Understanding the Establishment of TWIST Neural Crest Gene Regulatory Systems and Their Evolutionary Significance in Chordates

By: Austin Wolfgang Katzer

Department of Ecology and Evolutionary Biology, University of Colorado Boulder

A Thesis Submitted to the Honors College of The University of Colorado Boulder

Defended on April 6th, 2022

Thesis Advisor:

Daniel Meulemans Medeiros, PhD, Department of Ecology and Evolutionary Biology

Committee Members:

Katharine Anne Semsar, PhD, Miramontes Arts and Sciences Program

Barbara Demmig-Adams, PhD, Department of Ecology and Evolutionary Biology

Abstract

The evolution of the vertebrate head in chordate organisms involved genetic mutations that permitted new gene interactions of the existing developmental toolkit. Compared to other phyla, the vertebrate head contains complex somatosensory receptors, as well as innovative mechanical and sensory structures. The emergence of these structures resulted from the evolution of the neural crest, a transient cell population exclusive to vertebrates, with qualities such as pluripotency, the capacity for a cell to give rise to other cell types, as well as the ability to delaminate and migrate. This ectodermal cell type can give rise to a multitude of adult structures including those that are putative mesoderm derivatives, such as craniofacial bone and cartilage. This project focused on the evolutionary aspect of the neural crest in chordate organisms addressed the question of how the neural crest emerged. The *twist* gene is essential in chordate development and encodes a basic transcription factor that plays an important role in the development of embryos. An experiment was conducted to determine what kind of mutations may have integrated twist in the neural crest gene network. This project focuses specifically on mutations within *twist* cis-regulatory elements (CREs), non-coding DNA sequences that regulate transcription of genes, that may have facilitated its involvement in neural crest cells. Our current understanding is that the CREs controlling *twist* expression in neural crest cells are different from those regulating *twist* expression in other regions within the developing organism. However, whether they are completely new or emerged from existing enhancer sequences is unknown. A new CRE would consist of a novel sequence that is not present in invertebrate chordates and only acts in neural crest cells. Alternatively, an existing CRE could have been modified to permit twist expression in a novel tissue other than the ancestral mesodermal pattern. By identifying putative CREs using conserved teleost sequences outside of the coding region for

twist it is possible to functionally test these elements using standard transgenic techniques. The putative cis-regulatory element sequences of around 1000-2000 base-pairs in length were tested functionally within zebrafish using the gateway cloning method to introduce the fragments into the reporter construct pGreene, which has a cFos basal promoter and eGFP flanked by Tol2 recombination arms. Injecting these constructs along with transcriptase into single-cell embryos highlights *twist* CRE activity by driving fluorescent signals in a given region of the embryo revealing how the neural crest gene network was co-opted, the expression of new function derived from an old gene.

1. Introduction

1.1 Background

As chordate embryos develop, they reach a phylotypic stage where the embryos appear highly similar phenotypically to several other species and other genetically distant organisms found within the same phylum (Gans & Northcutt, 1983). Although, later on the developed organisms take on highly diversified morphologies as chordate evolution progressed, vertebrates diverged. The evolution of the head in chordate organisms involved several diverse and complex mutations unlike any others when compared to different phyla with a central area containing many somatosensory receptors (Gans & Northcutt, 1983). Head morphologies include a variety of innovative mechanical and sensory structures which are thought to, collectively, be an underlying source of the immense diversification within this abundant lineage. The emergence of these structures is thought to have resulted from the evolution of the neural crest, a transient cell population exclusive to vertebrates, with qualities such as pluripotency as well as the ability to delaminate and migrate (Theveneau & Mayor, 2014). This ectodermal cell type can give rise to a

multitude of adult structures including those that are putative mesoderm derivatives, such as craniofacial bone and cartilage (Theveneau & Mayor, 2014).

While there is much known about neural crest development, the genetic means through which it evolved remain unclear. This project focuses on the evolutionary aspect of the neural crest in chordate organisms and tries to help solve the question of how the neural crest emerged. The *twist* gene is very important in chordate development and encodes a basic transcription factor that plays an important role in the development of embryos (Gitelman, 2007). The *twist* gene is thought to have played a major role in chordate development and evolution as a mutation in the genome gave rise to the ability for neural crest cells to migrate in the neural border due to the activation of the *twist* gene in these regions (Gitelman, 2007; Horie et al., 2018). The main focus of the experimental approach of this thesis is determining what kind of mutation activated the *twist* gene in the neural border, and what may have happened in *twist* cis-regulatory sequences that gave way to this mutation. Cis-regulatory elements (CREs) are regions of noncoding DNA that regulate the transcription of genes surrounding them (Wittkopp and Kalay, 2011). Cis-regulatory elements are very important in the management of gene expression and in turn the morphogenesis crucial in the process of embryonic development (Wittkopp and Kalay, 2011).

1.2 Neural Crest and the Bigger Picture

Neural crest cells (NCCs) are an embryonic cell type specific to vertebrates that arise from the neural plate border during neurulation, then migrate throughout the body (Huang & Saint-Jeannet, 2004). These NCCs are responsible for the development of many diverse cell lineages including craniofacial cartilage and bone, melanocytes, smooth muscle, glial cells, and

various neuron types (Huang & Saint-Jeannet, 2004). NCCs are viewed as crucial in the evolution of vertebrates particularly for their role in craniofacial development as many of the features occurring only in vertebrates are concentrated in the head (Gans & Northcutt, 1983). The diversity of facial structures in members of the vertebrate subphylum is an adaptive advantage that presumably directly contributed to their success. Assessing how the neural crest evolved, as a common feature of the vertebrate subphylum, is a natural first step (Huang, 2004). One example of an organism that may provide further insight into vertebrate evolution is *Ciona intestinalis*, a tunicate in the phylum chordata that is not a vertebrate based on its lack of NCCs (Horie et al., 2018). However, presence of a neural crest derivative in *Cionia* is consistent with the evolutionary importance of NCCs and raises questions about the development of gene regulatory networks that resulted in the evolution of the neural crest (Horie et al., 2018).

1.3 *twist* and its evolutionary significance

The *twist* gene is an important regulatory gene that codes for a transcription factor in embryonic development (U.S. National Library of Medicine, 2022). The TWIST transcription factor binds to regulatory gene enhancers that facilitate cellular differentiation of mesodermal tissue (U.S. National Library of Medicine, 2022). *Twist* is conserved from jellyfish to humans (U.S. National Library of Medicine, 2022), which highlights its importance in embryonic development through immense evolutionary conservation. Although *twist* plays an important developmental role, it is not a defining feature of the vertebrate subphylum (Gitelman, 2007). This thesis focuses on *twist* not only due to its importance in embryonic development, but even more so for its unique role in the evolution of vertebrate novelties. Normally expressed in the mesoderm during early development, twist exhibits ectodermal expression in vertebrates that is

linked to activation of neural crest-specific factors at the neural plate border (Meulemans & Bronner-Fraser, 2004). This occurrence appears to be a vertebrate novelty, which makes *twist* a gene of interest as a driver of vertebrate evolution (Meulemans & Bronner-Fraser, 2004). Revisiting the NCC-like mechanisms found in *Ciona intestinalis*, the reprogramming of *twist* into the already present "cephalic melanocyte lineage" results in a migratory melanocyte cell type behaving very similarly to NCCs (Horie et al., 2018). The development of neural crest acting cells from the addition of *twist* to *Cionia* suggests that co-option of *twist* expression into the neural crest gene regulatory system may have resulted in emergence of the vertebrate head and its ultimate high evolutionary success (Abitua et al., 2012).

1.4 Cis-Regulatory Elements and their potential as an evolutionary driver

My thesis focuses exclusively on cis-regulatory elements and their method for gene regulation. Cis-regulatory elements (CREs) are non-protein-coding regions of DNA that regulate the transcription of genes and fall into three main categories consisting of promoters, enhancers, and silencers (Wittkopp & Kalay, 2011). Promoters are found directly upstream of genes and are responsible for initiation of transcription for their respective gene targets (Wittkopp & Kalay, 2011). Enhancers and silencers increase or decrease, respectively, expression of a gene (Wittkopp and Kalay, 2011). The ability of CREs to cause changes in gene expression at varying levels is a driving force for morphological changes on spatial and temporal scales as well as evolutionary novelties. The concept of gene co-option is an explanation for spatial and temporal changes CREs cause on an evolutionary timescale (Prud'homme, 2007). My thesis focuses on the ability for CREs to alter expression on a spatial and temporal scale specific to that of *twist* and its co-option into the neural crest gene-regulatory system.

1.5 Research Questions

1.5.1 Problem: The main focus in the experiment is determining what kind of mutation activated the *Twist* gene in the neural border, and what may have happened in TWIST cis-regulatory sequences that gave way to this mutation.

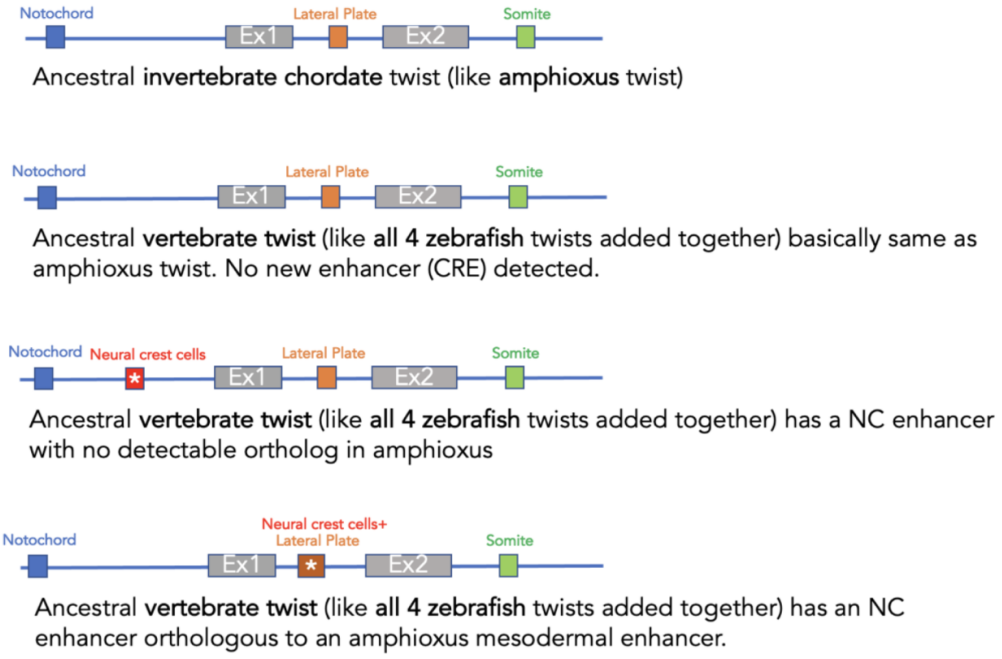

Figure 1: A visual hypothesis demonstrating a graphical example of the Ancestral chordate DNA sequence and its proposed evolutionary possibilities resulting in twist being expressed in the ectodermal layer

Figure 1 was acquired from Daniel Medeiros, PhD.

1.5.2 Hypotheses:

Change in Trans (null hypothesis) - No cis-regulatory element evolution occured. Potentially trans-regulation was at play and a trans-regulator of twist moved to the ectodermal layer and transferred the twist gene along with it.

De novo CRE hypothesis - A completely new neural crest cis-regulatory element appears near

the twist gene, perhaps by transposition or accumulated point mutations, resulting in its expression in the ectoderm becoming a part of the neural crest gene regulatory system.

Divergence hypothesis - A mutation in an old cis-regulatory element introduces twist expression in the neural crest gene regulatory network, while still maintaining its expression in the old mesodermal domain.

2. Material and Methods

2.1 Zebrafish as a model organism

The zebrafish, *Danio rerio*, was chosen for this study for several reasons, including availability of techniques for large-scale genome mutagenesis and gene mapping, transgenesis, and protein overexpression have increased the power of this model organism (Veldman, Lin., 2008). The clear embryos of *D. rerio* allow for easy microinjection, and external development further simplifies the process. Their very rapid embryonic development and time required to reach sexual maturity allow for easy transgenic studies and development of transgenic lines. Specific to identification of putative CREs, the zebrafish genome is exceptionally well mapped and its overall popularity decreases the cost for complex genetic studies.

2.2 Identification of putative Cis-regulatory elements

In order to visualize upstream and downstream areas of the *twist1b* gene in zebrafish, the University of California Santa Cruz genome browser's compiled data and interface was used looking at specific DNA tracks utilizing "DANIO-CODE: consortium for the functional annotation of the zebrafish (Danio rerio) genome" (Zfin, 2014). The data analyzed was sorted by two sequencing tracks, chromatin immunoprecipitation (ChIP) sequencing and Assay for

Transposase-Accessible Chromatin (ATAC) with high-throughput sequencing. ChIP sequencing analyzes interactions between proteins and DNA, with histones being the proteins in focus. By analyzing where histones have loosely wound DNA, these regions represent locations where transcription factors could potentially be attaching themselves to CREs within these sections of DNA. This means that the ChIP sequenced locations are likely hotspots of activity for cis-regulatory elements. The regions of DNA specifically focused on included: H3K27ac, H3K4me3, and H3K4me1 at specific times during embryonic development. These times these locations were sequenced included: bud, 5-9 somites, Prim-5, and Long-pec stages. Through a zebrafish genome alignment overlaid with ATAC and ChIP sequencing data, putative cis-regulatory elements upstream and downstream of the target gene were selected by hand. The selected fragments were attempted to be synthesized via the gene synthesis service "Twist Biosciences" , but many did not qualify for synthesis due to fragment length (maximum of 1800 BP) and long chains of repeating sequences. CRE1, CRE3, and CRE7 were successfully synthesized by Twist Biosciences, and primers were ordered for the rest.

2.3 Sample template, Gel Electrophoresis, and Purification

After the identification of putative cis-regulatory elements, a sample template of zebrafish genomic DNA was collected through a Phenol-chloroform DNA extraction of zebrafish tissues. In order to make DNA fragments of the desired sequences, synthetic DNA oligonucleotide primers were designed to flank the region of interest for each particular sample. These synthetic primers anneal to their antiparallel complementary sequence during polymerase chain reaction (PCR) resulting in an amplification of the desired DNA region. The primers for this experiment were ordered from Sigma-Aldrich chemicals company. Each DNA fragment was

given the addition of attB binding sites in order to be cloned into the pGreenE plasmid. After PCR amplification of the desired sequences, gel-electrophoresis of the PCR products were performed and the resulting bands that coincide with the desired length of DNA fragments are cut out of the agar gel and purified. The gel purification process was conducted using the Wizard SV Gel and PCR Clean-Up System according to manufacturer protocols.

2.4 DNA Purity and Fragment Sequencing I

After DNA purification is completed utilizing the Wizard SV Gel and PCR Clean-Up System, the purified samples were tested using a Thermo Scientific NANODROP 2000 spectrophotometer to verify their purity and concentration. The verified purified DNA samples need to be sequenced in order to determine if the resulting PCR product is the intended target fragment. The resulting fragments often range between 1,000 to 2,000 base pairs in length and are sent off to Qiagen to be sequenced and reported on.

2.5 Gateway Cloning

Once DNA fragments are verified as the desired sequences, they need to be inserted into a vector for injection prep. DNA fragments are ligated into a pGreenE bacterial vector through the use of DH5 alpha competent E. coli cells. The steps below indicate the ligation process:

1. PCR in 50 ul, ideally with Phusion, GoTaq is also fine to use
2. Gel extract, elute into 30 ul TE buffer
3. pGreenE, 37.5 ng
4. Add PCR eluate up to 2 ul
5. Add 0.5 ul BP Clonase
6. Incubate overnight at room temperature
7. Add 0.5 ul Proteinase K (from the clonase kit)
8. Incubate 10 min at 37°C overnight
9. Transform into DH5 alpha cells
10. Plate on Ampicillin agarose gel plates, 37°C overnight

After the colonies are left to grow overnight, two colonies were selected from each plate sample and set up to grow again overnight at 37°C in a shaking incubator. The resulting colonies go through Omega BIO-TEK E.Z.N.A. Plasmid DNA Miniprep Kit per manufacturer protocols.

2.6 DNA Purity and Fragment Sequencing II

Following the purification of plasmids through Omega BIO-TEK E.Z.N.A. Plasmid DNA Miniprep Kit per manufacturer protocols, the purified plasmids were tested using a Thermo Scientific NANODROP 2000 spectrophotometer to verify their purity and concentration. The verified purified plasmids need to be sequenced in order to determine if the resulting products from the Gateway cloning process successfully incorporated the desired DNA fragments into the plasmids. The plasmids are sent off to Qiagen to be sequenced and reported on.

2.7 Single-cell Embryonic Injection

In order to functionally test the putative cis-regulatory elements, an in vivo study needs to be conducted through the insertion of the verified vector into a single cell zebrafish embryo. The cocktail injected into the single cell embryos contains three ingredients, transposase, phenol red, and the sequence verified vector. Transposase is an enzyme that recognizes the tol2 sites in the pGreenE plasmid and cuts the plasmid at those two loci and flips it into the genome of the injected cell. Phenol red is a pigment that allows for easy visibility of the cocktail and approximate measurement when injecting it into the embryo based on droplet size. The cocktail is injected into single cell zebrafish embryos using a Picospritzer III device and a pulled glass needle.

2.8 GFP Imaging of Embryos

Following injection of fertilized zebrafish eggs, embryos are analyzed every 12 hours for expression of Green Fluorescent Protein (GFP) which is a visible marker of selected putative cis-regulatory element expression both spatially and temporally. Embryos were anesthetized using 1 ml of tricaine per dish at a concentration of 4g/L. Following anesthetization, embryos were suspended in a methylcellulose solution on a glass microscope slide and observed under a Zeiss Discovery.V8 SteREO Dissecting Microscope and illuminated with an EXFO X-Cite Series 120Q Ultraviolet laser. The resulting visible expression of GFP was photographed using a Zeiss AxioCam MRc5 5 megapixel color microscope camera.

3.0 Results

3.1 Identification of putative Cis-regulatory elements and Gel Electrophoresis

After analyzing the compiled ChIP and ATAC sequencing data provided from the Zfin Danio Code project, numerous putative CREs were identified, and seven were ultimately selected to be functionally tested *in vivo*.

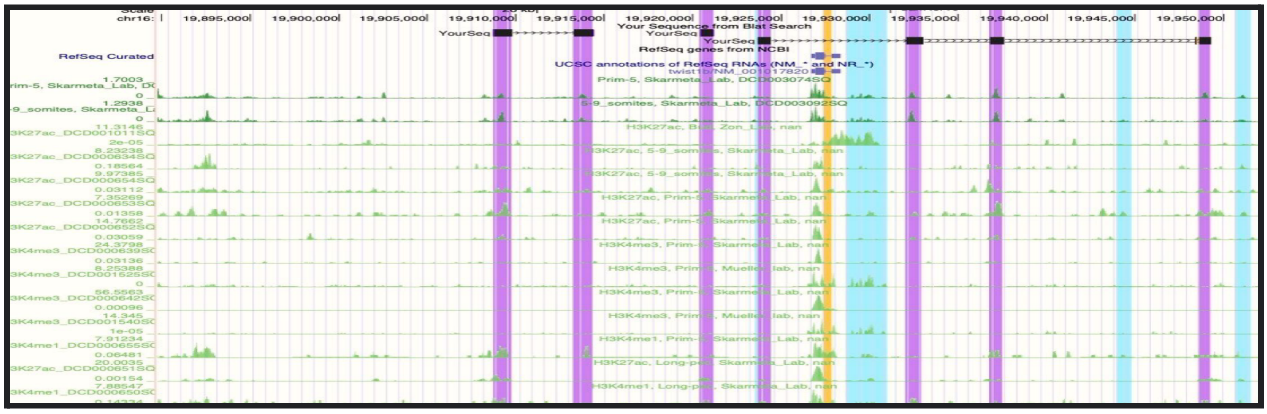

Figure 2: Screen capture from the UCSC Genome Browser illustrating the zebrafish twist 1b gene with the gold color highlighted column and multiple putative cis-regulatory sequences both upstream and downstream of the gene. The green horizontal tracks are an example of the Histone acetylation markers with peaks demonstrating areas where there is loosely bound DNA.

Putative cis-regulatory sequences selected ranged from around 1000-2000 base pairs in length and were denoted CRE1 through CRE7. CRE1 is 1742 base pairs in length and can be located on the UCSC genome browser utilizing the BLAT function through the marker "chr16:19907971-19909712". CRE2 is approximately 1200 base pairs in length and can be located at "chr16:19913310-19914473". CRE3 is 1500 base pairs in length and can be located at "chr16:19919841-19921283". CRE4 is 1500 base pairs in length and can be located at "chr16:19926457-19928051". CRE5 is 1900 base pairs in length and can be located at "chr16:19928651-19930498". CRE6 is 1200 base pairs in length and can be located at "chr16:19931943-19933030". CRE7 is 1298 base pairs in length and can be located at "chr16:19936499-19937738".

3.2 Injection, Development, and Imaging

Injected fish exhibited variation of expression and appeared to be consistent with that of NCC expression. Despite some background, or apparent non-specific, expression unrelated to that of the selected putative CREs in focus, images also show interpretable features . GFP expression was not seen until around 24 hours post fertilization, and was only seen in putative CRE3, CRE6, and CRE7 (refer to figures 3, 4, and 5).

Figure 3 shows a zebrafish embryo injected with putative CRE6 and indicates potential specific expression of *twist* in the neural crest, particularly near the notochord (labeled Nc) and in the frontonasal process (labeled FnP). The ovular links starting at the arrowhead and moving upward along the trunk toward the head show GFP expression outlining vertebrae in the embryo (Figure 3).

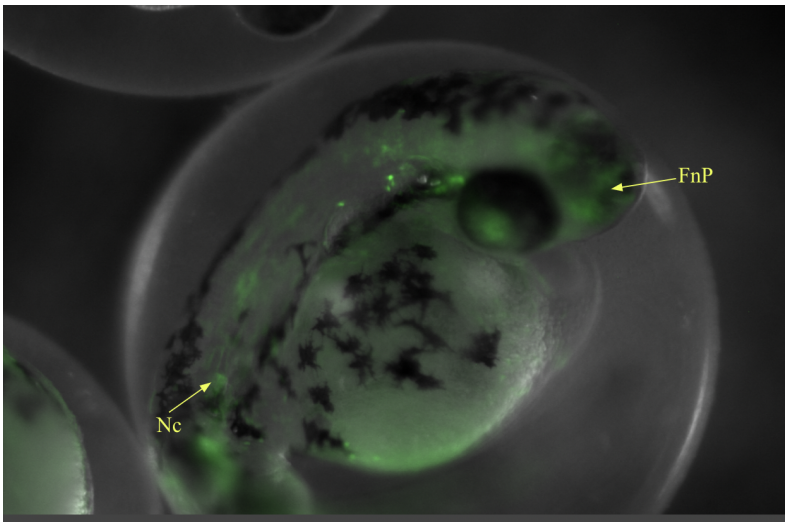

Figure 3: A zebrafish embryo injected with putative CRE6 imaged 24 hpf displaying GFP expression. The arrow labeled Nc is pointing at potential expression alongside the notochord, and FnP is highlighting potential expression in the frontonasal process. Both of these are consistent with NCC activity.

C and A in figure 5 both demonstrate potential neural crest expression around the eyes, particularly in C. This could be a non-specific expression, but the complete surrounding of the eye in C appears much more like that of NCC expression than if it were random based on the patterning seen. B and D both show potential NCC expression in the frontonasal process, with expression looking particularly strong in D, which could be due to the image angle when compared to B. These embryos earlier on in development may display more active NCC expression when compared to more developed embryos where expression may be residual. Figure 6 displays a unique image where the zebrafish is angled so that the lower part of the head is visible, showing off the bottom jaw. In this case there appears to be a very visible GFP expression in the jaw meaning potential NCC expression is visible. Neural crest cells are responsible for craniofacial bone and cartilage development which suggests that there is a high potential that this expression is indeed NCC. Outside of specific NCC expression seen in the

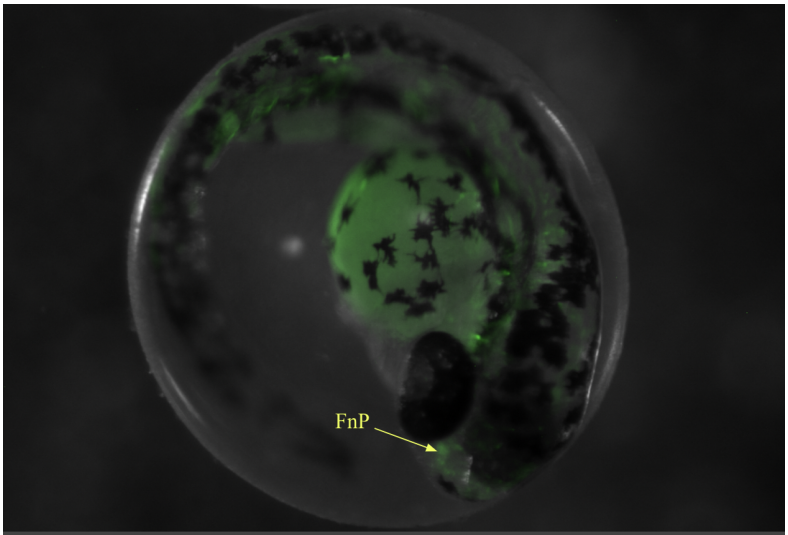

Figure 4: A zebrafish embryo injected with putative CRE3 imaged 24 hpf displaying GFP expression. FnP is highlighting potential expression in the frontonasal process. Both of these are consistent with NCC activity.

developing zebrafish embryos, there was a high potential for *twist* expression based on two main tissue types that had recurring expression. When thinking about NCCs these expression patterns such as those in figures 7 and 8 would be dismissed as non-specific, but in the case of *twist* it is not only involved in the neural crest gene regulatory system, but also plays a large role during its mesodermal expression.

Both muscle and vertebrae development are related to *twist* expression in the mesoderm meaning that there is visible twist expression likely occurring in these regions, but the CREs selected may not be ones that play a role in *twist* involvement in the neural crest gene regulatory system.

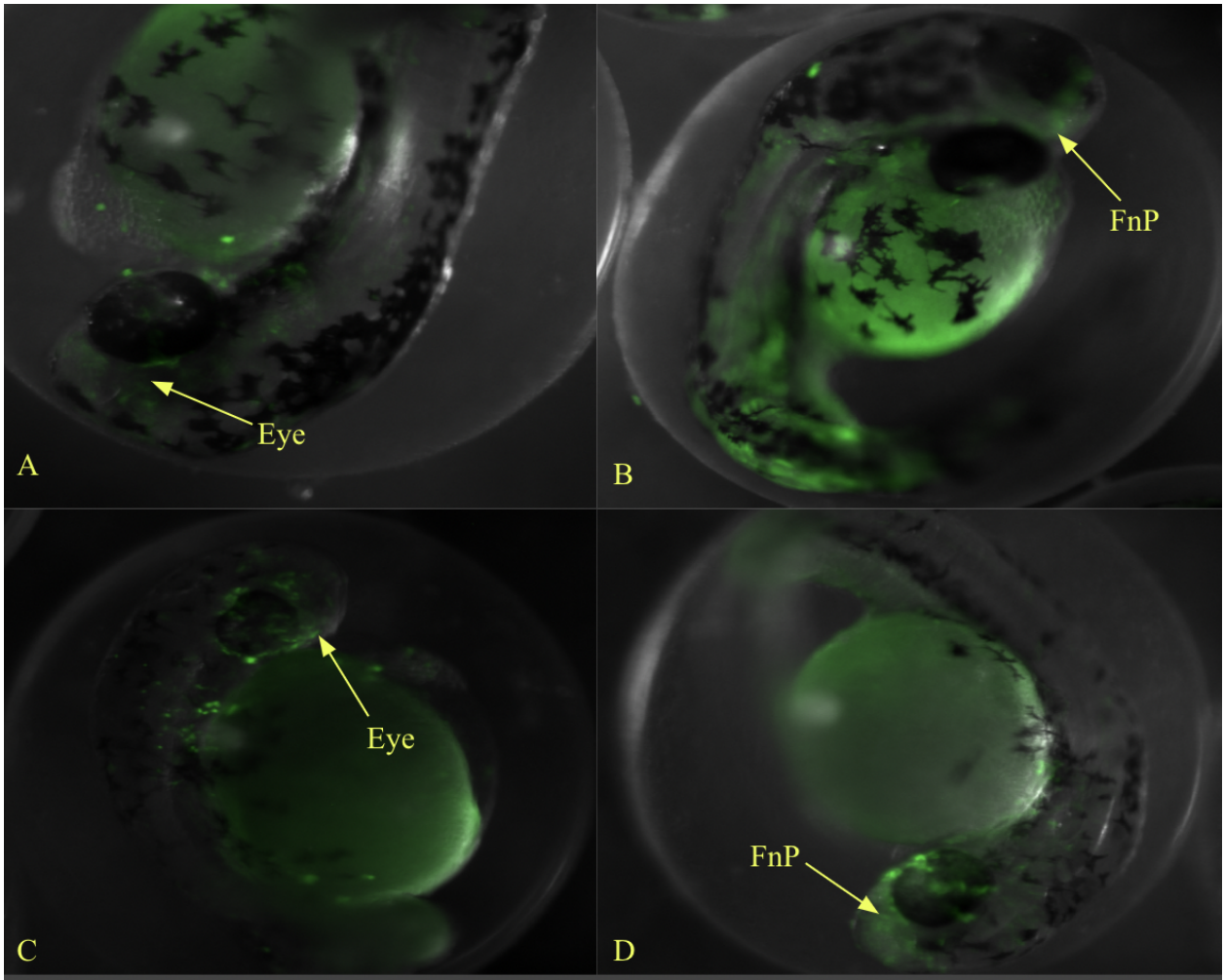

Figure 5: A series of zebrafish embryos injected with putative CRE7 imaged 24 hpf displaying GFP expression. FnP is highlighting potential NCC expression in the frontonasal process particularly strongly in D. Eye is showing potentially specific NCC expression around the eyes, particularly strongly in C. Both of these are consistent with NCC expression and activity.

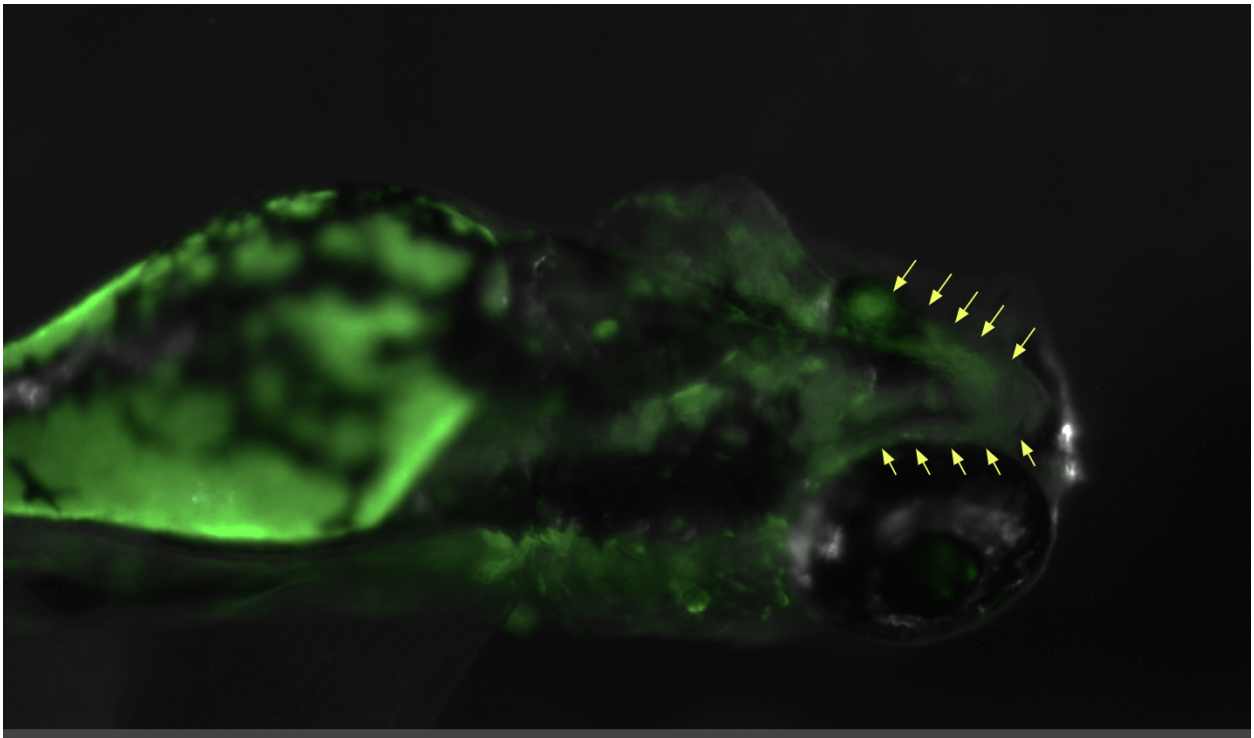

Figure 6: This image demonstrates potential specific signal in the lower jaw of a zebrafish 48 hpf. The lower jaw is NCC derived but could be residual expression from an earlier stage. The arrows outline the developing lower jaw. The pictured embryo was injected with putative CRE1, and appears to show significant potential for NCC expression patterning.

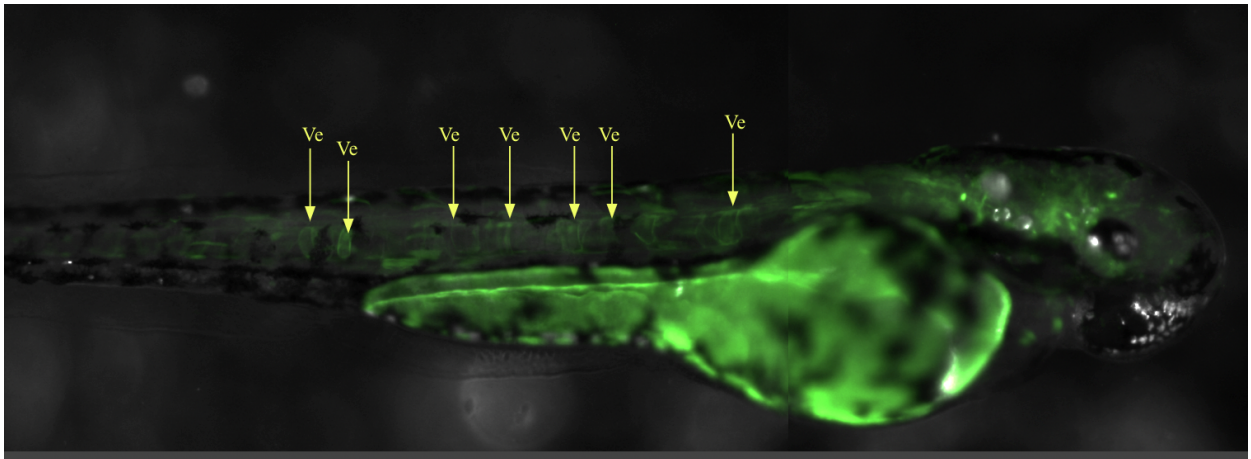

Figure 7: This image was taken 48 hpf of a embryo injected with putative CRE1. Although this does not show any expression consistent with that of NCC, the vertebral column shows very apparent expression which could be attributed to *twist1b* because it is expressed in somites at earlier stages, and GFP may be persisting in differentiating vertebral tissues.

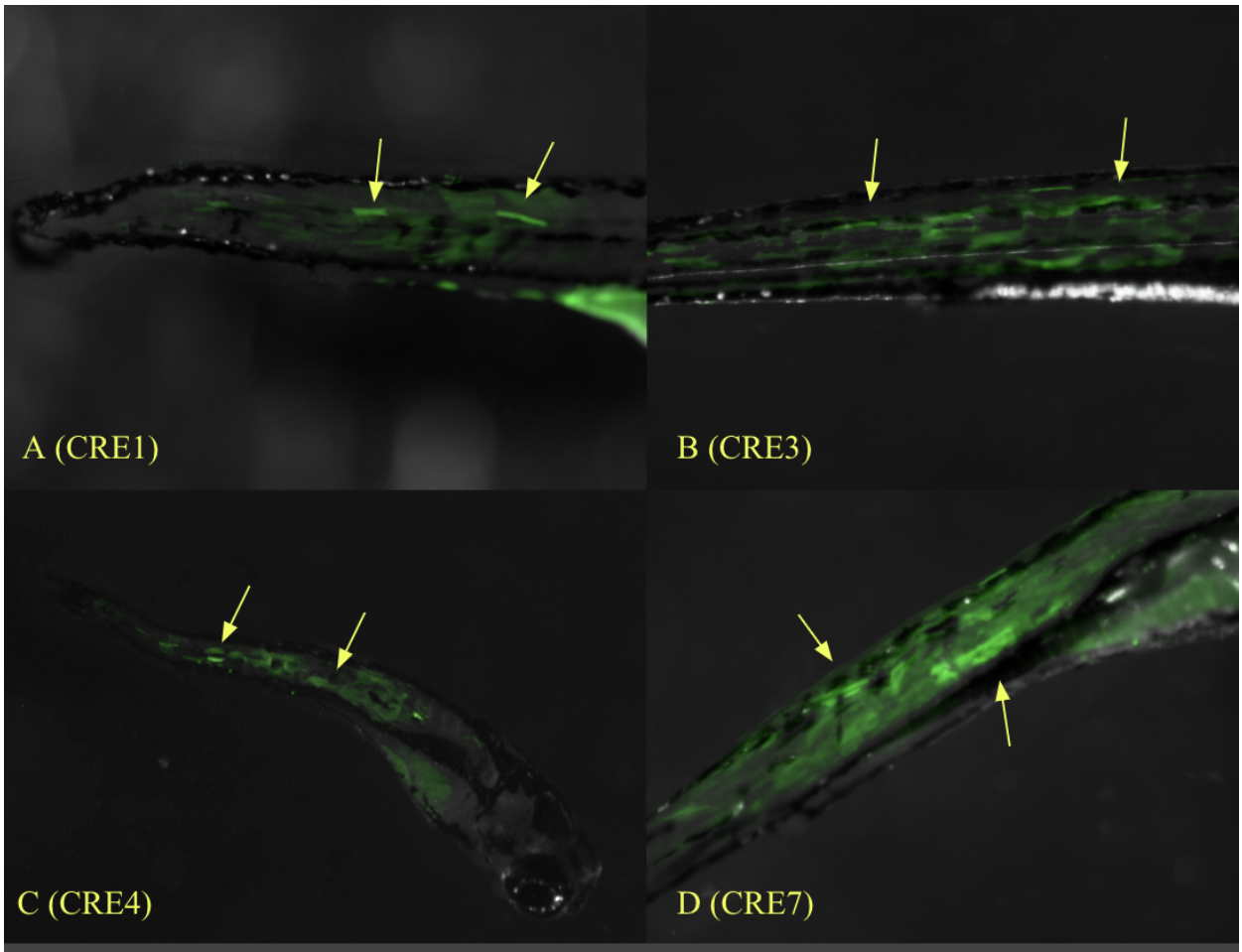

Figure 8: These images show GFP expression in muscle, The horizontal lines are muscle fiber expression. This is often non-specific, but may be specific in this case because Twist1b is expressed in somites at earlier stages, and GFP may be persisting in differentiating muscle fibers.

From the expression patterning in injected zebrafish embryos there were multiple cases of mesodermal and ectodermal expression in the same embryo as well as examples of isolated expression. Table 1 below shows which putative cis-regulatory elements demonstrated mesodermal or ectodermal expression.

Table 1: Expression Patterning in Dermal Tissue Layers

Putative CRE #	Mesodermal Expression	Ectodermal Expression
1	✓	✓
3	✓	✓
4	✓	
6		✓
7	✓	✓

Table 1 shows that only CRE6 expressed only ectodermal expression patterning, while CRE1, CRE3, and CRE7 demonstrated observable expression in both the mesoderm and ectoderm. The only putative cis-reulgatory element functionally tested that did not demonstrate any ectodermal expression was CRE4.

4. Discussion

4.1 Findings and Limitations

Of the five putative CREs injected into single-cell stage zebrafish embryos, all showed expression activity confined to specific tissue types and locations, including tissues that represent general mesoderm, general ectoderm, or are neural crest. At this time, it cannot be excluded that expression recorded as non-specific is due to *twist* expressive patterns in both the mesoderm and ectoderm. The present findings, suggesting potential expression of *twist1b* in both ectodermal and mesodermal tissues during zebrafish development, need to be further tested with attention to whether or not the enhancers demonstrate *twist* specific expression. For a key developmental

gene such as *twist,* elucidation of expression patterns is complicated by the vast number of CREs involved in regulating developmental genes.

Those CREs that exhibit expression strictly limited to ectodermal tissue, such as CRE6, or NCCs are unlikely to be highly conserved and may have developed more recently in an evolutionary timescale. If they were older evolutionarily, they would be involved in twist regulation prior to its co-option in the neural crest. Further analysis of these CREs may provide insight into the question of the co-option of *twist* occurred. If these putative sequences were not conserved, this would suggest the possibility that a completely new element appears in the *twist* gene.

The CREs expressed in both mesoderm and neural crest may be the most interesting for future assessment of whether an existing CRE may have mutated, resulting in twist expression in the ectoderm while still also maintaining its mesodermal expression. Such a scenario would also not exclude the possibility of a combination of conservation of an old CRE accompanied with a newly developed one would still result in the expression of both layers we see today.

Ultimately, my findings suggest that the co-option of twist into the neural crest gene regulatory system may have occurred via more than one of the hypothesized pathways.

4.2 Recommendations for Future Research

There are several approaches that can be taken both in the short term and long term that would give provide more clarity in the results and lead to the potential for more confidence both in the support of a particular hypothesis and in the specific function of the tested putative cis-regulatory elements.

A problem faced when analyzing images taken during fluorescent microscopy was the

mesodermal expression seen in many of the embryos. This mesodermal expression patterning could not be dismissed as non-target due to the putative enhancers being connected to the *twist* gene. Since *twist* is activated in both the mesoderm and ectoderm, seeing both expression patterns leads to the belief that perhaps the putative cis-regulatory element being tested is responsible for *twist* activation in both tissue types. The incorporation of a control into experimentation may highlight where non-specific expression is occurring and make it easier to identify non-target expression in transgenic embryos. By identifying non-target expression patterning associated with the vector and not the putative CRE the expression patterns not associated with the CRE can more easily be removed and clarify the observed data. In order to incorporate a control into experimentation, an empty pGreenE vector containing no DNA fragment incorporated between the Tol2 arms would show where GFP expression is occurring that is non-specific.

Mosaic expression occurs when not all cells within the transgenic organism incorporate the injected information into their genome. This results in incomplete expression patterning and higher amounts of non-target expression. In order to mitigate the non-target expression seen in the results and clarify the observable data, transgenic lines from injected embryos need to be formed. By raising injected embryos and crossing those with the same cis-regulatory elements incorporated into their genome, transgenic lines can be developed. Through the development of transgenic lines, mosaic expression is mitigated due to the complete incorporation of the target sequence in the entirety of the organism. Imaging embryos from transgenic lines should show significantly clearer specificity in the GFP expression of putative cis-regulatory elements and help in determining whether the mesordermal expression observed was non-specific or instead associated with *twsit1b* expression outside of the neural crest.

For CREs exhibiting specific mesodermal expression, use of a target species such as tunicates would allow further assessment of whether or not these enhancers are evolutionarily conserved. Moreover, genome alignments of multiple vertebrates, chordates, and other phyla could be used to assess whether these enhancer sequences may potentially be heavily conserved. By incorporating the suggested control, raising transgenic lines, and using the clarified data to conduct further experimentation in other species, a more confident conclusion could be drawn toward which hypothesis is responsible for the co-option of twist into the neural crest gene-regulatory system.

Acknowledgments:

Thank you to all of the members of my thesis committee, including Drs. Daniel Medeiros, Barbara Demmig-Adams, and Kate Semsar. You have each been important mentors during my time at the University of Colorado Boulder, and I greatly appreciate your insights and patience. Thank you to the Medeiros lab for letting me to work alongside them for the past two years allowing me to learn more about genetics and find my passions for graduate school. Thank you to David Jandzik for the immense hands on-work conducted with me during the Honors thesis process.

References

Abitua, P. B., Wagner, E., Navarrete, I. A., & Levine, M. (2012). Identification of a rudimentary neural crest in a non-vertebrate chordate. *Nature*, *492*(7427), 104–107. https://doi.org/10.1038/nature11589

Fan, X., Masamsetti, V. P., Sun, J. Q. J., Engholm-Keller, K., Osteil, P., Studdert, J., Graham, M. E., Fossat, N., & Tam, P. P. L. (2021). TWIST1 and chromatin regulatory proteins interact to guide neural crest cell differentiation. *ELife*, *10*. https://doi.org/10.7554/elife.62873

Gans, C., & Northcutt, R. G. (1983). Neural crest and the origin of vertebrates: A new head. *Science*, *220*(4594), 268–273. https://doi.org/10.1126/science.220.4594.268

Germanguz, I., Lev, D., Waisman, T., Kim, C.-H., & Gitelman, I. (2007). Four twist Genes in Zebrafish, Four Expression Patterns . *Developmental Dynamics*, *236*(9), 2615–2626. https://doi.org/10.1002/dvdy.21267

Gitelman, I. (2007). Evolution of the vertebrate twist family and synfunctionalization: A mechanism for differential gene loss through merging of expression domains. *Molecular Biology and Evolution*, *24*(9), 1912–1925. https://doi.org/10.1093/molbev/msm120

Horie, R., Hazbun, A., Chen, K., Cao, C., Levine, M., & Horie, T. (2018). Shared evolutionary origin of vertebrate neural crest and cranial placodes. *Nature, 560*(7717), 228–232. https://doi.org/10.1038/s41586-018-0385-7

Huang, X., & Saint-Jeannet, J.-P. (2004). Induction of the neural crest and the opportunities of life on the edge. *Developmental Biology*, *275*(1), 1–11. https://doi.org/10.1016/j.ydbio.2004.07.033

Irie, N., & Sehara-Fujisawa, A. (2007). The vertebrate phylotypic stage and an early bilaterian-related stage in mouse embryogenesis defined by genomic information. *BMC Biology*, *5*(1). https://doi.org/10.1186/1741-7007-5-1

Martik, M. L., & Bronner, M. E. (2021). Riding the crest to get a head: Neural Crest Evolution in vertebrates. *Nature Reviews Neuroscience*, *22*(10), 616–626. https://doi.org/10.1038/s41583-021-00503-2

Meulemans, D., & Bronner-Fraser, M. (2004). Gene-regulatory interactions in neural crest evolution and development. *Developmental Cell*, *7*(3), 291–299. https://doi.org/10.1016/j.devcel.2004.08.007

Prud'homme, B., Gompel, N., & Carroll, S. B. (2007). Emerging principles of regulatory evolution. *Proceedings of the National Academy of Sciences*, *104*(suppl_1), 8605–8612. https://doi.org/10.1073/pnas.0700488104

Theveneau, E., & Mayor, R. (2014). Neural Crest Cell migration. *Neural Crest Cells*, 73–88. https://doi.org/10.1016/b978-0-12-401730-6.00004-1

U.S. National Library of Medicine. (n.d.). *Twist1 twist family bHLH transcription factor 1 [Homo sapiens (human)] - gene - NCBI*. National Center for Biotechnology Information. Retrieved March 4, 2022, from

https://www.ncbi.nlm.nih.gov/gene?Db=gene&Cmd=ShowDetailView&TermToSearch=7291

Veldman, M. B., & Lin, S. (2008). Zebrafish as a developmental model organism for pediatric research. *Pediatric Research*, *64*(5), 470–476. https://doi.org/10.1203/pdr.0b013e318186e609

Wittkopp, P. J., & Kalay, G. (2011). Cis-regulatory elements: Molecular mechanisms and evolutionary processes underlying divergence. *Nature Reviews Genetics*, *13*(1), 59–69. https://doi.org/10.1038/nrg3095

Zfin Code DCC. DANIO. (n.d.). Retrieved March 15, 2022, from https://danio-code.zfin.org/trackhubDescription/