2014

Elucidating the Transcriptional Regulation of col2a1a in Zebrafish

Sonja Dabizljevic
Loyola University Chicago

LOYOLA UNIVERSITY CHICAGO

ELUCIDATING THE TRANSCRIPTIONAL REGULATION OF *COL2A1A* IN ZEBRAFISH

A THESIS SUBMITTED TO

THE FACULTY OF THE GRADUATE SCHOOL

IN CANDIDACY FOR THE DEGREE OF

MASTER OF SCIENCE

PROGRAM IN BIOLOGY

BY

SONJA DABIZLJEVIC

CHICAGO, IL

MAY 2016

ACKNOWLEDGMENTS

I would like to thank my thesis advisor, Dr. Rodney M. Dale, for his mentorship during my graduate studies. His passion for research is contagious and his guidance has been instrumental in my advancement as a scientist. I am especially grateful for his continuous support during my continued academic endeavors. Thank you to my fellow Dale lab members who have been valuable assets to have in the laboratory, especially Peter Lyson, Daniel Brissette, and Stephanie Erhard. I am also very grateful to the members of my committee, Dr. Bryan F. Pickett and Dr. Catherine Putonti, for their support and advice during the completion of my project.

TABLE OF CONTENTS

ACKNOWLEDGMENTS iii

LIST OF TABLES vi

LIST OF FIGURES viii

CHAPTER ONE: INTRODUCTION 1
Evolution and Genetic Regulation 1
Cartilage and Collagens 2
Zebrafish Model Organism 7
Transcriptional Regulation of *col2a1a* 7

CHAPTER TWO: MATERIALS AND METHODS 10
Construction of Minimized R2 Entry Plasmids 10
Amplification of the R2 minimized plasmids 10
Insertion of amplified sequence in plasmid 12
Purification of plasmid DNA from transformation 14
Confirmation of accurate plasmid 14
Preparation of larger volume of plasmids 15
Construction of mutated putative transcription factor binding sites 16
Generation and Analysis of Minimized R2 Reporter Transgenic Zebrafish 17
Construction of destination vector 17
Zebrafish embryo microinjections 19
Imaging and analysis of transgenic zebrafish lines 20
Construction of Transcription Factor Expression Plasmids 22
Isolation of transcription factor sequences 22
Transcription factor sequence inserted into middle-entry vector 24
Isolation of middle-entry plasmids 25
Transcription factor expression vectors 25
Generate transcription factor mRNA 26
Zebrafish embryo microinjections 28
Zebrafish Husbandry 28

CHAPTER THREE: RESULTS 29
Identification of R2-DE as the *col2a1a* Regulatory Region 29
Tissue Specificity of R2-DE Regulatory Region 43
Identification of Transcription Factors Binding to Regulatory Region 51
Mutagenesis of Putative R2-DE Transcription Factor Binding Sites 53
Increase Activity in Mutated Ets-family Binding Site 56
Identification of Transcription Factors 72

CHAPTER FOUR: DISCUSSION 75

REFERENCE LIST 78

VITA 83

LIST OF TABLES

Table 1. Primers for isolation of R2 minimized plasmids 11

Table 2. Primers for isolation of transcription factors 23

Table 3. Primers with attB sites for isolation of transcription factors 24

Table 4. Results of R2 injections 33

Table 5. Results of zebrafish injected with R2-AB plasmid 35

Table 6. Results of zebrafish injected with R2-BC plasmid 36

Table 7. Results of R2-CD driving expression of EGFP 37

Table 8. Results of R2-DE driving expression of EGFP reporter gene 38

Table 9. Results of R2-EF driving expression of EGFP 40

Table 10. Results of injections with D-E1 driving expression 46

Table 11. Results of injections with D2-E driving expression 47

Table 12. Results of injections with D2-E1 driving expression 48

Table 13. Results of injections with D region driving expression 49

Table 14. Results of injections with E region driving expression 50

Table 15. Results of injections with DE plasmid containing Ets-binding site mutated 63

Table 16. Results of injections with the Runx-binding site mutated in the DE plasmid 64

Table 17. Results of injections of the DE plasmid with the Sox binding site mutated 65

Table 18. Results of injections with Ets and Sox binding sites mutated in DE plasmid 66

Table 19. Results of Ets and Runx binding sites mutated in DE plasmid injections 67

Table 20. Results of Runx and Sox binding sites mutated in the DE plasmid 68

Table 21. Results of injections with Fli binding site mutated in the DE plasmid 69

Table 22. Results of injections with mutations in Fli and Runx binding site of DE 70

Table 23. Results of DE plasmid with mutations in the Fli and Sox binding sites 71

LIST OF FIGURES

Figure 1. R2 Regulatory Region Sequence and Results 33

Figure 2. Minimization of R2 regulatory region 34

Figure 3. Results of R2-AB plasmid injections 35

Figure 4. Results of R2-BC plasmid injections 36

Figure 5. Results of R2-CD plasmid injections 37

Figure 6. Results of R2-DE plasmid injections 38

Figure 7. Results of R2-EF plasmid injections 40

Figure 8. Results of R2-DE stable line 41

Figure 9. Minimization of R2-DE Sequence 45

Figure 10. Results of R2: D-E1 plasmid injections 46

Figure 11. Results of R2: D2-E plasmid injections 47

Figure 12. Results of R2: D2-E1 plasmid injections 48

Figure 13. Results of R2: D plasmid injections 49

Figure 14. Results of R2: E plasmid injections 50

Figure 15. Sequence Conservation Schematic 58

Figure 16. Ets-Binding Site Mutation Schematic 59

Figure 17. Runx-Binding Site Mutation Schematic 60

Figure 18. Sox-Binding Site Mutation Schematic 61

Figure 19. Fli-Binding Site Mutation Schematic 62

Figure 20. Results of Injections with Ets Binding Site Mutated in Plasmid 63

Figure 21. Results of Injections with Runx Binding Site Mutated in Plasmid 64

Figure 22. Results of Injections with Sox Binding Site Mutated in Plasmid 65

Figure 23. Results of Injections with Ets & Sox Binding Sites Mutated in Plasmid 66

Figure 24. Results of Injections with Ets & Runx Binding Sites Mutated in Plasmid 67

Figure 25. Results of Injections with Runx & Sox Binding Sites Mutated in Plasmid 68

Figure 26. Results of Injections with Fli Binding Site Mutated in Plasmid 69

Figure 27. Results of Injections with Fli & Runx Binding Sites Mutated in Plasmid 70

Figure 28. Results of Injections with Fli & Sox Binding Sites Mutated in Plasmid 71

Figure 29. Results of Transcription Factor Injections 74

CHAPTER ONE

INTRODUCTION

Evolution and Genetic Regulation

Even though different vertebrate species look quite diverse in their juvenile and adult forms, they all begin as a single fertilized egg and use similar developmental processes and genes to reach their mature state, which can be seen from comparisons of zebrafish to humans. There are early structural similarities among most vertebrate embryos during certain stages, including the presence of a head, pharyngeal arches, and a tail. Many of these early structures seem to be conserved among a variety of species, and it is the differential regulation of genetic networks that helps specific tissues of each organism differentiate according to evolutionarily and anatomically integrated patterns of gene expression (Gilbert, 2010). Particular genes are turned on and off at the appropriate place and time in development to regulate downstream targets and ensure the proper development of tissues for that particular organism. From the use of HOX genes in early developmental processes, to Netrins in axon guidance, many of these genetic networks are conserved throughout evolution (Manzanares *et al.*, 2000, Rajasekharan *et al.*, 2009). Even though organisms are constantly evolving in part due to the mutations that are randomly occurring in their genome, it makes sense that the amino acid sequence of the coding region of genes are highly conserved in order to maintain protein function (Bejerano *et al.*, 2004; Pruitt K, 2009). It is more interesting that cis elements associated closely with genes tend to be functionally conserved through evolution, although relative

position and location in relation to the transcriptional start site appear to vary. This has led modern biology to focus on the evolution and function of these putative cis regulatory regions (Maston *et al.*, 2006; ENCODE Project Consortium, 2012). The conservation of these non-coding sequences over evolution suggests that they are not altered because they play an important role for the fitness of the organism, particularly in the regulation of the genome (Woolfe *et al.*, 2005). Appropriate genetic regulation allows for normal cell growth, but if this process is interrupted it can have detrimental effects on the organism. In order to understand the components of this regulation it is necessary to identify the transcription factors essential for the transcriptional regulation. The presence of these regulatory proteins can accelerate or inhibit the rate of transcription by binding to DNA, which affects the amount of mRNA produced by the gene (Spitz and Furlong, 2012; Alberts, 2007). These transcription factors can bind upstream, downstream, near, or far from the transcriptional start site of a gene because it is the sequence of nucleotides that determine whether it is capable of binding (Palstra, 2012). The binding of these transcription factors emphasizes the necessity of organisms to conserve the function of non-coding portions of the genome based on their role in gene regulation.

Cartilage and Collagens

The requirement for abundant cartilage in a vast number of processes in vertebrate species suggests the genetic regulation of the genes critical for its formation is particularly important. Cartilage is a connective tissue evolved to be a support structure for organisms that also provides flexibility during movement. The elasticity of this tissue is essential in articular surfaces to support the movement of adjacent bones, in the

intervertebral discs to cushion the force and strain of the vertebrae, and the ear of mature animals (Fox *et al.*, 2009; Roberts *et al.*, 2006). Musculoskeletal issues, such as back pain caused by intervertebral disc degeneration and degradation of joint cartilage and osteoarthristis, can naturally develop as an individual ages due to wear and tear on the body. These conditions are becoming more prevalent and leading to a major healthcare concern in society (Le Maitre, 2007; Freemont, 2009; Goldring, 2012).

Cartilage also has a significant role in the developing embryo because the endochondral bones that form many parts of the skeleton, like the lower jaw structure of gnathostomes, are initially composed of cartilage and will provide the framework for future bone formation (Kuratani, 2005; Kimmel *et al.*, 1998). The differentiated cartilage cells, called chondrocytes, secrete extracellular matrix representative of cartilage. When the ossification process begins the chondrocytes undergo hypertrophy and a change in transcription factor regulation, which will begin the process of calcification of the chondrocytes and the mineralization of the matrix to form the skeleton (Mackie et al, 2008; Vega *et al.*, 2004). For example, it is the initial configuration of the cartilage that lays the framework for the lower jaw. This structure is neural crest derived and these cells migrate from the midbrain and hindbrain region into the pharyngeal arches to form the craniofacial cartilage, and eventually the bony lower jaw (Minoux and Rijli, 2010; Knight *et al.*, 2006). This process shows that the tightly regulated spatiotemporal expression of transcription factors is not only important for the initial differentiation of chondrocytes, but also continues to play a role in the endochondral ossification process (Ding *et al.*, 2012; Yang and Karsenty, 2002; Gilbert, 2010).

Cartilage is mainly composed of an extracellular matrix of collagen and proteoglycans, which is secreted from proliferating chondrocytes (Gao *et al.*, 2014). Collagen is an abundant fibrillar protein composed of a triple helix repeat, which can then link with other collagen molecules in the extracellular matrix (Exposito *et al.*, 2010; Kadler, 1996). There are twenty-eight different collagen types that are present in a wide array of tissues including skin, eye, muscle, and cartilage; of these collagen types it is the type II collagen α1 (*COL2A1*) gene that is crucial in the cartilage, notochord, intervertebral disc, and vitreous humor of the eye (Cheah *et al.*, 1991; Alberts *et al.*, 2007). Mutations to this gene can lead to congenital birth defects including Achondrogenesis, Stickler syndrome, and Spondyloperpheral dysplasia, and the detrimental results are characterized by spinal or skeletal deformities, hearing irregularities, or ocular abnormalities (Winterpacht *et al.*, 1993; Brown *et al.*, 1992; Kuivaniemi *et al.*, 1991).

Zebrafish have two orthologues of *col2a1*, referred to as *col2a1a* and *col2a1b*. Following the sequence analysis of the zebrafish homologues, our laboratory has shown overlapping expression patterns in the perichondrium of cartilage, but only *col2a1a* was expressed in the chondrocytes based on *in situ* hybridization (Dale and Topczewski, 2011). Our laboratory focuses on *col2a1a* because of this expression in the chondrocytes and perichondrium of the craniofacial cartilage and the notochord. The goal of our research is to understand the transcriptional regulation of the zebrafish orthologue *col2a1* in order to elucidate the components required for *col2a1a* expression. Studying the fundamental aspects of *col2a1a* transcriptional regulation will provide us with a basic

understanding of how this evolutionary conserved and critical vertebrate gene functions. Understanding the transcription factors that determine its expression and the tissue specific activity of our identified enhancer element may lead to novel ways to initiate transcription in areas of deteriorating cartilage. By having a minimal regulatory element that can initiate transcription only in specific cell types will have implications for future research in the fields of synthetic biology and gene therapy.

There are transcription factors that have already been shown to play an important role in chondrogenesis, including Sox9 (Yang and Karsenty, 2002; Ng *et al.*, 1997; Yan *et al.,* 2005). The mouse knockout of Sox9 results in a deficiency of cartilage and bone, to further emphasize the importance of the transcription factor in early development (Akiyama *et al.*, 2002). Morpholino injections that target *sox9a* in zebrafish embryos also produce zebrafish with disrupted and/or lost craniofacial cartilage structures, similar to what is observed in mice (Yan *et al.,* 2002). The injection of these targeted morpholinos was also performed by our laboratory to show the reduction of reporter protein levels specifically in cranial cartilage elements but not in the ear or notochord in our *col2a1a* reporter transgenic zebrafish line (Dale and Topczewski, 2011). While it is evident that Sox9a is an important transcription factor in the regulation of *col2a1a*, many other transcription factors certainly contribute to chondrogenesis. For example, Runx2 is a transcription factor known to be involved in osteoblast differentiation, but it can be inferred that this transcription factor has another role early in development because of its presence when chondrocytes begin to form (Otto *et al.*, 1997; Kerney *et al.*, 2007). In zebrafish, *runx2b* and *runx3* are expressed at high levels as early as 34 hours post

fertilization which correlates to the development of the craniofacial cartilage; this spatiotemporal correlation implies the transcription factor could be important for regulation of *col2a1a* by binding to the regulatory element (Flores *et al.*, 2006). These transcription factors are present at the correct time and place for cartilage formation, but it is most likely Runx2 that is responsible for this expression based on previous research (Ding *et al.*, 2012; Flores *et al.*, 2006). Similar to *sox9a*, the injection of antisense morpholinos for Runx proteins results in craniofacial deformities (Flores *et al.*, 2006). Continued research of this transcription factor family will continue to enrich our knowledge of their different roles in the development of cartilage and the endochondral ossification process. The Sox and Runx transcription factors are also known to interact with the Ets family of transcription factors, a large class of DNA binding proteins. Fli1 is a member of the ERG subfamily of Ets family of transcription factors. Yeast two-hybid interactions have demonstrated its interactions with Sox proteins (Deramaudt *et al.*, 2001). Interestingly, Fli1 and Runx1 proteins have been confirmed to bind by immunoaffinity studies and transcriptional reporter assays (Huang *et al.*, 2009). The interplay of these transcription factor proteins is particularly interesting because ERG is shown to have a role in cartilage differentiation and development (Dhordain *et al.*, 1995; Iwamoto *et al.*, 2001). While these transcription factors are known to play a significant role in cartilage development, our research hopes to identify whether these proteins are binding to the critical regulatory element of *col2a1a*.

Zebrafish Model Organism

Our laboratory is interested in studying the fundamental aspects of *col2a1* transcriptional regulation utilizing *Danio rerio*, commonly known as the zebrafish. Zebrafish are an excellent vertebrate model organism for our research because of their transparent embryo and reproductive capabilities (Westerfield, 2000). This transparency allows us to view development of the zebrafish without disrupting the embryo or mother, as well as visualize the presence of the fluorescent proteins using standard microscopy techniques. Their large clutch size and rapid development is also an advantage compared to other vertebrate model organisms. While these are beneficial for our research, it is the conservation of vertebrate signaling pathways that make zebrafish an ideal model organism for studying developmental pathways (Dooley and Zon, 2000). The ability to easily produce transgenic zebrafish will allow us to characterize the upstream regulatory element of collagen.

Transcriptional Regulation of col2a1a

A 182 bp regulatory element was identified in the first intron of the mouse *Col2a1* gene, and could drive expression of a reporter gene in chondrocytes (Zhou et al, 1995). Interestingly, the characterized mouse enhancer is quite weak and there have been suggestions in the field that there are other stronger enhancers elsewhere. Subsequently, our laboratory identified a 360 bp regulatory region in zebrafish as the region critical for cartilage, ear, and notochord expression of *col2a1a*, as determined by its ability to express enhanced green fluorescent protein (EGFP) at the proper time and place (Dale and Topczewski, 2011). The regulatory element important for this expression, referred to

as R2, was identified using comparative genomics of four distantly related teleost fish. Twelve kilobases (kb) of the zebrafish, medaka, fugu, and stickleback genome surrounding the transcriptional start site were compared, and four highly conserved sequences were discovered. When these conserved regions were placed upstream of a reporter gene to drive expression, only the R2 region drove EGFP expression in the appropriate tissues. The 360 bp regulatory region was approximately -1.7kb upstream of the transcriptional start site of *col2a1a* on chromosome 8, as indicated by the schematic in Figure 1. As previously published, the plasmid containing the R2 regulatory element driving expression of EGFP can recapitulate most of the endogenous expression of *col2a1a*, with the presence of this fluorescent protein in the cartilage, ear, and notochord.

The goal of our research was to determine what region of R2 was responsible for initiating transcription, and by isolating this portion of the regulatory region for *col2a1a* we could begin to infer the mechanisms of its regulation. Deletion analysis of the R2 regulatory region has been performed testing six 60 bp units designated as A through F, which allowed us to identify the critical region of the regulatory fragment that can drive expression of *col2a1a*. Gateway Cloning, zebrafish transgenesis, and fluorescence microscopy were used to determine which section of the R2 region was critical for the expression of *col2a1a* (Kawakami, 2004). Subsequently, our aim was to identify tissue specificity within the R2 subunits. Testing segments of the R2 regulatory region in isolation demonstrated whether these individual fragments were responsible for the specific expression in the cartilage, ear, or notochord. This process allowed for an inference as to whether the reporter protein expression could be activated specifically in

different tissues. The last step was to determine what transcription factors may be required for the expression pattern in each specific tissue. The isolation of a minimal regulatory region was crucial in the identification of transcription factor binding sites that were necessary for expression in that specific tissue. *In silico* analysis using comparative genomics to assess the sequences of distantly related teleosts identified highly conserved sequences, and potential transcription factor binding sites of interest. The importance of these putative transcription factor binding sites were determined using site-directed mutagenesis of the identified sequences. This process investigated of the interactions between the DNA sequence in the regulatory region and the transcription factors that were bound to allow for transcriptional initiation. In particular if the smaller regulatory units were crucial for expression in different tissues of the zebrafish, we could begin to identify the transcription factors and their networks that were essential for expression in the different units of the regulatory region.

CHAPTER TWO

MATERIALS AND METHODS

Construction of Minimized R2 Entry Plasmids

Amplification of the R2 minimized plasmids

The homologous recombination technology of the Multisite Gateway vector system was used to construct the expression plasmids containing the dissected regulatory elements (Petersen and Stowers, 2011). The primers used for the amplification of the minimized R2 units were derived from the complementary sequence of the regulatory element. Additional sequences, referred to as attB sequences, were added to the flanking 5' end of the primers to subsequently assist with their recombination into the pDONR vectors. The PCR primer sets used for the amplifications are listed below.

Table 1. Primers for isolation of R2 minimized plasmids

Regulatory Elements	**Forward Primer**	**Reverse Primer**
AB	GGGGACAACTTTGTATAG AAAAGTTGCTGCCCTCTG ACACCTGATGCCAATTGC	GGGGACTGCTTTTTTGTA CAAACTTGCCAGGGGTG TGTAGGGTGGCTGGG
BC	GGGGACAACTTTGTATAG AAAAGTTGATCCAATGGC CAGGCCCCTCATCATC	GGGGACTGCTTTTTTGTA CAAACTTGCGGGTGTGG ATGGAGGGAGAGTGCG
CD	GGGGACAACTTTGTATAG AAAAGTTGCTGAGCCTCT CCGTGTTCTCCTCATCC	GGGGACTGCTTTTTTGTA CAAACTTGCTCTGTGTGC AGACCTGAGGAATGTG
DE	GGGGACAACTTTGTATAG AAAAGTTGCTGCGGCTCT CTTCTCCCCCACTGCC	GGGGACTGCTTTTTTGTA CAAACTTGCGTGTGTGTG TCCGAAATGAGCCC
EF	GGGGACAACTTTGTATAG AAAAGTTGCTGCCGCATT GTGTGTGTGTCTTACAG	GGGGACTGCTTTTTTGTA CAAACTTGCAGGGATAT GTGTATGTGTGTGTACG
D	GGGGACAACTTTGTATAG AAAAGTTGCTGCGGCTCT CTTCTCCCCCACTGCC	GGGGACTGCTTTTTTGTA CAAACTTGCTCTGTGTGC AGACCTGAGGAATGTG
E	GGGGACAACTTTGTATAG AAAAGTTGCTGCCGCATT GTGTGTGTGTCTTACAG	GGGGACTGCTTTTTTGTA CAAACTTGCGTGTGTGTG TCCGAAATGAGCCC
D-E1	GGGGACAACTTTGTATAG AAAAGTTGCTGCGGCTCT CTTCTCCCCCACTGCC	GGGGACTGCTTTTTTGTA CAAACTTGAGTGCTCTGT AAGACACAC
D2-E	GGGGACAACTTTGTATAG AAAAGTTGAGTCTCTCAC ATTCCTCAGGTCTGC	GGGGACTGCTTTTTTGTA CAAACTTGCGTGTGTGTG TCCGAAATGAGCCC
D2-E1	GGGGACAACTTTGTATAG AAAAGTTGAGTCTCTCAC ATTCCTCAGGTCTGC	GGGGACTGCTTTTTTGTA CAAACTTGAGTGCTCTGT AAGACACAC

The previously published plasmid in Dale and Topczewski 2011 containing the R2 region in the 5'-entry vector was used to amplify the subunits of R2 (A-F). The PCR reactions contained dNTPS (200 μM each), 20 μl I- Buffer, 10 ng DNA, 2 U iProof DNA polymerase (BioRad), and forward and reverse primers (1 μM each) for the desired

subunit. The PCR program used for the amplification of the attB-flanked product is detailed below.

Step 1	98°C	1 min
Step 2	98°C	45 sec
Step 3	60°C	45 sec
Step 4	72°C	30 sec
Repeat Step 2-4 x35 times		
Step 5	72°C	10 min
Step 6	4°C	10 min
Step 7	15°C	∞

In order to verify the amplification of the correct sequence, agarose gel electrophoresis was performed to confirm the size of the PCR product. The confirmed band length was cut from the agarose gel, and the Zymoclean Gel DNA Recovery Kit was used to purify the DNA. Three volumes of the agarose dissolving buffer were added to the gel fragment and incubated at 50°C to liquefy the solid gel to release the desired DNA fragment. The solution was subsequently centrifuged in a column to remove the agarose, while the DNA remained bound to the filter. The DNA binding filter was rinsed with a DNA wash buffer, and then the purified DNA was eluted. This process isolated, amplified, and purified the R2 minimized units with the flanking attB sequences.

Insertion of amplified sequence in plasmid

The purified PCR products for the subunits of R2 (AB, BC, CD, DE, EF, D, E, D-E1, D2-E, D2-E1) had the attB sites required for recombination into the 5'- entry vector in a Multisite Gateway BP reaction. A mixture with 150 ng of the PCR product, 150 ng of the pDONR4-1 vector, TE Buffer, and 2 µl of BP clonase enzyme was made and incubated overnight at room temperature. The following day the reaction was terminated

by the addition of 1 μl of proteinase K to the reaction. The pDONR4-1 vector had a *ccdB* gene that eliminated potential bacterial cells that contained the original pDONR plasmid, and did not recombine out the *ccdB* gene. This gene is recombined out of the vector if our sequences are successfully integrated. The homologous sequences between the pDONR vector and the attB sites on the PCR product allowed the recombination of the PCR product into the 5'-entry vector, which removed the *ccdB* gene. Bacterial cells that contained the entry vector with the appropriate insert proliferated, while those cells with the original *ccdB* gene were destroyed. The plasmid was added to chilled TOP10 One Shot chemical competent cells (Invitrogen), which was set on ice for 30 minutes. These cells were heat-shocked at 42°C for 90 seconds to allow for plasmid uptake into the bacterial cell. Super Optimal Broth with Catabolite Repression (SOC) medium was added to the cells and provided nutrients for the replication of bacterial cells, and in turn produced copies of the desired vector. After an hour in a 37°C shaking incubator, the bacterial cultures were centrifuged and the pellet was resuspended in 125 μl of the SOC medium. This concentrated bacterial culture was plated on a kanamycin antibiotic plate and grown overnight at 37°C to selectively allow growth of bacteria that had taken up the desired plasmid. The 5'-entry plasmids had a kanamycin resistance gene, so bacterial cells with the entry plasmids were resistant to this antibiotic. For our transformation, kanamycin eliminated the bacterial cells without our desired plasmid. The individual colonies on the agar plates were instances of bacterial cells with the minimized R2 units in the 5'-entry plasmid, which provided antibiotic resistance.

Purification of plasmid DNA from transformation

The Zyppy Plasmid Miniprep Kit (Zymo Research) was used for the purification of the 5'-entry plasmids, which separated the plasmids from the components of the bacterial cell. At least five individual colonies from each transformation plate were chosen for inoculum of 2 ml liquid cultures, which were grown for 6-16 hours in a shaking incubator at 37°C. At least six hundred microliters of each culture were placed in an eppendorf tube and lysis buffer was added. The bacterial cells were lysed to release the plasmids into the supernatant, while the remnants of the bacterial cells formed a precipitate. The addition of refrigerated neutralization buffer ceased the reaction. Subsequently the mixture was spun down for the compression of the cellular components into a pellet, while the supernatant was easily separated into a spin-column. The replicated plasmid in the supernatant was bound to the filter, while the remaining liquid was drained through the spin column. The filter on the spin column was rinsed with wash buffers to be certain only the plasmid was bound to the filter, and then the plasmids were released with the addition of an elution buffer and subsequent centrifugation. The final product of the Zyppy Plasmid Miniprep Kit was approximately 30 µl of the 5'-entry plasmid with the minimized R2 units.

Confirmation of accurate plasmid

The miniprep culture with the correct product was initially confirmed with the digestion of the plasmid with restriction enzymes. A mixture that included 400 ng of the plasmid, 10 units of SacI (New England Biolabs), 2 µl of Buffer 1, and 2 µl 10x BSA were incubated to allow for the plasmid to be cut at two sites. Gel electrophoresis

confirmed positive colonies that had the expected bands for the AB-EF plasmids. Each was predicted to give bands of 2377 bp and 387 bp on a 1% agarose gel stained with ethidium bromide. The correct band length on the agarose gel was visualized under ultraviolet light, and at least five hundred nanograms of the miniprep product were sent to GeneWiz Inc (South Plainfield, NJ) to be sequenced using the M13 Forward and M13 Reverse primers. The following day the sequencing results were downloaded from GeneWiz, where they were analyzed using the bioinformatics tool Geneious (Biomatters Ltd.), to compare the sequenced results to the original R2 regulatory region sequence. The program displayed a visualization of the nucleotide sequences of the amplified and control R2 region, and was an assurance that the appropriate nucleotides were present in the 5'-entry vector.

Preparation of larger volume of plasmids

Once the sequencing results were confirmed, a greater amount and concentration of the DNA was accomplished with Zymo Research's Zyppy Plasmid Maxiprep Kit. A 1:1000 dilution of kanamycin was added to 150 µl of LB broth, and one hundred microliters of the miniprep culture with the confirmed band length was used as an inoculum of the broth. The inoculated culture was grown at 37°C overnight in a shaking incubator. After sixteen hours of incubation, the maxiprep culture was spun down in a centrifuge at 4700g for twenty minutes, which produced a bacterial pellet. The pellets were resuspended in 15 mL of buffer, and a lysis buffer was subsequently added. Very similar to the miniprep kit, the bacterial cells were lysed to release the contents of the cells, including the plasmid DNA, into the supernatant. The neutralization buffer was

added and the mixture was stored on ice for 10 minutes. The separation of the precipitate eliminated the components of the bacterial cell, but left the plasmids in the supernatant. Using the vacuum manifold method, the column attached to the vacuum removed the supernatant and left the plasmids attached to the filter. The column was rinsed with a wash buffer, which eliminated any residues that inadvertently stuck to the filter. The addition of three milliliters of the Zyppy Elution Buffer and centrifugation of the column released the plasmid. The final product of the maxiprep procedure was a higher concentration of the 5'-entry plasmids with the minimized R2 units.

Construction of mutated putative transcription factor binding sites

In a process similar to the identification of R2, comparative genomics was utilized to determine which nucleotides should be altered using site directed mutagenesis. By comparing individual nucleotides in the DE region of zebrafish, stickleback, medaka, and tetradon, highly conserved sequences were identified. These sequences were potential transcription factor binding sites, which happened to correlate to the Ets, Runx, and Sox families of transcription factors. Site-directed mutagenesis by PCR changed the specific nucleotides that were important for the binding of these transcription factors. The PCR primers contained the altered nucleotides within the DE sequence and amplified outward, which copied the entire entry vector. These primers were also phosphorylated at the 5' end, which was essential for the ligation of the vector. A mutated vector was created for each potential transcription factor binding site individually, as well as each of the sites in parallel with each other. The addition of DpnI restriction enzyme (New England Biolabs) digested the entry vectors that were methylated. A digest with this restriction enzyme cut

the original methylated DE vector, while the new un-methylated mutated vectors created via PCR remained intact. Gel electrophoresis confirmed the correct band length of the PCR product was 2764 base pairs. Pending correct results, the PCR product was purified with Zymo Research's Zyppy DNA Clean and Concentrator Kit. The Quick Ligation Kit (New England Biolabs) was subsequently used in the recircularization of the phosphorylated ends of the PCR product. The combination of the PCR product, deionized water, 10 µl 2x Buffer, and 1 µl Quick Ligase was incubated at room temperature for 10 minutes. The reaction was put on ice until the standard transformation procedure was completed with kanamycin antibiotic plates. Following the transformation, cultures were started from individual colonies and plasmids were extracted with the Zyppy Plasmid Miniprep Kit. The isolated plasmids were digested with SacI (New England Biolabs) to confirm the correct band lengths were 2377 bp and 387 bp. The plasmids with the appropriate band lengths were sent out for sequencing with an M13 forward primer. When the results confirmed the sequence was correctly mutated, a maxiprep culture was prepared using the miniprep culture as the inoculum.

Generation and Analysis of Minimized R2 Reporter Transgenic Zebrafish

Construction of destination vector

The final expression vector was formed from the recombination of three entry vectors into the final destination vector. The dissected R2 regulatory region was placed in 5'-entry vector as described above. An adenovirus derived E1b transcriptional initiator with a Carp TATA-box were used as a minimal promoter was placed in the middle-entry vector. An enhanced green fluorescent protein (EGFP) gene was placed in the 3'-entry

vector, and it produced the reporter protein that was visualized by fluorescent microscopy. The 5'-entry vectors were constructed in the Dale laboratory; the middle-entry and 3'-entry plasmids were obtained from the Topczewski laboratory at Lurie Children's Research Center and used in previous experiments (Dale and Topczewski 2011). Ten fentomoles of each entry plasmid, 20 fentomoles of the destination vector, tris-EDTA buffer, and 2 μl of LR clonase were added together for the LR reaction and incubated at room temperature overnight. The addition of the LR clonase enzyme assisted with the recombination of the entry vectors into the appropriate orientation in the final destination vector; the reaction was based on the attR sequences present in the plasmids. Proteinase K was added to the reaction and placed at 37°C for 15 minutes, which terminated the reaction by destroying the reaction enzymes. The final expression vector contained the minimized R2 regulatory unit, which drove expression of the EGFP reporter protein. The final expression vector was transformed into Top10 chemical competent bacterial cells (Invitrogen) by heat-shock, and the plasmids penetrated the cell wall of the bacterial cells. Five hundred microliters of LB broth was added to the transformed cells, and placed in a 37°C shaking incubator for one hour, which aided in the division of the bacterial cells and therefore generation of the transgenic plasmid. The bacterial culture was centrifuged into a pellet and resuspended with 125 μl of the culture, which gave a concentrated bacterial culture. The transformation mixture was spread on an ampicillin antibiotic plate and incubated overnight to selectively permit growth of the bacterial cells that took up the expression plasmid. The pDest final destination vector had an ampicillin resistant gene, which was passed onto the bacterial cells that took up the

plasmid. The colonies on the agar plates were the reproduction of individual cells that contained the expression vector with antibiotic resistance.

The procedures used for the isolation of the entry vector plasmids were repeated for the production of a high concentration of the final expression vector. The Zyppy Miniprep Kit isolated the expression vector from the miniprep cultures with ampicillin. The plasmids were cut at two sites when digested with SacI (New England Biolabs), and the expected band lengths for gel electrophoresis were 1700 bp and 3500 bp. The miniprep culture with the correct digest was used for the inoculation of 150 mL of LB broth with ampicillin, and grown overnight at 37°C in a shaking incubator. Zymo Research's Zyppy Maxiprep Kit, was used to extract the plasmid DNA, and provided 30 µl of the final expression vector. This process was also done for the generation of the R2 mutagenesis enhancer elements as well. An LR reaction inserted the mutated region, the E1b-TATA sequence, and the EGFP reporter gene into the pDestTol2 vector. The LR reaction was transformed and placed on ampicillin plates, and the Zyppy Plasmid Miniprep Kit procedure was performed on their cultures to isolate the DNA. The plasmids were digested with SacI (New England Biolab), and confirmed the band lengths were 3512 bp and 1785 bp. The miniprep culture was used to inoculate a maxiprep culture, and a larger volume of the plasmid was isolated with the Zyppy Plasmid Maxiprep Kit

*Zebrafish embryo microinjection*s

The DNA plasmids were injected into embryos at the one-cell stage; the expectation was that the transgenic plasmid would be present in each of the dividing cells

of the developing embryo. The final expression plasmids were diluted to 80 ng/µl of DNA, and added to 1 µl of 4x Danio solution, 1 µl of transposase, and 1 µl of phenol red. The transposase assisted with the integration of the reporter sequence into the zebrafish genome, which aided in the establishment of a stable line for the zebrafish. Phenol red was essential for the visualization of the injection solution into the zebrafish embryos. Capillary tubes were pulled and formed a sharp needle, which penetrated the chorion and embryo for the injection process. The injection solution was loaded into the needles and placed in a microinjector. The zebrafish embryos were organized into rows on agar plates, and microscopes were used to observe the insertion of the needle into the embryo for the injection. Even though the EGFP expression pattern was visualized in the transient zebrafish in a mosaic fashion, a stable line was established in the next generation with the transposition of the transgenic sequence into the germ line cells. The zebrafish that were positive for the EGFP reporter protein were grown to adulthood, and then crossed with wild-type zebrafish. The second generation with EGFP in the tissues provided a confirmation of the results and displayed a clear expression pattern of the reporter protein.

Imaging and analysis of transgenic zebrafish lines.

Following the injections, the zebrafish were prepared for fixation at 5 days post fertilization (dpf). Ice was added to the water in the petri dish containing the zebrafish, and the larval fish were subsequently collected in eppendorf tubes. The water was removed from the eppendorf tubes and replaced with 4% paraformaldehyde (PFA) (Sigma) in phosphate buffered saline (PBS). The zebrafish in PFA solution were stored

overnight on a compact rocker in a 4°C laboratory refrigerator. After about 16 hours in the cold, the 4% PFA solution was removed from the eppendorf tubes and replaced with PBS supplemented with 0.02% Triton-X 100 (PBS-TX) and placed on a laboratory rocker for 15 minutes. The PBS-TX was removed and replaced with fresh PBS-TX solution at least 5 times, which rinsed and removed the remainder of the PFA. The rinsed zebrafish were ready for scoring, and stored at 4°C when not being examined.

The zebrafish tissues were inspected for EGFP expression under the fluorescent microscope. A scoring sheet was designed for the inventory of the EGFP expression in the tissues of the zebrafish, and one sheet was dedicated to each construct injected from the clutch of embryos. The top half of the scoring sheet contained images of the lateral view and ventral view of a zebrafish and focused on their anatomy (Haffter et. al, 1996). These images were used as the basis of the identification of expression in the tissues, and they were also marked for expression in unexpected tissues. The bottom half of the scoring sheet was a table used to tally the expression pattern of the zebrafish, particularly in the expected tissues. The table consisted of columns for expression in the notochord, craniofacial cartilage, ear, heart, skin, muscle, and other tissues. Based on an examination of each zebrafish under the fluorescent microscope, marks were made under the appropriate column when EGFP was present in particular tissues. The scoring permitted a calculation of the total number of zebrafish with EGFP in each of the tissues, which determined the percentage of fish that expressed the reporter gene. The zebrafish were also mounted on slides, and high quality images were captured with an Olympus Fluoview 1200 laser-confocal microscope. The images captured were at a 10x

magnification and a 1:1 aspect ratio with 1024 x 1024 resolution. A Z-scan captured a series of images of the zebrafish at different depths, which allowed for the visualization of EGFP expression in the various tissue layers. Images of the ventral view of the zebrafish head displayed the craniofacial cartilage, while the lateral view images revealed the ear, notochord, and craniofacial cartilage. Each construct was imaged with at least three fish in the ventral view, as well as a lateral view of three fish from the anterior to posterior end.

Construction of Transcription Factor Expression Plasmids

Isolation of transcription factor sequences

The construction of the transcription factor expression vector began with the isolation of the RNA from a zebrafish embryo digest. Thirty-five zebrafish embryos were euthanized according to our IACUC approved protocol, and placed in a microcentrifuge tube and measured to determine the total weight. One milliliter of Invitrogen TRIzol was added per 50 mg of tissue, and the solution was homogenized with a syringe and spun down for one minute. An equivalent volume of ethanol was added to homogenate, and the mixture was loaded onto the spin column and spun down for one minute. Four hundred microliters of Direct-zol RNA pre-wash from Zymo was added to the column and spun down for one minute. The flow through was discarded and this process was repeated. Two hundred microliters of the RNA Wash Buffer was added to the column and the flow-through was once again discarded. The column was transferred to a new microcentrifuge tube, and 30 µl DNase/RNase-free water were added to the column and spun down for the final RNA product. The complementary DNA sequences of the

transcription factors were created from the zebrafish RNA using Life Technologies' Superscript III Reverse Transcriptase Kit. The total RNA, primers, a dNTP mix, and deionized water were added to a nuclease-free microcentrifuge tube and heated for 5 min at 65°C, followed by an incubation for one minute at 4°C. First strand buffer, 0.1 DTT, and superscript III reverse transcriptase were added to each microcentrifuge tube and incubated at 25°C for 5 min, 50°C for 60 min, and 70°C for 15 min. The addition of RNase H and an incubation of 37°C for 20 minutes degraded the RNA, which left all the complementary DNA strands, including the sequences for our transcription factors. Primers were designed to isolate the sequence for the specific transcription factors from the total zebrafish cDNA.

Table 2. Primers for isolation of transcription factors

Transcription Factors	**Forward Primer**	**Reverse Primer**
Ets1	CGTTTGAATGCGTGACCAT GACGGC	GCAGGATTTATCCGTCAGG AGCTCC
Sox9a	CCATCTACGGTGTTACCAT GAATC	CATTCAGGCGTGCTCATGG TCTGG
Runx3	CTG TAG CCT ACT CAA CCA ACT G	ATC ATG CGC AAC TCT TCT GGT C

A PCR reaction using this cDNA, 1.25 µl (1µM) designated primers (Table 2), 10 µl iProof Buffer, 1 µl (200 µM) dNTPs, 0.5 µl (1 U) iProof polymerase, and deionized water were used to isolate the desired transcription factor sequence with the following protocol:

Step 1	98°C	1 min
Step 2	98°C	30 sec
Step 3	60°C	30 sec
Step 4	72°C	45 sec
Repeat Step 2-4 x34 times		
Step 5	72°C	10 min
Step 6	4°C	10 min
Step	15°C	∞

The product of this PCR reaction was run on a 1% low melt agarose gel, and subsequently separated and purified using the Zymoclean Gel DNA Recovery Kit. The isolated DNA was used in a second PCR reaction, which amplified the same transcription factor sequence with different primers that contained attB sites (Table 3). The PCR product was again run on a 1% low melt gel, and the correct bands were once again separated using Zymoclean Gel DNA Recovery Kit.

Table 3. Primers with attB sites for isolation of transcription factors

Transcription Factors	**Forward Primer**	**Reverse Primer**
Ets1*	GGGGACAAGTTTGTACAAAAA AGCAGGCTAGCGTTTGAATGC GTGACCATGACGGC	GGGGACCACTTTGTACAAGAA AGCTGGGTAGCAGGATTTATC CGTCAGGAGCTCC
Sox9a*	GGGGACAAGTTTGTACAAAAA AGCAGGCTAGCCATCTACGGT GTTACCATGAATC	GGGGACCACTTTGTACAAGAA AGCTGGGTACATTCAGGCGTG CTCATGGTCTGG
Runx3*	GGGGACAAGTTTGTACAAAAA AGCAGGCTAGTTGTGATGCAT ATTCCCGTAGACC	GGGGACCACTTTGTACAAGAA AGCTGGGTATCTTAGTACGGC CTCCAGACAGAC
Fli1a	GGGGACAAGTTTGTACAAAAA AGCAGGCTAGTTGTGATGCAT ATTCCCGTAGACC	GGGGACCACTTTGTACAA GAAAGCTGGGTATCTTAGTAC GGCCTCCAGACAGAC

Transcription factor sequence inserted into middle-entry vector

Using Gateway Technology, the isolated DNA sequence of the transcription factors were inserted into the middle-entry vector using a BP reaction. The gel purified PCR products had the appropriate attB sites for recombination into the middle-entry vector. A mixture of 150 ng of the PCR product, 150 ng of the pDONR 221 vector, TE Buffer, and 1 µl of BP clonase enzyme were added and incubated at room temperature overnight. The following day 1 µl of proteinase K was added to terminate the reaction with a 15 minute incubation at 37°C. This reaction generated middle-entry vectors with

the newly inserted PCR product, which was then transformed into bacterial cells. The middle-entry plasmids had a kanamycin resistance gene, therefore the colonies on the antibiotic- agar plates were indicative of bacterial cells with the desired plasmid.

Isolation of middle-entry plasmids

The Zyppy Plasmid Miniprep Kit standard procedure isolated the middle-entry plasmids for the DNA sequence of *sox9a, ets1*, and *runx3*. The middle-entry vector for *ets1* was digested with SacI (New England Biolabs) for the expected band sizes of 29998 bp and 585 bp. The expected bands for *sox9a* when digested with ApaLI (Fermentas) were 2295 bp and 1671 bp. The middle-entry vector for *runx3* was digested with HincII (New England Biolabs), and the expected band lengths were 3154 bp and 675 bp. The bands that travel the correct distance according to the ladder had the potential to be accurate, so they were sent out for sequencing. Five hundred nanograms of the plasmids were sent out to GeneWiz for sequencing with the M13 forward and M13 reverse primers to determine the accuracy of the sequences. When the sequence accuracy was confirmed, 100 µl of the corresponding miniprep culture was used to inoculate 150 mL of LB broth with kanamycin. Standard maxiprep procedures with the Zyppy Plasmid Maxiprep Kit isolated a large volume of the desired plasmid.

Transcription factor expression vectors

Gateway technology was once again employed to insert the transcription factor sequence into the final expression vector with homologous recombination. A mixture of 150 ng of the middle-entry vector with the transcription factor sequence, 150 ng of the pCSDest vector, TE Buffer, and 1 µl LR clonase enzyme were added together and

incubated at room temperature overnight. The following day, 1 µl of proteinase K was added to the reaction with a subsequent 15-minute incubation at 37°C to terminate the reaction. The expression vector was transformed into bacterial cells, and the ampicillin plates were incubated overnight at 37°C. Individual, isolated colonies from the bacterial plates were used to make 2 mL bacterial cultures and the Zyppy Plasmid Miniprep Kit was used for the isolation of the plasmids. The plasmids from the miniprep protocol were digested with various restriction enzymes to check the lengths of the bands. The expression vector for *runx3* was digested with HincII (New England Biolabs), with the correct band lengths of 3154 bp and 675 bp. A PvuII (New England Biolabs) digested the expression vector of *ets*1a, which formed two bands of 3661 bp and 1531 bp. Also, the expression vector of *sox9a* was digested with ApaLI (Fermentas), with accurate bands lengths of 2579 bp, 1750 bp, and 1246 bp. For the DNA with the correct band lengths, the corresponding miniprep culture was used to inoculate 150 mL of LB broth with ampicillin. This culture was used for the Zyppy Maxiprep Plasmid Kit and the isolation of the desired plasmids.

Generate transcription factor mRNA

The transcription factor expression vectors were digested with the appropriate restriction enzymes, purified with phenol:chloroform, and finally RNA was synthesized using the mMessage mMachine Kit (Life Technologies). Ten micrograms of the expression vectors for Runx3, Ets1, and Sox9a were digested with NotI into 100 µl reactions, and incubated for three hours. After digestion, the volume was brought up to 200 µl with RNA-free water. An equal amount of phenol:chloroform:isoamyl alcohol was

added to the solution. Each tube was centrifuged at 16,000g for 5 minutes. When the aqueous layer was recovered, 200 µl of chloroform was added to the solution, centrifuged again for 3 minutes and repeated. The upper aqueous layer was once again recovered and the DNA was precipitated with the addition of 20 µl of sodium acetate and 500 µl of ethanol. This solution was mixed and incubated at 20°C for 30 minutes and subsequently spun down at 4°C for 20 minutes at 16,000g. The supernatant was discarded and the tube was left to dry. Five hundred microliters of 70% ethanol was added to the tubes and spun down at 4°C for 10 minutes at 16,000g. The sample was resuspended with 25 µl of RNAse free water. Finally, the sample was run on a 1% agarose gel to verify the presence and accuracy of the results. This process provided the purified transcription factor expression vectors, but the synthesis of the RNA can be completed with the mMessage mMachine Kit (Life Technologies). The following solutions were added to an eppendorf tube and incubated at 37°C for 2 hours: 10 µl 2x NTP/CAP, 2 µl 10x reaction buffer, 1 µg isolated DNA, 2 µl of the enzyme mix, and RNAse free water if necessary. Any remaining DNA was eliminated with the addition of 1 µl TURBO DNAse (Ambion), which was added to each tube and incubated at 37°C for 15 minutes. The RNA product was purified using NucAway spin column (Ambion), which placed the sample on the hydrated gel bed and spun the column in a centrifuge at 750g for 2 minutes. This completed the synthesis of the transcription factor RNA, which was prepared for microinjections.

Zebrafish embryo microinjections

The RNA of the transcription factors was injected into the embryos at the one-cell stage, which was at an earlier stage than the genes endogenous period of transcription. The transcription factor mRNAs were diluted to 300 ng/µl of RNA, and added to 4x Danio solution and phenol red. The mixture of transcription factor RNA was injected into the stable line R2-DE embryos, which would produce the protein and theoretically initiated transcription of the EGFP reporter gene. The binding of the transcription factors to the DE region would be confirmed with the presence of the reporter protein prior to normal expression.

Zebrafish Husbandry

Zebrafish housing and care was maintained as per LeClair *et al.* (2009) for wild-type and transgenic zebrafish lines in the vertebrate animal care facility at Loyola University Chicago (Chicago, IL). A natural spawning technique was used to obtain embryos for these experiments, and they were staged according to Kimmel *et al.* (1995). Zebrafish for these experiments were scored based on the morphology labeled by Kimmel *et al*. (1995). The protocols and care of the zebrafish were approved by IACUC of Loyola University Chicago (Chicago, IL).

CHAPTER THREE

RESULTS

Identification of R2-DE as the col2a1a Regulatory Region

Our laboratory has previously identified a critical zebrafish regulatory sequence of the *col2a1a* gene, referred to as R2, allowing for the expression of Col2a1a protein in the craniofacial cartilage, ear, and notochord utilizing zebrafish transgenic report analysis (Dale and Topczewski 2011). To identify what segment of the R2 unit is critical for tissue specificity, we set out to dissect the previously identified sequence. To be able to compare my work to that of the previous transgenic reporter analysis, I repeated the original Dale and Topczewski 2011 transient transgenic R2 experiments to demonstrate my technique was similar to the original work. During my study, the injections completed with the original R2 transgenic plasmid had EGFP reporter expression in about 30% of the embryos (Table 4). When compared, there seemed to be no difference in expression pattern compared to the original work, allowing us to set out to dissect the R2 element.

While the identification of the R2 regulatory element was originally based on its ability to recapitulate expression in critical tissues, it was also based on a teleost fish conserved sequence of 360 bp, which could contain many overlapping transcription factor binding sites. We set out to minimize the R2 regulatory element to determine exactly which bases of the original 360 are responsible for the transcriptional regulation

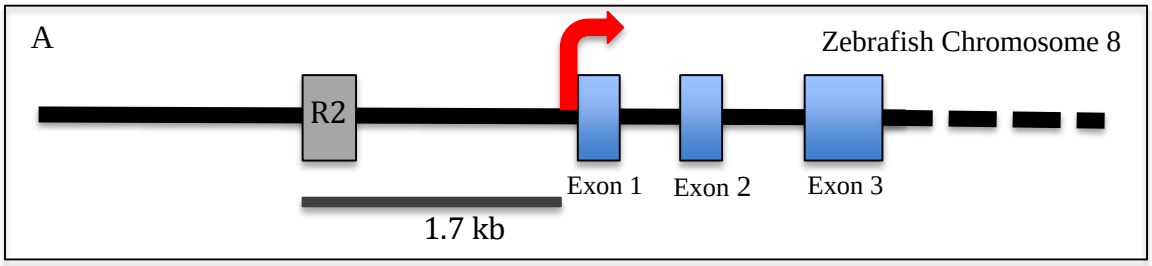
of *col2a1a*. To do this we utilized a combination of Gateway cloning, zebrafish transgenesis, and fluorescence microscopy to identify the critical sequence for gene expression. The R2 regulatory region was divided into 5 overlapping fragments (AB, BC, CD, DE, EF), and placed upstream of an EGFP reporter gene. The first 120 bp of the R2 regulatory region, AB, was not able to drive expression of the reporter protein when examined under the fluorescent microscope (Figure 3). Of the 233 embryos that were injected and scored at 5 days post fertilization (dpf), none of the juvenile zebrafish expressed the reporter protein in the craniofacial cartilage and only one EGFP positive cell was observed in the ear and notochord in one zebrafish, suggesting that zebrafish to be a statistical outlier when the overall totals were compiled (Table 5). The embryos injected with the plasmid driven by the BC subunit also lacked expression of the reporter protein (Figure 4). There were no embryos with expression in the notochord of the 191 embryos injected (Table 6). There were two zebrafish with slight expression in the craniofacial cartilage and one with expression in the ear, but again this was not statistically significant. These injections eliminated the first 180 bp of the R2 element responsible for the majority of *col2a1a* expression. Of the 150 embryos injected with the reporter plasmid that was driven by the CD subunit, there was no evident EGFP expression when examined and scored under the fluorescent microscope (Table 7; Figure 5). EGFP expression was present in all expected tissues of the zebrafish injected with the DE subunit and recapitulated the original R2 expression pattern seen (Figure 6). Of the 221 total zebrafish embryos injected with the DE construct, about 30% of the zebrafish expressed the EGFP reporter in the craniofacial cartilage, ear, and notochord (Table 8). I

also did notice reporter expression in cranial neurons in these transient transgenic lines as was seen in the original transient R2 element characterization, but this expression was lost in the stable lines. The results of these DE injections paralleled the expression patterns of my initial experiments of the entire R2 region. Finally, when the expression of the reporter plasmid was driven by the EF subunit, there was again no fluorescent seen in 109 of the zebrafish examined and scored (Table 9; Figure 7). Even though the embryos injected with the CD and EF subunit vectors had no visible fluorescent protein under the dissecting microscope, when these embryos were imaged using a laser-confocal microscope, there seemed to be only a faint outline of possible expression in the craniofacial cartilage. This was interesting because the CD and EF reporter plasmids each contained half of the DE subunit, which seems sufficient to phenocopy R2 activity in the desired tissues. Even though the CD region contained the D region, and EF contained the E region, it was not sufficient to drive expression like the DE region.

To test if either the D or E region could alone drive reporter activity and therefore be the critical sequence of the regulatory region, the DE regions were split into 60 bp enhancer fragments regulating the expression of EGFP (Figure 9). Reporter plasmid expression driven with the R2-D unit did not yield GFP positive tissues in the 44 embryos injected (Table 13). Of the 44 embryos injected with the R2-E unit only 2% had EGFP expression in the notochord, 5% in the craniofacial cartilage, and no expression in the ear (Table 14). Neither unit D nor unit E was sufficient to drive expression of the EGFP reporter protein (Figure 13, 14). This suggests that when this DE regulatory region

was split in half, and formed these D and E units, it eliminated important potential transcription factors binding affinities.

Unlike the DE subunit, the reporter plasmids with the remaining R2 conserved genomic sequence (AB, BC, CD, and EF) were not able to drive expression of the EGFP reporter. The minimal R2 regulatory region was 120 bp in the DE region, and was sufficient to drive expression of the reporter protein in most of the tissues that *col2a1a* has been observed (Figure 6). These results reduce the functional transcriptional regulation element region of R2 to 120 bp. Multiple stable transgenic lines were generated with the DE transcriptional element regulating EGFP expression to confirm our initial findings. I found that the stable R2-DE transgenic zebrafish expression was the same as the full R2 stable transgenics (Figure 8). These results demonstrated the critical region for driving *col2a1a* in the cartilage, notochord, and ear resided in the R2-DE sequence.



B **R2**

GCCCTCTGACACCTGATGCCAATTGCGGTCAGTGTTTTGCTGGCGACACA
GATTCTTGTGCCAATGGCCAGGCCCCTCATCATCTGATCCGCAGCAACCC
AGCCACCCTACACACCCCTGGAGCCTCTCCGTGTTCTCCTCATCCCTCTAC
CTTTCCGCACTCTCCCTCCATCCACACCCGCGGCTCTCTTCTCCCCCACTG
CCCGGTGCTCTCTCACATTCCTCAGGTCTGCACACAGAGCCGCATTGTGTG
TGTGTCTTACAGAGCACACAGTCAGGGCTCATTTCGGACACACACACACA
TCCAACAGGGTGTGTGCACAGTCGCAGCGATGCGTACACACACATACACA
TATCCCT

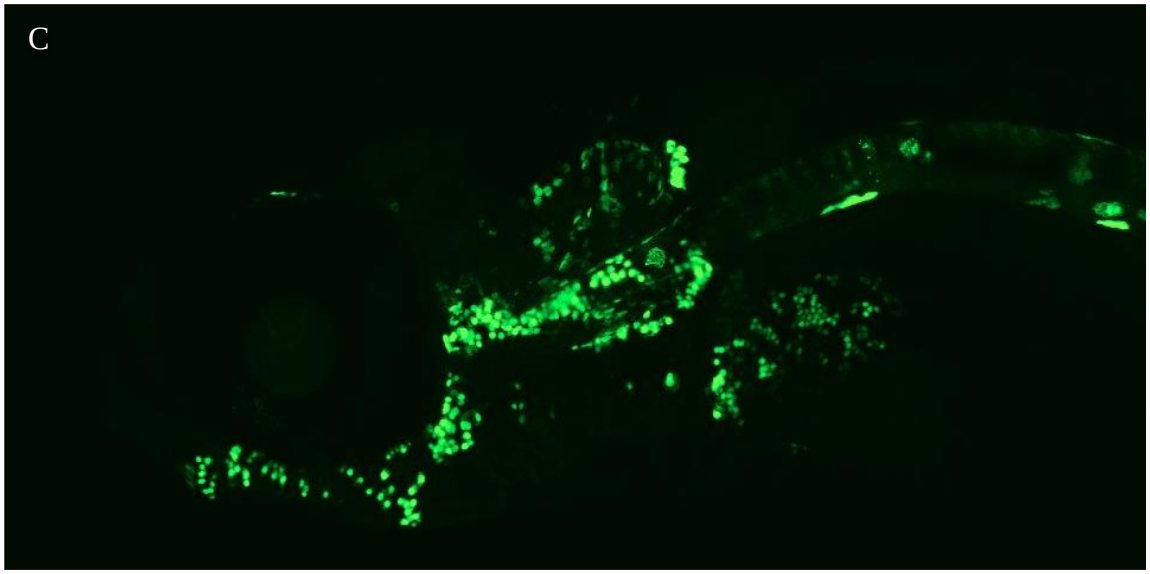
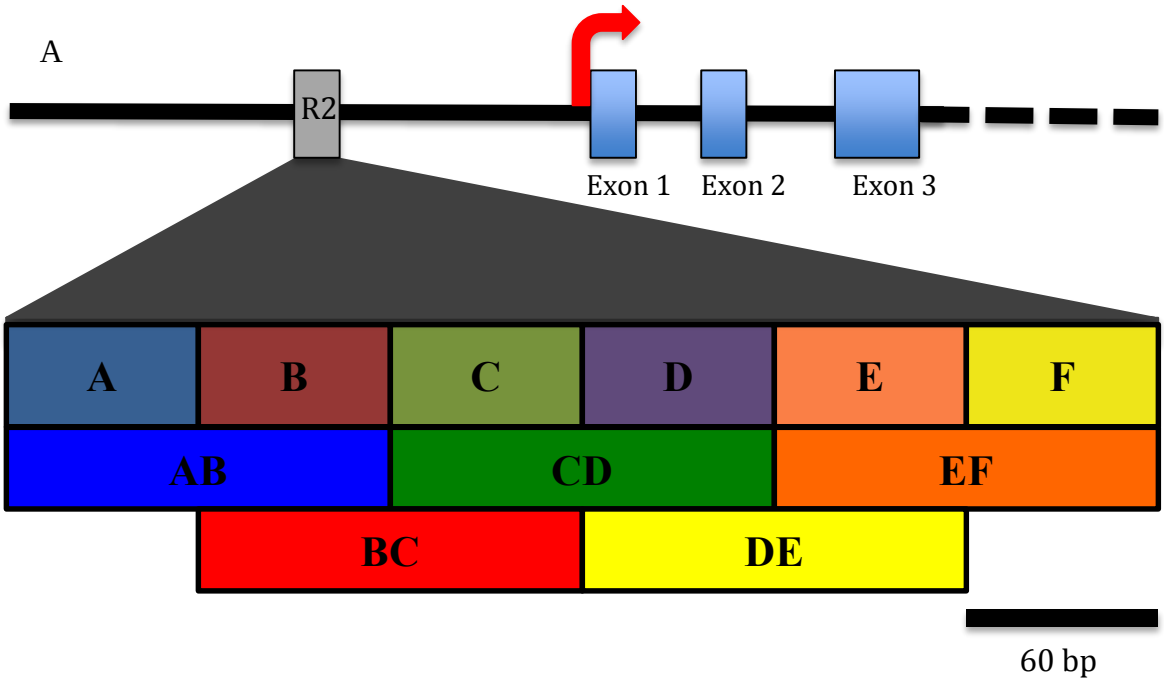


Figure 1. R2 Regulatory Region Sequence and Results. (A) Schematic of the location of the R2 regulatory region on zebrafish chromosome 8. (B) Genomic sequence of R2 regulatory region. (C) Injection of R2 driving expression of EGFP reporter gene, lateral view of zebrafish at 5dpf.

Table 4. Results of R2 injections.

R2	Date Injected	Fish Injected	Notochord		Craniofacial Cartilage		Ear	
			# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive
	3/20	4	2	50%	2	50%	1	25%
	2/28	67	20	30%	17	25%	10	15%
	Total	71	22	31%	19	27%	11	15%



B **R2**

GCCCTCTGACACCTGATGCCAATTGCGGTCAGTGTTTTGCTGG
CGACACAGATTCTTGTGCCAATGGCCAGGCCCCTCATCATCTG
ATCCGCAGCAACCCAGCCACCCTACACACCCCTGGAGCCTCTC
CGTGTTCTCCTCATCCCTCTACCTTTCCGCACTCTCCCTCCATC
CACACCCGCGGCTCTCTTCTCCCCCACTGCCCGGTGCTCTCTCA
CATTCCTCAGGTCTGCACACAGAGCCGCATTGTGTGTGTGTCT
TACAGAGCACACAGTCAGGGCTCATTTCGGACACACACACAC
ATCCAACAGGGTGTGTGCACAGTCGCAGCGATGCGTACACAC
ACATACACATATCCCT

Figure 2. Minimization of R2 regulatory region. (A) Schematic depicting the dissection of the R2 regulatory region, including 60 bp fragments and 120 bp overlapping fragments. (B) Genomic sequence of R2 regulatory region with colors based on the smaller R2 fragments (Blue-A; Red-B; Green-C; Purple-D; Orange-E; Yellow-F)

Table 5. Results of zebrafish injected with R2-AB plasmid.

	Date Injected	Fish Injected	Notochord		Craniofacial Cartilage		Ear	
			# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive
AB	2/21	17	1	6%	0	0%	1	6%
	2/27	36	0	0%	0	0%	0	0%
	3/22	34	0	0%	0	0%	0	0%
	3/28	146	0	0%	0	0%	0	0%
	Total	233	1	0%	0	0%	1	0%

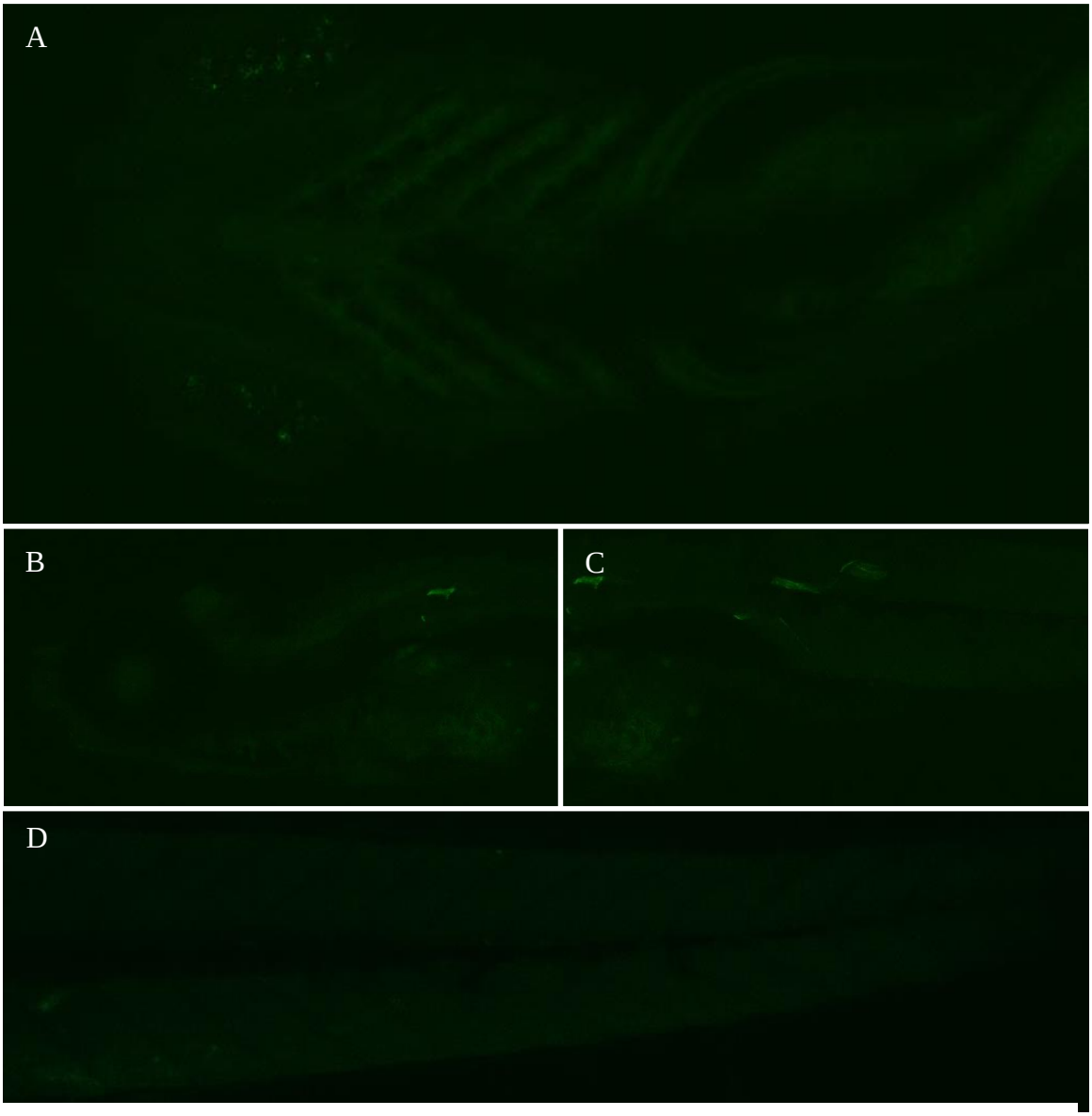


Figure 3. Results of R2-AB plasmid injections. No EGFP expression with R2-BC driving expression. (A) Ventral view of craniofacial cartilage at 5 dpf. (B,C,D) Lateral view of zebrafish head, mid-section, and trunk respectively at 5 dpf.

Table 6. Results of zebrafish injected with R2-BC plasmid.

	Date Injected	Fish Injected	Notochord		Craniofacial Cartilage		Ear	
			# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive
BC	2/27	34	0	0%	0	0%	0	0%
	3/22	45	0	0%	1	2%	0	0%
	3/24	112	0	0%	1	1%	1	1%
	Total	191	0	0%	2	1%	1	1%

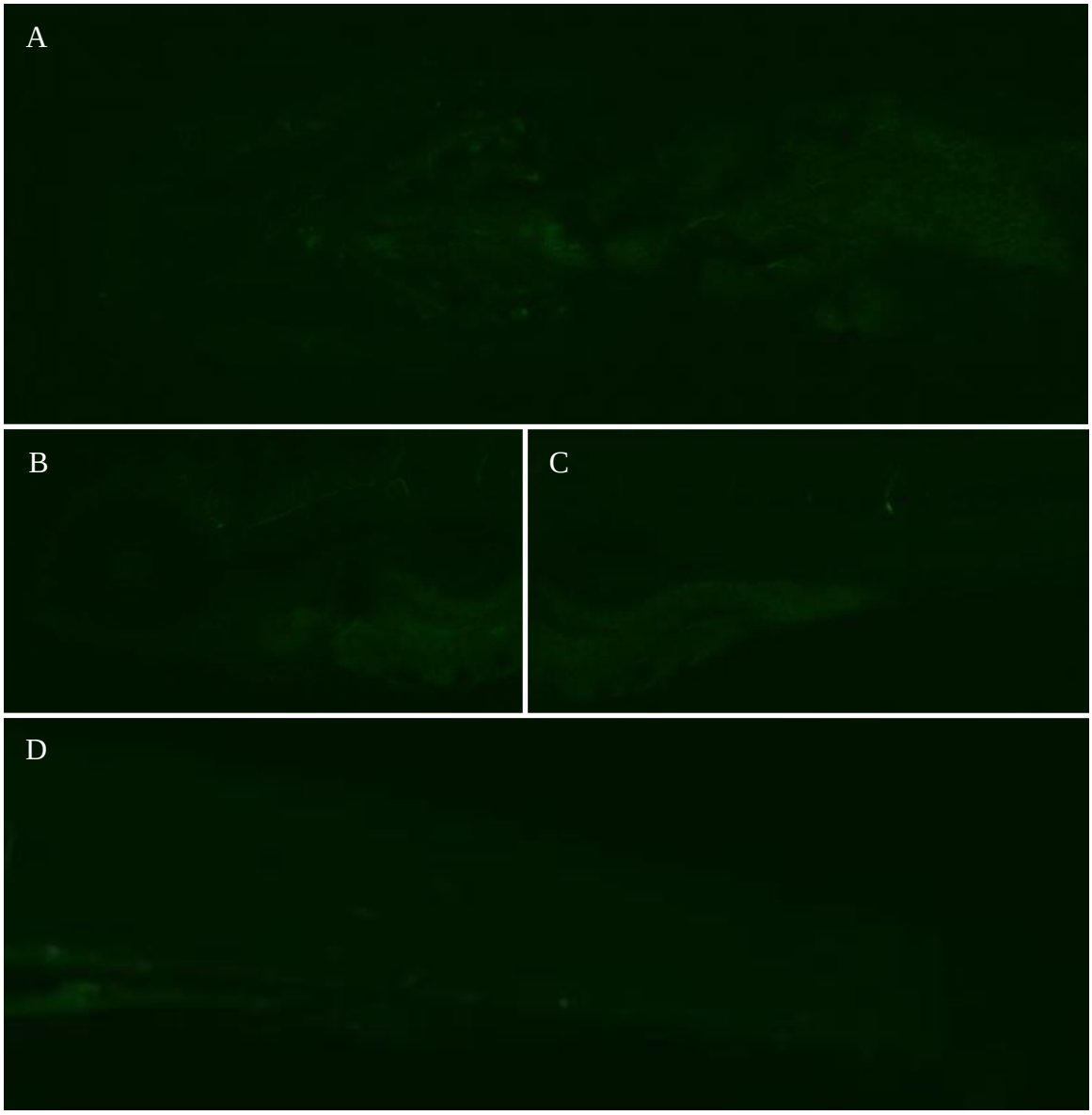


Figure 4. Results of R2-BC plasmid injections. No EGFP expression with R2-BC driving expression. (A) Ventral view of craniofacial cartilage at 5 dpf. (B,C,D) Lateral view of zebrafish head, mid-section, and trunk respectively at 5 dpf.

Table 7. Results of R2-CD driving expression of EGFP.

CD	Date Injected	Fish Injected	Notochord		Craniofacial Cartilage		Ear	
			# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive
	2/21	117	0	0%	0	0%	0	0%
	2/28	33	0	0%	1	3%	0	0%
	Total	191	0	0%	1	1%	0	0%

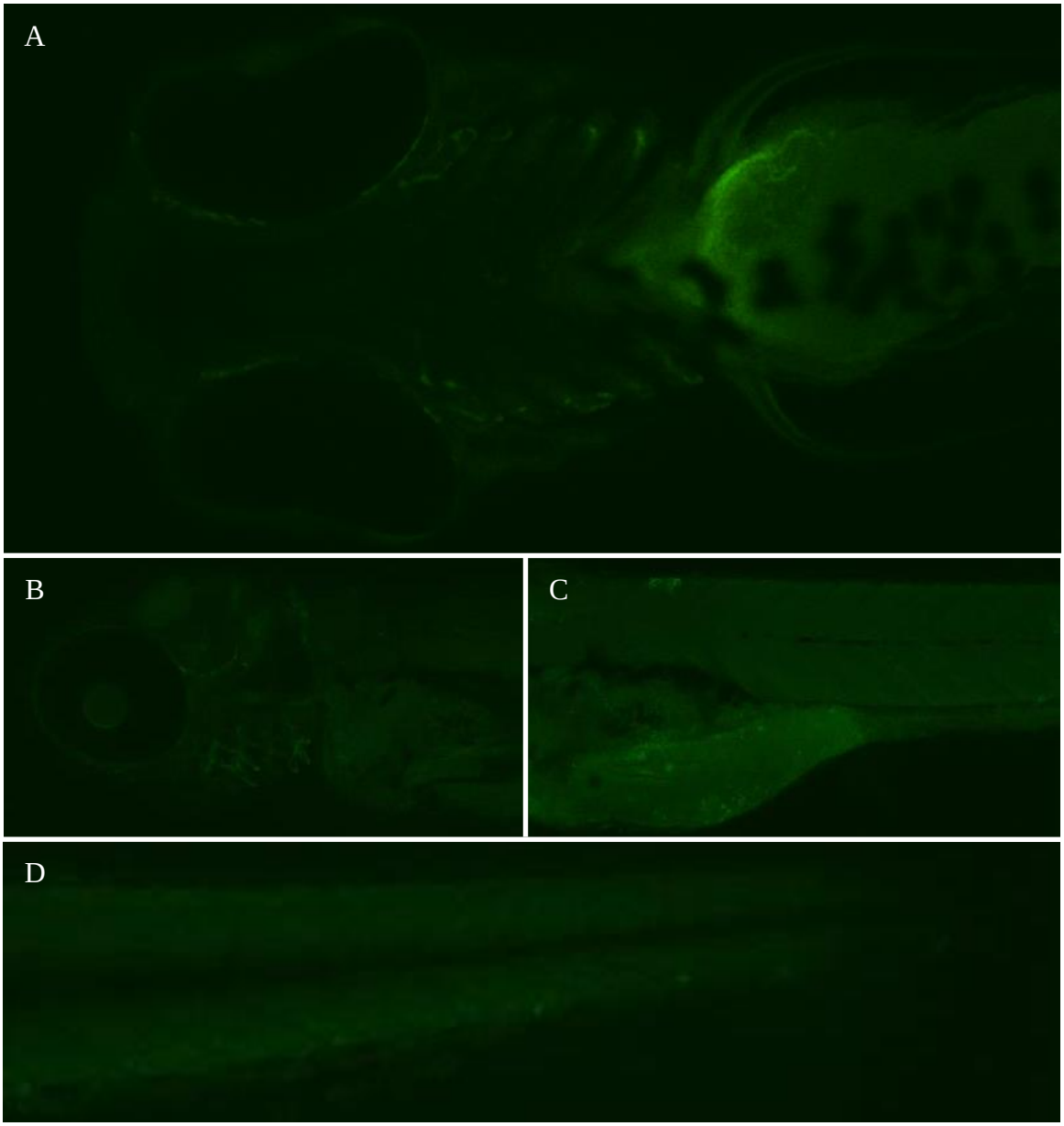


Figure 5. Results of R2-CD plasmid injections. No of EGFP expression with R2-CD driving expression. (A) Ventral view of craniofacial cartilage at 5 dpf. (B,C,D) Lateral view of zebrafish head, mid-section, and trunk respectively at 5 dpf.

Table 8. Results of R2-DE driving expression of EGFP reporter gene.

	Date Injected	Fish Injected	Notochord		Craniofacial Cartilage		Ear	
			# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive
DE	2/21	67	6	9%	6	9%	6	9%
	2/27	62	18	29%	18	29%	17	27%
	2/28	92	43	47%	49	53%	49	53%
	Total	221	67	30%	73	33%	72	33%

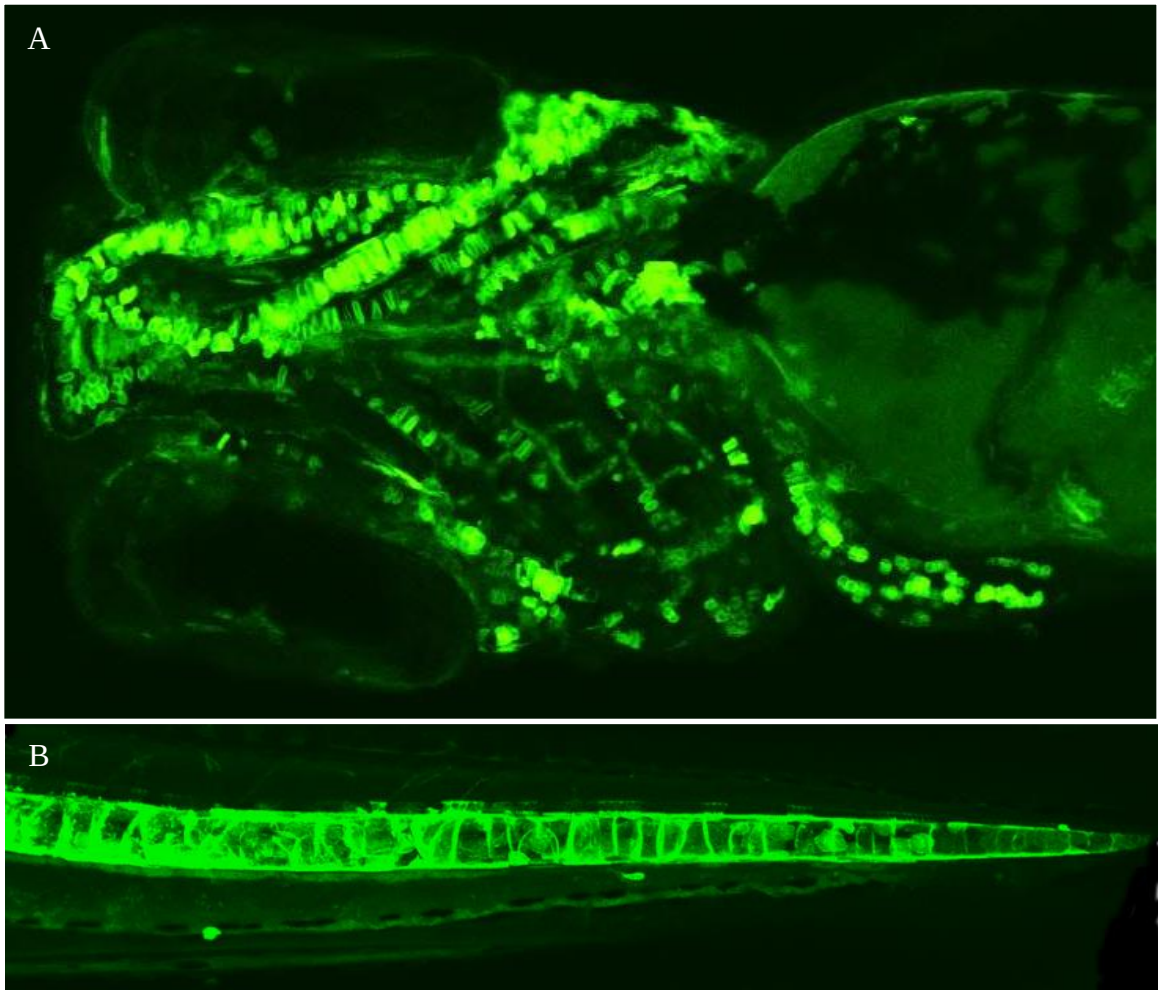


Figure 6. Results of R2-DE plasmid injections. EGFP present in craniofacial cartilage, ear, and notochord. (A) Chondrocytes highlighted by EGFP expression in craniofacial cartilage. Ventral view of zebrafish at 5 dpf. (B) Lateral view of zebrafish tail at 5 dpf depicting expression of EGFP in notochord. (C) EGFP expression in craniofacial cartilage, ear, and notochord visible in lateral view of zebrafish head at 5 dpf. (continued on next page) (D) Zebrafish midsection at 5 dpf showing expression of EGFP in notochord.

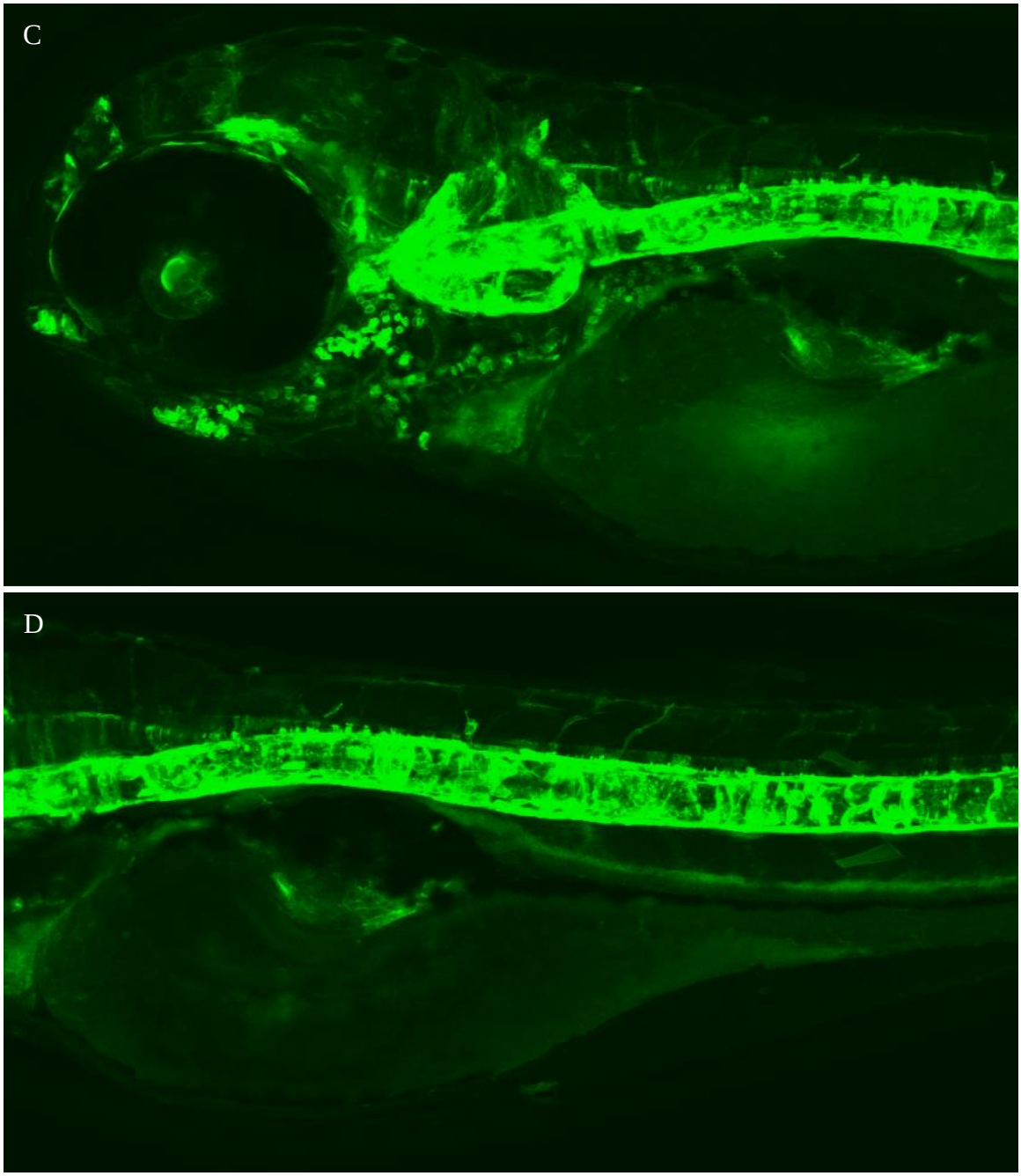


Figure 6. (continued) Results of R2-DE plasmid injections. EGFP present in craniofacial cartilage, ear, and notochord. (A) Chondrocytes highlighted by EGFP expression in craniofacial cartilage. Ventral view of zebrafish at 5 dpf. (B) Lateral view of zebrafish tail at 5 dpf depicting expression of EGFP in notochord. (C) EGFP expression in craniofacial cartilage, ear, and notochord visible in lateral view of zebrafish head at 5 dpf. (continued on next page) (D) Zebrafish midsection at 5 dpf showing expression of EGFP in notochord.

Table 9. Results of R2-EF driving expression of EGFP.

	Date Injected	Fish Injected	Notochord		Craniofacial Cartilage		Ear	
			# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive
EF	2/21	41	0	0%	1	2%	0	0%
	2/27	22	0	0%	1	5%	0	0%
	2/28	46	1	2%	0	0%	0	0%
	Total	109	1	1%	2	2%	0	0%

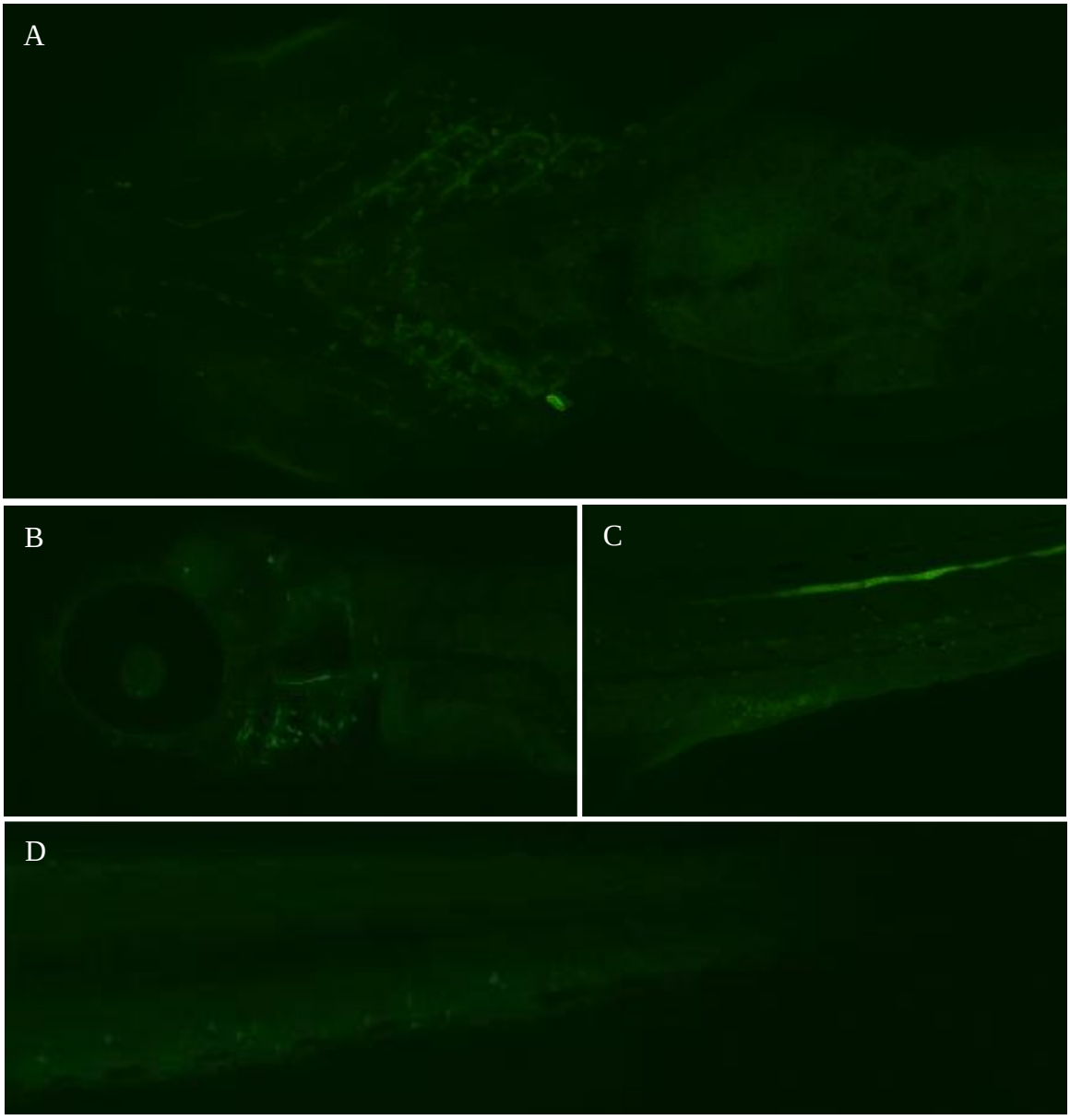


Figure 7. Results of R2-EF plasmid injections. No of EGFP expression with R2-EF driving expression. (A) Ventral view of craniofacial cartilage at 5 dpf. (B,C,D) Lateral view of zebrafish head, mid-section, and trunk respectively at 5 dpf.

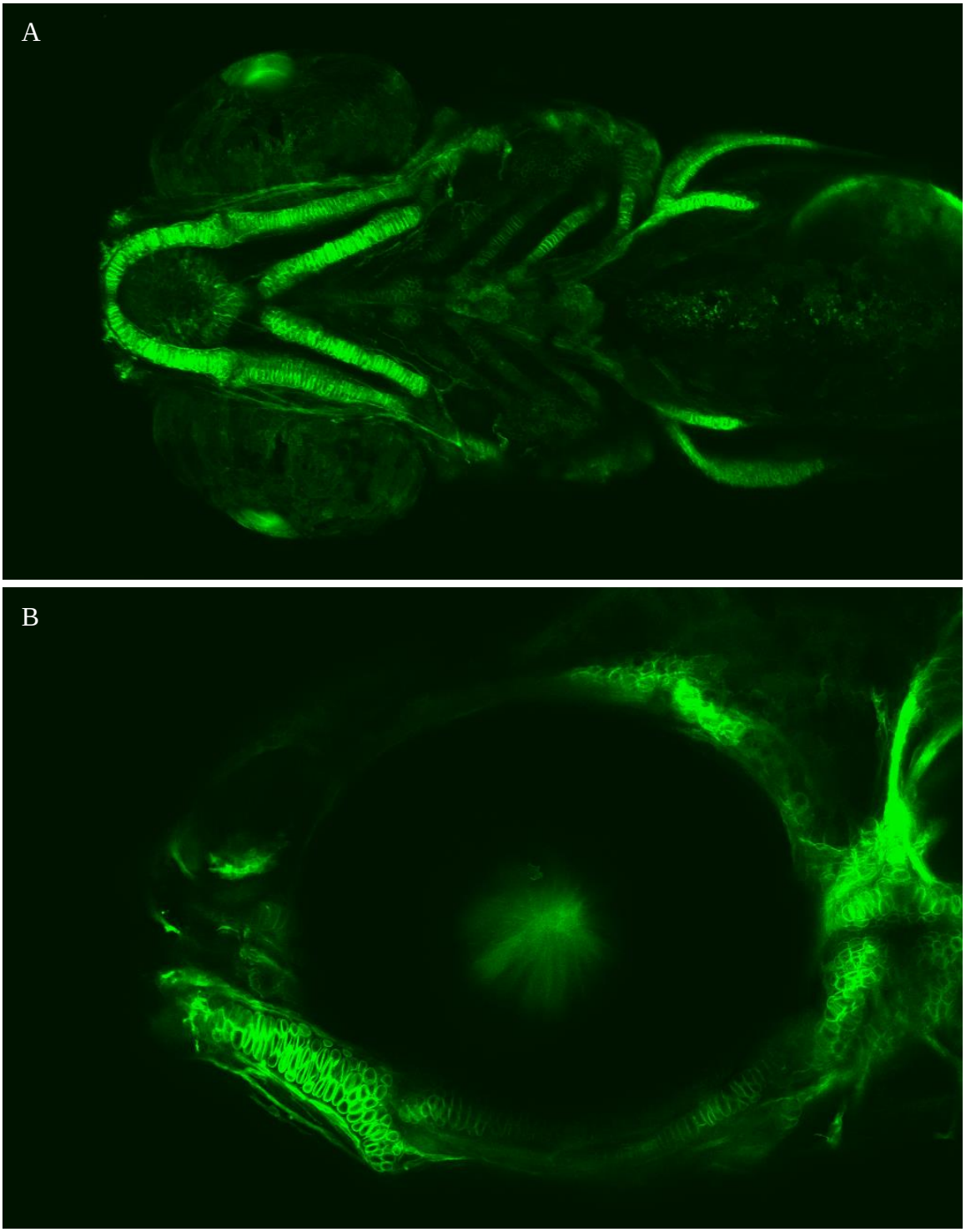


Figure 8. Results of R2-DE stable line. EGFP expression in R2-DE stable line zebrafish. (A) Ventral view of craniofacial cartilage at 5 dpf. (B) Lateral view of zebrafish craniofacial cartilage with EGFP highlighting the chondrocytes.

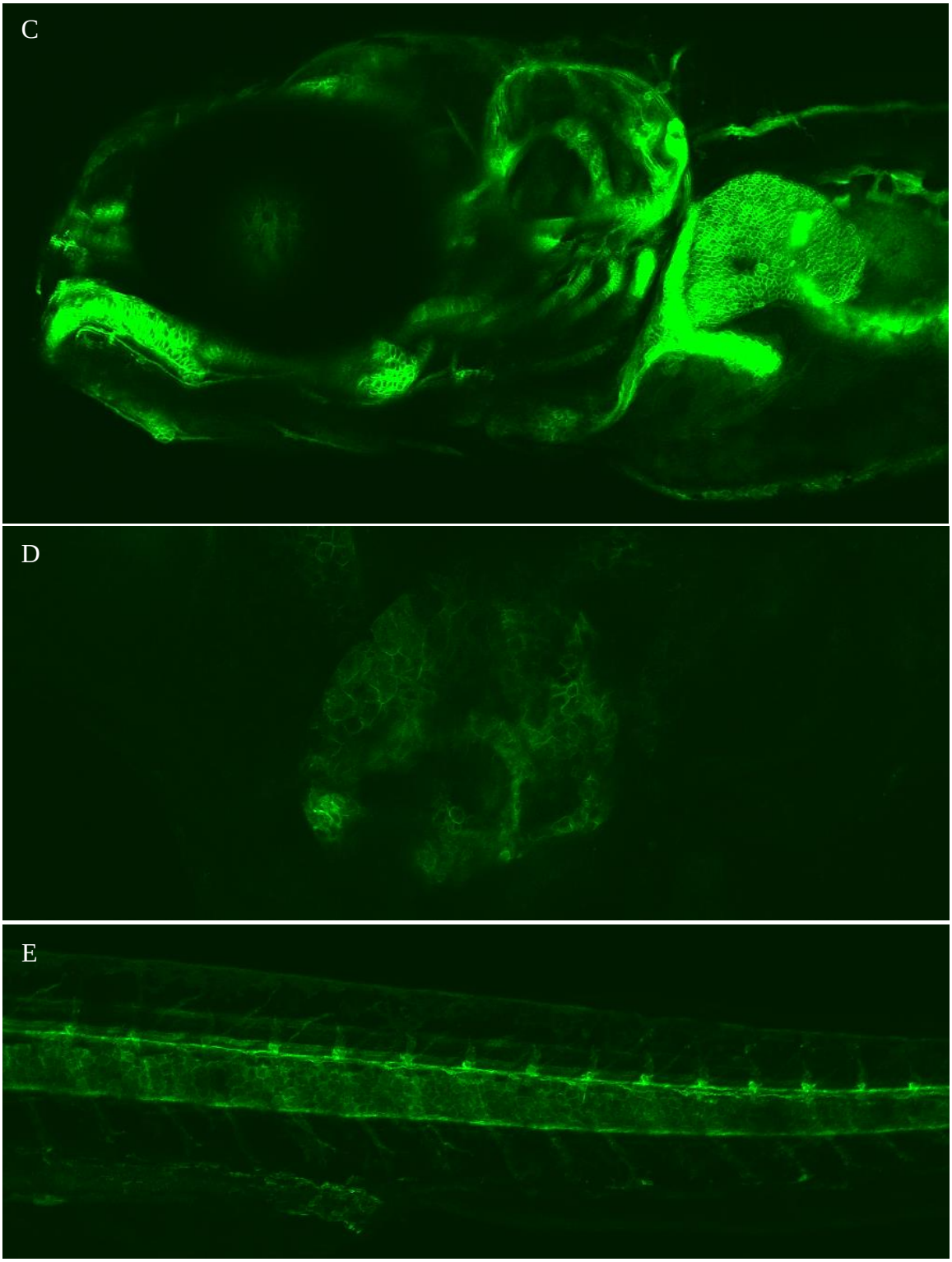


Figure 8. (continued) Results of R2-DE stable line. Images of the EGFP expression in the R2-DE zebrafish stable line. (C) Lateral view of zebrafish head at 5 dpf. Expression of EGFP present in the craniofacial cartilage, ear, pec fin. (D) 20x magnification of zebrafish ear expressing reporter. (E) Lateral view of EGFP expression in the zebrafish notochord.

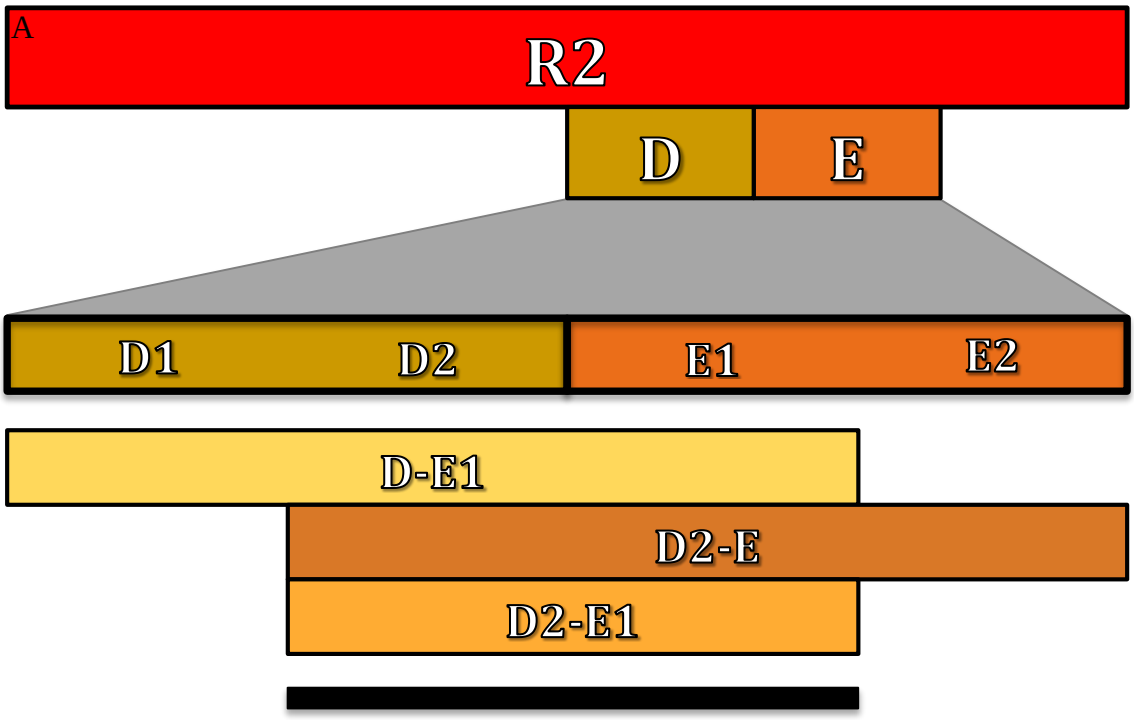
Tissue Specificity of R2-DE Regulatory Region

The identification of the 120 bp *col2a1a* R2-DE regulatory region that recapitulates the cartilage, ear, and notochord expression of the full R2 allows us to try and determine what parts of this 120 bp sequence is critical for its tissue specificity. To do this we divided the DE region into smaller fragments which will provided us with the ability to dissect the tissue specificity in this regulatory unit. The DE unit was divided into three fragments (Figure 9). The first regulatory fragment, D-E1, encompassed all of D and the first 30 bp of the E region. Of the 131 embryos injected, the expression plasmid driven by this fragment showed a decline in the expression level of the notochord to 7% and ear expression was down to 8% (Figure 10). The craniofacial cartilage expression was more than double compared to the other tissues, with expression of the reporter protein in about 20% of the zebrafish. With the elimination of the last 30 bp of the E regulatory region, the expression levels of the notochord and ear were decreased (Table 10). The second regulatory fragment, D2-E, encompassed the last 30 bp of D region and all of the E region. Two hundred and fifteen embryos were injected with the D2-E reporter and subsequently scored for expression (Table 11). Of the zebrafish embryos injected, about 25% of the zebrafish had fluorescence in the notochord, 62% in the craniofacial cartilage, and 45% in the ear. Even though the notochord expression was slightly lower than expected, the expression of the craniofacial cartilage and ear was maintained with this 90 bp region (Figure 11). The third regulatory element consisting of the center region of the DE region, D2-E1 (Figure 9). Even though the D2-E1 region was minimized to only 60 bp, there continued to be EGFP expression in the craniofacial

cartilage and ear similar to what was seen with the full R2-DE regulatory region (Table 12). The EGFP reporter protein was present in the craniofacial cartilage for about 49% of the embryos injected and scored, and also expressed in the ear for about 25% of the zebrafish (Figure 12). While expression in these two tissues was maintained, reporter protein expression was completely lost in the notochord, which was down to only 3% of the zebrafish.

The minimization of the DE region to 60 bp still maintained reporter fluorescence in the tissues, which implied the regulatory element of *col2a1a* expression in the craniofacial cartilage and ear is in the R2-D2-E1 sequence.

A



B

DE

GCGGCTCTCTTCTCCCCCACTGCCCGGTGCTCTCTCACATTC
CTCAGGTCTGCACACAGA**GCCGCATTGTGTGTGTGTCTTAC**
AGAGCACACAGTCAGGGCTCATTTCGGACACACACAC

Figure 9. Minimization of the R2-DE Sequence. (A) Schematic depicting where DE is located in relation to R2 regulatory region, as well as the relative location of the minimized R2-DE regions (D-E2; D2-E; D2-E1). (B) The sequence of the DE region. The gold sequence is the D region, while the orange sequence is the E region. The D1 and the E1 region are bolded.

Table 10. Results of injections with D-E1 driving expression.

	Date Injected	Fish Injected	Notochord		Craniofacial Cartilage		Ear	
D-E1			# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive
	3/27	51	5	10%	15	29%	6	12%
	3/28	80	4	5%	11	14%	4	5%
	Total	131	9	7%	26	20%	10	8%

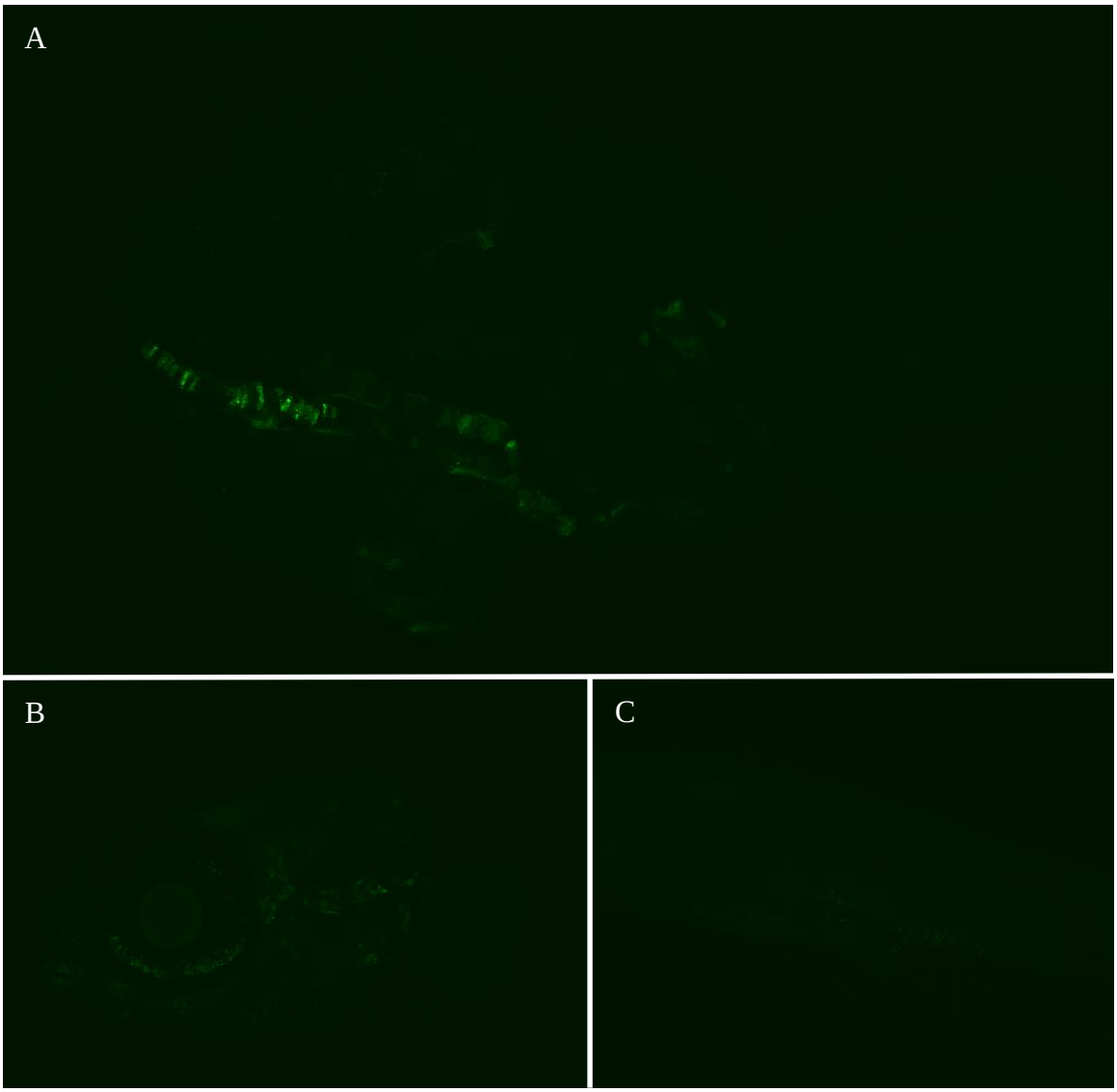


Figure 10. Results of R2: D-E1 plasmid injections. (A) Ventral view of zebrafish head depicting EGFP expression in the craniofacial cartilage. (B) Lateral view of the zebrafish head showing faint expression in the craniofacial cartilage and ear of the zebrafish. (C) Lateral view of zebrafish trunk with no EGFP expression (5 dpf).

Table 11. Results of injections with D2-E driving expression.

	Date Injected	Fish Injected	Notochord		Craniofacial Cartilage		Ear	
			# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive
D2-E	3/27	50	3	6%	19	38%	3	6%
	3/28	7	1	14%	2	29%	1	14%
	4/4	63	8	13%	40	63%	32	51%
	5/29	95	30	32%	73	77%	61	64%
	Total	215	42	20%	134	62%	97	45%

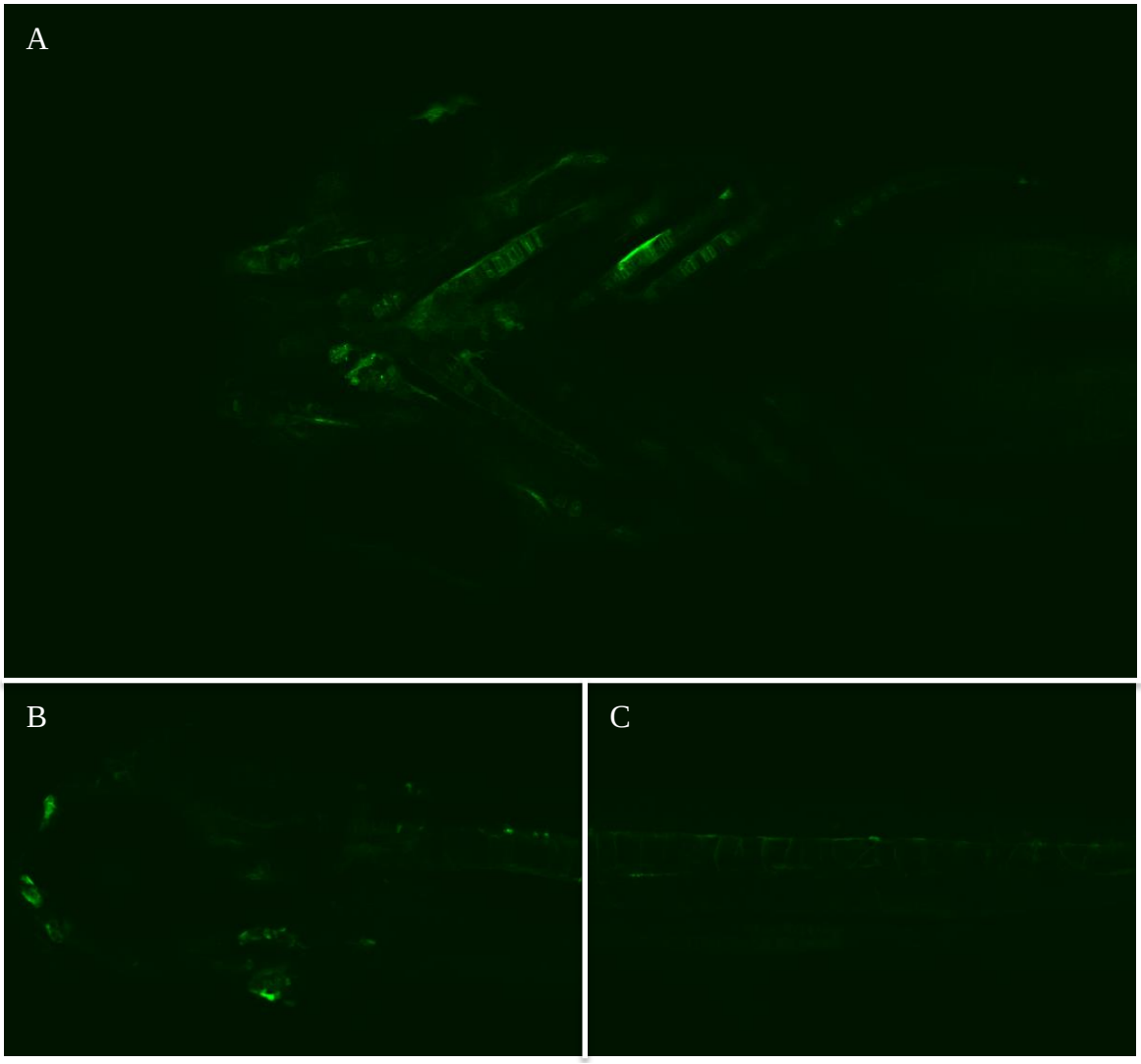


Figure 11. Results of R2: D2-E plasmid injections. (A) Ventral view of zebrafish head depicting EGFP expression in the craniofacial cartilage. (B) Lateral view of the zebrafish head showing faint expression in the craniofacial cartilage and ear of the zebrafish. (C) Lateral view of zebrafish trunk with EGFP expression (5 dpf).

Table 12. Results of injections with D2-E1 driving expression.

	Date Injected	Fish Injected	Notochord		Craniofacial Cartilage		Ear	
			# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive
D2-E1	4/4	140	9	6%	65	46%	57	41%
	5/29	196	5	3%	50	26%	27	14%
	6/12	198	2	1%	146	74%	70	35%
	6/12	161	3	2%	82	51%	22	14%
	Total	695	19	3%	343	49%	176	25%

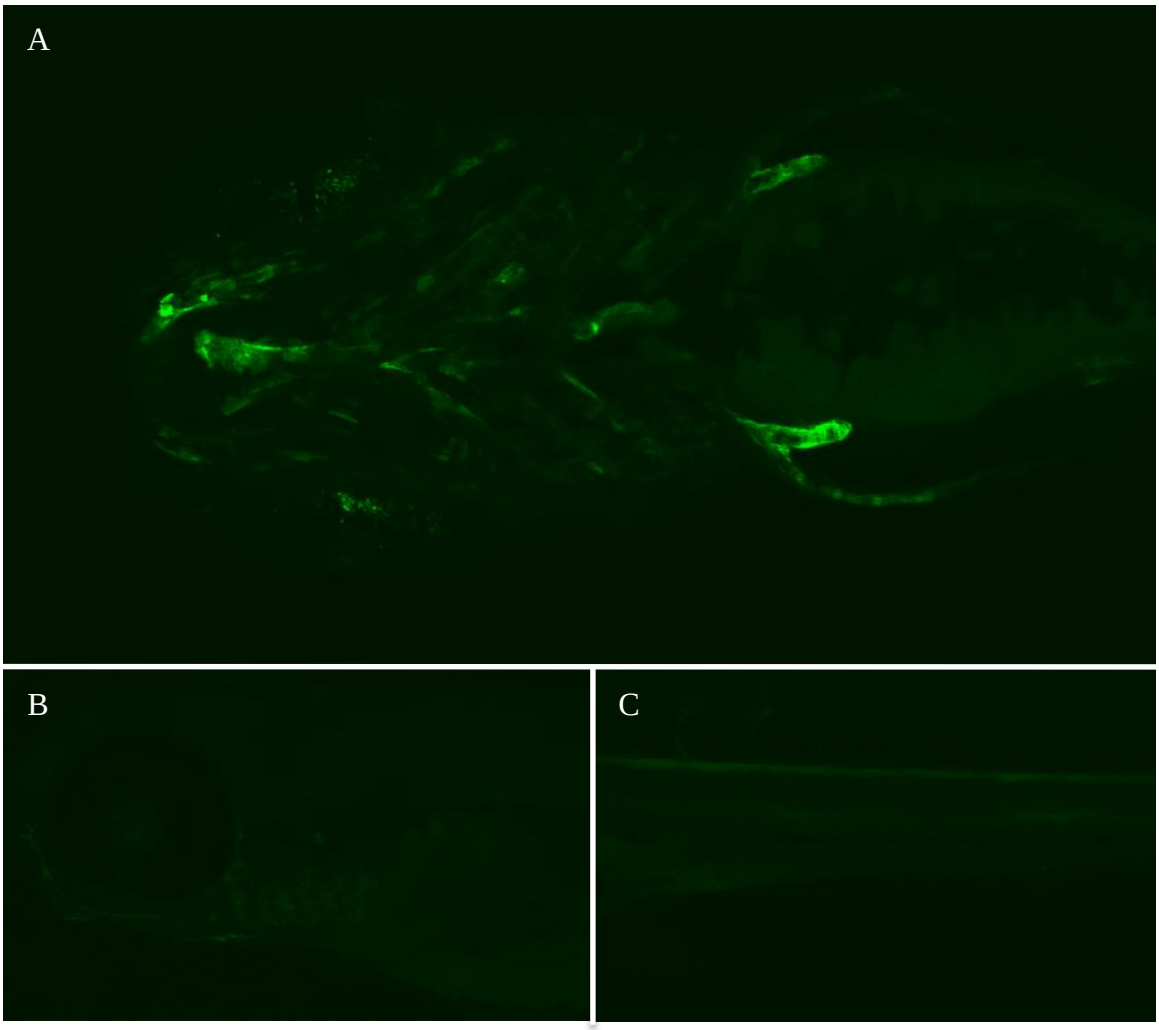


Figure 12. Results of R2: D2-E1 plasmid injections. (A) Ventral view of zebrafish head depicting EGFP expression in the craniofacial cartilage. (B) Lateral view of the zebrafish head showing faint expression in the craniofacial cartilage and ear of the zebrafish. (C) Lateral view of zebrafish trunk with no EGFP expression in the notochord (5 dpf).

Table 13. Results of injections with D region driving expression.

	Date Injected	Fish Injected	Notochord		Craniofacial Cartilage		Ear	
			# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive
D	3/20	5	0	0%	0	0%	0	0%
	3/22	39	0	0%	0	0%	0	0%
	Total	44	0	0%	0	0%	0	0%

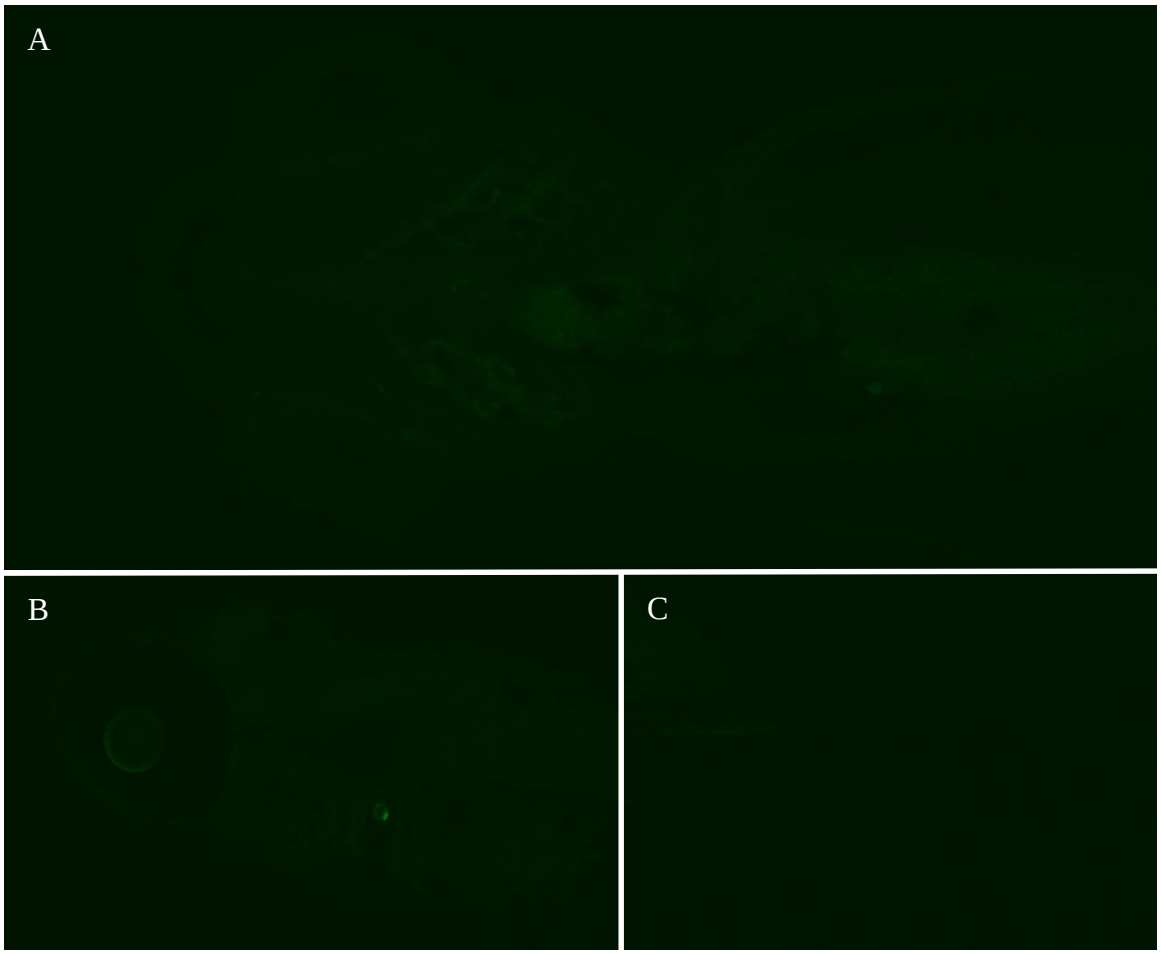


Figure 13. Results of R2: D plasmid injections. (A) Ventral view of zebrafish head depicting no EGFP expression in the craniofacial cartilage. (B) Lateral view of the zebrafish head showing no expression in the craniofacial cartilage and ear of the zebrafish. (C) Lateral view of zebrafish trunk with no EGFP expression in the notochord (5 dpf).

Table 14. Results of injections with E region driving expression.

E	Date Injected	Fish Injected	Notochord		Craniofacial Cartilage		Ear	
			# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive
	3/27	44	1	2%	2	5%	0	0%
	Total	44	1	2%	2	5%	0	0%

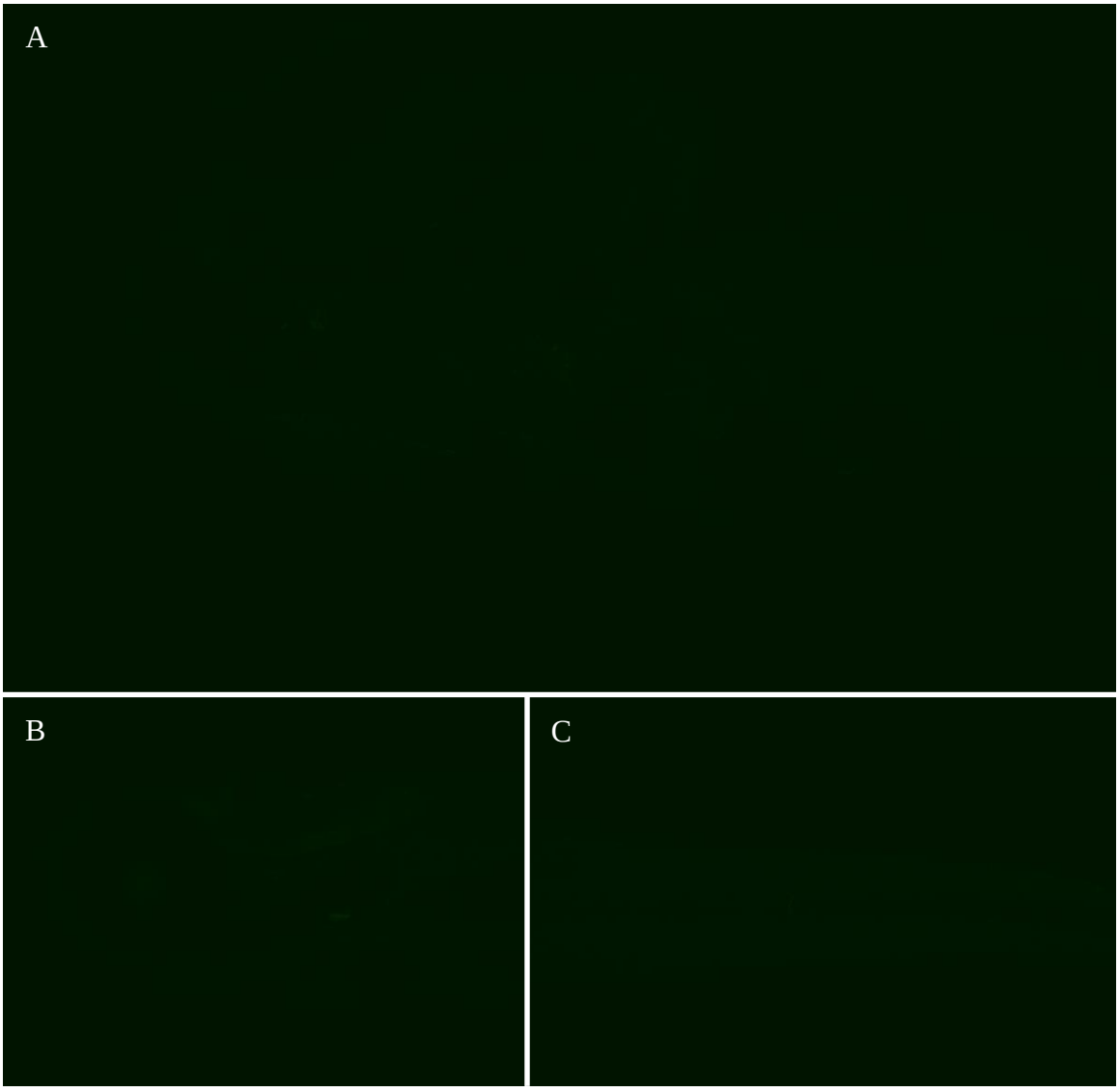


Figure 14. Results of R2: E plasmid injections. (A) Ventral view of zebrafish head depicting no EGFP expression in the craniofacial cartilage. (B) Lateral view of the zebrafish head showing no expression in the craniofacial cartilage and ear of the zebrafish. (C) Lateral view of zebrafish trunk with no EGFP expression in the notochord (5 dpf).

Identification of Transcription Factors Binding to Regulatory Region

It is well known that when one compares homologous regions of genomes from related animals that the protein coding exons can be well conserved at rates upwards of 75% due in part to mutations in exons leading to non-functional proteins critical for survival. Previous work has shown that there are regions of the genome not in exons that seem to be conserved over evolution at rates similar to exons, suggesting that these non-coding DNA sequences are also critical for survival (Sakabe *et al.*, 2012; ENCODE Project Consortium, 2012*).* Many of these conserved non-coding DNA sequences have been found to have required transcription factor binding sites that regulate gene expression. We originally identified the R2 regulatory region by comparison of 12kb of genomic sequence surrounding the transcriptional start site of *col2a1a* gene of various teleost fish. Of the highly conserved sequences, the R2 region was able to recapitulate most of the expression pattern of *col2a1a* in the cartilage, ear, and notochord. The reduction of the R2 regulatory element determined that R2-DE was the most important region for the regulation of this gene. A comparative genomics approach of teleosts was again utilized but this time concentrated on short nucleotide sequences, which could determine exactly which transcription factors are binding to these regions. Since the 120 bp DE region was sufficient to recapitulate the expression seen with the entire 360 bp R2 regulatory region, we focused on this region to be compared among four teleost fish: medaka, stickleback, fugu, and zebrafish. Within the 120 bp sequence there were six regions of highly conserved nucleotide sequences, with many of them located within the D2-E1 region. Because of our focus on cartilage development we concentrated our

analysis on the D2-E1 region. The conservation of these sequences suggested they were the potential transcription factor binding sites regulating the expression of the *col2a1a* gene. The preference of transcription factors to bind to specific sequences allowed for *in silico* identification of potential proteins that would bind to the known conserved sequence.

Using basic bioinformatic tools for identifying transcription factor binding sites, I analyzed the R2-DE sequence. This resulted with a substantial amount of potential transcription factor binding sites in this region. In order to filter through the results, I focused on the three conserved regions completely in the D2-E1 sequence and analyzed them using the JASPAR database to identify high scoring transcription factor binding profiles that also where shown to play a role in cartilage formation. I used JASAR to compare the conserved DNA sequence against known transcription factor binding matrices. This allowed JASPAR to compute a relative score for putative candidates. The closer the score is to 1, the higher the conservation/possible binding of the transcription factors. The first highly conserved sequence, CATTCCTCAGGTC, is linked to the Ets family of transcription factors with a relative score of 0.9160. The second highly conserved sequence, TGCACACAGAG, was associated with the Runx family of transcription factors with a relative score of 0.8137. The third highly conserved sequence, ATTGTGTGTG, was associated with the Sox family of transcription factor with a relative score of 0.8430. While JASPAR does suggest what exact transcription factor has the best match to the analyzed sequence, we and others have found that these programs tend to better identify families because of transcription factors being slightly flexible

(transcription factor wobble) in what sequence they bind (Spitz and Furlong, 2012; Ko and Engel, 1993). Based on the results of the database, Ets, Runx, and Sox transcription factor families were potentially key regulators of the R2-D2E1 region. As previously discussed, Runx2 and Sox9 transcription factors play an important role in chondrogenesis in many animals (Flores *et al.*, 2006; Yan *et al.*, 2005). The identification of the specific Ets transcription factor is more difficult to decipher because of the size of the family and multiple members are expressed at the right time and place, but Fli1a was our primary candidate based on its expression pattern in the pharyngeal arches at the appropriate developmental stage.

Mutagenesis of Putative R2-DE Transcription Factor Binding Sites

I hypothesized that the R2-DE region contained transcription factor binding sites for Fli1a, Sox9a, and Runx2b, and these transcription factors played important roles in the regulation of *col2a1a*. This candidate gene approach was based off the current literature and known expression patterns. The sequence logos generated by JASPAR allowed us to identify nucleotides of the conserved genome sequences that must be conserved for each transcription factor binding site, as well as locations of flexibility. We next set out to mutate the key bases in our R2-DE EGFP reporter plasmids to determine if our identified conserved sequences are necessary for R2-DE gene regulatory activity. By using these mutated plasmids to generate transient transgenic reporter zebrafish a lack of EGFP in the proper tissue would demonstrate the importance of the original sequence in the binding of transcription factors.

I first mutated the putative Ets-family binding site, from the core TTCC to GGAA (Figure 16). The TTCC nucleotide sequence seemed to be particularly important for Ets family member binding based on its sequence logo. Of the 280 fish injected, 75% of the embryos had EGFP expression in the craniofacial cartilage and 59% showed expression in the ear (Table 15). The notochord also had EGFP expression in about 33% present of the embryos. The four nucleotides mutated did not seem to negatively affect the expression level in the notochord or the cartilage (Figure 20), but potentially enhanced expression. The mutagenized Runx-family binding site altered five nucleotides changing the sequence from ACACAGA to GTCTAGC. Once again the expression levels of the notochord and craniofacial cartilage were not affected in the 240 embryos injected (Figure 17). Forty percent of the zebrafish expressed the reporter protein in the notochord, and 33% expressed the protein in the craniofacial cartilage (Table 16) comparable to normal R2-DE activity. The expression in the tissue of the ear was present in about 16% of the zebrafish, which was slightly lower than the tissues of notochord and craniofacial cartilage (Figure 21). The Sox-family core binding site was mutated from TGTGT to CTCGG to prevent the family of transcription factors from binding (Figure 18). Similar to the other mutagenized plasmids described above, the notochord, craniofacial cartilage, and ear expression levels of the Sox-family mutated binding site remained steady (Figure 22). One hundred and seventy-eight embryos were injected with this mutagenized plasmid and there was EGFP expression in the notochord of 45% of the zebrafish, 51% in the craniofacial cartilage, and 25% of the ear (Table 17).

The mutagenesis of the putative transcription factor binding sites individually did not affect the expression level of EGFP in the cartilage, ear, and notochord tissues. This was not surprising with the close proximity of these three conserved putative transcription factor binding sites, and it is highly possible that protein-protein interactions could still recruit the appropriate transcription factors, even with their inability to bind the DNA anymore. Research by other groups has demonstrated that there are protein-protein interactions between the Ets, Runx, and Sox family of transcription factors (Deramaudt *et al.*, 2001; Huang *et al.*, 2009). To destabilize this putative three part regulation complex, we next mutated different combinations of the three conserved binding sites. It was not until two of the putative binding sites were mutated together, that a decrease in the expression level of EGFP was seen in the tissues of interest. When the potential Sox and Ets binding sites were mutated in conjunction with each other, the EGFP expression levels were decreased in cartilage, notochord, and ear (Figure 23). Of the 311 embryos injected, only 13% of the zebrafish expressed EGFP in the craniofacial cartilage and 5% in the ear (Table 18). The zebrafish injected with the plasmid expressed the reporter protein in the notochord for only 22 fish, or 7% of the zebrafish. Similar to these results, the mutagenesis of the Ets and Runx binding sites also decreased the expression levels of the reporter protein (Figure 24). The EGFP expression level in the notochord was only in 9% of zebrafish, 18% in the craniofacial cartilage, and 7% in the ear (Table 19). By mutating these sequences of the potential binding sites we were able to disrupt the effective binding of the transcription factors to the DNA, and in turn minimized the likelihood of the R2-DE regulatory region to initiate transcription of the reporter gene.

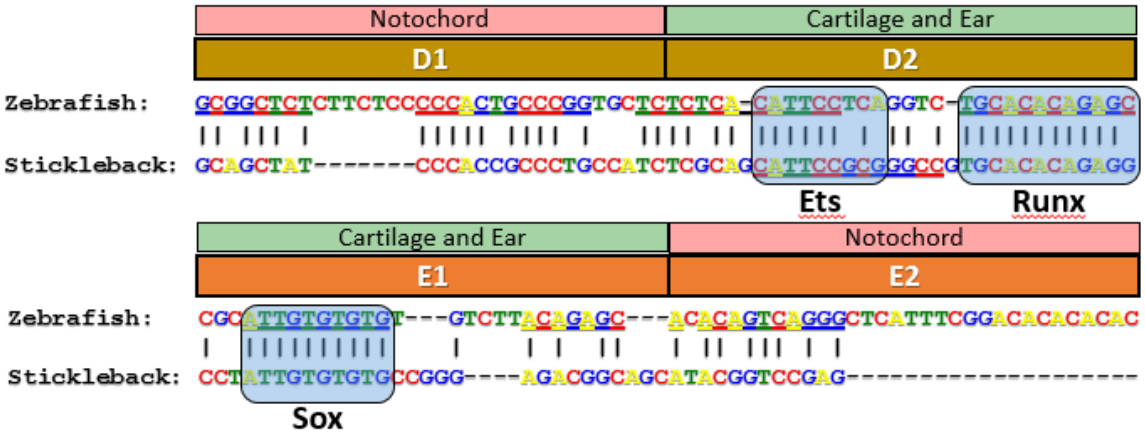
When the Runx and Sox binding sites were mutated in the plasmid, EGFP was present in the notochord of 29% of fish, 31% in the craniofacial cartilage, and 11% in the ear (Figure 25; Table 20). The percentage of zebrafish that expressed EGFP in the appropriate tissues was not significantly decreased by the mutagenesis of the plasmids at these two sites.

Increase Activity in Mutated Ets-family Binding Site

The percentage of zebrafish with expression in the craniofacial cartilage was much higher for zebrafish injected with the plasmids mutated at the Ets binding site. Initially *in silico* analysis and literature searches suggested the possibility of Ets1 binding to the DE region, which was how the mutations were designed. Further analysis revealed the possibility of Fli1a, an Ets family member, as a potential candidate also due to its expressed in the correct tissue at the right time during development. A review of the Ets mutagenesis revealed the transgenic reporter plasmids possibly contained a better binding sequence for Fli1a. This finding was a potential explanation for the large percentage of zebrafish that expressed EGFP in the craniofacial cartilage when injected with this plasmid. The plasmid with the Ets mutated binding site seemed to up-regulate the transcription of the reporter protein because it was better able to bind to the regulatory region. In order to account for this, the mutation was repeated at the same site, but altered to different nucleotides. The CATTCC sequence was mutated to TGCATG, which altered the nucleotides to theoretically no longer allow the binding of Ets1 or Fli1a. Plasmids were created in combination as well as alone. The Runx and Sox binding site mutations remained the same from the previous mutagenesis. The plasmid with the Ets/Fli1a

binding site mutated was injected into 410 zebrafish, and 66% expressed the reporter protein in the craniofacial cartilage, 47% in the ear, and 36% percent in the notochord (Figure 26; Table 21). Once again, the binding site mutation did not significantly decrease the percentage of embryos that expressed EGFP. Similar to the previous mutagenesis, when Fli1a and Runx binding sites were mutated simultaneously, it decreased the percentage of zebrafish with reporter plasmid expression (Figure 27). There were 677 zebrafish injected with this mutated plasmid; only 1% expressed the reporter protein in the notochord, 3% in the craniofacial cartilage, and 5% in the ear (Table 22). This data shows that the two mutations to the Ets family binding site, and the Ets binding site in conjunction with the Runx site had very similar results. A change came with a mutation to the Ets and Sox binding sites. The mutation to the Ets/Fli1a binding site, together with the mutated Sox binding site, was injected into 168 zebrafish (Table 23). Thirty nine percent of the zebrafish expressed the reporter protein in the craniofacial cartilage, 15% in the notochord, and 18% in the ear (Figure 28). The initial mutation decreased the amount of zebrafish expressing the reporter plasmid, but the new mutation to the Ets/Fli1a and the Sox binding site did not have the same results.

A



B

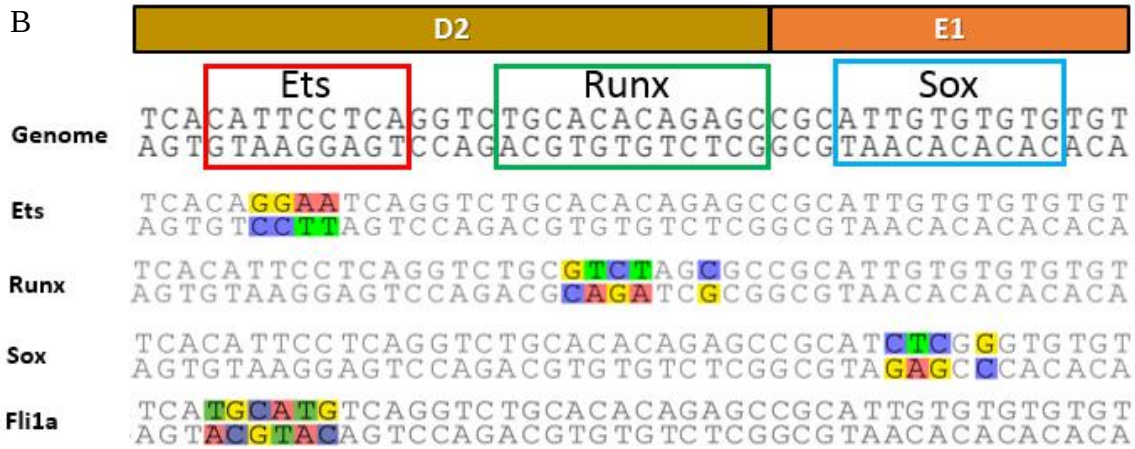


Figure 15. Sequence Conservation Schematic. Images depicting sequence conservation and putative transcription factor binding sites. (A) The image is a comparison of zebrafish and stickleback, but analysis also includes comparisons with tetradon and medaka. There seems to be Ets, Runx, and Sox Transcription factor binding sites. (B) Image of the binding sites based on the zebrafish genome, as well as the mutations that were made to the sequence.

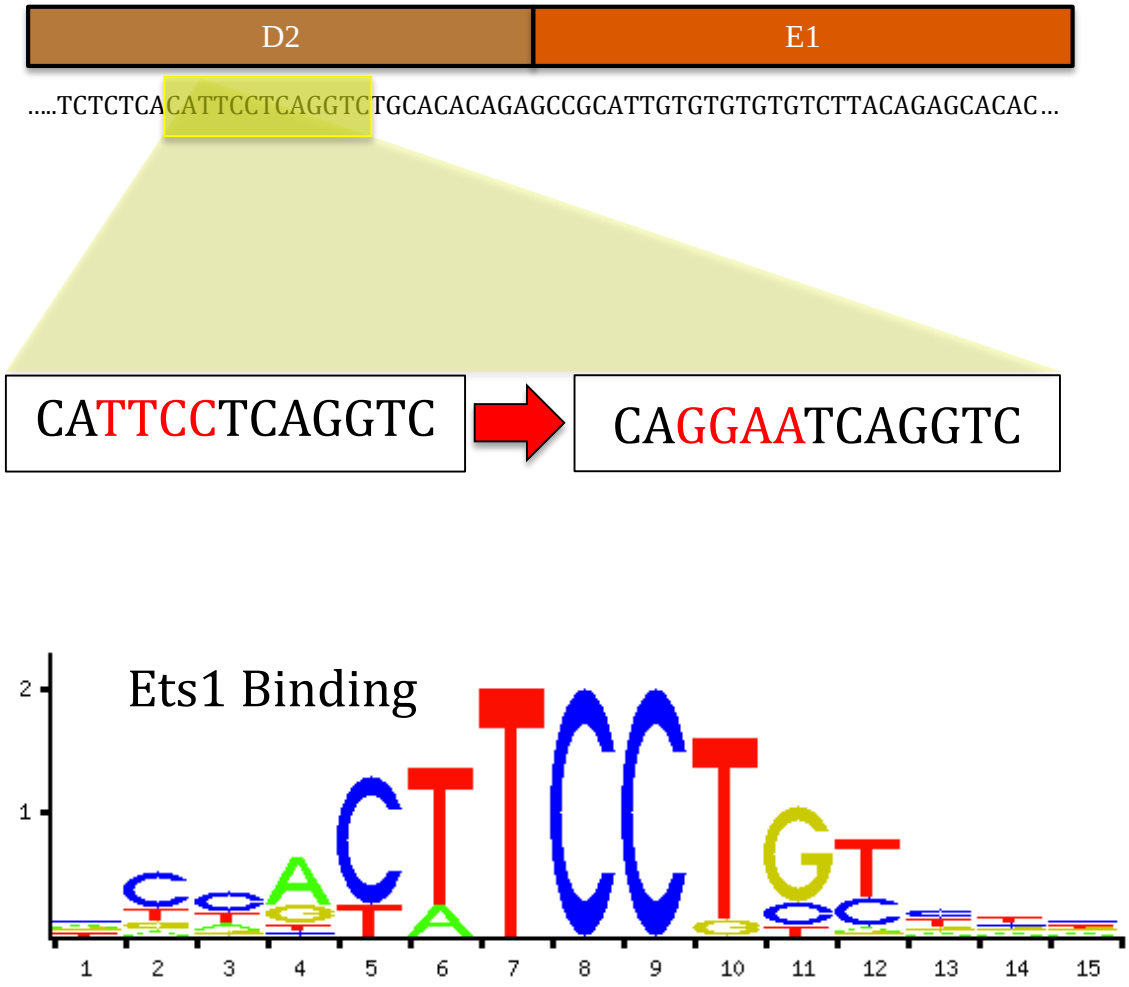


Figure 16. Ets-Binding Site Mutation Schematic. The potential Ets binding site is highlighted to show its location in the D2 region of DE. Based on the sequence logo, mutations were made to alter the binding sequence to prevent the protein from binding.

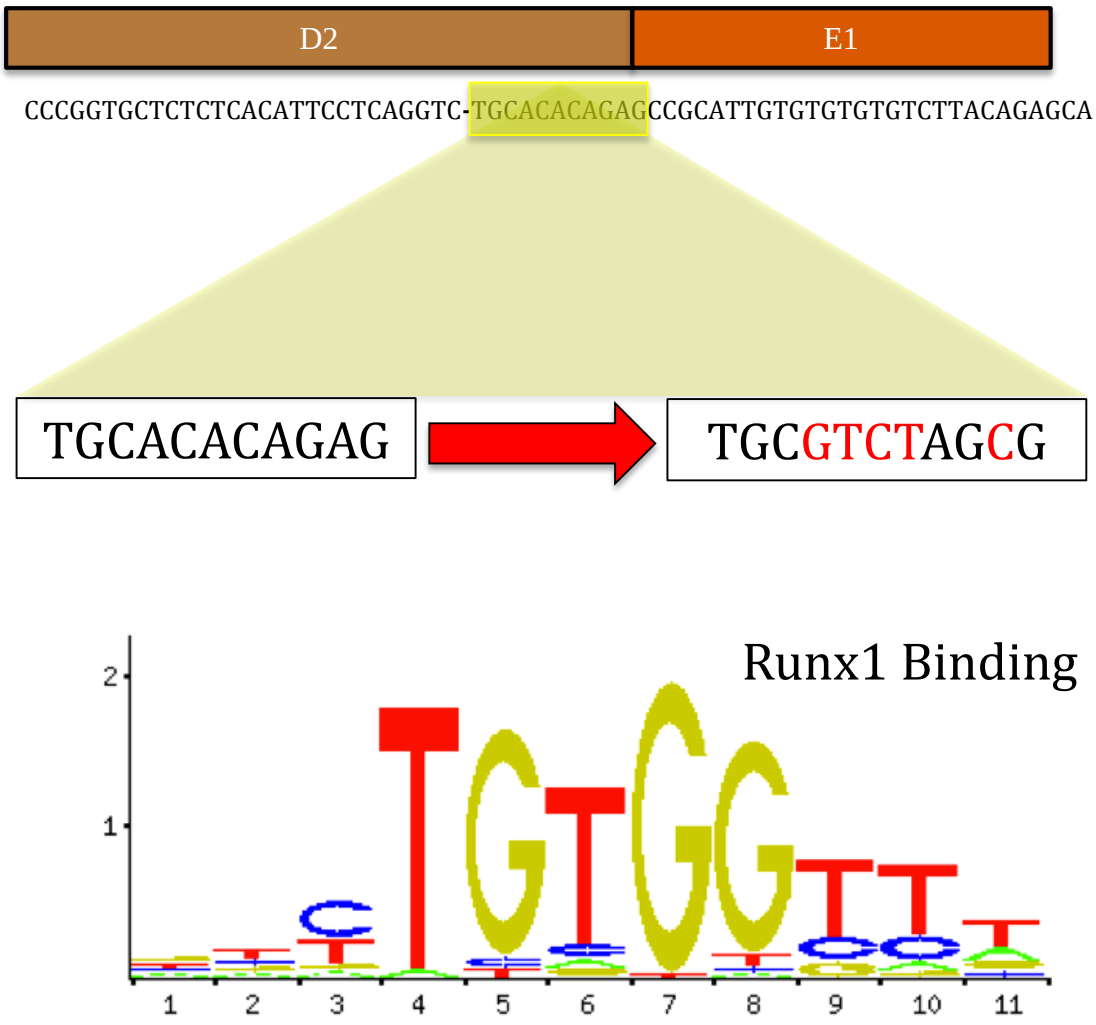
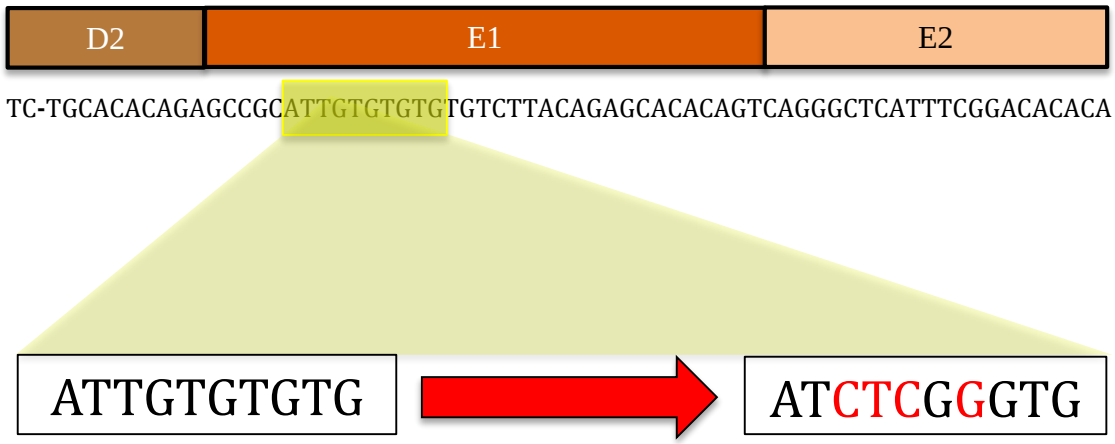


Figure 17. Runx-Binding Site Mutation Schematic. The region of DE highlighted is the potential Runx binding site. The sequence logo for Runx provided the nucleotides that are mutated to prevent binding of the transcription factor.



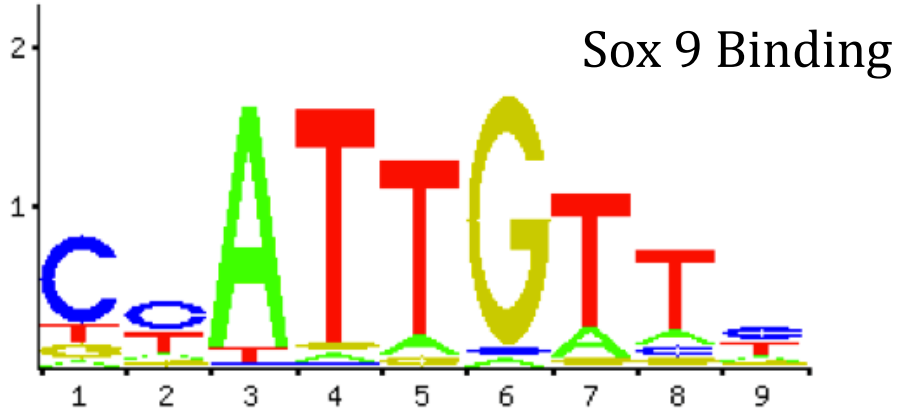
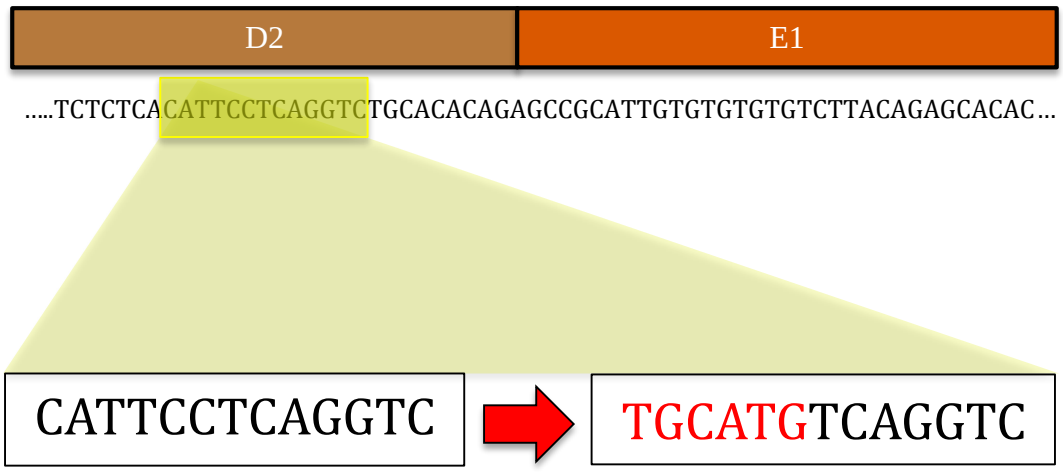


Figure 18. Sox-Binding Site Mutation Schematic. The potential Sox binding site is located in the E1 region of DE as depicted above. The sequence was mutated in accordance to the sequence logo for Sox 9.



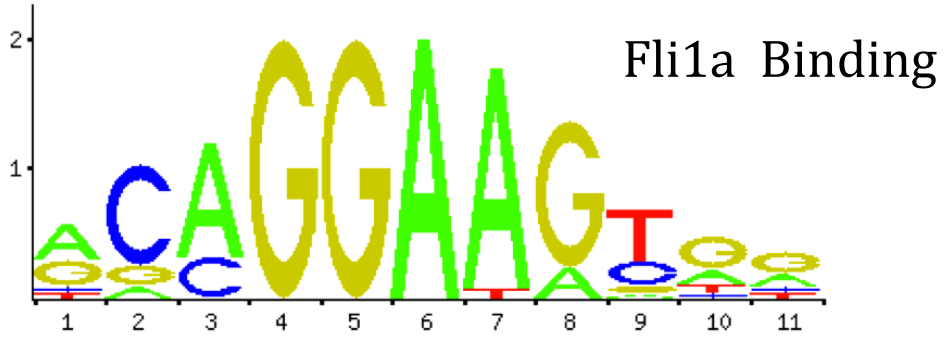


Figure 19. Fli-Binding Site Mutation Schematic. After literature review and analysis of the data, the original Ets binding site mutation may have provided a better Fli1a binding site. More plasmids were created to mutate this potential Fli1a binding site according to the sequence logo above.

Table 15. Results of injections with DE plasmid containing Ets-binding site mutated.

Mutation Ets	Date Injected	Fish Injected	Notochord		Craniofacial Cartilage		Ear	
			# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive
	10/24	146	41	28%	103	71%	87	60%
	12/4	51	17	33%	43	84%	35	69%
	12/12	83	33	40%	64	77%	43	52%
	Total	280	91	33%	210	75%	165	59%

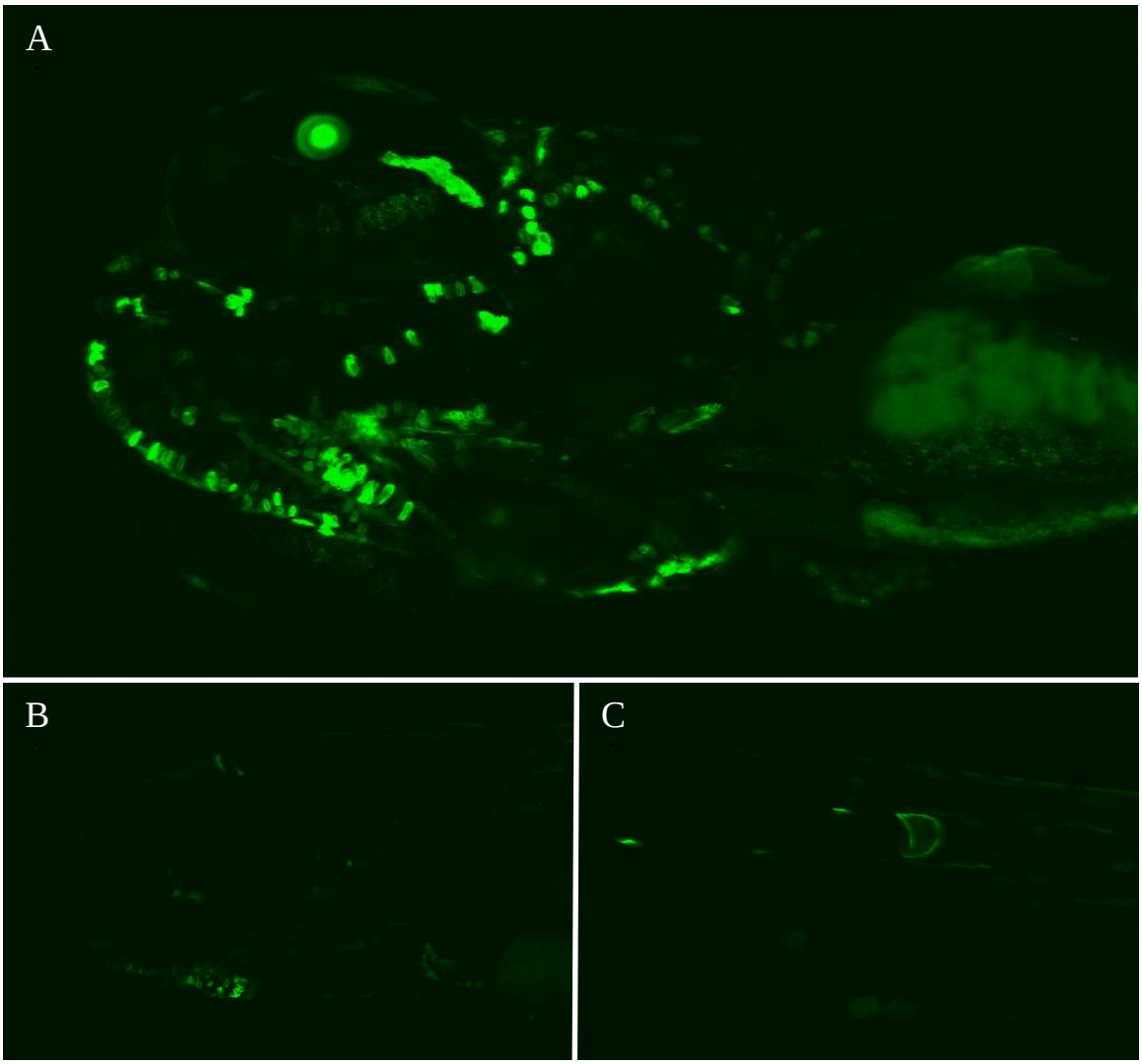


Figure 20. Results of Injections with Ets Binding Site Mutated in Plasmid. Zebrafish at 5 dpf depicting the EGFP expression pattern of Ets binding site mutated in the DE plasmid. (A) Ventral view of craniofacial cartilage. Reflection of fluorescent protein may be visible in the eye. (B) Lateral view of zebrafish head. EGFP present in craniofacial cartilage. (C) Lateral view of trunk depicting the notochord (5dpf).

Table 16. Results of injections with the Runx-binding site mutated in the DE plasmid.

Mutation Runx	Date Injected	Fish Injected	Notochord		Craniofacial Cartilage		Ear	
			# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive
	10/8	28	7	25%	18	64%	4	14%
	11/6	23	9	39%	4	17%	3	13%
	11/13	82	21	26%	13	16%	7	9%
	12/12	107	60	56%	45	42%	24	22%
	Total	240	97	40%	80	33%	38	16%

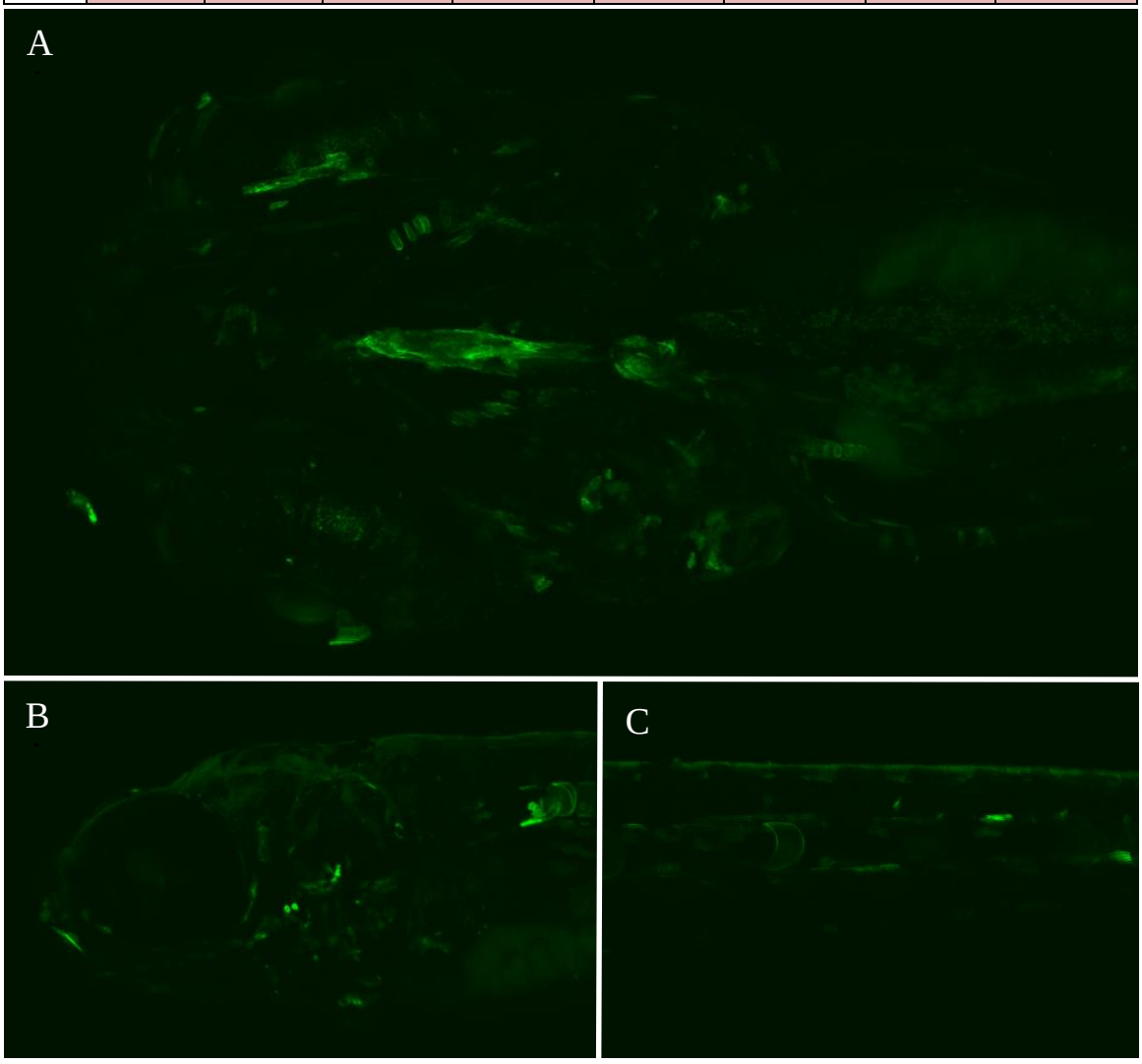


Figure 21. Results of Injections with Runx Binding Site Mutated in Plasmid. Zebrafish at 5 dpf depicting the EGFP expression pattern of Runx binding site mutated in the DE plasmid. (A) Ventral view of zebrafish showing EGFP expression in the craniofacial cartilage. (B) EGFP present in craniofacial cartilage and ear in lateral view of head. (C) Lateral view of the notochord in the trunk region.

Table 17. Results of injections of the DE plasmid with the Sox binding site mutated.

Mutation Sox	Date Injected	Fish Injected	Notochord		Craniofacial Cartilage		Ear	
			# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive
	10/24	72	37	51%	46	64%	26	36%
	11/20	44	18	41%	23	52%	9	20%
	12/3	62	25	40%	22	35%	9	15%
	Total	178	80	45%	91	51%	44	25%

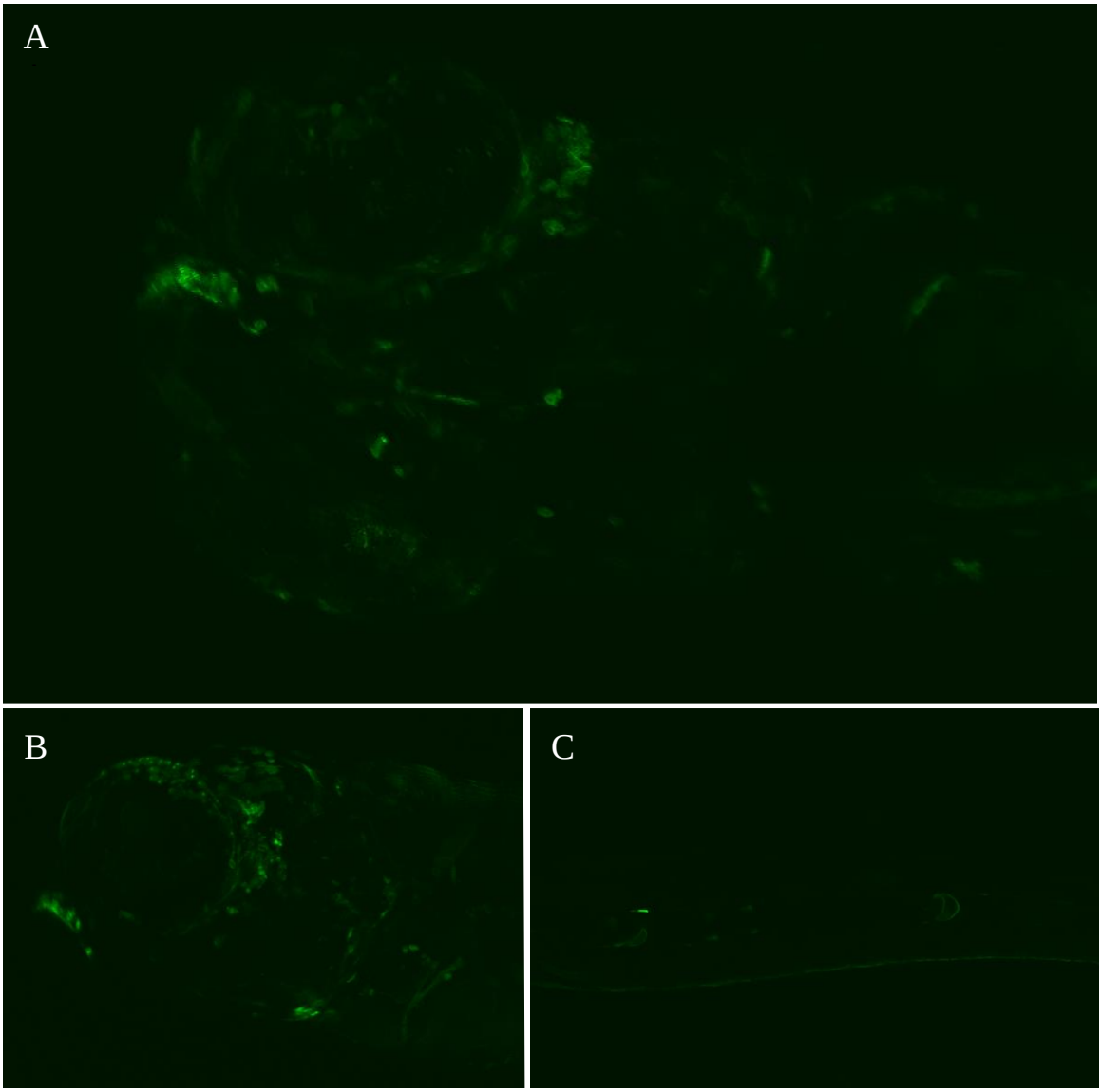


Figure 22. Results of Injections with Sox Binding Site Mutated in Plasmid. Images of zebrafish at 5 dpf following injections of the DE plasmid with the Sox binding site mutated. (A) Ventral view of craniofacial cartilage. (B) Lateral view of zebrafish head. (C) Lateral view of zebrafish head, showing craniofacial cartilage and ear. (D) Lateral view of trunk to show notochord.

Table 18. Results of injections with Ets and Sox binding sites mutated in DE plasmid.

Mutation EtsSox	Date Injected	Fish Injected	Notochord		Craniofacial Cartilage		Ear	
			# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive
	10/17	98	0	0%	3	3%	2	2%
	11/7	88	15	17%	12	14%	4	5%
	11/21	125	7	6%	25	20%	10	8%
	Total	311	22	7%	40	13%	16	5%

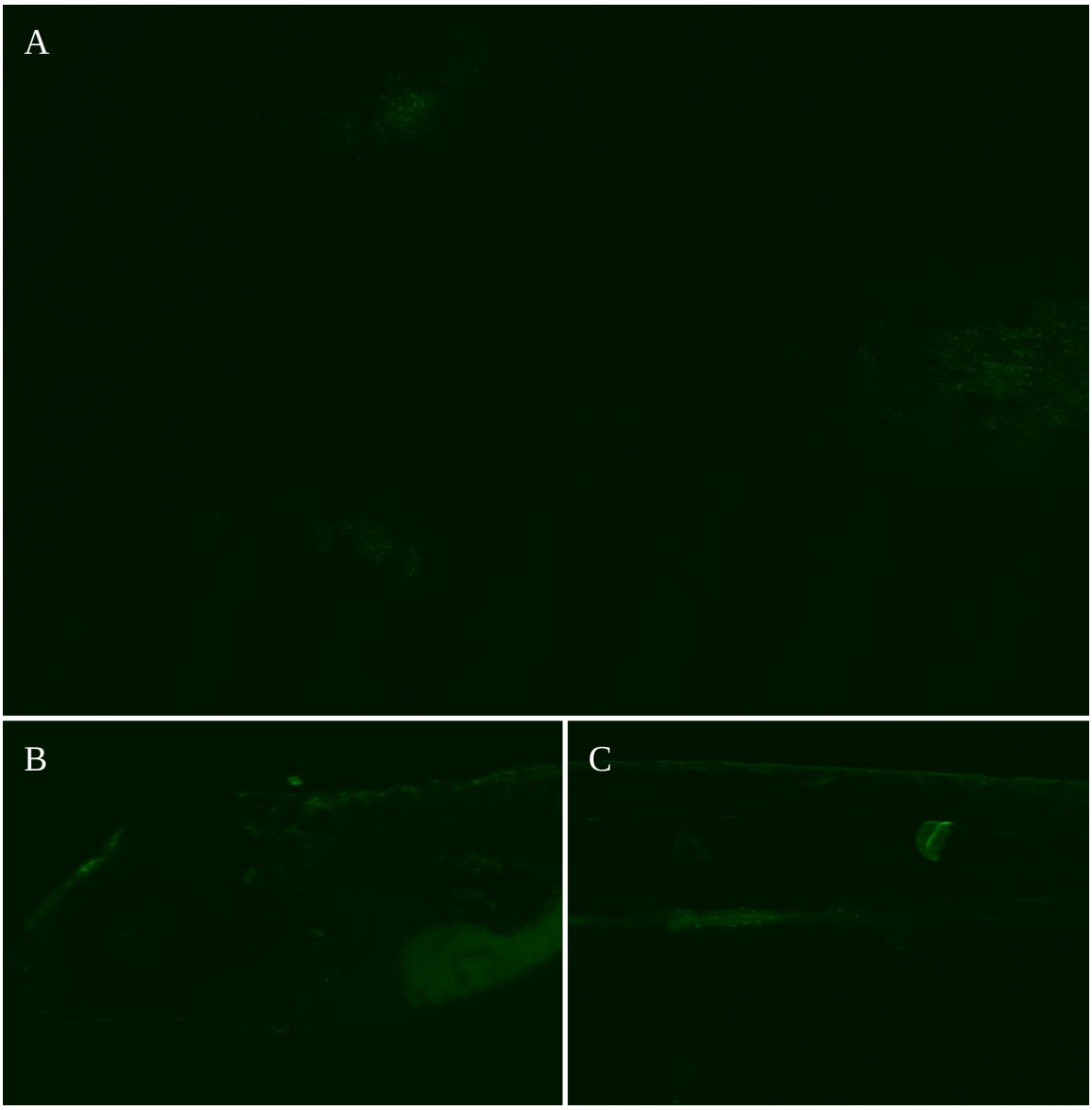


Figure 23. Results of Injections with Ets & Sox Binding Sites Mutated in Plasmid. Low or no EGFP expression in zebrafish injected with mutated binding sites of Ets and Sox in DE plasmid. (A) Ventral view of craniofacial cartilage at 5 dpf. (B, C) Lateral view of head and trunk of zebrafish at 5 dpf.

Table 19. Results of Ets and Runx binding sites mutated in DE plasmid injections.

Mutation EtsRunx	Date Injected	Fish Injected	Notochord		Craniofacial Cartilage		Ear	
			# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive
	10/17	36	3	8%	16	44%	6	17%
	10/31	211	17	8%	33	16%	12	6%
	12/10	10	2	20%	2	20%	1	10%
	12/11	107	12	11%	16	15%	6	6%
	Total	364	34	9%	67	18%	25	7%

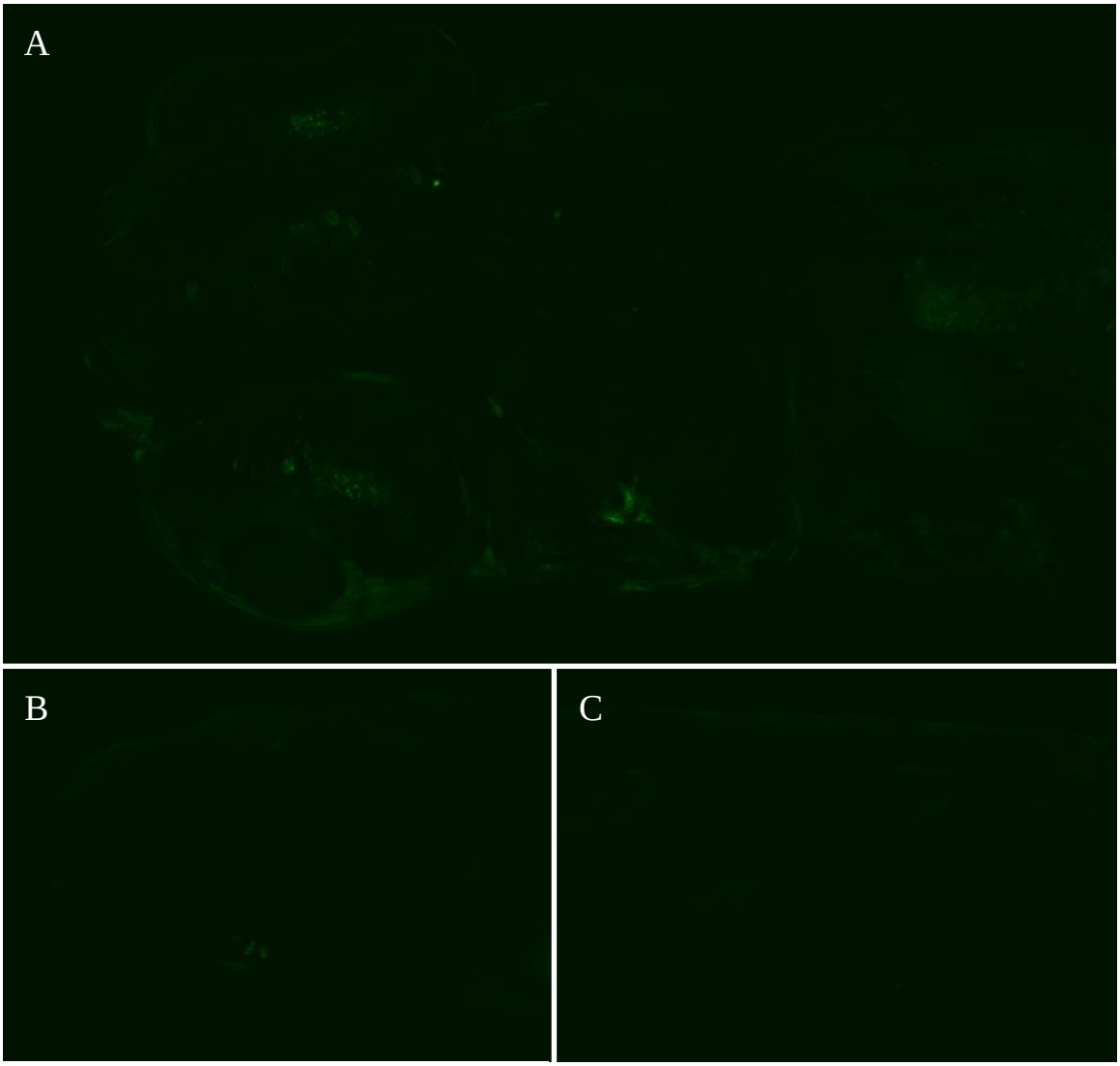


Figure 24. Results of Injections with Ets & Runx Binding Sites Mutated in Plasmid. Zebrafish at 5 dpf, following injections of DE plasmid with Ets and Runx binding sites mutated. (A) Ventral view of craniofacial cartilage. (B) Lateral view of zebrafish head. (C) Lateral view of trunk to show notochord.

Table 20. Results of Runx and Sox binding sites mutated in the DE plasmid.

Mutation RunxSox	Date Injected	Fish Injected	Notochord		Craniofacial Cartilage		Ear	
			# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive
	10/16	9	5	56%	3	33%	1	11%
	10/30	93	30	32%	24	26%	10	11%
	11/21	113	25	22%	35	31%	13	12%
	12/11	78	25	32%	28	36%	8	10%
	Total	293	85	29%	90	31%	32	11%

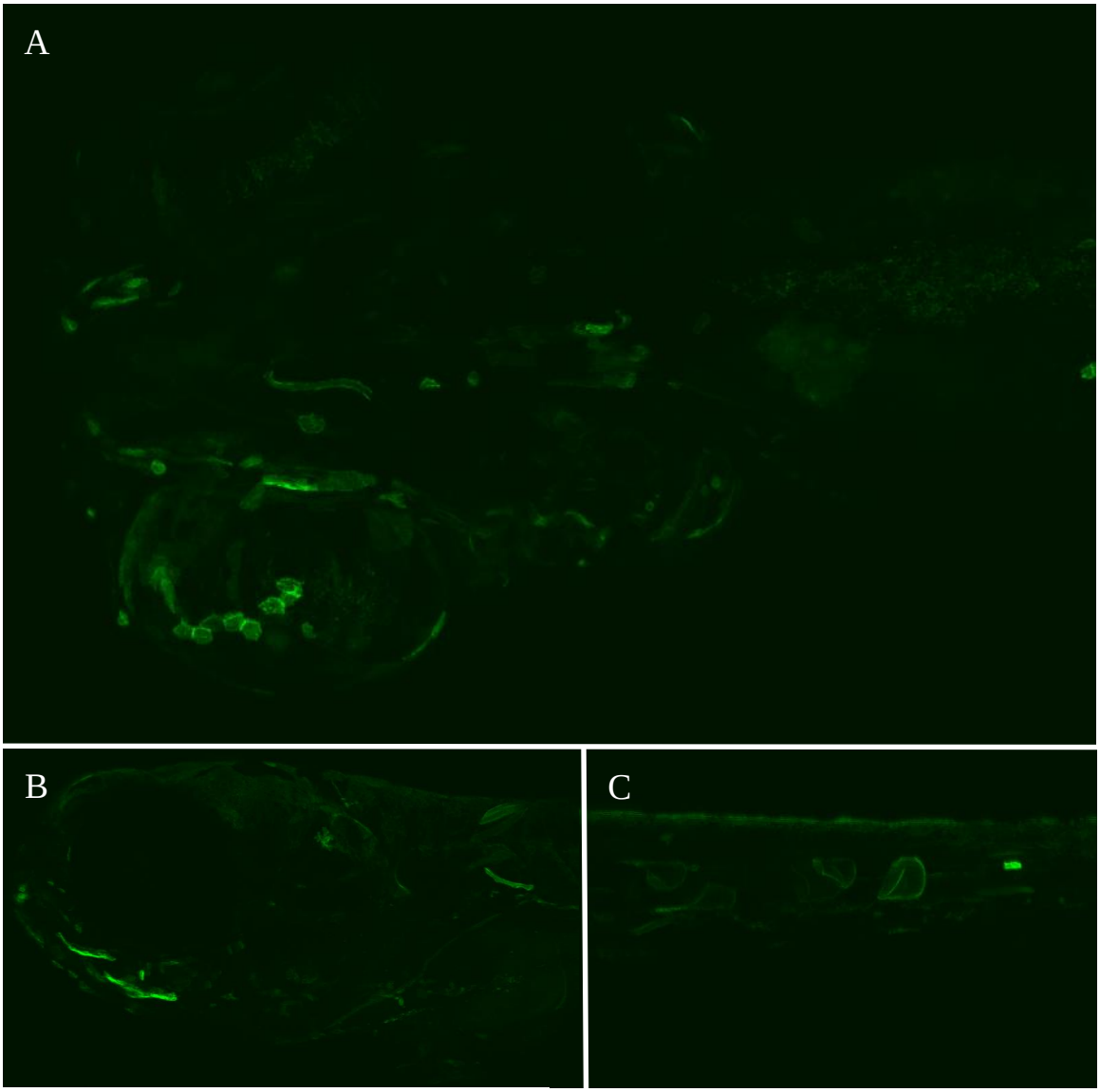


Figure 25. Results of Injections with Runx & Sox Binding Sites Mutated in Plasmid. Images of zebrafish at 5 dpf depicting EGFP expression pattern. (A) Ventral view of craniofacial cartilage. (B,C) Lateral view of zebrafish head and trunk, respectively.

Table 21. Results of injections with Fli binding site mutated in the DE plasmid.

Mutation Fli	Date Injected	Fish Injected	Notochord		Craniofacial Cartilage		Ear	
			# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive
	4/16	57	28	49%	41	72%	35	61%
	4/17	192	52	27%	118	61%	78	41%
	4/24	78	25	32%	52	67%	34	44%
	4/30	83	42	51%	60	72%	45	54%
	Total	410	147	36%	271	66%	192	47%

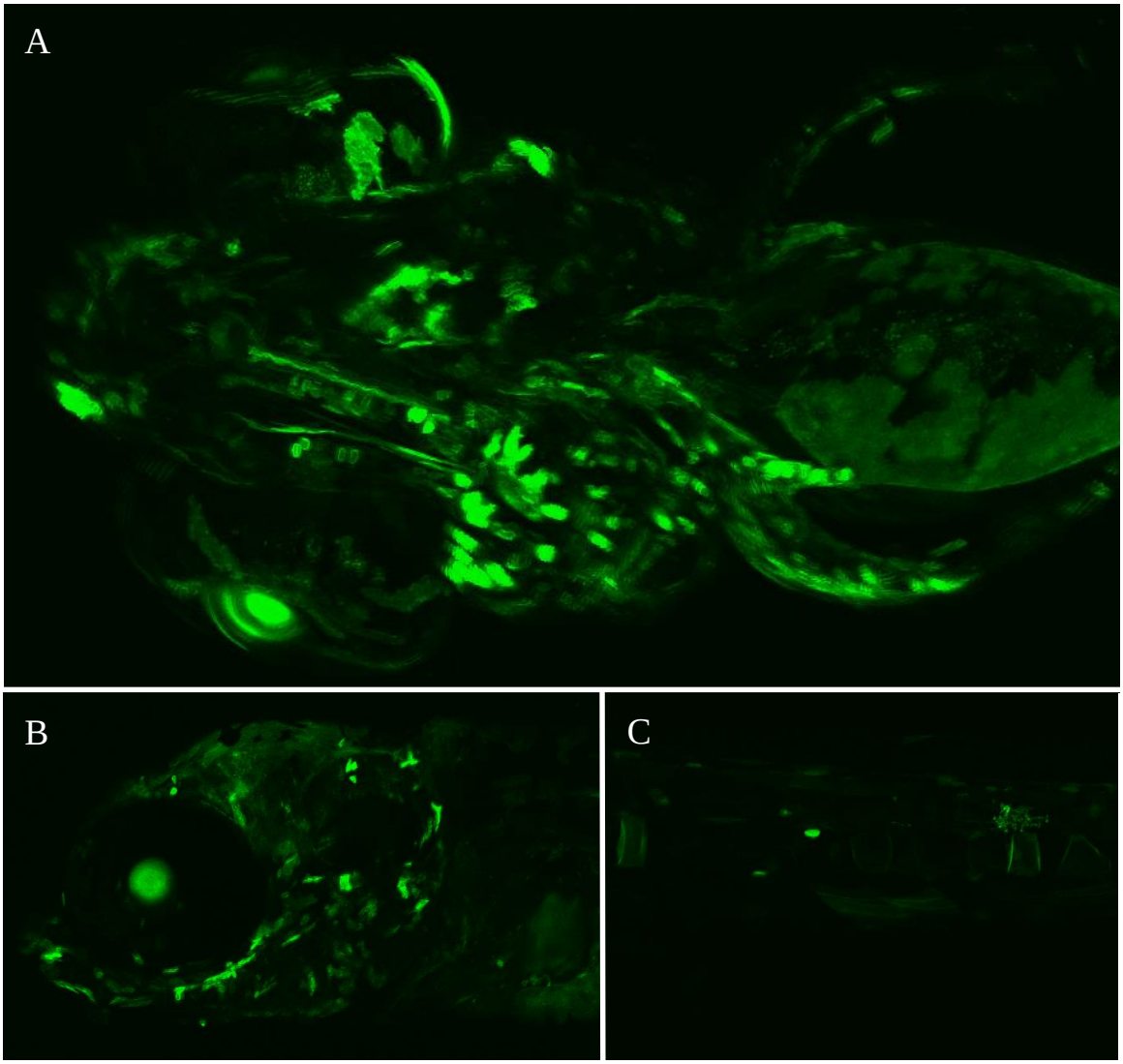


Figure 26. Results of Injections with Fli Binding Site Mutated in Plasmid. EGFP expression pattern in zebrafish with Fli binding site mutated. (A) Ventral view displaying the EGFP expression in the chondrocytes of the craniofacial cartilage. (B) Lateral view of head to display expression in the craniofacial cartilage and ear. (C) Lateral view of trunk to visualize the notochord.

Table 22. Results of injections with mutations in Fli and Runx binding site of DE.

Mutation FliRunx	Date Injected	Fish Injected	Notochord		Craniofacial Cartilage		Ear	
			# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive
	4/10	331	4	1%	16	5%	14	4%
	4/23	186	0	0%	2	1%	15	8%
	4/24	116	1	1%	2	2%	3	3%
	4/30	44	1	2%	2	5%	1	2%
	Total	677	6	1%	22	3%	33	5%

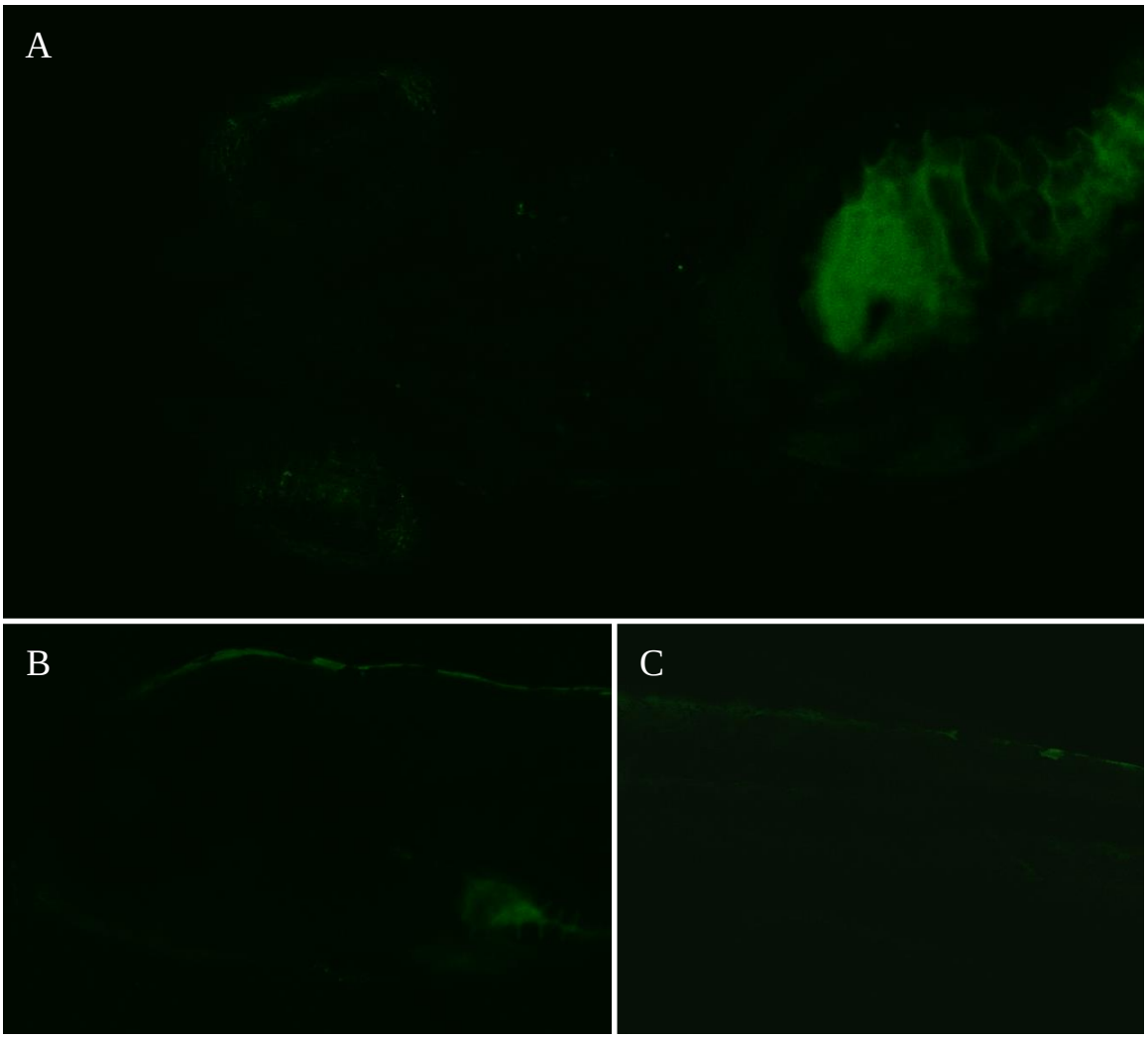


Figure 27. Results of Injections with Fli & Runx Binding Sites Mutated in Plasmid. No expression of EGFP in the zebrafish at 5 dpf. (A) Ventral view of craniofacial cartilage. (B) Lateral view of zebrafish head to show side view of craniofacial cartilage and ear. (C) Lateral view of trunk to show lack of EGFP expression in the notochord.

Table 23. Results of DE plasmid with mutations in the Fli and Sox binding sites.

Mutation FliSox	Date Injected	Fish Injected	Notochord		Craniofacial Cartilage		Ear	
			# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive	# of GFP Positive	% of GFP Positive
	5/2	120	24	20%	65	54%	30	25%
	5/22	48	2	4%	0	0%	0	0%
	Total	168	26	15%	65	39%	30	18%

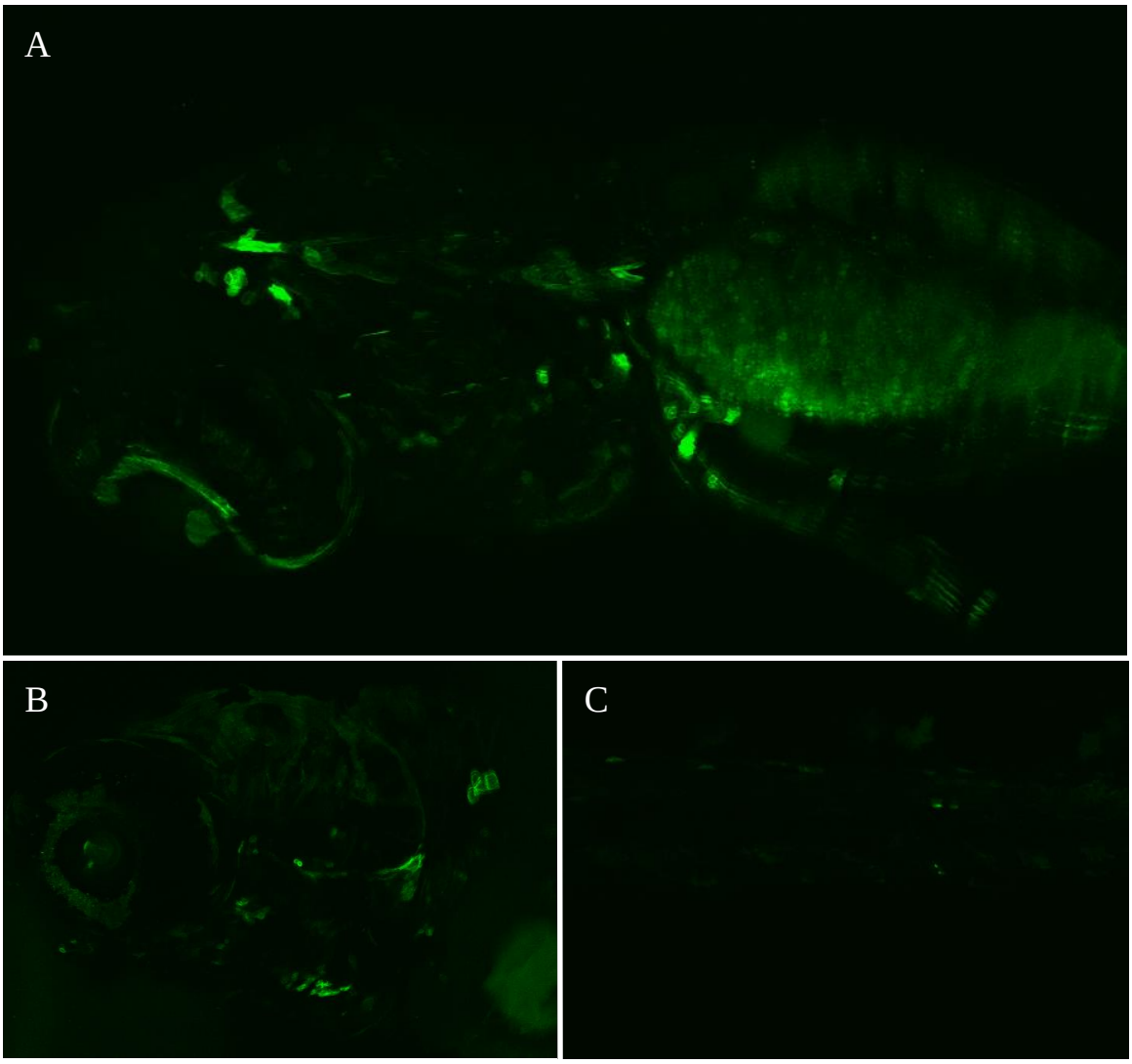


Figure 28. Results of Injections with Fli & Sox Binding Sites Mutated in Plasmid. Images depicting the EGFP expression in zebrafish injected with mutations to the Fli and Sox binding sites of the DE plasmid. (A) Ventral view of zebrafish head at 5 dpf displaying craniofacial cartilage. (B) Lateral view of zebrafish head to show expression in cartilage and ear. (C) Lateral view of trunk to visualize the notochord.

Identification of Transcription Factors

The identification of putative transcription factor binding sites was completed using bioinformatic analysis, and this focused my research on the Ets, Sox, and Runx families of transcription factors. We established the importance of the potential transcription factor binding sites based on the decreased reporter expression with mutated binding sites. Next we wanted to establish the exact member of the transcription factor family that was bound to the R2-DE region using overexpression analysis. To do this, we utilized RT-PCR to clone the zebrafish mRNA homologs for each of the candidate transcription factor proteins into in vitro transcription vectors. This allowed us to produce mRNA for each transcription factor to be injected into our R2-DE transgenic lines. These RNAs were injected into the embryo at the one-cell stage resulting in the early production of the transcription factors at a much earlier stage than normal in zebrafish. If the injected proteins were indeed bound to the regulatory region, they could prematurely produce the reporter protein that would be visible under fluorescence. The embryos of the stable line of the R2 and R2-DE zebrafish were injected with the RNA of *Fli1a*, *Sox9a*, and *Runx3* simultaneously in order to account for all of the theorized binding sites. As it can take up to 2 hours for protein to be made from injected mRNA, we chose to test for reporter activity from 6 to 10 hpf, which covers most of gastrulation but before somatogenesis when *col2a1a* expression is first seen by *in situ* hybridization (Yan et al 1995). When the injected embryos were inspected during gastrulation, the animal half of the embryo appeared to be fluorescent green. Unfortunately, the uninjected embryos were inspected they also appeared to have green fluorescence. There was an inability to distinguish

between the injected and uninjected embryos. In order to determine whether the green glow in the embryos was from EGFP or auto-fluorescence, an antibody was used to label EGFP in the cells. Unfortunately, once again there was no discernable difference between the injected and uninjected embryos because the fluorophore was present in both the control and injected embryos (Figure 29). Future experiments would be necessary to confirm the binding of the candidate transcription factors.

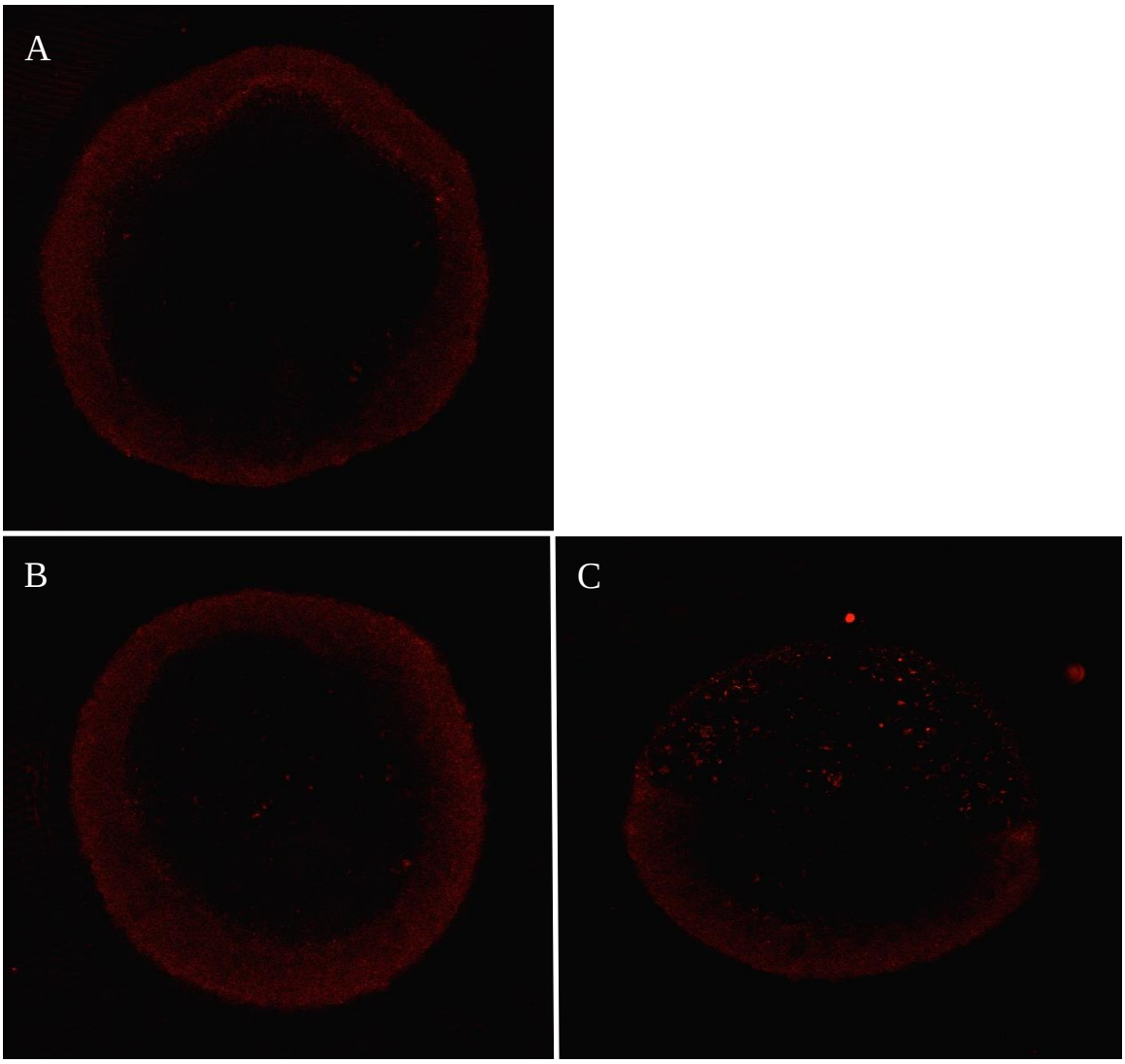


Figure 29. Results of Transcription Factor Injections. Images of stable line embryo injections with RNA of the transcription factors. Embryos stained with antibody for EGFP expression. (A) Wild-type embryos injected with the RNA of Fli1a, Sox9a, and Runx3. (B) Stable line of R2 embryos injected with the RNA of Fli1a, Sox9a, and Runx3. (C) Stable line of DE embryos injected with the RNA of Fli1a, Sox9a, and Runx3.

CHAPTER FOUR

DISCUSSION

Our experiments have identified the element of the R2 regulatory region that was responsible for the transcriptional initiation of *col2a1a*. The minimization of this region has allowed for the identification of specific transcription factor binding sites using comparative genomics. Site-directed mutagenesis analysis demonstrated the importance of those nucleotide sequences for the activity of R2, but it did not confirm the exact member of the transcription factor family binding there. Even though candidate transcription factors suspected to bind to these regions based on *in silico* analysis were identified, future experiments would be necessary to confirm our results. Based on the conserved nucleotide sequence *Fli1a*, *Runx2b*, and *Sox9a* were believed to bind to the R2-DE regulatory element. Future experiments will be to confirm the hypothesis, and whether the correct proteins of the transcription factor families were identified. This confirmation could be accomplished by chromatin immunoprecipitation assay and or an electromobility shift assay in order to confirm an interaction between the DNA sequence and transcription factor proteins. This interaction confirmation would provide evidence that the transcription factors are sufficient to drive expression of the *col2a1a* gene in zebrafish. These experiments would confirm the DNA binding interactions, which leaves the potential for future research to elucidate other protein-protein interactions in this complex that are necessary for the regulation of *col2a1a*.

The transcription factor binding sites we have focused on were identified based on the cartilage expression present in the R2:D2-E1 region of the *col2a1a* gene. It was interesting that the reporter plasmid with the entire DE region maintained expression in the tissues specific to *col2a1a* expression, but once it was minimized to the D2-E1 unit the notochord expression was lost. The loss of the first 30 bp and last 30 bp of the DE region have eliminated reporter expression in the notochord, to show that there must have been an important binding site for expression in this region. Notochord expression was low with the D-E1 region, completely eliminated in the D2-E1 region, but present in 20% of the embryos at the D2-E region. For this reason, I hypothesized there was an important transcription factor binding site for notochord expression in the E2 region. Once again comparative genomics and computer analysis could be used to identify the potential transcription factor binding sites, which would demonstrate the sequence could be important for notochord expression. Even though we have identified potential transcription factor binding sites for cartilage expression, it would be compelling to see which transcription factors differ in the notochord expression region. This process could identify an enhancer element that is responsible for isolated expression in the notochord, which would allow for transcriptional initiation in this specific tissue. This could be appealing for many biomedical reasons because of the ability to drive expression of any gene in a spatiotemporal manner with this enhancer sequence

Investigating the transcriptional regulation of zebrafish *col2a1a* gene entailed identifying the components sufficient to drive expression of this gene. The paralogue of this gene in zebrafish is *col2a1b*, which differs in its expression pattern based on *in situ*

hybridization analysis (Dale and Topczewski, 2011), could be another interesting avenue to explore. For example, *col2a1a* is present in the chondrocytes and the perichondrium of the pharyngeal arches while *col2a1b* is restricted to the perichondrium. The experiments completed on the regulation of *col2a1a* identified three potential transcription factor sites, but because the homologues have some overlap in tissue specificity their regulatory elements could be similar, especially if both paralogues were derived from an earlier gene duplication that would have maintained a version of the *col2a1a* R2 element. Similar to the process already completed, the identification of the regulatory element and transcription factor binding sites could be achieved with comparative genomics. The importance of transcription factor binding sites may again be confirmed with site-directed mutagenesis on the reporter plasmids. This research would allow a comparison into how the regulation of *col2a1a* and *col2a1b* has evolved.

REFERENCE LIST

Akiyama H, Chaboissier MC, Martin JF, Schedl A, de Crombrugghe B. The transcription factor Sox9 has essential roles in successive steps of the chondrocyte differentiation pathway and is required for expression of Sox5 and Sox6. *Genes Dev*. 2002;16(21):2813-2828.

Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. The Molecular Biology of the Cell. New York: Garland Science, 2007.

Andersson GB. Epidemiological features of chronic low-back pain. *Lancet*. 1999;354(9178):581-585.

Aszodi A, Chan D, Hunziker E, Bateman JF, Fassler R. Collagen II is essential for the removal of the notochord and the formation of intervertebral discs. *J Cell Biol*. 1998;143(5):1399-1412.

Aszodi A, Hunziker EB, Olsen BR, Fassler R. The role of collagen II and cartilage fibril-associated molecules in skeletal development. *Osteoarthritis Cartilage*. 2001;9 Suppl A:S150-9.

Bejerano G, Pheasant M, Makunin I, *et al.* Ultraconserved elements in the human genome. *Science*. 2004;304(5675):1321-1325.

Brown DM, Nichols BE, Weingeist TA, Sheffield VC, Kimura AE, Stone EM. Procollagen II gene mutation in stickler syndrome. *Arch Ophthalmol*. 1992;110(11):1589-1593.

Cheah KS, Lau ET, Au PK, Tam PP. Expression of the mouse alpha 1(II) collagen gene is not restricted to cartilage during development. *Development*. 1991;111(4):945-953.

Crombrugghe, Benoit de, Lefebvre V, Nakashima K. Regulatory mechanisms in the pathway of cartilage and bone formation. *Current Opinion in Cell Biology*. 2001; 13(6):721-728.

Dale RM, Topczewski J. Identification of an evolutionarily conserved regulatory element of the zebrafish col2a1a gene. *Dev Biol*. 2011;357(2):518-531.

Ding M, Lu Y, Abbassi S, *et al.* Targeting Runx2 expression in hypertrophic chondrocytes impairs endochondral ossification during early skeletal development. *J Cell Physiol*. 2012;227(10):3446-3456.

Dooley K, Zon L. Zebrafish: a model system for the study of human disease. *Current Opinion in Genetics & Development*. 2000; 10: 252-256

Elgar G, Vavouri T. Tuning in to the signals: Noncoding sequence conservation in vertebrate genomes. *Trends Genet*. 2008;24(7):344-352.

ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature.* 2012; 489: 57-74.

Exposito J.Y, Valcourt U, Cluzel C, Lethias C. The fibrillar collagen family. *Int. J. Mol. Sci*; 11 (2010): 407–426

Flores MV, Lam EY, Crosier P, Crosier K. A hierarchy of runx transcription factors modulate the onset of chondrogenesis in craniofacial endochondral bones in zebrafish. *Dev Dyn*. 2006;235(11):3166-3176.

Fox, AJS, Bedi A, Rodeo SA. The Basic Science of Articular Cartilage. *Sports Health*. 2009; 1(6): 461-468

Freemont AJ. The cellular pathobiology of the degenerate intervertebral disc and discogenic back pain. *Rheumatology (Oxford).* 2009;48(1):5-10.

Gao Y, Liu S, Huang J, *et al.* The ECM-cell interaction of cartilage extracellular matrix on chondrocytes. *Biomed Res Int*. 2014;2014:648459.

Gilbert, Scott F. *Developmental Biology.* Sunderland: Sinauer Associates, Inc., 2010. Print

Glickman NS, Kimmel CB, Jones MA, Adams RJ. Shaping the zebrafish notochord. *Development*. 2003;130(5):873-887.

Goldring MB. Chondrogenesis, chondrocyte differentiation, and articular cartilage metabolism in health and osteoarthritis. *Ther Adv Musculoskelet Dis*. 2012;4(4):269-285.

Gustafsson E, Aszodi A, Ortega N, *et al.* Role of collagen type II and perlecan in skeletal development. *Ann N Y Acad Sci*. 2003;995:140-150.

Iwamoto M, Higuchi Y, Enomoto-Iwamoto M, Kurisu K, Koyama E, Yeh H, Rosenbloom J, Pacifici M. The role of ERG (ets related gene) in cartilage development. *Osteoarthritis and Cartilage*, 2001;14(Suppl A):S41–47.

Kadler KE, Holmes DF, Trotter JA, Chapman JA. Collagen fibril formation. *Biochem J*. 1996;316:1-11.

Kawakami K. Transgenesis and gene trap methods in zebrafish by using the Tol2 transposable element. *Methods Cell Biol*. 2004;77:201-222.

Kerney R, Gross JB, Hanken J. Runx2 is essential for larval hyobranchial cartilage formation in xenopus laevis. *Dev Dyn*. 2007;236(6):1650-1662.

Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF. Stages of embryonic development of the zebrafish. *Dev Dyn*. 1995;203(3):253-310.

Kimmel, Charles K., Miller CT, Kruze G, Ullmann B, BreMiller RA, Larison KD, Snyder HC. The Shaping of Pharyngeal Cartilages during Early Development of the Zebrafish. *Developmental Biology.* 1998;203 (2): 245-263.

Kimmel CB, DeLaurier A, Ullmann B, Dowd J, McFadden M. Modes of developmental outgrowth and shaping of a craniofacial bone in zebrafish. *PLoS One*. 2010;5(3): e9475.

Knight RD, Schilling TF. Cranial neural crest and development of the head skeleton. *Adv Exp Med Biol*. 2006;589:120-133.

Ko LJ, Engel JD. DNA Binding Specificities of the GATA Transcription Factor Family. *Molecular and Cellular Biology*. 1993; 13(7):4011-22.

Kuivaniemi H, Tromp G, Prockop DJ. Mutations in collagen genes: cause of rare and some common diseases in humans. *FASEB J*. 1991;5(7): 2052-60.

Kuratani S. Developmental studies of the lamprey and hierarchical evolutionary steps towards the acquisition of the jaw. *J Anat*. 2005;207(5):489-499.

LeClair EE, Mui SR, Huang A, Topczewska JM, Topczewski J. Craniofacial skeletal defects of adult zebrafish glypican 4 (knypek) mutants. *Dev Dyn*. 2009;238(10):2550-2563.

Le Maitre CL, Pockert A, Buttle DJ, Freemont AJ, Hoyland JA. Matrix synthesis and degradation in human intervertebral disc degeneration. *Biochemical Society Transaction.* 2007;35(4):652-655.

Mackie EJ, Ahmed YA, Tatarczuch L, Chen KS, Mirams M. Endochondral ossification: How cartilage is converted into bone in the developing skeleton. *Int J Biochem Cell Biol*. 2008;40(1):46-62.

Masago Y, Hosoya A, Kawasaki K, *et al.* The molecular chaperone Hsp47 is essential for cartilage and endochondral bone formation. *J Cell Sci*. 2012;125(Pt 5):1118-1128.

Maston, Glenn A., Sara K. Evans, and Michael R. Green. "Transcriptional Regulatory Elements in the Human Genome." *Annual Review of Genomics and Human Genetics Annu. Rev. Genom. Human Genet.* 2006; 29-59.

Minoux M, Rijli FM. Molecular mechanisms of cranial neural crest cell migration and patterning in craniofacial development. *Development*. 2010;137(16):2605-2621.

Ng LJ, Wheatley S, Muscat GE, *et al.* SOX9 binds DNA, activates transcription, and coexpresses with type II collagen during chondrogenesis in the mouse. *Dev Biol*. 1997;183(1):108-121.

Otto F, Thornell AP, Crompton T, *et al.* Cbfa1, a candidate gene for cleidocranial dysplasia syndrome, is essential for osteoblast differentiation and bone development. *Cell*. 1997;89(5):765-771.

Palstra RJ, Grosveld F. Transcription factor binding at enhancers: Shaping a genomic regulatory landscape in flux. *Front Genet.* 2012;3:195.

Pruitt KD, Harrow J, Harte RA, *et al.* The consensus coding sequence (CCDS) project: Identifying a common protein-coding gene set for the human and mouse genomes. *Genome Res*. 2009;19(7):1316-1323.

Roberts S, Evans H, Trivedi J, Menage J. Histology and pathology of the human intervertebral disc. *J Bone Joint Surg Am*. 2006;88 Suppl 2:10-14.

Sakabe NJ, Savic D, Nobrega M. Transcriptional enhancers in development and disease. *Genome Biol.* 2012;13(1): 238.

Schilling, Thomas F.Genetic analysis of craniofacial development in the vertebrate embryo. *Bioessays*. 1997; 19(6):459-468.

Siepel A, Bejerano G, Pedersen JS, *et al.* Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res*. 2005;15(8):1034-1050.

Spitz F, Furlong EE. Transcription factors: From enhancer binding to developmental control. *Nat Rev Genet*. 2012;13(9):613-626.

Vega RB, Matsuda K, Oh J, *et al.* Histone deacetylase 4 controls chondrocyte hypertrophy during skeletogenesis. *Cell*. 2004;119(4):555-566.

Westerfield, M. (2000). The zebrafish book. A guide for the laboratory use of zebrafish (Danio rerio). 4th ed., Univ. of Oregon Press, Eugene.

Whitfield TT, Riley BB, Chiang MY, Phillips B. Development of the zebrafish inner ear. *Dev Dyn*. 2002;223(4):427-458.

Winterpacht A, Hilbert M, Schwarze U, Mundlos S, Spranger J, Zabel BU. Kniest and stickler dysplasia phenotypes caused by collagen type II gene (COL2A1) defect. *Nat Genet*. 1993;3(4):323-326.

Woolfe A, Goodson M, Goode DK, *et al.* Highly conserved non-coding sequences are associated with vertebrate development. *PLoS Biol*. 2005;3(1):e7.

Yan YL, Hatta K, Riggleman B, Postlethwait JH. Expression of a type II collagen gene in the zebrafish embryonic axis. *Dev Dyn*. 1995;203(3):363-376.

Yan YL, Miller CT, Nissen RM, *et al.* A zebrafish sox9 gene required for cartilage morphogenesis. *Development*. 2002;129(21):5065-5079.

Yan YL, Willoughby J, Liu D, *et al.* A pair of sox: Distinct and overlapping functions of zebrafish sox9 co-orthologs in craniofacial and pectoral fin development. *Development*. 2005;132(5):1069-1083.

Yang C, Li SW, Helminen HJ, Khillan JS, Bao Y, Prockop DJ. Apoptosis of chondrocytes in transgenic mice lacking collagen II. *Exp Cell Res*. 1997;235(2):370-373.

Yang X, Karsenty G. Transcription factors in bone: Developmental and pathological aspects. *Trends Mol Med*. 2002;8(7):340-345.

Zhang G, Eames BF, Cohn MJ. Chapter 2. Evolution of vertebrate cartilage development. Curr Top Dev Biol 2009; 86: 15–42. doi: 10.1016/s0070-2153(09)01002-3.

Zhou G, Garofalo S, Mukhopadhyay K, *et al.* A 182 bp fragment of the mouse proα1(II) collagen gene is sufficient to direct chondrocyte expression in transgenic mice *J Cell Sci*. 1995(106):3677.

VITA

Sonja Dabizljevic of Chicago, Illinois, graduated from Loyola University Chicago in 2009 with a Bachelor of Science degree in Biology. After working at Dow Chemical as a laboratory technician for almost three years, she returned to the Master of Science program at Loyola University Chicago in 2012 where she joined the laboratory of Dr. Rodney M. Dale. Ms. Dabizljevic is currently continuing her education at the University of Chicago as a doctoral student in the Committee on Development, Regeneration, and Stem Cell Biology.