TRANSCRIPTIONAL REGULATION OF *wnt1* AND *wnt10b* BY EVOLUTIONARILY CONSERVED CIS-REGULATORY REGIONS

A Thesis

by

AVANTIKA SINGH

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Arne Lekven
Committee Members,	Gary Kunkel
	Alvin Yeh
Head of Department,	Tom McKnight

August 2017

Major Subject: Biology

ABSTRACT

This study assesses the importance of enhancers of *wnt* genes- *wnt1* and *wnt10b* during midbrain hindbrain boundary formation in zebrafish. These cis-regulatory elements are hypothesized to regulate *wnt1* and *wnt10b* transcription both spatially and temporally. A secreted signaling molecule is encoded by *wnt1* that is first expressed in the presumptive midbrain region of the neural plate and is essential for midbrain and hindbrain development. An understanding of *wnt* regulation would, therefore, shed light on transcriptional mechanisms that underlay regional organization of the vertebrate CNS.

In mice it has been shown that certain non-coding regions work as enhancer elements and trigger the full *wnt1* expression in the early murine embryo. These enhancers are proved to be both necessary and sufficient component of endogenous regulatory machinery of *wnt1* expression. The *wnt1* and *wnt10b* locus in fish shows many conserved non-coding regions. As compared to spotted gar (Lepisosteus oculatus), zebrafish underwent an additional round of whole-genome duplication but these cis-regulatory regions have still been conserved. That indicates they might be of key importance in regulating *wnt1* and *wnt10b* during brain development. There is a possibility that these regions represent an evolutionarily conserved regulatory module controlling the *wnt1* and *wnt10b* expression.

This project targets a comparative sequence analysis of non-coding regions from zebrafish and spotted gar to check their ability to activate reporter gene expression in zebrafish midbrain and midbrain-hindbrain boundary. Our goal is to understand the transcriptional regulation of *wnt* in detail. ECRs are critical in understanding the regulation of *wnt* genes around midbrain hindbrain boundary. Overall, this study helped us to better understand the role of *wnt* in anterior posterior patterning in vertebrates as wells as complex control of brain development.

ACKNOWLEDGEMENTS

First, I would like to thank my committee chair, Dr. Arne Lekven who gave me the chance to work with him. I would like to thank my committee members, Dr. Kunkel and Dr. Yeh for their guidance and support throughout the course of this research.

I would also like to thank all Lekven lab people for their help and useful suggestions: Jo-Ann for always clearing my doubts, Amy for helping with my experiments, David for taking such a beautiful pictures and teaching me how to use fluorescent microscope and finally my lab sister Saurav for your friendship, advises and for always being there for me. I would like to thank Jennifer Bradford for helping me with all the administrative responsibilities. I am grateful to my colleagues and the department faculty and staff for making my time at Texas A&M University a great experience. Thanks to all my friends, especially Sharath for his patience and love and for making this journey such a joyful experience. Finally, thanks to my mother and father for their encouragement, constant support and motivation without which I could have never done this.

CONTRIBUTORS AND FUNDING SOURCES

This work was supervised by a thesis committee consisting of Professor Dr. Arne Lekven of Department of Biology and Dr. Gary Kunkel of the Department of Biochemistry and Dr. Alvin Yeh of the Department of Biomedical Engineering.

All work for the thesis was completed by the student, under the advisement of Dr. Arne Lekven of the Department of Biology.

There are no outside funding contributions to acknowledge related to the research and compilation of this document.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I
INTRODUCTION

Neural plate patterning and the midbrain-hindbrain boundary

The neural plate is the basic structure from which the complex vertebrate brain arises. A number of early signals play important roles in diving the neural plate into distinct compartments. Later on during development, these separate compartments are further specialized morphologically and at the molecular level. This division starts from the early gastrula stage. One of the initial divisions is the separation of neural plate into broad compartments that serve as the basis for the embryonic brain. The first molecular subdivision divides the neural plate into rostral and caudal regions that set the boundary between future midbrain and hindbrain, the midbrain-hindbrain boundary (MHB) (Raible et al., 2004). The MHB is a constriction in the neural tube at the junction between the midbrain and hindbrain. This structure serves as an important signaling center for patterning the midbrain and anterior hindbrain (Rhinn and Brand, 2001; Wassef and Joyner, 1997).

The cells in this boundary have the potential to send instructive signals to their neighboring cells and tissues to direct their developmental fates. The midbrain-hindbrain organizer instructs the formation of crucial territories. On the rostral side, it patterns the tectum, whereas on the caudal side it patterns the cerebellum (Joyner et al., 2000; Rhinn et al., 2001; Wurst and Bally-Cuif, 2001).

A number of genes, including *otx, gbx, fgf8, pax2* and *wnt1* are expressed during neural pattering. The expression of these genes is stage specific. During early neural patterning, adjacent expression of Otx and Gbx class homeodomain transcription factors marks the anterior and posterior compartments of the brain, respectively (Millet et al., 1999; Broccoli et al., 1999; Katahira et al., 2000). Following this, the transcription factors and

signaling molecules necessary for MHB development are expressed in this region (Joyner et al., 2000; Rhinn and Brand 2001; Wurst and Bally-Cuif, 2001; Crossley et al., 1996; Chi et al., 2003; Ja´szai et al., 2003). After Otx and Gbx, transcription of *fgf8* and *pax2* occurs in this region. This expression of *fgf8* and *pax2* is independent of Otx and Gbx expression. This indicates that Otx2 and Gbx are not important for MHB initiation but only for the spatial refinement of MHB gene expression.

There are two main phases of gene regulation at the MHB i.e. establishment phase and maintenance phase. During the establishment phase, MHB specific genes are expressed independently of each other. In contrast, during the maintenance phase, genes expressed are interdependent and exhibit cross-regulatory activities. Wnt and Fgf are the secreted proteins expressed at the MHB junction.

Wnt genes: History and functions

The term Wnt is the combination of wingless (wg) and *int.* The First Wnt genes were cloned as proto- oncogenes from the mouse genome and are named *wnt1* and *wnt3* (Nusse and Varmus, 1982; Roelink et al., 1990), as a gene near the cystic fibrosis locus (human Writ-2) (Wainwright et al., 1988), and as human, Drosophila, and Xenopus homologs of mouse Writ-7 (Van't et al., 1984; Rijsewijk et al., 1987; Noordermeer et al., 1989). Most of the Wnt genes isolated in the past few years are from diverse species and are important because of the roles they play during development. Wnt genes are shown to induce axis duplication in frog embryos (McMahon and Moon, 1989), mammary cancer in mice, normal brain development of mammals (McMahon and Bradley, 1990; Thomas and Capecchi, 1990) as well as normal segmentation of insect embryos (Rijsewijk et al., 1987).

Wnt genes are involved in anterior-posterior axis specification in vertebrates. Several Wnt ligands, such as Wnt3a, Wnt5a, Wnt8, and Wnt11, are expressed in ventral or posterior regions of the embryo (Sokol et al., 1991; Krauss et al., 1992; Christian et al., 1993; Kelly et al., 1995). Whereas multiple Wnt antagonists like Frzb/Sfrp3, Crescent, Shisa, and Dkk1, are expressed in the head region (Leyns et al., 1997; Wang et al., 1997; Glinka et al., 1998; Shibata et al., 2005; Yamamoto et al., 2005). This complimentary expression pattern of Wnt ligands and antagonists set up the anterior-posterior axis.

Similarly, the Wnt family of secreted glycoproteins plays important roles in patterning the vertebrate nervous system at many points. During neural induction, Wnt signaling is induced in the paraxial mesoderm, providing posteriorizing signals that establishes anterior-posterior polarity (Bang et al., 1999; McGrew et al., 1995). Following neural plate polarization, further steps of vertebrate neural patterning and morphogenesis require at least three Wnt genes. Now the refinement of initial A/P domains into smaller subdivisions takes place. In mouse, the *wnt3a* locus is the major contributor in forming the hippocampus area (Bang et al., 1999; McGrew et al., 1995). Also, studies have

shown the importance of *wnt8b* in addition to *wnt3a* in dorsal telencephalon formation (Theil et al., 2002). Further, in the zebrafish, the *wnt8b* locus (Kelly et al., 1995) has been shown to be involved in establishing subdivisions within the anterior neural plate (Houart et al., 2002; Kim et al., 2002). Genetic studies in mouse have shown that *wnt1* locus is important for formation of MHB (McMahon and Bradley, 1990; Thomas and Capecchi, 1990; Thomas et al., 1991). This implies that Wnt signaling is required to pattern the early neurectoderm and also to establish a signaling center that further patterns brain subdivisions.

The Wnt genes encode secretory signaling proteins, i.e. glycoproteins that stimulate cell proliferation and differentiation. Wnt proteins can induce different biological responses in different cellular contexts, a well-known phenomenon for secretory proteins. The Wnt gene family is very large and evidence suggests that there is functional redundancy among different gene members. Wnt genes are conserved well enough to permit recognition of the homologs of individual mouse genes in different orders of vertebrates (Christian et al., 1991) and even in insects or echinoderms (Sidow, 1992).

Before the divergence of protostomes and deuterosomes, an ancient cluster of three paralogs *wnt1, wnt6* and *wnt10* existed (Nusse, 2001). It has been shown that zebrafish *wnt1* maps to linkage group 23, in close proximity to another Wnt paralog, *wnt10b* (Krauss et al., 1992a; Postlethwait et al., 1998). This indicates that the linkage of *wnt1* and *wnt10b* paralogs has been conserved in the vertebrate lineage. Since *wnt1* and *wnt10b* are conserved spatially, there is a chance that their expression is regulated in a coordinated fashion. It has been shown that the expression pattern of zebrafish *wnt10b* and *wnt1* substantially overlaps (Lekven et al., 2003). Also, *wnt1* and *wnt10b* are shown to be partially functionally redundant during the regulation of gene expression at the MHB. Unlike mouse *wnt1,* zebrafish *wnt1* and *wnt10b* are not required for midbrain and cerebellum formation, but they have notable role in maintenance of MHB and expression of several genes at the MHB (Lekven et al., 2003).

Transcriptional regulation of Wnt genes

Many studies have shown the significance of Wnt genes for the formation of the embryonic midbrain during development (McMahon and Bradley, 1990; Thomas and Cappechi, 1990). *wnt*1 transcripts are initially detected in the broad region that demarcated the presumptive midbrain. Later expression of *wnt1* transcripts is restricted to the dorsal and ventral midlines of the midbrain and caudal diencephalon and the MHB. Analysis of *wnt1* transcriptional regulation may provide clues as how the midbrain becomes specified as being different from neighboring forebrain and hindbrain. However, molecular interactions that initiate and maintain *wnt1* expression have not yet been identified. The role of cis- acting transcription factors, as the necessary components of this regulation, has been a subject of extensive analysis.

Previous studies in murine embryos have identified various regions within the *wnt1* locus that are necessary for activation and maintenance of *wnt1* transcription. These regions act as enhancers of *wnt1* (Echelard et al., 1994). More specifically, it has been shown that there are evolutionary conserved regions within these enhancers that are minimal regulatory regions for expression of *wnt1*. A 5.5kb cis- acting regulatory region that lays 3' of the *wnt1* transcription unit was shown to be sufficient to direct the expression of *wnt1* gene in mouse. Also, the sequences within this enhancer are proven to act independent of orientation, on a heterologous promoter to activate, modify and maintain expression of *wnt1* (Echelard et al., 1994). Later, studies found a 110 base pair evolutionary conserved regulatory region within this 5.5kb enhancer that governs the *Wnt1*expression pattern (Rowitch et al., 1998). Altogether, these findings demonstrate that correct temporal and spatial expression of mouse *wnt1* in the CNS depends upon a cis- regulatory region 3' of the *wnt1* locus.

Transcriptional regulation by enhancers

Studies on differential gene expression have been used to retrieve genetic information and regulatory mechanisms controlling vertebrate development. The importance of transcription can be judged by the presence of abundant genes in the genome that have a function in transcription. Around 3303 genes in the zebrafish genome encode proteins with an inferred role as transcription factors; chromatin modifying factors, or proteins involved in the general process of transcription (Armant et al., 2013). The human genome comprises of similar number of genes with a function in transcription regulation (Vaquerizas et al., 2009).

The first transcriptional enhancer was discovered in a virus, followed by the beta-globin locus of the rabbit (Banerji et al., 1983; Banerji et al., 1981). Since that time, our knowledge of mechanisms underlying transcriptional regulation has expanded significantly (Andersson et al., 2015; Kolovos et al., 2012; Ong & Corces, 2011). In past noncoding DNA was frequently referred to as junk DNA of genome (Ohno, 1972). It has become clear that noncoding DNA is not at all irrelevant garbage, but it can regulate transcription by encoding additional information required to retrieve genetic information. These cis-regulatory regions comprise enhancers, silencers, and insulators and also it includes gene proximal regulatory elements, i.e, the promoter (Ferg et al., 2014; Kolovos et al., 2012).

In general, genes have multiple enhancers that control their transcription in spatially and temporally distinct but overlapping domains of expression in the embryo. Hence these enhancers can be partially redundant (Blader et al., 2004; Blader et al., 2003; Ertzer et al., 2007). These cis- regulatory regions harbor clusters of binding sites for transcription factors forming an open chromatin domain (Levine, 2010). Open chromatin is transcriptionally active, or is poised for activation, and generally has specific histone H3 modifications (Maston et al., 2012; Ong & Corces, 2011; Shlyueva et al., 2014). The

nucleus has an inherent structure that reflects the transcriptional state of genes. In general, genes present at the periphery tend to be repressed whereas genes in the center are actively transcribed (Schneider & Grosschedl, 2007).

There are various ways in which enhancers and promoters interact. One model of this interaction is a looping model. It requires the intervening DNA to be looped out or to be organized in such a manner to permit the enhancer-promoter interaction (Bulger and Groudine, 1999; Blackwood and Kadonaga, 1998; Guelen et al., 2008). Alternative models differ from the conventional looping model only in how the interaction between promoter and enhancer is established. It can either be by facilitated diffusion or free diffusion within the nucleus or by an active scanning or tracking mechanism. In Scanning or tracking mechanism, enhancer diffuses one dimensionally along the chromatin fiber in search of a promoter (Blackwood and Kadonaga, 1998).

There are several other indirect models of interaction, including “oozing” or “linking”. In such models, a complex is nucleated at the enhancer and then polymerizes along the chromatin fiber bidirectionally until it reaches a promoter (Ptashne, 1986; Dorsett, 1999; Bulger and Groudine, 1999). In other variations of this model, RNA polymerase II or other complexes are loaded at the enhancer and then actively move along the DNA until reaching a promoter. Chromatin looping involving the mediator complex and cohesins brings together enhancers and promoters in close vicinity, after which domains of active transcription- factories are formed by the interaction of enhancers and promoters (Cook, 1999; Rieder, Trajanoski, & McNally, 2012).

Evolutionary sequence conservations can identify enhancers

Identification of the sequences that direct spatial and temporal expression of genes is a significant challenge in annotation of vertebrate of genomes. Identification of such distant acting gene regulatory sequences has been limited despite their established roles in development, phenotypic diversity and human diseases (Stathopoulos et al., 2003; Emison et al., 2005; Lettice et al., 2003). Recent advancement in computational and molecular biology has allowed the application of genome wide tools to analyze potential enhancer sequences and their function. The results emerged from these studies are considerably complex and projects enhancers as sequences carrying epigenetic information.

Gene regulatory sequences can be identified by comparative genomic-based approaches. This approach is used for the organisms that are separated by varying evolutionary distances. The sequence from the genomic interval from one organism is compared to orthologous regions of another. These studies have shown that ancient conservation, between humans and fish, and ultra conservation among mammals (sequences at least 200 base pairs in length that are 100% identical among human/mouse/rat) (Bejerano et al., 2004) may be useful indicator of sequences with gene regulatory activity.

Previous experimental results have revealed the high propensity of extremely conserved human non- coding regions to behave as transcriptional enhancers in vivo. This model supports both ancient human-fish conservation and human-rodent ultra conservation as highly effective filters to identify such functional elements. These studies suggested that under extreme constraints throughout vertebrate evolution, tissue-specific transcriptional enhancer activity is one of the predominant functions of non-coding genomic regions (Pennacchio et al., 2006)

Conservation of wnt1 regulatory elements

wnt1 homologues have been found in all vertebrates and the regulation of expression of *wnt1* in the developing CNS, specifically midbrain, by regulatory elements is proven to be conserved during the course of evolution (McMahon, 1992). Various studies have identified relevant cis- acting regulatory regions by comparing enhancer sequences from different species. Comparative vertebrate genome sequencing with Pufferfish (Fugu rubripes) identified enhancer elements for Hoxb-4 (Arparcio et al., 1995) and Hoxb-1 (Marshall et al., 1994). Pufferfish is a vertebrate having a compact genome i.e. one-eighth the size of mouse genome.

David Rowitch et. al in 1998 screened genomic regions flanking *wnt1* in the pufferfish and compared it to murine *wnt1* enhancer sequences. The study aligned the murine and pufferfish genomic regions flanking *wnt1* and identified a 200bp region located in *Fugu wnt1* with nucleotide identity of about 70 percent compared to the mouse *wnt1* enhancer. This region in pufferfish was found to be inversely oriented with respect to the mouse sequence. These sequences conserved between pufferfish and mice are most likely to be involved in the normal activation of *wnt1* in the midbrain region. However, few differences that are present between these species can be due to sequence divergence of the region and potentially be important for refinement or maintenance of the *wnt1* expression pattern.

The pufferfish and mammalian lineages were separated approximately 450 million years ago. This large evolutionary distance separating pufferfish and mammals ensures that most of the sequences have diverged, except where there are some functional importance i.e. coding and regulatory elements. Also, the human *wnt1* regulatory region shares significant identity with the murine region, indicating that mechanisms involved in regulation of *wnt1* are most likely conserved (Song et al., 1996; Zec et al., 1997).

Zebrafish and spotted gar as an effective genomic comparison pair

Some 450 million years ago, bony vertebrates separated into two major groups. Lobe-finned fishes formed one category whereas ray-finned fishes formed the other. Humans and other limbed creatures emerged along the lobe-finned line. The ray-finned line underwent an additional round of genome duplication resulting in teleost fishes including zebrafish, stickleback, tuna, salmon and almost every fish found in home aquarium. They are the most common fishes known today.

The gar, also known as spotted gar (*Lepisosteus oculatus*), split off the ray-finned fish lineage before the genomic duplication took place. As a result, gar maintained a conserved genetic makeup, including many entire chromosomes similar to those in the ancestor of bony vertebrates. Compared to zebrafish, gar does not have extra gene copies that zebrafish gained from the teleost genome duplication. Hence, comparison of human genome to gar's is less confusing, and the gar genome serves as a convenient middle point to identify potential functional conservation between mammalian and teleost genomes that cannot be identified by direct sequence comparison.

Zebrafish as a model system

Zebrafish (*Danio rerio)* is a tropical freshwater fish mainly found in rivers of South Asia. It belongs to the family of the cyprinids (Cyprinidae) in the class of ray-finned fishes (Actinopterygii) and within this class to the bony fishes (teleosts or Teleostei). It is a very popular vertebrate model organism, especially for studies focusing genetic control of embryonic development. This is mainly because zebrafish combine various experimental and embryological advantages (Fig.1). Not only are they easy to maintain and breed (Brand et al., 2002) but also their embryos are transparent making it easy for experimental manipulations, such as microinjections and cell transplantation experiments (Carpio et al., 2006). Also, transparency of the zebrafish chorion and embryo and early larval stages allows easy visualization of internal processes. These also facilitate tracking of expression of fluorescently tagged transgenes and monitoring reporter gene activity.

Embryogenesis in zebrafish takes only about 24 hours, and oraganogenesis is almost completed after day 5 of development. Their rapid development enables the observation of defined aspects of development as well as the completion of experiments generally within a few hours to days (Nusslein- Volhard. 2002). Zebrafish can spawn throughout the year in the laboratory, and a single female can lay up to 200 eggs per week, which is very high as compared to other organisms. The constant supply of large number of offspring from defined pairs makes the zebrafish excellent for genetic approaches (Nusslein- Volhard., 2002). However, among the fishes, zebrafish are known to be highly derived because they have evolved many traits that are not thought to have been present in ancestral species.

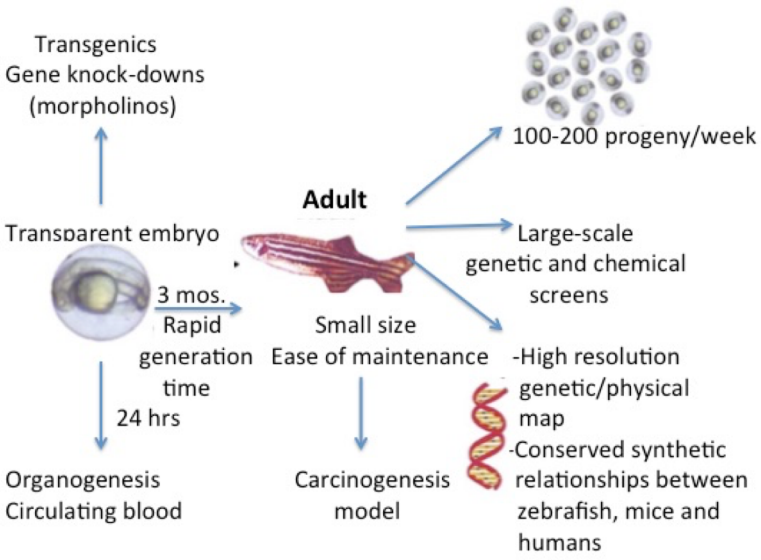

Fig 1. Zebrafish as a genetic model organism. Adapted from: Amatruda, James F., et al. "Zebrafish as a cancer model system." *Cancer cell*1.3 (2002): 229-231

Tol2 a transposable element useful for zebrafish transgenesis

Genetic approaches have been used to study the development of vertebrates for a very long time. One of the important means to study genes in vivo is transgenesis. Different methods have been developed to create transgenic zebrafish. The first approach was injecting plasmid DNA into fertilized eggs. Plasmid DNA integrates into the genome and is transmitted through the germ lineage (Stuart et al., 1988). Transgenic fish expressing green fluorescent protein (GFP) in specific tissues and organs can be created using this method. However, the frequency of germline transmission of the injected DNA is very low with this method (Higashijima et al., 1997). Another method used is injection of pseudotyped retrovirus into blastula stage embryos that generates chromosomal integration of its cDNA, and the proviral insertion is transmitted to F1 offspring (Gaiano et al., 1996). Disadvantages associated with this method are laborious process of handling and modifying retroviral vectors.

These limitations led to the development of transgenesis methods using Tol2. The vertebrate genome contains large number of DNA transposons but rarely any of these sequences have been shown to be naturally active. The Tol2 element is the exception that has been found to be autonomously active (Kawakami et al., 1999). Tol2 element was identified from genome of a small fresh water teleost-Medaka fish (*Oryzias latipes)*. This element is 4-7 kilobases in length and consists of a gene that encodes transposase protein (Kawakami et al., 1999). Transcription of this gene results in formation of mRNA that can synthesize a protein of 649 amino acids. This transposase protein is fully functional and it can catalyze the transposition of non-autonomous Tol2 construct, construct having a deletion in the transposase coding region but retaining the Tol2 ends (Kawakami et al., 1999; Kawakami et al., 2000). Through a simple cut and paste mechanism, Tol2 integrates as a single copy. It does not cause rearrangement or modification at the target site except for a creation of 8 basepair duplication.

Studies have focused on minimal Tol2 cis-sequences that are essential for transposition (Urasaki et al., 2006; Balciunas et al., 2006). Transposon constructs having 200bp and 150bp of DNA from left and right ends, respectively, is capable of transposing, whereas anything smaller than this are incapable of transposition (Urasaki et al., 2006). Thus, these 200bp and 150bp sequences containing 12bp terminal inverted repeats and sub terminal regions are enough for transposition. Any foreign DNA can be cloned between these sequences, for transgenesis.

For transgenesis using Tol2, first a transposon donor plasmid and synthetic mRNA encoding the transposase are microinjected into fertilized eggs. The transposase protein is then translated from the injected mRNA and leads to excision of transposon construct from the donor plasmid. This results in the stable integration of excised DNA into the genome. The injected mRNA and translated transposase protein gradually degrade and Tol2 insertions become stable after transposase activity ends (Kawakami, 2007). These integration events occur during the early stages of embryonic development, in fact some of them occur in the cells that produce germ cells. Germline transmission of Tol2 insertions can be selected in the F1 generation by outcrossing the injected fishes. The Tol2 method of transgenesis can produce transgenic F1 offspring from 50-70% of the injected fish, at a frequency of around 3-100%. These frequencies are significantly greater than those observed with transgenesis by plasmid microinjection (Urasaki et al., 2006; Kawakami 2004). Even though this frequency is lower than those obtain from retroviral vectors, the Tol2 transposon system is much easier to handle and manipulate when compared to retroviral vectors.

Transient expression assay in Zebrafish

When DNA is injected into fertilized zebrafish eggs, in most cases, the injected DNA is distributed unevenly and remains as extra- chromosomal DNA without getting integrated into the genome. As a result of which, embryos injected with plasmid DNA containing GFP under the control of ubiquitous promoter shows transient expression only in a small number of cells (Amsterdam et al., 1995). Whereas, when the same GFP cassette, cloned in Tol2 transposase vector, is injected with transposase mRNA in fertilized eggs, GFP is expressed ubiquitously throughout the body (Kawakami et al., 2004). This difference in expression can be explained by the fact that the transposon got integrated in the genome in many cells during early stages of embryonic development through transposition. This implies that if a cell type specific enhancer/ promoter is placed upstream of GFP on a Tol2 construct, GFP will be expressed in cell specific manner in injected F0 embryos.

Functional analysis of non-coding sequences and cis-regulatory elements in the vertebrate genome is one of the important applications of Tol2 mediated transient expression assay. No cost-effective method has been developed to conduct large-scale analysis of cis elements in regulation of genes. By using Tol2 transposon constructs containing a minimal c-fos promoter and the GFP reporter gene, the regulatory activity of genomic DNA fragments can be analyzed. Since this method can detect activities of non- coding genomic DNA regions, it has been widely used to identify discrete enhancer elements from genomic regions of different organisms.

Goal

This study focuses on determining how anteroposterior patterning is transcriptionally controlled in the vertebrate nervous system. The basic approach is to analyze the transcriptional regulation of *wnt1* and *wnt*10b genes in zebrafish. *wnt1* and *wnt10b* are adjacent genes expressed in nearly identical patterns in midbrain- hindbrain boundary, midbrain roofplate, hindbrain, and spinal cord. The goal of this project is to assess the importance of putative *wnt1/wnt10b* enhancers during midbrain hindbrain boundary formation in zebrafish. These cis-regulatory elements are hypothesized to regulate *wnt1* and *wnt10b* transcription both spatially and temporally.

Previous studies on mouse *wnt1* regulation have shown that the 3' enhancer is necessary and sufficient to recapitulate normal *wnt1* expression (Danielian et al., 1997). Also, there are studies providing evidence that within this 3' enhancer, there is a 110bp sequence that is conserved between mammals and fish (Rowitch et al., 1998). These factors along with the fact that *wnt1* cis- regulatory elements have not been studied much in other organisms made this locus an interesting area of research.

Zebrafish and gar separated around 400 millions years ago during the course of evolution and zebrafish underwent one extra round of genomic duplication along with other teleosts. Even after these events, genomic regions surrounding the *wnt*1 locus are conserved between these two fishes (Fig.2), suggesting that there has to be some functional importance of these regions.

We compared the zebrafish and garfish *wnt1* and *wnt10b* genomic regions and identified four conserved non- coding elements, which we named ECR 20, 25, 26 and 27. These ECRs were tested for their transcriptional activities. They will be further tested for their transcriptional activities in midbrain hindbrain boundary region at the time of *wnt1* and *wnt10b* expression. We complemented these with one additional potential regulatory

sequence, ECR37, identified in a zebrafish-stickleback genomic comparison

This study revolves around regulation of *wnt* genes that are significantly important in vertebrate brain patterning. The study addresses expression of *wnt1* and *wnt10b* in different model organism other than mouse. Therefore it represents an advance in our understanding of the dynamics of *wnt1/wnt10b* transcriptional regulation that would help us to better understand the role of Wnt genes in anterior posterior patterning in vertebrates as wells as complex control of brain development.

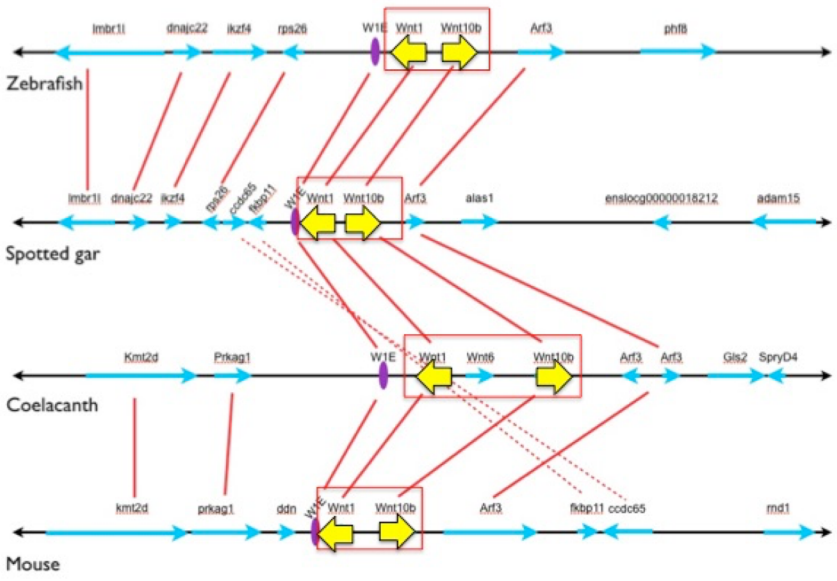

Fig 2. Conservation of spatial arrangement of *wnt1* and *wnt10b* among various organisms

CHAPTER II

ENHANCER ACTIVITY OF CIS- REGULATORY EVOLUTIONARY CONSERVED SEQUENCES

Previous studies from our lab have shown that there are number of evolutionary conserved regions (ECRs) surrounding the *wnt1* and *wnt10b* locus. The study found that these ECRs are conserved between zebrafish and stickleback genomic regions (Fig. 3). Transgenic expression assays and in situ hybridizations were done to locate the expression pattern stimulated by these ECRs in zebrafish embryos at 24hpf. The result from this study indicated that zebrafish ECR20 and ECR27 might act as shadow enhancers to regulate the transcription of *wnt1* and *wnt10b.* Shadow enhancers are the cis- regulatory regions that carry out the regulation of target genes in a similar manner (Cannavò et al., 2016). They show redundancy in their function. Since both zebrafish ECR20 and ECR27 were found to be expressed in the midbrain and MHB regions (Fig.4), it is possible that they act as shadow enhancers for regulation of *wnt1* and *wnt10b*.

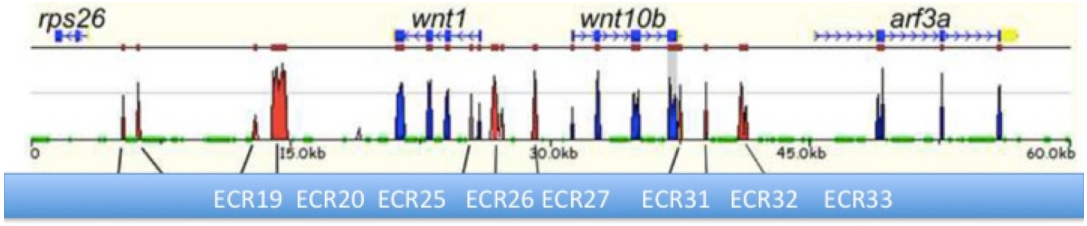

Fig 3. Evolutionarily conserved elements at the teleost *wnt1/wnt10b* locus. (A) Vista diagram of a comparison of the zebrafish and stickleback *wnt1/wnt10b* genomic intervals. Coding regions are indicated above the graph. Red peaks represent conserved non-coding elements while blue peaks are exons, which are numbered as indicated. Source: http://genome.lbl.gov/vista/index.shtml

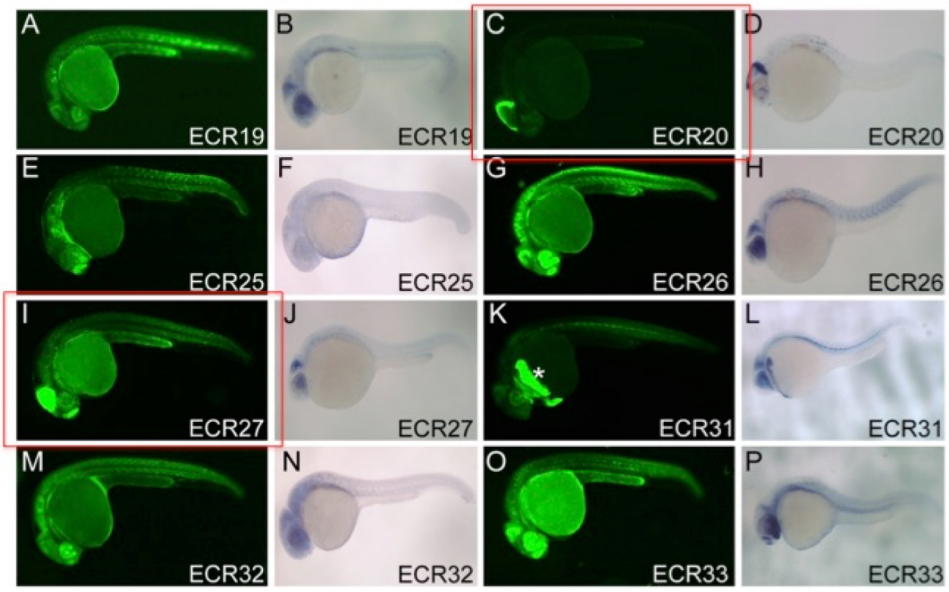

Fig 4. Evolutionarily conserved regions (ECRs) show transcriptional activity in zebrafish. All images lateral views (A, C, E, G, I, K, M, O) EGFP fluorescence. (B, D, F, H, J, L, N, P) in situ hybridizations for EGFP transcripts. EGFP expression domain of ECR20 (C) and ECR27 (I) is specifically present in MHB and midbrain domain indicating their role as shadow enhancers

This study indicated the functional importance of evolutionarily conserved cis-regulatory regions that might be controlling the expression of *wnt1* and *wnt10b* genes at the transcriptional level. We decided to test if this regulation is conserved among related fishes separated long back in evolution. For this study, we selected the cis- regulatory regions conserved between zebrafish, stickleback, and gar and checked how these ECRs regulate transcription of *wnt1* and *wnt10b* in zebrafish.

Materials and methods

Identification and cloning of evolutionarily conserved regions

There are several cis- regulatory sequences conserved between spotted gar and zebrafish, which diverged from their last common ancestor approximately 400 MYA. To identify these cis-regulatory modules that drive *wnt1* and *wnt10b* expression patterns, we compared intervals spanning this region from the zebrafish (genome version Zv9) and gar (genome assembly LepOcu1, Ensembl.org) genomes. zPicture alignment visualization tool in rVista 2.0 ECR browser (rvista.dcode.org) was used to align and compare genomic interval of these two fishes. The sequence comparison of these genomic regions identified several conserved non-coding sequences around the *wnt1/wnt10b* locus and between the flanking genes, *rps26* and *arf3a*. We selected one zebrafish ECR37 and four ECRs from the spotted gar genome (ECR20, 25, 26 and 27) for amplification and cloning.

Reporter plasmid vector construction

To determine whether these conserved non-coding elements possess transcriptional regulatory activity, we amplified the identified zebrafish and gar ECRs and ligated them upstream of a cfos minimal promoter in an EGFP reporter plasmid pT2AW2K (Fig.5). Not all the ECRs were successfully ligated directly into pT2AW2K, so these ECRs were first ligated directly into pGEM- T easy vector. Upon their successful ligation into pGEMT-easy vector, ECR fragments were isolated by restriction digestion and ligated into pT2AW2K .

pT2AW2K contains Tol2 elements for transgenesis on both the ends between which the fragment to be cloned was constructed. ECRs were cloned upstream of the minimal cfos promoter region. ECRs from zebrafish and gar were amplified by polymerase chain reaction. Restriction enzyme sites for BamHI and HindIII were included in the 5'& 3' primers respectively during amplification. Both the amplified ECRs and pT2AW2 vector were cut with BamHI and HindIII restriction enzymes and ligated together afterwards. Unfortunately, only zebrafish ECR37 showed successful ligation directly into pT2AW2K. So, rest of the ECRs was cloned into pGEM- T easy vector (Fig.6).

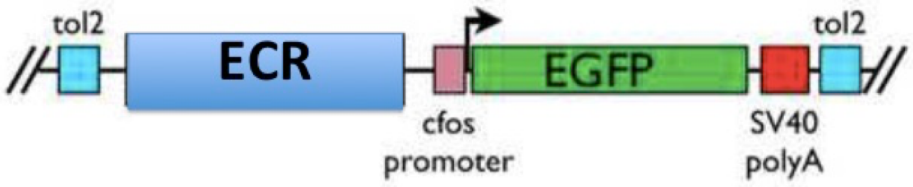

Fig 5. Diagram of the pT2AW2K reporter plasmid used for expression analysis. ECRs are cloned upstream of the minimal cfos promoter and EGFP. The plasmid contains Tol2 elements for transgenesis, SV40 polyA: poly-adenylation signal sequence

All four gar ECRs (20, 25, 26 and 27) were ligated into pGEM- T easy vector. After successful ligation, vector carrying the desired insert fragment was transformed into *E. coli* competent cells. Recombinant DNA from the bacterial colonies was amplified using mini preparation. The insertion of the fragment of interest into the vector system was verified by digesting the mini prepared DNA with EcoRI and NotI restriction enzymes. Once the insertion was confirmed, Sanger sequencing was performed to check the sequence of ECRs into the plasmid vector.

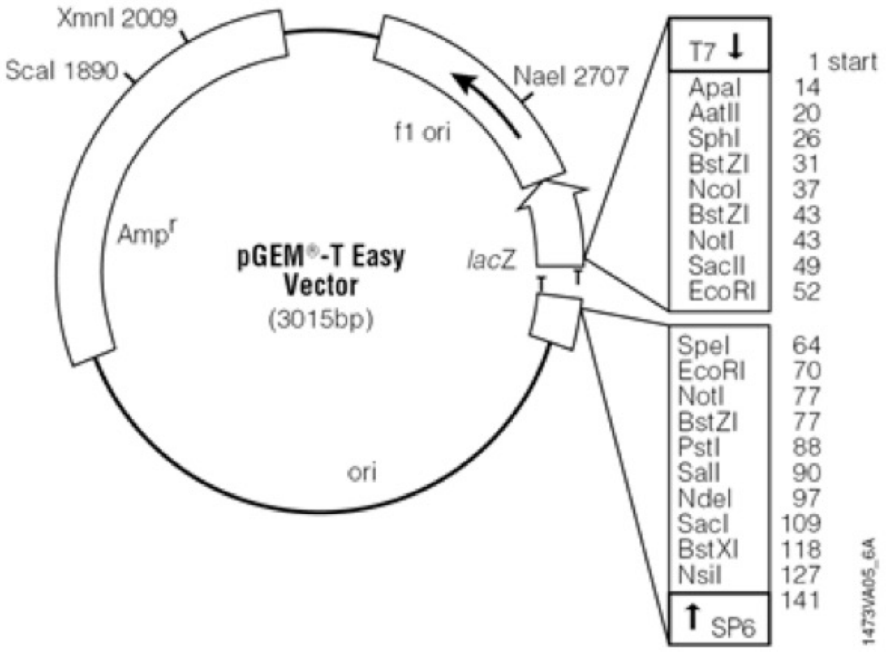

Fig 6. pGEMT Easy Vector Construct

After construction of vectors containing the gar ECRs, my next step was to clone the ECR fragments into pT2AW2K vector. So, the pGEM-T easy gar ECRs was digested with BamHI and HindIII restriction enzymes in order to release them free from the

vector. Afterwards the insert fragments were isolated and purified from the gel. Isolated inserts were ligated into pT2AW2K that was already digested with BamHI and HindIII. All pT2AW2 vectors carrying zebrafish and gar ECRs were transformed into *E.coli* competent cells. Recombinant DNA from the bacterial colonies was amplified by mini and maxed out mini preparations. Insert fragments were again verified by digesting the mini prepared DNA with BamHI and HindIII restriction enzymes.

Transient expression assay

pT2AW2 zebrafish ECR 37 and pT2AW2 gar ECR 20, 25, 26 and 27 reporter plasmids along with synthetic mRNA encoding the transposase were microinjected into wild type fertilized zebrafish eggs at one cell stage. The transposase protein is then translated from the injected mRNA and leads to excision of transposon construct from the donor plasmid. This results in the stable integration of excised DNA into the genome. Zebrafish were maintained according to standard procedures (Westerfield, 2000).

Transgenic lines were generated by co-injection of 1-3nl of transgene plasmids (25 ng/µl) with Tol2 transposase mRNA (25 ng/µl) into wild-type 1-cell stage embryos. For transient expression analysis, injected embryos were examined by fluorescence microscopy.

Stable expression assay

To determine more accurately the activity pattern of the ECRs, we attempted to generate stable transgenic lines for three ECR constructs (ZebECR37, GarECR20 and GarECR27). Transient embryos were raised to adulthood, and then outcrossed to wild type. Progeny were scored for EGFP fluorescence at 24hpf. Embryos showing transgenic expression were raised to establish lines.

Results and discussion

Amplification of gar ECR20, 25, 26, 26 and 27 and zebrafish ECR37

To analyze transcriptional regulation activity of cis- regulatory evolutionarily conserved sequences, one zebrafish ECR37 and four gar ECRs (20,25,26 and 27) (Fig.7) were amplified by PCR (Fig.8), using primers as shown in Table 1. After amplification, all ECRs were purified. Note that gar ECR20 is not indicated in Figure 7, because the sequence contig used for this analysis had a gap for that sequence interval. Gar ECR20 was identified in an unattached contig, and amplification verified the authenticity of this sequence.

Fish	ECR	Size	Forward Primer	Reverse Primer
Zeb	ECR37	437 bp	attggatccTTATCAATCCTCACACTAATGTCC	attaagcttTGCAGCTCTTGGTCTTGGAG
Gar	ECR20	857 bp	attggatccCACTCAGCCCATTCACAGGC	attaagcttGTGCAGCCGTGTCCTGAG
			atatggatccGCCCATTCACAGGCGACAG	atataagcttGCCGTGTCCTGAGCCTGG
Gar	ECR25	668 bp	attggatccCGCCCACATTTAGTGCGGAG	attaagcttAGGGAGTCCCTCGGCTATC
Gar	ECR26	384 bp	attggatccAAGTGTCCACGGCGCGAG	attaagcttCCGCTTTCCTGGTTCTTTCC
Gar	ECR27	550 bp	attggatccGGTGGTGCAGATGACGCAC	attaagcttAGAGAAGCGAGAAGACACGC
Gar	ECR33	1058 bp	attggatccCTGGAGGGCATGTCACCTG	attaagcttGCTCTTTGCCCTCCATCGC
				attaagcttATGCAGGGCTGTGTGTGTGG

Table 1. Primers for gar ECRs and zebrafish ECR37

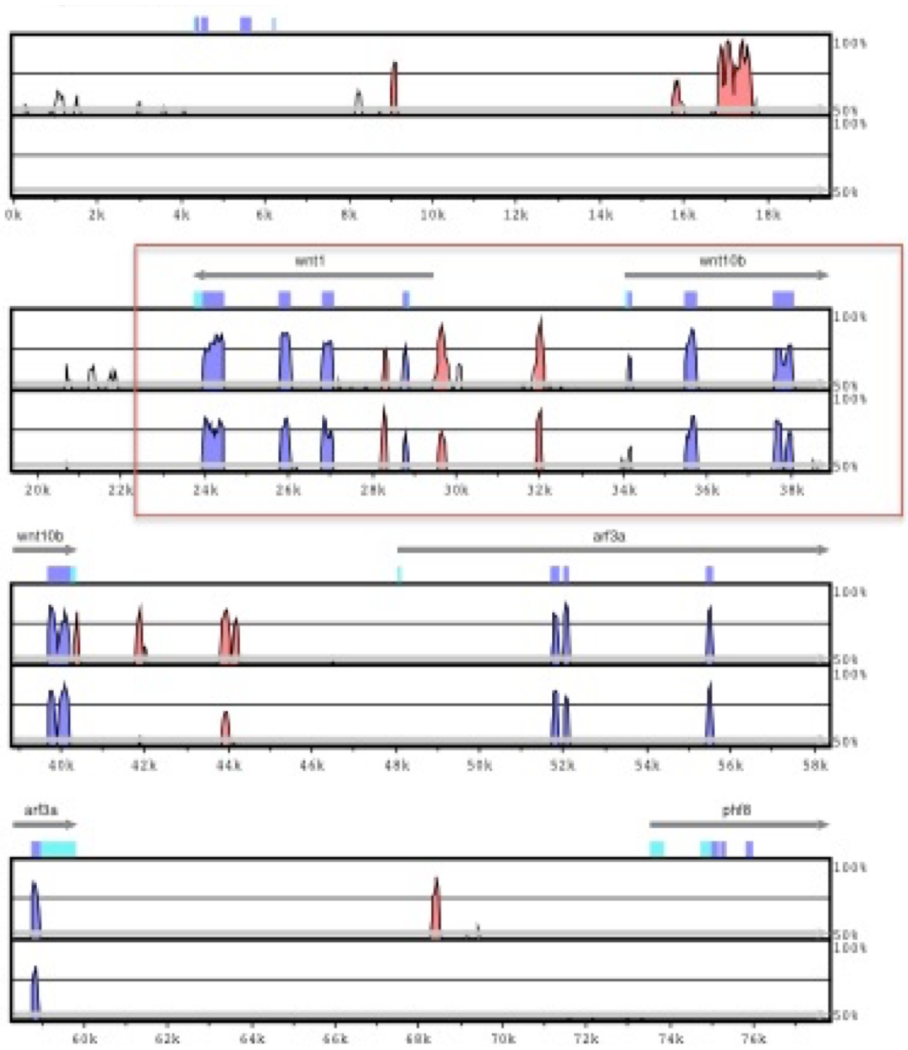

Fig 7. Evolutionarily conserved elements at the teleost *wnt1/wnt10b* locus

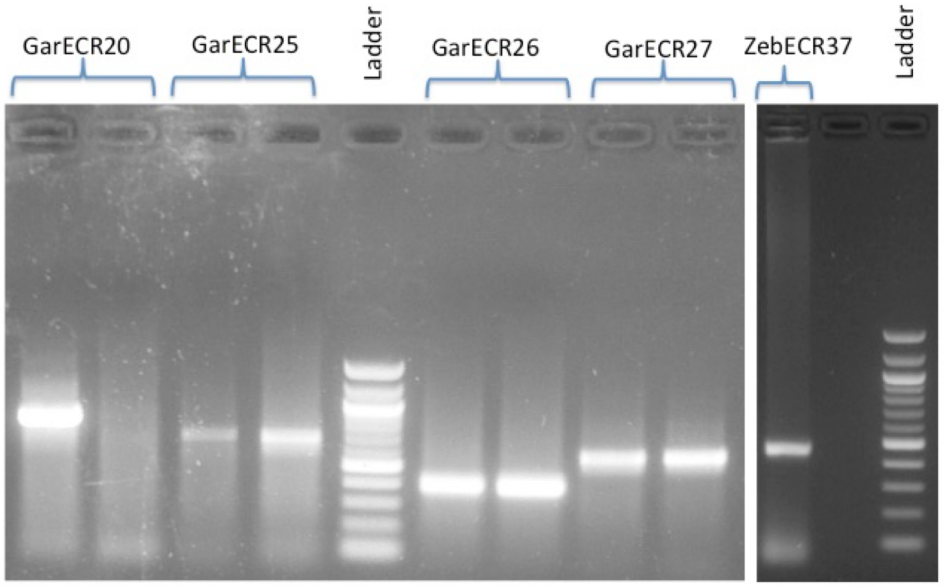

Fig 8. PCR amplification of zebrafish evolutionarily conserved regions. A. Gar ECR20, 25, 26 and 27. B. Zebrafish ECR37. Ladder: NEB 100 bp ladder

Cloning of zebrafish ECR37 into pT2AW2K

All the purified amplified ECRs were then cloned into pT2AW2K reporter plasmid construct. Surprisingly, only zebrafish ECR37 was successfully ligated into pT2AW2K vector whereas rest of the ECRs failed to show successful ligation with pT2AW2K vector (Fig.9).

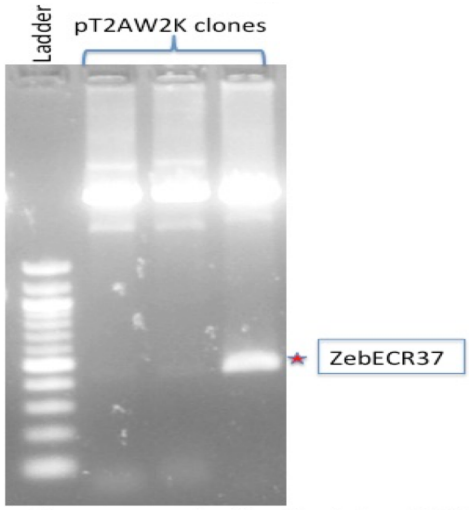

Fig 9. Ligation of ECRs into pTW2AW2K vector

Cloning of gar ECRs into pGEMT- easy and pT2AW2K

Gar ECRs were then successfully cloned into pGEMT- easy vector (Fig.10). Upon successful cloning, pGEMT- easy gar ECRs were released from the pGemT-easy clones by digestion with BamHI and HindIII restriction enzymes, and then ECR inserts were successfully ligated back to pTW2AW2K plasmid (Fig.11).

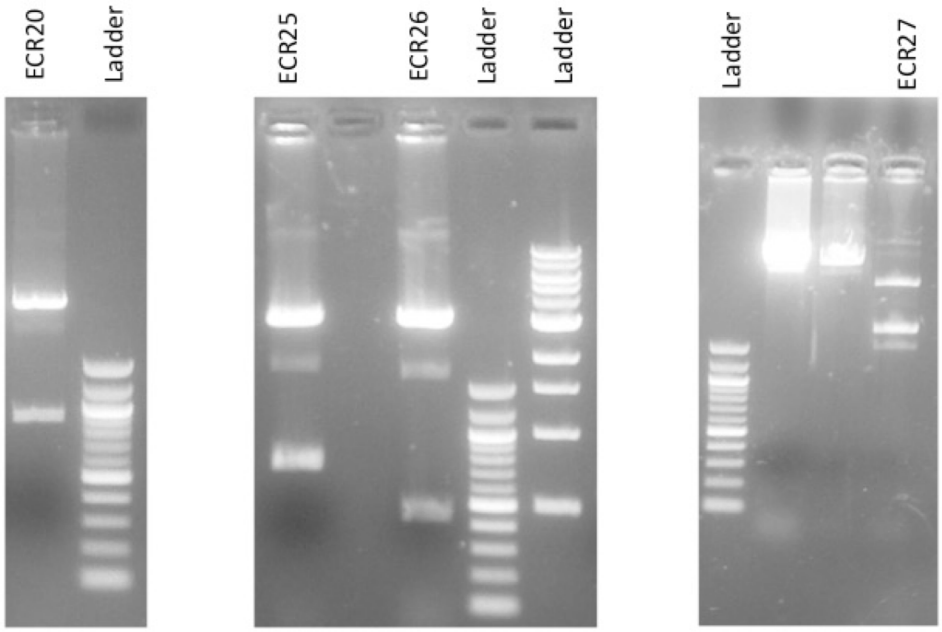

Fig 10. Ligation of gar ECRs into pGEMT- easy vector**.**

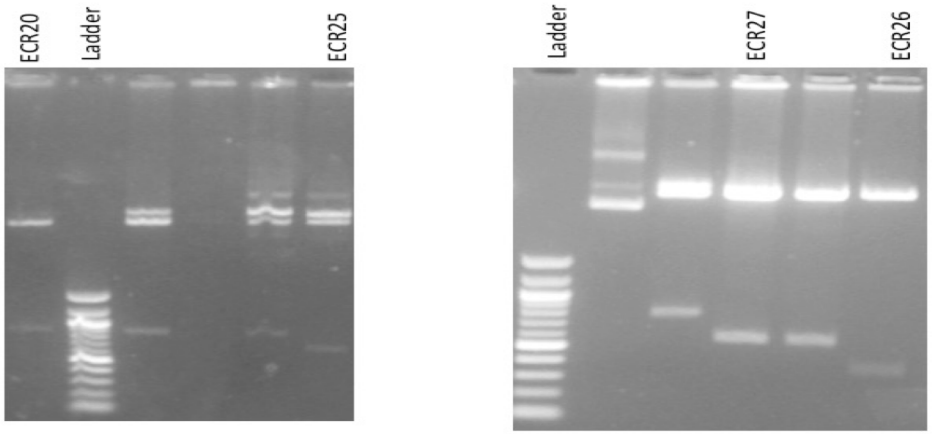

Fig 11. Ligation of gar ECRs into pTW2AW2K vector**.**

ECR transient expression analysis

To test ECR transcriptional activity, we initially performed transient expression assays. As shown in Table 2, only three ECRs, one zebrafish ECR37 and two gar ECRs (20 and 27) element drove expression that could be observed at 24 hours post fertilization (hpf), albeit expression levels varied significantly between elements. The other two gar ECRs (25 and 26) failed to drive any observable expression pattern at 24hpf.

Constructs	No. Injected fish	No. GFP +	% GFP +	Expression Domain
pZebECR37	160	70	43.75	Broad neural expression
pGarECR20	110	38	34.54	MHB
pGarECR25	120	0	0	Nil
pGarECR26	155	0	0	Nil
pGarECR27	189	55	29.10	Midbrain

Table 2. Transgenic expression assay for pT2AW2K ECR constructs

In general all three elements drove predominantly broad neural expression, but some were observed to drive higher levels in restricted domains (Fig.12). For example, zebrafish ECR37 driven expression was concentrated in midbrain, midbrain- hindbrain

boundary and faded down the tail bud. Gar ECR20 and 27 driven expression was similar to that of zebrafish ECR37 except that gar ECR27 expression was broader and intense in midbrain region.

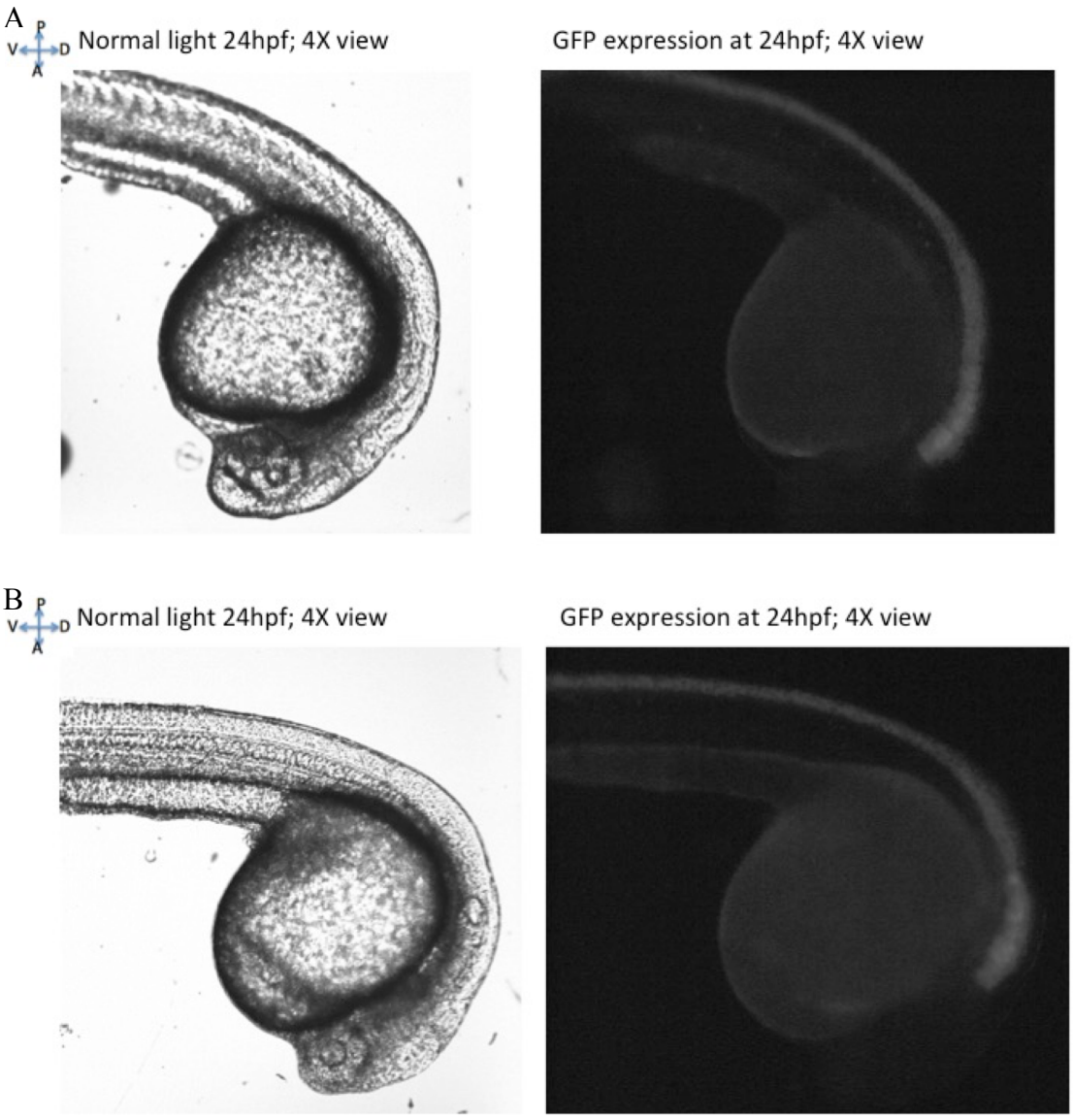

Fig 12. Transgenic expression of ECRs. A. Zebrafish ECR37 B. Gar ECR27

ECR stable expression analysis

To determine more accurately the activity patterns of the CNE elements, we generated stable transgenic lines for these three constructs. Stable integration of ECRs into the zebrafish genome was analyzed by back crossing the previously injected fish (G0, aka "founder fish") to wild type, and screening the progenies (F1) at 24hpf for EGFP fluorescence (Table 3). Surprisingly, only one founder fish carrying gar ECR20 drove expression at 24hpf (Fig.13). In contrast to the general transient expression pattern observed for this element, stable expression pattern was found to be more specific. This expression was predominantly in midbrain and showed a sharp pattern lining the midbrain- hindbrain boundary. This gar ECR20 transgenic is being raised to generate a stable line.

Founderfish#	**Number of F1 fish screened for stable expression for each construct:**		
	ZebECR37	**GarECR27**	**GarECR20**
1	180	100	10
2	100	20	65
3	20	20	30
4	10	100	50
5	20	20	65
6	10	40	50
7	75	60	110
8	20	75	130
9	10	100	**80** Only one Gar ECR20 founder fish showed stable expression pattern
10	10	20	150

Table 3. Stable expression assay for pT2AW2K ECR constructs

These three ECRs drove broad neural expression pattern specifically in midbrain and midbrain- hindbrain boundary region at 24hpf. These observations suggest that these ECRs play some important functional role during the formation of midbrain and midbrain-hindbrain boundary. Also, as Wnt genes have important role in brain patterning esp. midbrain, midbrain- hindbrain boundary formation, zebrafish ECR37 and gar ECR20 and 27 might play a role in regulating Wnt gene expression around MHB region in zebrafish. Since gar ECR25 and 26 failed to show any transient expression, their role in regulation of transcription is still ambiguous. Out of the three ECRs, only gar ECR20 was able to integrate in the zebrafish genome and show stable expression pattern, which indicates its conserved function in regulation. Also, since ECR20 and ECR27 showed same expression pattern i.e. in midbrain and MHB region, it is suggested that they might act as shadow enhancers to carry out transcriptional regulation *wnt1* and *wnt10b* during zebrafish midbrain and MHB formation.

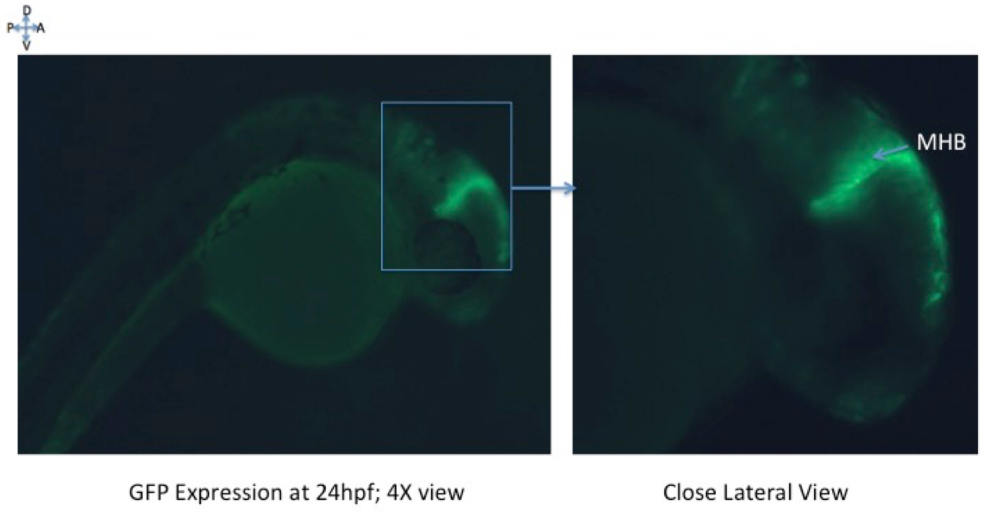

Fig 13. Stable expression of GFP for pT2AW2K ECR20 construct

Summary

The Wnt paralogs *wnt1* and *wnt10b* are closely linked on linkage group 23 and this linkage has been conserved in all vertebrate lineages. The evolutionarily conserved spatial arrangement raises a possibility that the expression of these genes is regulated in a coordinated manner as shown for *vox* and *vent* loci in Xenopus (Rastegar et al., 1999). Previous studies on mouse (Echelard et al., 1994) have shown that cis- regulatory evolutionarily conserved regions regulate expression of *wnt1*. Also, recent studies comparing stickleback and zebrafish have found evolutionarily conserved cis- regulatory regions that might be potential enhancers for *wnt1* and *wnt10b* during neural patterning, especially with regards to MHB establishment. This study focused on ECRs from gar and zebrafish to determine their enhancer activity.

Out of the five ECRs studied, only one gar ECR20 was able to show a stable expression pattern in zebrafish neural region. This expression driven by gar ECR20 was specifically present in MHB and midbrain region, which was exactly similar to expression shown by zebrafish ECR20 in previous study. Sharp expression pattern in MHB is an indicator that ECRs have significant role in MHB specification and since this expression is similar to thae one driven by zebrafish ECR20 in previous study, it can be concluded that potential enhancers might have conserved function of regulation

Unfortunately, other ECRs were not recovered in stable transgenic lines, so their role in regulation is still ambiguous. Also, previous studies comparing zebrafish and stickleback ECRs have found ECR20 and ECR27 to be expressed in similar expression pattern indicating to a possibility that they are shadow enhancers. It would be interesting to test if gar ECR27 drives an expression pattern similar to that of gar ECR20 in zebrafish, to confirm their activity as shadow enhancers.

CHAPTER III

INTERACTION OF ECR20 AND ECR27 WITH *wnt1* AND *wnt10b* PROMOTERS

A 5kb inter- genic region separates *wnt1* and *wnt10b* that are located together on linkage group 23. A number of studies have shown that there are many cis- regulatory enhancers present in this 5kb inter- genic fragment as well as around the *wnt1* and *wnt10b* genomic region (Fig.14), responsible for regulating transcriptional activity of *wnt1* and *wnt10b. wnt1* and *wnt10b* show redundant expression pattern in anterior epiphysis, cerebellum and extend to posterior rhombomere (Lekven at al., 2003). These findings suggest that these two genes are under the control of same enhancer elements. If true, then the relevant enhancers should be equally functional in combination with either gene promoter. But how these enhancers interact with *wnt1* and *wnt10b* promoters to regulate their transcription is still unknown. As the regulation of *wnt1* and *wnt10b* is unknown, it is unclear whether these neighboring loci are controlled by separate transcriptional regulatory mechanisms or by a single mechanism such as a locus control region or a shared upstream regulatory element. We determined the interaction mechanism of zebrafish *wnt1/wnt10b* enhancers (ECR 20 and 27) with *wnt1* and *wnt10b* promoter regions.

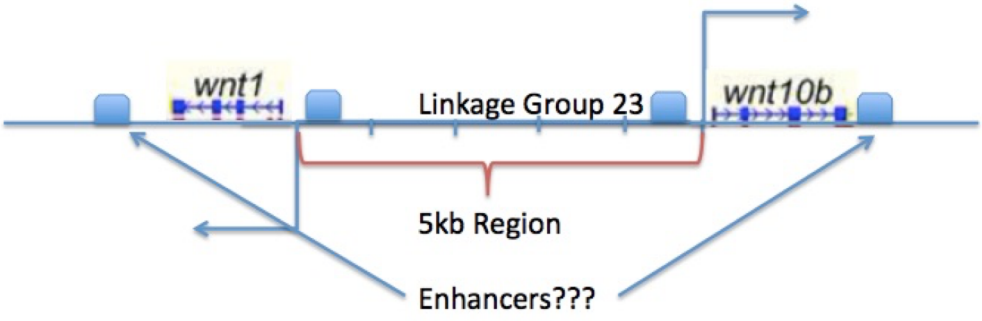

Fig 14. *wnt1* and *wnt10b* paralogs.

Materials and methods

Generation of wnt1 and wnt10b reporter promoter constructs

We tested the 5kb inter- genic fragments by constructing dual-reporter plasmids having ECR20 and ECR27 along with RFP and GFP reporting *wnt10b* and *wnt1*, respectively. pALW9 plasmid construct was designed to have ECR27 region flanked with *wnt10b* and *wnt1* separated by 5kb intergenic region. Further, GFP and RFP were ligated to be expressed from the *wnt1* and *wnt10b* translation start sites, respectively. pALW10 plasmid was constructed differently from pALW9 plasmid by ligating ECR20 further downstream of *wnt1*.Whereas in pALW11 plasmid construct ECR20 was ligated further downstream of *wnt10b* along with RFP at the distal end (Fig.15).

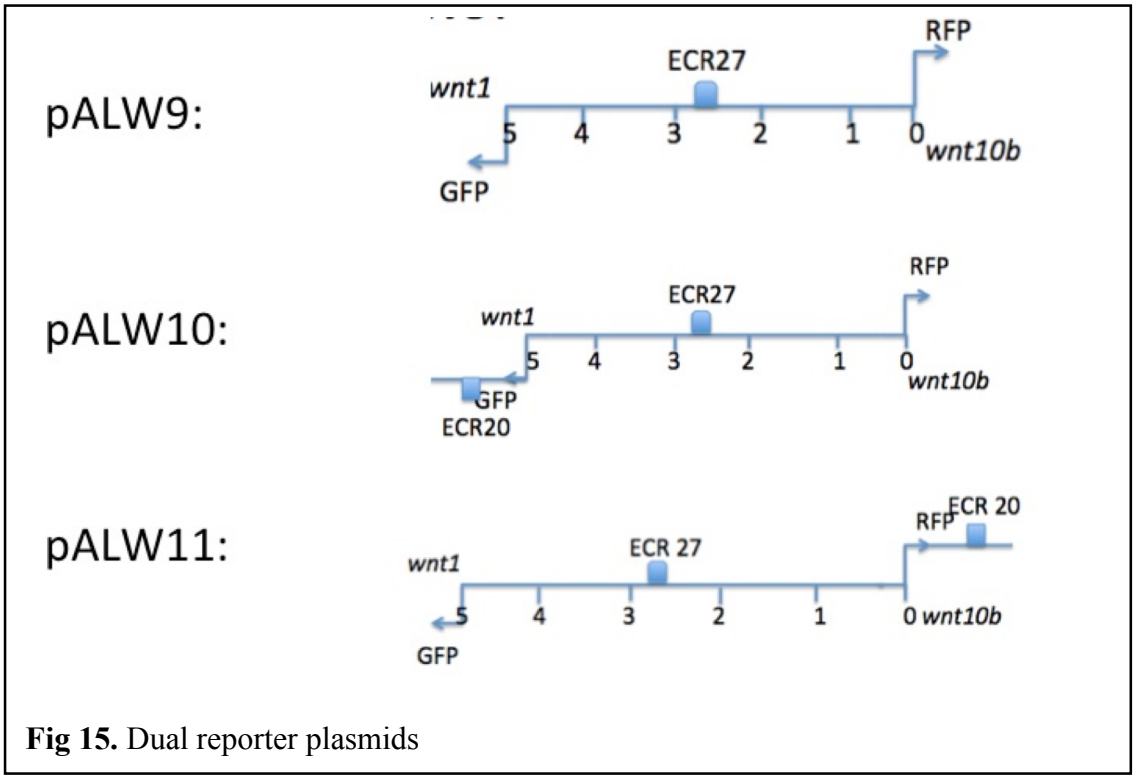

Fig 15. Dual reporter plasmids

Different subregions of *wnt10b* promoter were tested in combination with ECR20 and ECR27, a 1kb proximal sequence, and a 2kb proximal sequence, to determine whether differences in the *wnt10b* promoter exist (Fig.16). To determine whether there are functional interactions between zebrafish *wnt1/wnt10b* enhancers (ECR20 and 27) and the *wnt*1 and *wnt10b* promoters, we generated nine different plasmid constructs having Tol2-based plasmids with the ECR20 and ECR27 placed upstream of these 2kb and 1kb region of *wnt10b* (Fig.17). pALW15 served as control construct with just 2kb *wnt10b* region in Tol2 plasmid. Whereas pALW5 and pALW16 had zebrafish ECR27 and ECR20 respectively, ligated upstream of 2kb region of *wnt10b* region. Similarly, pALW6 served as control construct for 1kb *wnt10b* region. Whereas pALW14 and pALW4 had zebrafish ECR27 and ECR20, respectively, ligated upstream of 1kb region of *wnt10b* region.

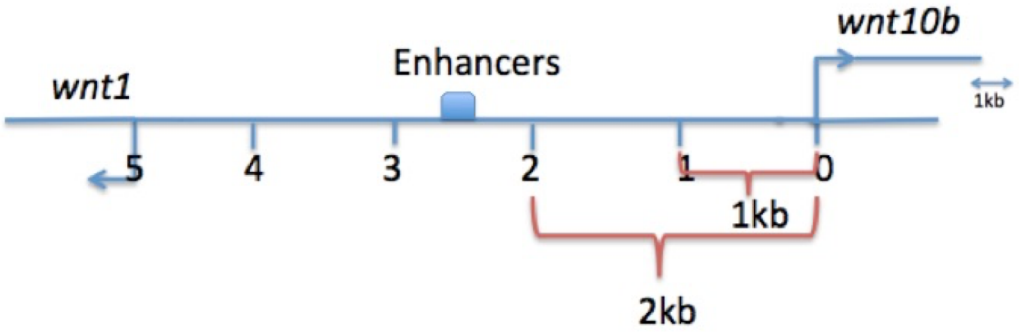

Fig 16. 1kb and 2kb of 5kb inter- genic region fragment tested are highlighted in red.

Plasmid DNA was co-injected with transposase mRNA, and transient expression was scored at 24hpf.

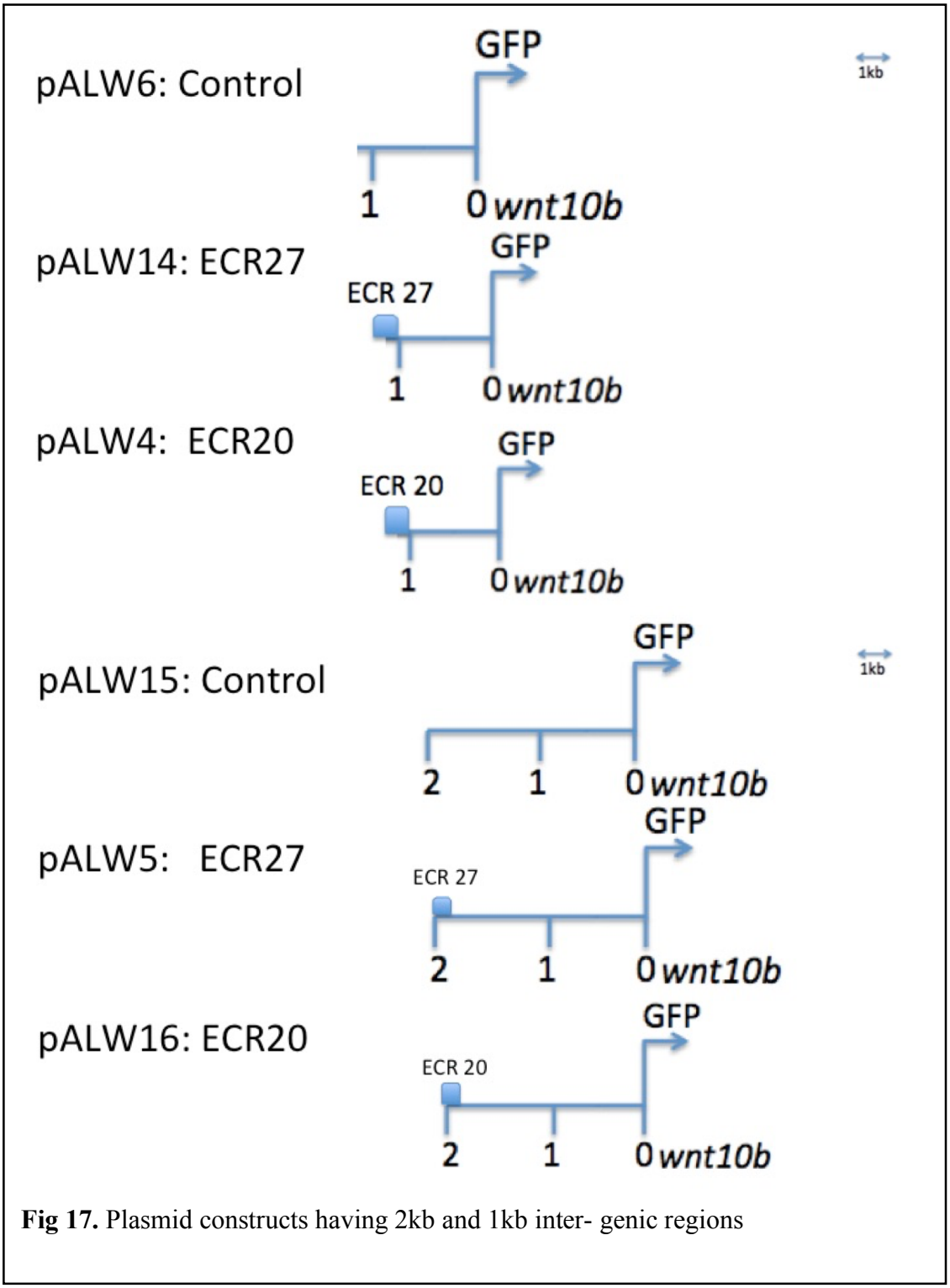

Fig 17. Plasmid constructs having 2kb and 1kb inter- genic regions

Results and discussion

To determine the interaction mechanism of ECR20 and ECR27 with *wnt*1 and *wnt10b* promoters, three dual- reporter plasmid constructs carrying zebrafish ECR20 and ECR27 interacting with *wnt1* and *wnt10b* in different orientations were created. In all these three constructs, the *wnt1* promoter and GFP flanked ECR27 on one side and the *wnt10b* promoter and RFP on the other. We performed transient expression assays to determine the mechanism of interaction. As shown in Table 4, all the plasmid constructs drove varying levels of neural expression at 24hpf. The pALW9 plasmid construct having the basic arrangement drove GFP expression utilizing *wnt1* promoter region but failed to drive any RFP expression (Fig. 18). Similarly, for other two constructs, pALW10 and 11 having additional ECR20 ligated near *wnt1* and *wnt10b* end respectively, no RFP expression was observed at 24hpf (Fig.18). Since all three constructs drove GFP expression utilizing *wnt1* promoter but failed to express RFP using *wnt10b* region indicates these enhancer elements are not functionally interacting with *wnt10b* promoter, or sequence elements may somehow suppress the enhancer interactions with the *wnt10b* promoter.

To further analyze the interaction of *wnt10b* promoter with zebrafish ECR20 and ECR27, we constructed reporter plasmids having two different fragments of the *wnt10b* region: a 1kb proximal sequence, and a 2kb proximal sequence along with ECR20 and ECR27. This helped us to determine if there are differences in the *wnt10b* promoter region. Transient expression driven by pALW5 and pALW16 constructs having 2kb fragment of *wnt10b* region along with ECR27 and ECR20, respectively, was very low or null, indicating that 2kb fragment is not functionally interacting or suppressing the enhancer activity. Whereas pALW14 and pALW4 which have the 1kb *wnt10b* region along with ECR27 and ECR20, respectively, drove bright and intense expression pattern in neural tissue (Fig.19). These results suggest that both the 1kb and 2kb fragment of *wnt10b* region can functionally interact with the enhancers.

Constructs	No. Injected fish	No. GFP+	%GFP+	Expression domain
pALW15	80	46	57.5	Broad neural expression
pALW5	60	32	53.3	Broad neural expression
pALW16	70	0	0	Nil
pALW6	65	40	61.5	Midbrain region
pALW14	78	44	56.4	Midbrain region
pALW4	60	28	46.6	Midbrain region
pALW9	75	52	69.3	Low MHB
pALW10	80	48	60	MHB
pALW11	75	40	53.3	MHB

Table 4. Transgenic expression assay for pALW constructs.

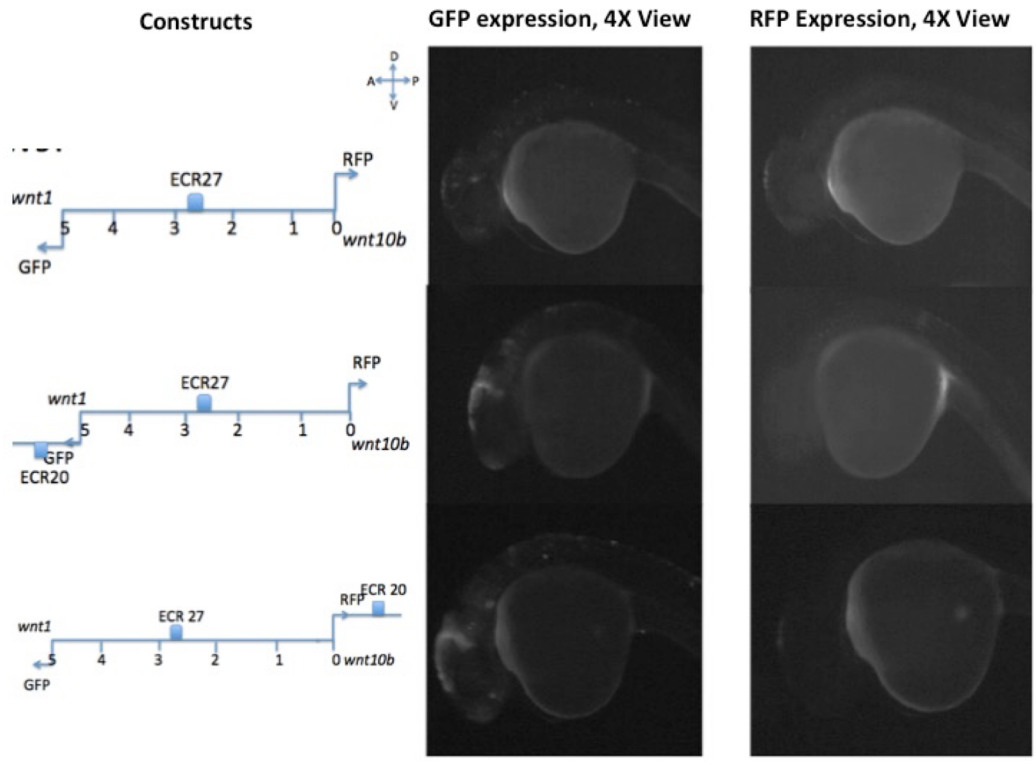

Fig 18. Transgenic expression of GFP and RFP for dual reporter plasmid

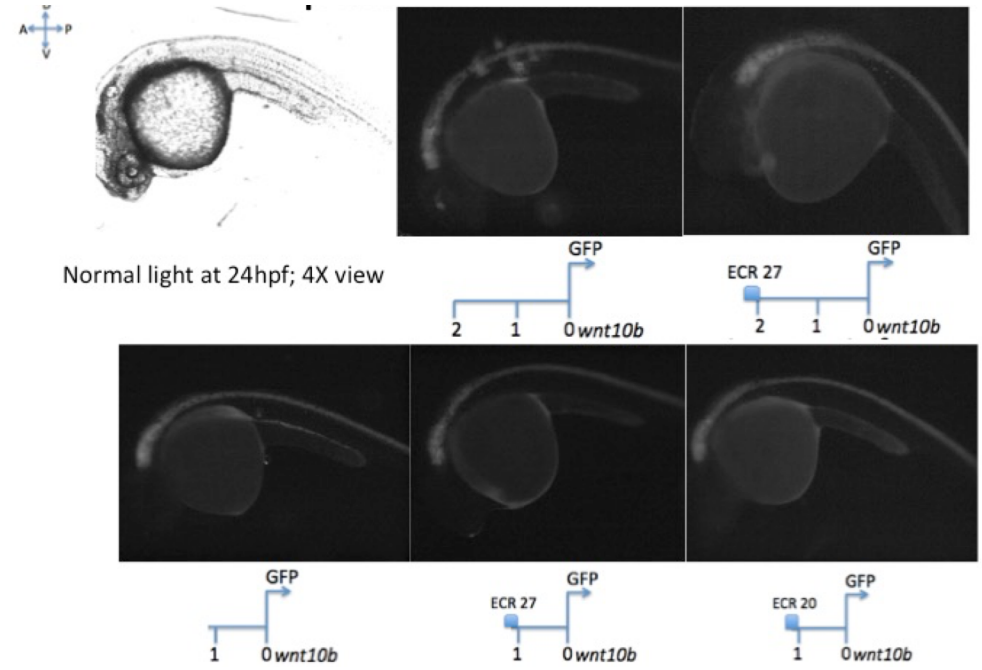

Fig 19. Transgenic expression of GFP for pALW construct

Summary

The conservation of the spatial arrangement of *wnt1* and *wnt10b* paralogs has also conserved a 5kb inter-genic fragment between them. One of the potential enhancers of *wnt1,* ECR27, is located in this fragment. The interaction mechanism of these evolutionarily conserved potential enhancer elements with *wnt1* and *wnt10b* to regulate their expression is still unknown. We tested 1kb fragment and 2kb fragments from the *wnt1* and *wnt10b* inter- genic region to determine how ECRs, specifically zebrafish ECR20 and ECR27 interact with these *wnt* genes.

Dual- plasmid constructs having GFP near *wnt1 and* RFP near *wnt10b* region failed to show any RFP expression pattern but drove a very sharp GFP expression pattern in the MHB. These observations lead to the conclusion that there is something that is suppressing the interaction of enhancers with *wnt10b* promoter. Transgenic expression from the plasmid constructs cloned with different combination of 1kb and 2kb fragments showed that the ECRs were interacting with both 1kb and 2kb inter- genic fragments. So something other than this 1kb and 2kb fragment is responsible of suppression of interaction.

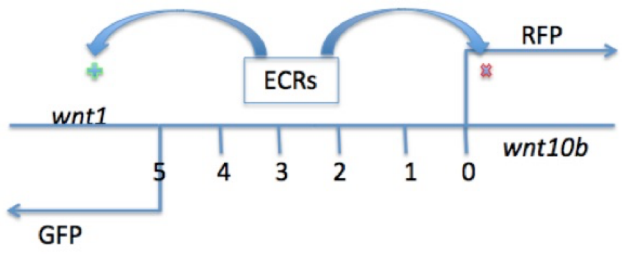

Fig 20. Proposed model for interaction of ECRs with *wnt1* and *wnt10b*

CHAPTER IV
CONCLUSION

This thesis has addressed the transcriptional regulation of the linked wnt1 and wnt10b genes in fish. These two genes are essential for development of the vertebrate midbrain-hindbrain domain, but little is understood about their transcriptional regulation. In this thesis, I provide evidence that two enhancers function as "shadow enhancers" to regulate these two genes in fish, and I provide data that furthers our understanding of how these different enhancers might interact differentially with either wnt1 or wnt10b promoters.

Future work will be required to substantiate and further this research project. For example, while I recovered one transgenic line comprising the spotted gar ECR20 element driving EGFP, we will require additional effort to recover stable transgenic lines for the other gar and zebrafish elements and constructs. This is required because transient expression assays are difficult to interpret due to variable injection dosages and variable expression from extra-chromosomal DNA. After obtaining stable transgenic lines for each, these can be assayed by in situ hybridization and fluorescence microscopy to determine whether our hypotheses are supported. If so, we anticipate that the gar ECR20 and ECR27 enhancer elements will both drive initiation of expression in zebrafish similar to endogenous wnt1, but ECR27 will not be expressed in the MHB.

Further studies will reveal the molecular nature of their regulation, for example whether Otx2 is a direct upstream regulator for either element. We suspect that alternative mechanisms are responsible for driving early wnt1 and wnt10b expression, as no mutations were ever recovered from the large scale Tubingen and Boston mutagenesis screens.

REFERENCES

Amsterdam, Adam, Shuo Lin, and Nancy Hopkins. "The Aequorea victoria green fluorescent protein can be used as a reporter in live zebrafish embryos." *Developmental biology* 171.1 (1995): 123-129.

Andersson, Robin, Albin Sandelin, and Charles G. Danko. "A unified architecture of transcriptional regulatory elements." *Trends in Genetics* 31.8 (2015): 426-433.

Aparicio, Samuel, et al. "Detecting conserved regulatory elements with the model genome of the Japanese puffer fish, Fugu rubripes." *Proceedings of the National Academy of Sciences* 92.5 (1995): 1684-1688.

Armant, Olivier, et al. "Genome-wide, whole mount in situ analysis of transcriptional regulators in zebrafish embryos." *Developmental biology* 380.2 (2013): 351-362.

Balciunas, Darius, et al. "Harnessing a high cargo-capacity transposon for genetic applications in vertebrates." *PLoS genet* 2.11 (2006): e169.

Banerji, Julian, Laura Olson, and Walter Schaffner. "A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes." *Cell* 33.3 (1983): 729-740.

Banerji, Julian, Sandro Rusconi, and Walter Schaffner. "Expression of a β-globin gene is enhanced by remote SV40 DNA sequences." *Cell* 27.2 (1981): 299-308.

Bang, Anne G., et al. "Expression of Pax-3 in the lateral neural plate is dependent on a Wnt-mediated signal from posterior nonaxial mesoderm." *Developmental biology* 212.2 (1999): 366-380.

Bejerano, Gill, et al. "Ultraconserved elements in the human genome." *Science* 304.5675 (2004): 1321-1325.

Blackwood, Elizabeth M., and James T. Kadonaga. "Going the distance: a current view of enhancer action." *Science* 281.5373 (1998): 60-63.

Blader, Patrick, Charles Plessy, and Uwe Strähle. "Multiple regulatory elements with spatially and temporally distinct activities control neurogenin1 expression in primary neurons of the zebrafish embryo." *Mechanisms of development* 120.2 (2003): 211-218.

Blader, Patrick, et al. "Conserved and acquired features of neurogenin1 regulation." *Development* 131.22 (2004): 5627-5637.

Brand, Michael, Michael Granato, and Christiane Nüsslein-Volhard. "Keeping and raising zebrafish." *Zebrafish* 261 (2002): 7-37.

Broccoli, Vania, Edoardo Boncinelli, and Wolfgang Wurst. "The caudal limit of Otx2 expression positions the isthmic organizer." *Nature* 401.6749 (1999): 164-168.

Bulger, Michael, and Mark Groudine. "Looping versus linking: toward a model for long-distance gene activation." *Genes & development* 13.19 (1999): 2465-2477.

Bulger, Michael, and Mark Groudine. "Looping versus linking: toward a model for long-distance gene activation." *Genes & development* 13.19 (1999): 2465-2477.

Cannavò, Enrico, et al. "Shadow enhancers are pervasive features of developmental regulatory networks." *Current Biology* 26.1 (2016): 38-51.

Carpio, Yamila, and Mario P. Estrada. "Zebrafish as a Genetic Model organism." *Biotecnología Aplicada* 23.4 (2006): 265-270.

Chi, Candace L., et al. "The isthmic organizer signal FGF8 is required for cell survival

in the prospective midbrain and cerebellum." *Development* 130.12 (2003): 2633-2644.

Christian, Jan L., and Randall T. Moon. "Interactions between Xwnt-8 and Spemann organizer signaling pathways generate dorsoventral pattern in the embryonic mesoderm of Xenopus." *Genes & development* 7.1 (1993): 13-28.

Christian, Jan Louise, et al. "Xwnt-8, a Xenopus Wnt-1/int-1-related gene responsive to mesoderm-inducing growth factors, may play a role in ventral mesodermal patterning during embryogenesis." *Development* 111.4 (1991): 1045-1055.

Crossley, Philip H., Salvador Martinez, and Gail R. Martin. "Midbrain development induced by FGF8 in the chick embryo." *Nature* 380.6569 (1996): 66.

Danielian, Paul S., et al. "A 5.5-kb Enhancer Is both Necessary and Sufficient for Regulation ofWnt-1Transcriptionin Vivo." *Developmental biology* 192.2 (1997): 300-309.

Dorsett, Dale. "Distant liaisons: long-range enhancer–promoter interactions in Drosophila." *Current opinion in genetics & development* 9.5 (1999): 505-514.

Echelard, Yann, Galya Vassileva, and Andrew P. McMahon. "Cis-acting regulatory sequences governing Wnt-1 expression in the developing mouse CNS." *Development* 120.8 (1994): 2213-2224.

Emison, Eileen Sproat, et al. "A common sex-dependent mutation in a RET enhancer underlies Hirschsprung disease risk." *Nature* 434.7035 (2005): 857-863.

Ertzer, Raymond, et al. "Cooperation of sonic hedgehog enhancers in midline expression." *Developmental biology* 301.2 (2007): 578-589.

Gaiano, Nicholas, et al. "Highly efficient germ-line transmission of proviral insertions in zebrafish." *Proceedings of the National Academy of Sciences* 93.15 (1996): 7777-7782.

Glinka, Andrei, et al. "Dickkopf-1 is a member of a new family of secreted proteins and functions in head induction." *Nature* 391.6665 (1998): 357-362.

Guelen, Lars, et al. "Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions." *Nature* 453.7197 (2008): 948-951.

Higashijima, Shin-ichi, et al. "High-frequency generation of transgenic zebrafish which reliably express GFP in whole muscles or the whole body by using promoters of zebrafish origin." *Developmental biology* 192.2 (1997): 289-299.

Houart, Corinne, et al. "Establishment of the telencephalon during gastrulation by local antagonism of Wnt signaling." *Neuron* 35.2 (2002): 255-265.

Jászai, József, et al. "Isthmus-to-midbrain transformation in the absence of midbrain-hindbrain organizer activity." *Development* 130.26 (2003): 6611-6623.

Joyner, Alexandra L., Aimin Liu, and Sandrine Millet. "Otx2, Gbx2 and Fgf8 interact to position and maintain a mid–hindbrain organizer." *Current opinion in cell biology* 12.6 (2000): 736-741

Katahira, Tatsuya, et al. "Interaction between Otx2 and Gbx2 defines the organizing center for the optic tectum." *Mechanisms of development* 91.1 (2000): 43-52.

Kawakami, Koichi, Akihiro Shima, and Noriko Kawakami. "Identification of a functional transposase of the Tol2 element, an Ac-like element from the Japanese medaka fish, and its transposition in the zebrafish germ lineage." *Proceedings of the National Academy of Sciences* 97.21 (2000): 11403-11408.

Kawakami, Koichi, and Akihiro Shima. "Identification of the Tol2 transposase of the medaka fish Oryzias latipes that catalyzes excision of a nonautonomous Tol2 element in zebrafish Danio rerio." *Gene* 240.1 (1999): 239-244.

Kawakami, Koichi, et al. "A transposon-mediated gene trap approach identifies developmentally regulated genes in zebrafish." *Developmental cell*7.1 (2004): 133-144.

Kawakami, Koichi. "Tol2: a versatile gene transfer vector in vertebrates." *Genome biology* 8.1 (2007): S7.

Kelly, Gregory M., Deniz F. Erezyilmaz, and Randall T. Moon. "Induction of a secondary embryonic axis in zebrafish occurs following the overexpression of β-catenin." *Mechanisms of development* 53.2 (1995): 261-273.

Kim, Kwonseop, Zifan Lu, and Elizabeth D. Hay. "Direct evidence for a role of β-catenin/Lef-1 signaling pathway in induction of EMT." *Cell biology international* 26.5 (2002): 463-476.

Kolovos, Petros, et al. "Enhancers and silencers: an integrated and simple model for their function." *Epigenetics & chromatin* 5.1 (2012): 1.

Krauss, Stefan, et al. "Expression of four zebrafish wnt-related genes during embryogenesis." *Development* 116.1 (1992): 249-259.

Krauss, Stefan, et al. "Zebrafish pax [b] is involved in the formation of the midbrain-hindbrain boundary." *Nature* 360.6399 (1992): 87.

Lekven, Arne C., et al. "Wnt1 and wnt10b function redundantly at the zebrafish midbrain–hindbrain boundary." *Developmental biology* 254.2 (2003): 172-187.

Lettice, Laura A., et al. "A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly." *Human molecular genetics* 12.14 (2003): 1725-1735.

Levine, Arnold J., and Anna M. Puzio-Kuter. "The control of the metabolic switch in cancers by oncogenes and tumor suppressor genes." *Science* 330.6009 (2010): 1340-1344.

Leyns, Luc, et al. "Frzb-1 is a secreted antagonist of Wnt signaling expressed in the Spemann organizer." *Cell* 88.6 (1997): 747-756.

Marshall, Christopher J. "MAP kinase kinase kinase, MAP kinase kinase and MAP kinase." *Current opinion in genetics & development* 4.1 (1994): 82-89.

Maston, Glenn A., et al. "Characterization of enhancer function from genome-wide analyses." *Annual review of genomics and human genetics* 13 (2012): 29-57.

McGrew, L. Lynn, Cheng-Jung Lai, and Randall T. Moon. "Specification of the Anteroposterior Neural Axis through Synergistic Interaction of the Wnt Signaling Cascade withnogginandfollistatin." *Developmental biology* 172.1 (1995): 337-342.

McMahon, Andrew P., and Allan Bradley. "The Wnt-1 (int-1) proto-oncogene is required for development of a large region of the mouse brain." *Cell* 62.6 (1990): 1073-1085.

McMahon, Andrew P., and Randall T. Moon. "Ectopic expression of the proto-oncogene int-1 in Xenopus embryos leads to duplication of the embryonic axis." *Cell* 58.6 (1989): 1075-1084.

Millet, Sandrine, et al. "A role for Gbx2 in repression of Otx2 and positioning the

mid/hindbrain organizer." *Nature* 401.6749 (1999): 161-164.

Noordermeer, Jasprien, et al. "Isolation of the Xenopus bomolog of int-1/wingless and expression during neurula stages of early development." *Nucleic Acids Research* 17.1 (1989): 11-18.

Nusse, Roel, and Harold E. Varmus. "Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome." *Cell* 31.1 (1982): 99-109.

Nusse, Roel. "An ancient cluster of Wnt paralogues." *Trends in Genetics*17.8 (2001): 443.

Nusslein-Volhard, Christiane, and Ralf Dahm. *Zebrafish*. Oxford University Press, 2002.

Ohno, Susumu. "So much" junk" DNA in our genome." *Brookhaven symposia in biology*. Vol. 23. 1972.

Ong, Chin-Tong, and Victor G. Corces. "Enhancer function: new insights into the regulation of tissue-specific gene expression." *Nature Reviews Genetics*12.4 (2011): 283-293.

Ong, Chin-Tong, and Victor G. Corces. "Enhancer function: new insights into the regulation of tissue-specific gene expression." *Nature Reviews Genetics*12.4 (2011): 283-293.

Pennacchio, Len A., et al. "In vivo enhancer analysis of human conserved non-coding sequences." *Nature* 444.7118 (2006): 499-502.

Postlethwait, John H., et al. "Vertebrate genome evolution and the zebrafish gene map." *Nature genetics* 18.4 (1998): 345-349.

Ptashne, Mark. "Gene regulation by proteins acting nearby and at a distance." *Nature* 322 (1986): 697-701.

Raible, Florian, and Michael Brand. "Divide et Impera–the midbrain–hindbrain boundary and its organizer." *Trends in neurosciences* 27.12 (2004): 727-734.

Rhinn, Muriel, and Michael Brand. "The midbrain–hindbrain boundary organizer." *Current opinion in neurobiology* 11.1 (2001): 34-42.

Rieder, Dietmar, Zlatko Trajanoski, and James McNally. "Transcription factories." *Frontiers in genetics* 3 (2012): 221.

Rijsewijk, Frans, et al. "The Drosophila homology of the mouse mammary oncogene int-1 is identical to the segment polarity gene wingless." *Cell* 50.4 (1987): 649-657.

Roelink, Henk, et al. "Wnt-3, a gene activated by proviral insertion in mouse mammary tumors, is homologous to int-1/Wnt-1 and is normally expressed in mouse embryos and adult brain." *Proceedings of the National Academy of Sciences* 87.12 (1990): 4519-4523.

Rowitch, David H., et al. "Identification of an evolutionarily conserved 110 base-pair cis-acting regulatory sequence that governs Wnt-1 expression in the murine neural plate." *Development* 125.14 (1998): 2735-2746.

Schneider, Robert, and Rudolf Grosschedl. "Dynamics and interplay of nuclear architecture, genome organization, and gene expression." *Genes & development* 21.23 (2007): 3027-3043.

Shibata, Mikihito, et al. "Role of crescent in convergent extension movements by modulating Wnt signaling in early Xenopus embryogenesis." *Mechanisms of development* 122.12 (2005): 1322-1339.

Shlyueva, Daria, Gerald Stampfel, and Alexander Stark. "Transcriptional enhancers: from properties to genome-wide predictions." *Nature Reviews Genetics* 15.4 (2014): 272-286.

Sidow, Arend. "Diversification of the Wnt gene family on the ancestral lineage of vertebrates." *Proceedings of the National Academy of Sciences* 89.11 (1992): 5098-5102.

Sokol, Sergei, et al. "Injected Wnt RNA induces a complete body axis in Xenopus embryos." *Cell* 67.4 (1991): 741-752.

Song, Dong-Li, et al. "Two Pax-binding sites are required for early embryonic brain expression of an Engrailed-2 transgene." *Development* 122.2 (1996): 627-635.

Stathopoulos, Angelike, and Michael Levine. "Genomic regulatory networks and animal development." *Developmental cell* 9.4 (2005): 449-462.

Stuart, Gary W., James V. McMurray, and Monte Westerfield. "Replication, integration and stable germ-line transmission of foreign sequences injected into early zebrafish embryos." *Development* 103.2 (1988): 403-412.

Theil, Thomas, et al. "Wnt and Bmp signalling cooperatively regulate graded Emx2 expression in the dorsal telencephalon." *Development* 129.13 (2002): 3045-3054.

Thomas, Kirk R., and Mario R. Capecchi. "Targeted disruption of the murine int-1 proto-oncogene resulting in severe abnormalities in midbrain and cerebellar development." *Nature* 346.6287 (1990): 847.

Thomas, Kirk R., et al. "Swaying is a mutant allele of the proto-oncogene Wnt-1." *Cell* 67.5 (1991): 969-976.

Thorpe, Kevin W., et al. "Aerial application of the viral enhancer Blankophor BBH with reduced rates of gypsy moth (Lepidoptera: Lymantriidae) nucleopolyhedrovirus." *Biological Control* 16.2 (1999): 209-216.

Urasaki, Akihiro, Ghislaine Morvan, and Koichi Kawakami. "Functional dissection of the Tol2 transposable element identified the minimal cis-sequence and a highly repetitive sequence in the subterminal region essential for transposition." *Genetics* 174.2 (2006): 639-649.

van 't Veer, Laura J., et al. "Molecular cloning and chromosomal assignment of the human homolog of int-1, a mouse gene implicated in mammary tumorigenesis." *Molecular and cellular biology* 4.11 (1984): 2532-2534.

Vaquerizas, Juan M., et al. "A census of human transcription factors: function, expression and evolution." *Nature Reviews Genetics* 10.4 (2009): 252-263.

Wainwright, B. J., et al. "Isolation of a human gene with protein sequence similarity to human and murine int-1 and the Drosophila segment polarity mutant wingless." *The EMBO journal* 7.6 (1988): 1743.

Wang, Shouwen, et al. "Frzb, a secreted protein expressed in the Spemann organizer, binds and inhibits Wnt-8." *Cell* 88.6 (1997): 757-766.

Wurst, Wolfgang, and Laure Bally-Cuif. "Neural plate patterning: upstream and downstream of the isthmic organizer." *Nature Reviews Neuroscience* 2.2 (2001): 99-108.

Yamamoto, Akihito, et al. "Shisa promotes head formation through the inhibition of receptor protein maturation for the caudalizing factors, Wnt and FGF." *Cell* 120.2 (2005): 223-235.