THE ROLE OF Sp1 FAMILY GENES IN WNT MEDIATED NEURAL PLATE PATTERNING

A Thesis

by

SAURAV MOHANTY

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,	Thomas McKnight

December 2017

Major Subject: Biology

ABSTRACT

Wnt signaling has a conserved role in anterior-posterior (A-P) patterning of the remarkably complex vertebrate brain from a very basic primordium, the neural plate, by differentially regulating downstream target genes to induce distinct neural fates. Wnt proteins originate from the blastoderm margin and diffuse anteriorly to re-pattern the default state of the anterior neural tissue to more posterior fates. How this graded Wnt ligand in the neural plate is translated by responding cells to establish specific cell fate zones along the A-P axis is not well established.

In order to understand neural patterning by Wnt better, we need to identify the downstream components of this pathway, for instance, the transcriptional targets of Wnt signaling and their subsequent roles. Two well known but less studied targets of Wnt/β-catenin signaling in Zebrafish, the *sp5a* and *sp5-like* (*sp5l*) genes, are good models to study induction of transcription factors and gene regulatory network that drives neural posteriorization downstream of Wnt signaling because these paralogous genes potentially play an important role in mediating Wnt-dependent posterior neural plate patterning.

Results from our recent RNA sequencing analysis identifies another gene, *sp8b*, along with the *sp5a* and *sp5l* genes, that may also play a role in Wnt dependent patterning in the nervous system. These Wnt target genes encode zinc finger transcription factors (TFs) belonging to the *Sp1* gene family and transcriptional regulation of *sp5* orthologs by Wnt signaling is known to be conserved in vertebrates. In spite of the compelling evidence of a significant and evolutionarily conserved relation between the *sp5* genes and Wnt

signaling in vertebrates, the function of these genes and how they are regulated by the Wnt pathway to pattern the vertebrate brain is not understood properly. This thesis is aimed at understanding the in-vivo function of the *sp5* an *sp8b* genes by generating zebrafish genetic mutants.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Arne Lekven, for his guidance, support and continuous optimism, my committee members, Dr. Gary Kunkel and Dr. Alvin Yeh, for their helpful comments and insights throughout the course of this research. I would also like to thank all the present and former members of the Lekven lab- Amy, David, Jo-Ann and Avantika, for providing useful feedback on my experiments and helping me grow as a researcher.

Thanks to all my friends, specially Vibhuti, for years of joyful friendship and my music teacher Ms. Suchitra Acharya for introducing me to a calming and beautiful form of art. Finally, I am grateful to have a wonderful and supportive family. Thanks to my mother for always believing in me, my father for never letting me give up, my sister for unceasing care, Rohan for being the cool brother I never had and Buddy for always cheering me up.

CONTRIBUTORS AND FUNDING SOURCES

This work was supervised by a thesis committee consisting of Dr. Arne Lekven of the Department of Biology, 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.

Graduate study was supported by a teaching assistantship from the Biology Department at Texas A&M University. Partial support came from a grant from the NIH (R01NS088564).

NOMENCLATURE

Cas9	CRISPR associated protein 9
CRISPR	Clustered regularly interspaced short palindromic repeats
CrRNA	CRISPR RNA
Dkk1b	Dickkopf related protein 1b
EGFP	Enhanced green fluorescent protein
FGF	Fibroblast growth factor
Hpf	Hours post fertilization
HRMA	High resolution melting analysis
LEF	Lymphoid enhancer-binding factor 1
sgRNA	small guide RNA
TCF3	Transcription Factor 3
TF	Transcription Factor
TracrRNA	Trans-activating CrRNA

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

Wnt signaling

The Wnt Signaling Pathway is an ancient and evolutionarily conserved signal transduction pathway in animals. It comprises several Wnt ligands, a large family of secreted glycoproteins, that pass a signal into a cell through the cell surface receptors thereby inducing changes that control multiple aspects of development, including the proliferation, fate specification, polarity, and migration of cells (Cadigan et al. 1997). The word 'Wnt' is derived from the fusion of the names of two homologous genes discovered in the 1980s- the Drosophila segment polarity gene Wingless and the mouse proto-oncogene Int-1 (Nusse et al. 1982, Rijsewijk et al. 1987).

In different developmental contexts, the downstream effectors of Wnt define two major branches of the Wnt pathway: (i) canonical (β-catenin dependent); and (ii) the non-canonical (β-catenin independent), which has various branches (Niehrs 2012). This study is focused on the canonical Wnt/β-catenin signaling, where a cytoplasmic protein, β-catenin, is the primary downstream effector molecule. Wnt proteins act on target cells by binding to the Frizzled (Fz)/low density lipoprotein (LDL) receptor-related protein (LRP) complex at the cell surface (Schweizer et al. 2003). These receptors transduce a signal to several intracellular proteins that include Dishevelled (Dsh), glycogen synthase kinase-3β

(GSK-3), Axin, Adenomatous Polyposis Coli (APC), CK1 (Casein kinase 1) and the transcriptional regulator, β-catenin (Clevers 2006).

In the absence of Wnt signal, cytoplasmic β- catenin levels are maintained at low levels through continuous proteasome-mediated degradation, which is controlled by a complex containing GSK-3/APC/CK1/Axin (Logan et al. 2004). In the nucleus, transcription factors such as lymphoid enhancer-binding factor 1/T cell-specific transcription factor (TCF/LEF) bind regulatory regions of Wnt-responsive genes and interact with other factors (e.g. Groucho, histone deacetylase) to repress transcription (Eastman et al. 1999). When cells receive Wnt signals, the degradation pathway is inhibited, and consequently β-catenin accumulates in the cytoplasm and nucleus. Nuclear β-catenin interacts with LEF/TCF factors to activate transcription of Wnt target genes (Figure 1).

Wnt signaling and nervous system patterning

Nervous system development has been a classic model to study pattern formation in vertebrate embryos because a very basic primordium, the neural plate, that develops from a group of ectodermal cells in the embryo undergoes rigorous patterning in the developing embryo to form a remarkably complex brain. The central nervous system, which has a fairly conserved morphology in all vertebrates (Northcutt 2002), is subdivided into four distinct regions, with a distinct A-P character, the most anterior forebrain (telencephalon and diencephalon), the midbrain (mesencephalon), the hindbrain (rhombencephelon) which is divided into seven segments called rhombomeres, and, finally, the spinal cord that forms posterior to the brain and extends to the rear end of the body (Figure 2).

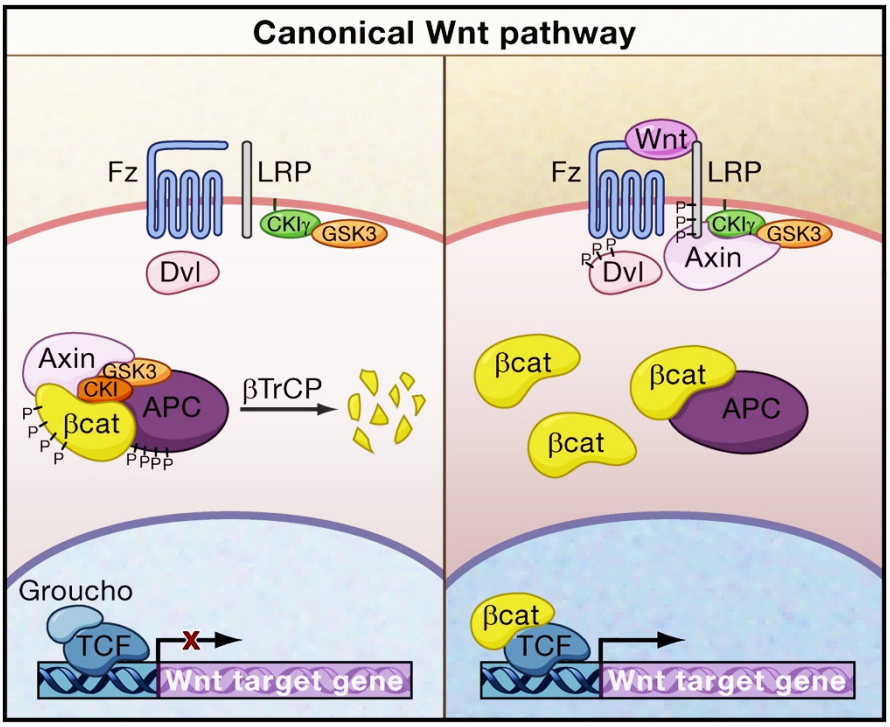

Figure 1. **Canonical Wnt signaling pathway.** Left: In the absence of Wnt, cytoplasmic β-catenin is sequestered and degraded in the cytoplasm and Wnt target genes are not transcribed in the nucleus. Right: Upon binding of Wnt to Fz/LRP receptor complex, destruction complex is inhibited, which results in accumulation and entry of β-catenin into the nucleus, resulting in transcription of Wnt target genes (Clevers 2006).

The initial neural tissue (neuroectoderm), which is induced by a cluster of cells called the embryonic organizer, has an anterior forebrain nature. The posterior mesoderm secretes posteriorizing molecules in a gradient along the A-P axis, that re-pattern the default state of anterior neural tissue to more posterior cell fates (Nieuwkoop 1952, Toivonen et al. 1968). Wnt/β-catenin signaling is involved in multiple steps of vertebrate neural patterning, including the initial specification of the posterior regions of the early nervous system, the fine-tuned patterning of each region, cell fate determination and proliferation

and differentiation of a specific cell populations. Within the posterior nervous system, posteriorizing molecules such as Wnt along with other signaling pathways such as the Retinoic Acid (RA) and Fibroblast Growth Factor (FGF), induces midbrain-hindbrain border (MHB), hindbrain, spinal cord and neural crest cell fates during late gastrulation stages of the embryo (McGrew et al. 1997, Woo et al. 1997, Kiecker et al. 2001, Villanueva et al. 2002).

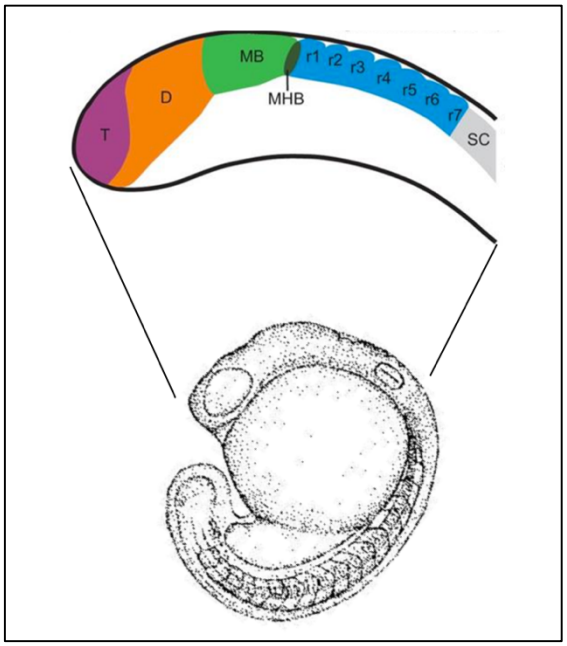

Figure 2. Vertebrate CNS subdivisions. Schematic diagram of an 18 hpf zebrafish embryo (anterior to left) showing the early subdivisions of the brain: From anterior to posterior- T= Telencephalon, D= Diencephalon (Forebrain); MB=Midbrain; MHB= Mid-hindbrain boundary; r= rhombomere (r1-r7, hindbrain); SC=spinal cord. Modified from (Kimmel et al. 1995, Šestak et al. 2014).

The role of Wnt in posteriorization of the neuroectoderm is studied is various vertebrate models. Wnt1 and Wnt3a deficient mouse embryos fail to develop posterior neural structures such as the midbrain, hindbrain and spinal cord properly (McMahon et al. 1992,

Augustine et al. 1995). Expression of dominant negative Wnt proteins causes the suppression of mid and hindbrain fates in *Xenopus* (McGrew et al. 1997). In zebrafish, inactivation of Wnt3 and Wnt8 abrogates formation of posterior brain fates and spinal cord (Erter et al. 2001, Lekven et al. 2001, Buckles et al. 2004). In contrast, activation of the Wnt pathway by overexpression of β-catenin or Tcf, induces posterior neural markers and suppresses anterior character (Domingos et al. 2001, Wu et al. 2005).

The gradient model of Wnt/β-catenin activity suggests that increasing activity of Wnt induces progressively more posterior fates (Nordström et al. 2002). In zebrafish, Wnt proteins originate from the blastoderm margin and diffuse anteriorly in a graded manner to re-pattern the default state of the anterior neural tissue to more posterior fates. How this graded Wnt pattern is interpreted in different regions along the neural plate to establish different cell fates is not clearly understood.

Transcriptional factors that are direct targets of the Wnt/β-catenin pathway must act downstream to mediate its morphogenetic activity. Therefore, studying how Wnt targets mediate downstream Wnt activity is crucial for understanding the mechanism of the Wnt pathway in nervous system patterning in vertebrates.

sp5 and *sp5l* are direct Wnt targets (Weidinger et al. 2005) and thus good model genes for understanding how differential Wnt signaling across the A/P neural axis is translated by receiving cells to form distinct A/P subdivisions of the brain. Recent studies also indicate that the *sp8b* gene is another potential Wnt target that regulates Wnt mediated neural patterning along with the *sp5a* and *sp5l* genes.

Expression of *sp5a, sp5l* and *sp8b* genes during embryogenesis

The *sp5a* gene was discovered in a screen for differentially expressed genes during mouse gastrulation and was found to encode a transcription factor containing DNA-interacting Zinc finger domains, characteristic of the Sp1 family of TFs (Harrison et al. 2000). During mouse development, it shows a highly dynamic expression pattern and is observed in the posterior neural plate, underlying mesoderm, notochord and the tail bud (Treichel et al. 2001), indicating its diverse transcriptional regulatory functions. Zebrafish *sp5a* is first observed in the ventrolateral blastoderm margin and in the yolk syncytial layer before gastrulation (Figure 3). Initially it is seen in the posterior epiblast, sharply in the midbrain-hindbrain boundary (MHB) during gastrulation and later the expression gets confined to MHB and the tail bud towards the end of gastrulation (Tallafuß et al. 2001).

sp5l is expressed in the neuroectoderm, adjacent hypoblast during gastrulation and later in the tail bud (Zhao et al. 2003, Sun et al. 2006). Overlapping expression of *sp5a* and *sp5l* in the posterior neuroectoderm indicates their potentially cooperative role in Wnt dependent A/P patterning during embryogenesis. Also, the distinct expression of *sp5a*, particularly in *wnt8a* expressing margin and the MHB organizer, point towards its additional potential function in neural ectoderm and ventrolateral mesoderm specification. *sp8b* is expressed in the neural plate during epiboly and strongly in the diencephalon, hindbrain and spinal cord at later stages (somite stage to 24hpf) of development (Thisse et al. 2004).

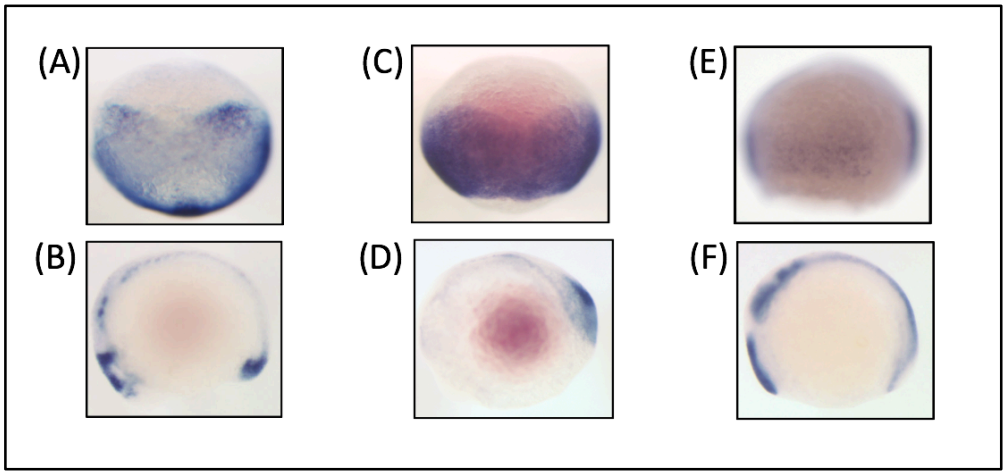

Figure 3. Expression of *sp5a*, *sp5l* and *sp8b* during gastrulation and somitogenesis. In situ hybridization for *sp5a* (A) 80% epiboly, (B) 14-somite stage; *sp5l* (C) 80% epiboly, (D) 3-somite stage and *sp8b* (E) 70% epiboly, (F) 14-somite stage. (A, C) Dorsal views, (B, D, F) Lateral views, anterior to left, (E) side view, anterior to top.

sp5a, *sp5l*, *sp8b* genes and Wnt signaling

Zebrafish *sp5* genes are direct targets of Wnt/β-catenin signaling and have been shown to act downstream of *wnt8a* to pattern the mesoderm and neuroectoderm (Weidinger et al. 2005), but the mechanism of their transcriptional regulation by Wnt and their downstream effects influencing neural patterning are unknown. *sp5* orthologs have been recognized as Wnt target genes in several model organisms such as *Xenopus*, mouse and humans (Takahashi et al. 2005, Park et al. 2013, Dunty Jr et al. 2014). The promoter of *sp5a* is highly conserved across different vertebrate species with a striking identity of 88% within a 500 bp region in the promoters of mouse and zebrafish *sp5a* (Figure 4). This region contains 5 consensus binding sites for Lef/Tcf proteins, which are the effectors of Wnt/β-catenin signaling. Although the induction of *sp5* in response to Wnt signaling seems to be a conserved feature in the vertebrates, their precise role in Wnt signaling mediated A/P

patterning in vertebrates is not studied well. *sp8b* is expressed in the neural plate and regulation of *sp8* by Wnt during limb development and neuropore closure, is a conserved feature in vertebrates (Bell et al. 2003, Kawakami et al. 2004).

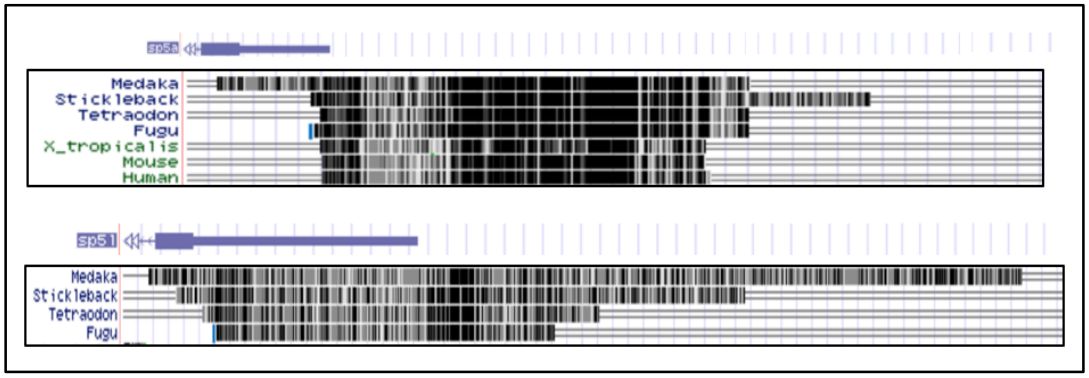

Figure 4. Sequence conservation in *sp5a* and *sp5l* upstream regions. Upstream regions of *sp5a* (top) and *sp5l* (bottom) Blue boxes represent Exon 1 and black lines represent sequence conservation. (UCSC genome browser)

The exact mechanism of activation of these genes by Wnt signaling is unknown. For instance, *sp5* genes can either be directly activated by Wnt signaling above a threshold value by recruitment of activating proteins such as CBP/p300 (Li et al. 2007) or de-repressed by Wnt signaling, wherein they are indirectly activated by removal of repression imposed by *tcf3* (Dorsky et al. 2003), to allow activation by other signaling pathways. A study indicates that *sp5l* expression in the dorsal neuroectoderm requires both Fgf and Wnt signaling and its expression in the epiblast is partially dependent on Fgf (Zhao et al.

2003). Therefore, the connection between Wnt signaling, *sp5* and *sp8b* genes expression is unsettled.

Transcriptional role of *sp5* and *sp8b* genes

sp5 genes and *sp8b* encode transcription factors belonging to the Sp1 family, characterized by a highly conserved zinc finger domain (Treichel et al. 2001), so they must be modulating the expression of their target genes to effect Wnt mediated neural patterning. *sp5* genes can act as transcriptional activators or repressors in a context dependent manner (Tallafuß et al. 2001, Fujimura et al. 2007). A recent study shows that the Sp5 recruits β-catenin and Lef/Tcf proteins to the enhancers of Wnt target genes in mouse ESCs to activate their transcription; the target genes include *pax2*, *cdx1*, *gbx2* and hoxb1 (Kennedy et al. 2016), which are also expressed in the posterior neural plate.

Whether *sp5* genes activate posterior neural genes or repress anterior neural genes to transform Wnt signals to induce neural posteriorization is unclear. Also, their repressor activity may be direct or indirect. For example, *sp5l* ectopic expression represses *tcf3* (*headless*) transcription (Weidinger et al. 2005) and they might cause loss of anterior neural fates not by directly repressing anterior neural genes but by indirectly repressing the repressor of posterior neural genes i.e., *tcf3* (Kim et al. 2000). Therefore, it is important to identify the function of these genes and determine how Sp5 and Sp8b TFs induce changes in the neural plate gene expression pattern to mediate Wnt dependent neural patterning.

The work presented in this thesis was designed to generate *sp5a*, *sp5l* and *sp8b* mutants to study the role of these genes in Wnt-mediated patterning.

Zebrafish as a model organism to study vertebrate nervous system patterning

The research presented here utilized zebrafish as a system to generate *sp5a*, *sp5l* and *sp8b* mutants and determine the function of these genes. Zebrafish is an ideal model organism for this study because of their relatively short life cycle, rapid external development and easy genetic manipulation of the embryos. Adult fish are very easy to maintain and breed, and it is possible to get 100-200 offspring from a defined breeding pair of fish at once, making it a suitable system for genetic analysis.

sp5 genes have been studied in other vertebrates such as *Xenopus* and mouse, in several contexts. But the longer generation time of *Xenopus* makes generation of stable transgenic lines or mutant lines a slow process and the in utero development of mouse embryos makes it difficult to analyze early patterning of neural plate. Zebrafish embryos are completely transparent up to 24 hpf (organogenesis) and at this point the different subdivisions of the brain are well defined. So, it is possible to follow the impact of an early genetic manipulation on brain A/P patterning in live embryos, using simple imaging techniques.

Generation of zebrafish mutants using CRISPR-Cas9 system

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)- CRISPR-associated protein 9 (Cas9) is a genome editing tool that is derived from a naturally

occurring defense mechanism in bacteria and archaea (Horvath et al. 2010, Jinek et al. 2012). It utilizes a RNA guided DNA endonuclease, Cas9, that induces site specific double stranded cleavage in DNA when complexed with a single guide RNA (sgRNA). An sgRNA is made up of a 20 nucleotide targeting sequence called the Crispr RNA (crRNA) which is complementary to the targeted genomic region and a trans-activating crRNA (tracrRNA) that is partially complementary to the crRNA and required for processing of the crRNA (Chylinski et al. 2013, Doudna et al. 2014). When the complex is formed, Cas9 is guided by the sgRNA to the target crRNA complementary sequence in the genome that is adjacent to a PAM (Protospacer-associated motif) sequence (Figure 5).

The double stranded endonuclease activity of Cas9 requires PAM, a short conserved 2-5 nucleotide sequence (usually NGG) that follows immediately the 3' end of the genomic target site (Swarts et al. 2012). The double stranded break in genomic DNA caused by Cas9 is repaired by cell repair mechanisms and can repaired by two different mechanisms – (1) Non-homologous end joining (NHEJ), which results in insertion or deletion of several nucleotides at the target site (INDEL mutation) that can disrupt the reading frame with a premature stop. (2) Homology directed repair (HDR) where a desirable donor homologous sequence guides the repair, resulting in the replacement of the target locus by another sequence of interest (Ran et al. 2013).

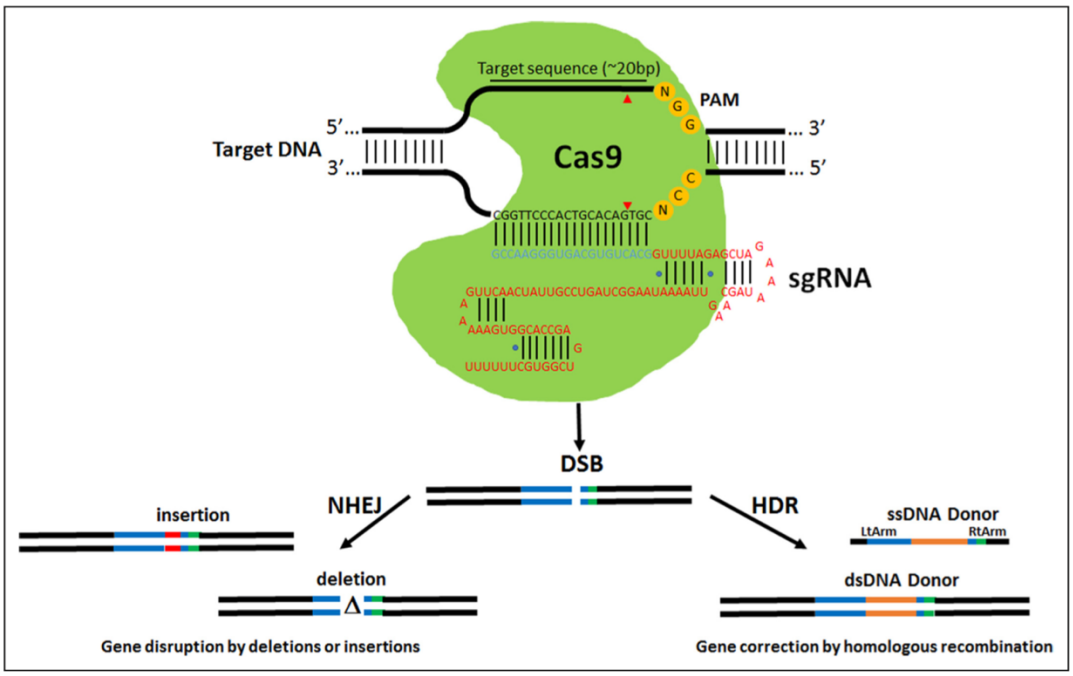

Figure 5. CRISPR-Cas9 genome editing. Schematic representation of the Cas9 nuclease (green) targeted to genomic DNA by a single-guide RNA (sgRNA) consisting of a 20 nucleotide crRNA sequence (blue) and the tracrRNA (red). The guide sequence is directly upstream of the protospacer adjacent motif (PAM), NGG (orange circles). Cas9 mediates a double-strand DNA break (DSB) upstream of the PAM (red triangles). The break is repaired by either nonhomologous end joining (NHEJ) that creates random insertions or deletions at the target site or homology-directed repair (HDR), that replaces the target DNA with the donor DNA containing homology arms at either end (Doetschman et al. 2017).

Detecting CRISPR-Cas9 induced mutations using high resolution melting analysis

High resolution melting analysis (HRMA) is a post-PCR analysis used to identify variations in nucleic acid sequences. To identify sequence variation in a target region, the region of interest is amplified using primers that amplify a product of size 100-300 bp. Amplification takes place in presence of saturation dyes that only fluoresce when bound to double stranded DNA. As the amplicon concentration increases, the fluorescence exhibited by the double stranded amplified product also increases. Next, the amplified

product (the target region) is gradually denatured by increasing temperature in small increments, to generate a melting profile. As it denatures, the dye is released, causing a drop in fluorescence. Different nucleic acid sequences will have different melting profiles. HRMA is sensitive enough to detect single nucleotide difference between two otherwise identical sequences.

This technique has advantages over many other methods of genotyping in terms of sensitivity, cost and total running time. It is a cheaper alternative to sequencing when hundreds of samples are required to be genotyped, it eliminates the use of restriction enzymes and agarose gels, and overall is time efficient because all of the necessary reagents can be combined together to perform the PCR in a single tube (Wittwer 2009, Thomsen et al. 2012). HRMA has been used extensively to detect sequence polymorphisms in zebrafish including detection of point mutations, and deletions and insertions of 3-20 bp in the target locus (Parant et al. 2009, Dahlem et al. 2012, Talbot et al. 2014, Yu et al. 2014).

The CRISPR-Cas9 genome editing tool and HRMA were the two main techniques that were utilized to complete the first aim of this project, which was to generate and identify the *sp5a*, *sp5l* and *sp8b* mutants.

CHAPTER II

GENERATION AND RECOVERY OF *sp5a*, *sp5l* AND *sp8b* MUTANT ALLELES

Identification of *sp5a*, *sp5l* and *sp8b* as potential Wnt mediators

We conducted RNA-sequencing on HS:dkk zebrafish lines in which the canonical wnt inhibitor, *dkk1b* gene, is under the control of an inducible heat shock (HS) promoter. We inhibited Wnt/β-catenin signaling during early (5hpf) and late (7hpf) gastrulation, time points crucial for neural plate patterning, and looked for differentially expressed genes at these two time points. Three Sp1 family genes- *sp5a*, *sp5l*, and *sp8b* showed significant reduction in expression upon suppression of Wnt signaling (Figure 6).

This indicates that the Sp1 family genes may play a significant role in mediating transcriptional regulation of neural genes by Wnt signaling. This result reinforces the idea that the *sp5* genes are mediators of Wnt during neural patterning and also identifies *sp8b* as a novel potential Wnt target in neural plate.

In order to study the function of the *sp5* and *sp8b* genes, we generated *sp5a*, *sp5l* and *sp8b* mutants using the CRISPR-Cas9 genome editing tool. We recovered the mutant alleles for each gene using High resolution melting assays and confirmed the exact nature of the mutation by sequencing. The goal of this part of the study was to study the independent roles of *sp5a*, *sp5l* and *sp8b* in embryonic brain patterning and together to determine whether these genes act redundantly to pattern the brain.

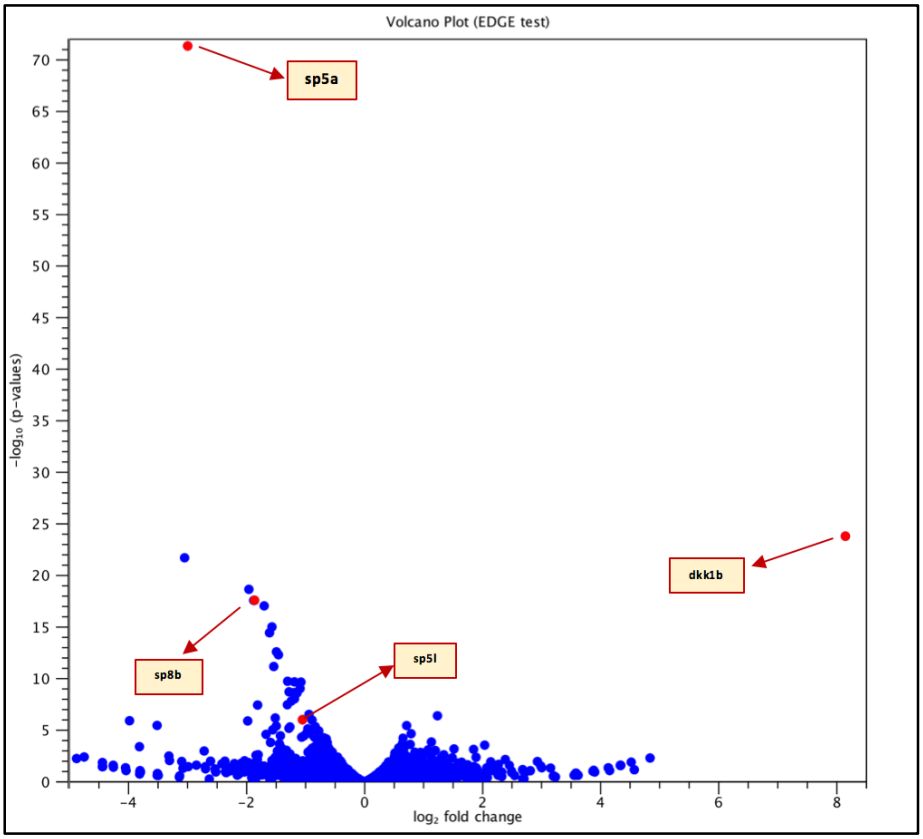

Figure 6. Differentially expressed genes upon Wnt/β-catenin suppression. Volcano plot showing Sp1 family genes- *sp5a*, *sp5l* and *sp8b*, are significantly downregulated when Wnt signaling is suppressed at 7 hpf. *dkk1b* under a heat shock promoter is used as canonical Wnt inhibitor. X-axis represents fold change in gene expression and Y-axis represents p-values. Both axes are represented in log scale.

Materials and methods

Generation of mutant alleles using CRISPR-Cas9 system

To design the CRISPRs we utilized the online tool CRISPRscan (Moreno-Mateos et al. 2015). To synthesize sgRNA targeting *sp5a* and *sp5l*, we first selected a pair of oligonucleotides targeting exon 2 of *sp5a* and *sp5l*. A BsaI recognition site was added to the 5'end of the oligos. Next, to make the crRNA template insert, the oligos were annealed by heating to 90-95°C and slow cooling to room temperature. Zebrafish sgRNA expression vector pDR274 (Hwang et al. 2013) was digested with BsaI and purified by gel extraction using the Zymoclean™ gel DNA recovery kit. The crRNA template insert was ligated to the digested vector by T4 DNA ligase.

After cloning of the insert, the plasmid was transformed into competent *E.coli* cells and colonies were grown overnight at 37°C on LB/Kanamycin plates. Ten random white colonies were picked and cultured for plasmid DNA preparation. The plasmid DNA samples were digested with BsaI to confirm the presence of insert because the BsaI sites are lost after successful ligation. Plasmid with insert was linearized using DraI, purified, and sgRNA was synthesized from the plasmid using T7 RNA polymerase. MEGAscript® T7 kit was used to transcribe sgRNA from the T7 promoter positioned upstream of the sgRNA insert site in the vector. The sgRNA was purified by the phenol chloroform (PCl) extraction method.

cas9 mRNA was synthesized from pCS2-nCas9n (Jao et al. 2013). To prepare the vector for transcription, it was linearized with NotI and purified with the Zymo DNA Clean & Concentrator™-5 kit. *cas9* mRNA was synthesized with the mMESSAGE mMACHINE® SP6 Transcription Kit. mRNA was purified using the PCl extraction method after it was transcribed.

To synthesize *sp8b* sgRNA we utilized an 80 base constant oligonucleotide, which contains the tracrRNA part of the guide RNA. We designed a gene specific oligonucleotide (crRNA) using the online tool CRISPRscan. The gene specific oligonucleotide has the T7 promoter sequence at its 5' end, 20 nt sequence that is complementary to the genomic target i.e., exon 2 of *sp8b* and 20-22 nt sequence at its 3' end that is complementary to the constant oligonucleotide. The gene specific oligonucleotide is annealed to the constant oligonucleotide and filled in using T4 DNA polymerase to generate the complete sgRNA template (Gagnon et al. 2014). The template was purified using Zymo DNA Clean & Concentrator™-5. MEGAscript® T7 kit was then used to synthesize the guide RNA from the template and the sgRNA was purified using the PCl extraction method.

We injected 30 pg of sgRNA and 300 pg of Cas9 mRNA into 1-cell zebrafish wild type embryos, in separate experiments for each gene, using a microinjector. Injected embryos were observed, genotyped and raised to recover mutants. We injected equivalent amounts of *tyrosinase* (*tyr*) CRISPR to serve as a positive control for our microinjections (Jao et al. 2013). The *tyr* gene is responsible for pigment synthesis in zebrafish and disrupting this gene would result in decreased pigment synthesis, a phenotype that can be easily

scored at 2 days post fertilization (dpf). We also injected embryos with only Cas9 mRNA to serve as a negative control for our microinjections.

Identification of G0 fish using high resolution melting assay

We performed HRMA after the CRISPR microinjection and while raising injected fish in order to confirm the efficiency of CRISPR injection in G0 embryo or adult fin clips and mutation frequency in F1 embryos respectively. The nature of mutation in F1 embryos, F1 adults, and F2 adults were confirmed by sequencing.

To determine the ability of the CRISPRs to cut at the target site after microinjection, we dechorionated and pooled 10 injected embryos (G0 embryos) from each CRISPR injection at 24 hpf. We extracted genomic DNA from the pool of embryos in a 50 µl DNA extraction buffer (10mM Tris-HCl, 1mM EDTA, 50mM KCl, 0.3% Tween-20, 0.3% NP-40) and 10mg/ml proteinase K solution in a thermocycler (55°C for 2h and 99°C for 5 min).

HRMA was performed on each pool in triplicate. The genomic DNA was used to perform amplification of the target and melt analysis. For this, primers flanking the target locus were designed and used with the MeltDoctor™ high-resolution melting reagents to generate melt curves for our control genomic DNA samples and genomic DNA from the CRISPR-injected G0 embryos. Control samples consisted of i) wildtype uninjected DNA, ii) Cas9-only injected G0 DNA, and iii) DNA from *sp5a*mut:EGFP transgenic embryos. Samples (i) and (ii) served as negative controls, while sample (iii) is a positive control because these embryos have both the wild-type *sp5a* target sequence as well as a corresponding sequence with multiple base changes in the transgene. Once the injected

embryos were confirmed to have sequence variation at the target site, the remaining G0 embryos from the same batch were raised to adulthood.

To identify the G0 adults carrying germline mutations, 3-month-old G0 fish were outcrossed to wild type fish and F1 embryos were genotyped using HRMA and confirmed by sequencing. We used Poisson's distribution to calculate the number of embryos to be screened to be 95% certain of finding at least one F1 mutant embryo, assuming 10% germline transmission. Following this rule, 28 embryos were screened from each G0 outcross in four pools of 7embryos each. The remaining F1 embryos from G0 fish carrying germline mutations were raised. This experiment gave us an estimate of the number of F1 embryos that we needed to raise from the outcross to have at least one male and female F1 mutant fish.

Identification of F1 heterozygous fish using HRMA and sequencing

When the F1 fish were 3 months old, a small portion of their anal fins were clipped carefully after they are temporarily anesthetized using a 10% Tricane solution. Genomic DNA was extracted from the fins of individual G0 fish using a method described earlier. The genomic DNA from each F1 fin clip was used to perform amplification of the target and melt analysis. If a melt curve shift was observed on HRMA plots that was typical of a heterozygous DNA sequence, that F1 fish was assumed to have certain mutation on the target locus and was thus isolated till it was confirmed by sequencing.

In order to perform sequencing, 2 pairs of nested primers were designed for each target locus. The outer primer pair was used to amplify the target region of the gene. The

amplicon from identical reactions were pooled, expected amplicon size was checked on a 1% agarose gel and pooled reactions were cleaned using Zymo DNA Clean & Concentrator™-5. The cleaned amplification products were used as template for sequencing reactions. Inner primer pairs were used for sequencing reactions. The sequences from individual F1 fish were aligned against an aligned annotated genomic wild-type reference and analyzed to confirm the exact nature of mutation.

Identification of homozygous mutants by PCR genotyping

Once the F1 fish were identified, two heterozygous fish carrying the same mutation were crossed and the F2 embryos were observed and raised to obtain homozygous mutants. F2 fish were identified using the PCR genotyping method. For this, the F2 adults were fin clipped to obtain genomic DNA and primers were designed that specifically amplified either the wild type or the mutant sequence from the genomic DNA samples. Therefore, by using these primer sets in two different reactions for every F2 sample, we could identify and isolate homozygous mutants fish for *sp5a* and *sp5l*. The mutation was reconfirmed by sequencing.

Results and discussion

Guide RNA synthesis for sp5a and sp5l genes

We designed *sp5a* and *sp5l* CRISPRs using the CRISPRScan tool online (Figure 7). The CRISPRs selected for the two genes had a high score and targeted exon 2 for each gene. BsaI restriction site sequence were added to the ends of the oligos for subsequent cloning. Utilizing the plasmid based method to clone the crRNA sequence into pDR274 (Figure 8), we successfully cloned the insert between the BsaI sites in the plasmid (Figure 9).

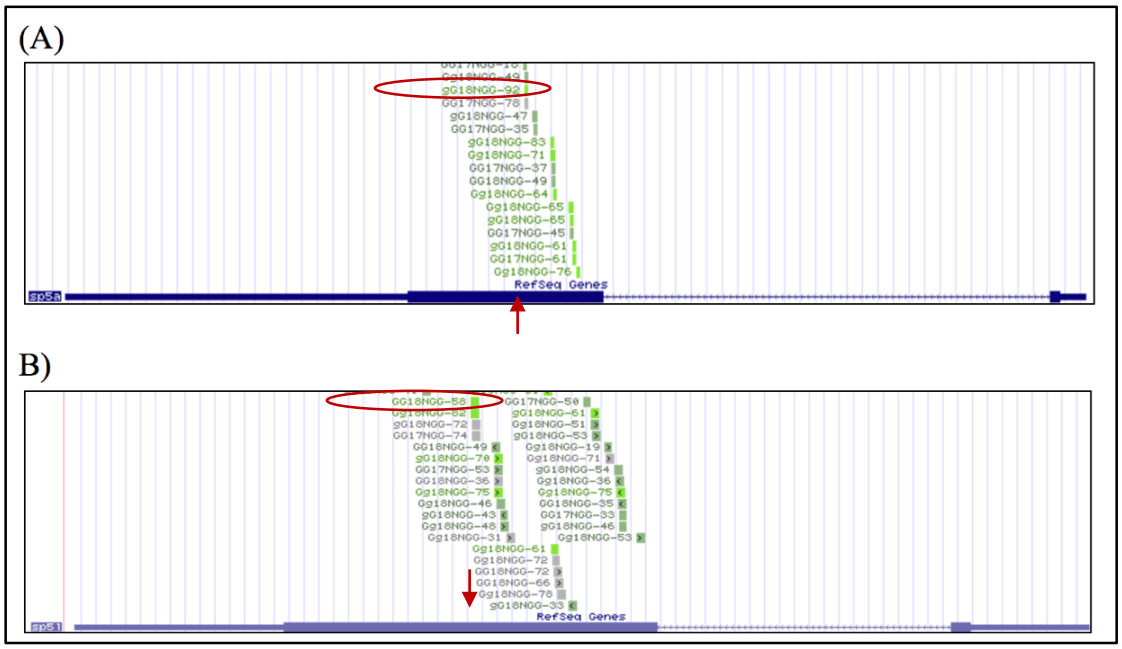

Figure 7. CRISPR target sites. Location of guide RNA target site on Exon 2 (longer exon, total of 2 exons in both genes) of (A) *sp5a* gene and (B) *sp5l* genes. Arrows and circles indicate the location and scores of the CRISPRS. (CRISPRscan UCSC tracks).

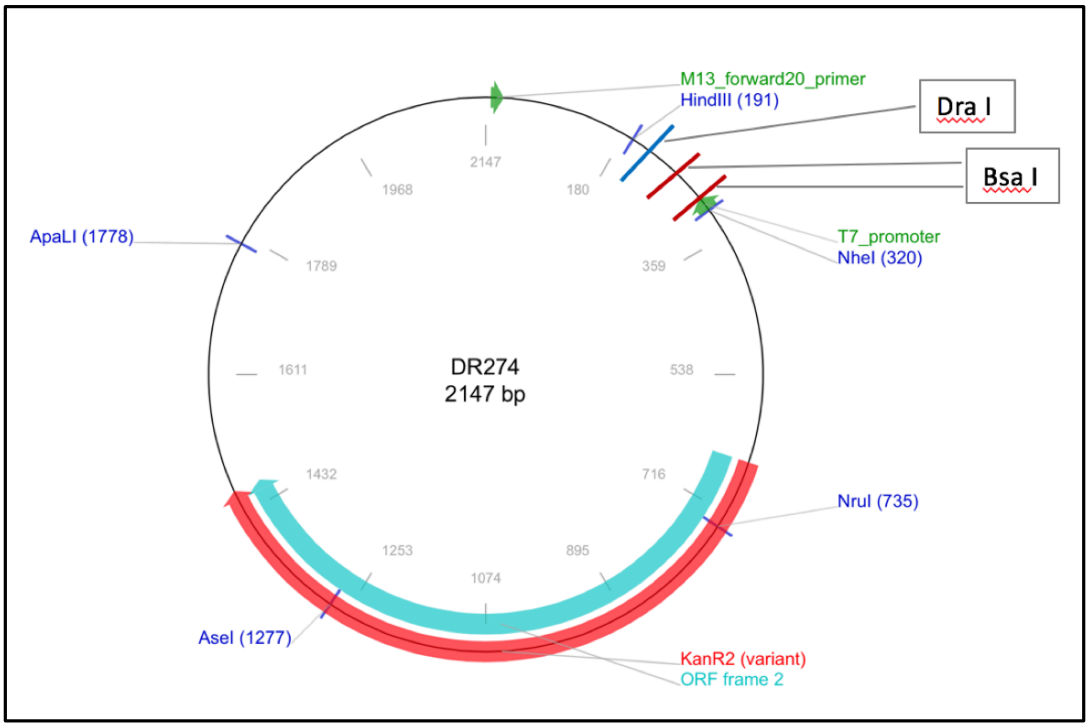

Figure 8. Plasmid map of vector DR274. One DraI site and two BsaI sites are labelled. The plasmid contains a T7 promoter binding site and kanamycin resistance.

The presence of an insert in independent bacterial colonies was confirmed by digestion with BsaI. We isolated ten clones for each guide RNA, which were unaffected by BsaI digestion because the BsaI sites are lost after successful ligation of the insert. Following transcription with T7 polymerase, 1.2% agarose gels of purified *sp5a* and *sp5l* guide RNA were run. A diffused band of about 100 bp length was observed on the gel for each sgRNA. The guide RNA on the gel does not form a sharp band as RNA is less stable and degraded during longer runs. (Figure 10).

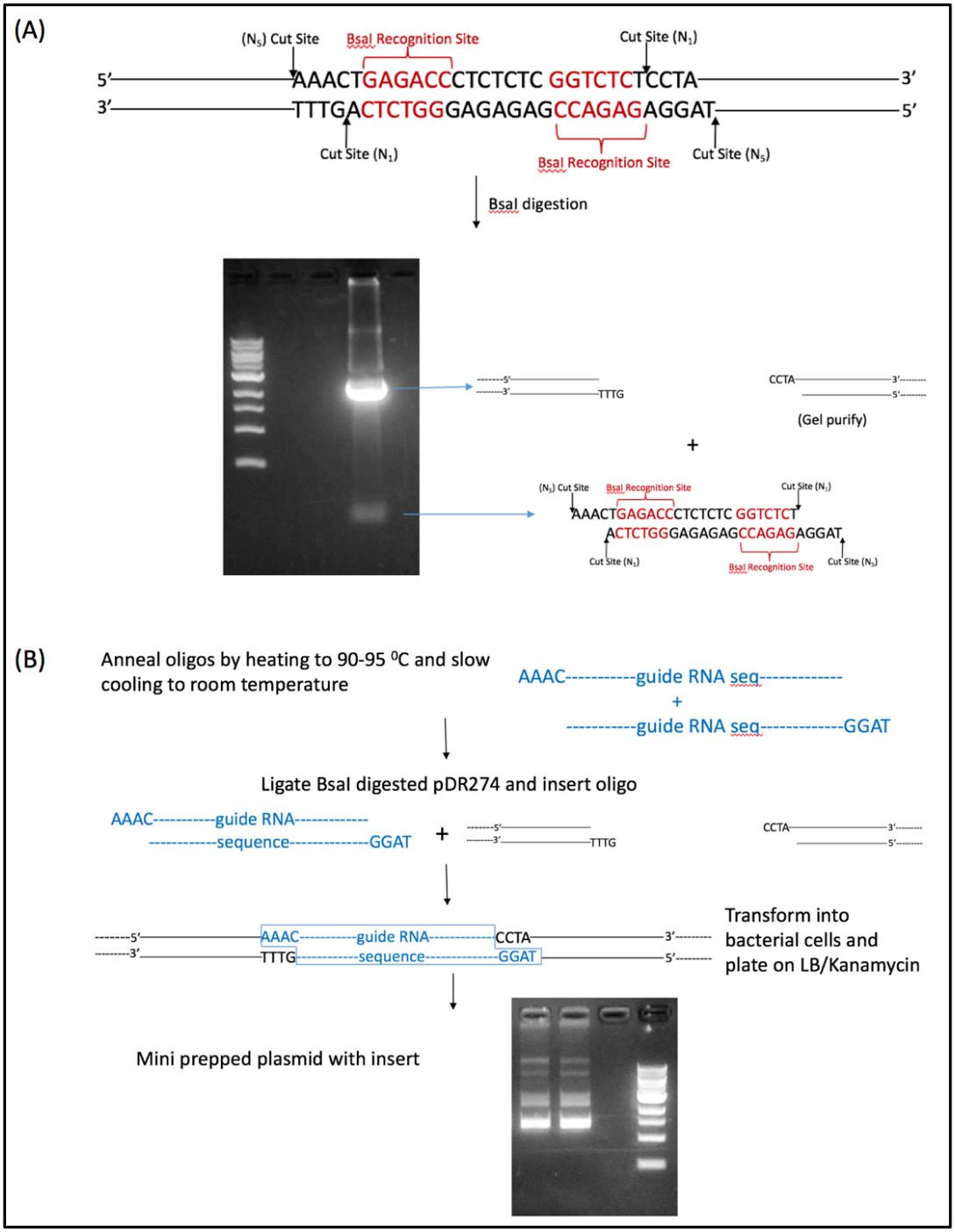

Figure 9. Plasmid based cloning method to construct guide RNA template. (A) Preparation of pDR274 for cloning- BsaI digestion and purification of longer template fragment by gel extraction. (B) The insert was prepared by annealing two 20 base-sgRNA oligonucleotides. The insert was successfully ligated to the vector, transformed and guide RNA template was prepared. 1% agarose gels were run to confirm fragment lengths by comparing with 100bp NEB ladder (right most lane).

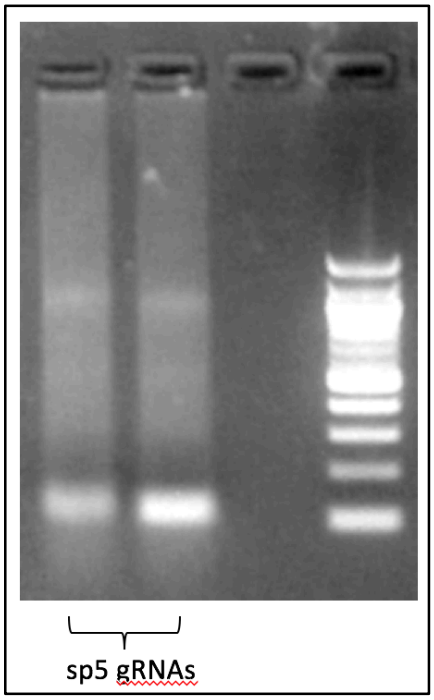

Figure 10. *sp5a* guide RNA synthesis. 1.2% Agarose gel of purified *sp5a* guide RNA run at 150V for 15 min. Diffused bands are seen as RNA is less stable and degrades during longer runs. (Right most lane is 100 bp NEB ladder).

Guide RNA synthesis for sp8b

1% agarose gel electrophoresis confirmed the size of the annealed and purified *sp8b* guide RNA template to be 140 bp (figure 11). 1.2% agarose gel analysis of purified *sp8b* guide RNA, showed a diffuse band of about 100 bp length. A diffused band is seen as RNA is less stable and degraded during the run. (Figure 11).

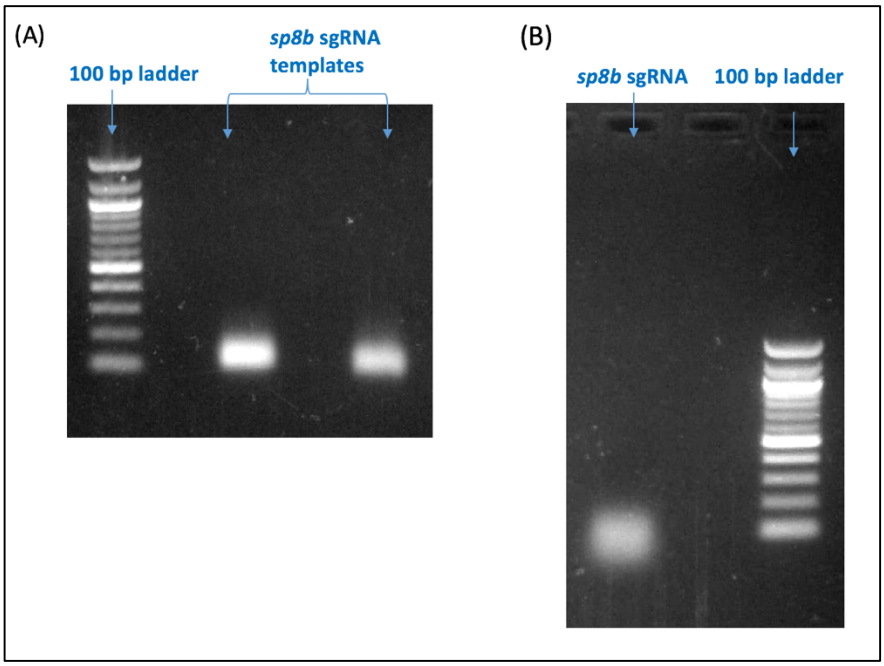

Figure 11. *sp8b* guide RNA synthesis. (A)1% Agarose gel electrophoresis of *sp8b* sgRNA template, (B) 1.2% Agarose gel of purified *sp8b* guide RNA run at 150V for 20 min. A diffused band is seen as RNA is less stable and degrades during longer runs.

Synthesis of cas9 mRNA

The length of the *cas9* mRNA is 4 kb. 1.2% agarose gel of purified mRNA, shows a diffused band through the gel (Figure 12). To test if our synthesized *cas9* worked properly, we co-injected it with a *tyrosinase* sgRNA. 90% of the injected embryos showed some degree of pigmentation loss since the injected embryos are expected to be mosaics with only a proportion of their cells with the *tyr* mutation (Figure 13).

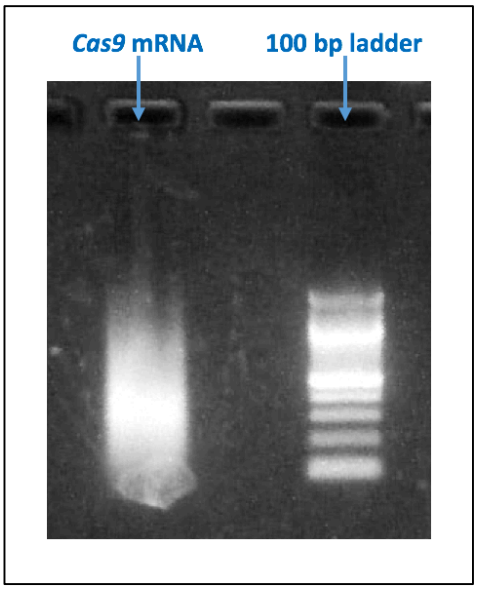

Figure 12. cas9 mRNA synthesis. 1.2% agarose gel of purified cas9mRNA, run at 150V for 20 min.

Microinjection of sgRNA, tyr gRNA and cas9 mRNA into wild type embryos

30 ng/µl of *sp5a* or *sp5l* or *sp8b* sgRNA and 300 ng/µl of *Cas9* mRNA mixed with colored injection dye was injected into wild type embryos at the 1 cell stage. The *sp5a*, *sp5l* or *sp8b* guide RNA injected embryos were observed from time of injection to 24 hpf. No abnormal phenotypes were observed which we hypothesized to be due to mosaicism, since only a proportion of cells in any embryo will be mutant in both alleles (Figure 13-B). Embryos injected with only 300 ng/µl *Cas9* mRNA showed no abnormal phenotype, as expected, proving that *Cas9* mRNA has no deleterious activity when injected alone (Figure 13-E). Embryos injected with 30 ng/µl of tyr sgRNA and 300ng/µl of Cas9 mRNA showed a decrease in pigmentation when observed 30 hpf, proving that our microinjection procedure and *Cas9* mRNA are effective (Figure 13-D).

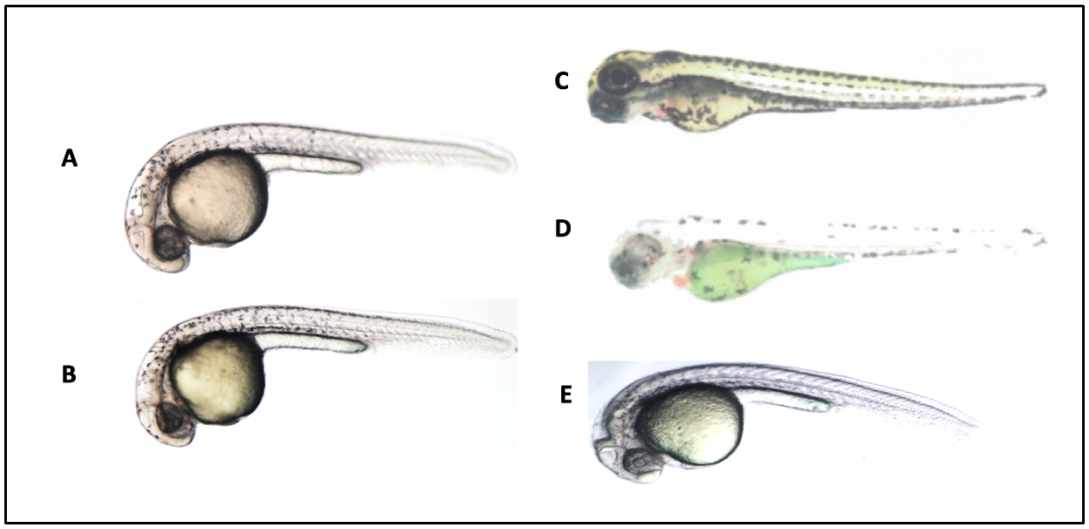

Figure 13. Microinjection of *sp5a* and control CRISPRs. (A,C) Uninjected wild type embryo; wild type embryo injected with (B) 30 ng/µl *sp5a* sgRNA + 300 ng/µl *Cas9* mRNA, (D) 30 ng/µl *tyr* sgRNA + 300 ng/µl *Cas9* mRNA, (E) 300 ng/µl *Cas9* mRNA. (A, B) 30 hpf, (C, D) 3 dpf, (E) 24 hpf.

HRMA of injected G0 embryos

Melting curves from each CRISPR injection showed a shift for the injected sample when aligned with a reference sample. Since the melt curve of the G0 sample was below the reference wild type sample, this suggested that there is most likely a deletion at the targeted locus in at least one out of the ten embryos. Based on this observation, it was implied that the CRISPR creates lesions at the target site and therefore a proportion of the injected G0 embryos would be expected to carry deletion mutations in their germline. The melt curve of the *sp5a*mut:EGFP sample served as a positive control proving that the HRMA reaction setup and reagents worked properly. The melt curve of the *Cas9* injected sample overlapped with the WT reference melt curve, suggesting that *cas9* mRNA

injection alone did not induce any sequence alterations in the injected embryo. (Figure 14).

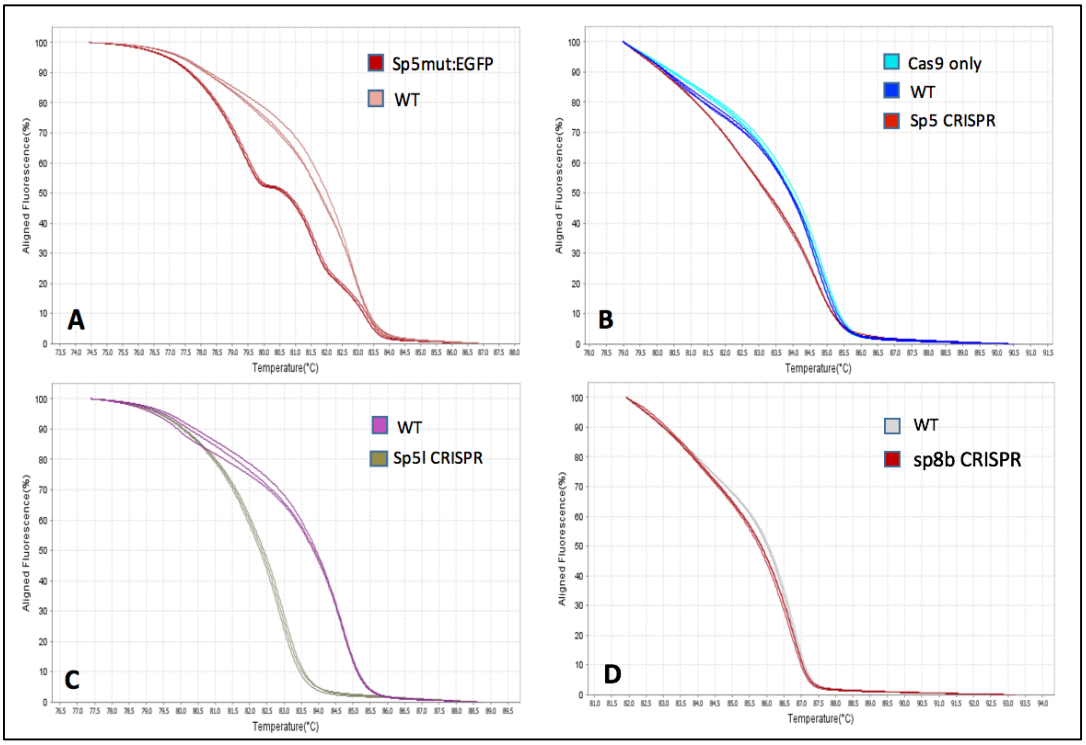

Figure 14. High resolution melting assay of CRISPR injected G0 embryos. Aligned Melt Curves of (A) *sp5a*mut:EGFP positive control (B) *sp5a* CRISPR injected G0 (C) *sp5l* CRISPR injected G0 (D) *sp8b* CRISPR injected G0. X-axis represents aligned fluorescence in percentage and Y-axis represents temperature in degree Celsius.

HRMA of F1 embryos from G0 outcross

Aligned melt curves typical of heterozygous sequence were observed for heterozygous F1 samples (Figure 15), and the exact mutant sequence was determined by sequencing. Using this assay along with sequencing, we identified the the nature of the *sp5a* and *sp5l* mutant alleles.

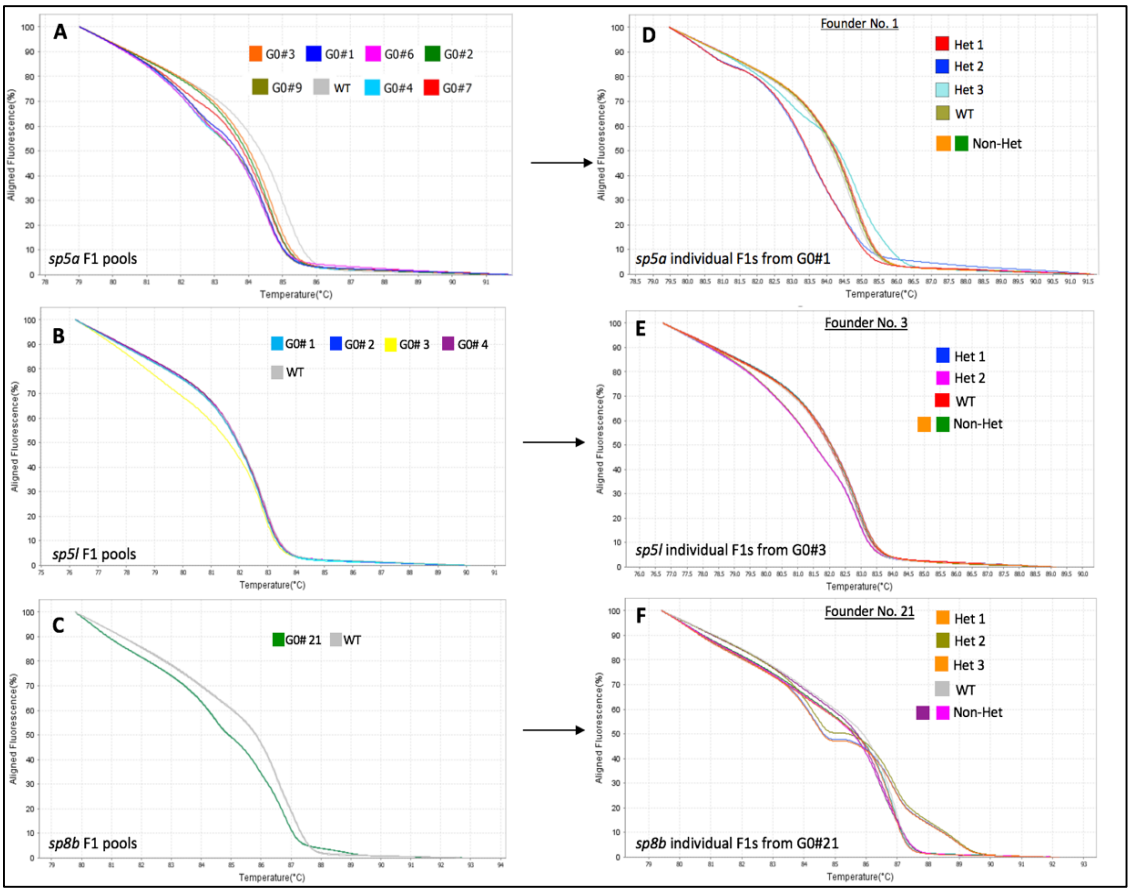

Figure 15. High resolution melting assay of F1 embryo pools and individual embryos from G0 outcross. Aligned melt curves of pools of 7 F1 embryos from (A) *sp5a* G0 fish, (B) *sp5l* G0 fish, (C) *sp8b* G0 fish, and individual embryos from identified (D) *sp5a* G0#1, (E) *sp5l* G0#3, (F) *sp8b* G0#21. X-axis represents aligned fluorescence in percentage and Y-axis represents temperature in degree Celsius.

The mutations in the *sp5a* gene in different heterozygous F1embryos were mostly deletions ranging from 2 bp to 19 bp. Mutations in the *sp5l* gene were also mostly deletions ranging from 2 bp to 16 bp. The only mutation that we found in one out of 28 *sp8b* G0s was a 3 bp deletion (Table 1).

Target gene	Alleles identified
sp5a	Δ 2 bp, Δ 6 bp, Δ 7 bp, Δ 8 bp, Δ 13 bp, Δ 19 bp
sp5l	Δ 2 bp, Δ 3 bp, Δ 5 bp, Δ 8 bp, Δ 16 bp
sp8b	Δ 3 bp

Table 1. List of different mutant alleles identified for *sp5a*, *sp5l* and *sp8b*.

HRMA of F1 adult fish

To identify the adult F1 heterozygotes, we performed HRMA on the genomic DNA samples obtained from anal fin clips of the adult fish. We observed typical heterozygote aligned melt curves for some samples, indicating that the F1 fish carried a mutant allele for the desired gene, which was confirmed by sequencing (Figure 16). We successfully recovered an 8 bp deletion mutant allele of *sp5a* and a 5 bp deletion mutant allele of *sp5l*. Both these deletions are frame shift mutations, so that an early stop appears in the ORF. Both alleles result in a truncated protein lacking more than half of the protein including the zinc finger domains (Figure 17).

We could recover a 3 bp deletion allele of *sp8b*, which leads to the deletion of a single amino acid (a glycine residue) and is not predicted to negatively affect protein function (Figure 17).

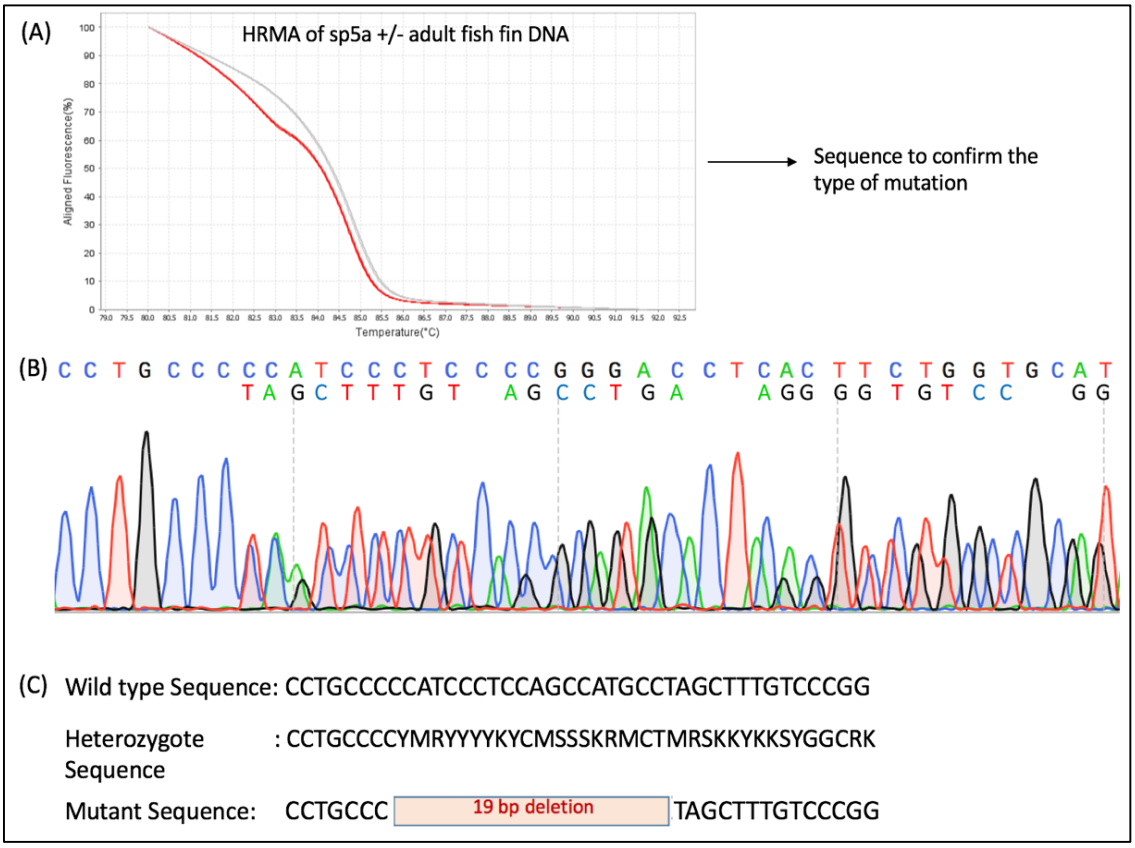

Figure 16. Identifying and confirming a heterozygous mutation. (A) HRMA of F1 adult fish fin clip DNA (red curve- *sp5a*+/- heterozygote, Grey curve- WT reference), (B) Sequencing chromatogram of 40 bp region around the CRISPR target site of *sp5a*+/- fish (overlapping chromatogram is typical of a heterozygous sample, where each base at an overlapping locus represents sequence of one allele, (C) Interpretation of B- *sp5a*Δ19/+ fish. Letters represent defined bases (A, T, C, G) and ambiguous bases (Y= C or T, M = A or C, R= A or G, S= G or C, K= G or T).

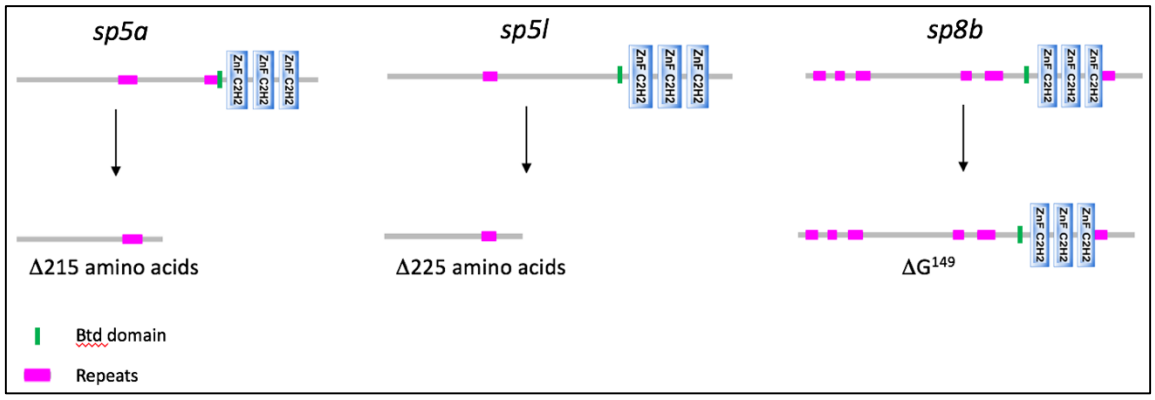

Figure 17. Result of protein truncations in mutant alleles of *sp5a*, *sp5l* and *sp8b*. Frame shift deletion mutation in *sp5a* and *sp5l* results in truncation of C-terminal half of the proteins including the conserved DNA binding zinc finger domains and btd domains. In frame deletion mutation in *sp8b* results in deletion of a single glycine residue at position 149 in Sp8b protein. (SMART web tool)

Generation of homozygous mutants

Identified F1 male and female carriers for the 8bp *sp5a* deletion allele or for the 5 bp *sp5l* deletion allele were intercrossed, and the embryos from this cross were raised to obtain F2 homozygous mutants. From a cross of two heterozygotes, we expect to obtain 25% homozygous mutants. We used PCR based genotyping to identify mutants from F2 adult fin clip samples. About 25% of fish from a *sp5a* F1 intercross were homozygous mutants (Figure 18). The *sp5a* mutants have no observable phenotypes and they develop normally. We are currently maintaining a line of these homozygous mutants. This suggested that this gene might function redundantly with *sp5l* and/or *sp8b* genes.

For *sp5l* gene (figure 19), we could only obtain one homozygous mutant (5% of the total fish genotyped), which implied that this mutation may have some negative effect on the survival of fish post 7dpf, as we did not observe any lethality during that period of

development. *sp5l* homozygous mutants are normal up to 7 dpf, which suggests it may function redundantly with other genes to effect patterning.

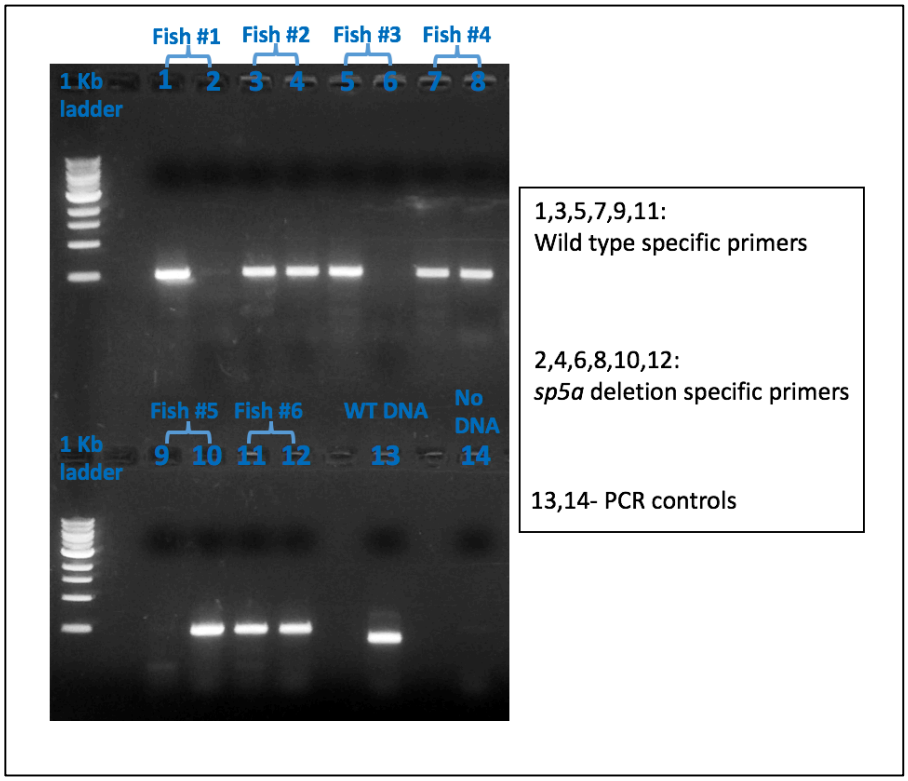

Figure 18. 1% agarose gel electrophoresis of *sp5a* PCR genotyping. Fish #1, #3- WT; Fish #2, #4, #6- *sp5a*Δ8/+ (heterozygotes); Fish #5- *sp5a*Δ8/Δ8 (homozygous mutant).

We were able to identify only one *sp8b* G0 which carried only a 3 bp deletion in the germline, so all the heterozygotes obtained from this G0 outcross had a 3 bp deletion mutation, which is an in frame mutation that is not predicted to alter the reading frame or production of *sp8b* protein.

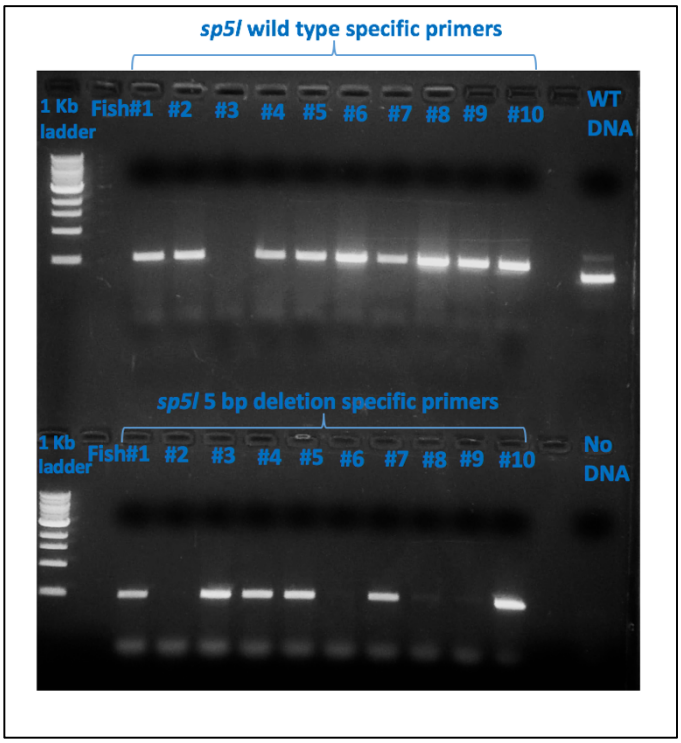

Figure 19. 1% agarose gel electrophoresis of *sp5l* PCR genotyping. Fish #2, #6, #8, #9- WT; Fish #1, #4, #5, #7, #10- *sp5l*Δ5/+ (heterozygotes); Fish #3- *sp5l* Δ5/Δ5 (homozygous mutant).

Summary

We recovered male and female fish homozygous for the *sp5a* 8 bp exon 2 deletion. For *sp5l*, we recovered a homozygous mutant female with a 5 bp deletion in exon 2. These deletions produce frameshift mutations. Early stop codons in the ORF results in predicted formation of a truncated protein lacking almost the entire C-terminal half of the protein, including their zinc finger domains.

Since these two genes encode transcription factors, we expected that deletion of the DNA binding zinc finger domains would result in loss of function alleles. For *sp8b*, we recovered an in frame mutation (3 bp deletion), which results in deletion of one amino acid of Sp8b protein. We expect this mutation to have no impact on the functionality of the Sp8b protein because the amino acid does not lie in DNA binding domain of the protein.

The homozygous *sp5a* or *sp5l* mutants show no phenotypic abnormalities during development and neural patterning. Since these two genes are paralogs, we hypothesized that they act redundantly during developmental patterning. *sp8b* heterozygotes have normal phenotypes because they have one functional copy of gene. We expect the F2 homozygous mutants from the intercross of two sp8bΔ3/+ will appear phenotypically normal, reflecting the predicted normal nature of the mutant protein.

CHAPTER III

FUNCTIONAL ANALYSIS OF MUTANT ALLELES

Redundant and cooperative roles of the *sp5a*, *sp5l* and *sp8b* genes

From our independent mutant analysis of *sp5a* and *sp5l*, we concluded that *sp5a* and *sp5l* mutants, individually, display no phenotype. The lack of visible phenotypes may be due to possible genetic redundancies, for example with *sp8b*. This hypothesis may be supported by previously documented evidence, such as in mouse, where *sp5* and *sp8* function redundantly for neuromesodermal cell development and potentially in anterior neural development (Dunty Jr et al. 2014). It has also been shown that in mouse embryos and differentiating embryonic stem cells that Sp5/8 are gene-specific transcriptional coactivators in the Wnt/β-catenin pathway (Kennedy et al. 2016).

The *sp5a* and *sp5l* genes are paralogous genes which have originated as a result of a duplication event during the teleost-specific whole genome duplication that occurred at the root of the teleost lineage about 350 million years ago (Pasquier et al. 2016). Genetic redundancies arise most often due to gene duplication where loss-of-function mutations of individual genes show weak or no phenotypic effects in the organism (Wagner 1996). Therefore, we hypothesized that these the two *sp5* genes might act redundantly to control some aspect of nervous system patterning. To this effect, we generated double mutants of *sp5a* and *sp5l* and aim to make triple loss of function mutants of *sp5a*, *sp5l* and *sp8b* genes.

Materials and methods

Generation of sp5a/sp5l double mutants

To generate *sp5a*+/-;*sp5l*+/-, we crossed a *sp5a*Δ8/+ female fish with a male *sp5l*Δ5/+ fish, collected and raised all the embryos from this cross. We fin-clipped the adult fish from the above cross and genotyped them using the PCR based method to identify the double heterozygous fish (*sp5a*Δ8/+; *sp5l*Δ5/+). Two *sp5a*Δ8/+; *sp5l*Δ5/+ fish were crossed, the embryos from this cross were raised, observed and genotype to confirm *sp5a*Δ8/Δ8;*sp5l*Δ5/Δ5 fish.

Transient injections to knock down sp5a, sp5l and sp8b genes

We injected *sp5a*, *sp5l* and *sp8b* guide RNAs (30 ng/µl each) together with 300ng/µl of Cas9 mRNA into WT embryos and observed the phenotypes at 24 hpf. This experiment was performed to see the potential effect of triple knockdown.

Results and discussion

Analysis of sp5a/sp5l double mutants

From an intercross of a *sp5a*Δ8/+ fish and *sp5l*Δ5/+ fish, we expect a quarter of the resulting embryos to be double heterozygotes (*sp5a*Δ8/+; *sp5l*Δ5/+). To genotype double heterozygous fish, we used the wild type and deletion-specific primers as described earlier, in four different PCR reactions per fish (Figure 20). We isolated fish which showed fin DNA amplification with the all four set of primers as double heterozygotes for both the genes.

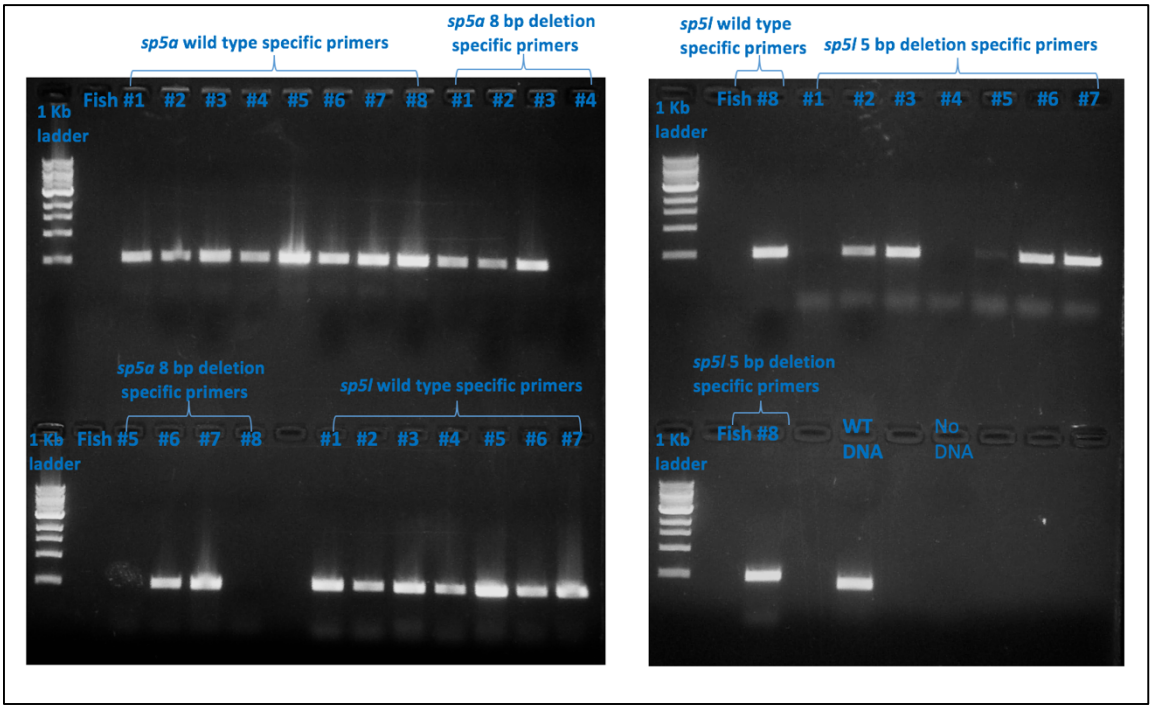

Figure 20. 1% agarose gel electrophoresis of *sp5a* and *sp5l* PCR genotyping. Fish #1-*sp5a*Δ8/+; *sp5l*+/+, Fish #2, #3, #6, #7- *sp5a*-Δ8/+; *sp5l*Δ5/+ (double heterozygotes), Fish #4, #5- *sp5a*+/+; *sp5l*+/+, Fish #8- *sp5a*+/+; *sp5l* Δ5/+

Two identified double heterozygotes were crossed to obtain homozygous double mutants. From a cross of two *sp5a*Δ8/+; *sp5l*Δ5/+ fish, we expect 1 out of 16 embryos (approximately 6% embryos) to be double homozygous mutants. We observed a total of 387 embryos from three independent cross of the double mutant fish and found all embryos to be phenotypically normal. From our observations, we concluded that *sp5a* and *sp5l* double mutants develop normally. One possible explanation could be redundancy with *sp8b*.

sp5a, sp5l and sp8b triple knock out microinjection

We co-injected three CRISPRs into wild type fish to knock down *sp5a*, *sp5l* and *sp8b* together. The injected embryos were mosaics because only a proportion of their cells will have mutations in the three target genes. We observed that 12 out of 44 injected fish (27 % of the injected fish) showed defects in the posterior body development with short body axis and a kinked tail (Figure 21). This indicates that *sp8b* could act redundantly with the *sp5* genes to regulate some aspects of patterning. Since these embryos were mosaic, we might not see an effect on the neural patterning until we have the triple loss of function mutants for these three genes.

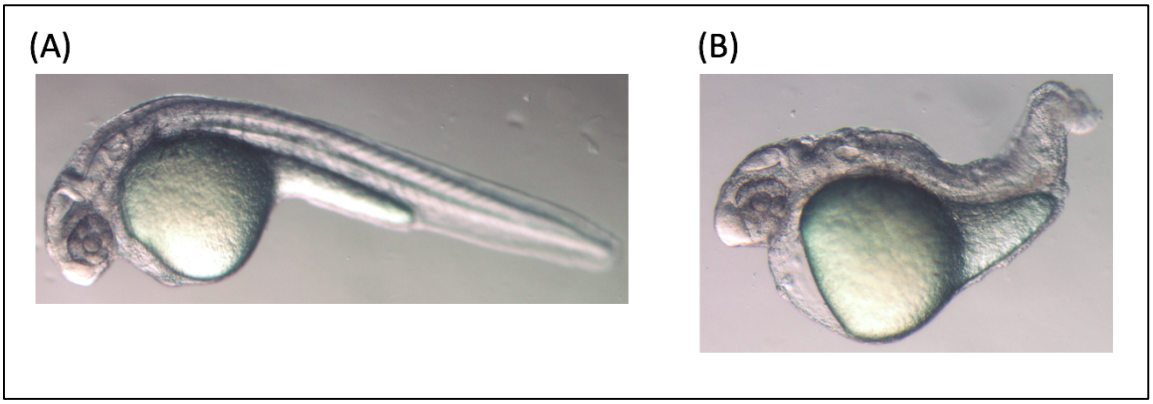

Figure 21. 24 hpf mosaic embryos microinjected with *sp5a*, *sp5l*, *sp8b* sgRNA mix. 30 ng/µl of each sgRNA and 300 ng/µl *cas9* mRNA. Injected embryo with (A) WT phenotype (n=32/44) (B) Posterior body defects (n=12/44).

Summary

Mouse *sp5* homozygotes are known to be viable and fertile with no defects due to functional overlap with *sp8*. This is a potential explanation for the lack of phenotype in our double loss-of-function *sp5/5l* mutants. Mouse *sp5*;*sp8* double mutants have been shown to have posterior body defects resembling wnt3a knockouts (Dunty Jr et al. 2014), which may explain our observation of tail development abnormalities after microinjection of CRISPRs targeting *sp5a*, *sp5l* and *sp8b* genes. To fully understand the role of these genes in Wnt-mediated patterning during development, generation of triple loss of function mutants is required.

sp5a, *sp5l* and *sp8b* are expressed in the neural plate during gastrulation, when the neural plate is being patterned. There is evidence from previous work that *sp5a*, *sp5l* and *sp8b* are Wnt responsive and act downstream of the Wnt pathway to regulate several processes during development (Bell et al. 2003, Zhao et al. 2003, Kawakami et al. 2004, Weidinger et al. 2005). Therefore, we expect a knock out of these three genes to result in patterning defects during zebrafish development.

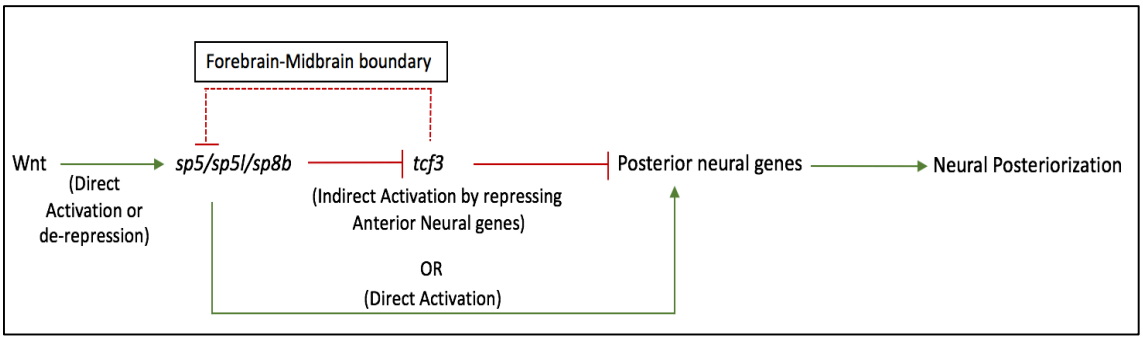

Figure 22. Proposed model for mechanisms by which *sp5a*, *sp5l* and *sp8b* integrate into the Wnt signaling pathway to effect neural posteriorization. Graded Wnt signaling de-represses the Tcf3 repressed posterior neural genes or directly activates *sp5* genes in the posterior neural plate. The *sp5* genes then contribute to the Wnt mediated de-repression of the posterior genes by antagonizing *tcf3* or alternatively activating the posterior neural genes directly.

Based on our observations from the mutant analyses and evidence from previous studies, we hypothesize that Wnt activity gradient determines the level of Sp1 TFs in the neural plate, which in turn induce the neural A/P fate by either directly or indirectly activating posterior neural genes by acting as activator or repressor respectively (Figure 22).

CHAPTER IV

CONCLUSION

This thesis is aimed at understanding the roles of *sp5* and *sp8b* genes in neural patterning during zebrafish development. These genes are targets of the Wnt pathway and potentially play a role in Wnt mediated patterning of the vertebrate nervous system. From our independent zebrafish mutant analysis, we observed that *sp5a*-/- and *sp5l*-/- fish are phenotypically normal. The double mutants, *sp5a*-/-; *sp5l*-/-, do not show any developmental defects.

Based on previous work, one potential explanation for double mutants being phenotypically normal is redundancy between sp5 and *sp8b* genes. Therefore, future work will involve generating homozygous *sp8b* loss of function mutants and triple homozygous mutant fish- *sp5a*-/-; *sp5l*-/-; *sp8b*-/- from the existing lines. Another potential explanation for normal phenotypes of *sp5a*/*sp5l* mutants is genomic compensation through altered mRNA processing which can lead to formation of functional transcripts in mutant lines (Anderson et al. 2017). Therefore, future work would involve analysis of mRNA in the mutant lines by sequencing to confirm that the loss of function phenotype for a mutation is due to a specific genomic alteration (Table 2).

Mutant line	Phenotype	Effect on protein	Explanation
sp5a Δ8/Δ8	Normal	Frame shift mutation- Truncated protein lacking zinc finger domains of sp5a	Potential Redundancy with other *sp* family genes
sp5l Δ5/Δ5	Normal	Frame shift mutation- Truncated protein lacking zinc finger domains of Sp5l	
sp5a Δ8/Δ8; *sp5l* Δ5/Δ5	Normal	Frame shift mutation- Truncated protein lacking zinc finger domains of both genes	
sp8b Δ3/+	Normal	In frame mutation- Mutant protein differs from WT protein by one amino acid	Heterozygous; has one functional copy of gene; Homozygous mutant will have functional protein

Table 2. Summary of mutant lines recovered to date.

REFERENCES

Anderson, J. L., et al. (2017). mRNA processing in mutant zebrafish lines generated by chemical and CRISPR-mediated mutagenesis produces potentially functional transcripts. bioRxiv: 154856.

Augustine, K. A., et al. (1995). Interactions of Wnt-1 and Wnt-3a are essential for neural tube patterning. Birth Defects Research Part A: Clinical and Molecular Teratology **51**(2): 107-119.

Bell, S. M., et al. (2003). Sp8 is crucial for limb outgrowth and neuropore closure. Proceedings of the National Academy of Sciences **100**(21): 12195-12200.

Buckles, G. R., et al. (2004). Combinatorial Wnt control of zebrafish midbrain–hindbrain boundary formation. Mechanisms of development **121**(5): 437-447.

Cadigan, K. M. and R. Nusse (1997). Wnt signaling: a common theme in animal development. Genes & development **11**(24): 3286-3305.

Chylinski, K., et al. (2013). The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. RNA biology **10**(5): 726-737.

Clevers, H. (2006). Wnt/β-catenin signaling in development and disease. Cell **127**(3): 469-480.

Dahlem, T. J., et al. (2012). Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. PLoS genetics **8**(8): e1002861.

Doetschman, T. and T. Georgieva (2017). Gene editing with CRISPR/Cas9 RNA-directed nuclease. Circulation research **120**(5): 876-894.

Domingos, P. M., et al. (2001). The Wnt/β-catenin pathway posteriorizes neural tissue in Xenopus by an indirect mechanism requiring FGF signalling. Developmental biology **239**(1): 148-160.

Dorsky, R. I., et al. (2003). Two tcf3 genes cooperate to pattern the zebrafish brain. Development **130**(9): 1937-1947.

Doudna, J. A. and E. Charpentier (2014). The new frontier of genome engineering with CRISPR-Cas9. Science **346**(6213): 1258096.

Dunty Jr, W. C., et al. (2014). Transcriptional profiling of Wnt3a mutants identifies Sp transcription factors as essential effectors of the Wnt/β-catenin pathway in neuromesodermal stem cells. PloS one **9**(1): e87018.

Eastman, Q. and R. Grosschedl (1999). Regulation of LEF-1/TCF transcription factors by Wnt and other signals. Current opinion in cell biology **11**(2): 233-240.

Erter, C. E., et al. (2001). Wnt8 is required in lateral mesendodermal precursors for neural posteriorization in vivo. Development **128**(18): 3571-3583.

Fujimura, N., et al. (2007). Wnt-mediated down-regulation of Sp1 target genes by a transcriptional repressor Sp5. Journal of Biological Chemistry **282**(2): 1225-1237.

Gagnon, J. A., et al. (2014). Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. PloS one **9**(5): e98186.

Harrison, S. M., et al. (2000). Sp5, a new member of the Sp1 family, is dynamically expressed during development and genetically interacts with Brachyury. Developmental biology **227**(2): 358-372.

Horvath, P. and R. Barrangou (2010). CRISPR/Cas, the immune system of bacteria and archaea. Science **327**(5962): 167-170.

Hwang, W. Y., et al. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. Nature biotechnology **31**(3): 227-229.

Jao, L.-E., et al. (2013). Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. Proceedings of the National Academy of Sciences **110**(34): 13904-13909.

Jinek, M., et al. (2012). A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. Science **337**(6096): 816-821.

Kawakami, Y., et al. (2004). Sp8 and Sp9, two closely related buttonhead-like transcription factors, regulate Fgf8 expression and limb outgrowth in vertebrate embryos. Development **131**(19): 4763-4774.

Kennedy, M. W., et al. (2016). Sp5 and Sp8 recruit β-catenin and Tcf1-Lef1 to select enhancers to activate Wnt target gene transcription. Proceedings of the National Academy of Sciences: 201519994.

Kiecker, C. and C. Niehrs (2001). A morphogen gradient of Wnt/β-catenin signalling regulates anteroposterior neural patterning in Xenopus. Development **128**(21): 4189-4201.

Kim, C.-H., et al. (2000). Repressor activity of Headless/Tcf3 is essential for vertebrate head formation. Nature **407**(6806): 913-916.

Kimmel, C. B., et al. (1995). Stages of embryonic development of the zebrafish. Developmental dynamics. **203**(3): 253-310.

Lekven, A. C., et al. (2001). Zebrafish wnt8 encodes two wnt8 proteins on a bicistronic transcript and is required for mesoderm and neurectoderm patterning. Developmental cell **1**(1): 103-114.

Li, J., et al. (2007). CBP/p300 are bimodal regulators of Wnt signaling. The EMBO journal **26**(9): 2284-2294.

Logan, C. Y. and R. Nusse (2004). The Wnt signaling pathway in development and disease. Annu. Rev. Cell Dev. Biol. **20**: 781-810.

McGrew, L. L., et al. (1997). Wnt and FGF pathways cooperatively pattern anteroposterior neural ectoderm in Xenopus. Mechanisms of development **69**(1): 105-114.

McMahon, A., et al. (1992). The midbrain-hindbrain phenotype of Wnt-1− Wnt-1− mice results from stepwise deletion of engrailed-expressing cells by 9.5 days postcoitum. Cell **69**(4): 581-595.

Moreno-Mateos, M. A., et al. (2015). CRISPRscan: designing highly efficient sgRNAs for CRISPR/Cas9 targeting in vivo. Nature methods **12**(10): 982.

Niehrs, C. (2012). The complex world of WNT receptor signalling. Nature reviews. Molecular cell biology **13**(12): 767.

Nieuwkoop, P. (1952). Activation and organization of the central nervous system in amphibians. Part III. Synthesis of a new working hypothesis. Journal of Experimental Zoology Part A: Ecological Genetics and Physiology **120**(1): 83-108.

Nordström, U., et al. (2002). Progressive induction of caudal neural character by graded Wnt signaling. Nature neuroscience **5**(6): 525.

Northcutt, R. G. (2002). Understanding vertebrate brain evolution. Integrative and comparative biology **42**(4): 743-756.

Nusse, R. and H. E. Varmus (1982). Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. Cell **31**(1): 99-109.

Parant, J. M., et al. (2009). A rapid and efficient method of genotyping zebrafish mutants. Developmental Dynamics **238**(12): 3168-3174.

Park, D. S., et al. (2013). Role of Sp5 as an essential early regulator of neural crest specification in xenopus. Developmental dynamics **242**(12): 1382-1394.

Pasquier, J., et al. (2016). Gene evolution and gene expression after whole genome duplication in fish: the PhyloFish database. BMC genomics **17**(1): 368.

Ran, F. A., et al. (2013). Genome engineering using the CRISPR-Cas9 system. Nature protocols **8**(11): 2281-2308.

Rijsewijk, F., et al. (1987). The Drosophila homology of the mouse mammary oncogene int-1 is identical to the segment polarity gene wingless. Cell 50(4): 649-657.

Šestak, M. S.,et al. (2014). Phylostratigraphic profiles in zebrafish uncover chordate origins of the vertebrate brain. Molecular biology and evolution. **32**(2): 299-312.

Schweizer, L. and H. Varmus (2003). Wnt/Wingless signaling through β-catenin requires the function of both LRP/Arrow and frizzled classes of receptors. BMC cell biology **4**(1): 4.

Sun, Z., et al. (2006). Sp5l is a mediator of Fgf signals in anteroposterior patterning of the neuroectoderm in zebrafish embryo. Developmental dynamics **235**(11): 2999-3006.

Swarts, D. C., et al. (2012). CRISPR interference directs strand specific spacer acquisition. PloS one **7**(4): e35888.

Takahashi, M., et al. (2005). Identification of SP5 as a downstream gene of the β-catenin/Tcf pathway and its enhanced expression in human colon cancer. International journal of oncology **27**(6): 1483-1487.

Talbot, J. C. and S. L. Amacher (2014). A streamlined CRISPR pipeline to reliably generate zebrafish frameshifting alleles. Zebrafish **11**(6): 583-585.

Tallafuß, A., et al. (2001). The zebrafish buttonhead-like factor Bts1 is an early regulator of pax2. 1 expression during mid-hindbrain development. Development **128**(20): 4021-4034.

Thisse, B. and C. Thisse (2004). Fast release clones: a high throughput expression analysis. ZFIN direct data submission **2**.

Thomsen, N., et al. (2012). High resolution melt analysis (HRMA); a viable alternative to agarose gel electrophoresis for mouse genotyping. PloS one **7**(9): e45252.

Toivonen, S. and L. Saxén (1968). Morphogenetic interaction of presumptive neural and mesodermal cells mixed in different ratios. Science **159**(3814): 539-540.

Treichel, D., et al. (2001). The novel transcription factor gene Sp5 exhibits a dynamic and highly restricted expression pattern during mouse embryogenesis. Mech Dev **101**(1): 175-179.

Villanueva, S., et al. (2002). Posteriorization by FGF, Wnt, and retinoic acid is required for neural crest induction. Developmental biology **241**(2): 289-301.

Wagner, A. (1996). Genetic redundancy caused by gene duplications and its evolution in networks of transcriptional regulators. Biological cybernetics **74**(6): 557-567.

Weidinger, G., et al. (2005). The Sp1-related transcription factors sp5 and sp5-like act downstream of Wnt/beta-catenin signaling in mesoderm and neuroectoderm patterning. Curr Biol **15**(6): 489-500.

Wittwer, C. T. (2009). High-resolution DNA melting analysis: advancements and limitations. Human mutation **30**(6): 857-859.

Woo, K. and S. E. Fraser (1997). Specification of the zebrafish nervous system by nonaxial signals. Science **277**(5323): 254-257.

Wu, J., et al. (2005). Neural crest induction by the canonical Wnt pathway can be dissociated from anterior–posterior neural patterning in Xenopus. Developmental biology **279**(1): 220-232.

Yu, C., et al. (2014). A PCR based protocol for detecting indel mutations induced by TALENs and CRISPR/Cas9 in zebrafish. PloS one **9**(6): e98282.

Zhao, J., et al. (2003). An SP1-like transcription factor Spr2 acts downstream of Fgf signaling to mediate mesoderm induction. The EMBO journal **22**(22): 6078-6088.