleotides and other loss of function studies, anterior neural structures were expanded towards posterior regions and posterior neural tissues in the hindbrain were disturbed (Choe et al., 2002; Dibner et al., 2001). Thus, it appears as if Meis3 is extremely important in hindbrain development as well as hindbrain differentiation. The *Meis3* gene, like the other two previously described *Meis* genes play a role in pancreatic development and in the regulation of insulin production by controlling endoderm patterning and pancreas gene expression (diIorio et al., 2007). The *Meis3* gene has also been implicated in histone acetylation as well as helping in the organization of the neural plate (Choe et al., 2009; Elkouby et al., 2010). Although they have been found in zebrafish, the *Meis4* genes have not been exclusively analyzed for expression or function since being identified previously (Waskiewicz et al., 2001).

Thus, it is evident how extensive the *Meis* genes expression patterns are in embryos and how important those same genes are during the processes of embryonic development. Although much has been revealed concerning the *Meis* gene's expression patterns and

developmental roles, little is known about *Meis* gene regulation. In order to gain a better understanding of *Meis* gene regulation, putative *cis*-regulatory elements must first be identified (Fisher et al., 2006; Hughes et al., 2005). Due to the selective evolutionary pressure and the subsequent sequence conservation that characterizes *cis*-regulatory elements, phylogenetic footprinting can be used to identify and locate HCNEs which have been shown to be candidates for *cis*-regulatory elements (Fisher et al., 2006; Hughes et al., 2005; Kikuta et al., 2007a; Kikuta et al., 2007b; Woolfe et al., 2005). The identification of HCNEs that potentially possess *cis*-regulatory function and direct the expression of the *meis2a* gene was done previously in the Zerucha lab (Wellington and Zerucha, unpublished). Making use of the fully published and publicly available human genome, approximately one million base pairs upstream and downstream of the human *MEIS2* sequence was analyzed for any known Hox transcription factor binding sites. After identifying the numerous binding sites, each one was analyzed via NCBI's BLAST to see if the Hox binding sites and the local sequences surrounding the binding sites were conserved across human (*Homo sapiens*), mouse (*Mus musculus*), chicken (*Gallus gallus*), zebrafish (*Danio rerio*), and pufferfish (*Takifugu rubripes*). Furthermore, multiple sequence alignments were executed to classify each HCNE and elucidate the levels of evolutionary conservation present between the orthologous elements.

The phylogenetic footprinting and analysis of the human *MEIS2* sequence resulted in the discovery of 4 HCNEs that have been named *Meis2* downstream elements (m2de) 1-4 (Wellington and Zerucha, unpublished). The first HCNE, named m2de1, is found 220 Kb downstream of *MEIS2* in humans, 220 Kb downstream of *Meis2* in mice, 120 Kb downstream of *Meis2* in chickens, and 40 Kb downstream of *meis2a* in zebrafish (Fig. 1).

The other three HCNEs that were discovered (m2de2, m2de3, and m2de4) were not found in zebrafish, but were found to be conserved downstream of *Meis2* in humans, mice, and chickens indicating tetrapod lineage specificity (Fig. 1). The second HCNE, named m2de2, is found 210 Kb downstream of *MEIS2* in humans, 210 Kb downstream of *Meis2* in mice, and 130 Kb downstream of *Meis2* in chickens (Fig. 1). The third HCNE, named m2de3, is found 250 Kb downstream of *MEIS2* in humans, 260 Kb downstream of *Meis2* in mice, and 150 Kb downstream of *Meis2* in chickens (Fig. 1). The fourth HCNE, named m2de4, is found 440 Kb downstream of *MEIS2* in humans, 415 Kb downstream of *Meis2* in mice, and 215 Kb downstream of *Meis2* in chickens (Fig. 1). Although four HCNEs were found, the research presented in this thesis concerns the m2de1 element only.

For the m2de1 element, it was found that a 450 nucleotide region is conserved within which exists a 260 nucleotide region that is highly conserved having 68% sequence conservation between humans and zebrafish (Fig. 2). It cannot be ruled out that there may be important regions that are less well conserved that extend beyond this as well (Fig. 2). Within the 260 nucleotide m2de1 element core, there are two conserved transcription factor binding sites for Hox proteins (ATTA / TAAT) (Popperl and Featherstone, 1992), a Pbx transcription factor binding site (ATCA / TGAT) (Chang et al., 1997; Fujino et al., 2001; Huang et al., 2005), and a conserved Meis transcription factor binding site (CTGTC / GACAG) (Chang et al., 1997; Fujino et al., 2001; Zhang et al., 2006) (Fig. 2). Additionally, three potential Hox transcription factor binding sites (ATTA / TAAT) have been identified within adjacent regions as well (Fig. 2).

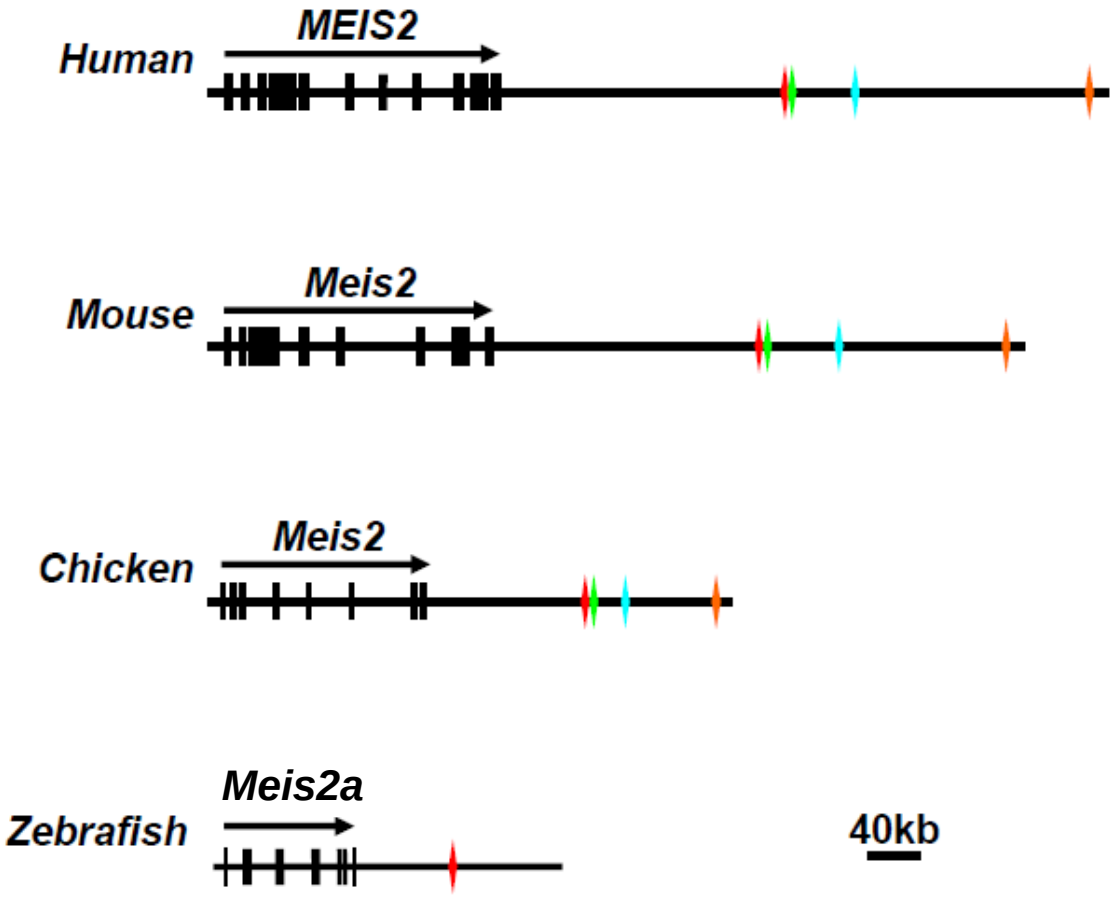


Fig. 1. Schematic diagram of the relative positions of the m2de1 elements. The relative positioning of the discovered m2de (1-4) elements in comparison to the *Meis2* gene and to each other in the genomes of humans, mouse, chicken and zebrafish is revealed. The diamonds shown above mark the positioning of the m2de (1-4) elements: red represents the m2de1 element, green represents the m2de2 element, blue represents the m2de3 element, and orange represents the m2de4 element.

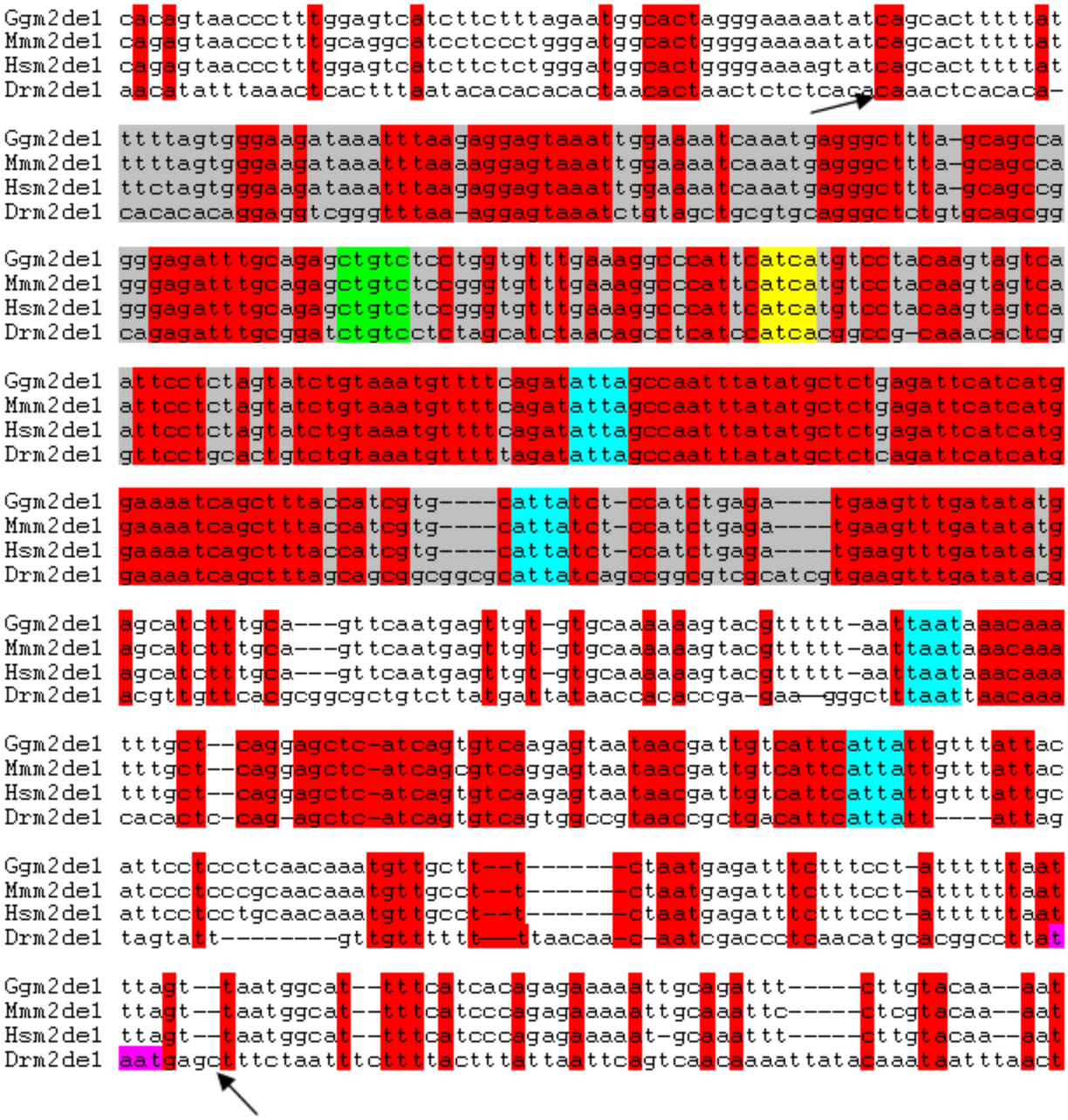


Fig. 2. Multiple sequence alignment for the m2de1 element. The multiple sequence alignment was executed in an effort to elucidate areas of conservation within the m2de1 element across chicken (*Gallus Gallus*; Gg), mouse (*Mus musculus*; Mm), human (*Homo sapiens*; Hs), and zebrafish (*Danio rerio*; Dr). In all, 585 base pairs of sequence that may possess *cis*-regulatory function are shown aligned. Areas highlighted in red are conserved across all vertebrates analyzed. Areas highlighted in grey represent the 260 base pairs of sequence that show the highest amount of sequence conservation. Arrows point to the 2 nucleotides that form the boundaries for the 450 base pair sequence chosen to be used in the m2de1 vector construction. The area highlighted in green is a potential Meis transcription factor binding site. The area highlighted in yellow is a potential Pbx transcription factor

binding site. The areas highlighted in light blue are potential Hox transcription factor binding sites. The area highlighted in purple is another potential Hox transcription factor binding site that was mutated using PCR driven nucleotide mutagenesis. Any binding sites that are not within the sequence indicated by the arrow boundaries have been purposely ignored as they were never included in the sequence chosen for vector construction.

In order to test the functionality of the newly discovered m2de1 element, it must be coupled with a reporter gene and a minimal promoter, introduced into developing embryos, and analyzed for reporter gene expression (Fisher et al., 2006; Kwan et al., 2007; Li et al., 2010; Pauls et al., 2012; Zerucha et al., 2000). The minimal promoter serves to drive reporter gene expression if the adjacent putative *cis*-regulatory element binds transcriptional proteins (Fisher et al., 2006; Li et al., 2010). If the isolated m2de1 element is a functional *cis*-regulatory element, the spatiotemporal characteristics of reporter gene expression patterns should recapitulate the endogenous expression of any gene the HCNE serves to regulate (Fisher et al., 2006; Li et al., 2010; Parker et al., 2011).

The organism *Danio rerio*, commonly known as the zebrafish, is a perfect model organism for transgenic studies (Udvadia and Linney, 2003). The zebrafish is extremely fecund, laying numerous eggs in single fertilization events that are not only fertilized externally, but also develop externally as well (Grunwald and Eisen, 2002; Parichy et al., 2009; Shin and Fishman, 2002). Also, the embryos are transparent and develop rapidly making it easier to visualize internal development, microinject, and ultimately view the embryos for reporter gene expression through confocal microscopy (Udvadia and Linney, 2003). The gene Green Fluorescent Protein (GFP) is the reporter of choice when doing transgenic zebrafish studies due to its stability and ease of visualization within the relatively clear embryos (Amsterdam et al., 1995; Chalfie et al., 1994).

The purpose of this study was to examine the functionality of a putative regulatory element that we hypothesize potentially directs the expression of the *Meis2* gene. Here we propose that the previously discovered m2de1 element has putative *cis*-regulatory function, directing reporter gene expression in a manner consistent with *meis2a* expression in *Danio rerio*.

MATERIALS AND METHODS

Zebrafish Husbandry

The zebrafish used for study were housed in a Marine Biotech Z-mod (Aquatic Habitats, Apopka, FL) closed system. The system was maintained at a constant temperature of 27ºC and was programmed for a regimented 14 hour light, 10 hour dark cycle. Genetically controlled zebrafish strains AB, AB*, and TU (Zebrafish International Resource Center) as well as non-genetically controlled wild-type strains (Carolina Biological, Burlington, NC) were kept in the system. The water quality of the system was checked on a daily basis to insure that the pH of the water was kept between 7.0 and 7.4 and that the conductivity stayed between 450 and 600 milliSiemens per meter (mS/m). The adult zebrafish were housed in individual, 1 L aquaria housing a maximum of six fish per aquarium. Attempts were made to keep the number of males and females in relatively equal ratios within the individual aquariums as well.

To maintain each genetic line and in order to microinject our reporter constructs for screening, adult zebrafish first had to be bred to obtain embryos. Fish of the same genetic line were placed into breeding chambers in a 1 L aquarium the day before crossing such that males were separated from females through the utilization of a specialized tank divider (Aquatic Habitats). The tank divider served two purposes: to prevent earlier than desired breeding and to increase sexual aggressiveness. The number of adult males and females used in each cross varied, but it was determined that two males and three females resulted in the most successful spawning. The next morning, on the day microinjections were to be carried

out, the tank divider was pulled 30 minutes after the system lights had come on to wake the fish. Once in physical proximity to each other, the females would release their eggs and the males would fertilize them externally. Adults were generally allowed to breed for 30 minutes in order to maximize the amount of embryos received. This time constraint was especially important for obtaining embryos that needed to be within the 1-2 cell stages for successful microinjection. Fertilized eggs fell through the porous bottom of the breeding chambers and were harvested using a fine mesh filter to catch the embryos while allowing the excess breeding chamber water to flow through. The adult fish were placed into a separate 1L aquarium and put back into the system. Embryos were then washed using Reverse Osmosis (RO) water from a squirt bottle and placed into a large glass dish where they were raised in 1X Danieau Buffer (58 mM NaCl, 0.7 mM KCl, 0.4 mM $MgSO_4$, 0.6 mM $Ca(NO_3)_2$, 5 mM HEPES pH 7.6). The embryos were placed in a VWR Mini Incubator set at 27ºC for 5 days with dead embryos and debris removed daily.

At 5 days post fertilization (dpf), the fry were placed into 1 L aquariums in the system that were approximately 40% full with 1x Danieau Buffer. Also at this time, feeding with fine particulate dry food (ZM-50; Zeigler) began twice daily. Daily, all dead fry, debris, and 25% of the 1x Danieau Buffer was removed from the aquarium using a turkey baster and wide bore pipette. Fresh 1x Danieau Buffer was then added to replenish dissolved oxygen levels, and to prevent a nitrogenous waste buildup. Fry were kept in this manner until 20 dpf.

At 20 dpf, the majority of the 1x Danieau Buffer was removed from the 1L aquariums and a gentle drip of system water was introduced to the tank in order to acclimate the young fish to the system water. At this point, feedings were reduced to once daily, but the size of the dry food was continually increased as development of the fry progressed (ZM-100, ZM-

200, ZM-300, ZM-400; Zeigler). Once the young zebrafish had begun readily eating the ZM-200 dry food, they were allowed to begin consuming 2 day old live brine shrimp (INVE Aquaculture, Salt Lake City, UT). As the fry continued to grow, they were separated into multiple 1L aquariums such that each individual aquarium had a maximum of 6 zebrafish as noted previously. This was done simply to prevent overpopulation and promote further, optimal development. The zebrafish reached adulthood and sexual maturity around 3 to 4 months post fertilization. Adults were fed Zeigler Adult Zebrafish Complete Diet (Zeigler) once daily while also receiving daily feedings of 2 day old live brine shrimp.

Zebrafish and Mouse HCNE Isolation from Genomic DNA

The identified m2de1 element found in zebrafish was isolated from the total genomic DNA through Polymerase Chain Reaction (PCR) amplification. First, total genomic DNA had to be isolated from a zebrafish of the AB* strain. The selected zebrafish was euthanized using a Tricaine solution (300 ng/L). The fish was then cut up into small pieces using a razor blade and the procedure for "Mammalian Tissue and Mouse/Rat Tail Lysate Protocol" found within the Invitrogen PureLink Genomic DNA Mini Kit (Invitrogen) was followed. 25 mg fish tissue placed in a 1.5 ml microcentrifuge tube with 180 µl PureLink Genomic Digestion Buffer and 20 µl Proteinase K; tube was placed on a heating block set at 55ºC overnight; lysate was centrifuged for 3 minutes at room temperature and transferred to a new 1.5 ml microcentrifuge tube; 20 µl RNase A added to lysate and vortexed; incubated at room temperature for 2 minutes; 200 µl PureLink Genomic Lysis/Binding Buffer added and vortexed; 200 µl 95% ethanol added and vortexed; lysate added to the PureLink Spin Column placed in a Collection Tube and centrifuged at 10000 x g for 1 minute; the Collection Tube containing the elution was discarded and a new Collection Tube was placed

under the spin column; 500 µl Wash Buffer 1 added to the Spin Column and centrifuged at 10000 x g for 1 minute; the Collection Tube containing the elution was discarded and a new Collection Tube was placed under the spin column; 500 µl Wash Buffer 2 added to the Spin Column and centrifuged at 10000 x g for 3 minutes; the Collection Tube containing the elution was discarded and a new 1.5 ml microcentrifuge tube was placed under the spin column; 100 µl RO water was added to the Spin Column and incubated at room temperature for 1 minute; Spin Column centrifuged at 10000 x g for 1 minute; another 100 µl RO water was added to the Spin Column and then centrifuged at 10000 x g for 90 seconds; genomic DNA was quantified by spectrophotometry and stored at -20ºC. Genomic DNA concentration was quantified by NanoDrop® ND-1000 spectrophotometry.

The zebrafish m2de1 element was then isolated from the total genomic DNA by PCR. The two primers, Dr-m2de1-3 and Dr-m2de1-5b (Table 1), used for the PCR protocol were designed bearing in mind the sequence and species that the HCNE was being isolated from in order to insure that correct and efficient amplification occurred.

Table 1. The nomenclature used to identify the various primers as well as the exact sequences of the primers used.

Primer Name	Primer Sequence
Dr-m2de1-5b	TATACCATGGAGGTCGGGTTTAAAGGA
Dr-m2de1-3	GCTCATTATAAGGCCGTGCATG
M13 Forward	GTAAAACGACGGCCAG
M13 Reverse	CAGGAAACAGCTATGAC
5'-attB2-TOPO primer	GGGGACCACTTTGTACAAGAAAGCTGGGT/GAGCTCGG ATCCACTAGTAAC
3'-attB1-TOPO primer	GGGGACAAGTTTGTACAAAAAAGCAGGCT/TCACTATA GGGCGAATTGGG
Sense: 5'-Gateway Seq	GCAATCCTGCAGTGCTGAAA
Antisense: 3'-Gateway Seq	GGACTTCCTACGTCACTGGA

The PCR protocol from the Invitrogen TOPO TA Cloning Procedure was used as a guideline when performing the HCNE isolation. The zebrafish HCNE sequence specific primers listed above (Dr-m2de1-3 and Dr-m2de1-5b) were used in conjunction with Phusion® High-Fidelity DNA Polymerase (NEB M0530L, Ipswich, MA) and the general reaction mixture associated with Phusion® based amplification [1.0 µl zebrafish genomic DNA (20.4 ng/µl), 1.0 µl 3' Dr-m2de1-3 primer (50 pmol/µl), 1.0 µl 5' Dr-m2de1-5b primer (50pmol/µl), 5.0 µl Phusion® 5X HF Buffer, 0.5 µl dNTPs (10 mM), 15.5 µl RO water, and 1.0 µl Phusion® High-Fidelity DNA Polymerase (NEB M0530L)]. The PCR mixture was then placed into a GeneAmp® PCR System 9700 (Applied Bioscience) set at parameters designated for optimal amplification according to the selected primers and HCNE length [one initial DNA melt at 95ºC for 30 minutes, 35 cycles (each cycle consisting of a 30 second melt

at 95ºC, a 30 second annealing at 61ºC, and a 30 second extension at 72ºC), 10 minutes at 72ºC post-cycle completion, and 4ºC for an indefinite amount of time].

Once the PCR reaction was complete, 10 µl of the zebrafish m2de1 product was run on a 1% agarose TBE gel containing ethidium bromide (0.3 µg/ml) set at 115 Volts for 40 minutes. This was done to insure that correct and sufficient amplification had actually occurred. The zebrafish m2de1 PCR product concentration was quantified by NanoDrop® ND-1000 spectrophotometry. The zebrafish m2de1 element was also sent off for Sanger DNA sequencing [performed by Cornell University's Life Sciences Core Laboratories using the Dr-m2de1-3 and Dr-m2de1-5b primers (Table 1)]. HCNEs found in the mouse genome were previously isolated and amplified using a similar protocol as detailed above involving PCR (Nelson, 2011).

HCNE TOPO® TA Cloning®

The isolated m2de1 elements from zebrafish and mouse were then cloned into the pCR®2.1-TOPO® cloning vector. This particular process was carried out for both the zebrafish and mouse HCNEs following standard Invitrogen TOPO TA Cloning® protocol (Invitrogen 45-0641). The addition of 3' A's, the ligation of each HCNE PCR product into the pCR®2.1-TOPO® cloning vector, the heat shock transformation of the vector into One Shot® TOP10 Chemically Competent *E. coli* cells, the subsequent plasmid isolation through miniprep procedure, and eventual analysis of the resulting plasmid through restriction digestion and gel electrophoresis testing was all carried out by a previous member of the Zerucha lab (Nelson, 2011). The pCR®2.1-TOPO® cloning vector containing the zebrafish m2de1 element was also sent off for Sanger DNA sequencing [performed by Cornell

University's Life Sciences Core Laboratories using universal M13 Forward and M13 Reverse primers (Table 1)].

Isolating HCNEs from pCR®2.1-TOPO®

The development of the reporter constructs containing the transgenic expression cassettes was done using Gateway® Technology's Cloning System. Vector construction was the result of a two-part subcloning process where the HCNEs were subcloned into the pDONR221 vector and then the pGW_*cfos*EGFP vector. Four sets of reporter constructs were created: construct with the zebrafish m2de1 element in forward orientation (pDr-m2de1F-*cfos*-EGFP), construct with the zebrafish m2de1 element in reverse orientation (pDr-m2de1R-*cfos*-EGFP), construct with mouse m2de1 element in forward orientation (pMm-m2de1F-*cfos*-EGFP), and construct with mouse m2de1 element in reverse orientation (pMm-m2de1R-*cfos*-EGFP). In order to produce the transgenic reporter construct containing the zebrafish element in its reverse orientation, the zebrafish HCNE had to first be isolated out of the pCR®2.1-TOPO® plasmid and prepared for reverse orientation insertion by PCR. The zebrafish m2de1 element was previously prepared for forward orientation insertion (Nelson, 2011). The isolation and amplification of the HCNE was executed using pCR®2.1-TOPO® plasmid sequence specific primers, 5'-attB2-TOPO and 3'-attB1-TOPO (Table 1), and Phusion® DNA polymerase [1.0 µl zebrafish plasmid DNA, 1.0 µl 3'-attB1-TOPO primer (50pmol/µl), 1.0 µl 5'-attB2-TOPO primer (50pmol/µl), 5.0 µl Phusion® 5X HF Buffer, 0.5 µl dNTPs (10 mM), 16.3 µl RO water, and 0.2 µl Phusion® High-Fidelity DNA Polymerase (NEB M0530L)]. The PCR mixture was then placed into a GeneAmp® PCR System 9700 (Applied Bioscience) set at parameters designated for optimal amplification according to the selected primers and HCNE length [one initial DNA melt at 98ºC for 90

seconds, 35 cycles (each cycle consisting of a 30 second melt at 98ºC, a 30 second annealing at 62.9ºC, and a 30 second extension at 72ºC), 7 minutes at 72ºC post-cycle completion, and 4ºC for an indefinite amount of time]. The primers used in the PCR reaction were developed such that the 3'-attB1-TOPO primer had an adjacent attB1 site and the 5'-attB2-TOPO primer had an adjacent attB2 site. Both attB1 and attB2 sequences were sequence extensions added onto the primers and subsequently, the exponentially produced HCNE sequences via the PCR mechanism. The attB1 and attB2 sequences were not necessary for actual primer binding. Furthermore, the 5'primer bound to the pCR®2.1-TOPO® plasmid 5' polyclonal sequence and the 3' primer bound to the pCR®2.1-TOPO® plasmid 3' polyclonal sequence. The resulting PCR product contained the zebrafish m2de1 element flanked by attB1 and attB2 sites (Fig. 3). Once the PCR reaction was complete, 5 µl of the zebrafish m2de1 product was run on a 1% agarose TBE gel containing ethidium bromide (0.3 µg/ml) set at 115 Volts for 45 minutes. This was done to insure that correct and sufficient amplification had actually occurred.

Zebrafish m2de1 Element Ready for Forward Orientation Insertion into the Donor Vector

Zebrafish m2de1 Element Ready for Reverse Orientation Insertion into the Donor Vector

Mouse m2de1 Element Ready for Forward Orientation Insertion into the Donor Vector

Mouse m2de1 Element Ready for Reverse Orientation Insertion into the Donor Vector

Fig. 3. Zebrafish and mouse m2de1 PCR products post pCR®2.1-TOPO® isolation. Representations of the spatial orientations of the zebrafish HCNEs and their respective attB1 and attB2 sites prior to cloning into the donor vector. One of the primers used in the PCR reaction for preparing the HCNE in its forward orientation had an attB2 sequence adjacent to its 3' primer (3'-attB2-TOPO) and the other had an attB1 sequence adjacent to its 5' primer (5'-attB1-TOPO). One of the primers used in the PCR reaction for preparing the HCNE in its reverse orientation had an attB1 sequence adjacent to its 3' primer (3'-attB1-TOPO) and the other had an attB2 sequence adjacent to its 5' primer (5'-attB2-TOPO). Representations of the spatial orientations of the mouse HCNEs and their respective attB1 and attB2 sites prior to cloning into the donor vector are shown as well. In all the sequences above, the m2de1 element is in its forward orientation (5' to 3'). For the elements primed for reverse orientation insertion, not until donor vector recombination via the BP reaction will the HCNE be adapted to its reverse orientation.

The zebrafish m2de1 PCR product was then cleaned up using the Wizard® SV Gel and PCR Clean-Up System (Promega TB308) using provided instructions involving purification by centrifugation. 1 volume of Membrane Binding Solution was added to an equal volume of PCR reaction solution; the PCR mixture was put in an SV Minicolumn that was placed in a provided Collection Tube and allowed to incubate at room temperature for 1 minute; the SV Minicolumn was centrifuged at maximum speed for 1 minute; the liquid in

the Collection Tube was discarded and the Collection Tube was placed back under the SV Minicolumn; 700 µl of Membrane Wash Solution was added to the SV Minicolumn and centrifuged at maximum speed for 1 minute; the liquid in the Collection Tube was discarded and the Collection Tube was placed back under the SV Minicolumn; 500 µl of Membrane Wash Solution was added to the SV Minicolumn and centrifuged at maximum speed for 5 minutes; the Collection Tube was emptied and placed back under the SV Minicolumn; centrifuged at maximum speed for 1 minute; the SV Minicolumn was placed over a 1.5 ml microcentrifuge tube and 50 µl of Nuclease-Free Water was added; SV column was incubated with the water for 1 minute and centrifuged at maximum speed for 1 minute; eluted DNA was stored at -20ºC. The zebrafish m2de1 PCR product concentration was quantified by NanoDrop® ND-1000 spectrophotometry. The mouse m2de1 PCR product was isolated out of the pCR®2.1-TOPO® plasmid and prepared for both forward and reverse orientation insertion via PCR previously (Nelson, 2011). The final mouse m2de1 PCR products possessed the same attB site spatial orientations as the final zebrafish m2de1 PCR products.

Construction of Middle Entry Vectors

Next, the zebrafish m2de1 element was inserted in its reverse orientation into the previously mentioned donor plasmid, pDONR221 in a process known as the BP reaction (Fig. 4). The BP reaction utilizes Gateway® BP Clonase™ II Enzyme Mix (Invitrogen 865071, Grand Island, NY) and enzyme assisted recombination to construct middle entry vectors. The resulting middle entry vector containing the zebrafish HCNE in its reverse orientation was named pME-Dr-m2de1R.

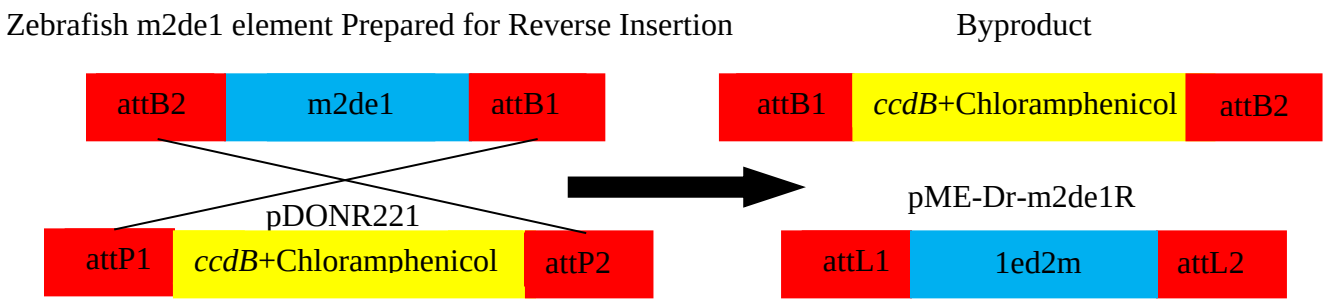


Fig. 4. Schematic diagram of the Gateway® BP reaction. In this reaction, the zebrafish m2de1 element flanked by aatB2 and attB1 cloning sites recombined with the donor vector (pDONR221). With the guidance of the BP Clonase™ II, compatible sequences between attB1-attP1 and attB2-attP2 are recognized, and the HCNE sequence between the attB2 and attB1 sites is translocated into the pDONR221 vector creating the middle entry vector, pME-Dr-m2de1R. The orientation of the attB1-attP1 and attB2-attP2 sites upon recombination results in the reverse insertion of the HCNE into the pDONR221 vector.

The pDONR221 donor plasmid was transformed into One Shot® *ccdB* Survival™ 2 T1R Competent Cells (Invitrogen A10460) by heat shock, subsequently grown on LB-Chloramphenicol (30 mg/µl) + Kanamycin (50 mg/µl) plates, and prepared for use. Half reaction amounts were utilized for the BP reaction for a total reaction mixture of 5.0 µl [0.33 µl pDONR221 (25 fmol), 0.92 µl zebrafish HCNE PCR product (25 fmol), 1.0 µl BP Clonase™II, 2.75 µl TE (10 mM Tris-Cl, 1 mM EDTA pH 7.4)]. The Gateway® Technology Cloning Manual provided a formula to determine the amount of zebrafish m2de1 element prepared for reverse insertion was needed for BP reaction success. Once the amount of BP product in nanograms (ng) needed for the reaction was determined, the volume necessary was calculated using the known concentration of the BP reaction product determined via the spectrophotometry procedure. The formula and numbers used for calculation are shown below:

$$ng = (x \text{ fmoles}) \bullet (\text{HCNE size in base pairs}) \bullet (660 \text{ fg/fmoles}) \bullet (1 \text{ ng}/10^6 \text{ fg})$$

$$ng = (25 \text{ fmoles}) \bullet (645 \text{ bp}) \bullet (660 \text{ fg/fmoles}) \bullet (1 \text{ ng}/10^6 \text{ fg})$$

The BP reaction was performed as indicated by the Gateway® manufacturer instructions. The mixture was vortexed briefly; 1.0 µl BP Clonase™ II Enzyme Mix was added to the reaction mixture; reaction was again vortexed briefly and then centrifuged quickly at low speed to pellet the reaction; and the reaction was run at room temperature overnight. The reaction mixture was then transformed via heat-shock transformation into chemically competent DH5α *E. coli*. 2.5 µl of BP reaction mixture was incubated on ice with 50 µl of chemically competent DH5α *E. coli* cells for 15 minutes; the mixture was then incubated in a 42ºC water bath for 45 seconds; the mixture was incubated on ice for 2 minutes; the cells were transferred to 1 ml room temperature SOC (0.5 g/L NaCl, 20 mM glucose, 20 g/L bacto-tryptone (BD 211705), 5 g/L bacto-yeast extract (BD 212750)); mixture was placed in a shaking incubator at 200 RPM and 37ºC for 90 minutes; the transformed cells were placed on an LB + Kanamycin (50 mg/µl) plate and incubated at 37ºC overnight.

After the transformation protocol, three colonies were chosen from the plate and used to inoculate cultures containing 3 ml LB + Kanamycin (50 mg/µl) and were grown overnight with shaking at 200RPM at 37ºC. After the cultures had been given time to reach maximum growth, the plasmid DNA was isolated from the DH5α *E. coli* using the Promega Wizard Plus SV Miniprep DNA Purification System (Promega N2511). 3.0 ml of each culture was pelleted in a microcentrifuge for 5 minutes at 10000 x g; the cells were resuspended with 250 µl of Cell Resuspension Solution; 250 µl of Cell Lysis Solution was added and inverted to mix; 10 µl Alkaline Protease Solution was added and inverted to mix; mixture was incubated at room temperature for 5 minutes; 350 µl Neutralization Solution was added and inverted to mix; mixture was centrifuged at 10000 x g for 10 minutes; the resulting cleared lysate was

put into the kit provided Spin Column that was placed into the provided Collection Tube; lysate was centrifuged at 10000 x g for 1 minute and the flowthrough was discarded; the Spin Column was placed back in the Collection Tube; 750 µl of Wash Solution was added to the Spin Column and then centrifuged at 10000 x g for 1 minute with the flowthrough being discarded; the Spin Column was placed back in the Collection Tube; 250 µl of Wash Solution was again added to the Spin Column and then centrifuged at 10000 x g for 1 minute with the flowthrough being discarded; the Spin Column was placed back in the Collection Tube; the column was centrifuged at 10000 x g for 2 minutes to remove any residual ethanol and the Collection Tube containing the flowthrough was discarded; the Spin Column was placed in a 1.5 ml microcentrifuge tube and 100 µl of Nuclease-Free Water was added to the Spin Colum; Spin Column was centrifuged at 10000 x g for 1 minute; DNA was stored at -20 ºC. The concentration of the prepared middle entry plasmid was quantified by NanoDrop® ND-1000 spectrophotometry. The plasmid was also subjected to restriction digest testing for 1 hour in a 37ºC water bath using the Restriction Enzymes BglII and EcoRI [6.0 µl middle entry plasmid DNA, 2.0 µl 10X EcoRI Buffer (NEB), 0.5 µl BglII (NEB R0144S), 0.5 µl EcoRI (NEB R0101S), and 11.0 µl RO water]. The digested plasmid was run on a 1% agarose TBE gel containing ethidium bromide (0.3 µg/ml) set at 130 Volts for 1 hour. This was done to insure that the BP reaction had worked successfully and that insertion of the zebrafish HCNE in its reverse orientation had occurred. The middle entry vector containing the zebrafish m2de1 element in its forward orientation (pME-Dr-m2de1F), the middle entry vector containing the mouse m2de1 element in its forward orientation (pME-Mm-m2de1F), and the middle entry vector containing the mouse m2de1 element in its reverse orientation (pME-Mm-m2de1R) were previously constructed (Nelson, 2011).

Construction of Transgenic Reporter Constructs

Next, the middle entry plasmid was used to relocate the reverse oriented zebrafish m2de1 element into the previously mentioned destination vector, pGW_*cfos*EGFP, in a process known as the LR reaction (Fig. 5). The LR reaction utilizes Gateway® LR Clonase™ II Plus Enzyme Mix (Invitrogen 12538120, Grand Island, NY) and enzyme assisted recombination to construct transgenic reporter constructs. The resulting reporter construct containing the zebrafish HCNE in its reverse orientation was named pDr-m2de1R-*cfos*-EGFP.

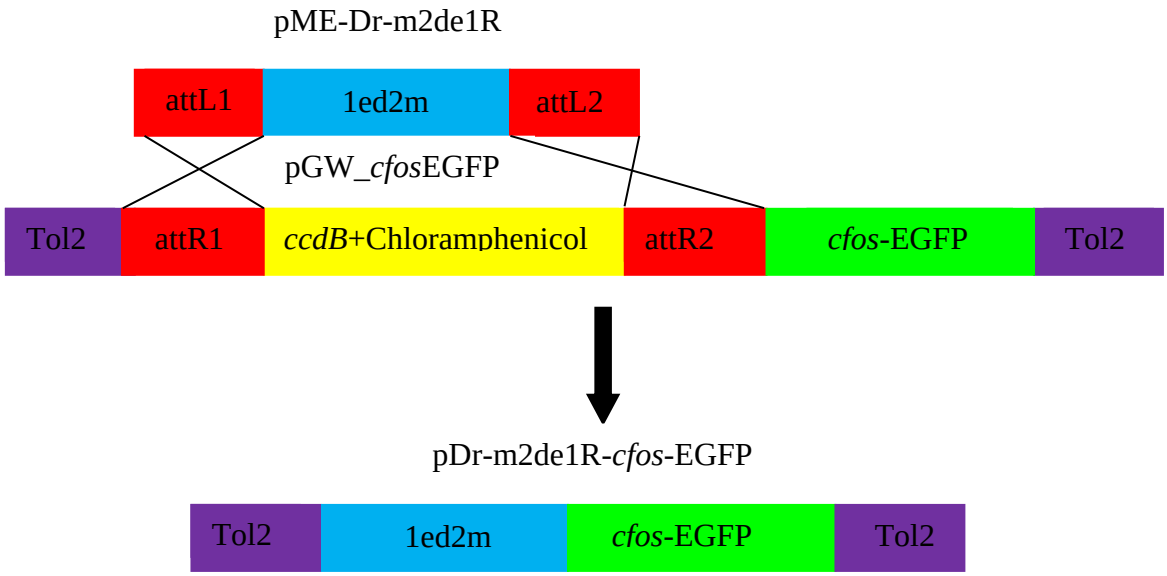


Fig. 5. Schematic diagram of the Gateway® LR reaction. In this reaction, the middle entry vector (pME-Dr-m2de1R) flanked by aatL1 and attL2 cloning sites recombined with the destination vector pGW_*cfos*EGFP). With the guidance of the LR Clonase™ II Plus, compatible sequences between attL1-attR1 and attL2-attR2 are recognized, and the m2de1 sequence between the attL1 and attL2 sites is translocated in its reverse orientation into the pGW_*cfos*EGFP vector creating the final construct, pDr-m2de1R-cfos-EGFP.

The pGW_*cfos*EGFP destination vector was transformed into One Shot® ccdB Survival™ 2 T1R Competent Cells (Invitrogen A10460) by heat shock, subsequently grown

on LB-Chloramphenicol (30 mg/µl) + Ampicillin (100 mg/µl) plates, and prepared for use. Half reaction amounts were utilized for the LR reaction for a total reaction mixture of 5.0 µl [0.64 µl pGW_*cfos*EGFP (10 fmol), 0.35 µl BP reaction product (10 fmol), 1.0 µl LR Clonase^{TM}II Plus Enzyme Mix, 3.01 µl TE (10 mM Tris-Cl, 1 mM EDTA pH 7.4)]. The Gateway® Technology Cloning Manual provided a formula to determine the amount of BP reaction product needed for LR reaction success. Once the amount of BP product in nanograms (ng) needed for the reaction was determined, the volume necessary was calculated using the known concentration of the BP reaction product determined via the spectrophotometry procedure. The formula and numbers used for calculation are shown below:

$$ng = (x \text{ fmoles}) \bullet (\text{HCNE size in base pairs}) \bullet (660 \text{ fg/fmoles}) \bullet (1 \text{ ng}/10^6 \text{ fg})$$

$$ng = (10 \text{ fmoles}) \bullet (645 \text{ bp}) \bullet (660 \text{ fg/fmoles}) \bullet (1 \text{ ng}/10^6 \text{ fg})$$

The LR reaction was performed as indicated by the Gateway® manufacturer instructions [The mixture was vortexed briefly; 1.0 µl LR Clonase^{TM} II Plus Enzyme Mix was added to the reaction mixture; the reaction was again vortexed briefly and then centrifuged quickly at low speed to pellet the reaction; and the reaction was run at room temperature overnight]. The reaction mixture was then transformed via heat-shock transformation into One Shot® TOP10 Chemically Competent *E. coli* (Invitrogen C4040-10). 2.5 µl of LR reaction mixture was incubated on ice with 50 µl of One Shot® TOP10 Chemically Competent *E. coli* cells for 15 minutes; the mixture was then incubated in a 42ºC water bath for 45 seconds; the mixture was incubated on ice for 2 minutes; the cells were transferred to 1 ml room temperature SOC (0.5 g/L NaCl, 20 mM glucose, 20 g/L bacto-tryptone (BD 211705), 5 g/L bacto-yeast extract (BD 212750)); mixture was placed in a

shaking incubator at 200 RPM and 37ºC for 90 minutes; and the transformed cells were placed on an LB + Ampicillin (100 mg/µl) plate and incubated at 37ºC overnight.

After the transformation protocol, three colonies were chosen from the plate and used to inoculate cultures containing 3 ml LB + Ampicillin (100 mg/µl) and were grown overnight with shaking at 200 RPM at 37ºC. After the cultures had been given time to reach maximum growth, the plasmid DNA was isolated from the One Shot® TOP10 Chemically Competent *E. coli* using the Promega Wizard Plus SV Miniprep DNA Purification System (Promega N2511) as previously described. The concentration of the prepared transgenic reporter construct was quantified by NanoDrop® ND-1000 spectrophotometry. The plasmid was also subjected to restriction digest testing for 1 hour in a 37ºC water bath using the Restriction Enzyme EcoRI [6.0 µl transgenic reporter DNA, 2.0 µl 10X EcoRI Buffer (NEB), 1.0 µl EcoRI (NEB R0101S), and 11.0 µl RO water]. The digested plasmid was run on a 1% agarose TBE gel containing ethidium bromide (0.3 µg/ml) set at 130 Volts for 1 hour and 15 minutes. This was done to insure that the LR reaction had worked successfully and that insertion of the zebrafish m2de1 element had occurred. For storage and future utilization, glycerol stocks were made for the transgenic reporter construct containing the zebrafish m2de1 in its reverse orientation. 750 ml culture and 750 ml LB + 30% Glycerol added to Nalgene Cryogenic Vials (Nalgene 5000-0020); mixture was vortexed to mix; the vials were placed in the -80ºC freezer. The reporter construct containing the zebrafish m2de1 element in its forward orientation (pDr-m2de1F-*cfos*-EGFP), the reporter construct containing the mouse m2de1 element in its forward orientation (pMm-m2de1F-*cfos*-EGFP), and the reporter construct containing the mouse m2de1 element in its reverse orientation (pMm-m2de1R-*cfos*-EGFP) were previously constructed (Nelson, 2011). The pDr-m2de1F-*cfos*-EGFP

construct was also sent off for Sanger DNA sequencing [performed by Cornell University's Life Sciences Core Laboratories using the Dr-m2de1-3 and Dr-m2de1-5b primers (Table 1)].

Mutation of a Potential Hox Binding Site

In order to analyze the impacts of a mutated potential transcription factor binding site within the zebrafish m2de1 element, PCR mutagenesis was performed using the QuikChange® II XL Site-Directed Mutagenesis Kit (Stratagene #200521). Furthermore, previous research revealed that a sequence comparable to m2de1 drives *Meis2* expression as well (Parker et al., 2011). Sequence comparison between the zebrafish m2de1 element and the *cis*-regulatory element utilized by Parker et al. showed that the m2de1 element contained an additional Hox binding site, the previously stated TAAT sequence. By eradicating this Hox binding site, it would then be possible to compare expression profiles driven by the two different elements. The binding site mutated was a known Hox binding site bearing the sequence TAAT.

First, primers were designed according to the primer design guidelines included in the QuikChange® II XL Site-Directed Mutagenesis Kit. Both primers (Table 2) were constructed so that they contained the desired mutation (TAAT to TCGA), they annealed to the same sequence on opposite strands of the pDr-m2de1F-*cfos*-EGFP plasmid, and they had a melting temperature higher than 75ºC. Furthermore, an additional guanine located directly adjacent to the Hox binding site was mutated as well (TAATG to TCGAT) via the primers. This TCGAT mutation sequence was chosen because it not only eradicates any Hox binding affinity the TAAT sequence may have had, but it also possesses a ClaI restriction site (5' AT▼CGAT 3') not seen in the original TAAT sequence as an adenine is located adjacent to

the 5' end of the Hox binding site. The additional ClaI site makes it relatively easy to determine if in fact the PCR mutagenesis was successful.

Table 2. The nomenclature used to identify each primer implemented in the mutation of the pDr-m2de1F-*cfos*-EGFP plasmid as well as the exact sequences of the primers. The nucleotides in bold and underlined are those possessing the implemented mutations. The black triangles indicate the recognition and cut site for the ClaI restriction enzyme.

Primer Name	Primer Sequence
Dr-m2de1-mut-R	GAGTTGTACGTGCCGGAAT**AGC▼T**ATCGTTCCCGCTTAAGGTCG
Dr-m2de1-mut-F	CTCAACATGCACGGCCTTAT**▼CGAT**AGCAAGGGCGAATTCCAGG

The pDr-m2de1F-*cfos*-EGFP plasmid was then retrieved from the glycerol stock. Cells from glycerol stocks for the pDr-m2de1F-*cfos*-EGFP plasmid were streaked onto LB + Ampicillin (100 mg/µl) plates and incubated at 37ºC overnight; two colonies were chosen from each plate and used to inoculate cultures containing 3 ml LB + Ampicillin (100 mg/µl) and were grown overnight with shaking at 200 RPM at 37ºC. The plasmid DNA was isolated from the One Shot® TOP10 Chemically Competent *E. coli* using the Promega Wizard Plus SV Miniprep DNA Purification System (Promega N2511) as previously described. The concentration of the prepared plasmid was quantified by NanoDrop® ND-1000 spectrophotometry.

The PCR mutagenesis was then executed following the protocol included with the QuikChange® II XL Site-Directed Mutagenesis Kit utilizing the primers listed in Table 2 (Dr-m2de1-mut-R and Dr-m2de1-mut-F), *PfuUltra* HF DNA Polymerase (Stratagene) and the general reaction mixture associated with *PfuUltra* based amplification [0.15 µl pDr-

m2de1F-*cfos*-EGFP plasmid DNA (10 ng), 0.125 µl Dr-m2de1-mut-R primer (125 ng), 0.125 µl Dr-m2de1-mut-F primer (125 ng), 5.0 µl 10X reaction buffer, 1.0 µl dNTP mix, 3.0 µl of QuikSolution, 40.6 µl RO water, and 1.0 µl *PfuUltra* HF DNA Polymerase (Stratagene)]. The PCR mixture was then placed into a GeneAmp® PCR System 9700 (Applied Bioscience) set at parameters designated for optimal amplification according to the QuikChange® II XL Site-Directed Mutagenesis Kit [1 initial DNA melt at 95ºC for 1 minute, 18 cycles (each cycle consisting of a 50 second melt at 95ºC, a 50 second annealing at 60ºC, and a 8 minute extension at 68ºC), 7 minutes at 68ºC post-cycle completion, and 4ºC for an indefinite amount of time].

After the PCR steps, the QuikChange® II XL Site-Directed Mutagenesis Kit protocol was utilized to retrieve the actual mutated m2de1 element. The PCR reaction tube was incubated on ice for 2 minutes; 1.0 µl of the DpnI restriction enzyme was added directly to the PCR tube, and the solution was mixed by pipetting up and down several times; the PCR tube was spun down in a microcentrifuge at 10000x g for 1 minute; the PCR tube was incubated in a 37ºC water bath for 1 minute; during the incubation, the XL10-Gold ultracompetent (Stratagene) cells were thawed on ice; 45.0 µl of the XL10-Gold ultracompetent cells were put into a prechilled 14 ml Falcon® polypropylene tube (Falcon® 2059); 2.0 µl of the β-Mercaptoethanol mix was added to the XL10-Gold ultracompetent cells and the contents were swirled gently; the XL10-Gold ultracompetent cells were then incubated on ice for 10 minutes, while gently swirling them every 2 minutes; 2.0 µl of the DpnI digested DNA was added to the cells and the reaction was swirled gently and incubated on ice for 30 minutes; during the incubation, NZY+ broth (10.0 g NZ amine, 5.0 g yeast extract, 5.0 g NaCl, 12.5 ml of 1 M $MgCl_2$, 12.5 ml of 1 M $MgSO_4$, 20 ml of 20% glucose,

pH adjusted to 7.5 with NaOH, and adjusted to 1.0 L with RO water) was preheated in a 42ºC water bath; after the incubation, the reaction tube was placed in a 42ºC water bath for 30 seconds; the reaction tube was then placed on ice for 2 minutes; 0.5 ml of the preheated NZY+ broth was added to the reaction tube and the mixture was incubated at 37ºC for 1 hour with shaking at 245 RPM; and after the incubation, 500 µl of the reaction mixture was spread evenly onto two LB + Ampicillin (100 mg/µl) plates and incubated overnight at 37ºC.

After the overnight incubation, two colonies were chosen from the plates and used to inoculate cultures containing 3 ml LB + Ampicillin (100 mg/µl) and were grown overnight with shaking at 200 RPM at 37ºC. Plasmid DNA was isolated from the XL10-Gold ultracompetent cells using the Promega Wizard Plus SV Miniprep DNA Purification System (Promega N2511) as previously described. The concentration of the prepared mutant transgenic reporter construct (pDr-m2de1F-mut-*cfos*-EGFP) was quantified by NanoDrop® ND-1000 spectrophotometry. The pDr-m2de1F-mut-*cfos*-EGFP plasmid and the unmutated pDr-m2de1F-*cfos*-EGFP plasmid were subjected to restriction digest testing overnight in a 37ºC water bath using the Restriction Enzyme ClaI [10.0 µl transgenic reporter construct DNA, 2.0 µl 10X Buffer 4 (NEB), 0.2 µl 100X BSA (NEB),1.0 µl ClaI (NEB R0197S), and 6.8 µl RO water]. The digested plasmids were run on a 1% agarose TBE gel containing ethidium bromide (0.3 µg/ml) set at 115 Volts for 1 hour and 10 minutes. This was done to make sure that actual mutation of the TAAT site and introduction of a new ClaI restriction site had occurred. The pDr-m2de1F-mut-*cfos*-EGFP construct was also sent off for Sanger DNA sequencing [performed by Cornell University's Life Sciences Core Laboratories using the Sense: 5'-Gateway Seq and Antisense: 3'-Gateway Seq primers (Table 1)].

Preparing the Constructs for Transgenic Analysis

In order to retrieve sufficient quantities of the transgenic reporter constructs, purification of all the constructs was performed using the QIAGEN® Plasmid *Plus* Maxi Kit (QIAGEN® 12963). In all, six constructs were prepared for microinjection: construct with the zebrafish m2de1 element in forward orientation (pDr-m2de1F-*cfos*-EGFP), construct with the zebrafish m2de1 element in reverse orientation (pDr-m2de1R-*cfos*-EGFP), construct with mouse m2de1 element in forward orientation (pMm-m2de1F-*cfos*-EGFP), construct with mouse m2de1 element in reverse orientation (pMm-m2de1F-*cfos*-EGFP), positive control construct (pDest-Sox10-mCherry-Tol2CG2; graciously provided by Dr. Chi-Bin Chien), and the construct containing the zebrafish m2de1 element in forward orientation but with a potential binding site sequence mutation (pDr-m2de1F-mut-*cfos*-EGFP). Cells from glycerol stocks for each of the six constructs were streaked onto LB + Ampicillin (100 mg/µl) plates and incubated at 37ºC overnight; three colonies were chosen from each plate and used to inoculate cultures containing 3 ml LB + Ampicillin (100 mg/µl) and were grown 8 hours with shaking at 200 RPM at 37ºC; the starter cultures were subsequently added to Erlenmeyer Flasks containing 500 ml LB + Ampicillin (100 mg/µl) and were grown overnight with shaking at 200 RPM at 37ºC.

The plasmid DNA was then harvested following the procedure as outlined in the QIAGEN® Plasmid *Plus* Maxi Kit (QIAGEN 12963, Hilden, Germany). One hundred ml of culture were pelleted at 6000 x g for 15 minutes at 4ºC and the supernatant was discarded; the pelleted bacteria cells were resuspended in 5 ml of Buffer P1 by pipetting up and down until no clumps were visible; 5 ml Buffer P2 was added to the mixture and gently mixed by inversion until viscous; the viscous mixture was incubated at room temperature for 3 minutes; 5 ml of Buffer S3 was added to the mixture and mixed by inverting six times;

immediately, the mixture was transferred to the QIAfilter™ Maxi Cartridge and incubated at room temperature for 10 minutes; during the incubation, a QIAGEN® Plasmid *Plus* Maxi Spin Column with a Tube Extender was attached to a vacuum manifold hooked up to a vacuum supply source; a QIAfilter™ Plunger was inserted into the QIAfilter™ Maxi Cartridge and the lysate was filtered into a new tube; 5 ml Buffer BB was added to the cleared lysate and was mixed by inverting six times; the lysate was transferred to the QIAGEN® Plasmid *Plus* Maxi Spin Column with a Tube Extender attached to the vacuum manifold; the vacuum source was turned on drawing the liquid through the column; the vacuum source was turned off once all the liquid had been drawn through; 0.7 ml Buffer ETR was added to the QIAGEN® Plasmid *Plus* Maxi Spin Column with a Tube Extender attached to the vacuum manifold; the vacuum source was turned on drawing the liquid through the column; the vacuum source was turned off once all the liquid had been drawn through; 0.7 ml Buffer PE was added to the QIAGEN® Plasmid *Plus* Maxi Spin Column with a Tube Extender attached to the vacuum manifold; the vacuum source was turned on drawing the liquid through the column; the vacuum source was turned off once all the liquid had been drawn through; the QIAGEN® Plasmid *Plus* Maxi Spin Column was removed from the vacuum source and centrifuged in a microcentrifuge at 10000 x g for 1 minute with the excess liquid being discarded; the QIAGEN® Plasmid *Plus* Maxi Spin Column was placed into a 1.5 ml microcentrifuge tube; 400 µl Nuclease-Free Water was added to the center of the QIAGEN® Plasmid *Plus* Maxi Spin Column and incubated for 2 minutes; the mixture was centrifuged for 1 minute at 10000 x g; the supernatant containing the transgenic reporter construct was separated into five 1.5 ml microcentrifuge tubes and placed in the -20ºC

freezer. The concentrations of the prepared transgenic reporter constructs were quantified by NanoDrop® ND-1000 spectrophotometry.

Making Transposase mRNA

The transposase mRNA utilized in the microinjection process was created from glycerol stock DNA graciously provided by Dr. Chi-Bin Chien. Cells from glycerol stocks containing pCS2FA Tol2 transposase plasmid were streaked onto LB + Ampicillin (100 mg/µl) plates and incubated at 37ºC overnight; two colonies were chosen from each plate and used to inoculate cultures containing 3 ml LB + Ampicillin (100 mg/µl) and were grown 8 hours with shaking at 200 RPM at 37ºC; the starter culture with the most growth was subsequently added to Erlenmeyer Flasks containing 500 ml LB + Ampicillin (100 mg/µl) and were grown overnight with shaking at 200 RPM at 37ºC. The plasmid DNA was then harvested following the procedure as outlined previously within the QIAGEN® Plasmid *Plus* Maxi Kit (QIAGEN 12963). After the concentrations of the prepared pCS2FA Tol2 transposase plasmid were quantified by NanoDrop® ND-1000 spectrophotometry, the DNA template was linearized for transcription via digestion with the NotI restriction enzyme (NEB R0189S) at 37ºC overnight [15.0 µl pCS2FA plasmid DNA (20 µg), 10.0 µl 10X Buffer 3 (NEB), 1.0 µl BSA, 8.0 µl NOTI (NEB R0189S), and 66.0 µl DEPC water]. After complete digestion, the NotI enzyme was heat killed by incubating the solution in a heating block at 65ºC for 25 minutes. The now linearized plasmid was purified using the Wizard® SV Gel and PCR Clean-Up System (Promega TB308) as described previously. The concentration of the linearized pCS2FA Tol2 transposase plasmid were quantified by NanoDrop® ND-1000 spectrophotometry, and 2.0 µl plasmid was also run on a 1% agarose TBE gel containing

ethidium bromide (0.3 µg/ml) set at 115 Volts for 1 hour and 10 minutes. This was done to insure that complete digestion had occurred.

The linearized pCS2FA Tol2 transposase template was then transcribed using the mMESSAGE mMACHINE® SP6 RNA Transcription Kit (Ambion® AM1340M, Garden Island, NY). The 10X Reaction Buffer, the 2X NTP/CAP, and the Enzyme Mix were spun briefly to remove RNases; the Enzyme Mix was placed on ice, while the 10X Reaction Buffer and the 2X NTP/CAP were thawed at room temperature; once thawed, the 10X Reaction Buffer and the 2X NTP/CAP were vortexed to insure that any solid particles were completely dissolved; the 10X Reaction Buffer was then kept at room temperature while the 2X NTP/CAP was placed on the ice; the transcription reaction was then assembled in a 1.5 ml microcentrifuge tube: 4.73 µl DEPC water, 1.27 µl pCS2FA Tol2 transposase DNA (1 µg), 10.0 µl 2X NTP/CAP, 2.0 µl 10X Reaction Buffer, and 2.0 µl Enzyme Mix; the reaction was mixed by gently flicking the tube, spun briefly to pellet the reaction, and incubated at in 37ºC water bath for 2 hours; after the incubation, 1.0 µl of TURBO DNase was added to the reaction, the reaction was mixed well, and then incubated at 37ºC for 15 minutes; 30 µl of DEPC water and 30 µl of LiCl Precipitation Solution were added, the reaction was mixed well, and then precipitated at -20ºC for 1 hour; after the precipitation, the reaction was centrifuged at 4ºC for 15 minutes at maximum speed; the supernatant was removed and 1 ml of 75% ethanol was added; the reaction was spun at 4ºC for 10 minutes at maximum speed; the supernatant was removed and the tube was inverted over a paper towel for 30 minutes to insure complete removal of the supernatant; the resulting pellet was resuspended in 20 µl of DEPC water; the mRNA was diluted to 125 ng/µl and aliquoted into fifteen 10 µl aliquots; mRNA was subsequently stored at -20ºC. The concentration of the pCS2FA Tol2

transposase mRNA was quantified by NanoDrop® ND-1000 spectrophotometry, and 2.0 µl mRNA was also run on a 1% agarose TBE gel containing ethidium bromide (0.3 µg/ml) set at 115 Volts for 1 hour and 5 minutes to make sure that the mRNA was the appropriate size.

Microinjections

In order to obtain embryos for microinjection, fish were allowed to mate for 30 minutes as previously described. The 1-2 cell embryos were washed using RO water from a squirt bottle, and placed into a large glass dish with 1X Danieau Buffer (58 mM NaCl, 0.7 mM KCl, 0.4 mM $MgSO_4$, 0.6 mM $Ca(NO_3)_2$, 5 mM HEPES pH 7.6). Embryos for microinjection needed to be within the 1-2 cell stages.

During the 30 minutes that the zebrafish were mating, the microinjection solution was prepared (125 ng transgenic reporter construct, 175 ng pCS2FA Tol2 transposase mRNA, 2.0 µl 0.5% Phenol Red (Sigma P0290), and RNase-free water to a final solution volume of 5.0 µl). The needles utilized during the microinjection process were also constructed during this time. The microinjection needles were created using a David Kopf Instruments Vertical Pipette Puller (Model 700C) with its heat set at 54 and its solenoid set at 10, and RNase free 3.5 nanoliter (nl) capillary tubes (World Precision Instruments 4878). The 3.5 nl capillary tubes were baked prior to pulling at 260ºC to inactivate RNases and were pulled so that they had an average taper and an outer diameter of approximately 8 micrometers (µm). The injection solution was kept on ice prior to injection.

Microinjections were executed using the Nanoliter 2000 Microinjector (World Precision Instruments Model B203XVY) housed on a Marhauser MMJR Micromanipulator (World Precision Instruments). Using watchmaker forceps (size 5), the tips of the injection needles were broken so that the needle tip was beveled. One needle was then filled with

mineral oil and mounted onto the Nanoliter 2000 Microinjector. The mineral oil was expelled so that approximately one-fifth of the needle still contained oil. This process was executed via the metal pushrod which extended from the microinjector apparatus into the needle mounted onto the machine itself. The needle was then filled with microinjection injection solution by drawing the metal pushrod back into the Nanoliter 2000 Microinjector, thus producing a needle ready for use inside of which, the microinjection solution was separated from the metal pushrod by a buffering, inert layer of incompressible mineral oil.

Next, embryos retrieved from crosses described previously were prepared for microinjection. An injection stage was made previously by taping a 1.0 millimeter thick VWR micro slide (VWR International 48300-025) onto the bottom of a small, plastic Petri dish. Approximately 50 zebrafish embryos at the 1-2 cells stages were placed onto the Petri dish so that they rested against the lip of the micro slide. This positioning not only kept the embryos lined up for more efficient injection, but also provided a backing support for the embryos during the actual injection process. The injection stage was put under a dissecting microscope and the injection needle was positioned at a 25º incline so that it was possible to observe the stage, embryos, and injection needle through the eyepiece. The stage was moved so that the tip of the needle pierced through a single embryo's chorion, passed through the yolk sac, and came to rest just under the blastomeres. The Nanoliter 2000 Microinjector was then used to inject 4 nl of the desired microinjection solution into the embryo's yolk sac. The dark red coloration of the 0.5% Phenol Red made it relatively easy to determine if the solution actually entered the yolk sac or the space surrounding the embryo. The needle was then removed from the embryo by backing the injection stage away from the needle, completing the process of microinjecting a single zebrafish embryo. This process was

repeated for all of the embryos positioned on the injection stage, after which another selection of 50 embryos were collected and placed onto the stage for a second round of injections. Generally, two rounds of injections, approximately 100 embryos, were all that could be injected in a single sitting due to the rapid pace of zebrafish embryonic development.

After microinjection, the embryos were placed in a large glass bowl filled with fresh 1X Danieau Buffer, and placed into a VWR Mini Incubator (VWR International 97025-630) set at 27ºC. Careful attention was paid to the embryos to insure they continued to develop properly and that the number of casualties was kept at a minimum. Two to three times daily, the embryos were removed from the incubator, the dead embryos and those possessing developmental malformations were removed, and fresh 1X Danieau Buffer was added. This process was carried out daily until the embryos had reached 48 or 54 hours post fertilization depending on the injected construct and the time of desired imaging. Some embryos injected with the pDr-m2de1F-*cfos*-EGFP construct were raised to adulthood to establish stable transgenic lines.

This particular study was centered on the process of microinjections and throughout the course of the project numerous problems were encountered and troubleshooting was a must. The first predicament faced was that, during the microinjection process, approximately 40% of the embryos died upon or right after actual injection. It has been reported that when utilizing the Tol2 system, along with enhancer containing reporter constructs in zebrafish, only a 10-20% success rate can be expected when looking for reporter gene expression (Fisher et al., 2006). Thus, the high death rates upon microinjection were only making it that much more difficult to retrieve successful, transgenic embryos. Repetition of the

microinjection process allowed for practice of the actual injection technique making it increasingly easy to inject the 1-2 cell embryos in the center of the yolk, just below the blastomeres. It is in this region that the literature instructs the solution to be deposited for the highest chances of successful transgenic element insertion (Fisher et al., 2006). It also became evident that modifying the injection needles was necessary. When the microinjections were first started, the needles used had a rather sharp taper and a somewhat large width, thus making it difficult to pierce the chorions of the embryos without bursting the embryo itself. It was decided to change the settings on the David Kopf Instruments Vertical Pipette Puller (Model 700C). The heat was decreased from 85 to 54 and the solenoid was increased from 0 to 10 in the hope of making the pulling process more gradual, thereby stretching the glass pipette to give a longer taper and smaller needle width. After doing so, the embryo survival rate jumped from 60% to approximately 90%.

Another major problem arose when, upon completion and analysis of the first several rounds of injections involving the m2de1 containing reporter constructs, no EGFP expression was seen in the transgenic embryos. It was decided that the transposase mRNA needed to be handled much more carefully as mRNA is a very delicate molecule susceptible to degradation from various RNases (Alberts, 2008). Without proper handling and care, the transposase mRNA was probably degrading, which in turn disallowed the insertion of the m2de1 element into the zebrafish genome. Upon careful consideration, it was determined that it would be beneficial to bake the 3.5 nl capillary tubes (World Precision Instruments 4878) in an oven at 260ºC overnight to destroy any residual RNases present in the tubes prior to pulling. This single step kept the mRNA intact, and soon after EGFP expression was noted.

Immobilization of Transgenic Embryos

Once the remaining injected embryos reached the desired developmental stage, they were then immobilized and positioned for viewing using laser scanning confocal microscopy. A way in which to position the embryos in the orientation needed to best reveal the EGFP expression, but at the same time keep them alive so as to not degrade the EGFP, was soon devised using a Tricaine solution to anesthetize the embryos and an agarose solution to further immobilize and position the embryos. Injected embryos were placed under a dissecting microscope, and were dechorionated using watchmaker forceps. At 48 and 54 hours post fertilization, many of the embryos had shed their chorions as part of their natural developmental process. Dechorionated embryos were placed into a 1.5 ml microcentrifuge tube along with 0.5 mL of the 1X Danieau Buffer. A solution of 1X Danieau Buffer + 0.4% Tricaine (Sigma-Aldrich Life Science A5040) was added to anesthetize the embryos. The injected zebrafish were exposed to the 1X Danieau Buffer + 0.4% Tricaine solution for approximately 5 minutes or until no visible voluntary movements were observed from the zebrafish embryo. At first, a 1X Danieau Buffer + 0.016% Tricaine (Sigma-Aldrich Life Science A5040) solution was implemented, but this concentration was too low to successfully anesthetize the embryos. After gradually increasing the concentration over multiple trials, it was concluded that a 1X Danieau Buffer + 0.4% Tricaine (Sigma-Aldrich Life Science A5040) solution worked best to anesthetize the zebrafish.

A solution of 1X Danieau Buffer + 0.8% GenePure LE Agarose (ISC BioExpress C404595) was then melted. Using a disposable glass pipette, 0.25 mL of the melted agarose solution was placed into a 1.5 ml microcentrifuge tube. Before the agarose solidified, one of the anesthetized, injected embryos was placed into the 1.5 ml microcentrifuge tube

containing the agar. The embryo, along with a sizable amount of the melted agarose solution, was removed from the microcentrifuge tube using a disposable glass pipette and deposited into the well of a deep-welled, VWR Culture Slide (VWR International 48327-000). The embryo was positioned in the desired orientation within the slide prior to the agarose solidifying. At first, a 1X Danieau Buffer + 0.4% GenePure LE Agarose (ISC BioExpress C404595) solution was used to set the anesthetized embryos into the deep welled slides for confocal imagery. However, such a low concentration of agarose did not fully set and made it difficult to keep the embryos in the orientation wanted for imaging. The agarose concentration was gradually increased as well and it was found that a 1X Danieau Buffer + 0.8% GenePure LE Agarose (ISC BioExpress C404595) solution worked best to set the embryos.

Generally, two orientations were utilized for viewing the embryos: a dorsal (top-down) orientation in which the upper portion of the zebrafish head was visible, and a sagittal orientation in which the lateral areas of the zebrafish head could be analyzed. This whole process was repeated for approximately six to eight of the injected, anesthetized embryos. Once the agarose solution had solidified around the embryos, a disposable glass pipette was used to add a few drops of 1X Danieau Buffer to the wells, and a Corning Cover Glass #1 (Corning Life Sciences 2845-18) slip was added over the tops of the wells.

Imaging Transgenic Embryos

The embryos were viewed for reporter gene expression using a Zeiss LSM 510 Confocal Microscope. Slides were positioned one at a time under the microscope and examined using the 10X objective, the Argon laser, and the FITC filter to image the localization and positioning of EGFP expression within the developing zebrafish embryos. If

an embryo showed EGFP expression, a Z-Stack was taken of that particular embryo in the hopes of imaging all layers and depth of the reporter gene expression. The Z-Stack was generally set at an average of 15 slices of approximately 8 µm thickness, a scan speed of 2, and a pixel quality of 1024 X 1024. The Z-Stacks were then modified using Zeiss software's built in Projection Tool in order to overlay each image slice on top of each other to give a composite image for each embryo.

RESULTS

Zebrafish and Mouse HCNE Isolation from Genomic DNA

The identified m2de1 element found in zebrafish was isolated from total genomic DNA through Polymerase Chain Reaction (PCR) amplification. The two sequence and species specific primers (Table 1) used for the PCR protocol were designed so that they flanked the chosen portion of the m2de1 sequence within the zebrafish genome (Fig. 2). The PCR reaction was predicted to produce a PCR product of around 450 nucleotides in length, which is many nucleotides longer than the previously identified 260 base pairs of the m2de1 element that were highly conserved (Fig. 6). This was by design as the primers were constructed to amplify extended regions adjacent to the core m2de1 element to ensure isolation of the areas potentially possessing *cis*-regulatory function. Upon successful isolation from the zebrafish genome, the m2de1 element was used for cloning the sequence into the pCR®2.1-TOPO® TA cloning vector. The isolation and amplification of the identified m2de1 element found in the mouse genome was done as previously described (Nelson, 2011).

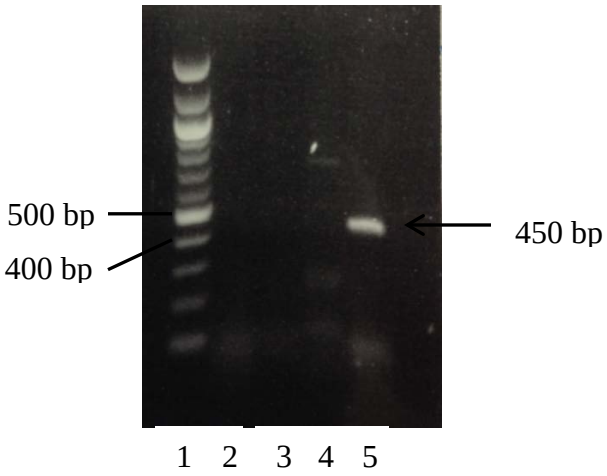


Fig. 6. Isolation of the zebrafish m2de1 element from genomic DNA. Gel electrophoresis was performed to insure correct and sufficient amplification of the m2de1 element. Lane 1 contains 100 base pair ladder (NEB N3231L). Lanes 2, 3, and 4 contain control samples. Lane 5 contains the m2de1 element PCR product. The arrow points to the approximately 440 nucleotide long element, indicating successful isolation.

HCNE TOPO® TA Cloning®

The isolated m2de1 elements from zebrafish and mouse were then cloned into the pCR®2.1-TOPO® cloning vector creating the Dr-m2de1-pCR2.1-TOPO vector and the Mm-m2de1-pCR2.1-TOPO vector, respectively, as previously described (Nelson, 2011). This particular process was carried out for both the zebrafish and mouse HCNEs following standard Invitrogen TOPO TA Cloning® protocol (Invitrogen 45-0641). The Dr-m2de1-pCR2.1-TOPO (Fig. 7A) and the Mm-m2de1-pCR2.1-TOPO (Fig. 7B) clones were grown up and used for the subsequent construction of transgenic constructs.

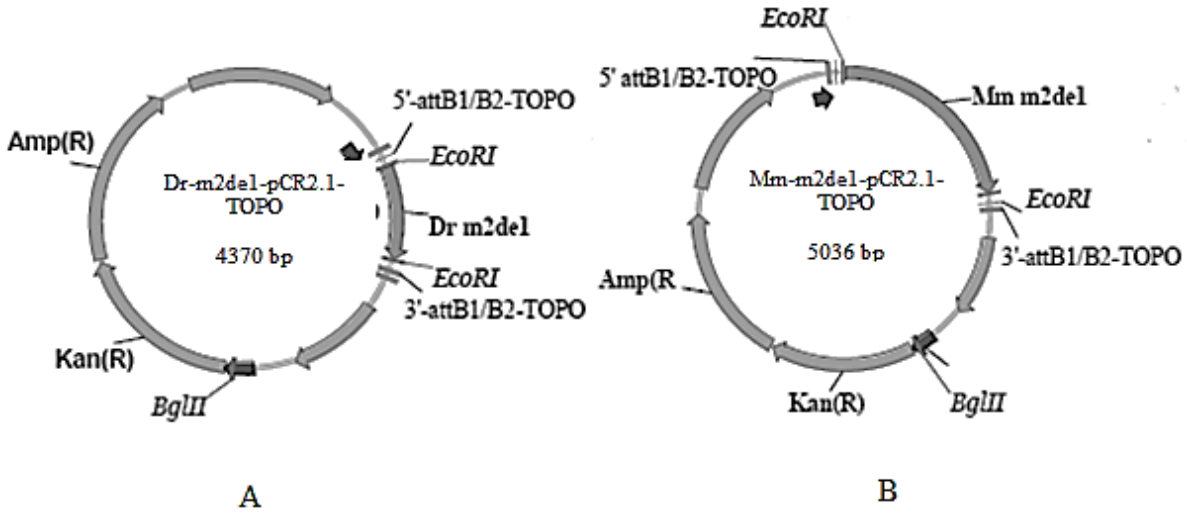


Fig. 7. Cloning the zebrafish and mouse m2de1 elements into pCR®2.1-TOPO®. Above are schematic diagrams depicting the Dr-m2de1-pCR2.1-TOPO vector (A) and the Mm-m2de1-pCR2.1-TOPO vector (B) after successful Invitrogen TOPO TA Cloning® protocol (Invitrogen 45-0641). The EcoRI sites which flank the m2de1 elements, the BglII sites, and the attB1 and attB2 sites are shown for both vectors. The generation of both vectors was done previously (Nelson, 2011).

Isolating HCNEs from pCR®2.1-TOPO®

The development of the reporter constructs containing the transgenic expression cassettes was done using Gateway® Technology's Cloning System. As stated previously, the construction of the vectors was the result of a two-part subcloning process where the m2de1 element was subcloned into the pDONR221 vector and then the pGW_*cfos*EGFP vector. In total, four sets of reporter constructs were created: construct with the zebrafish m2de1 element in forward orientation (pDr-m2de1F-*cfos*-EGFP), construct with the zebrafish m2de1 element in reverse orientation (pDr-m2de1R-*cfos*-EGFP), construct with mouse m2de1 element in forward orientation (pMm-m2de1F-*cfos*-EGFP), and construct with mouse m2de1 element in reverse orientation (pMm-m2de1R-*cfos*-EGFP). I was responsible for preparing the zebrafish m2de1 element for reverse orientation insertion. The preparation of the zebrafish m2de1 element for forward orientation insertion, the mouse m2de1 element

for forward orientation insertion, and the mouse m2de1 element for reverse orientation insertion were done previously by a member of the Zerucha lab (Nelson, 2011).

In order to produce the transgenic reporter construct containing the zebrafish m2de1 element in its reverse orientation, the zebrafish element was first isolated out of the Dr-m2de1-pCR2.1-TOPO (Fig. 7A) plasmid and prepared for reverse orientation insertion by PCR using the primer set 5'-attB2-TOPO and 3'-attB1-TOPO (Table 1). The PCR product was predicted to be 528 nucleotides in length due to the addition of the attB1 and attB2 sites, which is what was observed (Fig. 8). The PCR product was subsequently cleaned for utilization in the BP reaction.

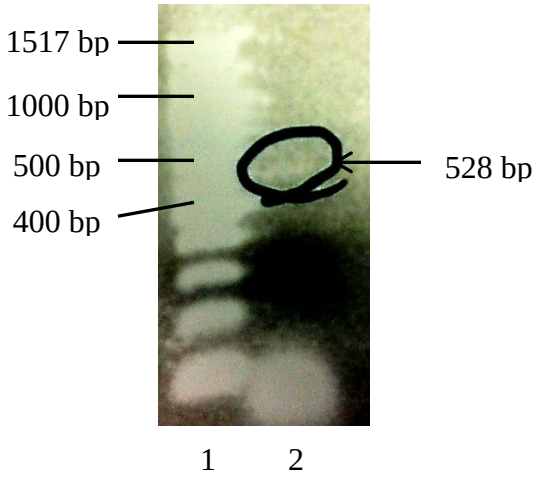


Fig. 8. Isolating the zebrafish m2de1 element out of Dr-m2de1-pCR2.1-TOPO. Gel electrophoresis was performed to determine if the zebrafish m2de1 element was successfully isolated from the Dr-m2de1-pCR2.1-TOPO vector for subsequent insertion into the pDONR221 donor plasmid (Fig. 7A). Lane 1 contains 100 base pair ladder (NEB N3231L). Lanes 2 contains the m2de1 element PCR product post isolation with the flanking attB1 and attB2 sites. The arrow points to the 528 nucleotide long element, indicating successful isolation.

Construction of Middle Entry Vectors

Next, the zebrafish m2de1 element was inserted in its reverse orientation into the previously mentioned donor plasmid, pDONR221 in a process known as the BP reaction

(Fig. 4). As mentioned previously, the BP reaction utilizes Gateway® BP Clonase™ II Enzyme Mix (Invitrogen 865071) and enzyme assisted recombination to construct middle entry vectors. The resulting middle entry vector containing the zebrafish HCNE in its reverse orientation was named pME-Dr-m2de1R. I was responsible for preparing the middle entry vector containing the zebrafish m2de1 element in its reverse orientation (pME-Dr-m2de1R). The preparation of the middle entry vector containing the zebrafish m2de1 element in its forward orientation (pME-Dr-m2de1F), the middle entry vector containing the mouse m2de1 element in its forward orientation (pME-Mm-m2de1F), and the middle entry vector containing the mouse m2de1 element in its reverse orientation (pME-Mm-m2de1R) was done by a previous member of the Zerucha lab (Nelson, 2011).

After the BP reaction had been completed, the reaction was then transformed into chemically competent DH5α *E. coli* and grown on LB + Kanamycin (50 mg/µl) plates. The *ccdB* gene found within the pDONR221 donor plasmid was removed upon successful insertion of the reverse oriented zebrafish m2de1 element into the donor plasmid due to BP reaction recombination (Fig. 2). The *ccdB* gene serves as an intracellular control element by killing cells that lack resistance to its resultant protein. The DH5α *E. coli* cells that the BP reaction was transformed into lack ccdB resistance meaning that any bacterial colonies that grew would have successfully undergone recombination and ccdB gene removal. Colonies were grown up in cultures containing 3 ml LB + Kanamycin (50 mg/µl), prepared via miniprep, and examined using BglII and EcoRI restriction digestion analysis. Both the Dr-m2de1-pCR2.1-TOPO (Fig. 7A) and pME-Dr-m2de1R (Fig. 9B) plasmids have Kanamycin resistance, necessitating the screening of the pME-Dr-m2de1R with the two different restriction enzymes to rule out the possibility of residual Dr-m2de1-pCR2.1-TOPO

background transformation. The Dr-m2de1-pCR2.1-TOPO construct contains 1 BglII site and 2 EcoRI sites (Fig. 7A). The pME-Dr-m2de1R construct contains two EcoRI sites (Fig. 9B). If the Dr-m2de1-pCR2.1-TOPO construct is present, restriction digestion will result in three bands of 2931 base pairs, 982 base pairs, and 457 base pairs. If the correct middle entry vector, pME-Dr-m2de1R, is present, two bands of 2656 base pairs and 457 base pairs will be seen as the construct does not contain the BglII site. Gel electrophoresis analysis of the restriction digest showed a band in lane 1 of 457 base pairs but no band of 982 base pairs (Fig. 10), indicating that the desired plasmid, pME-Dr-m2de1R, had been constructed.

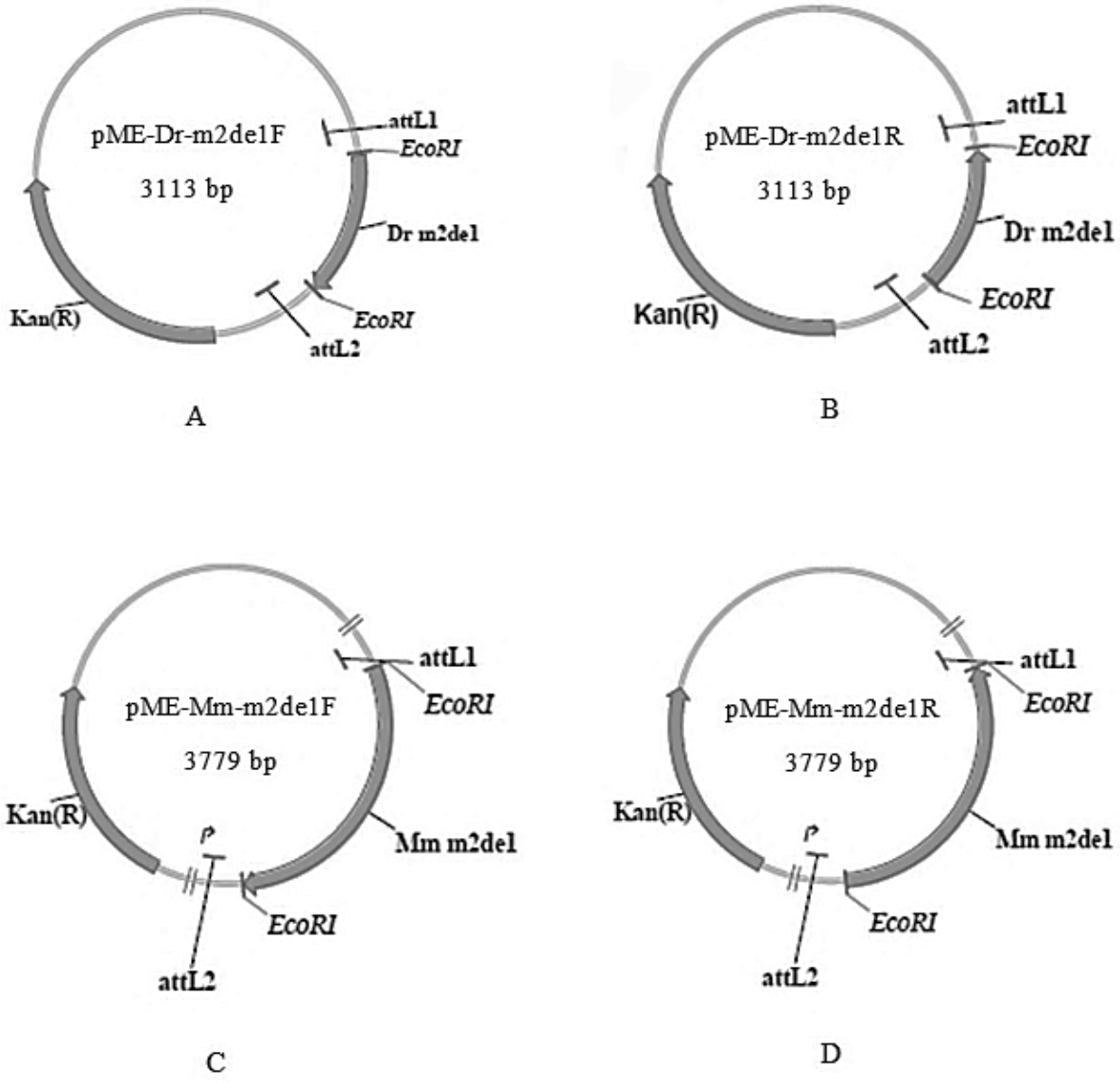


Fig. 9. Constructing middle entry vectors for the mouse and zebrafish m2de1 elements. Above are schematic diagrams depicting the pME-Dr-m2de1F construct (A), the pME-Dr-m2de1R construct (B), the pME-Mm-m2de1F construct (C), and the pME-Mm-m2de1R construct (D) after successful BP reaction. The EcoRI sites which flank the m2de1 elements and the attL1 and attL2 sites are shown for the constructs.

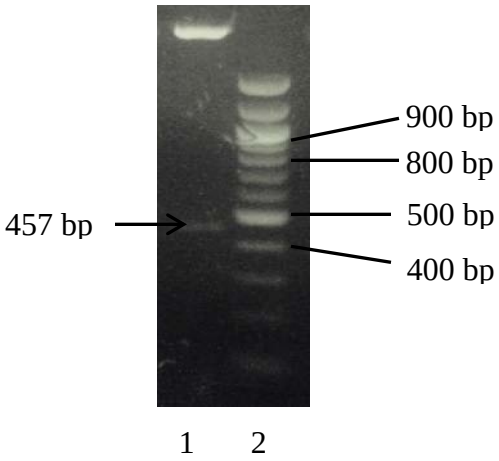


Fig. 10. Gel electrophoresis analysis of the pME-Dr-m2de1R restriction digest. Gel electrophoresis was performed to determine if the BP reaction was successful in creating the pME-Dr-m2de1R construct (Fig. 9B). Lane 1 contains pME-Dr-m2de1R plasmid cut with EcoRI (NEB R0101S) and BglII (NEB R0144S), testing for successful translocation. Lane 2 contains 100 base pair ladder (NEB N3231L). The arrow points to the 457 base pair long m2de1 fragment. The lack of a 982 base pair fragment indicates the lack of a BglII site meaning only the correct middle entry vector, pME-Dr-m2de1R, was formed.

Construction of Transgenic Reporter Constructs

Next, the middle entry plasmid was used to relocate the reverse oriented zebrafish m2de1 element into the previously mentioned destination vector, pGW_*cfos*EGFP, in a process known as the LR reaction (Fig. 5). The LR reaction utilizes Gateway® LR Clonase™ II Plus Enzyme Mix (Invitrogen 12538120) and enzyme assisted recombination to construct transgenic reporter constructs. The resulting reporter construct containing the zebrafish HCNE in its reverse orientation was named pDr-m2de1R-*cfos*-EGFP. I was responsible for preparing the reporter construct containing the zebrafish m2de1 element in its reverse orientation (pDr-m2de1R-*cfos*-EGFP). The preparation of the reporter constructs containing the zebrafish m2de1 element in its forward orientation (pDr-m2de1F-*cfos*-EGFP), the reporter constructs containing the mouse m2de1 element in its forward orientation (pMm-m2de1F-*cfos*-EGFP), and the reporter constructs containing the mouse m2de1 element in its

reverse orientation (pMm-m2de1R-*cfos*-EGFP) was done by a previous member of the Zerucha lab (Nelson, 2011). All the reporter constructs contain the expression cassette (m2de1 element upstream from the minimal *cfos* promoter and the reporter gene, EGFP) flanked by Tol2 recognition sites.

After the LR reaction had been completed, the reaction was then transformed into One Shot® TOP10 Chemically Competent *E. coli* (Invitrogen C4040-10) and grown on LB + Ampicillin (100 mg/µl) plates. As shown earlier, the pME-Dr-m2de1R vector has Kanamycin resistance (Fig. 9B) while the pDr-m2de1R-*cfos*-EGFP construct has Ampicillin resistance (Fig. 11B). Any bacterial colony that grew on the LB + Ampicillin (100 mg/µl) plates should be the reporter construct and not the middle entry vector. Furthermore, the *ccdB* gene found within the pDONR221 donor plasmid was removed upon successful insertion of the reverse oriented zebrafish m2de1 element into the donor plasmid due to BP reaction recombination (Fig. 3). The One Shot® TOP10 Chemically Competent *E. coli* (Invitrogen C4040-10) cells that the LR reaction was transformed into lack *ccdB* resistance meaning that any bacterial colonies that grew would have successfully undergone recombination and ccdB gene removal. Colonies were grown up in cultures containing 3 ml LB + Ampicillin (100 mg/µl), prepared via miniprep, and examined using EcoRI restriction digestion analysis. The pME-Dr-m2de1R middle entry construct contains two EcoRI sites (Fig. 9B). The pDr-m2de1R-*cfos*-EGFP reporter construct contains two EcoRI sites as well (Fig. 11B). If the pME-Dr-m2de1R middle entry construct is present, restriction digestion will result in two bands of 2656 base pairs and 457 base pairs. If the correct reporter construct, pDr-m2de1R-*cfos*-EGFP, is present, two bands of 7484 base pairs and 457 base pairs will be seen. Therefore, although both constructs contain the same restriction sites, the

large variation in size between the top bands that are formed with restriction should make it possible to determine if the LR reaction was successful. Gel electrophoresis analysis of the restriction digest showed two bands in lane 2 of 7484 base pairs and 457 base pairs (Fig. 12), indicating that the desired reporter construct, pDr-m2de1R-*cfos*-EGFP, had been constructed.

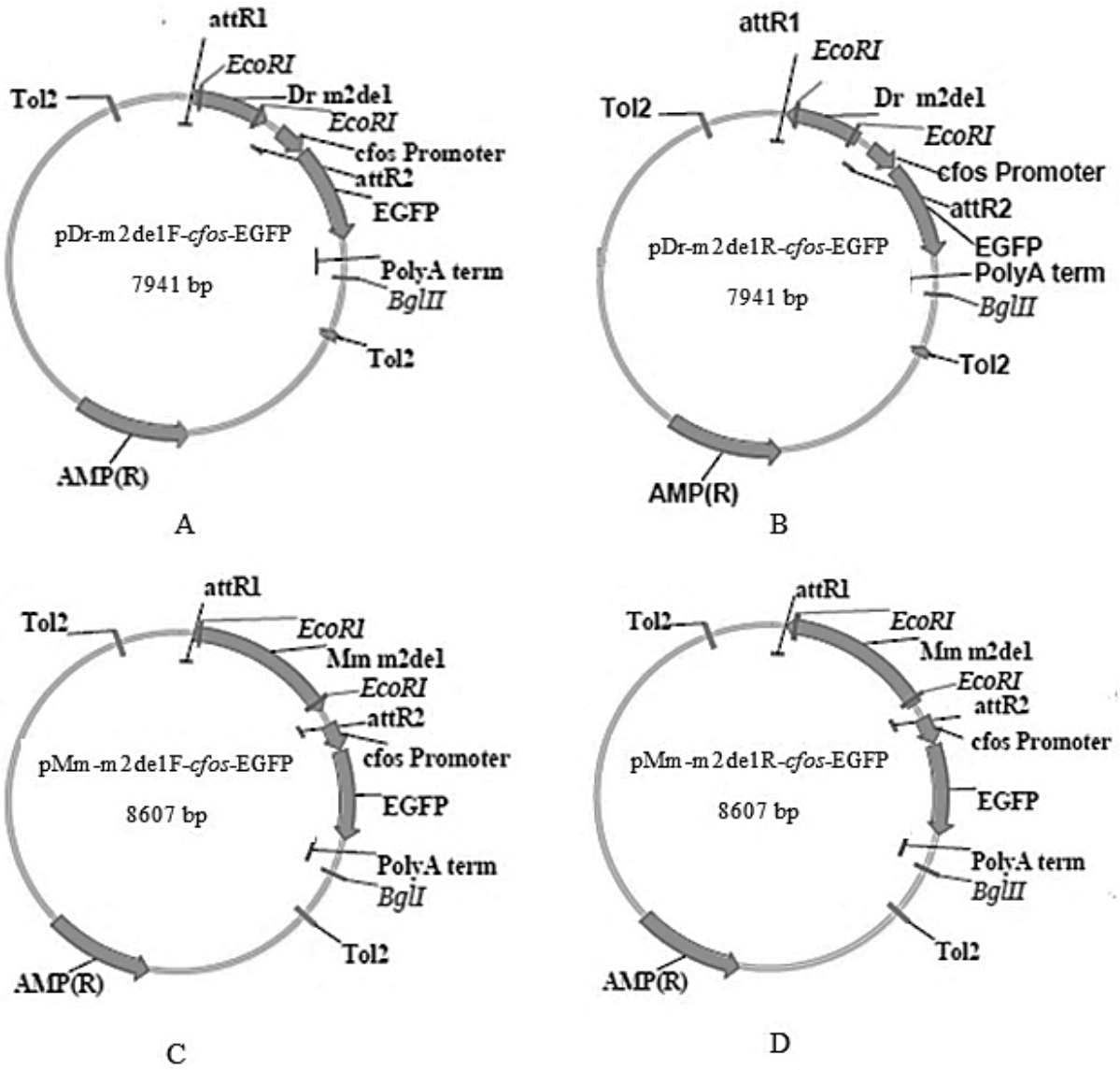


Fig. 11. Constructing the reporter constructs for the mouse and zebrafish m2de1 elements. Above are schematic diagrams depicting the pDr-m2de1F-*cfos*-EGFP construct (A), the pDr-m2de1F-*cfos*-EGFP construct (B), the pMm-m2de1F-*cfos*-EGFP construct (C), and the pMm-m2de1R-*cfos*-EGFP construct (D) after successful LR reaction. The EcoRI sites which flank the m2de1 elements, the attR1 and attR2 sites, the minimal *cfos* promoter, the Tol2 sites, and the EGFP gene are shown for the constructs.

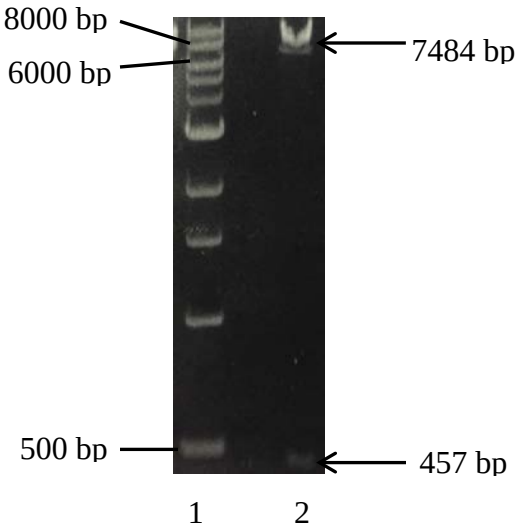


Fig. 12. Gel electrophoresis analysis of the pDr-m2de1R-*cfos*-EGFP restriction digest. Gel electrophoresis was performed to determine if the LR reaction was successful in creating the pDr-m2de1R-*cfos*-EGFP construct (Fig. 11B). Lane 1 contains 1 Kb base pair ladder (NEB N3232L). Lane 2 contains pDr-m2de1R-*cfos*-EGFP construct cut with EcoRI (NEB R0101S), testing for successful translocation. The arrow points to the 457 base pair long m2de1 fragment and the larger 7484 base pair long fragment. The correct size of both bands indicates that the transgenic reporter construct, pDr-m2de1R-*cfos*-EGFP, was formed.

Sequencing Results

As mentioned throughout the Materials and Methods section, several of the constructs were sequenced during the course of vector construction to ensure that random point mutations had not occurred to the m2de1 element itself. By definition, *cis*-regulatory elements are conserved and changes to their sequence could very well alter function. In all, three different constructs were sent off for m2de1 sequencing at various stages in transgenic reporter construct development. The zebrafish m2de1 element was sequenced after being amplified from the zebrafish genomic DNA [performed by Cornell University's Life Sciences Core Laboratories using the Dr-m2de1-3 and Dr-m2de1-5b primers (Table 1)]. The pCR®2.1-TOPO® cloning vector containing the zebrafish m2de1 element was sequenced [performed by Cornell University's Life Sciences Core Laboratories using universal M13

Forward and M13 Reverse primers (Table 1)]. Lastly, the pDr-m2de1F-*cfos*-EGFP construct was sequenced as well [performed by Cornell University's Life Sciences Core Laboratories using the Dr-m2de1-3 and Dr-m2de1-5b primers (Table 1)].

The sequencing analysis of the three separate constructs revealed that two individual nucleotides within the m2de1 element are different than in the published zebrafish genome (Fig. 13). In one case, a cytosine was identified as a thymine (Fig. 13). In the other case, a guanine was identified as an adenine (Fig. 13). All the sequencing analyses support this case as all three sequenced constructs not only contained nucleotide differences in the exact same locations within the m2de1 sequence, but all three constructs possessed the exact same nucleotides in those positions as well.

Fig. 13. Summary of sequencing analysis resulting in a corrected m2de1 sequence. Areas highlighted in red represent the 260 base pairs of sequence that most likely possess *cis*-regulatory function and show the highest amount of sequence conservation. The area highlighted in purple is the Hox binding site which was not seen in the element used by Parker and was subsequently mutated (Parker et al., 2011). The 2 nucleotides highlighted in green represent the 2 nucleotides that form the boundaries for the m2de1 element used in our vector construction as shown previously (Fig. 2). The nucleotide highlighted in yellow is the

cytosine that was miscalled as a thymine in the published sequence. The nucleotide highlighted in blue is the guanine that was miscalled as an adenine in the published sequence.

Mutation of a Potential Hox Binding Site

In order to detail the effects of a potential Hox binding site, TAAT (Fig. 5) within the zebrafish m2de1 element, PCR mutagenesis was performed using the QuikChange® II XL Site-Directed Mutagenesis Kit (Stratagene #200521) to create the pDr-m2de1F-mut-*cfos*-EGFP reporter construct. Previous research demonstrated that a sequence similar to m2de1 drives *Meis2* expression (Parker et al., 2011). Sequence comparison between the zebrafish m2de1 element and the *cis*-regulatory element utilized by Parker et al. revealed that the m2de1 element included an additional Hox binding site not used by Parker; the previously mentioned TAAT sequence (Fig. 14). By eradicating this Hox binding site, it would then be possible to compare expression profiles driven by the two different elements. The objective was to mutate the Hox binding site (TAAT) into a radically different sequence (TCGA) believed to possess no association with Hox binding (Fig. 14). The TCGA mutation sequence should eliminate any Hox binding affinity the TAAT sequence may have had. The mutation primers were designed so that an additional guanine adjacent to the Hox binding site was mutated as well (TAATG to TCGAT). Due to the convenient location of an adenine directly 5' to the Hox binding site, the mutation introduced a ClaI restriction site (5' AT▼CGAT 3') within the zebrafish m2de1 element not seen in the original form of the element. The introduction of such a restriction site makes it relatively easy to determine if in fact the PCR mutagenesis was successful.

After the QuikChange® II XL Site-Directed Mutagenesis Kit mutagenesis procedure was carried out using the Dr-m2de1-mut-R primer (Table 2), the Dr-m2de1-mut-F primer

(Table 2), and the pDr-m2de1F-*cfos*-EGFP reporter construct retrieved from glycerol stock, the reaction was transformed into XL10-Gold ultracompetent (Stratagene) cells and grown on LB + Ampicillin (100 mg/µl) plates. As shown earlier, the pDr-m2de1F-*cfos*-EGFP construct has Ampicillin resistance (Fig. 11A). The mutagenesis protocol should not have altered the Ampicillin resistance gene meaning both the mutated (pDr-m2de1F-mut-*cfos*-EGFP) and the unmutated (pDr-m2de1F-*cfos*-EGFP) constructs should successfully grow on the plates. Additional measures were needed to confirm successful mutagenesis. Therefore, colonies were grown up in cultures containing 3 ml LB + Ampicillin (100 mg/µl), prepared via miniprep, and examined using ClaI restriction digestion analysis. The pDr-m2de1F-mut-*cfos*-EGFP construct contains two Cla1 sites, one site present prior to mutation and another site introduced through successful mutation (Fig. 15). The pDr-m2de1F-*cfos*-EGFP construct contains one ClaI site (Fig. 11A). If the pDr-m2de1F-mut-*cfos*-EGFP construct was created, restriction digestion will result in two bands of 6700 base pairs and 1241 base pairs. Restriction of the pDr-m2de1F-*cfos*-EGFP construct with ClaI should result in only one band representing the linearized construct. Gel electrophoresis analysis of the restriction digest showed two bands in lane 6 of 6700 base pairs and a larger than expected band of 1600 base pairs (Fig. 16). However, sequencing analysis of the plasmid confirmed that the TAATG binding site had been changed to TCGAT and that the pDr-m2de1F-mut-*cfos*-EGFP construct had been created successfully.

Fig. 14. Mutated m2de1 element within pDr-m2de1F-mut-*cfos*-EGFP reporter construct. As indicated previously, 450 bp of sequence were isolated to analyze for *cis*-regulatory function. Areas highlighted in red represent the 260 base pairs of sequence that show the highest amount of sequence conservation. The area highlighted in purple is the mutated sequence which once served as a potential Hox transcription factor binding site and is the binding site not seen in the element used by Parker and colleagues (Parker et al., 2011). The 2 nucleotides highlighted in green represent the 2 nucleotides that form the boundaries for the m2de1 element used in our vector construction as shown previously (Fig. 2). The 2 nucleotides highlighted in blue represent the 2 nucleotides that form the boundaries for the *cis*-regulatory element used by Parker and colleagues (Parker et al., 2011). The positioning of the additional Hox binding site which was subsequently mutated to the sequence highlighted in purple was such that it was not included in the element utilized by Parker and colleagues (Parker et al., 2011).

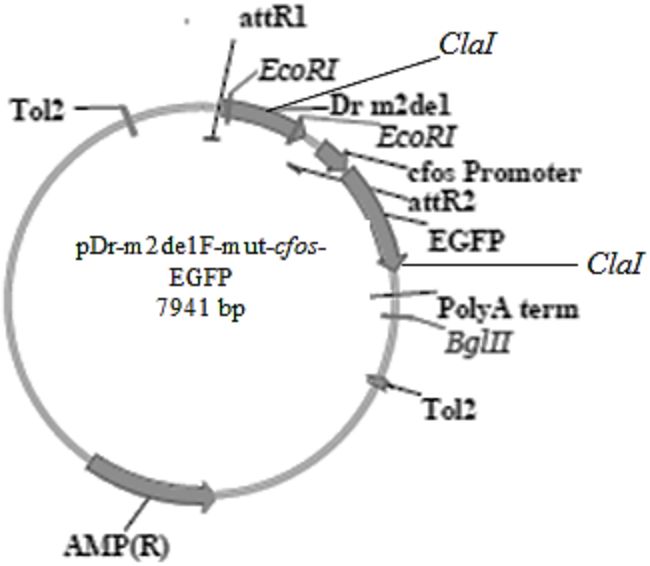


Fig. 15. Constructing the pDr-m2de1F-mut-*cfos*-EGFP reporter construct. Above is a schematic diagram depicting the pDr-m2de1F-mut-*cfos*-EGFP construct after successful PCR mutagenesis using the QuikChange® II XL Site-Directed Mutagenesis Kit (Stratagene #200521). The EcoRI sites which flank the m2de1 elements, the ClaI sites within the mutated zebrafish m2de1 element, the attR1 and attR2 sites, the minimal *cfos* promoter, and the EGFP gene are shown for the construct.

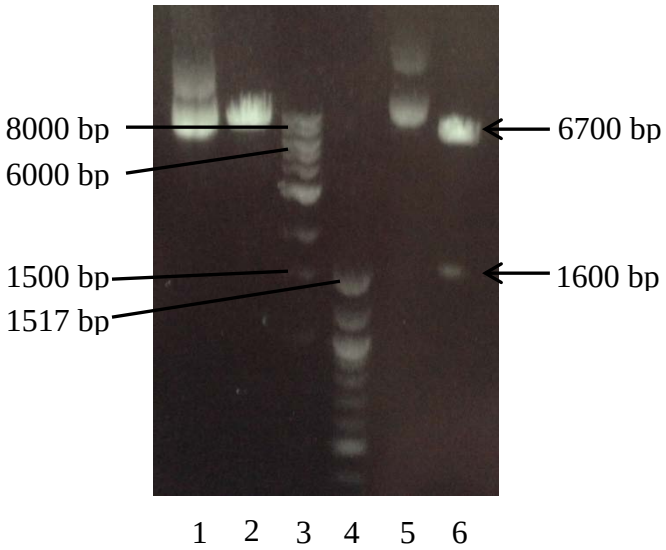


Fig. 16. Gel electrophoresis analysis of pDr-m2de1F-mut-*cfos*-EGFP restriction digest. Gel electrophoresis was performed to determine if the mutagenesis procedure was successful in creating the pDr-m2de1F-mut-*cfos*-EGFP construct (Fig. 15). Lane 1 contains uncut pDr-m2de1F-*cfos*-EGFP construct and Lane 2 contains ClaI (NEB R0197S) restricted pDr-m2de1F-*cfos*-EGFP construct. Lane 3 contains 1 Kb base pair ladder (NEB N3232L) and Lane 4 contains 100 base pair ladder (NEB N3231L). Lane 5 contains uncut pDr-m2de1R-

mut-*cfos*-EGFP construct cut and Lane 6 contains ClaI restricted (NEB R0197S) pDr-m2de1R-mut-*cfos*-EGFP construct. The arrows point to the 6700 base pair long fragment and the 1600 base pair long fragment produced by the restriction analysis. Although these fragment sizes are not correct, sequencing analysis proved that the mutagenesis protocol was successful.

Making Transposase mRNA

The pCS2FA Tol2 transposase mRNA utilized in the microinjections was created from glycerol stock DNA graciously provided by Dr. Chi-Bin Chien (Kwan et al., 2007). The plasmid was first harvested with the QIAGEN® Plasmid *Plus* Maxi Kit (QIAGEN 12963) and then linearized with NotI, resulting in a linear construct ready for transcription. The SP6 promoter would bind necessary transcriptional proteins, beginning transcription of the transposase template which would stop shortly after the end of the SV40 PolyA tail due to linearization (Fig. 17). The linearized plasmid was then subjected to gel electrophoresis to insure complete NotI digestion. The restriction digest resulted in 1 thick band of around 6,300 base pairs in length, the size of the plasmid (Fig. 18). The linearized pCS2FA Tol2 transposase template was then transcribed using the mMESSAGE mMACHINE® SP6 RNA Transcription Kit (Ambion® AM1340M). Once the transcription reaction had been completed, the mRNA product was purified and analyzed by gel electrophoresis to determine if the mRNA was the right size (Fig. 19). The Tol2 transposase mRNA is 2,251 bases long, and the gel analysis revealed a single band of approximately that size. However, closer analysis of the gel reveals a decent amount of smearing which could in fact be the result of mRNA degradation. This mRNA breakdown could have occurred prior to running the gel or during the actual electrophoresis process. Fortunately, the intensity of the band and exactness of its size indicates that transcription resulted in enough Tol2 mRNA for successful microinjection.

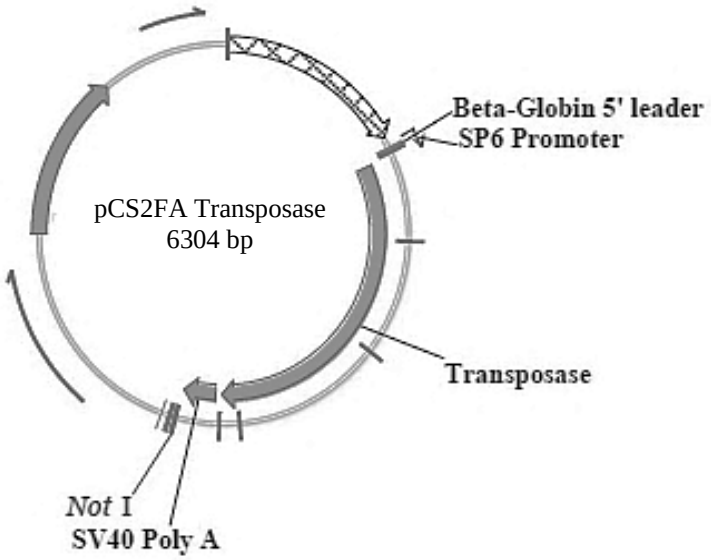


Fig. 17. The pCS2FA Tol2 transposase plasmid. Above is a schematic diagram depicting the pCS2FA Transposase plasmid. The NotI site, the SP6 promoter, the transposase sequence, and the SV40 Poly A tail are all shown.

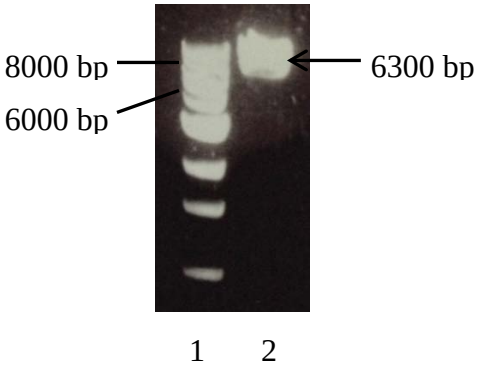


Fig. 18. Digestion of the pCS2FA Tol2 transposase plasmid for linearization. After restriction digestion of the pCS2FA Tol2 Transposase Plasmid with NotI (NEB R0189S), gel electrophoresis was performed to determine if the digestion was successful. Lane 1 contains 1 Kb base pair ladder (NEB N3232L). Lane 2 contains linearized pCS2FA Tol2 Transposase plasmid. The arrow points to the 6300 base pair long linearized plasmid.

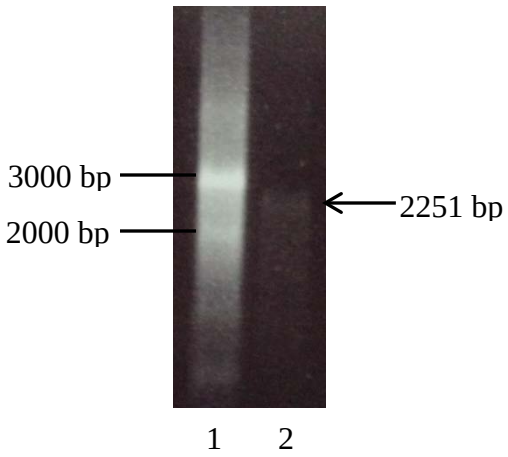
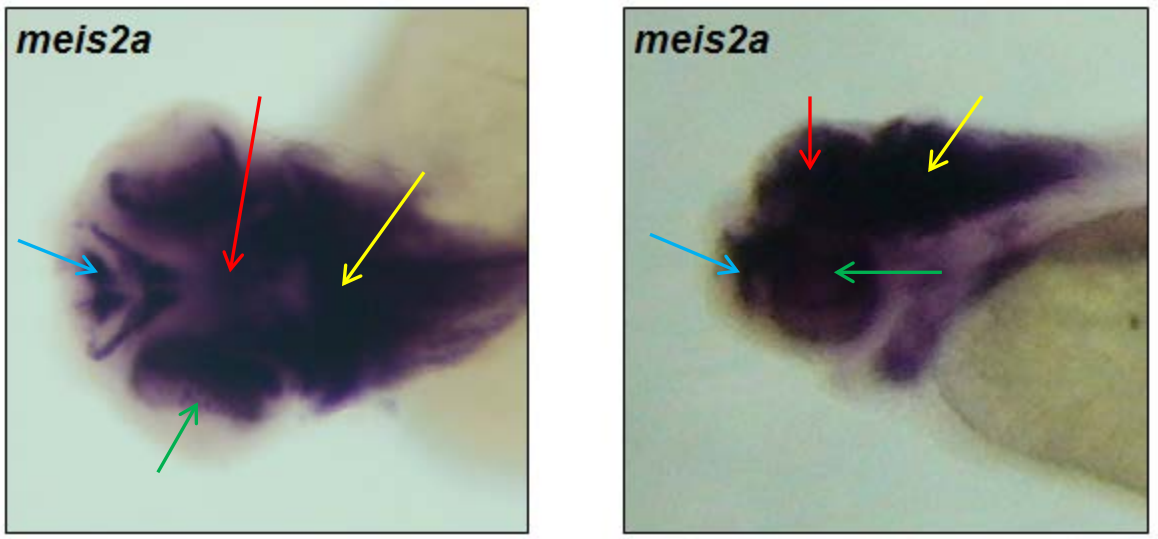


Fig. 19. Production of pCS2FA Tol2 transposase mRNA. After transcription of pCS2FA Tol2 Transposase Plasmid with the mMESSAGE mMACHINE® SP6 RNA Transcription Kit (Ambion® AM1340M), gel electrophoresis was performed to determine if transcription was successful. Lane 1 contains ssRNA ladder (NEB N0362). Lane 2 contains pCS2FA Tol2 Transposase mRNA. The arrow points to the 2251 base pair long mRNA product.

Transgenic Results Overview

Microinjections were performed to determine if the m2de1 putative *cis*-regulatory element was able to direct expression of the reporter gene in a manner consistent with the known expression pattern of *meis2a* (Fig. 20). Once microinjected into the zebrafish embryos, the pCS2FA Tol2 Transposase mRNA is translated into transposase protein, the expression cassette flanked by Tol2 recombination sites is cleaved from the reporter construct by Tol2 transposase and ultimately inserted into the embryos genome. Once in the zebrafish genome, if in fact the m2de1 element functions as an enhancer element, the transcription factors bound to this element should direct the minimal *cfos* promoter to initiate transcription of the reporter gene, EGFP. Without the binding of additional transcriptional proteins, the minimal promoter cannot begin transcription of the reporter gene thus correlating EGFP expression with element functionality. It is believed that due to the proximity between the m2de1 element and the *meis2a* gene *meis2a* expression is under the

direction of m2de1. Therefore, the expression pattern of EGFP should recapitulate at least a portion of the expression pattern of endogenous *meis2a* as expression of the reporter gene is under the influence of m2de1.



A B

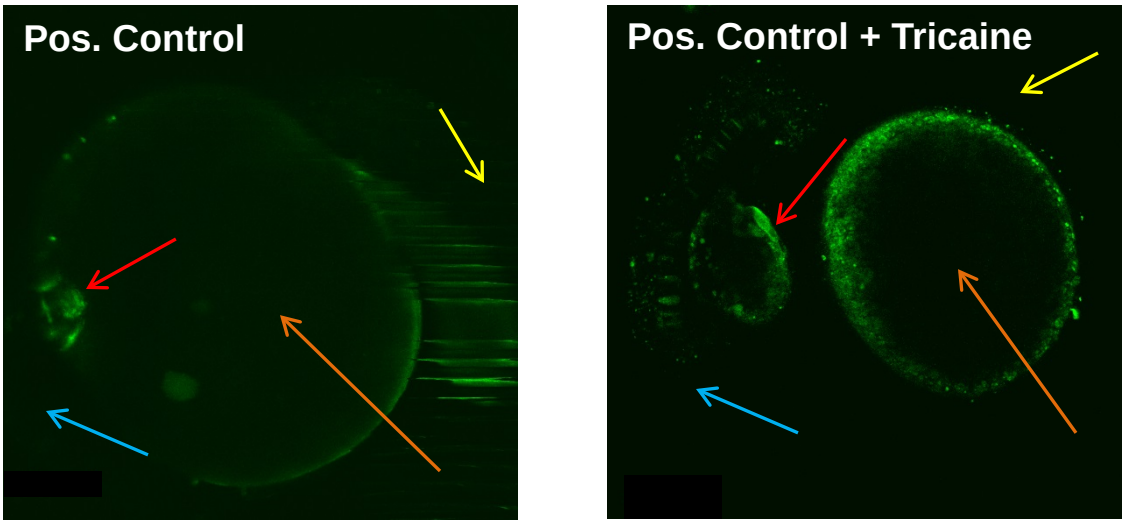
Fig. 20. Whole mount *in situ* hybridization for *meis2a* in zebrafish. *In situ* hybridization was performed on 48 hours post fertilization zebrafish embryos to localize and define the expression of *meis2a* in a spatial and temporal manner (Carpenter, 2010). Two orientations were utilized for viewing the embryos: a dorsal (top-down) orientation in which the upper portion of the zebrafish head was visible (A), and a sagittal orientation in which the lateral areas of the zebrafish head could be analyzed (B). The expression of *meis2a* was found throughout the forebrain (blue arrow), midbrain (red arrow), hindbrain (yellow arrow), and eye of the zebrafish (green arrow).

A Nanoliter 2000 Microinjector (World Precision Instruments Model B203XVY) housed on a Marhauser MMJR Micromanipulator (World Precision Instruments) was utilized to inject 1-2 cell embryos with one of the six various constructs previously developed: positive control construct (pDest-Sox10-mCherry-Tol2CG2; graciously provided by Dr. Chi-Bin Chien), construct with the zebrafish m2de1 element in forward orientation (pDr-m2de1F-*cfos*-EGFP), construct with the zebrafish m2de1 element in reverse orientation (pDr-

m2de1R-*cfos*-EGFP), construct with mouse m2de1 element in forward orientation (pMm-m2de1F-*cfos*-EGFP), construct with mouse m2de1 element in reverse orientation (pMm-m2de1F-*cfos*-EGFP), and the construct containing the zebrafish m2de1 element in forward orientation but with a putative Hox binding site sequence mutation (pDr-m2de1F-mut-*cfos*-EGFP). The embryos were injected at the 1-2 cell stage to ensure the presence of the expression cassette in all subsequently produced cells.

Expression Directed by the Positive Control Construct

The pDest-Sox10-mCherry-Tol2CG2 positive control construct has a known expression pattern in which EGFP is expressed in the heart of developing zebrafish embryos most notably at 24 hours post fertilization. The positive control construct also directs the expression of mCherry within the neural crest of developing embryos but this expression was not analyzed. Using a positive control construct allowed troubleshooting, practice of the microinjection technique, and the elimination of experimental variables to be carried out. A total of 463 embryos were injected with the positive control construct and analyzed for EGFP expression using the Zeiss LSM 510 Confocal Microscope. Fifteen of the 463 embryos exhibited reporter gene expression. As expected, EGFP expression was seen in the developing hearts of the zebrafish embryos (Fig. 21). This was a good indication that injection of the experimental constructs could ensue.



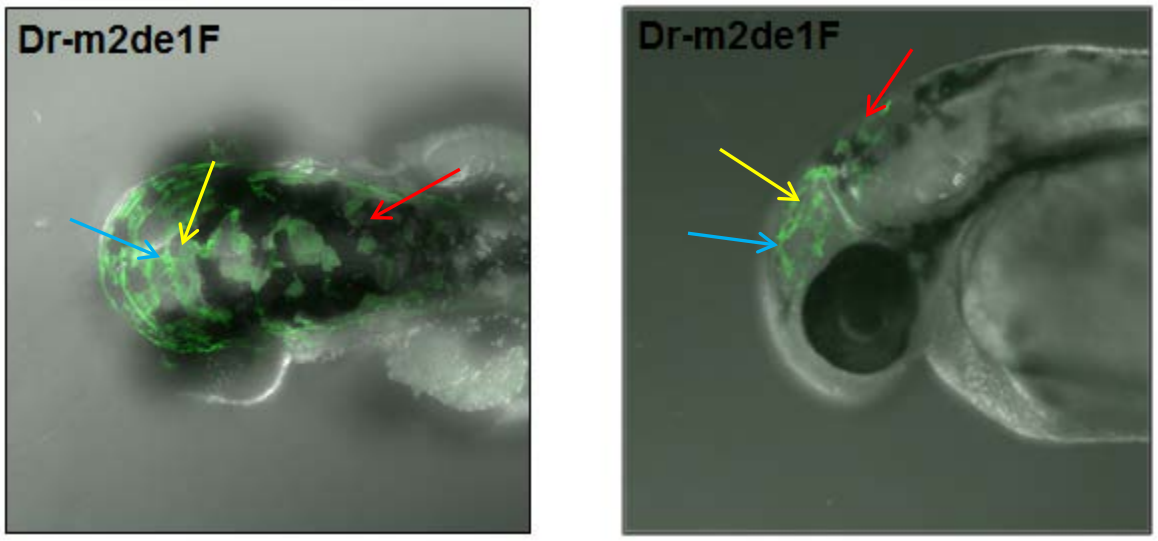
A B

Fig. 21. pDest-Sox10-mCherry-Tol2CG2 positive control reporter gene expression. EGFP expression is localized in the heart of 24 hours post fertilization zebrafish embryos previously injected with pDest-Sox10-mCherry-Tol2CG2 positive control construct. During the process of injecting the positive control construct, the immobilization technique was perfected. Embryos that were not anesthetized displayed a great deal of movement that impacted the quality of the confocal images (A). Embryos that were anesthetized displayed no movement which made positioning the embryos for confocal imagery much easier (B). The red arrows point to the developing zebrafish hearts showing abundant EGFP expression. The blue arrows point to the areas in which the head of the embryo would be found. The yellow arrows point to the areas in which the tail of the embryo would be found. The orange arrows point to the yolk sac.

Expression Directed by the Zebrafish m2de1 Element

The pDr-m2de1F-*cfos*-EGFP construct contains the zebrafish m2de1 element in its forward orientation. A total of 60 embryos were imaged for EGFP expression at 48 hours post fertilization using the Zeiss LSM 510 Confocal Microscope after injection with the pDr-m2de1F-*cfos*-EGFP construct. Fourteen of these injected embryos displayed reporter gene expression. Furthermore, 147 embryos were retrieved from crosses involving zebrafish previously injected with the pDr-m2de1F-*cfos*-EGFP construct and raised to adulthood to establish stable transgenic lines. In total, 100 of these 147 embryos exhibited reporter gene expression. As expected, EGFP expression was seen in anterior regions of the zebrafish

embryos (Fig. 22) mirroring endogenous *meis2a* expression (Fig. 20). More specifically, reporter gene expression appears to be located within the midbrain, neurons of the midbrain, portions of the hindbrain, as well as the tectum (Fig. 22). Little to no EGFP expression was seen in the developing eyes of the embryos (Fig. 22).

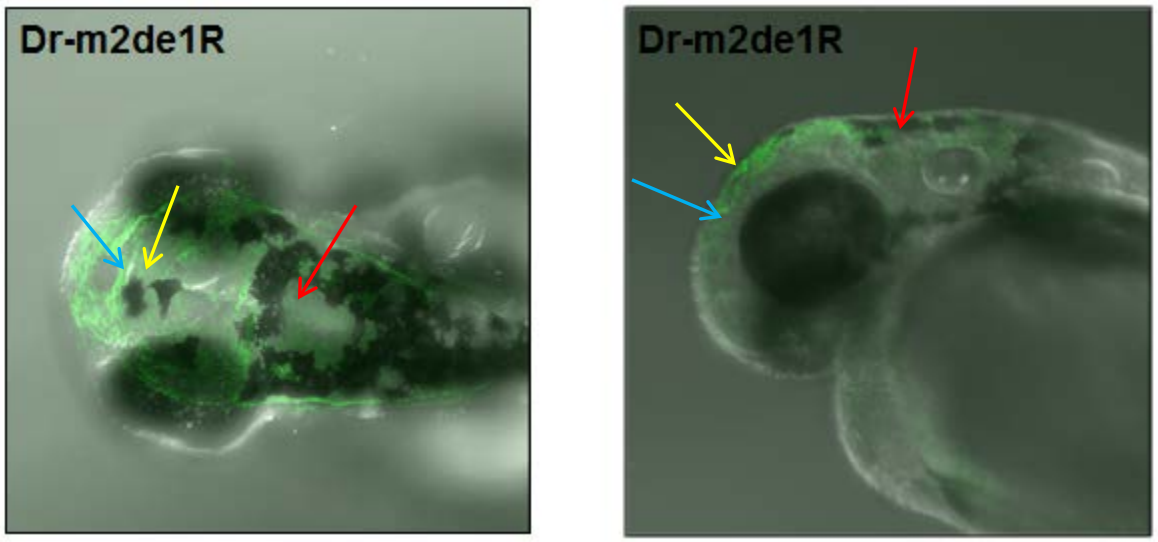


A B

Fig. 22. pDr-m2de1F-*cfos*-EGFP reporter gene expression. The images above depict EGFP expression localization in 48 hours post fertilization zebrafish embryos previously injected with pDr-m2de1F-*cfos*-EGFP construct. Two orientations were utilized for viewing the embryos: a dorsal (top-down) orientation in which the upper portion of the zebrafish head was visible (A), and a sagittal orientation in which the lateral areas of the zebrafish head could be analyzed (B). The expression of EGFP was found throughout the midbrain (blue arrow), portions of the hindbrain (red arrow), and tectum (yellow arrow).

The pDr-m2de1R-*cfos*-EGFP construct contains the zebrafish m2de1 element in its reverse orientation. A total of 16 embryos were imaged for EGFP expression at 48 hours post fertilization using the Zeiss LSM 510 Confocal Microscope after injection with the pDr-m2de1R-*cfos*-EGFP construct. Nine of these injected embryos displayed reporter gene expression. As expected, EGFP expression was seen in anterior regions of the zebrafish embryos (Fig. 23) mirroring endogenous *meis2a* expression (Fig. 20). More specifically,

reporter gene expression appears to be located within the midbrain, portions of the hindbrain, as well as the tectum (Fig. 23). Little to no EGFP expression was seen in the developing eyes of the embryos (Fig. 23).



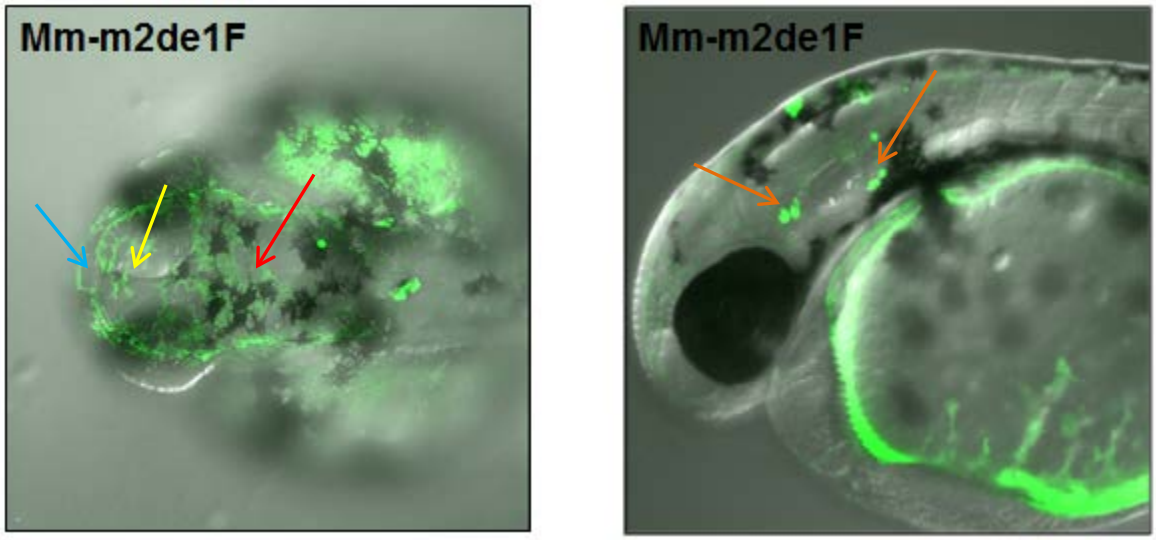
A B

Fig. 23. pDr-m2de1R-*cfos*-EGFP reporter gene expression. The images above depict EGFP expression localization in 48 hours post fertilization zebrafish embryos previously injected with pDr-m2de1R-*cfos*-EGFP construct. Two orientations were utilized for viewing the embryos: a dorsal (top-down) orientation in which the upper portion of the zebrafish head was visible (A), and a sagittal orientation in which the lateral areas of the zebrafish head could be analyzed (B). The expression of EGFP was found throughout the midbrain (blue arrow), portions of the hindbrain (red arrow), and tectum (yellow arrow).

Expression Directed by the Mouse m2de1 Element

The pMm-m2de1F-*cfos*-EGFP construct contains the mouse m2de1 element in its forward orientation. A total of 14 embryos were imaged for EGFP expression at 48 hours post fertilization using the Zeiss LSM 510 Confocal Microscope after injection with the pMm-m2de1F-*cfos*-EGFP construct. All 14 of these injected embryos displayed reporter gene expression. As expected, EGFP expression was seen in anterior regions of the zebrafish

embryos (Fig. 24) mirroring endogenous *meis2a* expression (Fig. 20). More specifically, reporter gene expression appears to be located within the midbrain, portions of the hindbrain, as well as the tectum (Fig. 24). However, the sagittal orientation of the embryo revealed some novel EGFP localizations within neurons that appear to be associated with the otic cup and developing ear (Fig. 24B). As seen before, little to no EGFP expression was seen in the developing eyes of the embryos (Fig. 24).

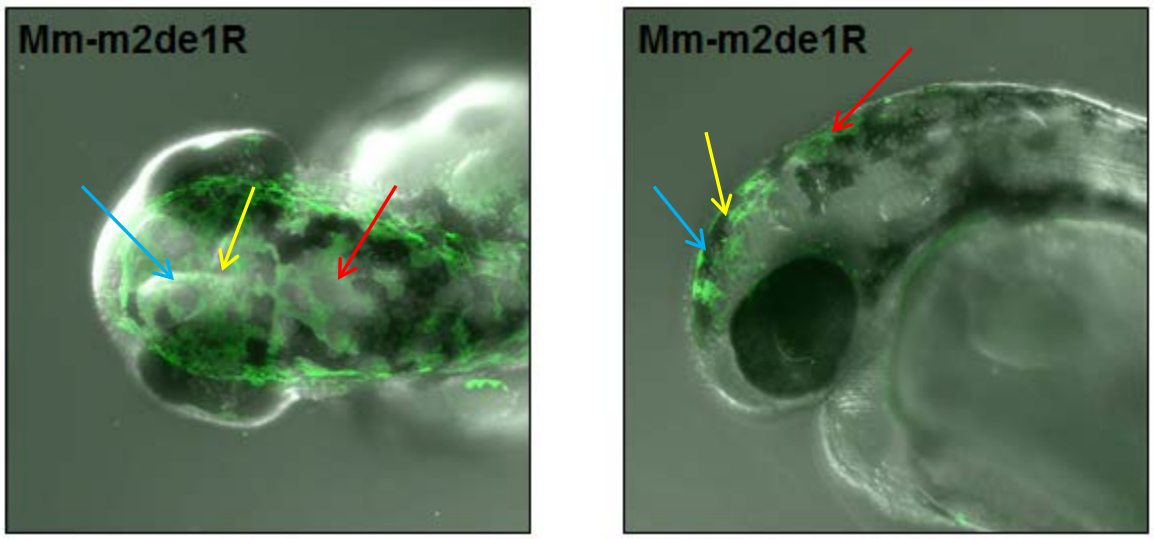


A B

Fig. 24. pMm-m2de1F-*cfos*-EGFP reporter gene expression. The images above depict EGFP expression localization in 48 hours post fertilization zebrafish embryos previously injected with pMm-m2de1F-*cfos*-EGFP construct. Two orientations were utilized for viewing the embryos: a dorsal (top-down) orientation in which the upper portion of the zebrafish head was visible (A), and a sagittal orientation in which the lateral areas of the zebrafish head could be analyzed (B). The expression of EGFP was found throughout the midbrain (blue arrow), portions of the hindbrain (red arrow), and tectum (yellow arrow). Additionally, neurons associated with the developing ear appear to possess EGFP expression (orange arrows).

The pMm-m2de1R-*cfos*-EGFP construct contains the mouse m2de1 element in its reverse orientation. A total of 13 embryos were imaged for EGFP expression at 48 hours post fertilization using the Zeiss LSM 510 Confocal Microscope after injection with the

pMm-m2de1R-*cfos*-EGFP construct. Twelve of these 13 injected embryos displayed reporter gene expression. As expected, EGFP expression was seen in anterior regions of the zebrafish embryos (Fig. 25) mirroring endogenous *meis2a* expression (Fig. 20). More specifically, reporter gene expression appears to be located within the midbrain, portions of the hindbrain, as well as the tectum (Fig. 25). However, unlike the pMm-m2de1F-*cfos*-EGFP, the sagittal orientation of the embryo did not reveal any novel EGFP localizations, but instead directed EGFP expression in regions highly similar to those directed by the zebrafish m2de1 elements (Fig. 25B). Therefore, it appears as if a simple orientation change of the element may have been enough to alter reporter gene expression, although other explanations are feasible. As seen before, little to no EGFP expression was seen in the developing eyes of the embryos (Fig. 25).



A B

Fig. 25. pMm-m2de1R-*cfos*-EGFP reporter gene expression. The images above depict EGFP expression localization in 48 hours post fertilization zebrafish embryos previously injected with pMm-m2de1R-*cfos*-EGFP construct. Two orientations were utilized for viewing the embryos: a dorsal (top-down) orientation in which the upper portion of the zebrafish head was visible (A), and a sagittal orientation in which the lateral areas of the zebrafish head could be analyzed (B). The expression of EGFP was found throughout the midbrain (blue arrow), portions of the hindbrain (red arrow), and tectum (yellow arrow).

Mutation of the Potential Hox Binding Site Alters Expression

The pDr-m2de1F-mut-*cfos*-EGFP construct contains the zebrafish m2de1 element in forward orientation but with a potential binding site sequence mutation. A total of six embryos were imaged for EGFP expression using the Zeiss LSM 510 Confocal Microscope after injection with the pDr-m2de1F-mut-*cfos*-EGFP construct at 48 hours post fertilization. In all, four of the embryos displayed reporter gene expression. Also, 16 embryos were imaged for EGFP expression using the Zeiss LSM 510 Confocal Microscope after injection with the pDr-m2de1F-mut-*cfos*-EGFP construct at 54 hours post fertilization. The 54 hours post fertilization time frame was not randomly chose but was the exact time that Parker et al. chose to view embryos that had been injected with their respective element. Seven of these 16 injected embryos displayed reporter gene expression.

By mutating the single putative Hox binding site, the expression pattern of EGFP was dramatically altered (Fig. 26). This altered expression does not resemble the expression patterns seen by Parker et al. at 54 hours post fertilization, but instead novel localizations were identified (Fig. 26C-D). Although reporter gene expression is still seen within the developing midbrain and hindbrain, EGFP appears to have become much more localized to specific neurons within the developing brain of the zebrafish embryos (Fig. 26). The 48 hours post fertilization embryos and the 54 hours post fertilization embryos possess varying EGFP expression patterns as well (Fig. 26). At 48 hours post fertilization, EGFP appears in a clump of neurons right behind the eye of the embryo, while at 54 hours post fertilization this clump is not present (Fig. 26B and Fig. 26D). Instead, EGFP appears strongly in the forebrain of the embryo at 54 hours post fertilization (Fig. 26D). Forebrain expression is

seen at 48 hours post fertilization in embryos injected with the pDr-m2de1F-mut-*cfos*-EGFP construct but its expression is very faint (Fig. 26B).

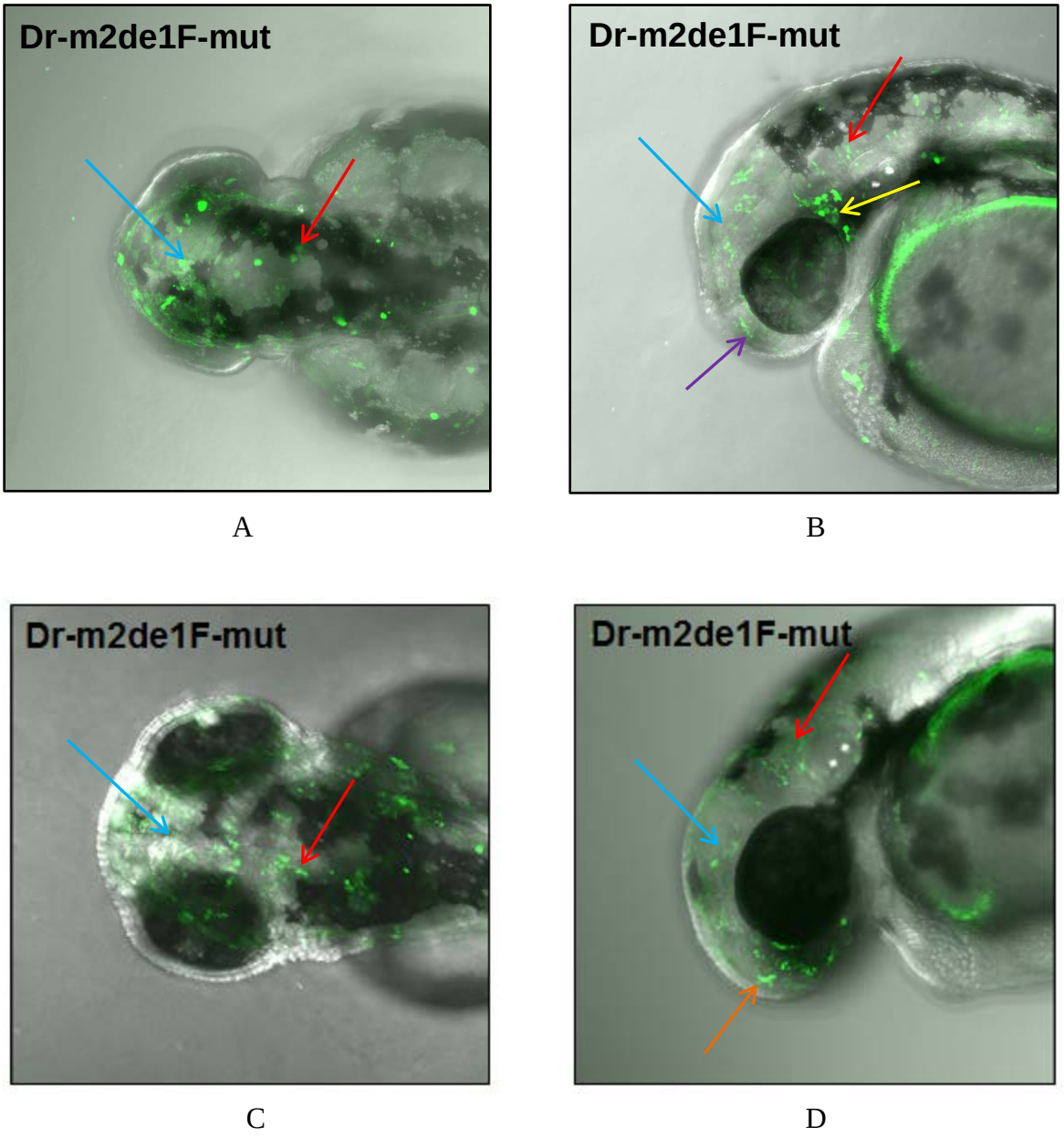


Fig. 26. pDr-m2de1F-mut-*cfos*-EGFP reporter gene expression. The images above depict EGFP expression localization in 48 hours post fertilization zebrafish embryos (A-B) and in 54 hours post fertilization zebrafish embryos (C-D) previously injected with pDr-m2de1F-mut-*cfos*-EGFP construct. Two orientations were utilized for viewing the embryos: a dorsal (top-down) orientation in which the upper portion of the zebrafish head was visible (A and C), and a sagittal orientation in which the lateral areas of the zebrafish head could be analyzed (B and D). The expression of EGFP was found throughout the midbrain (blue

arrow) and portions of the hindbrain (red arrow). At 48 hours post fertilization, EGFP expression is seen in an unidentified clump right behind the eye of the zebrafish (yellow arrow) and faintly in the forebrain (purple arrow). At 54 hours post fertilization, EGFP expression is seen prominently in the forebrain of the zebrafish (orange arrow).

DISCUSSION

To date, much has been revealed concerning the diverse roles that the *Meis* genes play in embryonic development, and the expression patterns of the *Meis* genes have been well documented; however, while the developmental importance of the *Meis* genes has been addressed, very little is known about the regulation of their expression. The *meis2a* gene in zebrafish is no exception. Of recent interest concerning transcriptional regulation of developmentally significant genes are *cis*-regulatory elements (Alberts, 2008; Kikuta et al., 2007b; Navratilova and Becker, 2009). *Cis*-regulatory elements are sequences of DNA that, while not expressed, help to direct the expression of a gene by serving as additional binding sites for transcription related proteins (Kikuta et al., 2007b; Woolfe et al., 2005). In this study, we have identified a *cis*-regulatory element that directs the expression of the *meis2a* gene in *Danio rerio*.

Due to their importance in gene expression and their relation with developmental genes, the same *cis*-regulatory elements are often found in widely divergent organisms with remarkably similar sequences (Engstrom et al., 2007; Navratilova and Becker, 2009). As a result of this conservation, phylogenetic footprinting is often used to find potential *cis*-regulatory elements (Fisher et al., 2006; Kikuta et al., 2007a; Kikuta et al., 2007b; Woolfe et al., 2005). Previously in the Zerucha lab, Phylogenetic footprinting of the human *MEIS2* sequence resulted in the detection of 4 HCNEs that have been named m2de (1-4) (Wellington and Zerucha, unpublished). All four of the m2de1 elements possess the characteristics that would be expected of functional *cis*-regulatory elements. First, the four elements are highly

conserved across the variety of organisms analyzed and in some extended areas even show 100% sequence conservation suggesting functional significance. Secondly, binding sites for transcriptional proteins, most notably Hox, Pbx and Meis factors, have been identified within the four m2de1 elements. The presence of binding sites indicates the potential for the four elements to recruit additional, yet necessary proteins for the direction of gene expression. Lastly, the elements were discovered in relative close proximity to *MEIS2*; a vital developmental gene. Only one of the identified m2de1 elements was found to be present in zebrafish; the m2de1 element (Wellington and Zerucha, unpublished). The m2de1 element possesses approximately 450 nucleotides that are conserved while 260 of those nucleotides appear to be highly conserved showing 68% sequence conservation between humans and zebrafish. Analysis of adjacent regions revealed that an additional 325 nucleotides may possess *cis*-regulatory function as well, in turn contributing to the overall *meis2a* expression pattern.

However, just identifying HCNEs does not guarantee that the conserved sequences possess actual *cis*-regulatory function. A proven method in which to test for such functionality is to couple the HCNE with a minimal promoter and a reporter gene creating an expression cassette, placing the cassette into the genome of developing embryos, and then analyzing the organism for HCNE driven reporter gene expression (Fisher et al., 2006; Li et al., 2010; Pauls et al., 2012; Zerucha et al., 2000). The minimal promoter serves to drive reporter gene expression only if the adjacent *cis*-regulatory element is bound by transcription factors (Fisher et al., 2006; Li et al., 2010). As stated previously, if the HCNE functions as a *cis*-regulatory element, the spatiotemporal characteristics of the reporter gene expression patterns should resemble the expression of the gene the HCNE works to direct (Fisher et al.,

2006; Li et al., 2010). In order to create transgenic reporter constructs containing the expression cassettes, a series of step-wise procedures were carried out in order to build the constructs. In total, four unmutated transgenic reporter constructs were developed: construct with the zebrafish m2de1 element in forward orientation (pDr-m2de1F-*cfos*-EGFP), construct with the zebrafish m2de1 element in reverse orientation (pDr-m2de1R-*cfos*-EGFP), construct with mouse m2de1 element in forward orientation (pMm-m2de1F-*cfos*-EGFP), and construct with mouse m2de1 element in reverse orientation (pMm-m2de1R-*cfos*-EGFP).

In order to determine if in fact the m2de1 element remained unaltered throughout the transgenic reporter construct construction process, the zebrafish m2de1 element amplified from the zebrafish genomic DNA, the pCR®2.1-TOPO® cloning vector containing the zebrafish m2de1 element, and the pDr-m2de1F-*cfos*-EGFP construct were sent off for Sanger DNA sequencing. The sequencing analysis of the three constructs showed that two nucleotides within the m2de1 element differ from the sequence identified in the published zebrafish genome. In the first case, a cytosine was mislabeled as a thymine and in the second case, a guanine was mislabeled as an adenine. All three of the sequencing analyses support this finding as all three sequenced constructs contained nucleotide variances in the same positions within the m2de1 sequence, and all three constructs had the same nucleotides in those locations also. It is possible this is due to differences between the individual fish used for the zebrafish genome project and for our isolation of the m2de1 region, or perhaps due to errors in the zebrafish genome sequence. As these variant nucleotides were determined to not be within any obvious transcription factor binding sites, they should not have been detrimental to the project. Nonetheless, such a finding is rather interesting and warrants further study.

Research published during the construction of the aforementioned unmutated transgenic reporter constructs revealed that a *cis*-regulatory element containing an almost identical sequence to the m2de1 element directed reporter gene expression in a different way than the pDr-m2de1F-*cfos*-EGFP construct did (Parker et al., 2011). The element described in this published study was identical to m2de1 except that the element lacked 17 additional base pairs, including a Hox binding site (TAAT) that was incorporated into the m2de1 element. In order to determine if this 17 base pair region was responsible for the differences observed between our experiments and those described in the Parker paper, I mutated the potential Hox binding site within the m2de1 element that was not present in the Parker element. The final construct contained the mutated zebrafish m2de1 element in forward orientation (pDr-m2de1F-mut-*cfos*-EGFP). After the pDr-m2de1F-mut-*cfos*-EGFP reporter construct was made, restriction with ClaI resulted in a plasmid band of 6700 base pairs and a larger than expected insert band of 1600 base pairs. Sequencing of the plasmid established that the TAATG sequence containing the Hox binding site had indeed been altered to the sequence TCGAT and that the insert was correct. The incorrect size of the smaller band is puzzling as the total size of the intact construct itself should only be 7941 base pairs long, not 8300 base pairs long as the restriction analysis suggests. One possible explanation for this discrepancy is that the ClaI enzyme remained attached to the fragment during gel electrophoresis and resulted in the size ambiguity.

Microinjections were subsequently performed to determine if the m2de1 element is able to direct expression of EGFP in a manner that is consistent with the known expression pattern of *Meis2*. After microinjection into the zebrafish embryos yolk, the pCS2FA Tol2 Transposase mRNA was translated into transposase protein that subsequently recognized the

Tol2 recombination sites flanking the m2de1 expression cassette. The m2de1 element was cleaved from the reporter construct by the Tol2 transposase and relocated into the embryos genome where *cis*-regulatory functionality would be tested.

The analysis of the transgenic embryos revealed that the m2de1 element does in fact drive expression in a manner consistent with the *meis2a* gene in zebrafish and potentially the *Meis2* gene in mouse. In total, all five experimental constructs as well as a positive control construct were injected and analyzed: positive control construct (pDest-Sox10-mCherry-Tol2CG2; graciously provided by Dr. Chi-Bin Chien), construct with the zebrafish m2de1 element in forward orientation (pDr-m2de1F-*cfos*-EGFP), construct with the zebrafish m2de1 element in reverse orientation (pDr-m2de1R-*cfos*-EGFP), construct with mouse m2de1 element in forward orientation (pMm-m2de1F-*cfos*-EGFP), construct with mouse m2de1 element in reverse orientation (pMm-m2de1R-*cfos*-EGFP), and the construct containing the zebrafish m2de1 element in forward orientation but with a potential Hox binding site sequence mutation (pDr-m2de1F-mut-*cfos*-EGFP). I decided to analyze the transgenic embryos at varying time frames post fertilization as m2de1 is a novel, putative enhancer. Although m2de1 was predicted to drive *meis2a* expression at 24 hours post fertilization due to the expression patterns previously observed, enhancers are known to vary in function based on their cellular context and molecular environment and the transcriptional proteins that bind to them (Carpenter, 2010; Woolfe et al., 2005). Furthermore, many developmentally important genes utilize multiple enhancers to drive their expression (Kikuta et al., 2007; Navratilova and Becker, 2009). Taking these two ideas into account, it is easy to see that by testing the putative enhancers out of their natural cellular and molecular "environment," it is possible to have expression differ spatially and temporally from

endogenous expression (Woolfe et al., 2005). EGFP expression was soon seen at 48 hours post fertilization, a rather unexpected discovery.

The expression patterns of the EGFP resulting from the injection of the unmutated constructs were extremely similar to the *meis2a* expression patterns. This not only shows that the m2de1 element has *cis*-regulatory function, but reveals that the m2de1 element may potentially be involved in the direction of *meis2a* expression. Reporter gene expression was seen confined to anterior regions, most notably in the midbrain and parts of the hindbrain, but not directly within the developing eye or retina where *meis2a* is known to be expressed (Carpenter, 2010). Expression appears to be rather non-localized and spread out through the anterior areas. Although expression was diffuse throughout neurons in these regions and not necessarily localized to any particular identifiable structure, expression was seen within the tectum corroborating with the *meis2a* expression profile (Carpenter, 2010). This is a rather striking discovery because, as stated earlier, *Meis2* is believed to play a role in eye development. The tectum is responsible for visual reflexes and processing; therefore m2de1 seems to at least be partially responsible for the expression of *Meis2* within the tectum, implicating m2de1 with the *Meis2* gene's association with vision.

Only one of the unmutated constructs resulted in a divergent EGFP expression profile; the pMm-m2de1F-*cfos*-EGFP construct. For some reason, this element when injected in zebrafish drives EGFP expression within neurons associated with the otic cup and the developing ear. Conversely, pMm-m2de1R-*cfos*-EGFP drove EGFP expression in the same way that pDr-m2de1F-*cfos*-EGFP and pDr-m2de1R-*cfos*-EGFP did. A logical explanation for this anomaly could be that the expression cassette within the pMm-m2de1F-*cfos*-EGFP construct was inserted nearby other *cis*-regulatory elements that direct expression

of genes expressed near the zebrafish ear. These *cis*-regulatory elements could have sufficiently mobilized the minimal promoter resulting in EGFP expression in a rather unusual location. Regardless, the fact that the mouse element in the pMm-m2de1R-*cfos*-EGFP construct drove EGFP expression in the same way as the zebrafish m2de1 element after insertion into the zebrafish genome further shows that the m2de1 element functions as a *cis*-regulatory element. Such expression pattern similarity indicates conservation in function between the elements from zebrafish and mouse strengthening the claim that the m2de1 element has resisted evolutionary change due to inherent developmental importance as a *cis*-regulatory element.

The mutated construct, pDr-m2de1F-mut-*cfos*-EGFP, resulted in an EGFP expression pattern that was dramatically altered. Oddly, this different expression does not resemble the expression patterns seen by Parker et al. as eradicating the additional putative Hox binding site should have eradicated any functional differences that potentially exists between the m2de1 elements and the elements utilized by Parker and colleagues. Instead, injection of the mutated reporter construct resulted in the localization of EGFP in previously unseen areas. In fact, the mutated m2de1 element appears to have directed the EGFP expression to more specific neurons and locations within the brain of the zebrafish embryos, although this localized reporter gene expression is still seen within the developing midbrain and hindbrain. The 48 hours and the 54 hours post fertilization zebrafish embryos have different EGFP expression patterns also. For the embryos injected with the pDr-m2de1F-mut-*cfos*-EGFP construct, many were viewed at 54 hours post fertilization for comparison purposes because that was the time in development that Parker and colleagues viewed their transgenic embryos (Parker et al., 2011). At 48 hours post fertilization, EGFP is seen in a cluster of neurons

directly posterior from the eye of the embryo, while at 54 hours post fertilization this mass is not seen. The EGFP also appears intensely in the forebrain of the embryo at 54 hours post fertilization, while in the 48 hours post fertilization embryos injected with the pDr-m2de1F-mut-*cfos*-EGFP construct, EGFP expression is very faint in the forebrain. Therefore, it is readily apparent how the eradication of a single binding site can completely alter the expression profiles associated with a specific gene. These differential expression patterns stemming from the mutation could be the result of numerous potential dynamics. One of the most probable explanations is that by eradicating the Hox binding site, we could have inadvertently removed a repressor site within the m2de1 element. This could in turn allow for expression in locales such as the forebrain and in the area posterior to the eye where expression is normally repressed by the *cis*-regulatory mechanism. Another explanation for the divergent expression patterns that resulted from injection of the mutated transgenic reporter construct is that we could have introduced another binding site in place of the Hox binding site. Although no literature exists stating that the TCGAT sequence serves as a transcription factor binding site, there exists an outside chance that it actually does. If the sequence is a binding site, the binding of novel proteins to the m2de1 element could alter the expression driven by the *cis*-regulatory element explaining the unusual EGFP localizations.

In conclusion, data resulting from this study concerning the functionality of a previously unidentified putative enhancer element indicates that m2de1 is a putative *cis*-regulatory element, potentially working to direct the expression of *meis2a* in zebrafish. However, the m2de1 element by itself and without its correct molecular surroundings is unable to recapitulate the *meis2a* expression profile in its entirety and it remains to be seen if a direct correlation actually exists between the m2de1 element and *meis2a* regulation. Taken

together, the work done for this study helped to characterize a novel *cis*-regulatory element possibly involved in the regulation of *Meis* expression.

REFERENCES

Akoulitchev, S., Makela, T. P., Weinberg, R. A., Reinberg, D., 1995. Requirement for TFIIH kinase activity in transcription by RNA polymerase II. Nature. 377**,** 557-560.

Alberts, B., 2008. Molecular biology of the cell. Garland Science, New York.

Amores, A., Force, A., Yan, Y. L., Joly, L., Amemiya, C., Fritz, A., Ho, R. K., Langeland, J., Prince, V., Wang, Y. L., Westerfield, M., Ekker, M., Postlethwait, J. H., 1998. Zebrafish hox clusters and vertebrate genome evolution. Science 282**,** 1711-1714.

Amsterdam, A., Lin, S., Hopkins, N., 1995. The Aequorea victoria green fluorescent protein can be used as a reporter in live zebrafish embryos. Dev. Biol. 171**,** 123-129.

Azcoitia, V., Aracil, M., Martinez, A. C., Torres, M., 2005. The homeodomain protein Meis1 is essential for definitive hematopoiesis and vascular patterning in the mouse embryo. Dev. Biol. 280**,** 307-320.

Balavoine, G., de Rosa, R., Adoutte, A., 2002. Hox clusters and bilaterian phylogeny. Mol. Phylogenet. Evol. 24**,** 366-373.

Becker, A., Bey, M., Burglin, T. R., Saedler, H., Theissen, G., 2002. Ancestry and diversity of BEL1-like homeobox genes revealed by gymnosperm (Gnetum gnemon) homologs. Dev. Genes Evol. 212**,** 452-457.

Bell, A. C., West, A. G., Felsenfeld, G., 2001. Insulators and boundaries: versatile regulatory elements in the eukaryotic genome. Science 291**,** 447-450.

Berthelsen, J., Kilstrup-Nielsen, C., Blasi, F., Mavilio, F., Zappavigna, V., 1999. The subcellular localization of PBX1 and EXD proteins depends on nuclear import and

export signals and is modulated by association with PREP1 and HTH. Genes Dev. 13**,** 946-953.

Berthelsen, J., Zappavigna, V., Mavilio, F., Blasi, F., 1998. Prep1, a novel functional partner of Pbx proteins. EMBO J. 17**,** 1423-1433.

Bessa, J., Tavares, M. J., Santos, J., Kikuta, H., Laplante, M., Becker, T. S., Gomez-Skarmeta, J. L., Casares, F., 2008. meis1 regulates cyclin D1 and c-myc expression, and controls the proliferation of the multipotent cells in the early developing zebrafish eye. Development 135**,** 799-803.

Biemar, F., Devos, N., Martial, J. A., Driever, W., Peers, B., 2001. Cloning and expression of the TALE superclass homeobox Meis2 gene during zebrafish embryonic development. Mech. Dev. 109**,** 427-431.

Blader, P., Lam, C. S., Rastegar, S., Scardigli, R., Nicod, J. C., Simplicio, N., Plessy, C., Fischer, N., Schuurmans, C., Guillemot, F., Strahle, U., 2004. Conserved and acquired features of neurogenin1 regulation. Development 131**,** 5627-5637.

Bomgardner, D., Hinton, B. T., Turner, T. T., 2003. 5' hox genes and meis 1, a hox-DNA binding cofactor, are expressed in the adult mouse epididymis. Biol. Reprod. 68**,** 644-650.

Bonner, J., Dahmus, M. E., Fambrough, D., Huang, R. C., Marushige, K., Tuan, D. Y., 1968. The Biology of Isolated Chromatin: Chromosomes, biologically active in the test tube, provide a powerful tool for the study of gene action. Science 159**,** 47-56.

Brand, A. H., Breeden, L., Abraham, J., Sternglanz, R., Nasmyth, K., 1985. Characterization of a "silencer" in yeast: a DNA sequence with properties opposite to those of a transcriptional enhancer. Cell 41**,** 41-48.

Brunet, F. G., Roest Crollius, H., Paris, M., Aury, J. M., Gibert, P., Jaillon, O., Laudet, V., Robinson-Rechavi, M., 2006. Gene loss and evolutionary rates following whole-genome duplication in teleost fishes. Mol. Biol. Evol. 23**,** 1808-1816.

Bumsted-O'Brien, K. M., Hendrickson, A., Haverkamp, S., Ashery-Padan, R., Schulte, D., 2007. Expression of the homeodomain transcription factor Meis2 in the embryonic and postnatal retina. J. Comp. Neurol. 505**,** 58-72.

Buratowski, S., Hahn, S., Guarente, L., 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., Felsenfeld, G., 2002. The insulation of genes from external enhancers and silencing chromatin. Proc. Natl. Acad. Sci. U.S.A. 99 (Suppl 4)**,** 16433-16437.

Burglin, T. R., 1997. Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals. Nucleic Acids Res. 25**,** 4173-4180.

Burglin, T. R., 1998. The PBC domain contains a MEINOX domain: coevolution of Hox and TALE homeobox genes? Dev. Genes Evol. 208**,** 113-116.

Bustos, M. M., Begum, D., Kalkan, F. A., Battraw, M. J., Hall, T. C., 1991. Positive and negative cis-acting DNA domains are required for spatial and temporal regulation of gene expression by a seed storage protein promoter. EMBO J. 10**,** 1469-1479.

Calo, E., Wysocka, J., 2013. Modification of enhancer chromatin: what, how, and why? Mol. Cell 49**,** 825-837.

Capdevila, J., Tsukui, T., Rodriquez Esteban, C., Zappavigna, V., Izpisua Belmonte, J. C., 1999. Control of vertebrate limb outgrowth by the proximal factor Meis2 and distal antagonism of BMPs by Gremlin. Mol. Cell. 4**,** 839-849.

Carbe, C., Hertzler-Schaefer, K., Zhang, X., 2012. The functional role of the Meis/Prep-binding elements in Pax6 locus during pancreas and eye development. Dev. Biol. 363**,** 320-329.

Carpenter, B.S., 2010. Identification and developmental expression of the zebrafish ZGC:154061 gene, a conserved yet uncharacterized maternally expressed *Meis2* linked gene. Department of Biology, Appalachian State University, Boone, NC.

Carroll, S. B., Prud'homme, B., Gompel, N., 2008. Regulating evolution. Sci. Am. 298**,** 60-67.

Castillo-Davis, C. I., Hartl, D. L., Achaz, G., 2004. cis-Regulatory and protein evolution in orthologous and duplicate genes. Genome Res. 14**,** 1530-1536.

Cecconi, F., Proetzel, G., Alvarez-Bolado, G., Jay, D., 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. Dev. Dyn. 210**,** 184-190.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., Prasher, D. C., 1994. Green fluorescent protein as a marker for gene expression. Science 263**,** 802-805.

Chang, C. P., Brocchieri, L., Shen, W. F., Largman, C., Cleary, M. L., 1996. Pbx modulation of Hox homeodomain amino-terminal arms establishes different DNA-binding specificities across the Hox locus. Mol. Cell Biol. 16**,** 1734-1745.

Chang, C. P., Jacobs, Y., Nakamura, T., Jenkins, N. A., Copeland, N. G., Cleary, M. L., 1997. Meis proteins are major in vivo DNA binding partners for wild-type but not chimeric Pbx proteins. Mol. Cell Biol. 17**,** 5679-5687.

Choe, S. K., Lu, P., Nakamura, M., Lee, J., Sagerstrom, C. G., 2009. Meis cofactors control HDAC and CBP accessibility at Hox-regulated promoters during zebrafish embryogenesis. Dev. Cell. 17**,** 561-567.

Choe, S. K., Vlachakis, N., Sagerstrom, C. G., 2002. Meis family proteins are required for hindbrain development in the zebrafish. Development 129**,** 585-595.

Chong, S. W., Korzh, V., Jiang, Y. J., 2009. Myogenesis and molecules - insights from zebrafish Danio rerio. J. Fish Biol. 74**,** 1693-1755.

Clark, A. R., Docherty, K., 1993. Negative regulation of transcription in eukaryotes. Biochem. J. 296 (Pt 3)**,** 521-541.

Conte, I., Carrella, S., Avellino, R., Karali, M., Marco-Ferreres, R., Bovolenta, P., Banfi, S., 2010. miR-204 is required for lens and retinal development via Meis2 targeting. Proc. Natl. Acad. Sci. U.S.A. 107**,** 15491-15496.

Coy, S. E., Borycki, A. G., 2010. Expression analysis of TALE family transcription factors during avian development. Dev. Dyn. 239**,** 1234-1245.

Crowley, M. A., Conlin, L. K., Zackai, E. H., Deardorff, M. A., Thiel, B. D., Spinner, N. B., 2010. Further evidence for the possible role of MEIS2 in the development of cleft palate and cardiac septum. Am. J. Med. Genet. A. 152A**,** 1326-1327.

Davidson, E. H., 2001. Genomic regulatory systems : development and evolution. Academic Press, San Diego, CA.

Davidson, E. H., 2006. The regulatory genome : gene regulatory networks in development and evolution. Academic Press, San Diego, CA.

Dibner, C., Elias, S., Frank, D., 2001. XMeis3 protein activity is required for proper hindbrain patterning in Xenopus laevis embryos. Development 128**,** 3415-3426.

diIorio, P., Alexa, K., Choe, S. K., Etheridge, L., Sagerstrom, C. G., 2007. TALE-family homeodomain proteins regulate endodermal sonic hedgehog expression and pattern the anterior endoderm. Dev. Biol. 304**,** 221-231.

Dorn, A., Affolter, M., Gehring, W. J., Leupin, W., 1994. Homeodomain proteins in development and therapy. Pharmacol. Ther. 61**,** 155-184.

Drapkin, R., Reardon, J. T., Ansari, A., Huang, J. C., Zawel, L., Ahn, K., Sancar, A., Reinberg, D., 1994. Dual role of TFIIH in DNA excision repair and in transcription by RNA polymerase II. Nature 368**,** 769-772.

Duboule, D., 1998. Vertebrate hox gene regulation: clustering and/or colinearity? Curr. Opin. Genet. Dev. 8**,** 514-518.

Dutton, J. R., Antonellis, A., Carney, T. J., Rodrigues, F. S., Pavan, W. J., Ward, A., Kelsh, R. N., 2008. An evolutionarily conserved intronic region controls the spatiotemporal expression of the transcription factor Sox10. BMC Dev. Biol. 8**,** 105.

Echelard, Y., Vassileva, G., McMahon, A. P., 1994. Cis-acting regulatory sequences governing Wnt-1 expression in the developing mouse CNS. Development 120**,** 2213-2224.

Ekker, S. C., Jackson, D. G., von Kessler, D. P., Sun, B. I., Young, K. E., Beachy, P. A., 1994. The degree of variation in DNA sequence recognition among four Drosophila homeotic proteins. EMBO J. 13**,** 3551-3560.

Elkouby, Y. M., Elias, S., Casey, E. S., Blythe, S. A., Tsabar, N., Klein, P. S., Root, H., Liu, K. J., Frank, D., 2010. Mesodermal Wnt signaling organizes the neural plate via Meis3. Development 137**,** 1531-1541.

Engstrom, P. G., Ho Sui, S. J., Drivenes, O., Becker, T. S., Lenhard, B., 2007. Genomic regulatory blocks underlie extensive microsynteny conservation in insects. Genome Res. 17**,** 1898-1908.

Fiering, S., Whitelaw, E., Martin, D. I., 2000. To be or not to be active: the stochastic nature of enhancer action. Bioessays 22**,** 381-387.

Fisher, S., Grice, E. A., Vinton, R. M., Bessling, S. L., Urasaki, A., Kawakami, K., McCallion, A. S., 2006. Evaluating the biological relevance of putative enhancers using Tol2 transposon-mediated transgenesis in zebrafish. Nat. Protoc. 1**,** 1297-1305.

French, C. R., Erickson, T., Callander, D., Berry, K. M., Koss, R., Hagey, D. W., Stout, J., Wuennenberg-Stapleton, K., Ngai, J., Moens, C. B., Waskiewicz, A. J., 2007. Pbx homeodomain proteins pattern both the zebrafish retina and tectum. BMC Dev. Biol. 7**,** 85.

Frith, M. C., Hansen, U., Weng, Z., 2001. Detection of cis-element clusters in higher eukaryotic DNA. Bioinformatics 17**,** 878-889.

Fujino, T., Yamazaki, Y., Largaespada, D. A., Jenkins, N. A., Copeland, N. G., Hirokawa, K., Nakamura, T., 2001. Inhibition of myeloid differentiation by Hoxa9, Hoxb8, and Meis homeobox genes. Exp. Hematol. 29**,** 856-863.

Geerts, D., Revet, I., Jorritsma, G., Schilderink, N., Versteeg, R., 2005. MEIS homeobox genes in neuroblastoma. Cancer Lett. 228**,** 43-50.

Gehring, W. J., 1987. Homeo boxes in the study of development. Science 236**,** 1245-1252.

Gehring, W. J., 1993. Exploring the homeobox. Gene 135**,** 215-221.

Gehring, W. J., Affolter, M., Burglin, T., 1994. Homeodomain proteins. Annu. Rev. Biochem. 63**,** 487-526.

Gerasimova, T. I., Corces, V. G., 2001. Chromatin insulators and boundaries: effects on transcription and nuclear organization. Annu. Rev. Genet. 35**,** 193-208.

Geyer, P. K., 1997. The role of insulator elements in defining domains of gene expression. Curr. Opin. Genet. Dev. 7**,** 242-248.

Gibcus, J. H., Dekker, J., 2013. The Hierarchy of the 3D Genome. Mol. Cell. 49**,** 773-82.

Gompel, N., Prud'homme, B., Wittkopp, P. J., Kassner, V. A., Carroll, S. B., 2005. Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in Drosophila. Nature 433**,** 481-487.

Gould, A., Morrison, A., Sproat, G., White, R. A., Krumlauf, R., 1997. Positive cross-regulation and enhancer sharing: two mechanisms for specifying overlapping Hox expression patterns. Genes Dev. 11**,** 900-913.

Greer, J. M., Puetz, J., Thomas, K. R., Capecchi, M. R., 2000. Maintenance of functional equivalence during paralogous Hox gene evolution. Nature 403**,** 661-665.

Griffiths, A. J. F., 2000. An introduction to genetic analysis. W.H. Freeman, New York.

Grunstein, M., 1997. Histone acetylation in chromatin structure and transcription. Nature 389**,** 349-352.

Grunwald, D. J., Eisen, J. S., 2002. Headwaters of the zebrafish -- emergence of a new model vertebrate. Nat. Rev. Genet. 3**,** 717-724.

Heine, P., Dohle, E., Bumsted-O'Brien, K., Engelkamp, D., Schulte, D., 2008. Evidence for an evolutionary conserved role of homothorax/Meis1/2 during vertebrate retina development. Development 135**,** 805-811.

Heine, P., Dohle, E., Schulte, D., 2009. Sonic hedgehog signaling in the chick retina accelerates Meis2 downregulation simultaneously with retinal ganglion cell genesis. Neuroreport 20**,** 279-284.

Hisa, T., Spence, S. E., Rachel, R. A., Fujita, M., Nakamura, T., Ward, J. M., Devor-Henneman, D. E., Saiki, Y., Kutsuna, H., Tessarollo, L., Jenkins, N. A., Copeland, N. G., 2004. Hematopoietic, angiogenic and eye defects in Meis1 mutant animals. EMBO J. 23**,** 450-459.

Holland, P. W., Garcia-Fernandez, J., 1996. Hox genes and chordate evolution. Dev. Biol. 173**,** 382-395.

Huang, H., Rastegar, M., Bodner, C., Goh, S. L., Rambaldi, I., Featherstone, M., 2005. MEIS C termini harbor transcriptional activation domains that respond to cell signaling. J. Biol. Chem. 280**,** 10119-10127.

Hughes, J. R., Cheng, J. F., Ventress, N., Prabhakar, S., Clark, K., Anguita, E., De Gobbi, M., de Jong, P., Rubin, E., Higgs, D. R., 2005. Annotation of cis-regulatory elements by identification, subclassification, and functional assessment of multispecies conserved sequences. Proc. Natl. Acad. Sci. U.S.A. 102**,** 9830-9835.

Irimia, M., Maeso, I., Burguera, D., Hidalgo-Sanchez, M., Puelles, L., Roy, S. W., Garcia-Fernandez, J., Ferran, J. L., 2011. Contrasting 5' and 3' evolutionary histories and frequent evolutionary convergence in Meis/hth gene structures. Genome Biol. Evol. 3**,** 551-564.

Jacobs, Y., Schnabel, C. A., Cleary, M. L., 1999. Trimeric association of Hox and TALE homeodomain proteins mediates Hoxb2 hindbrain enhancer activity. Mol. Cell Biol. 19**,** 5134-5142.

Juo, Z. S., Chiu, T. K., Leiberman, P. M., Baikalov, I., Berk, A. J., Dickerson, R. E., 1996. How proteins recognize the TATA box. J. Mol. Biol. 261**,** 239-254.

Kikuta, H., Fredman, D., Rinkwitz, S., Lenhard, B., Becker, T. S., 2007a. Retroviral enhancer detection insertions in zebrafish combined with comparative genomics reveal genomic regulatory blocks - a fundamental feature of vertebrate genomes. Genome Biol. 8 Suppl 1**,** S4.

Kikuta, H., Laplante, M., Navratilova, P., Komisarczuk, A. Z., Engstrom, P. G., Fredman, D., Akalin, A., Caccamo, M., Sealy, I., Howe, K., Ghislain, J., Pezeron, G., Mourrain, P., Ellingsen, S., Oates, A. C., Thisse, C., Thisse, B., Foucher, I., Adolf, B., Geling, A., Lenhard, B., Becker, T. S., 2007b. Genomic regulatory blocks encompass multiple neighboring genes and maintain conserved synteny in vertebrates. Genome Res. 17**,** 545-555.

Kim, Y., Geiger, J. H., Hahn, S., Sigler, P. B., 1993. Crystal structure of a yeast TBP/TATA-box complex. Nature 365**,** 512-520.

Kuraku, S., Meyer, A., 2009. The evolution and maintenance of Hox gene clusters in vertebrates and the teleost-specific genome duplication. Int. J. Dev. Biol. 53**,** 765-773.

Kurant, E., Pai, C. Y., Sharf, R., Halachmi, N., Sun, Y. H., Salzberg, A., 1998. Dorsotonals/homothorax, the Drosophila homologue of meis1, interacts with extradenticle in patterning of the embryonic PNS. Development 125**,** 1037-1048.

Kwan, C. T., Tsang, S. L., Krumlauf, R., Sham, M. H., 2001. Regulatory analysis of the mouse Hoxb3 gene: multiple elements work in concert to direct temporal and spatial patterns of expression. Dev. Biol. 232**,** 176-190.

Kwan, K. M., Fujimoto, E., Grabher, C., Mangum, B. D., Hardy, M. E., Campbell, D. S., Parant, J. M., Yost, H. J., Kanki, J. P., Chien, C. B., 2007. The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. Dev. Dyn. 236**,** 3088-3099.

Lawrence, H. J., Rozenfeld, S., Cruz, C., Matsukuma, K., Kwong, A., Komuves, L., Buchberg, A. M., Largman, C., 1999. Frequent co-expression of the HOXA9 and MEIS1 homeobox genes in human myeloid leukemias. Leukemia 13**,** 1993-1999.

Lee, D. Y., Hayes, J. J., Pruss, D., Wolffe, A. P., 1993. A positive role for histone acetylation in transcription factor access to nucleosomal DNA. Cell 72**,** 73-84.

Lemons, D., McGinnis, W., 2006. Genomic evolution of Hox gene clusters. Science 313**,** 1918-1922.

Li, Q., Ritter, D., Yang, N., Dong, Z., Li, H., Chuang, J. H., Guo, S., 2010. A systematic approach to identify functional motifs within vertebrate developmental enhancers. Dev. Biol. 337**,** 484-595.

Li, X. Y., Virbasius, A., Zhu, X., Green, M. R., 1999. Enhancement of TBP binding by activators and general transcription factors. Nature 399**,** 605-609.

Liu, Y., MacDonald, R. J., Swift, G. H., 2001. DNA binding and transcriptional activation by a PDX1.PBX1b.MEIS2b trimer and cooperation with a pancreas-specific basic helix-loop-helix complex. J. Biol. Chem. 276**,** 17985-17993.

Ludwig, M. Z., 2002. Functional evolution of noncoding DNA. Curr. Opin. Genet. Dev. 12**,** 634-639.

Maeda, R., Ishimura, A., Mood, K., Park, E. K., Buchberg, A. M., Daar, I. O., 2002. Xpbx1b and Xmeis1b play a collaborative role in hindbrain and neural crest gene expression in Xenopus embryos. Proc. Natl. Acad. Sci. U.S.A. 99**,** 5448-5453.

Mann, R. S., Affolter, M., 1998. Hox proteins meet more partners. Curr. Opin. Genet. Dev. 8**,** 423-429.

Mathis, D. J., Chambon, P., 1981. The SV40 early region TATA box is required for accurate in vitro initiation of transcription. Nature 290**,** 310-315.

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

McGinnis, W., Levine, M. S., Hafen, E., Kuroiwa, A., Gehring, W. J., 1984. A conserved DNA sequence in homoeotic genes of the Drosophila Antennapedia and bithorax complexes. Nature 308**,** 428-433.

Mercader, N., Leonardo, E., Azpiazu, N., Serrano, A., Morata, G., Martinez, C., Torres, M., 1999. Conserved regulation of proximodistal limb axis development by Meis1/Hth. Nature 402**,** 425-429.

Mercader, N., Leonardo, E., Piedra, M. E., Martinez, A. C., Ros, M. A., Torres, M., 2000. Opposing RA and FGF signals control proximodistal vertebrate limb development through regulation of Meis genes. Development 127**,** 3961-3970.

Mercader, N., Tanaka, E. M., Torres, M., 2005. Proximodistal identity during vertebrate limb regeneration is regulated by Meis homeodomain proteins. Development 132**,** 4131-4142.

Moens, C. B., Selleri, L., 2006. Hox cofactors in vertebrate development. Dev. Biol. 291**,** 193-206.

Mojsin, M., Stevanovic, M., 2010. PBX1 and MEIS1 up-regulate SOX3 gene expression by direct interaction with a consensus binding site within the basal promoter region. Biochem. J. 425**,** 107-116.

Morata, G., Lawrence, P. A., 1977. The development of wingless, a homeotic mutation of Drosophila. Dev. Biol. 56**,** 227-240.

Moskow, J. J., Bullrich, F., Huebner, K., Daar, I. O., Buchberg, A. M., 1995. Meis1, a PBX1-related homeobox gene involved in myeloid leukemia in BXH-2 mice. Mol. Cell Biol. 15**,** 5434-5443.

Mukherjee, K., Burglin, T. R., 2007. Comprehensive analysis of animal TALE homeobox genes: new conserved motifs and cases of accelerated evolution. J. Mol. Evol. 65**,** 137-153.

Nakamura, T., Jenkins, N. A., Copeland, N. G., 1996. Identification of a new family of Pbx-related homeobox genes. Oncogene 13**,** 2235-2242.

Nakatani, Y., Horikoshi, M., Brenner, M., Yamamoto, T., Besnard, F., Roeder, R. G., Freese, E., 1990. A downstream initiation element required for efficient TATA box binding and in vitro function of TFIID. Nature 348**,** 86-88.

Navratilova, P., Becker, T. S., 2009. Genomic regulatory blocks in vertebrates and implications in human disease. Brief Funct. Genomic Proteomic. 8**,** 333-342.

Navratilova, P., Fredman, D., Lenhard, B., Becker, T. S., 2010. Regulatory divergence of the duplicated chromosomal loci sox11a/b by subpartitioning and sequence evolution of enhancers in zebrafish. Mol. Genet. Genomics 283**,** 171-184.

Nelson, K.C., 2011. Identification of gene regulatory elements associated with the *Meis* family of homeobox genes. Depratment of Biology, Appalachian State University, Boone, NC.

Nikolov, D. B., Chen, H., Halay, E. D., Hoffman, A., Roeder, R. G., Burley, S. K., 1996. Crystal structure of a human TATA box-binding protein/TATA element complex. Proc. Natl. Acad. Sci. U.S.A. 93**,** 4862-4867.

O'Shea-Greenfield, A., Smale, S. T., 1992. Roles of TATA and initiator elements in determining the start site location and direction of RNA polymerase II transcription. J. Biol. Chem. 267**,** 6450.

Orphanides, G., Lagrange, T., Reinberg, D., 1996. The general transcription factors of RNA polymerase II. Genes Dev. 10**,** 2657-2683.

Otting, G., Qian, Y. Q., Muller, M., Affolter, M., Gehring, W., Wuthrich, K., 1988. Secondary structure determination for the Antennapedia homeodomain by nuclear magnetic resonance and evidence for a helix-turn-helix motif. EMBO J. 7**,** 4305-4309.

Oulad-Abdelghani, M., Chazaud, C., Bouillet, P., Sapin, V., Chambon, P., Dolle, P., 1997. Meis2, a novel mouse Pbx-related homeobox gene induced by retinoic acid during differentiation of P19 embryonal carcinoma cells. Dev. Dyn. 210**,** 173-183.

Pai, C. Y., Kuo, T. S., Jaw, T. J., Kurant, E., Chen, C. T., Bessarab, D. A., Salzberg, A., Sun, Y. H., 1998. The Homothorax homeoprotein activates the nuclear localization of another homeoprotein, extradenticle, and suppresses eye development in Drosophila. Genes Dev. 12**,** 435-446.

Pankotai, T., Soutoglou, E., 2013. Double strand breaks: Hurdles for RNA polymerase II transcription? Transcription 4**,** 34-38.

Parichy, D. M., Elizondo, M. R., Mills, M. G., Gordon, T. N., Engeszer, R. E., 2009. Normal table of postembryonic zebrafish development: staging by externally visible anatomy of the living fish. Dev. Dyn. 238**,** 2975-3015.

Parker, H. J., Piccinelli, P., Sauka-Spengler, T., Bronner, M., Elgar, G., 2011. Ancient Pbx-Hox signatures define hundreds of vertebrate developmental enhancers. BMC Genomics 12**,** 637.

Pauls, S., Smith, S. F., Elgar, G., 2012. Lens development depends on a pair of highly conserved Sox21 regulatory elements. Dev. Biol. 365**,** 310-318.

Pearson, J. C., Lemons, D., McGinnis, W., 2005. Modulating Hox gene functions during animal body patterning. Nat. Rev. Genet. 6**,** 893-904.

Phelan, M. L., Featherstone, M. S., 1997. Distinct HOX N-terminal arm residues are responsible for specificity of DNA recognition by HOX monomers and HOX.PBX heterodimers. J. Biol. Chem. 272**,** 8635-9643.

Popham, D. L., Szeto, D., Keener, J., Kustu, S., 1989. Function of a bacterial activator protein that binds to transcriptional enhancers. Science 243**,** 629-635.

Popperl, H., Featherstone, M. S., 1992. An autoregulatory element of the murine Hox-4.2 gene. EMBO J. 11**,** 3673-3680.

Prince, V. E., Joly, L., Ekker, M., Ho, R. K., 1998. Zebrafish hox genes: genomic organization and modified colinear expression patterns in the trunk. Development 125**,** 407-420.

Prohaska, S. J., Stadler, P. F., 2004. The duplication of the Hox gene clusters in teleost fishes. Theory Biosci. 123**,** 89-110.

Razin, A., Riggs, A. D., 1980. DNA methylation and gene function. Science 210**,** 604-610.

Rieckhof, G. E., Casares, F., Ryoo, H. D., Abu-Shaar, M., Mann, R. S., 1997. Nuclear translocation of extradenticle requires homothorax, which encodes an extradenticle-related homeodomain protein. Cell 91**,** 171-183.

Roeder, R. G., 1991. The complexities of eukaryotic transcription initiation: regulation of preinitiation complex assembly. Trends Biochem. Sci. 16**,** 402-408.

Rombauts, S., Florquin, K., Lescot, M., Marchal, K., Rouze, P., van de Peer, Y., 2003. Computational approaches to identify promoters and cis-regulatory elements in plant genomes. Plant Physiol. 132**,** 1162-1176.

Rottkamp, C. A., Lobur, K. J., Wladyka, C. L., Lucky, A. K., O'Gorman, S., 2008. Pbx3 is required for normal locomotion and dorsal horn development. Dev. Biol. 314**,** 23-39.

Ryoo, H. D., Marty, T., Casares, F., Affolter, M., Mann, R. S., 1999. Regulation of Hox target genes by a DNA bound Homothorax/Hox/Extradenticle complex. Development 126**,** 5137-5148.

Sagerstrom, C. G., Kao, B. A., Lane, M. E., Sive, H., 2001. Isolation and characterization of posteriorly restricted genes in the zebrafish gastrula. Dev. Dyn. 220**,** 402-408.

Saleh, M., Huang, H., Green, N. C., Featherstone, M. S., 2000. A conformational change in PBX1A is necessary for its nuclear localization. Exp. Cell Res. 260**,** 105-115.

Salzberg, A., Elias, S., Nachaliel, N., Bonstein, L., Henig, C., Frank, D., 1999. A Meis family protein caudalizes neural cell fates in Xenopus. Mech. Dev. 80**,** 3-13.

Sanchez-Guardado, L. O., Ferran, J. L., Rodriguez-Gallardo, L., Puelles, L., Hidalgo-Sanchez, M., 2011. Meis gene expression patterns in the developing chicken inner ear. J. Comp. Neurol. 519**,** 125-147.

Sawadogo, M., Sentenac, A., 1990. RNA polymerase B (II) and general transcription factors. Annu. Rev. Biochem. 59**,** 711-754.

Schnabel, C. A., Jacobs, Y., Cleary, M. L., 2000. HoxA9-mediated immortalization of myeloid progenitors requires functional interactions with TALE cofactors Pbx and Meis. Oncogene 19**,** 608-616.

Shanmugam, K., Green, N. C., Rambaldi, I., Saragovi, H. U., Featherstone, M. S., 1999. PBX and MEIS as non-DNA-binding partners in trimeric complexes with HOX proteins. Mol. Cell Biol. 19**,** 7577-7588.

Sharkey, M., Graba, Y., Scott, M. P., 1997. Hox genes in evolution: protein surfaces and paralog groups. Trends Genet. 13**,** 145-151.

Shen, W. F., Rozenfeld, S., Kwong, A., Kom ves, L. G., Lawrence, H. J., Largman, C., 1999. HOXA9 forms triple complexes with PBX2 and MEIS1 in myeloid cells. Mol. Cell Biol. 19**,** 3051-3061.

Shen, W. F., Rozenfeld, S., Lawrence, H. J., Largman, C., 1997. The Abd-B-like Hox homeodomain proteins can be subdivided by the ability to form complexes with Pbx1a on a novel DNA target. J. Biol. Chem. 272**,** 8198-8206.

Shin, J. T., Fishman, M. C., 2002. From Zebrafish to human: modular medical models. Annu. Rev. Genomics Hum. Genet. 3**,** 311-340.

Smith, J. E., Jr., Bollekens, J. A., Inghirami, G., Takeshita, K., 1997. Cloning and mapping of the MEIS1 gene, the human homolog of a murine leukemogenic gene. Genomics 43**,** 99-103.

Steelman, S., Moskow, J. J., Muzynski, K., North, C., Druck, T., Montgomery, J. C., Huebner, K., Daar, I. O., Buchberg, A. M., 1997. Identification of a conserved family of Meis1-related homeobox genes. Genome Res. 7**,** 142-156.

Stone, J. R., Wray, G. A., 2001. Rapid evolution of cis-regulatory sequences via local point mutations. Mol. Biol. Evol. 18**,** 1764-1770.

Struhl, K., 1999. Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. Cell 98**,** 1-4.

Svejstrup, J. Q., Vichi, P., Egly, J. M., 1996. The multiple roles of transcription/repair factor TFIIH. Trends Biochem. Sci. 21**,** 346-350.

Sydow, J. F., Cramer, P., 2009. RNA polymerase fidelity and transcriptional proofreading. Curr. Opin. Struct. Biol. 19**,** 732-739.

Szutorisz, H., Dillon, N., Tora, L., 2005. The role of enhancers as centres for general transcription factor recruitment. Trends Biochem. Sci. 30**,** 593-599.

Takahashi, K., Liu, F. C., Oishi, T., Mori, T., Higo, N., Hayashi, M., Hirokawa, K., Takahashi, H., 2008. Expression of FOXP2 in the developing monkey forebrain: comparison with the expression of the genes FOXP1, PBX3, and MEIS2. J. Comp. Neurol. 509**,** 180-189.

Thomas, M. J., Platas, A. A., Hawley, D. K., 1998. Transcriptional fidelity and proofreading by RNA polymerase II. Cell 93**,** 627-637.

Tolhuis, B., Palstra, R. J., Splinter, E., Grosveld, F., de Laat, W., 2002. Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol. Cell. 10**,** 1453-1465.

Toresson, H., Mata de Urquiza, A., Fagerstrom, C., Perlmann, T., Campbell, K., 1999. Retinoids are produced by glia in the lateral ganglionic eminence and regulate striatal neuron differentiation. Development 126**,** 1317-1326.

Tumpel, S., Cambronero, F., Wiedemann, L. M., Krumlauf, R., 2006. Evolution of cis elements in the differential expression of two Hoxa2 coparalogous genes in pufferfish (Takifugu rubripes). Proc. Natl. Acad. Sci. U.S.A. 103**,** 5419-5424.

Udvadia, A. J., Linney, E., 2003. Windows into development: historic, current, and future perspectives on transgenic zebrafish. Dev. Biol. 256**,** 1-17.

Valverde-Garduno, V., Guyot, B., Anguita, E., Hamlett, I., Porcher, C., Vyas, P., 2004. Differences in the chromatin structure and cis-element organization of the human and mouse GATA1 loci: implications for cis-element identification. Blood 104**,** 3106-3116.

Venkatesh, B., 2003. Evolution and diversity of fish genomes. Curr. Opin. Genet. Dev. 13**,** 588-592.

Vlachakis, N., Choe, S. K., Sagerstrom, C. G., 2001. Meis3 synergizes with Pbx4 and Hoxb1b in promoting hindbrain fates in the zebrafish. Development 128**,** 1299-1312.

Vlachakis, N., Ellstrom, D. R., Sagerstrom, C. G., 2000. A novel pbx family member expressed during early zebrafish embryogenesis forms trimeric complexes with Meis3 and Hoxb1b. Dev. Dyn. 217**,** 109-119.

Wagner, G. P., Amemiya, C., Ruddle, F., 2003. Hox cluster duplications and the opportunity for evolutionary novelties. Proc. Natl. Acad. Sci. U.S.A. 100**,** 14603-14606.

Walters, M. C., Fiering, S., Eidemiller, J., Magis, W., Groudine, M., Martin, D. I., 1995. Enhancers increase the probability but not the level of gene expression. Proc. Natl. Acad. Sci. U.S.A. 92**,** 7125-7129.

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

Weber, M., Hellmann, I., Stadler, M. B., Ramos, L., Paabo, S., Rebhan, M., Schubeler, D., 2007. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat. Genet. 39**,** 457-466.

Whiting, J., Marshall, H., Cook, M., Krumlauf, R., Rigby, P. W., Stott, D., Allemann, R. K., 1991. Multiple spatially specific enhancers are required to reconstruct the pattern of Hox-2.6 gene expression. Genes Dev. 5**,** 2048-2059.

Williams, T. M., Williams, M. E., Innis, J. W., 2005. Range of HOX/TALE superclass associations and protein domain requirements for HOXA13:MEIS interaction. Dev. Biol. 277**,** 457-471.

Wolpert, L., 2007. Principles of development. Oxford University Press, Oxford, U.K.

Woolfe, A., Goodson, M., Goode, D. K., Snell, P., McEwen, G. K., Vavouri, T., Smith, S. F., North, P., Callaway, H., Kelly, K., Walter, K., Abnizova, I., Gilks, W., Edwards, Y. J., Cooke, J. E., Elgar, G., 2005. Highly conserved non-coding sequences are associated with vertebrate development. PLoS Biol. 3**,** e7.

Wray, G. A., 2007. The evolutionary significance of cis-regulatory mutations. Nat. Rev. Genet. 8**,** 206-216.

Zawel, L., Kumar, K. P., Reinberg, D., 1995. Recycling of the general transcription factors during RNA polymerase II transcription. Genes Dev. 9**,** 1479-1490.

Zerucha, T., Prince, V. E., 2001. Cloning and developmental expression of a zebrafish meis2 homeobox gene. Mech. Dev. 102**,** 247-250.

Zerucha, T., Stuhmer, T., Hatch, G., Park, B. K., Long, Q., Yu, G., Gambarotta, A., Schultz, J. R., Rubenstein, J. L., Ekker, M., 2000. A highly conserved enhancer in the Dlx5/Dlx6 intergenic region is the site of cross-regulatory interactions between Dlx genes in the embryonic forebrain. J. Neurosci. 20**,** 709-721.

Zhang, X., Friedman, A., Heaney, S., Purcell, P., Maas, R. L., 2002. Meis homeoproteins directly regulate Pax6 during vertebrate lens morphogenesis. Genes Dev. 16**,** 2097-2107.

Zhang, X., Rowan, S., Yue, Y., Heaney, S., Pan, Y., Brendolan, A., Selleri, L., Maas, R. L., 2006. Pax6 is regulated by Meis and Pbx homeoproteins during pancreatic development. Dev. Biol. 300**,** 748-757.

VITA

Cody Evan Barrett was born in Shelby, North Carolina to Joey and Lori Barrett. After graduating from Kings Mountain High School in 2006 as Valedictorian of his senior class, he began his undergraduate studies at the University of North Carolina at Chapel Hill in the fall of 2006. He received his Bachelor of Science in Biology with a minor in Chemistry from the University of North Carolina at Chapel Hill in the spring of 2010. During the fall of 2010, he began graduate studies in Dr. Ted Zerucha's lab, receiving his Master of Science in Cell and Molecular Biology in the summer of 2013. In the fall of 2013, he enrolled at the University of North Carolina at Chapel Hill's School of Dentistry, and in the future hopes to obtain his D.D.S. degree and practice dental medicine in the southeast region of the United States.