The Dissertation Committee for Patrick Dooling McGurk certifies that this is the approved version of the following dissertation:

Genetic analysis of tendon development and musculoskeletal integration in the vertebrate head

Committee:

Johann Eberhart, Supervisor

John Wallingford

Steven Vokes

Vernita Gordon

Vishwanath Iyer

Genetic analysis of tendon development and musculoskeletal integration in the vertebrate head

by

Patrick Dooling McGurk

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin
August 2017

Dedication

To my daughter, Clara, you light up my life.

Acknowledgements

I owe a deep wealth of gratitude to my advisor, Dr. Johann Eberhart. Johann has given me unconditional patience as I figured out the direction of my research and career, support to pursue the lines of investigation that interest me. He is a phenomenal writer as well, and my grant and manuscript submissions have been much the better for it. Johann's contributions to my technical development are second only to those of Mary Swartz, whose talents and willingness to share them are abundant. I also want to thank many current and former Eberhart lab members whose diverse areas of expertise broadened my perspectives and techniques. I am grateful for the help and companionship of Dr. Kelly Sheehan-Rooney, Dr. Ben Lovely, Dr. Neil McCarthy, Dr. Desire' Buckley, Dr. Yohaan Fernandes, Ben Wells, Alfire Sidik, Tim Kuka, and Ranjeet Kar. Thanks also to the technicians who have kept the lab running and the fish healthy over the years, especially Briana Schroeder, Anna Percy, Angie Martinez, and Taylor Yamakawa. Special thanks to Rhonda Stanley and Luana Kohnke, whose eagerness to learn led them to serve as fine laboratory assistants, but more importantly for the lessons they taught me as a mentor.

My committee members have been a source of great generosity over the past several years. I am thankful for the wisdom and guidance of Drs. Steve Vokes, John Wallingford, Vishy Iyer, and Vernita Gordon. Steve and John especially have shown concern for my personal development above and beyond their positions as committee members, in addition to their stellar expertise in developmental biology.

I have had the privilege of collaborating with excellent investigators outside of UT and communicating and sharing resources with researchers locally and around the world. I am grateful to Dr. Jenna Galloway and Dr. Jessica Chen of Harvard Medical School for sharing their resources and passion for investigating zebrafish tendon development. I also

Genetic analysis of tendon development and musculoskeletal integration in the vertebrate head

Patrick Dooling McGurk, Ph.D.

The University of Texas at Austin, 2017

Supervisor: Johann K Eberhart

We investigated zebrafish mutants that affect craniofacial development to examine how those mutations affect the musculoskeletal patterning of the jaw. We discovered that loss of function in *cyp26b1* caused musculoskeletal patterning defects related to tendon condensation. Subsequently, we decided to use reverse genetics to interrogate regulation of tendon development in the vertebrate head and body. Needing ways of evaluating tendon and myotendinous junction phenotypes, we devised a protocol for inducing muscle strain *in vivo*. We generated mutants for *scxa* and *xirp2a,* determined whether homozygous mutants had developmental defects, and tested their muscle attachment stability under strain. In conjunction with our mutagenesis efforts, I developed a software program that would aid in the identification of mutant alleles in the F1 offspring of CRISPR mutagenized animals.

Table of Contents

List of Figures

Chapter 1: Introduction and Significance

The vertebrate head is an intricate wonder of development. Twenty-two bones make up the human skull, and over 100 skeletal muscles combine to move the face, jaw, eyes, ears, nose, tongue, and pharynx. It has long been known that the tendons of the head derive from neural crest, the multipotent precursors of the craniofacial skeleton (Noden, 1983b). Neural crest cells also provide positional cues that pattern the mesoderm-derived cranial muscles (Noden, 1983a; Tokita & Schneider, 2009) but the mechanisms that direct this patterning remain unclear. However, as cranial muscle extends from its original pharyngeal arch to attach elsewhere, so too do neural crest cells that form the muscle's distal attachment (Köntges & Lumsden, 1996). Together, these observations suggest that tendon precursors guide the elongation and attachment of cranial muscles, but to date no one has demonstrated such an ability.

Of the connective tissues that give structure and motility to vertebrate animals, tendon's development is the least well understood. Links between molecular markers and cytological milestones have been characterized for most developmental stages of cartilage/bone (reviewed in Kozhemyakina et al., 2015; Li & Dong, 2016) and muscle (reviewed in Buckingham & Vincent, 2009), but few are known for tendons (reviewed in Gaut & Duprez, 2016). We believe that to understand the patterning of skeletal muscles and their attachments we must know more about the genetics and morphogenesis of tendons.

The developmental origins and bauplan of the craniofacial musculoskeletal system are highly conserved among gnathostome (jawed vertebrate) species, though the exact morphology of individual elements varies widely among the many clades of gnathostomes. We have chosen to study the genetics of craniofacial development in the zebrafish, *Danio*

rerio, for several reasons. External fertilization and optical clarity of zebrafish embryos make them ideal subjects for *in vivo* microscopy, which expands the utility of transgenic labels. Zebrafish develop rapidly, which allows time-lapse experiments to cover more developmental stages in less time. Their rate of maturation and high fecundity also make zebrafish ideal for forward genetic screens. Though previously a weakness for the zebrafish model, reverse genetic approaches are now easily carried out thanks to technological advances in nuclease-based *in vivo* genome editing.

Vertebrate craniofacial development

The first vertebrates emerged around 525 million years ago (Shu et al., 1999), characterized by a bony, jointed spinal cord and articulated head with sensory placodes and an oropharynx, which together eventually allowed active predation to replace filter feeding. The "new head" that enabled this lifestyle was born out of the emergence of a cell type novel to vertebrates, the neural crest, called by some a "fourth germ layer" (Gans & Northcutt, 1983; Northcutt, 2005) Neural crest cells are specified from the ectoderm at the neural plate border, and undergo an epithelial-to-mesenchymal transition during neural tube closure (Fig 1.1A). Neural crest mesenchyme separates from the neural tube and surface ectoderm epithelia and migrates into regions throughout the body. Various neural crest populations then undergo differentiation to become pigment cells, peripheral and enteric nerves, and cardiomyocytes.

The most rostral, or cranial, neural crest migrates in three so-called streams, corresponding with the axial level of the brain at which the neural crest cells were specified (Fig 1.1B). Cranial neural crest cells gather on the ventral side of the head into serially reiterated transient structures called pharyngeal arches (Fig 1.1C). Fundamental similarities of craniofacial form exist throughout vertebrates due to their shared phylotypic stage, which is named the pharyngula, after the pharyngeal arches (Galis & Metz, 2001).

2
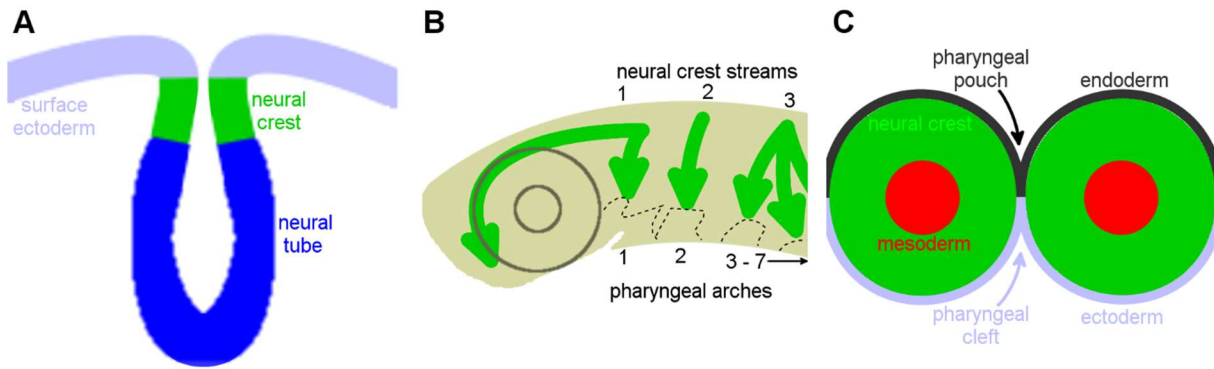
Fig 1.1 Neural crest specification and formation of the pharyngeal arches

A) Illustration of a transverse section through the neural tube during its closure, dorsal up. Neural crest cells are specified at the neural plate border. They will delaminate from the closing neural tube and the non-neural ectoderm will join together to form a surface epithelium. B) Illustration of a zebrafish embryo in lateral view, anterior to the left. Arrows indicates migration paths of each neural crest stream, pointing to the future pharyngeal arches they will populate. C) Illustration of a frontal section through two adjacent pharyngeal arches, medial up, lateral down.

For example, the first stream of cranial neural crest cells corresponds with the first pharyngeal arch, which gives rise to structures including the palate and the mandible and maxilla of the jaw. The second stream corresponds with the second arch and gives rise to skeletal elements including the jaw supports. The third stream corresponds with the posterior arches (3+), which each make specific skeletal contributions to the pharynx, thymus, and thyroid.

Germ layer interactions in the pharyngeal arches

Each pharyngeal arch is visibly defined by its neural crest component, but the other germ layers also contribute to each arch and undergo specific morphogenesis based on arch identity. Head mesoderm divides and migrates into the core of each pharyngeal arch, where it gives rise to the cranial muscles. Around each mesodermal core are the neural crest cells. The two populations of arch mesenchyme are situated between the surface ectoderm and pharyngeal endoderm epithelia. Endoderm evaginates to form pharyngeal pouches that

3

connect to surface ectoderm invaginations (pharyngeal clefts) and separate adjacent arches. In a sense, each pharyngeal arch is similar to a peanut M&M. The candy coating in this analogy is half endodermal and half ectodermal epithelia, the chocolate is mesenchymal neural crest, and the peanut at the center is mesoderm.

Much of what we know about patterning the pharyngeal arch derivatives pertains to skeletal development. Intrinsic and extrinsic patterning of the neural crest defines the skeletal elements of each pharyngeal arch. First, arch identity is established prior to neural crest migration (Noden, 1983b), which is reflected in the rostro-caudal expression pattern of Hox genes. The pharyngeal endoderm is also pre-patterned along the rostro-caudal axis, and this patterning is important later for the identity and orientation of cartilages derived from the neural crest (Couly et al., 2002; Ruhin et al., 2003). The pharyngeal arches are also subject to several interactions that differentiate fates along the dorsal-ventral axis. BMP and endothelin (Edn1) signals come from the ventral ectoderm (Lovely et al., 2016; reviewed in Medeiros & Crump, 2012). The pharyngeal endoderm and neural crest mesenchyme express Six1, which both promotes Jagged-Notch signaling between pharyngeal pouches and the dorsal neural crest and opposes dorsal Edn1 expression (Tavares et al., 2017). The opposing gradients of Jagged and Edn1 signals define dorsal, intermediate, and ventral regions of pharyngeal arches 1 and 2, demarcated by the nested expression of Distal-less homeobox (Dlx) transcription factors (Depew et al., 2002, 2005).

Some of the aforementioned factors that pattern the pharyngeal arch neural crest also promote myogenesis in the head. Expression of Dlx5 and Dlx6 in cranial neural crest cells is required for jaw muscles to form, as shown in mice (Heude et al., 2010). Six1a is expressed in zebrafish pharyngeal mesoderm, and knockdown of Six1a function leads to an almost complete loss of head muscles because Six1a promotes expression of the myogenic transcription factor MyoD (Lin et al., 2009). Though these transcription factors

4

participate in the specification of individual skeletal elements in the first and second arches, they do not control the specification of individual muscles.

Patterning the jaw musculoskeletal system

Schilling and Kimmel (1997) described the emergence of muscle elements in the zebrafish pharyngeal arches, noting that myosin and MyoD expression appears in different muscles at different times despite a common pharyngeal arch core origin. Fig 1.2 schematizes the musculoskeletal system of the larval zebrafish jaw. Schilling and Kimmel (1997) observed the appearance of ocular muscles early in embryonic day 3. Concurrently, the adductor mandibulae and sternohyal muscles appear, the former from the first arch and the latter from anterior somites. Subsequently, anterior and posterior muscle pairs appear in the second arch, the interhyal and hyohyal muscles. Later, a triangle of intermandibularis muscles appears from the first arch. One anterior muscle connects between the bilateral Meckel's cartilage condensations and two posterior muscles extend toward the medial tip of the ceratohyal condensations of the second arch. They also noted that cranial muscles typically appeared after the formation of pre-cartilage condensations that the muscles would later connect to.

A number of studies have described the influence of neural crest on the development of cranial muscles. The cranial neural crest gives rise to all of the skeletal elements of the jaw (Köntges & Lumsden, 1996; Eberhart et al., 2006), and also the tendons and connective tissues of the head (Noden, 1988; Couly et al., 1992). Furthermore, muscles derived from pharyngeal arch mesoderm make their attachments with connective tissue derived from neural crest from the same pharyngeal arch (Köntges & Lumsden, 1996). Noden (1983b) demonstrated that cranial neural crest cells grafted heterotopically confer their arch identity on both the skeletal and muscle derivatives of the arch that the crest cells migrate into. Beyond arch identity, neural crest cells grafted orthotopically between duck
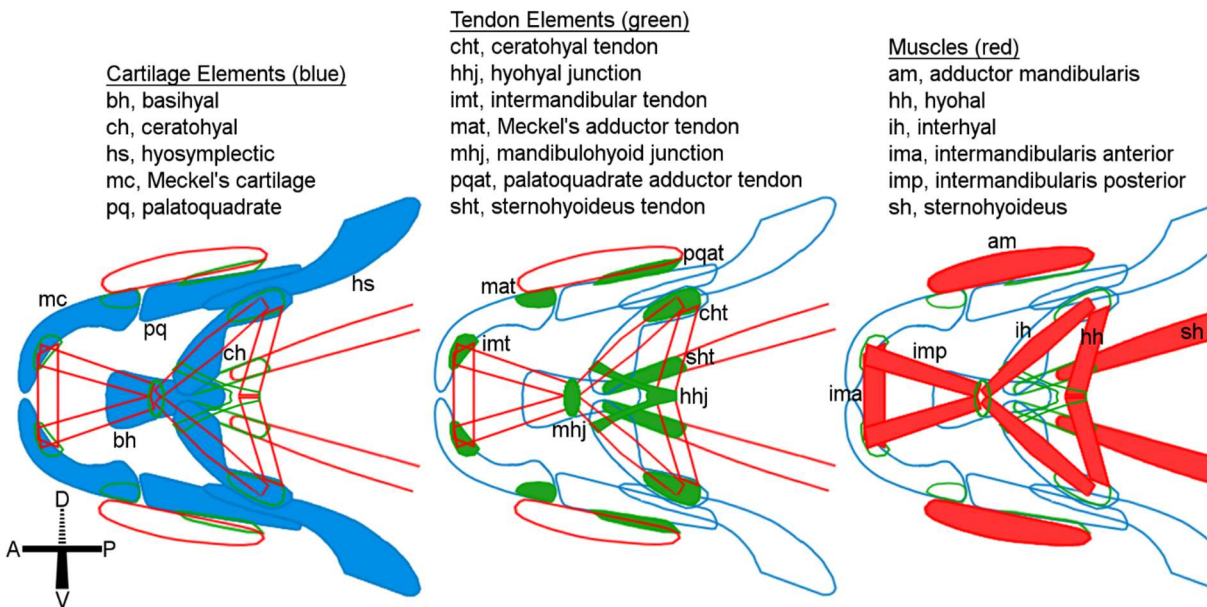
Fig 1.2. Musculoskeletal pattern of the larval zebrafish lower jaw.

Schematic modified from Chen and Galloway, 2014.

and quail embryos conferred both skeletal and muscular morphologies specific to the donor species (Tokita & Schneider, 2009). Conversely, Trainor et al. (1994) showed that heterotopic grafts of cranial paraxial mesoderm carry no axial fate specification with them, and grafted cells adopt muscle fates according to their new axial position. These findings suggest that pharyngeal arch mesodermal cores contain little in the way of intrinsic programming for shaping their muscle derivatives.

VERTEBRATE TENDON DEVELOPMENT

Specification of tenoblasts from multipotent precursors

The primary marker of tendon and ligament lineages is Scleraxis (Scx), a basic helix-loop-helix (bHLH) transcription factor expressed in progenitor cells and mature tissues alike (Cserjesi et al., 1995; Schweitzer et al., 2001; Pryce et al., 2007; Mendias et al., 2012). One of the main regulators of *Scx* is the fibroblast growth factor (FGF) family

of secreted ligands. Global inhibition of FGF signaling eliminates *Scx* expression in somites, limbs (Brent et al., 2005), and the head (J. W. Chen & Galloway, 2014). Brent et al. (2003) first identified a spatial correspondence in chick somites between *Fgf8* expression in the myotome and *Scx* expression in a compartment of sclerotomal cells abutting the myotome, and then demonstrated the sufficiency of FGF8 to induce *Scx* expression. FGF8 signals through two transcription factors, Etv4 (formerly Pea3) and Etv5 (formerly Erm), to induce *Scx* expression in the sclerotome (Brent & Tabin, 2004). Brent et al. (2005) carried this line of investigation over to a mouse model, where they observed *Scx* is not expressed in the sclerotome after myotome removal. *Scx* expression expands in the sclerotome of *Sox5;Sox6* double mutant mice, indicating that skeletal specification opposes tendon specification except where myotomal FGF8 is most concentrated (Brent et al., 2005). During mouse and chick limb development, muscles express *Fgf4* and *Fgf8*, and muscles and tendons both express *Etv4* and the Sprouty FGF signaling modulators *Spry1* and *Spry2* (Eloy-Trinquet et al., 2009). Perhaps a positive FGF feedback mechanism promotes the specification of tendon from mesenchyme surrounding developing muscles, as FGF4 drives its own effectors Etv4 and Sprouty in chick limb mesenchyme, and expression of those same effectors in mouse limb tendons depends on muscle (Eloy-Trinquet et al., 2009) Some controversy exists surrounding the time-, tissue-, or species-specific roles for FGF, though. FGF4 did not significantly affect tendon gene expression at the time of tendon progenitor specification in mouse limb explants, but blockade of ERK, which causes severe downregulation of *Etv4* and *Spry2*, actually significantly activated expression of *Scx* as well as *Col1a1, Col1a2,* and *Thbs2*, extracellular matrix proteins enriched in tendon (Havis et al., 2014).

The other major regulator of *Scx* is the transforming growth factor beta (TGFβ) family. TGFβ signaling is not required for the specification of tendon progenitors, but it

does directly regulate *Scx*. *Scx*-positive tendon progenitors gradually disappear in *Tgfb2;Tgfb3* double mutant mice and in the limbs of *Tgfbr2* conditional knockout mice (Pryce et al., 2009). TGFβ2 is sufficient to increase expression of *Scx, Col1a1, Col1a2, Thbs2, and Thbs4* via SMAD2/3 in early mouse limb bud explants (Havis et al., 2014). As tendon progenitors transition toward differentiation, they express high levels of TGFβ pathway components, and even higher levels in differentiated tendons (Havis et al., 2014). These findings suggest that TGFβ is necessary and sufficient for tendon development but unnecessary for specification.

Muscle-dependent tendon differentiation

The discovery that axial tendon tissue is induced by myotome begat a number of investigations into tendon development in mutant animals that fail to specify muscles. Interestingly, both head and limb tenoblasts express *Scx* initially in the absence of muscles, but subsequently lose *Scx* expression (Schweitzer et al., 2001; Edom-Vovard et al., 2002; Grenier et al., 2009; J. W. Chen & Galloway, 2014; A. H. Huang et al., 2015). This makes sense given the expression of FGF and TGFβ ligands by muscles. Muscles also provide contractile force, and mechanical stimulation synergizes with *Scx* to promote tendon differentiation from human mesenchymal stem cells in culture (X. Chen et al., 2012). To test the *in vivo* role of muscle contractile forces on tendon development, Havis et al. (2016) pharmacologically immobilized chick embryos and cultured chick limb explants. After immobilization or explant, limbs expressed significantly less FGF and TGFβ ligands and effectors (*Etv4* and *Spry2; Smad7* respectively), as well as *Scx, Thbs2,* and *Tnmd* (Havis et al., 2016)*. Tnmd* encodes a transmembrane protein (tenomodulin) that promotes tenoblast proliferation and proper extracellular matrix organization (reviewed in Dex et al., 2016). Exogenous FGF4 induced expression of *Etv4, Spry2,* and *Scx* in immobilized and explanted limbs, while exogenous TGFβ2 induced expression of *Smad7, Scx, Tnmd,* and *Thbs2* in

explanted limbs (Havis et al., 2016). Thus, growth factor signaling required for maintenance and differentiation of tendon progenitors depends on muscle contraction

Functional analysis of tendon markers

Though expressed elsewhere, *Scx* and *Tnmd* are so enriched in tendon that they have become *de facto* markers of tendon progenitors and differentiating tendons, respectively. Interestingly, Scx-null mice form tendons at all muscle attachment sites, though force-transmitting tendons have morphological and functional defects and tendons fail to express *Tnmd* (Murchison et al., 2007). Scx overexpression in mesenchymal stem cells induces *Tnmd* expression (Alberton et al., 2012), indicating that Scx is necessary and sufficient for *Tnmd* expression.

Scx also directly regulates *Col1a1* (Lejard et al., 2007) and *Col1a2* (Espira et al., 2009; Bagchi and Czubryt, 2012), but neither Scx nor Tnmd is necessary for the formation of tendons or tendon extracellular matrix (Murchison et al., 2007). *Tnmd* mutant mice have less uniform and generally larger collagen fibril diameters than wild-type mice, however (Docheva et al., 2005). Prior to the discovery of *Scx*, a matrix marker called Tenascin C (TNC) was similarly used to track tendon development, but mouse null mutants displayed no identifiable phenotype. Thus, no "master regulator" of tendon is known, but there are several useful markers of the tendon lineage. It is of great interest to the tendon research community to elucidate the function of *Scx* in making "proper" tendons given that it is not required for tendons to form.

Structure and composition of mature tendons

Mesenchymal tendon progenitors eventually become tendon tissue, which is hypocellular and mostly composed of tough, flexible extracellular matrix. For the purposes of this work it is worth noting that adult zebrafish tendons and ligaments resemble those of mammals in structure and makeup (Chen & Galloway, 2014). The dry mass of tendon is

86% collagen, almost all of which is type I collagen (Jozsa & Kannus, 1997). Collagen molecules combine to form 50-500 nm diameter fibrils, held together by proteoglycans and bundled into fibers by a sheath called the endotenon. Fibers are bundled hierarchically by endotenon sheaths to form 50-300 µm diameter fascicles in which elongated collagen-producing tenocytes reside between fibers. Fascicles are grouped together by a sheath called the epitenon to form 100-500 µm diameter tendons (Fratzl, 2003).

The structures of tendon attachments to bone and muscle are quite different. Tendon attaches to bone via an *enthesis.* The enthesis can be fibrous, in which mineralized collagen fibers connect directly to bone, or most often fibrocartilaginous, exhibiting a gradated transition from tendon to fibrocartilage to mineralized fibrocartilage to bone (reviewed in Lu & Thomopoulos, 2013). Tendon attaches to muscle via a *myotendinous junction* (MTJ). The MTJ consists of actin cytoskeletal elements within the end of the muscle, transmembrane proteins to connect cytoskeleton to the muscle basement membrane outside the sarcolemma, finger-like extensions of the sarcolemma, and protein linkers between the basement membrane and collagenous tendon matrix (reviewed in Charvet et al., 2012).

IN VIVO GENOME EDITING FOR REVERSE GENETIC ANALYSIS

In very recent memory, mouse was the ideal model organism for analyzing the developmental effects of mutating a gene of interest. The basic "knockout mouse" technique utilizes homologous recombination in embryonic stem cells to insert a positive selection cassette where it interrupts the targeted gene's sequence. The faults of this method are mostly procedural. There is the time involved in generating a homology construct or the expense of having one synthesized. There is the transfection and subsequent selection of ES cells, after which the selected cells are inserted into cultured mouse embryos at the blastocyst stage and the chimeric blastocysts implanted into the uterus of a female mouse. Of course, the mutated ES cells may or may not contribute to a chimeric mouse's germline,

so some individuals will not be able to transmit the mutation to subsequent generations. The process is highly technical but also adaptable. One can design a very specific mutation to introduce to their gene of interest, or express another gene under the targeted gene's promoter. Where knocking out a gene results in early embryonic lethality or deformity, the technique can also be used to generate conditional mutants in which the gene can be mutated in a specific tissue and/or at a specific developmental stage.

In vivo genome editing offers the broader genetics and development research communities the potential to generate such a range of genetic modifications with similar or greater efficiency than homologous recombination in mouse ES cells. The basic principle involves generating a double-strand break in a specific target within the genome. Sometimes these breaks will be fixed by homology-directed repair (HDR), which can be exploited with homology-linked recombinant DNA to create knock-ins. Sometimes no homology template is available (breaks in both chromosomes, for instance) or for some other reason the break is fixed by non-homologous end-joining (NHEJ). Genome editing relies on the random and error-prone nature of NHEJ to obtain insertion/deletion mutations (indels) that cause gene coding sequence frameshifts, though other types of mutations are also sometimes desirable. Often, genome editing will create a number of different mutations given efficient targeting of the nuclease.

The crucial technology to efficient targeting is an engineered nuclease, similar to the restriction endonucleases expressed by bacteria to protect against foreign DNA and used daily in molecular biology labs across the world. Most of these endonucleases recognize relatively short DNA sequences, but others, called meganucleases, recognize longer sequences, making them highly specific. For example, the well-known meganuclease I-SceI has an 18-base recognition site. Endonucleases with shorter sequences are liable to recognize several sites throughout an animal genome, but the odds

of any string of 18 nucleotides matching the I-SceI recognition sequence are 1 in 4^{18}. That's over 68.7 billion basepairs, or roughly twenty times the size of the human genome. To utilize that specificity for targeted mutagenesis, engineering by fusion of different meganucleases (Chevalier et al., 2002) or by mutagenesis has been used to generate novel sequence-specific meganucleases that exhibit low toxicity in cells (Smith et al., 2006). Engineering and testing these custom nucleases is a laborious and costly process, however, and more efficient and cost-effective methods have adapted the ability to target nucleases to highly specific DNA sequences.

In lieu of altering the sequence specificity of nucleases, researchers sought out nucleases with separate sequence recognition and DNA-cleaving activity. Such non-specific endonuclease domains were isolated, leading to the development of zinc-finger (ZFN) and transcription activator-like effector (TALEN) nucleases. One can search their gene or locus of interest for algorithmically-selected target sites and receive instructions for the assembly of the nuclease sequence. For ZFNs, recognizing long, specific binding sites requires an array of zinc-finger domains, which are encoded individually in plasmid vectors to be combined together and linked to the nuclease sequence. For TALENs, assembly is even more modular, arraying plasmid vectors for TALE repeat elements that each recognize a single base of the binding sequence. For both methods, a nuclease protein is designed for both strands of DNA, one on either side of the cut site. This is because at the cut site, the nuclease domains must dimerize to break open the DNA. This dimer-based approach greatly reduces the chances of an individual ZFN or TALEN creating off-target breaks, but also doubles the amount of plasmid engineering and protein testing. TALENs display less toxicity than ZFNs and the greatest specificity and efficiency of all engineered nucleases (Kim & Kim, 2014; Boglioli & Richard, 2015). Their main disadvantages,

however, are the cost and expertise required for assembly of the nuclease constructs and validation of the protein products.

By comparison, the cost-effectiveness and technical ease of the CRISPR/Cas9 system has led to an abundance of published papers utilizing the two-component system (Cas9 and guide RNA) reported in 2012 (Jinek et al., 2012). Additionally, CRISPR has become a household name due to worldwide press attention over the possibilities of widely available genome editing as well as patent disputes over the technology. The key to this technique's flexibility is the mechanism by which Cas9 recognizes its target sites. The nuclease enzyme, rather than binding to a specific DNA sequence, binds to specific CRISPR RNA (crRNA) molecules. The simplest CRISPR/Cas9 system utilizes a single guide RNA (sgRNA) with constant sequence aside from a ~20 nucleotide targeting sequence. The targeting sequence can be cloned into a gRNA-coding plasmid, or synthesized as part of a long oligonucleotide gRNA template. Part of the constant RNA sequence binds to Cas9, which unwinds the DNA double helix and compares the DNA sequence to see if it matches the targeting portion of the RNA sequence, at which point Cas9 creates a double-strand break.

The main restriction of Cas9's gene targeting capabilities is its requirement for a protospacer-adjacent motif (PAM). A target sequence must be followed by 5'-NGG-3'. The PAM tells Cas9 where to compare the DNA and RNA sequences, which is much faster than comparing against the entire genome. This of course restricts the possible target sites. However, other CRISPR-associated nucleases with different PAMs could accommodate when a target site with Cas9 PAM is unavailable (Zetsche et al., 2015; Fonfara et al., 2016). Cas9 is sufficient for a vast majority of cases, though, and readily available in plasmid or protein form. Researchers have even engineered codon-optimized Cas9 genes in which the peptide sequence is unchanged but the mRNA sequence is more efficiently translated by

utilizing the most abundant tRNAs for any given amino acid in the subject species, including human, zebrafish, and several plant species.

As ZFNs and then TALENs came into vogue, there was a sense that they would revolutionize reverse genetics in zebrafish. Then Cas9 made genome editing incredibly accessible. Zebrafish, known as an excellent forward genetic model organism, now has the tools to be excellent for reverse genetics, and researchers are working to maximize the potential of homology-directed and homology-independent repair. There's an ongoing renaissance in generating new mutant lines in zebrafish and other organisms in which reverse genetics were previously labor-intensive. This is an incredible boon to genetics overall. Zebrafish fecundity and rapid development allows new mutants to be generated and screened for phenotypes that might be of interest to mouse researchers, for instance, whose animals are expensive and slower-growing. Other advantages of the zebrafish model are enhanced by solid reverse genetics, like their amenability to *in vivo* time-lapse imaging and high-throughput screening approaches (Lessman, 2011). The humble zebrafish has perhaps never been more relevant as a model.

Chapter 2: *In vivo* zebrafish morphogenesis shows Cyp26b1 promotes tendon aggregation, musculoskeletal patterning in the embryonic jaw

[Adapted from an article by the same title submitted to PLoS Genetics, currently in revision. PDM designed and carried out experiments, curated and analyzed data, engineered the *503unc:mCherry* construct and zebrafish line, assembled figures and movies, and wrote and revised the manuscript.]

Integrated development of diverse tissues gives rise to a functional, mobile vertebrate musculoskeletal system. However, the genetics and cellular interactions that drive the integration of muscle, tendon, and skeleton are poorly understood. In the vertebrate head, neural crest cells, from which cranial tendons derive, pattern developing muscles just as tendons have been shown to in limb and trunk tissue, yet the mechanisms of this patterning are unknown. We determined that *cyp26b1* is critical for musculoskeletal integration in the ventral pharyngeal arches, particularly in the mandibulohyoid junction where first and second arch muscles interconnect. Using time-lapse confocal analyses, I detail musculoskeletal integration in wild-type and *cyp26b1* mutant zebrafish. In wild-type fish, tenoblasts migrate in apposition to elongating muscles and aggregate in discrete muscle attachment sites. In the absence of *cyp26b1*, tenoblasts are generated in normal numbers but fail to aggregate into nascent tendons within the ventral arches and, subsequently, muscles project into ectopic locales. These ectopic muscle fibers eventually associate with ectopic tendon marker expression. Genetic mosaic analysis demonstrates that neural crest cells require Cyp26b1 function for proper musculoskeletal development. Using an inhibitor, I find that Cyp26 function is required in a short time window that overlaps the dynamic window of tenoblast aggregation. However, *cyp26b1* expression is largely restricted to regions between tenoblast aggregates during this time. My results suggest that degradation of retinoic acid (RA) by this previously undescribed population

of neural crest cells is critical to promote aggregation of adjacent *scxa*-expressing tenoblasts and that these aggregates are subsequently required for proper musculoskeletal integration.

INTRODUCTION

The movements and functions of the human head depend on 150 individual muscles, and loss of tendon or of tendon-muscle interactions has been implicated in human craniofacial syndromes with muscle defects. Vertebrate craniofacial development is a complex process involving communication between muscles, tendons, cartilages and surrounding tissues. Much of this communication occurs in transient, reiterated, pharyngeal arches. Within each pharyngeal arch are neural crest cells that form a specific set of skeletal elements and mesoderm cells that form a specific set of muscles (Schilling & Kimmel, 1997). The developmental origins of the various cranial tendons remain unclear, save that they derive from the neural crest (Grenier et al., 2009; Chen & Galloway, 2014). Given that neural crest cells control the patterning of cranial muscles (Noden, 1983a; Tokita & Schneider, 2009), it is likely that cranial tendons provide the positional cues needed for this patterning. We know little, however, about the genetics and cellular interactions underlying the mechanisms of development at the myotendinous junction.

We have a limited number of tendon disease models from which to gain an understanding of tendon functions in development. Muscle elements form and elongate after surgical removal of tendon primordia in the developing avian hindlimb, but in these limbs, muscle fibers are ectopically localized (Kardon, 1998). This result suggests that tendons restrict the patterning of limb musculature. Genetic disruption of *Scx* in mice causes defects in force-transmitting tendons in the trunk, limbs, and tail (Murchison et al., 2007). However, Scx function is not necessary for tendons to develop and form functional muscle attachments in mutant mice. Scx is also not sufficient for tendon development

because tendon progenitors (tenoblasts) are specified in the pharyngeal arches but eventually disappear in muscle-less mutant models (Chen & Galloway, 2014; Brent et al., 2005). These findings demonstrate that in the head, unlike in the somites (Schweitzer et al., 2001), specification of tendons is independent of muscle. However, in all tendon populations muscle provides mechanical stimulation that promotes growth factor signaling and *Scx* expression, and muscle contraction is as necessary as the muscle tissue itself for tendon differentiation (Havis et al., 2016). Thus, to study tendon functions, it will be essential to discover genetic models in which muscle differentiation is normal but muscle attachments are defective.

Larval zebrafish have a well-characterized complex craniofacial musculoskeletal system (Fig 1.2; Schilling & Kimmel, 1997; Chen & Galloway, 2014). In the zebrafish head, the more ventral portion of the first and second pharyngeal arches gives rise to the lower jaw and its supports (Schilling & Kimmel; 1997). Seven muscles extend across the ventral surface of both arches in an hourglass-like shape whose center sits at the anterior tip of the second arch. This central attachment, or mandibulohyoid junction, is comprised of tendon and the ends of four of these muscles, two intermandibularis posterior muscles originating from the first arch and two interhyal muscles from the second arch (Chen & Galloway, 2014). Similarly, tendon develops between the two hyohyal muscles at the posterior of the second arch. At the remaining attachments, tendon connects muscle to cartilage. All of these tendons express the zebrafish *Scx* ortholog, *scxa*, and *xirp2a,* whose expression requires muscle (Chen & Galloway, 2014).

Zebrafish is also an ideal system for genetic analysis. Thus far, there are no *in vivo* models detailing the morphogenetic dynamics of tendon and muscle. Few genes have yet been determined necessary for tendon development, though tendon function is impaired by genetic knockdown of *Scx* or tendon extracellular matrix components (Murchison et al.,

17

2007; Subramanian & Schilling, 2014). In a zebrafish forward genetic screen for craniofacial mutants, the Kimmel lab discovered a novel allele of *cyp26b1* with a craniofacial skeletal phenotype consistent with previously described loss-of- function alleles (Piotrowski et al., 1996; Laue et al., 2008; Spoorendonk et al., 2008). Larvae homozygous for this mutation also fail to close their jaws. *In vivo* analysis of wild-type and *cyp26b1* mutant embryos revealed essential steps in the morphogenesis of the ventral first- and second-arch muscles responsible for jaw movement. We show that neural crest cells require the retinoic-acid-catabolizing enzyme Cyp26b1 to pattern jaw muscle attachments in the first and second pharyngeal arches. We propose that *cyp26b1* expression is required in a mass of non-tendon neural crest that separates adjacent tenoblast populations and promotes tendon aggregation necessary for normal muscle patterning.

RESULTS

Cranial muscles have defective attachments in b1024 mutants.

To investigate the mechanisms responsible for integrating the cranial musculoskeletal system, members of the Kimmel lab at the University of Oregon screened ENU-mutagenized zebrafish lines for musculoskeletal defects. After bringing the *b1024* line to the University of Texas, Mary Swartz found that mutants had defects in ventral first- and second-arch jaw muscles (see Fig 1.2 for schematics of these jaw muscles and associated cartilages and tendons). Though these muscles mostly retained their stereotypic pattern in mutants, muscle fibers often split from the main mass to project and terminate ectopically (Fig 2.1B-C).

Not all of the jaw muscles were equally defective in *b1024* homozygotes. Most commonly, the intermandibularis posterior muscles (Fig 2.1C, arrows, 76.5% of mutants) and/or interhyal muscles (Fig 2.1C, white arrowheads, 64.7% of mutants) sent ectopic
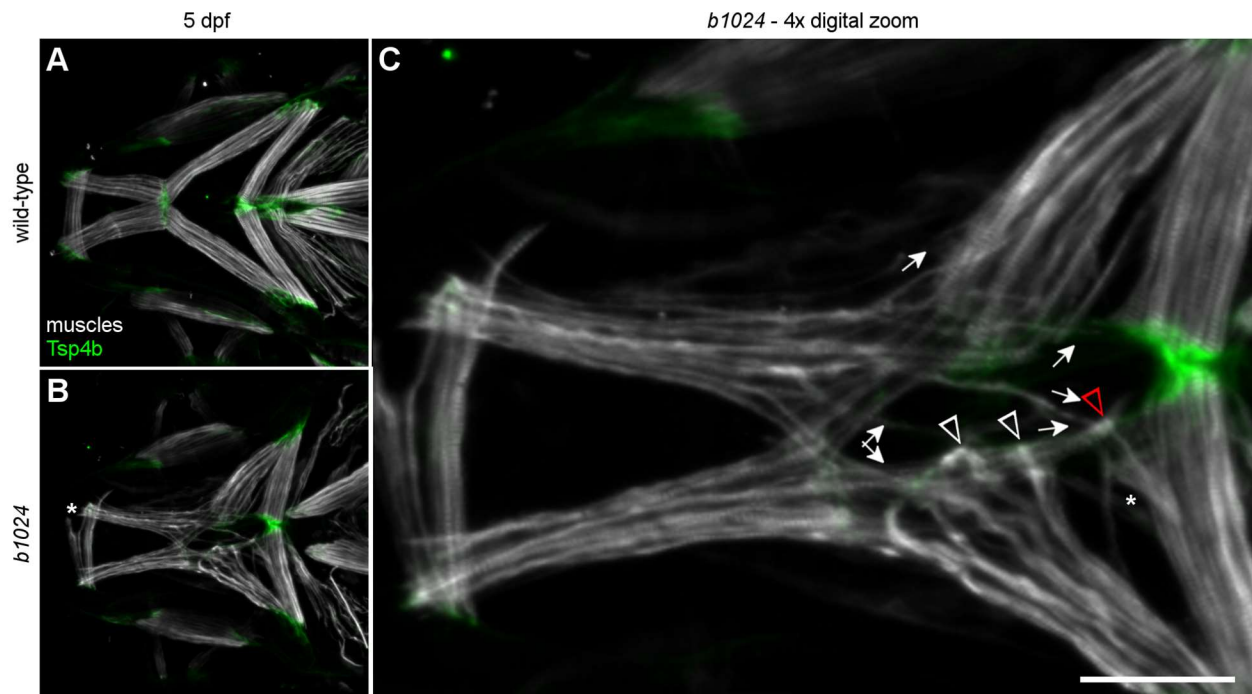
Fig 2.1. Homozygous carriers of *b1024* display second pharyngeal arch lower jaw muscle defects.

(A) Confocal image of jaw muscles in a 5 dpf wild-type larva. (B,C) Jaw muscles project ectopically in *b1024* mutants. Ectopic muscle fibers split off from the intermandibularis anterior muscle in a small percentage of mutants (asterisk in B). (C) 4x digital zoom of confocal image in B. Arrows show ectopic paths of intermandibularis posterior muscle fibers. Arrowheads show ectopic midline attachments of interhyal muscle fibers. Muscle fibers from both intermandibularis posterior muscles cross each other at the midline, extend past the mandibulohyoid junction, and insert alongside the adjacent interhyal muscle or at the hyohyal junction. Fibers from an interhyal muscle (white arrowheads in C) split off to intersect in the midline. A three-way junction (red arrowhead in C) forms between ectopic intermandibularis posterior, interhyal, and hyohyal (asterisk in C) muscle fibers. All images ventral view, anterior to the left. Scale bars = 50 μm.

projections into the medial second pharyngeal arch. Intermandibularis posterior muscle fibers overextended toward the midline and terminated at the junction between hyohyal muscles, on the interhyal, or at other positions on the medial second arch surface (Fig 2.1C, arrows). Muscle fibers split off from the hyohyal muscles (almost always toward other ectopic muscle fibers) in 36.4% of mutants (asterisk in Fig 2.1C). The intermandibularis

anterior muscle fibers split and formed ectopic attachments in less than 20% of mutants (asterisk in Fig 2.1B). The remaining, more dorsal, muscles of the head as well as the trunk musculature appeared normal in *b1024* mutants. These results indicate that the *b1024* mutation disrupts musculoskeletal connections, particularly within the second pharyngeal arch midline.

A novel mutation in *cyp26b1* causes craniofacial musculoskeletal defects in the *b1024* line.

In addition to muscle defects, *b1024* mutants had midline skeletal defects (Fig 2.2A-D). Homozygous embryos had a narrow ethmoid plate and parasphenoid bone (Fig 2.2C). In the posterior neurocranium, mutants displayed a gap in the parachordal cartilages (arrows in Fig 2.2A,C), medial to the ear, and a gap between the anterior and posterior basicranial cartilages (arrowheads in Fig 2.2A,C), lateral to the ear (Fig 2.2C). In the viscerocranium of *b1024* mutants, the ventral cartilage elements of the first and second pharyngeal arches, Meckel's cartilage and ceratohyals, respectively, were fused in the midline (arrowheads and insets in Fig 2.2B,D). The anterior end of the basihyal cartilage was narrow and hypoplastic compared to wild-type larvae, similar to the ethmoid plate defect (arrows in Fig 2.2B,D). Collectively, these results demonstrated that *b1024* mutants have prominent musculoskeletal defects largely localized to the ventral midline.

To identify the genetic lesion in *b1024*, Mary Swartz performed linkage analysis with simple sequence length polymorphisms. She genetically mapped the *b1024* lesion to an interval of chromosome 7 between z8889 and z1239 containing *cyp26b1*. While muscle phenotypes have not been previously analyzed in zebrafish *cyp26b1* mutants, the *b1024* skeletal phenotype fit with previous descriptions of *cyp26b1* mutants (Laue et al., 2008). Mary discovered a single base substitution (C > T) seven bases into exon 6 (Fig 2.2E-G, RefSeq NM_212666). This variant creates an early stop codon (p.Gln384*, RefSeq
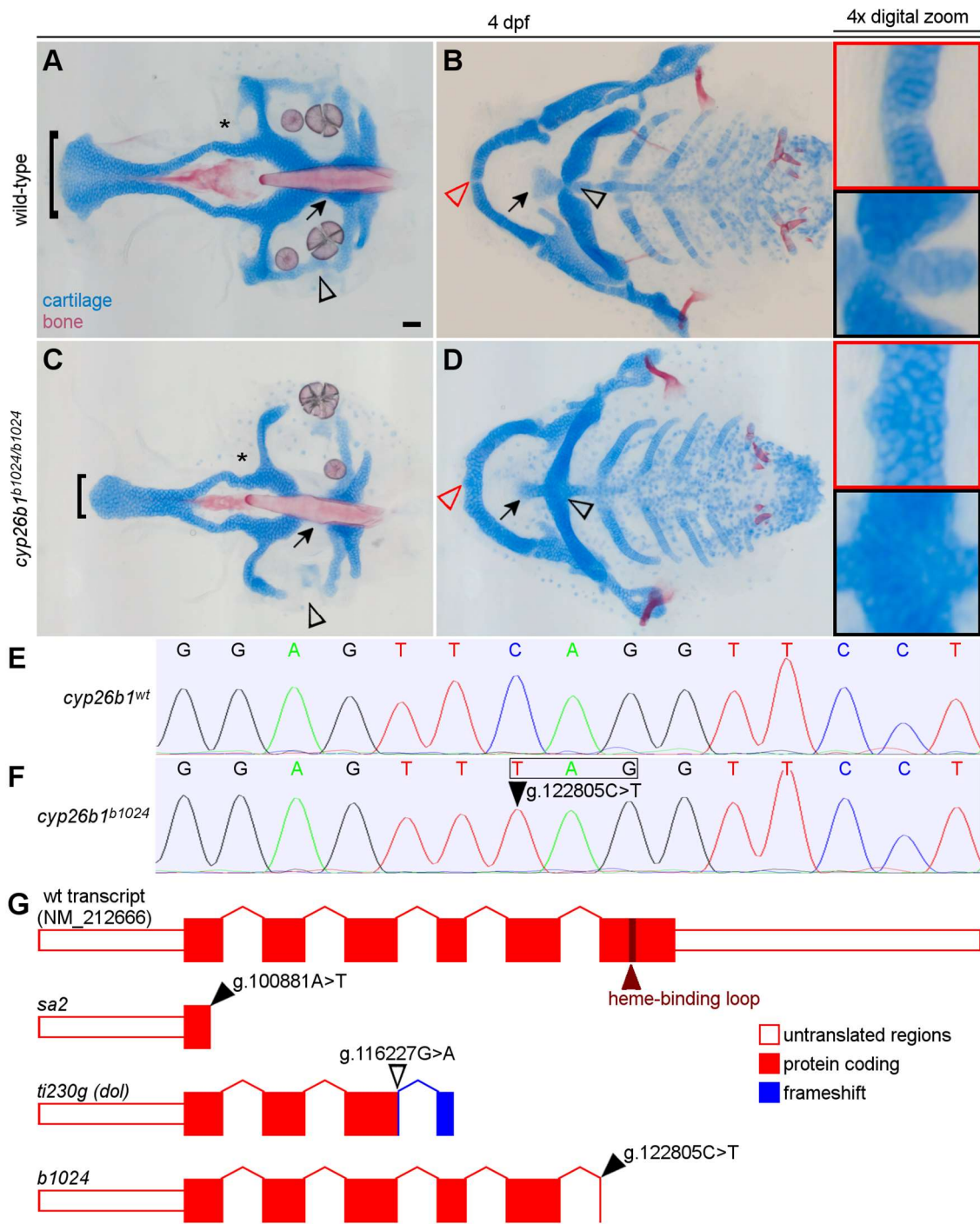
Fig 2.2. The *b1024* lesion is a nonsense mutation in *cyp26b1* that causes skeletal defects.

Fig 2.2. The *b1024* lesion is a nonsense mutation in *cyp26b1* that causes skeletal defects.

(A-B) Wild-type head skeleton stained with Alcian blue and Alizarin Red dyes. (C) In *b1024* mutants the ethmoid plate (brackets in A,C) and parasphenoid bone are narrow and the parachordal (arrows in A,C), basicranial commissure (arrowheads in A,C), and lateral commissure (asterisks in A,C) cartilages of the posterior neurocranium are hypoplastic. (D) The midline symphysis of Meckel's (red arrowheads and red insets in B,D) and ceratohyal (black arrowheads and black insets in B,D) cartilages are absent in *b1024* homozygotes, fusing the left and right cartilage elements, and the basihyal cartilage (arrows in B,D) is hypoplastic. (E-F) A single base substitution in *cyp26b1* (arrowhead) results in an in-frame stop codon (box). (G) *b1024* causes a later truncation of *cyp26b1* than previously characterized alleles, but still eliminates essential components of the cytochrome P450 enzyme domain encoded in exon 6. Panels A-D generated by Mary Swartz. All images ventral view, anterior to the left. Scale bars = 50 µm.

NP_997831) that is predicted to truncate the last 128, of 511 total, amino acids of Cyp26b1. A single cytochrome P450 superfamily domain comprises most of the protein sequence. This lesion in *cyp26b1* truncates 23% of the P450 domain, including the highly conserved heme-binding loop in exon 6, which is required for enzymatic activity. Mary concluded that *b1024* is a recessive loss-of-function, likely null, allele of *cyp26b1* like *sa2* and *ti230g*.

Loss of Cyp26b1 function causes abnormal jaw muscle morphogenesis.

To characterize the genesis of muscle defects in *cyp26b1* mutant zebrafish, I used *in vivo* time-lapse imaging to track muscle morphogenesis. I used a *503unc:mCherry; fli1:EGFP* double transgenic line for live fluorescence imaging of muscles and neural crest, respectively. By 48 hpf, the interhyal and hyohyal muscle masses emerged on the ventrolateral surface of the second pharyngeal arch and elongated toward the midline. Around 51 hpf, two bilateral muscle masses formed on the ventrolateral surface of the first pharyngeal arch. In the next few hours, muscle fibers extended between these masses across the midline to form the intermandibularis anterior muscle. The nascent intermandibularis posterior muscle fibers elongated posteriorly from each mass toward the second pharyngeal arch. By 53 hpf, all of these first and second pharyngeal arch jaw muscles were present in both wild-type and *cyp26b1* mutant embryos (Fig 2.3A,F; only *503unc:mCherry* expression shown for clarity). Thus, Cyp26b1 is dispensable for the initial formation of these muscle masses.

Loss of Cyp26b1 function had a profound effect on the connection of these muscles without altering the initial muscle dynamics. As in wild-type embryos, intermandibularis posterior muscles encountered the second arch neural crest and interhyal muscles around 55 hpf in Cyp26b1-deficient embryos (Fig 2.3B,G). However, the elongation of interhyal muscles toward the midline was retarded by loss of Cyp26b1 function, (S1 Movie, *cyp26b1*[-/-]). Their medial tips were further apart at 55 hpf (compare arrows in Fig 2.3B,G),
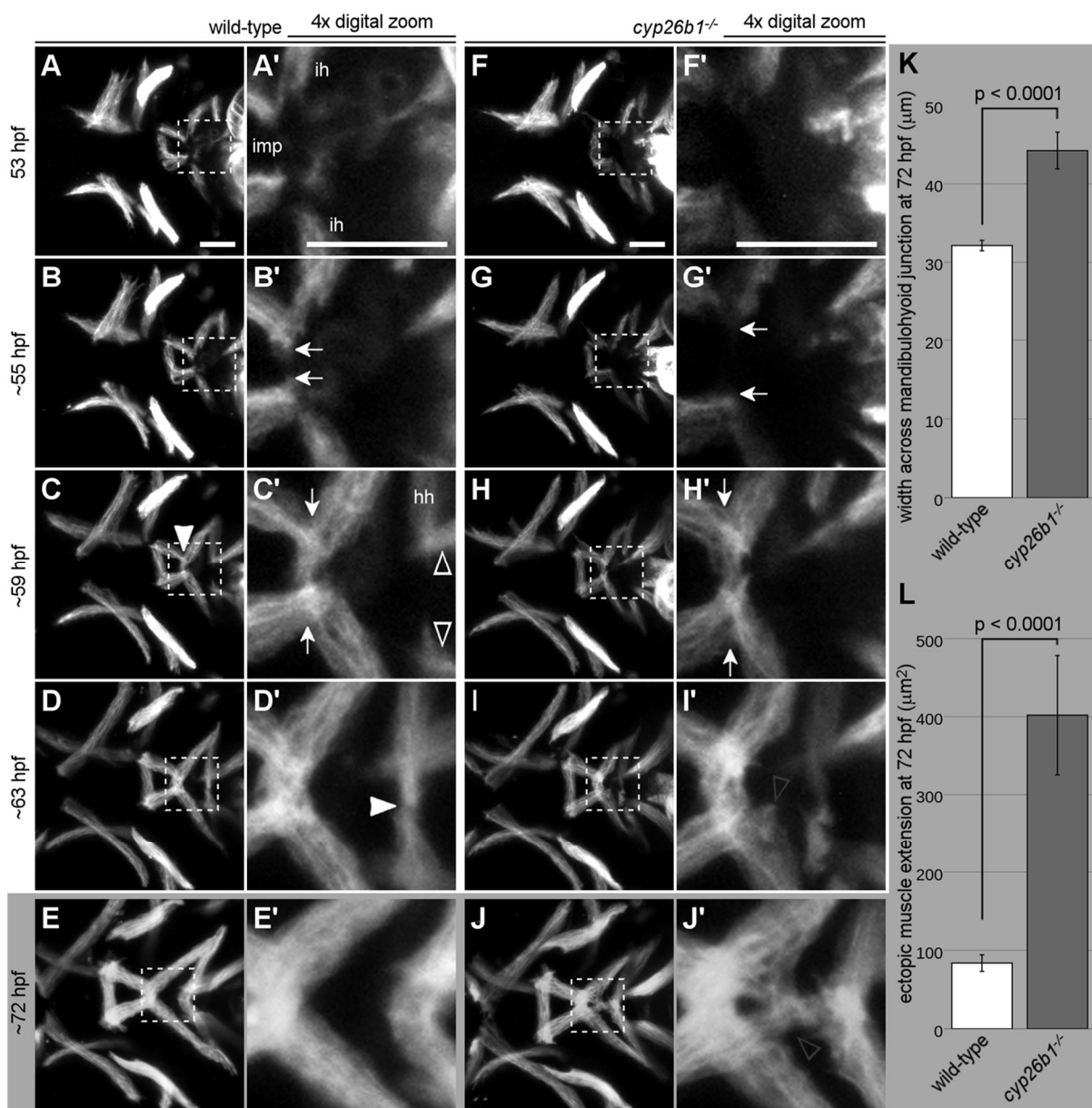
Fig 2.3. Loss of Cyp26b1 function disrupts morphogenesis of the mandibulohyoid junction.

Fig 2.3. Loss of Cyp26b1 function disrupts morphogenesis of the mandibulohyoid junction.

(A-J) Stills from S1 Movie. (A) All first and second pharyngeal arch muscles are present at 53 hpf. (A') Interhyal muscles (ih) extend toward the second pharyngeal arch midline while the intermandibularis posterior muscles (imp) extend from the first arch to the anterior edge of the second arch. (B,B') Around 55 hpf, the tips of intermandibularis posterior and interhyal muscles on each side of the head connect (arrows). (C) Before 60 hpf, the left and right sides meet in the midline, connecting the four muscles of the mandibulohyoid junction (arrowhead). (C') Hyohyal muscles (hh) extend toward the second pharyngeal arch midline while the tips of sternohyoideus muscles pass just dorsally (arrowheads). (D,D') Around 63 hpf, the hyohyal muscles meet end-to-end in the midline (arrowhead). (E,E') At 72 hpf, now broader muscles sit end-to-end at the mandibulohyoid and hyohyal junctions. In *cyp26b1* mutants, the timing of jaw muscle differentiation is normal (F,F'). (G,G') Elongation toward the midline by intermandibularis posterior and interhyal muscles is slower (compare arrows to those in B'). (H,H') The intermandibularis posterior and interhyal muscles connect poorly across the midline (distance between arrows = 46.5 μm, compare between arrows in C' = 38.2 μm). (I,I'J,J') Muscle fibers extend into the space bounded by the interhyal and hyohyal muscles (arrowheads). Both aberrant behaviors have phenotypic readouts at 72 hpf. (K) Where the muscles are narrowest at the junction between intermandibularis posterior and interhyal muscles (see arrows in C',H'), the mandibulohyoid junction is significantly wider in mutants than in wild-type siblings. (L) There is also a significantly larger surface area in which the ends of intermandibularis posterior and interhyal muscles extend ectopically. All images ventral view, anterior to the left. Scale bar = 50 μm.

and these and intermandibularis posterior muscles made no connection at the midline as late as 60 hpf. Meanwhile, intermandibularis posterior muscles continued to elongate posteriorly beyond the anterior edge of the second arch in Cyp26b1-deficient embryos (Fig 2.3I-J). Fibers could be observed extending over the neural crest cells of the medial second arch or overlapping with the medial tip of the interhyal muscle. Subsequently, muscle fibers became more disorganized rather than bundling tightly at the mandibulohyoid junction. Later muscle phenotypes indicated that some muscle fibers at the tips of interhyal muscles fan out or split off as well (see Fig 2.1B-D). These results demonstrated that the muscle cell movements necessary to integrate the ventral pharyngeal musculature, particularly at the mandibulohyoid junction, require Cyp26b1 function.

To quantify mandibulohyoid junction defects, I measured jaw muscles in 72 hpf embryos (Fig 2.3E,J). By this point the left and right intermandibularis posterior and interhyal muscles connected in the midline in nearly all mutants. However, the width across the mandibulohyoid junction (see arrows in Fig 2.3C',J') was 12.1 µm wider in the average mutant (44.2 µm, n = 11) relative to wild-type fish (32.1 µm, n = 49, Fig 2.3K). The bundling of muscle fibers at the mandibulohyoid junction in wild-type embryos meant that muscles rarely overlapped or projected ectopically. In the average *cyp26b1* mutant, overlapping muscles and ectopic muscle fibers covered an area of 401.9 μm^2 (n = 11), compared to 83.7 μm^2 in wild-type fish (n = 49, Fig 2.3L).Together, these findings suggested that Cyp26b1 functions to promote the proper extension of interhyal muscles into the midline and to stop the posterior elongation of intermandibularis posterior muscles at the anterior edge of the second arch.

Cyp26b1 functions in the neural crest to promote craniofacial musculoskeletal patterning.

Because loss of Cyp26b1 disrupts skeletal and muscle morphogenesis, Mary Swartz performed genetic mosaic experiments to test the hypothesis that Cyp26b1 functions in neural crest cells to mediate musculoskeletal development. She transplanted cells from wild-type embryos into *cyp26b1* mutant embryos at the onset of gastrulation (Fig 2.4A). At 24 hours post-fertilization, she selected individuals in which there were large contributions of wild-type crest to the first and second pharyngeal arches (Fig 2.4B). She subsequently evaluated skeletal (neural-crest-derived) and muscular (mesoderm-derived) phenotypes of these individuals at 4 dpf (Fig 2.4C-F).

Even though donated neural crest cells contributed to one side of the head, they were able to affect midline skeletal phenotypes in each of 11 mutant hosts. Almost half of neural crest transplants restored the flared shape of the ethmoid plate (Fig 2.4D, n=5/11) Donated crest efficiently rescued the symphysis of Meckel's (Fig 2.4E, red arrowhead and inset, n=7/11) and ceratohyal (Fig 2.4E, black arrowhead and inset, n=8/11) cartilages completely or partially. These results indicated that neural crest cells require Cyp26b1 function for ethmoid plate, Meckel's cartilage, and ceratohyal development, contrary to a previous report (Laue et al., 2008).

Consistent with our hypothesis, Cyp26b1 function within the neural crest restored proper jaw muscle patterning. Though donated wild-type neural crest cells could completely rescue mutant muscle phenotypes (n=1/6), more frequently individuals displayed unilateral improvement (Fig 2.4F, n=3/6). In these embryos, jaw muscles looked wild-type on the donor side of the head. These results demonstrated that Cyp26b1 function within the neural crest integrates musculoskeletal development in the face.
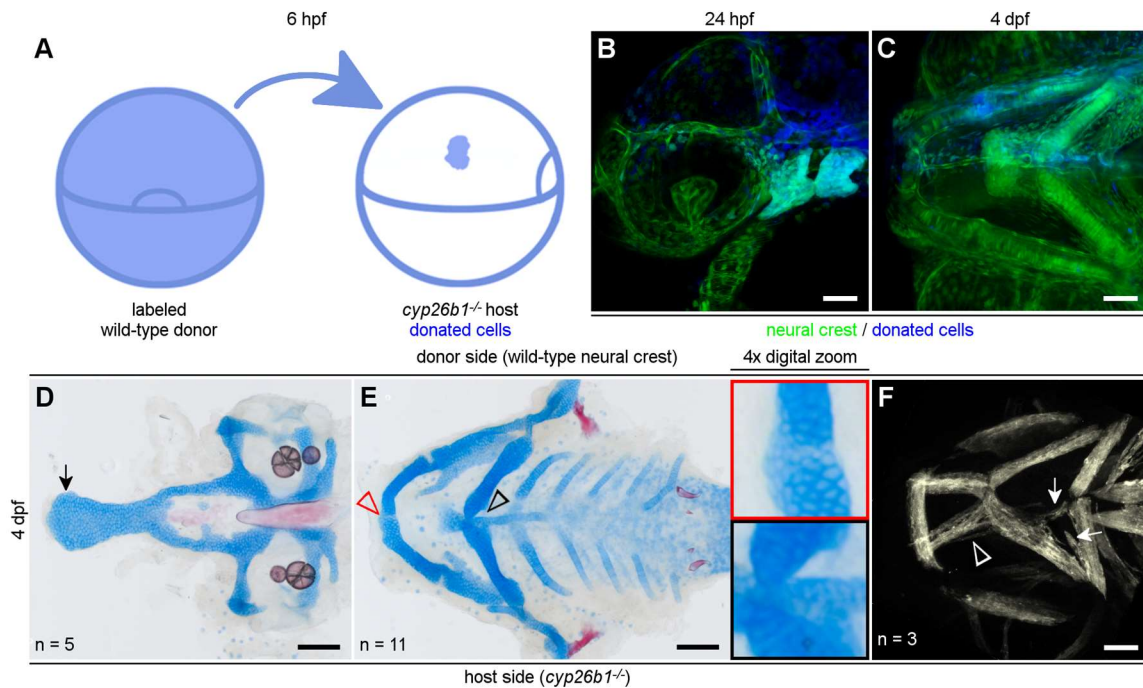
Fig 2.4. Cyp26b1 functions in the neural crest to promote craniofacial musculoskeletal patterning.

(A) Schematic for transplantation of embryonic cells at 6 hpf. The blue area shown on the right side of A will contribute to the neural crest on one side of the head (B, C). (D-E) Donated neural crest cells improve skeletal phenotypes in a *cyp26b1* mutant host. Mutant hosts had a wider palate and a flared ethmoid plate (arrow in D). (E) Depletion of Alcian-positive cartilage matrix indicates partial rescue of Meckel's cartilage fusion (red arrowhead/red inset) and ceratohyal morphogenesis and fusion at the midline is rescued on the donor side (black arrowhead/black inset). (F) On the side of the head without donated neural crest, muscles project ectopically (arrows) and muscle fiber bundles split (arrowhead), compared to normal phenotypes on the donor side. Figure panels generated by Mary Swartz. Lateral view in B. All other images ventral view, anterior to the left. Scale bars = 50 μm.

Loss of Cyp26b1 function disrupts cranial tendon differentiation.

The mandibulohyoid junction forms within a population of neural crest mesenchyme just ventral to the boundary between the ceratohyal cartilage condensations (S2 Fig). I tested the hypothesis that tendons, a neural crest derivative, are defective in *cyp26b1* mutants. In wild-type embryos, Tsp4b-positive tendon tissue formed at all cranial muscle attachment

points by 3 dpf (Fig 2.5A). Tsp4b-rich tendon matrix grew as jaw muscles matured in larvae (Fig 2.5B-C). In contrast, the only visible cranial deposits of Tsp4b-positive matrix in 3-day-old *cyp26b1* mutants labeled the inferior oblique and inferior rectus muscles of the eyes (Fig 2.5D). However, most of the jaw muscle attachment points displayed Tsp4b by 4 dpf in mutants (Fig 2.5E), though the mandibulohyoid junction (arrowheads in Fig 2.5B,E) and sternohyoideus-anchoring tendons (arrows in Fig 2.5B,E) were still difficult to detect with α-Tsp4b. In 5-day-old mutant larvae, all visible jaw muscle attachment points, including ectopic attachments, displayed distinct Tsp4b deposition (Fig 2.5F). These observations suggested that loss of Cyp26b1 function inhibits the initial differentiation of craniofacial tendons and disrupts the later morphology of these tendons.

Patterning of head tendon progenitors requires Cyp26b1.

The reduction of Tsp4b deposition in 3-day-old *cyp26b1* mutant embryos led me to question whether tenoblasts are disturbed during the period of jaw muscle morphogenesis. I observed that *scxa:mCherry*-positive neural crest cells were aggregating along the forming mandibulohyoid junction at 60 hpf (Fig 2.6A, arrow). Four other *scxa*-positive aggregates resided more posteriorly in the medial second pharyngeal arch, two slightly dorsal sternohyoideus tendons (arrows in Fig 2.6A') and two smaller aggregates labeling the medial tips of the hyohyal muscles (arrowheads in Fig 2.6A'). An elongated group of non-chondrogenic, non-tenogenic neural crest cells defined the insertion of the sternohyoideus muscles (S3 Fig, A-D). These brightly *fli1:EGFP*-positive, *scxa:mCherry*-negative cells separated the two ceratohyal cartilage condensations, and also separated anterior, ventral tenoblasts from the posterior, dorsal sternohyoideus tendon aggregates (arrowheads in S3 Fig, B,C). At 4 dpf, *scxa:mCherry* expression labeled all cranial tendons and ligaments (Fig 2.6C). The mandibulohyoid junction (arrow in Fig 2.6C) and a small population of *scxa*-positive mesenchyme on the periphery of the intermandibularis
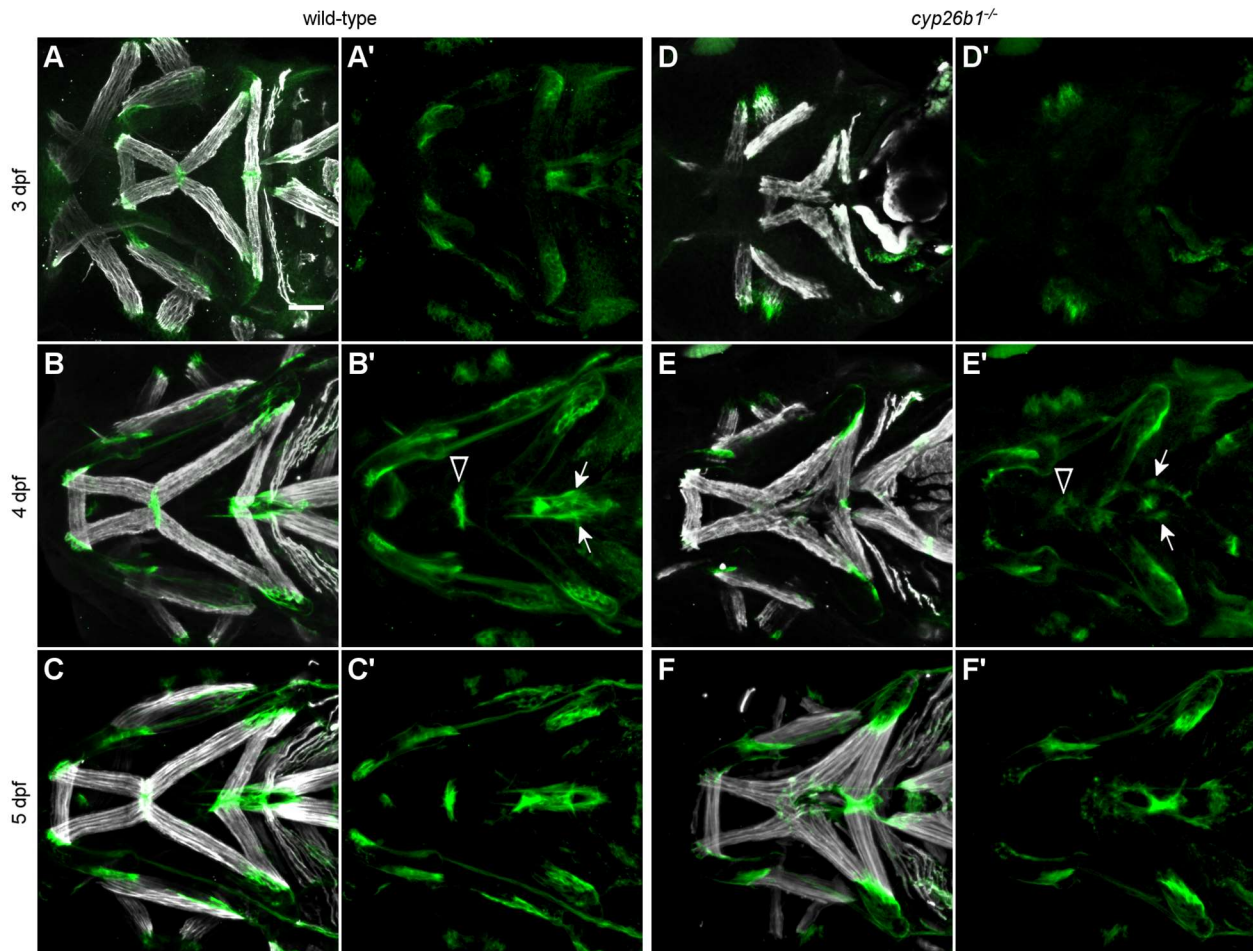
Fig 2.5. Loss of Cyp26b1 function disrupts cranial tendon differentiation.

(A-C) From 3 dpf onward, muscle attachment points are enriched in Tsp4b (green). (D) Where the intermandibularis, interhyal, hyohyal, sternohyoideus, and adductor mandibulae muscles attach to the viscerocranium, *cyp26b1* mutants fail to deposit tendon matrix by 3 dpf. (E) By 4 dpf, mutants display Tsp4b at jaw muscle attachments, though weakly at the mandibulohyoid junction (arrowheads in B',E') and sternohyoideus tendons (arrows in B',E'). (F) At 5 dpf, punctate deposits of Tsp4b can be seen at all ectopic points of jaw muscle attachment. All images ventral view, anterior to the left. Scale bar = 50 µm.

posterior muscles displayed the *scxa* reporter brightly. Expression of the *scxa* reporter was also strong at the hyohyal junction (Fig 2.6C', arrowhead in Fig 2.6C), and two *scxa*-positive spurs extended from between the hyohyal muscles (outlines in Fig 2.6C') to the anterior edge of each interhyal muscle (outlines in Fig 2.6E'). Notably, these *scxa*-positive
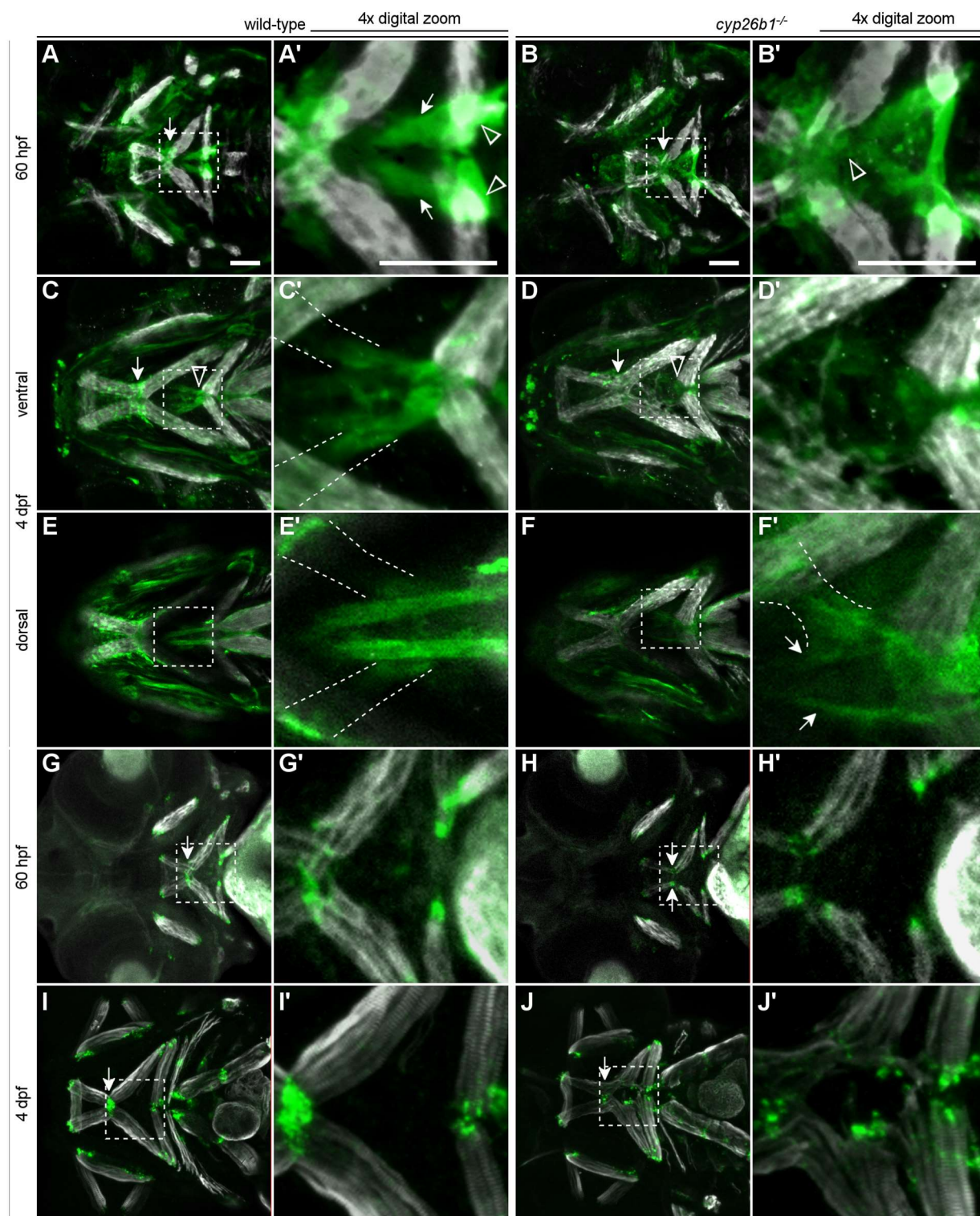
Fig 2.6. Patterning of *scxa*-expressing head tendon progenitors is disrupted in *cyp26b1* mutants.

31

Fig 2.6. Patterning of *scxa*-expressing head tendon progenitors is disrupted in *cyp26b1* mutants.

(A) At 60 hpf, bright puncta of aggregating scxa-positive tenoblasts form in a line between the intermandibularis posterior and interhyal muscles (arrow), and other tenoblasts populate a mass of midline cells from the oral ectoderm to the mandibulohyoid junction. (A') In the second pharyngeal arch, the ends of the hyohyal (arrowheads) and sternohyoideus (arrows) muscles are labeled by *scxa*-positive aggregates. (B) In *cyp26b1* mutants, the pattern of *scxa* expression suggests less aggregation at the mandibulohyoid junction (arrow) and less separation of anterior and posterior tenoblasts. (C) At 4 dpf, the mandibulohyoid (arrow) and hyohyal (arrowhead) junctions express *scxa* strongly, and most muscle attachment points are *scxa*-positive. Notably, two spurs extend from the hyohyal junction (C', outlines) and insert near the anterior edge of each interhyal muscle (E', outlines). (E,E') The sternohyoideus tendons insert in the midline, just dorsal to the hyohyal junction structures. (D) Four-day-old mutants display some *scxa* expression at most muscle attachment points, but expression associated with the intermandibularis muscles (arrow) and the hyohyal junction (arrowhead) is reduced. (D') The spurs of the hyohyal junction are highly dysmorphic (outlines). (F') The sternohyoideus tendons are rod-like but hypoplastic (arrows). (G,I) Cells expressing *xirp2a* are present at the tips of all cranial muscles at 60 hpf and 4 dpf. (H,J) Cells at muscle tips are also labeled with *xirp2a* in *cyp26b1* mutants. All images ventral view, anterior to the left. Scale bars = 50 μm.

spurs displayed little to no Tsp4b deposition in wild-type embryos, though Tsp4b was enriched on every other cranial tendon and ligament (compare Fig 2.5B and Fig 2.6C). Just dorsal to the hyohyal junction structures were the long, rod-shaped sternohyoideus tendons (Fig 2.6E). My findings suggested that tendon components of the mandibulohyoid junction develop from a population of anterior tenoblasts that is segregated, by at least 60 hpf, from those posterior tenoblasts that form the hyohyal junction and the sternohyoideus tendons.

In *cyp26b1* mutants, the population of *scxa*-positive neural crest cells was less organized. In mutants there was less clear separation of anterior and posterior tenoblasts in the second pharyngeal arch (Fig 2.6B, arrowhead in Fig 2.6B'). Though visually it appeared that the population of tenoblasts could be expanded in mutants, cell counts indicated that 60 hpf mutants (148.67 tenoblasts, SD = 2.89, n = 3) did not have significantly more or fewer tenoblasts than siblings (159.67 tenoblasts, SD = 16.4, n = 6). Instead, the *scxa*-positive cells in 60 hpf *cyp26b1* mutants were not concentrated at the mandibulohyoid junction, as they were in wild-type embryos (arrow in Fig 2.6B), though aggregates still formed at the sternohyoideus and hyohyal muscle tips. The structure separating the ceratohyal cartilage condensations was also missing in mutants (S3 Fig, E-H). Bright, non-chondrogenic neural crest cells were among the tenoblasts, but they formed a round shape ventral to the ceratohyals instead of elongating between them (green arrowhead in S3 Fig, F). Four-day-old *cyp26b1* mutants did not maintain strong *scxa* expression at the mandibulohyoid and hyohyal junctions (arrow and arrowhead in Fig 2.6D, respectively). In fact, mutants lacked defined spurs at the hyohyal junction. Instead, a poorly organized population of *scxa*-positive cells stretched between the tips of the hyohyal muscles and the adjacent interhyal muscles, as well as ectopic muscle projections from these (detail in Fig 2.6D'). Four day old mutants had developed sternohyoideus tendons, but these were hypoplastic and slightly curved (arrows in Fig 2.6F'). Loss of *scxa*

expression over time came in contrast to the observation of progressive buildup of Tsp4b-positive extracellular matrix at muscle attachment points (Fig 2.5). I therefore used *in situ* hybridization to visualize expression of a muscle tip marker, *xirp2a,* which labels all muscle attachments (Chen & Galloway, 2014). At both 60 hpf and 4 dpf, *xirp2a* labeled cells at the tips of cranial muscle fibers (Fig 2.6G,I). This remained true in *cyp26b1* mutant embryos, in which *xirp2a* labeled muscle fiber tips even at ectopic points of attachment (Fig 2.6H,J). Notably, in 60 hpf mutants there was no expression of *xirp2a* in the midline where the mandibulohyoid junction should be forming. Together, these findings suggested that the *cyp26b1* mutant muscle phenotype is coincident with a defective pattern of *scxa*-positive tenoblasts, but mutants showed tendon matrix deposition over time because muscles and *xirp2a*-positive cells still promote tendon differentiation.

Loss of Cyp26b1 function perturbs the morphogenesis of *scxa*-positive tendon progenitors.

Based on the disruption of mature tendon morphology in *cyp26b1* mutants, I hypothesized that loss of Cyp26b1 function disrupts tenoblast behaviors. Live imaging in *503unc:EGFP);scxa:mCherry* double transgenic control embryos revealed dynamic populations of tenoblasts in the ventral midline of the jaw (S4 Movie, control; Fig 2.7A-F). Tenoblasts surrounded the forming intermandibularis muscle fibers near the midline of the first pharyngeal arch around 51 hpf (arrows in Fig 2.7A'). Tenoblasts migrated toward the midline as the intermandibularis posterior muscles elongated (Fig 2.7B). Coincident with this migration, the intermandibularis posterior muscles angled toward the midline by 54 hpf, almost perpendicular to the intermandibularis muscle masses at 51 hpf (compare S5 Fig, A and B). By this time, two other tenoblast masses at the tips of the hyohyal and sternohyoideus muscles became visible (arrowheads in Fig 2.7B'). By 57 hpf, the movement of tenoblasts separated a bright aggregate at the anterior tip of the second
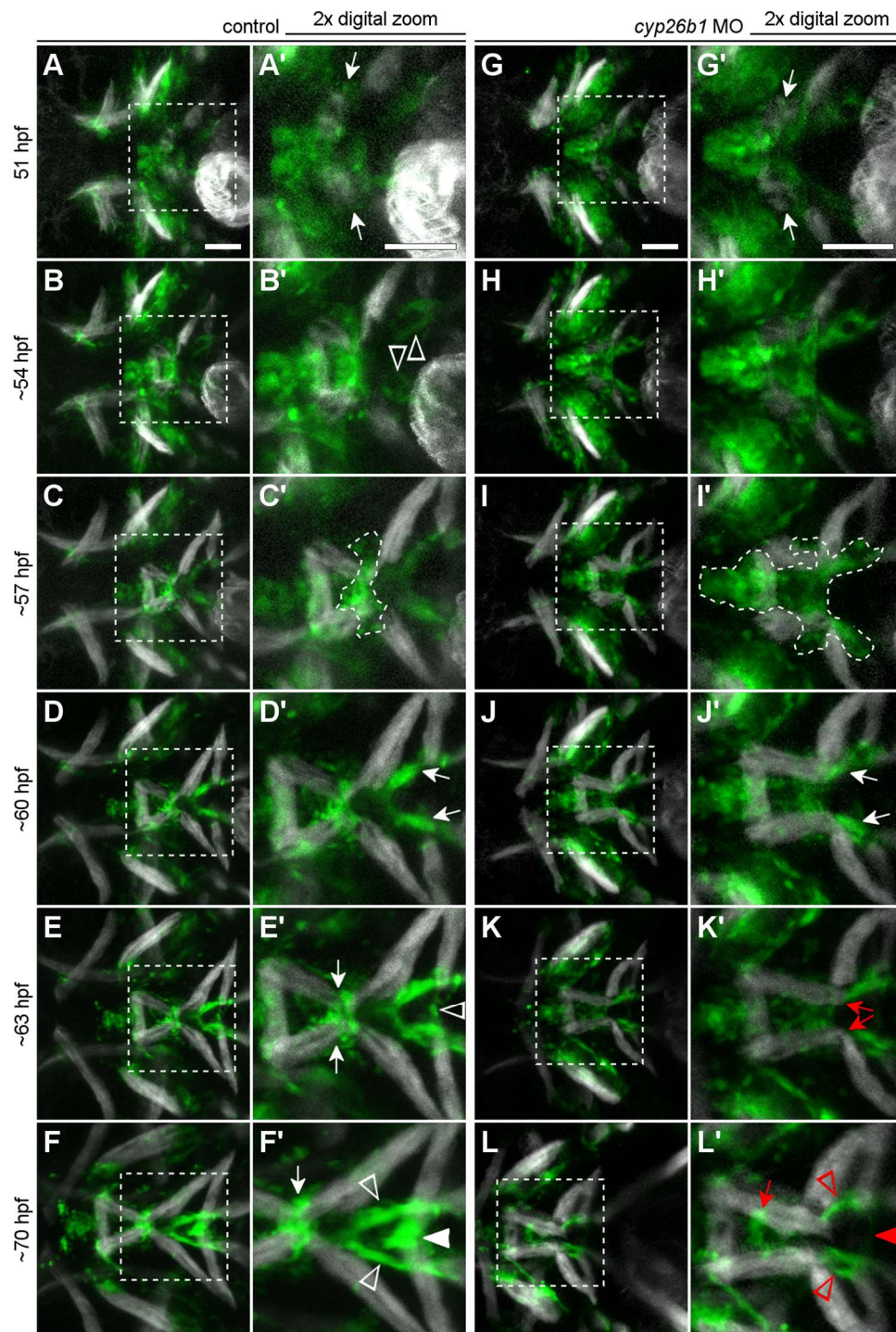
Fig 2.7. Loss of Cyp26b1 function disrupts tendon morphogenesis in the second pharyngeal arch midline.

Fig 2.7. Loss of Cyp26b1 function disrupts tendon morphogenesis in the second pharyngeal arch midline.

(A-L) Stills from S4 Movie. (A,A') The emerging intermandibularis muscles (arrows) are surrounded by tenoblasts at 51 hpf (B,B'). As tenoblasts converge on the midline, intermandibularis muscles point their posterior tips medially (see S5 Fig, B). Meanwhile, tenoblast masses (arrowheads) associated with the hyohyal and sternohyoideus muscles on each side have become visible. (C,C') By 57 hpf, aggregating cells are visible at the mandibulohyoid junction (outline), where *scxa*-positive cells have segregated from the more anterior group. (D,D') By 60 hpf, aggregation of the sternohyoideus tendons is apparent (arrows). (E,E') By 63 hpf, tenoblasts attached to each hyohyal muscle connect in the midline and an aggregate forms (arrowhead). (F,F') Approaching 3 dpf, tenoblasts at the mandibulohyoid (arrow) and hyohyal (white arrowhead) junctions are highly compact between the muscle tips, and sternohyoideus tendons are elongated (open arrowheads). (G,G") In morpholino-injected embryos, intermandibularis muscles (arrows) emerge normally among a field of tenoblasts around 51 hpf. (H,H') Tenoblasts migrate slowly toward the midline, and intermandibularis posterior muscles point posteriorly at 54 hpf (see S5 Fig, D). (I,I') Tenoblasts at the mandibulohyoid junction are neither aggregating nor separating from adjacent tenoblasts at 57 hpf (outline). (J,J') Despite surrounding abnormalities, sternohyoideus tendons initiate aggregation around 60 hpf (arrows). (K,K') Connections between intermandibularis posterior muscles and sternohyoideus tendon aggregates persist, and at 63 hpf the muscle tips (red arrows) now extend past mandibulohyoid tenoblasts. (L) Compared to wild-types, the sternohyoideus tendons have poorly aggregated and elongated by 70 hpf, and both intermandibularis posterior and both hyohyal muscles connect to either end of these tendon rudiments. (L') Mesenchymal tenoblasts sit between the intermandibularis posterior muscles (arrow) and still no aggregate has formed connecting the hyohyal muscles (red arrowhead). All images ventral view, anterior to the left. Scale bar = 50 µm.

pharyngeal arch (outline in Fig 2.7C'). Over the next few hours, this rudimentary mandibulohyoid junction took shape as tenoblasts continued to migrate and aggregate in the midline, and the intermandibularis posterior and interhyal muscles elongated toward the aggregate. Around 60 hpf, two more bright aggregates formed from tenoblasts extending from the tips of the sternohyoideus muscles (arrows in Fig 2.7D'). One final aggregate, connecting the hyohyal muscles, formed after tenoblasts extended from each muscle tip to make contact in the midline around 63 hpf (arrowhead in Fig 2.7E'). By 70 hpf, the sternohyoideus tendons and tendon elements of the mandibulohyoid and hyohyal junctions appeared bright and compact (Fig 2.7F). Tenoblast aggregation thus precedes the morphogenesis of the mandibulohyoid and hyohyal junctions.

In *cyp26b1* morpholino-injected embryos, clear tenoblast defects presaged the ectopic elongation of intermandibularis posterior muscles (S4 Movie, *cyp26b1* MO; Fig 2.7G-L). Loss of Cyp26b1 function did not appear to interfere with tenoblast specification, as tenoblasts surrounded the newly forming intermandibularis muscle fibers at 51 hpf (arrows in Fig 2.7G'). However, these tenoblasts failed to migrate into the midline and aggregate there (Fig 2.7H-L). At 54 hpf, the intermandibularis posterior muscles were not oriented toward the midline but rather toward the posterior or even slightly laterally (arrows in S5 Fig, D). Over time, muscle angle did turn toward the midline, but no tenoblast aggregate was apparent at the mandibulohyoid junction at 57 hpf (Fig 2.7I). Instead, a contiguous, diffuse mass of tenoblasts stretched from the oral ectoderm to the sternohyoideus muscles (outline in Fig 2.7I'). Aggregation of the sternohyoideus tendons initiated normally around 60 hpf (arrows in Fig 2.7J'), and the intermandibularis posterior muscles extended beyond the prospective mandibulohyoid junction to abut these aggregates. By 63 hpf, the sternohyoideus tendon aggregates were the only tenoblasts touching the tips of the intermandibularis posterior muscles (arrows in Fig 2.7K'). Loss of

37

Cyp26b1 inhibited tenoblast aggregation between the hyohyal muscles, and blocked their connection even out to 3 dpf (arrowhead in Fig 2.7L'). No tendon aggregate was visible between the tips of the intermandibularis posterior and interhyal muscles, but mesenchymal tenoblasts sat in the midline between the intermandibularis posterior muscles (arrow in Fig 2.7L'). These results indicated that Cyp26b1 function is required for the precise morphogenesis of several cartilage, tendon, and muscle elements that all occupy the ventral midline of the second pharyngeal arch.

Cyp26b1 function before 60 hpf is necessary and sufficient for mandibulohyoid junction morphogenesis.

Because jaw muscle morphogenesis occurs over several hours, I sought to ascertain when Cyp26b1 function is required for this process. I treated *scxa:mCherry* embryos with media containing talarozole, which inhibits Cyp26b1 and its orthologs Cyp26a1 and Cyp26c1, then assessed tendon and muscle morphology at 4 dpf. Compared to controls (Fig 2.8A), embryos treated with talarozole during the period of tenoblast migration to and aggregation at the mandibulohyoid junction (54-60 hpf) had severe phenotypes (Fig 2.8B). Of 64 individuals, 100% displayed ectopic muscle projections across the second pharyngeal arch, including 58.3% in which the intermandibularis posterior muscles extended past the location of the mandibulohyoid junction. Treatment during the period of sternohyoideus tendon aggregation and hyohyal junction formation (60-72 hpf) affected muscle morphology weakly (Fig 2.8C). Of 26 individuals, 76.9% displayed mild ectopic muscle projections from the interhyal or hyohyal muscles, but 25 of 26 developed a normal mandibulohyoid junction. These data show that for proper musculoskeletal patterning, Cyp26b1 function is required when mandibulohyoid tenoblasts are migrating and aggregating.
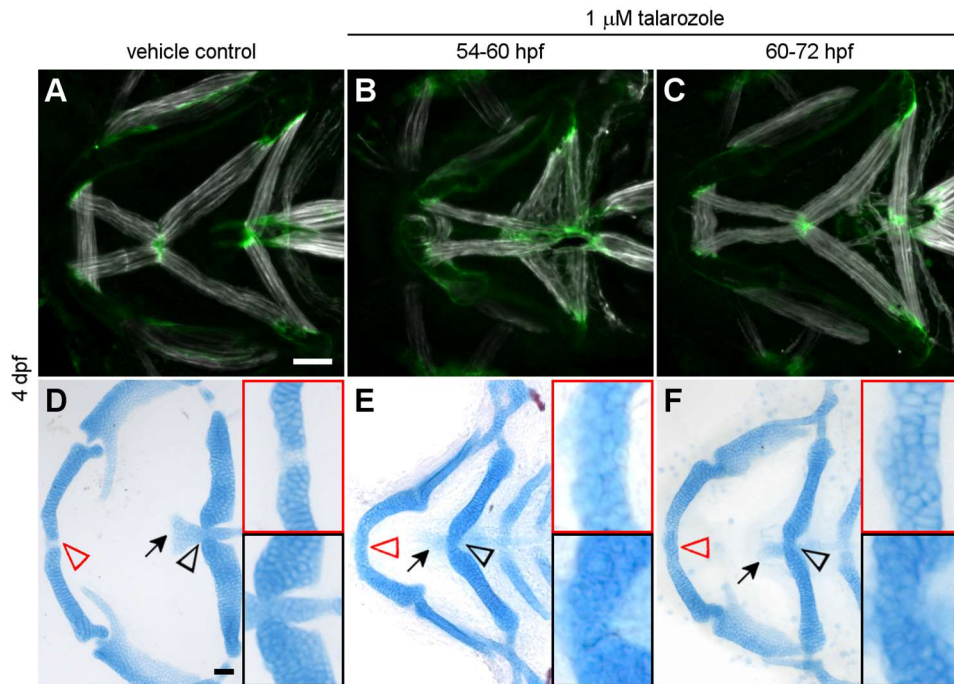
Fig 2.8. Cyp26b1 function before 60 hpf is necessary and sufficient for mandibulohyoid junction formation.

(A,D) Embryos raised in media containing 0.01% DMSO develop normally. (B) 100% of embryos treated with talarozole between 54 and 60 hpf displayed ectopic muscle projection in the second pharyngeal arch, and the intermandibularis posterior muscles overextended in 58.3% of embryos (n = 64). (C) Among embryos treated from 60 to 72 hpf, 76.9% had mild ectopic muscle projections, but 96.2% developed normal tendon and end-to-end muscle connections at the mandibulohyoid junction (n = 26). (E,F) Talarozole-treated embryos recapitulated *cyp26b1* mutant skeletal phenotypes of Meckel's cartilage fusion (red arrowheads and insets), ceratohyal fusion (black arrowheads and insets), and basihyal reduction (arrows). All images ventral view, anterior to the left. Scale bar = 50 µm.

We also evaluated the effects of Cyp26 inhibition on the cranial skeleton at 4 dpf. Control larvae had normal phenotypes (Fig 2.8D). Talarozole treatment from 54-60 hpf recapitulated much of the jaw skeletal phenotype of *cyp26b1* mutants. These embryos consistently displayed fused Meckel's (Fig 2.8E, red arrowhead and inset) and ceratohyal cartilages (Fig 2.8E, black arrowhead and inset), and loss or severe hypoplasia of the basihyal cartilage (Fig 2.8E, arrow). Embryos treated from 60-72 hpf also had midline

cartilage fusions and basihyal cartilage hypoplasia (Fig 2.8F). These findings suggested that Cyp26b1 continues to function in cartilage development after the mandibulohyoid junction forms, and that the mandibulohyoid junction can form normally despite deformity of the basihyal and ceratohyal cartilages.

Cyp26b1 expression separates anterior and posterior muscle attachments in the midline of the second pharyngeal arch.

Cyp26b1 function is required in the neural crest during tenoblast migration and aggregation. I turned to *in situ* hybridization to determine if there were neural crest subpopulations expressing *cyp26b1* during this period. Intriguingly, at 54 hpf, a band of *cyp26b1*-positive cells filled the midline of the posterior second arch between the posterior edge of the mandibulohyoid tenoblast mass (Fig 2.9B, arrow) and the sternohyoid tenoblast masses (Fig 2.9B, arrowheads). There were few, if any, *cyp26b1*-expressing tenoblasts. The mandibulohyoid tenoblasts appeared entirely *cyp26b1*-negative. At 60 hpf, the gap between aggregating mandibulohyoid tenoblasts (Fig 2.9F, arrow) and sternohyoid tenoblasts (Fig 2.9F, arrowheads) had grown. Strong *cyp26b1* expression labeled bilateral cell masses in the posterior second arch (Fig 2.9G, arrows). Again, there appeared to be little overlap between the *scxa* transgenic and *cyp26b1* expression. Together, our results indicated that Cyp26b1 functions in a mesenchymal population of neural crest cells to promote separation and aggregation of tendons in the ventral midline of the second pharyngeal arch.

DISCUSSION

Here we provide *in vivo* characterization of developing musculoskeletal attachments in a living embryo and demonstrate the involvement of retinoic acid in this process. We found that Cyp26b1 function is essential for ventral cranial muscle patterning,
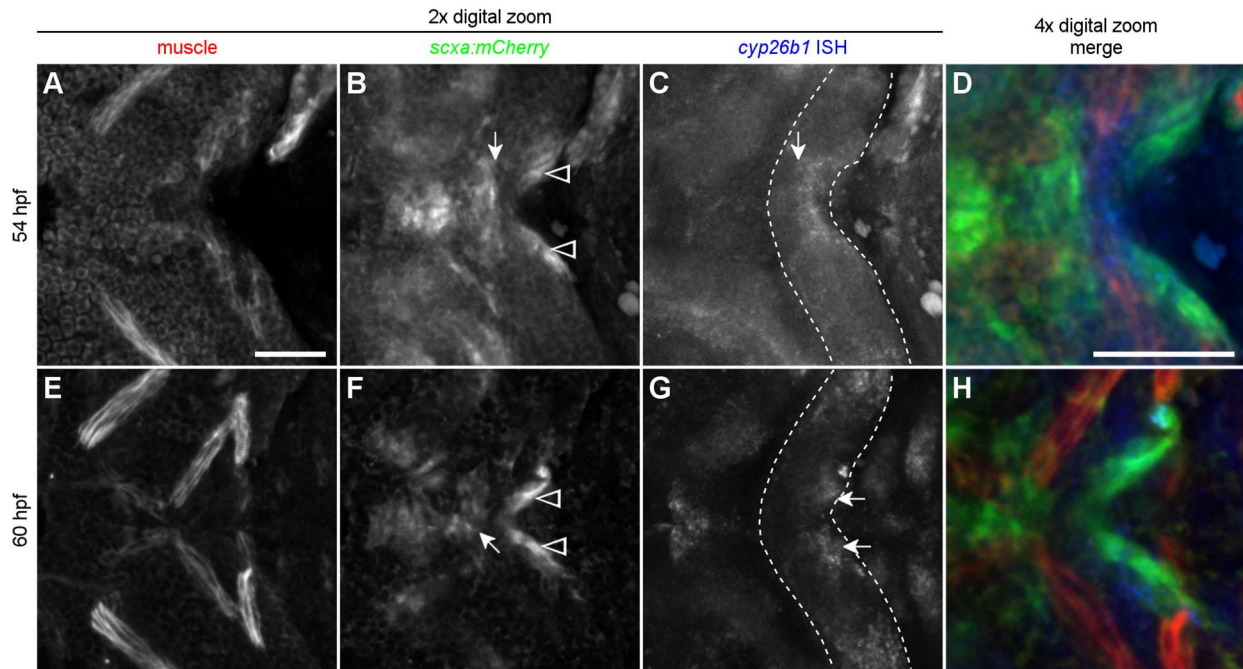
Fig 2.9. Cells expressing *cyp26b1* separate anterior and posterior tenoblast masses between 54 and 60 hpf.

(A-D) 54 hpf zebrafish embryo. (B) Tenoblasts of the future mandibulohyoid junction (arrow) display a flattened posterior boundary and maintain a small distance from tenoblasts lining the posterior edge of the second pharyngeal arch (arrowheads). (C) In the second arch (outline), cells expressing *cyp26b1* line the posterior edge of the arch midline and display a flattened anterior boundary abutting the mandibulohyoid tenoblasts (arrow). (D) Overlay, *cyp26b1* expression labels non-tenoblast cells separating tenoblasts that are mostly or entirely c*yp26b1*-negative. (E-H) 60 hpf zebrafish embryo. (F) Tenoblasts are more compact at the forming mandibulohyoid junction (arrow), and further separated from the sternohyoideus tendon condensations (arrowhead). (G) Strong *cyp26b1* expression now labels bilateral cell masses of the posterior edge of the second arch (arrows). (H) Overlay, Non-tenoblast cells expressing *cyp26b1* sit anterior to the sternohyoideus tendon condensations on each side, facing the nearby interhyal muscles. All images ventral view, anterior to the left. Scale bar = 50 μm.

but not specification. Genetic mosaic analysis demonstrated that neural crest cells required Cyp26b1 function for proper muscle patterning. Time-lapse confocal imaging showed that neural crest derived tenoblasts migrate to and aggregate in regions that predict muscle attachments. In *cyp26b1* mutants, tenoblasts are generated, yet fail to migrate to or aggregate appropriately at those muscle attachments that are disrupted. Furthermore,

41

crucial events of tendon and muscle morphogenesis appear to be orchestrated by non-tenoblast neural crest cells expressing *cyp26b1*.

Cranial muscle and tendon morphogenesis are concerted processes that require Cyp26b1 function.

Our work is consistent with that of others showing that neural crest have critical influences on muscle morphology in the face. Transplanted neural crest cells confer their axial (Noden, 1983b) and species-specific (Tokita & Schneider, 2009) identities upon cranial muscles. In both mouse and zebrafish, and presumably all vertebrates, neural crest gives rise to Scx-positive cranial tendons and ligaments (Grenier et al., 2009; Chen & Galloway, 2014), unlike other regions of the body where tendon, bone and muscle all derive from mesoderm populations (Chevallier et al., 1977; Brent et al., 2003). Our work with that of others demonstrates that, regardless of the source of the progenitors, tendons are critical for muscular patterning. In the chick, ablation of tendon primordia results in mispatterning of the limb musculature (Kardon, 1998). I find in *cyp26b1* mutants similar overextension of specific muscles associated with highly disrupted tendons. This work demonstrates that the presence of tenoblasts is not sufficient for muscular patterning and that aggregation appears to be a critical step in musculoskeletal integration.

In no system are the cell dynamics underlying musculoskeletal integration understood. Time-lapse images show that tendon progenitors migrate to muscle attachment points closely followed by the muscle fibers themselves. This behavior suggests that muscle patterning by neural crest (Noden, 1983b; Tokita & Schneider, 2009) is due to short range cues from tenoblasts. This conclusion is consistent with the finding that the mouse flexor digitorum superficialis muscles migrate from the autopod to the forelimb and require their associated tendons for this translocation (A. H. Huang et al., 2013). Tendon progenitors form in *cyp26b1* mutants but fail to aggregate appropriately – their migration

42

toward the ventral midline is disrupted. Muscle fibers, still in close contact with tenoblasts, fail to elongate toward the midline appropriately without tenoblast aggregation. Our results suggest that tenoblast aggregation controls the localization of muscle guidance cues or enhances their potency.

The nature of the cues that tenoblasts provide are unknown. However, in chicken, Wnt11 orients muscle fibers in the developing somites (Gros et al., 2009). It will be of great interest to determine if cranial tenoblasts utilize Wnt signaling for aggregation and/or directing muscle elongation. Several signals have been discovered as well that facilitate the organization of *Drosophila* muscles and tendons (reviewed in Schweitzer et al., 2010). The zebrafish genome contains a diverse array of homologs to these *Drosophila* genes. Given the dearth of knowledge of this process, the *cyp26b1* mutant is likely to provide key insights into musculoskeletal integration.

Though tenoblasts are present in *cyp26b1* mutants, they are delayed in maturation throughout the face. This is the case even in regions with fairly normal muscle morphology, indicating this defect is independent of the migration/aggregation deficit. Mutants initiate *scxa* expression normally and the specification of muscle appears to follow a normal developmental trajectory. This suggests that mutants do not suffer a general musculoskeletal developmental delay but, rather, that RA modulates the differentiation of tendons downstream of initial *scxa* expression. Muscle contractions maintain Fgf and Tgfβ signaling necessary for maturation of tendon tissue (Havis et al., 2016). Recovery of tendon differentiation in *cyp26b1* mutant larvae might indicate that excess RA competes with contractile forces. It will be of great interest to explore whether RA signaling modulates Fgf or Tgfβ signaling in the zebrafish head or if RA directly regulates Fgf and Tgfβ target genes in tenoblasts.

Retinoic acid regulates musculoskeletal integration of the second pharyngeal arch.

Our results imply a specific role for Cyp26b1 in mandibulohyoid junction formation that is not autonomous to tenoblasts or muscles. Expression of retinoic acid pathway components is highly dynamic, and for much of early pharyngeal arch development *cyp26b1* expression is broad (Laue et al., 2008). However, at 54-60 hpf, when Cyp26b1 function is required for tenoblast aggregation, c*yp26b1* expression in the second pharyngeal arch is restricted predominantly to non-tenogenic midline neural crest cells. This differs greatly from RA dynamics reported in chick limbs, where a Cyp26 enzyme is expressed in tendons themselves at the interface with muscle and RA promotes apoptosis of ectopic muscle fibers between adjacent tendons (Rodriguez-Guzman et al., 2007). It remains to be determined if this is a general difference between the role of RA signaling in musculoskeletal development of the head versus the limb.

Emerging work suggests that non-muscle tissues promote the differentiation of specific tendons. For example, a recent study showed that cartilage is required for tendon development in mouse limb digits (A. H. Huang et al., 2015). Our work strongly suggests that a novel population of neural crest modulates musculoskeletal integration in the second pharyngeal arch. I have found that neural crest cells expressing *cyp26b1* occupy a region that excludes *scxa*-positive tenoblasts. Cyp26b1 expression in this region of exclusion also defines regions of tendon aggregation anteriorly and posteriorly. In live imaging of *cyp26b1* depleted embryos, tenoblasts are spread widely around the mandibulohyoid junction and across this exclusion region. The intermandibularis posterior and interhyal muscles elongate and connect within this field, but neither these muscles nor their connection drives the tenoblasts to aggregate. Instead, mandibulohyoid tenoblasts remain a loose mesenchymal population, and the tips of the intermandibularis posterior and interhyal muscles migrate toward the sternohyoideus tendon aggregates. Furthermore, the

44

sternohyoideus tendons fail to aggregate at the same rate as in controls despite contact with supernumerary muscle tips. Thus, *cyp26b1* mutants have cranial muscles that maintain tenoblast populations, but these muscles are not sufficient for proper tendon aggregation and differentiation. We propose that signals from these neural crest cells are critical for promoting specific sites of tendon aggregates, which then organize the musculature.

How might excess RA inhibit cranial tendon aggregation? Much work is needed to answer this question, partly because the function of Cyp26b1 is likely to range beyond the cells in which it is expressed. The changes in local RA signaling resulting from loss of Cyp26b1 function could alter gene expression in some or all of the adjacent populations of neural crest cells and tenoblasts, and directly or indirectly affect tenoblast aggregation. Additionally, RA has also been shown to directly promote neurite outgrowth (Dmetrichuk et al., 2005, 2006) and act as a neurite chemoattractant, independent of its transcriptional regulatory mechanisms (Farrar et al., 2009). This suggests an alternative model in which focally elevated RA directly disrupts cell movements required for aggregation. While future research is required to distinguish these possibilities, inhibiting Cyp26 function precisely during the 6-hour time window when the mandibulohyoid junction is aggregating disrupts this morphogenesis, arguing for more direct regulation by RA.

The ability to, at least partly, disentangle the effects of Cyp26 on muscle versus skeletal morphogenesis is surprising. We hypothesized that abnormal formation of the mandibulohyoid junction subsequently limited muscle contractile forces necessary for joint formation (Kahn et al., 2009) and skeletal morphogenesis (Rot-Nikcevic et al., 2006; Hinits et al., 2011; Sharir et al., 2011; Shwartz et al., 2012). The basihyal and ceratohyal cartilages form immediately adjacent to the mandibulohyoid junction. Our hypothesis suggested that the reduction of the basihyal cartilage and the midline fusion of the ceratohyal cartilage in *cyp26b1* mutants could both be explained by the muscle phenotype. However, blocking

45

Cyp26 function from 60-72 hpf recapitulates the skeletal phenotypes while leaving the mandibulohyoid junction largely intact. We cannot exclude the possibility that the mandibulohyoid junction is necessary for proper joint and basihyal morphogenesis, but my results strongly suggest that Cyp26b1 has a separate function to promote joint and cartilage fates in the midline.

Our work also combines with other Cyp26b1 loss-of-function models to suggest an evolutionarily conserved RA signaling mechanism for patterning facial muscle. In mouse, palate and tongue development requires *Cyp26b1* (Okano et al., 2012). Loss of Cyp26b1 function increases RA signals to tongue muscles in the ventral midline and activates RA-responsive gene expression ectopically in neural crest cells surrounding these muscles. Besides the glossal muscle, all of the dysmorphic muscles are part of the anterior suprahyoid. Like the affected muscles in zebrafish *cyp26b1* mutants, they form a complex of first- and second-arch branchiomeric musculature that attaches to the anterior hyoid skeleton. The diverse morphologies of these mouse and fish muscles belie their similar origins and likely molecular conservation. Differential expression of *Fgf10* in the tongues of *Cyp26b1* mutant mouse embryos suggests specific Fgf signaling components that could be investigated in zebrafish *cyp26b1* mutants.

Collectively, our work provides novel inroads regarding the genetic underpinnings of craniofacial muscle patterning by neural crest cells. Furthermore, we detailed the dynamic morphogenesis of craniofacial muscles and tendons. Cyp26b1 promotes the morphogenesis of apparently several different neural crest derived structures in the second pharyngeal arch that serve to organize the muscles responsible for the operation of the zebrafish lower jaw. Our findings are illustrative of the interactivity and fine patterning of tissues in the vertebrate head.

METHODS

Danio rerio (zebrafish) care and husbandry

Zebrafish stocks were maintained and embryos were raised according to established protocols (Westerfield, 2007) with approval from the University of Texas at Austin Institutional Animal Care and Use Committee. The *cyp26b1^{b1024}* allele was recovered from a forward genetic screen at the University of Oregon. The *Tg(fli1:EGFP)y1* (Lawson & Weinstein, 2002) and *Tg(-0.5unc45b:EGFP)* (Berger & Currie, 2013) transgenic lines are referred to as *fli1:EGFP* and *503unc:EGFP*, respectively, throughout the text. The 503unc promoter fragment (Berger and Currie, 2013) and Tol2Kit materials and protocols (Kwan et al., 2007) were used to construct the *Tg(-0.5unc45b:mCherry)* transgenic line, referred to as *503unc:mCherry*. The *Tg(scxa:mCherry)* line (J. W. Chen et al., 2017) is referred to as *scxa:mCherry* throughout the text. For pharmacological treatments, embryos were bathed in embryo medium with 0.01% DMSO (Thermo Fisher Scientific, Waltham, MA, USA) as vehicle control or 1 µM talarozole (HY-14531, MedChem Express, Monmouth Junction, NJ, USA).

In situ hybridization and immunohistochemistry

Probes for *cyp26b1* (Laue et al., 2008) and *xirp2a* (cb1045; (Thisse et al., 2001) are described elsewhere. Color development for fluorescence imaging was performed with α-Digoxigenin-POD F$_{ab}$ fragments (11207733910, Roche Diagnostics, Indianapolis, IN, USA) and TSA Plus Fluorescein (NEL741001KT) or Cy3 (NEL744001KT, Perkin-Elmer, Inc., Waltham, MA, USA) System. In preparation for immunohistochemistry, embryos were fixed in 95% methanol/5% glacial acetic acid. The protocol for myosin heavy chain/Tsp4b staining is described previously (Subramanian & Schilling, 2014). Primary antibodies utilized include MF 20 (1:100 dilution; Developmental Studies Hybridoma

47

Bank, Iowa City, IA, USA), α-Thbs4b (1:200 dilution; GTX125869, GeneTex, Inc., Irvine, CA, USA), α-GFP (1:200 dilution; sc-9996, Santa Cruz Biotechnology, Inc., Dallas, TX, USA). Alexa Fluor™-conjugated secondary antibodies from Thermo Fisher Scientific (Waltham, MA, USA) were used at a 1:1000 dilution.

Microscopy and figure processing

Confocal *z*-stacks were collected using a Zeiss LSM 710 and ZEN software. *In situ* hybridization images were collected using a Zeiss Axio Imager.A1 equipped with an AxioCam HRc, which was operated through AxioVision release 4.9.1 SP1. Images were processed and measured in Fiji (Schindelin et al., 2012; Schneider et al., 2012). Figures were assembled in the GNU Image Manipulation Program (GIMP).

Cell counting

Fixed 60 hpf embryos from an incross of *scxa:mCherry;cyp26b1^{b1024}* were stained with TO-PRO®-3 Iodide (642/661) (T3605, Life Technologies, Carlsbad, CA, USA) and imaged as described. Confocal z-stacks were processed and analyzed using Imaris v.8.4.0. Briefly, the Spots module was used to detect nuclei in the TO-PRO channel with a diameter of 3 μm and an automated quality filter and a minimum intensity filter. The Surface module was used to generate a volume rendering of tenoblasts, and a region was selected containing the tenoblasts surrounding the intermandibularis muscles, extending from the oral ectoderm to the point best separating anterior tenoblasts from sternohyoideus tendon aggregates. Then, the Imaris XTension "Spots Split to Surface Objects" automatically counted the nuclei located inside the anterior tenoblast Surface object.

Cell Transplantations

Transplantation experiments targeting cranial neural crest were performed as described elsewhere (Eberhart et al., 2006; Sheehan-Rooney et al., 2013).

Morpholino injection

A previously described morpholino, Cyp26b1-SDEx3 MO (Laue et al., 2008), was used. Approximately 3 nl of morpholino (working concentration 0.9 mM) were injected into zebrafish embryos between the one-cell and four-cell stages.

Contributions

Patrick McGurk carried out *in vivo* time-lapse imaging, cloned the *503unc:EGFP* construct and established a transgenic line, performed analyses, assembled figures and wrote the first draft of the manuscript. For panels provided by Mary Swartz, Mary carried out genetic mosaic transplants, Alcian/Alizarin staining, antibody staining, and imaging. For all other fixed specimens, Patrick carried out *in situ* hybridization, antibody staining, Alcian/Alizarin staining, and imaging. Mary Swartz performed other genetic mosaic transplants with the help of Patrick and Johann Eberhart. Patrick, Mary Swartz, and Johann Eberhart designed the experiments and revised the manuscript.

Chapter 3: Examination of muscle attachment stability in zebrafish *scxa* and *xirp2a* mutants

We sought to evaluate an experimental paradigm for assaying the stability of the vertebrate myotendinous junction (MTJ) and investigate candidate factors that might contribute to MTJ stability. Using CRISPR/Cas9 mutagenesis we generated novel loss-of-function alleles of zebrafish *scxa* and *xirp2a*. Homozygous loss of either gene caused no overt developmental defects in larvae, nor did loss of both genes' functions. We induced neuromuscular seizure following the establishment of muscle pattern in live 3 day old embryos. A 30-minute treatment with the GABA-A receptor antagonist PTZ disrupted a varying number of muscle attachments in axial muscles in nearly all embryos treated. Muscle detachment was not significantly increased in mutants compared to their non-mutant siblings. We observed a non-significant increase of muscle detachment between control embryos and embryos deficient for Tsp4b. This finding comes in stark contrast to the results of experiments assessing MTJ stability by stimulating muscles with electrical pulses, which had severe effects on *tsp4b* morpholino injected embryos and no effect on uninjected siblings. We will continue to generate mutant lines and pursue a different line of experimentation for interrogating MTJ stability, concluding that our current PTZ treatment paradigm produces inconsistent, high-variance results.

INTRODUCTION

Myotendinous junctions (MTJs) are complex structures connecting skeletal muscle to tendon (reviewed in Charvet et al., 2012). Starting at the last sarcomere of a muscle fiber, actin filaments extend beyond the Z-disc. Cytoskeletal extension folds the sarcolemma (plasma membrane) to form finger-like projections that extend over the tendon's surface, increasing the surface area of the interface (Tidball & Lin, 1989). The interface has several

layers of connectivity. The actin cytoskeleton is connected to the sarcolemma by protein complexes. Transmembrane proteins (e.g., integrins) in turn connect to components of the muscle basement membrane (e.g., laminins) that bind on the other side to tendon extracellular matrix. These are some of the many known components employed by muscles at MTJs. Still others have unknown or only partly understood functions.

The Xin actin-binding repeat protein (Xirp) family is expressed in muscles near their attachments, but what roles these proteins play there are unclear. The family is named after the first member that was identified, Xin, named so for the Mandarin word for "heart." Loss of Xin function caused morphogenetic heart defects, but other muscle attachments labeled by *Xin* expression appeared normal (D. Z. Wang et al., 1999). Later, a related gene, *Xirp2,* was discovered to have sequence homology to Xin's actin-binding repeats (H. T. Huang et al., 2006). Zebrafish have three confirmed Xirp genes – *xirp1* (analogous to *Xin*), *xirp2a,* and *xirp2b*. Of these, *xirp2a* appears to be expressed at all muscle attachment sites, whereas *xirp1* and *xirp2b* are not clearly expressed in cranial muscles. Though they may have specific roles depending on the type of muscle, it seems probable that Xirp2a and its related proteins participate in MTJ formation or maintenance given their localized expression at muscle tips and their ability to bind actin and nebulin (Pacholsky et al., 2004; Cherepanova et al., 2006; Eulitz et al., 2013).

Tendon's physical properties are attributed to its rich extracellular matrix, composed mostly of type I collagen. What is known about its function at the MTJ is limited to its expression of extracellular matrix components that connect to the muscle basement membrane and sarcolemma. For example, Tsp4, a member of the thrombospondin family associated with extracellular activation of TGFβ signaling, is highly expressed at the MTJ, where it interacts with laminin and promotes integrin signaling (Subramanian & Schilling, 2014). Morpholino knockdown of zebrafish *tsp4b* weakens MTJ stability, as evidenced by

widespread somitic muscle detachment following electrical stimulation that does not disrupt muscles in untreated wild-types (Subramanian & Schilling, 2014). This report demonstrated that tendon matrix not only provides binding targets for muscles, but organizes MTJ formation.

Loss of tendon extracellular matrix proteins is deleterious to tendon function, but what of factors upstream of these? We know almost nothing about how regulation of tendon development contributes to MTJ formation and stability. The major regulator of tendon specification, *Scx*, is dispensable for tendon differentiation and establishment of muscle attachments (Murchison et al., 2007). *Scx*-null mouse pups survive to term, but they have altered tendon morphology and deficient limb and tail mobility (Murchison et al., 2007). The fact that expression of *Tnmd*, a widely used marker of differentiating tendons, is lost in these mutants suggests that Tnmd function is also dispensable for tendon differentiation. Tnmd is necessary for collagen fibril organization in tendons (reviewed in Dex et al., 2016), which may account for most or all of the functional and physiological defects of *Scx* mutants. Still, it remains unclear how loss of Scx or Tnmd affects tendon function in terms of physical properties or MTJ stability.

A simple, replicable assay for MTJ stability under strain would facilitate further analysis of the factors that assemble and maintain both sides of the MTJ. Electrical stimulation robustly causes muscle detachment in sensitive zebrafish embryos (Subramanian & Schilling, 2014), but requires specialized equipment and individual manipulation of live embryos. Here, I put forth a protocol for inducing muscle detachment in embryonic zebrafish by constitutive stimulation of muscle contractions in chemical-treated embryo medium. I describe novel zebrafish mutants of *scxa* and *xirp2a* generated by CRISPR/Cas9 mutagenesis and test the procedure on these mutants and *tsp4b* morpholino injected embryos. I find that loss of *scxa*, *xirp2a*, or both genes' functions

causes no apparent developmental defect in the larval zebrafish musculoskeletal system. Furthermore, I find that seizures induced by PTZ inconsistently disrupt MTJ stability. Work will continue toward finding an appropriate method that's affordable, easy, and effective. We will also target more tendon and muscle genes for CRISPR mutagenesis.

RESULTS

Generation of *scxa* and *xirp2a* mutant zebrafish

We injected single-cell stage AB zebrafish embryos with recombinant Cas9 protein and guide RNA for *scxa* or *xirp2a.* Once P0 fish grew to adulthood, we screened individuals by back-crossing them to wild-type ABs. We lysed day-old F1 offspring and genotyped them first by PCR and gel electrophoresis. Both gRNAs targeted exon sequences containing restriction endonuclease cut sites. We cut PCR products with AflIII or SacI for *scxa* or *xirp2a*, respectively, and selected samples that displayed a longer band around the predicted size of an uncut fragment. I then sequenced PCR products from these samples to discern the nature of the lesion at the restriction site. From the sequencing results I chose P0 individuals that transmitted a nonsense allele at relatively high frequency and other mutant alleles infrequently. I crossed these individuals to ABs and grew the F1 offspring to adulthood, then genotyped the adults by PCR and sequencing.

For *scxa*, I targeted an AflIII restriction site in the first exon, shortly downstream of the coding sequence start codon. I identified the allele *au97*, a five-basepair deletion (Fig 3.1A). I also identified *au98*, a combined deletion of two basepairs and insertion of three basepairs (Fig 3.1B). Both of these lesions shift the reading frame into the same position, immediately followed by a stop codon, resulting in coding sequences of 30 and 36 bases, respectively.

I targeted a SacI restriction site in the fifth exon of *xirp2a*. I chose exon 5 because it was included in all three predicted transcripts of the gene at the time (there are two now).

53
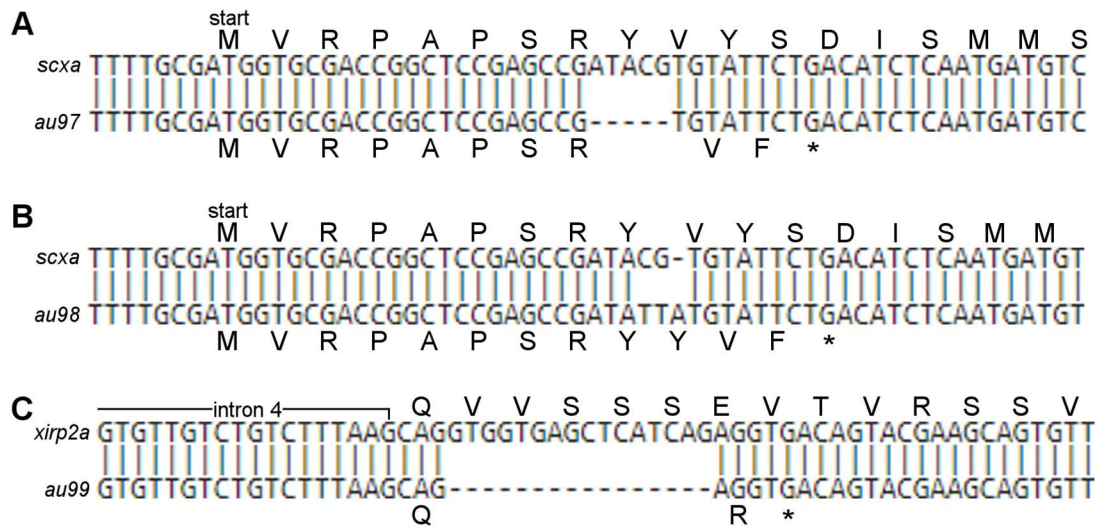
Fig 3.1. Novel alleles of *scxa* and *xirp2a*.

(A,B) The indel alleles *au97* and *au98* both cause a shift in the *scxa* reading frame close to the coding sequence start codon. (C) The 16-bp deletion allele *au99* shifts the *xirp2a* reading frame in exon 5, resulting in a stop codon at the 180th codon position.

I identified the allele *au99* (Fig 3.1C), a 16-basepair deletion. This deletion also shifts the reading frame to create a nearby stop codon, after 179 amino acids of 3315 total and well before the actin-binding Xin repeat domains.

Zebrafish *scxa* and *xirp2a* mutants are phenotypically wild-type

Next, we tried to assess whether loss of *scxa* made tendons more susceptible to muscle detachment, or if loss of *xirp2a* made muscles more susceptible to detachment. We incrossed *au97, au98,* and *au99* carriers as well as *au97;au99* double heterozygote carriers and treated a subset of the embryos with PTZ. Loss of either or both genes' function caused no overt developmental phenotype (data not shown). In *scxa* mutants, *xirp2a* mutants, double mutants, and sibling embryos there were no severe cranial muscle defects whether treated with PTZ (Fig 3.2A-D) or untreated (data not shown).

PTZ treatment caused varying amounts of trunk muscle detachment in mutants and siblings alike. Sibling embryos (Fig 3.2E) had, on average, 77.9 detached muscles per mm^2

Fig 3.2. Example muscle phenotypes in 3 dpf PTZ-treated embryos from a cross of *scxa;xirp2a* double heterozygotes

Fig 3.2. Example muscle phenotypes in 3 dpf PTZ-treated embryos from a cross of
scxa;xirp2a double heterozygotes

 (A-D) Ventral view of jaw muscles, anterior up. Jaw muscles are phenotypically wild-type for all genotypes whether PTZ-treated or untreated. (E-H) Lateral view of trunk muscles, anterior left. (E) PTZ treatment causes trunk muscles to detach sporadically in wild-type animals. Muscles are phenotypically wild-type, and PTZ causes similar levels of muscle detachment in _scxa_ mutants (F), _xirp2a_ mutants (G), and _scxa;xirp2a_ double mutants (H). Scale bars 50 μm.

(SEM = 14.7, n = 18). In *scxa* mutants (Fig 3.2F), the average was 92.7 (SEM = 12.6, n = 13). In *xirp2a* mutants (Fig 3.2G), the average was 68.2 (SEM = 11.2, n = 13). In *scxa;xirp2a* double mutants (Fig 3.2H), the average was 94.6 (SEM = 18.2, n = 6). There were no significant differences between groups (*p > 0.4)*. Within groups, counts were highly variable. At best, the standard deviation represented 47.2% of the mean for double mutants. At worst, it represented 79.9% of the mean for non-mutant siblings. These results suggest that, without insult, *scxa* and *xirp2a* mutant zebrafish have normal phenotypes. More work is required to assess MTJ stability in these mutants.

Tsp4b-deficient zebrafish are not significantly sensitized to PTZ-induced muscle detachment

Since morpholino knockdown of *tsp4b* caused severe muscle detachment to occur in electrically stimulated embryos, I used the same morpholino to test the chemical approach in a known model of MTJ deficiencies. We treated both morpholino-injected embryos and uninjected control embryos with the seizure-inducing GABA-A receptor antagonist PTZ for 30 minutes just before 3 days post-fertilization (dpf). Control and morpholino injected embryos alike twitched persistently in embryo medium containing 15mM PTZ. PTZ-treated and untreated embryos were subsequently fixed after 1 hour of recovery and stained for imaging. An antibody for Tsp4b strongly labels the myosepta in uninjected embryos (Fig 3.3A), but does not label *tsp4b* morpholino injected embryos (Fig 3.3B). I counted rounded myosin-positive cells and detached myofibers in the head and trunk. Head and trunk muscles did not detach in untreated embryos (data not shown). Muscles appeared normal in the heads of uninjected PTZ-treated embryos (Fig 3.3C), and trunk muscles had a few rounded cells (Fig 3.3D). I observed no apparent defects in the heads of morpholino injected PTZ treated embryos (Fig 3.3E). However, in the trunk (Fig 3.3F), there were, on average, 100.9 detached muscles per mm^2 (SEM = 42.2, n = 7),
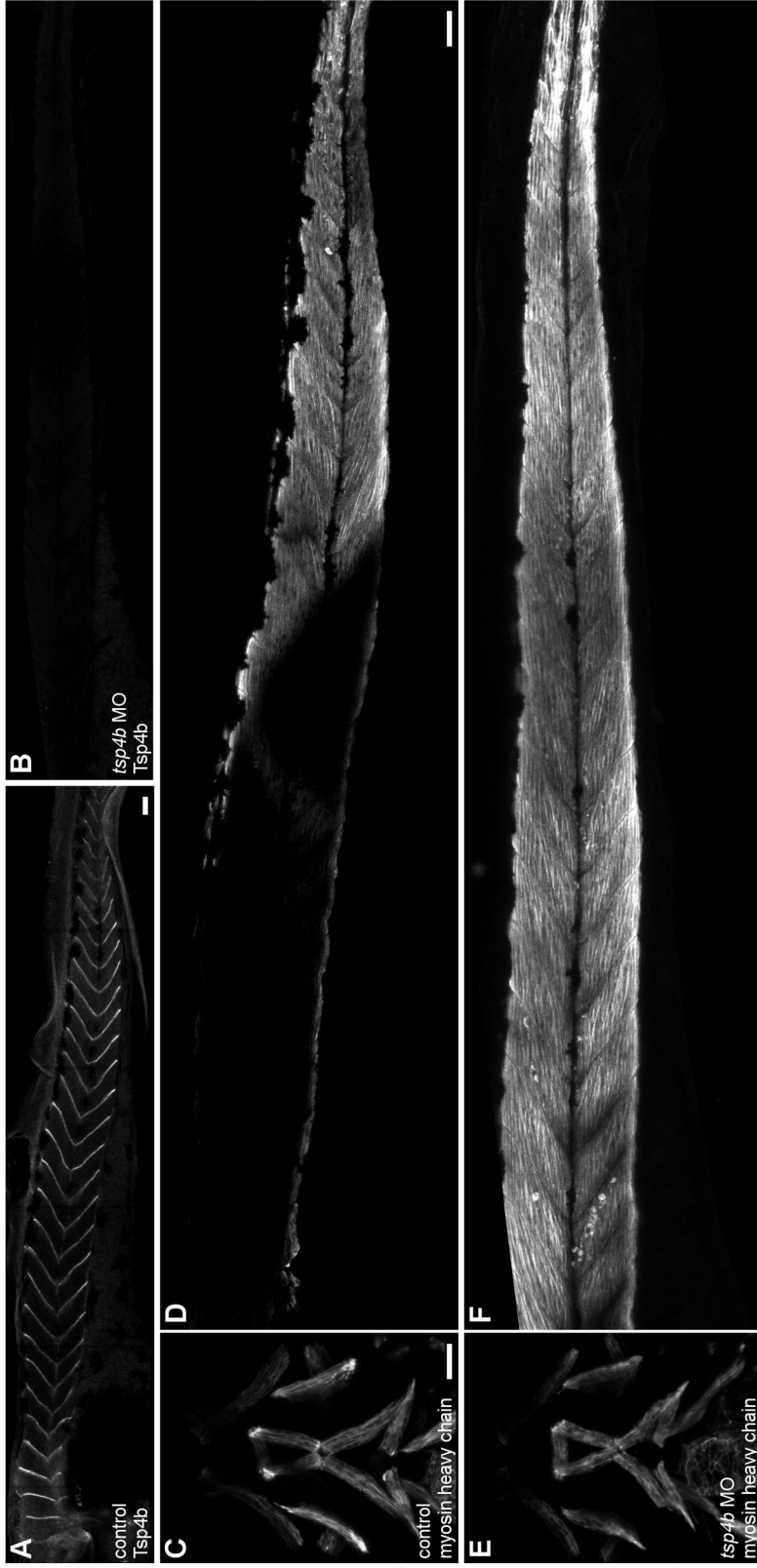
Fig 3.3. Trunk muscles and myosepta in PTZ-treated 3 dpf zebrafish embryos

(A) Normal myosepta are enriched with Tsp4b protein. (B) Injection with *tsp4b* MO effectively depletes Tsp4b protein levels *in vivo*. (C,D) No defect is seen in jaw muscles, but PTZ treatment causes a small number of trunk muscle detachments even in control embryos. (D) Jaw and trunk muscles have wild-type morphology in *tsp4b* MO injected embryos, but more trunk muscle fibers appear to detach from PTZ treatment. Scale bars 50 µm.

58

compared to 21.1 in uninjected embryos treated with PTZ (SEM = 8.3, n = 4). The increase in muscle detachment was not significant, however (p = 0.1981). Though the means differed between the two groups, the variance lessened any significance. The standard deviations of uninjected controls and *tsp4b* MO embryos represent 78.5% and 110.6% of their respective means. Quantification of muscle detachment for *tsp4b* MO and mutant incross experiments is shown in Fig 3.4. Together, these results indicate that the PTZ assay induces muscle detachments with inconsistent severity in zebrafish embryos with MTJ defects, controls, and *scxa* and *xirp2a* mutants.

DISCUSSION

This work introduces new alleles of *scxa* and *xirp2a* available to interested researchers. We report negative results for developmental defects in these mutants as well as for the method of causing muscle detachment by PTZ treatment. Described in the Methods section is an automated process to facilitate counting of detached muscles with ImageJ.

Both *scxa* and *xirp2a* are dispensable for muscle patterning and MTJ formation

Neither loss of *scxa* or *xirp2a* nor loss of both genes caused apparent morphological phenotypes. It must be noted that loss of function has yet to be conclusively demonstrated for these mutants. However, each lesion shifts the reading frame of the coding sequence to include an early stop codon. A combined approach of *in situ* hybridization and RT-PCR will verify whether nonsense-mediated decay of *scxa* and *xirp2a* mRNAs occurs in their respective mutants.

For *scxa*, a lack of larval phenotype is consistent with the *Scx*-null mouse model. Mice deficient for *Scx* do have defects, namely in limb function and ultrastructure of long force-transmitting tendons. Zebrafish fins bear little homology to the tetrapod limb, and zebrafish have few long tendons at larval stage. It remains an interesting question as to
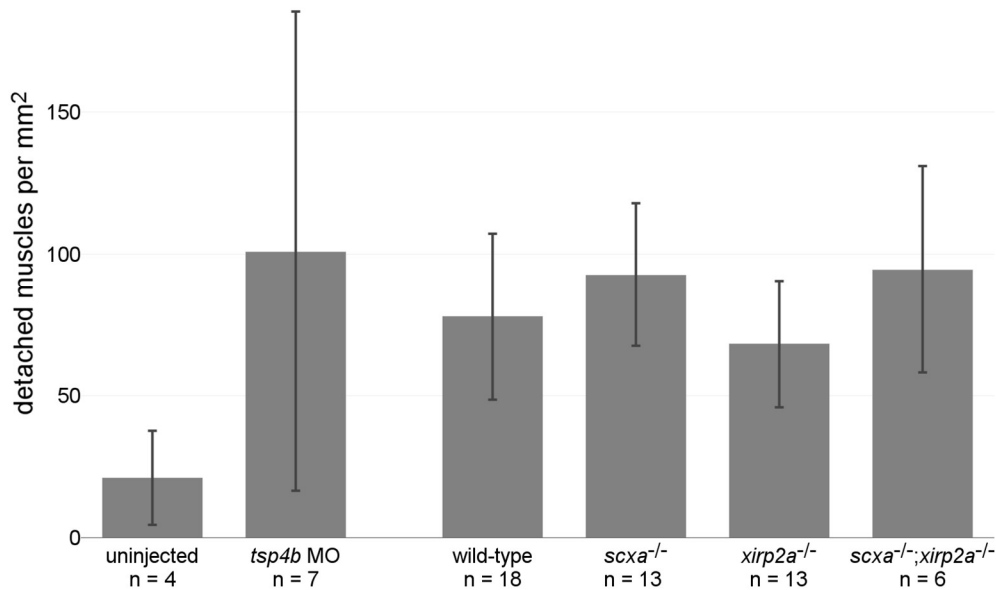
Fig 3.4. Comparison of muscle detachment rates across tested groups

For each trunk imaged, the number of detached muscles was divided by the measurable area. Error bars represent the 95% confidence interval of the mean.

whether tendon ultrastructure is disrupted in zebrafish *scxa* mutants. Additionally, this work does not rule out the possibility of genetic compensation for *scxa.* There is another zebrafish transcript paralogous to *scxa* that can be detected by RT-PCR, but *scxb* expression is not visible by *in situ* hybridization from embryonic to adult stages (Chen & Galloway, 2014). It is therefore difficult to conclude that *scxb* compensates for loss of *scxa,* but difficult also to discount the possibility. If *scxb* expression in *scxa* mutants was assayed by qPCR, multiple fold changes would be expected if it does indeed compensate for *scxa* loss.

On the other hand, *xirp2a* has a known paralog (*xirp2b*) and another family member (*xirp1*) that are both strongly expressed at muscle tips. Each of these genes has a specific expression profile in embryonic zebrafish, but may have overlapping compensatory roles. The earliest expressed is *xirp1*, and *xirp2a* is expressed most broadly and strongly (Otten

et al., 2012). Because *xirp2b* is expressed only after 24 hpf in a more restricted pattern, it doesn't seem to play a developmental role. Other Xirps are likely to compensate for loss of *xirp2a* in myosepta, but it remains unclear whether they do so in the head or if *xirp2a* is dispensable for cranial muscle development and attachment. The heart and somitic muscles develop normally without the functions of *xirp1* and *xirp2a* (Otten et al., 2012), which is strikingly different from the requirement for *Xirp1 and Xirp2* in mouse heart development (reviewed in Q. Wang et al., 2014). In mouse myoblast cell culture, Xirp1 and Xirp2 colocalize with an interacting partner, nebulin, in developing myofibers but not mature muscle (Eulitz et al., 2013). Similar to a developmental gene, zebrafish Xirps are upregulated in injured tissue (Otten et al., 2012). The same is true in cultured muscle samples from human patients with myopathy (Nilsson et al., 2013) and injured mouse muscle (Nissar et al., 2012). Together, these results suggest that zebrafish Xirps have some overlapping function with those in mammals and chick, but may also have disparate functions.

The utility of convulsant drugs for muscle strain assays remains unclear

This report demonstrates that somitic muscle fibers detach from myosepta in zebrafish larvae treated with the convulsant drug PTZ. However, the same analysis must be applied to untreated controls to verify quantitatively that muscle detachments were PTZ-related rather than stochastic. The variability of impacts seemingly caused by PTZ treatment is worrying, but the approach has its merits. With zebrafish, large numbers of embryos can be treated at once with any drug deliverable in embryo medium. Drug treatments usually require no special equipment. Furthermore, the muscle detachment phenotypes shown in this work are rather discrete, allowing for appropriate quantification of phenotypic severity and variance. I employed automated processing of whole zebrafish trunk images to rapidly count detached muscles and output an image with detached muscles

labeled. Additionally, the number of detached muscles on average would have been tedious to count manually, but the algorithm afforded speed and accuracy by providing a labeled map and a count for easy manual validation.

The weakest point of this approach for assessing MTJ stability under strain is perhaps the choice of convulsant. Many publications, mostly neurological, have utilized PTZ for a zebrafish model of seizures. However, PTZ is just one of several GABA-A receptor antagonists, which is just one class of convulsant drugs. Different convulsants have been shown to elicit different waveforms, amplitudes, and durations of electrical discharge in the zebrafish forebrain (Baraban, 2013). Screening various convulsants for muscle strain activity could be done with *tsp4b* MO injected embryos. With the right drug and optimized parameters, this methodology could realistically allow quantifiable assessment of MTJ stability with minimal technical labor.

METHODS

Danio rerio (zebrafish) care and husbandry

Zebrafish stocks were maintained and embryos were raised according to established protocols (Westerfield, 2007) with approval from the University of Texas at Austin Institutional Animal Care and Use Committee. Embryos were bathed in 15 mM PTZ (Sigma-Aldrich, St. Louis, MO, USA) in embryo medium just before 3 dpf. After 30 minutes in PTZ, embryos were given an hour to recover before euthanasia and fixing.

Morpholino injection

A previously described morpholino cocktail against *tsp4b* and *p53* was used (Subramanian & Schilling, 2014). Approximately 150 ng of each was injected into zebrafish embryos between the one-cell and four-cell stages.

CRISPR/Cas9 mutagenesis

Guide RNA synthesis from a PCR-amplified long oligonucleotide template was carried out as described previously (Hruscha et al., 2013; Talbot & Amacher, 2014). The target sequence for *scxa* was GGATGTCAGAATACACGTAT, and for *xirp2a* GGAGGTGGTGAGCTCATCAG. RNA solutions (200 ng/nL) were mixed 1:1 with recombinant NLS-tagged Cas9 protein (PNA Bio, Newbury Park, CA) just prior to injection. Single-cell embryos were injected with approximately 1 nL of Cas9/gRNA mix.

Immunohistochemistry

In preparation for immunohistochemistry, embryos were fixed in 95% methanol/5% glacial acetic acid. The protocol for myosin heavy chain/Tsp4b staining is described previously (Subramanian et al., 2014). Primary antibodies utilized include MF 20 (1:100 dilution; Developmental Studies Hybridoma Bank, Iowa City, IA, USA), α-Thbs4b (1:200 dilution; GTX125869, GeneTex, Inc., Irvine, CA, USA), α-GFP (1:200 dilution; sc-9996, Santa Cruz Biotechnology, Inc., Dallas, TX, USA). Alexa Fluor™-conjugated secondary antibodies from Thermo Fisher Scientific (Waltham, MA, USA) were used at a 1:1000 dilution.

Microscopy and figure processing

Confocal *z*-stacks were collected using a Zeiss LSM 710 and ZEN software. Images were processed and measured in Fiji (Schneider et al., 2012; Schindelin et al., 2012). Figures were assembled in the GNU Image Manipulation Program (GIMP).

Automated counting of detached muscles

The following process was designed and scripted using the Fiji distribution of ImageJ (Schneider et al., 2012; Schindelin et al., 2012) and the MorphoLibJ package (Legland et al., 2016). ImageJ macro code is available at github.com/pdmcgurk.

Pre-processing: Confocal z-stacks along the length of the zebrafish trunk were stitched together using the Plugins > Stitching > Pairwise Stitching function (Preibisch et al., 2009). 2D maximum intensity projections were made from the full-length stacks. If the maximum intensity projection was saturated, a standard deviation projection was used instead. Then, if needed, the image was rotated to make the trunk horizontal. For each projection, the channels were split into 8-bit grayscale images, and the muscle image was duplicated twice.

Convolution: The first copy was filtered to reduce the intensity of normal muscle fibers, which run close to horizontal, and enhance round spots that are slightly brighter than their surroundings. Filtering was done with the Process > Filters > Convolve… function, using the custom convolution kernel shown in Fig 3.5. Then the image is binarized with threshold values from 250 to 255 to isolate the brightest resulting spots. Finally, the smallest bright spots are eliminated by three iterations of the Process > Noise > Despeckle operation. This worked well to remove false positives in the last step of identifying detached muscles. Users should be mindful that the risk of creating false negatives increases per iteration, so in some instances three iterations may be excessive for their data.

Mask: The other muscle image copy was converted to a binary mask to fill out cell shapes and connect two or more bright spots that represent the same object. Steps follow:

1. Image > Adjust > Auto Local Threshold: Otsu method (Otsu, 1979), radius 20

2. Plugins > MorphoLibJ > Morphological Filters: opening, square element, radius 2

3. At this point, the script pauses. The previous two operations generate a binary representation of the original image with elongated objects outlining the specimen. A dialog prompts the user to use the paintbrush tool to connect these outline objects, select the interior of the outline with the wand tool, and click OK.

4. Edit > Clear Outside

```
-0.1  0    0    0    0    0    0    0    0    0    0    0    0    0   -0.1
-0.1 -0.1  0    0    0    0    0    0    0    0    0    0    0   -0.1 -0.1
-0.2 -0.1 -0.1  0    0    0    0    0    0    0    0    0   -0.1 -0.1 -0.2
-0.2 -0.2 -0.1 -0.1  0    0    0    0    0    0    0   -0.1 -0.1 -0.2 -0.2
-0.2 -0.2 -0.2 -0.1 -0.1  0    0    0    0    0   -0.1 -0.1 -0.2 -0.2 -0.2
-0.3 -0.3 -0.2 -0.2 -0.2 -0.1  0    1    0   -0.1 -0.2 -0.2 -0.2 -0.3 -0.3
-0.2 -0.2 -0.3 -0.3 -0.3 -0.2  1    2    1   -0.2 -0.3 -0.3 -0.3 -0.2 -0.2
-0.1 -0.1 -0.1 -0.1 -0.1 -0.3  2    3    2   -0.3 -0.1 -0.1 -0.1 -0.1 -0.1
-0.2 -0.2 -0.3 -0.3 -0.3 -0.2  1    2    1   -0.2 -0.3 -0.3 -0.3 -0.2 -0.2
-0.3 -0.3 -0.2 -0.2 -0.2 -0.1  0    1    0   -0.1 -0.2 -0.2 -0.2 -0.3 -0.3
-0.2 -0.2 -0.2 -0.1 -0.1  0    0    0    0    0   -0.1 -0.1 -0.2 -0.2 -0.2
-0.2 -0.2 -0.1 -0.1  0    0    0    0    0    0    0   -0.1 -0.1 -0.2 -0.2
-0.2 -0.1 -0.1  0    0    0    0    0    0    0    0    0   -0.1 -0.1 -0.2
-0.1 -0.1  0    0    0    0    0    0    0    0    0    0    0   -0.1 -0.1
-0.1  0    0    0    0    0    0    0    0    0    0    0    0    0   -0.1
```

Fig 3.5. Convolution kernel for enhancing bright, round spots over normal muscle fibers.

Each pixel's value in the image is recalculated based on this matrix. The pixels in a square 15 wide by 15 tall are each multiplied by the value in the corresponding location in the kernel. The adjusted values are summed, and the total assigned to the pixel at the center.

5. Plugins > MorphoLibJ > Binary Images > Connected Components Labeling: connectivity 4, 16 bit type (for up to 65,536 objects)

6. Plugins > MorphoLibJ > Analyze > Region Morphometry

7. Plugins > MorphoLibJ > Label Images > Assign Measure to Label (twice)

   • This step is done twice to generate images in which the objects are grayscale coded by their area in one and by their circularity in the other. If done manually, one simply chooses the appropriate measure from the morphometry data table in the dialog.

Lacking that ability, the macro saves the data table locally and then imports area and circularity values into individual data tables.

8. Process > Image Calculator: Image1 = Area, Operation = Divide, Image2 = Circularity

9. Edit > Invert

10. Image > Adjust > Threshold…: Max - 500 to Max

- Manually, click Reset in the Threshold dialog, then Set. Change the Lower Threshold Value to 500 less than the Upper Threshold Value, click OK, then Apply.

Reconstruction: The final convolution image and mask image were combined to generate a map of detached muscles. This was done using the Plugins > MorphoLibJ > Morphological Reconstruction tool. The convolution product was set as the marker image, and the final mask product as the mask image, with the type set to dilation and a connectivity of 4. In essence, wherever one or more marker objects intersect a mask object, the mask object was included in the output. Objects in the resulting image were counted with the Analyse > Analyze Particles… function. For the output, objects were then converted into outlines using the Plugins > MorphoLibJ > Morphological Filters tool by an External Gradient operation with a disk element of radius 3. The macro records the detached muscle count in a text window. It also saves images of the final map by itself as well as overlaid with the original muscle image in different channels on an RGB image.

Contributions

Patrick McGurk designed the experiments, identified genetic lesions, imaged specimens, performed analyses, assembled figures, and wrote the first draft of the manuscript. Patrick synthesized guide RNAs, injected P0 embryos, and genotyped F1 offspring with the help of Rhonda Stanley. Patrick treated embryos with PTZ and performed antibody staining with the help of Yohaan Fernandes. Patrick and Johann Eberhart revised the manuscript.

Chapter 4: ChromatVis, an all-purpose chromatogram viewer for Sanger sequencing of homozygous and heterozygous loci.

Sanger sequencing continues to be widely used by molecular biologists in the age of next-generation sequencing. The process is mostly automated now, including algorithmic detection of signal intensity peaks and the corresponding base sequence. Sanger sequencing is particularly useful in identifying mutant sequences generated in forward genetic screens or via reverse genetic approaches such as CRISPR/Cas9 that produce unpredictable indels. However, if the sequenced locus is heterozygous for a mutant SNP, automated sequencers fail entirely and produce either 'N' or one of two possible base calls for the double peak. Worse yet are indels because the size difference shifts the sequence of the mutant allele relative to the wild-type allele, resulting in failed base calls from the point of the mutation onward. The chromatogram contains both the wild-type and mutant sequence, but even the human eye struggles when trying to separate a known base sequence from unknown sequence encoded in doublet peaks. A few software tools have been developed to separate heterozygous sequences, mostly by using a base string as reference. Presented here is an all-purpose chromatogram viewing program, ChromatVis, which combines standard viewing functionality with an algorithm that aligns two chromatograms and calculates a visual representation of the differences between them. ChromatVis uses this "difference profile" to algorithmically separate the sequence encoded in the reference peaks from the alternate sequence in the doublet peaks of the query. This method accounts not only for called bases, but incorporates actual signal intensity, which varies within and across experiments.

INTRODUCTION

Dideoxynucleotide chain-termination sequencing, also known as the Sanger method (Sanger et al., 1977), remains an integral part of the molecular biology landscape. As of July, searching Google Scholar for "Sanger sequencing" yields 9,920 results published in the year 2017. The raw data of automated sequencers are four arrays of signal intensity values, one channel for each DNA base. The output files containing these data and their metadata are commonly referred to as chromatograms, which is also the name of a plot of these data. A great number of software tools have been developed and publicly released for viewing or analyzing chromatograms. One area getting considerable attention currently is decoding both alleles in heterozygous sequences. One of the major reasons for that focus is the recent popularity of CRISPR/Cas9 and other nuclease-based mutagenesis techniques. After mutagenizing animals, researchers can spend a considerable amount of time determining which alleles show promise and which individuals transmit a particular allele frequently. Sequencing the target locus in the F1 generation is a quick way to do so, provided the right tool to help sort the mutant allele from the wild-type sequence.

Currently, free and open-source tools that exist for this purpose focus on the function of parsing out the basecalls. PolyPeakParser (Hill et al., 2014) and Indigo (Rausch, 2017) are recent releases that perform excellently in that area. Here we describe a new multifunctional chromatogram viewer and analyzer called ChromatVis. This program was inspired by an older tool called SeqDoC (Crowe, 2005), which does not employ base calling algorithms, but rather aligns two chromatogram files to each other and generates a visual representation of the differences between them. This difference profile is flat where the sequence is the same, and may include small positive or negative peaks where intensity differs between peaks or small peaks of noise appear in other channels. Where the query sequence has a heterozygous or homozygous peak that differs from the reference peak,

strong peaks both positive and negative mark the position. The results produced by SeqDoC are intuitive and highly accurate, but leave the user responsible for manually annotating the base sequences from the bidirectional peaks in the difference profile. ChromatVis combines a fully featured chromatogram viewer and SeqDoC-style difference profiling with automated variant annotation.

ChromatVis has two modes of use. The first is to open a single chromatogram in its own window. This mode doesn't warrant further mention because the other mode runs all of the same algorithms and a few more of its own. In the latter mode, one chromatogram selected as the query is aligned to another selected as the reference. After aligning the sequence traces, the program calculates a difference profile. Then it looks to identify any bidirectional peaks in the difference profile and heterozygous peaks in the query sequence. The aligned sequence traces are drawn in the same window, stacked atop each other with the option to view the difference profile in between them. The difference profile highlights variants, which are marked in the chromatogram image as well by a base letter below the line of base letters that label each peak with the sequencer-provided basecall (Fig 4.1).

The ChromatVis project aims to offer user-friendly interactivity and utility. The program is an all-in-one platform for chromatogram viewing and automated analysis methods including peak detection, base calling, and variant annotation.

FEATURES

Interactive visualization

Foremost, ChromatVis is a chromatogram viewer with features that users would expect. Viewer windows scroll left and right in response to mouse wheel up and mouse wheel down, respectively. Buttons, checkboxes, sliders, and dialogs allow the user to manipulate the visualization (Fig 4.2). The first button (Fig 4.2A) opens a dialog that lets
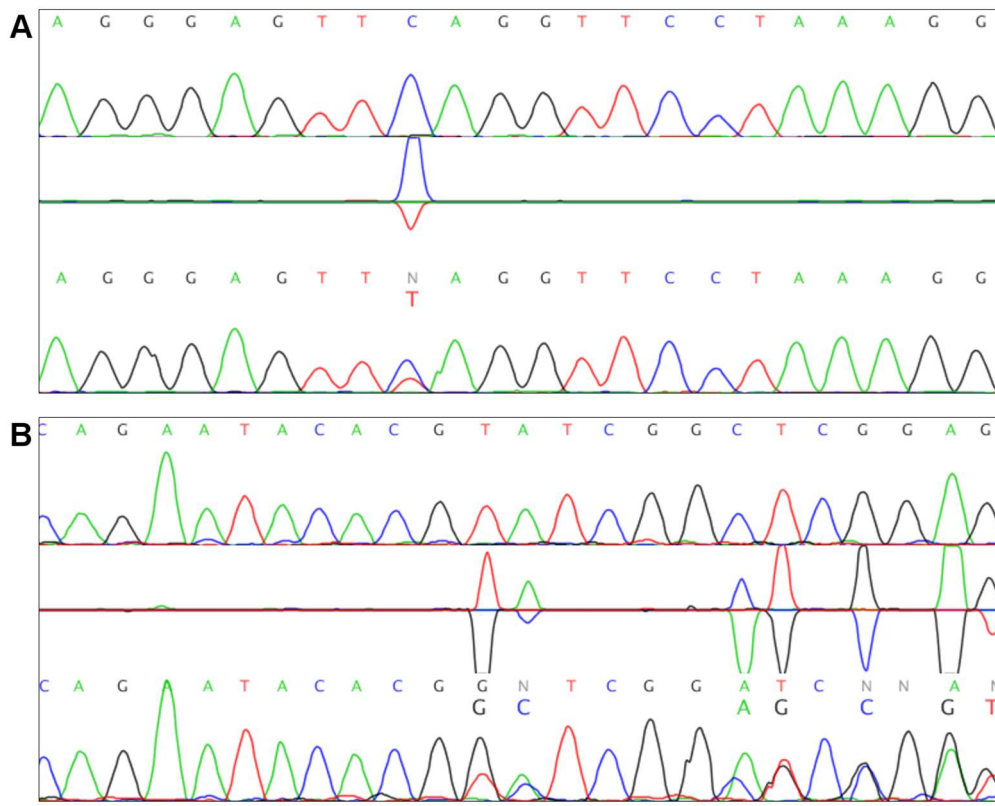
Fig 4.1 Images of aligned chromatograms with difference profiles.

(A) A single base mutation in zebrafish *cyp26b1* results in an early stop codon (TAG). (B) A 5 basepair deletion (TATCG) in zebrafish *scxa* shifts the encoded translation reading frame.

the user specify the pixel dimensions of the canvas (Fig 4.2F). Note that this is not the window size. The window is resized to accommodate the user interface (UI) elements and a canvas of the specified dimensions. There is also a minimum window width, so UI elements are not cut off. If the user specifies canvas dimensions more narrow than the minimum window width, the canvas will appear in the center of the window on a grey background. Checkboxes allow the user to toggle between forward and reverse-complement chromatogram images and toggle whether the difference profile is shown for an alignment (Fig 4.2D). A slider allows the user to change the scale of the X dimension
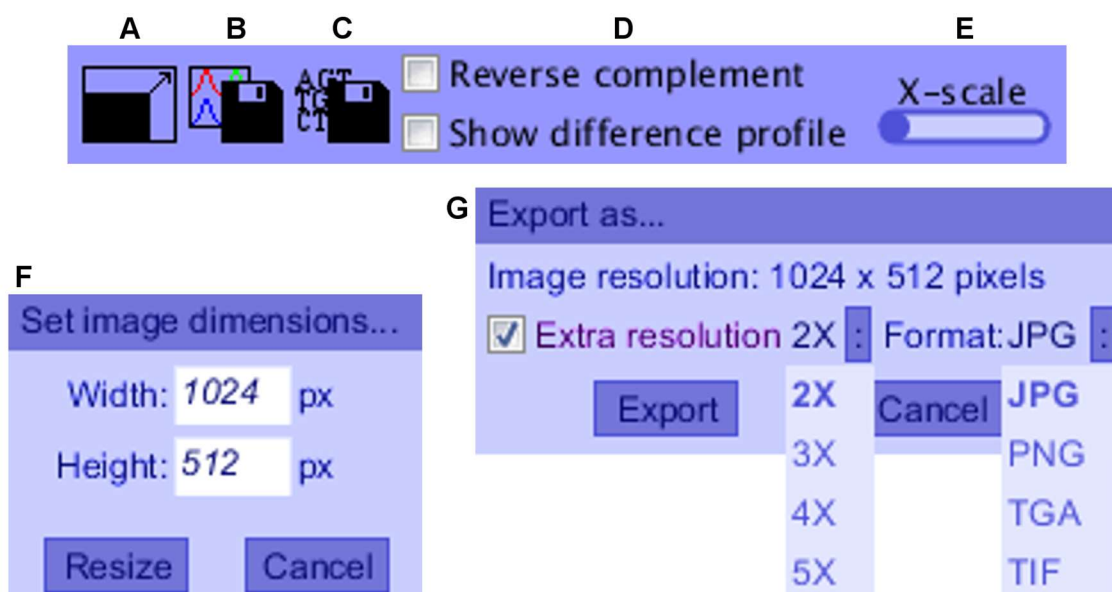
Fig 4.2 ChromatVis interface controls.

(A) Button for opening the image resize dialog (F). (B) Button for opening the image export dialog (G). (C) Button for exporting a text file of base sequences. (D) Checkboxes for toggling viewing modes. (E) Slider for adjusting peak width.

(Fig 4.2E). By default, chromatograms are rendered with a single data point being one pixel wide. The scale can be adjusted to broaden the peaks, up to a width of five pixels per data point.

Image export

In addition to standard chromatogram viewer functionality, generating publication-quality chromatogram images was another priority. The second button (Fig 4.2B) on the viewer window UI bar opens a dialog for image export (Fig 4.2G). This tool is designed to create an image of the current view of the canvas. The user chooses the canvas size, peak width, and the X position before opening the dialog.

The dialog displays the current canvas dimensions, which by default are the dimensions of the output. A checkbox toggles the pixel multiplier option, which makes a

dropdown list appear. The user can choose to scale the output image two to five times from the original dimensions. For example, to create an image 2 inches tall and 6 inches wide at 300 pixels per inch the canvas can be set at 200 pixels high and 600 pixels wide, and exported at 3X resolution.

The checkbox on the right lets the user select the format for the output image. The allowed formats for output from Processing are .jpg, .png, .tga, and .tif. Clicking the "Export" button opens a file chooser dialog, allowing the user to name their file and choose where it is saved, then saves the output image once a path is chosen.

Basecall export

The third button (Fig 4.2C) on the viewer window UI bar generates FASTA-formatted text output for called bases. In its current version, this tool works in one of two modes. For a single chromatogram, the output is the string of basecalls made by the sequencer and encoded in the chromatogram file. For aligned chromatograms, the sequencer basecalls for the reference chromatogram are written first. The program next writes the sequencer basecalls for the query chromatogram. Lastly, the program writes a string of basecalls for the query sequence parsed from the reference by ChromatVis. Clicking the base export button opens a file chooser dialog for selecting the text file output path and subsequently saves the output file once a path is chosen.

ALGORITHMS

Applied Biosystems sequencer output files are read in byte by byte using an algorithm translated from the open-source Python program SeqTrace (Life Technologies Corporation, 2009; Stucky, 2012). The signal traces for each base and sequencer-annotated base positions are stored in array attributes of a SeqRun object, and the basecall letters are stored in a string attribute.

The next part of instantiating a SeqRun is to process the traces using algorithms based on SeqDoC (Crowe, 2005). The first method trims low-value data points from each end of the traces, which reduces subsequent computational load in addition to improving the chromatogram's aesthetics. From the start, data points are skipped in all the traces until the value in one of the traces exceeds 50. The base position array is updated to reflect the shift in starting point, and any negative values result in that position and its corresponding letter being removed from their respective objects. Subsequent data points are appended to new arrays until the value in all traces drops below 50. At that point, the next 100 data points are queried, and if the value in all traces remains under 50 in that span, the trimmed arrays are returned in place of the original trace arrays.

The second method normalizes each trace to a local (1000 data points) mean of 100, which is another aesthetic improvement that's also necessary for comparison of traces between experiments with differing signal intensity. Data points are normalized against their own channel. For each data point, the previous 500 data points and the next 500 data points are averaged. For data points within 500 of one end of the trace, the remaining data points to the end and are used. A scaling factor is calculated by dividing the mean by 100, and the normalized value is determined by dividing the original value by the scaling factor. Normalized values are appended to new trace arrays, which are returned in place of the trimmed arrays.

Aligning query traces to reference traces

When directly comparing two chromatograms, slight variations between experiments can cause peaks to align out of register with reference peaks. SeqDoC has an algorithm for offsetting the starting position of alignment and nudging peaks to maintain register across the chromatogram (Crowe, 2005), which has been translated from Perl and modified slightly for ChromatVis. Briefly, the algorithm works on the following methods.

73

Scoring: A method compares sub-arrays of equal length between the reference and query traces. For each channel, the absolute differences between corresponding data points are summed, and the sum from all channels is the score. Lower scores indicate better alignment.

Fine adjustment: The alignment method scans through data points from both the query and reference sequences, and appends query data points to new traces representing the aligned sequence. For every third point, a baseline score is calculated for the next 30 points (n+1 to n+30) from the current position in each sequence. Two other scores are calculated, using the same 30 points for the reference sequence. If the score for the query points n to n+29 is lowest, the current point is appended twice for each channel to the output. If the score for the query data points n+2 to n+31 is lowest, the current point is not appended to the output. If the baseline score is lowest, the current point is appended like normal.

Initial offset: Another method selects the best offset for aligning two sequences by generating several short alignments. Scores are generated after aligning reference data points 201-1200 to 1000 query data points starting between 1 and 401 (in increments of 10). This requires the alignment method using fine adjustment as described above (i.e. starting at 1 does not necessarily mean the original data points 1-1000).

Calculating differences between aligned chromatograms

Once an aligned query sequence has been generated, the next step is to calculate a "difference profile." This is the essential function of SeqDoC (Crowe, 2005), and the ChromatVis method is a direct translation from Perl. Briefly, the algorithm subtracts every value in the query traces from its corresponding value in the reference traces. This produces an array for each channel with some positive numbers and some negative numbers, most of which are small in magnitude. Subsequently, all of the differences have their magnitudes

squared, which makes the large differences much larger with respect to the smaller ones. The squared differences in each channel are enhanced once more if there are differences in other channels that have the opposite sign. This multiplier takes the square root of the opposite-signed values, which again is biased toward large values getting larger while small ones stay small. If the differences for all bases share the same sign, however, the squared difference values are made zero. This reduces difference profile noise when one experiment has consistently stronger signal intensity than the other.

In terms of differences between normalized signals, a true heterozygous peak has a medium to large negative difference (the alternate peak) and a medium positive difference representing a reduction in signal from the homozygous reference. This results is a strong bidirectional peak in the difference profile. Elsewhere, peaks of noise or fluctuations of intensity in either chromatogram yield unidirectional peaks.

Peak detection

In mathematical terms, a peak is essentially a rapid change in the slope of a signal intensity function with a change in the slope's sign. The peak detection algorithm for ChromatVis is a simplification of an algorithm for Gaussian peaks, findpeaksG.m (O'Haver, 2014), translated from MATLAB. First, it calculates the first derivative of the traces for each channel, and then applies sliding average smoothing to the derivative arrays. Then it scans through the smoothed derivative arrays, identifying positions where the sign of the slope changes from +1 to either 0 or -1. When it encounters such a position, it checks that the change in slope exceeds a threshold value and that the trace value at that position exceeds an amplitude threshold. If both are true, a PeakPoint object is generated with position, value, and base attributes and stored in a list. After all the channels are scanned for peaks, the list is sorted by position.

A modified version of the same algorithm is applied to the difference profile traces to detect zero-crossing events in the derivative arrays from +1 to -1 or from -1 to +1. The program generates a list of both positive and negative peaks and sorts them by position. Then the program iterates over the list of PeakPoints for instances when a peak and the next one in order are opposite in sign and the differences in their positions are below a certain threshold. If so, a method is invoked to set these PeakPoints as each other's opposites, pointing each one to the other, and the first peak by position is added to a list of bidirectional peaks.

For the purposes of detecting double peaks in heterozygous sequences, the program scans through the list of peaks to find pairs below a minimum distance threshold. When a pair is found, its position is referenced against the list of bidirectional peaks. If the double peak and bidirectional peak align closely, a HetPeak object is instantiated and stored in a list. Each of the peaks in the doublet is classified as normal or alternate by matching bases with the positive or negative half of the bidirectional peak, and subsequently linked to their respective HetPeak attributes. If two chromatograms are being compared and a nearby bidirectional peak cannot be found, the doublet is assumed to be spurious and no HetPeak is called. When viewing a single chromatogram, doublets are assessed as to whether one peak value greatly exceeds the other. If so, the doublet is again assumed to be spurious. If not, a HetPeak is instantiated and flagged as an 'N' so only one letter is added to the basecalls for that position.

DISCUSSION

ChromatVis is yet another tool among many for viewing chromatograms and decoding ambiguous sequences. It combines elements from several existing tools and a few of its own into one open-source, all-purpose package. In short, ChromatVis:

- Determines the "alternate" sequence of a chromatogram with heterozygous peaks given a reference chromatogram, using a SeqDoC-based method.

- Provides a fully-featured interactive viewing experience for single chromatograms and aligned chromatograms and their difference profiles.

- Generates high-quality chromatogram images and FASTA sequence output for sequencer-encoded and program-parsed basecalls.

For users performing CRISPR/Cas9, TALEN, or similar site-directed mutagenesis in animals, ChromatVis aims to facilitate the process of identifying genetic lesions passed on by the original mutagenized animals. Identifying indels is particularly laborious by visual inspection of heterozygous chromatograms. SeqDoC makes the differences between chromatograms easily visible, but still requires manually documenting the alternate sequence. Meanwhile, plugging a chromatogram into a form that then returns the alternate base sequence requires faith in an algorithm without the assurance the human eye can provide. By combining algorithmic peak detection and parsing with user-friendly visualization, ChromatVis offers more confidence than either component can provide individually.

Planned features

There are a number of goals for the future of ChromatVis. Most importantly, it is only available currently as a .jar applet or as Processing files, either of which must be downloaded and run locally. Work is underway to turn ChromatVis into a Web applet and an ImageJ plugin including all of the current features. In addition to the ABI file format, ChroamtVis will support Standard Chromatogram Format (SCF) (Dear and Staden, 1992) and ZTR format Bonfield and Staden, (2002). Still other features are in planning currently, listed below.

Configurable options

Currently, there are a number of variables hard-coded into the algorithms that would better serve advanced users if they could be changed. For example, peak detection relies on several variables, such as the number of data points to use in the smoothing method and the number of times to repeat smoothing, as well as the thresholds for slope change and amplitude. Though quality testing has been performed to minimize false positive and false negative peaks, bidirectional peaks, and heterozygous peaks, it's doubtful that the hard-coded values will work perfectly for every user's data. Moving such variables into a text file with default and custom fields and adding a configuration window will be a simple upgrade.

Additionally, more graphical options would benefit users. In the main UI, more toggles are needed for whether to display sequencer-encoded basecalls, program-parsed basecalls, both, or neither. To that same point, users will be able to configure the font and line thickness on the canvas. Another possible feature would let users put custom labels into their images. Other options for the configuration window include UI skins and custom color choices for the UI and canvas. These are all relatively easily done in addition to the other configuration options. A possible, more ambitious upgrade is the addition of a tabbed view for users who want a tidier way of having more than one canvas open.

Parsing heterozygous sequence from base sequence

The initial focus of this project was to make a tool specifically for SeqDoC-style direct comparison of chromatograms. However, having only one way of decoding heterozygous sequences is unnecessarily restrictive for a program that aims to provide all-in-one functionality to its users. Two notable examples of programs that use base sequences to decode chromatograms with heterozygous sequence are PolyPeakParser (Hill et al., 2014) and Indigo (Rausch, 2017). Incorporating this functionality into ChromatVis would

certainly be feasible by adding two components. The first is a dialog for the user to upload a FASTA or text file, or paste a FASTA sequence or plaintext base string. The second is a global alignment algorithm, such as Needleman-Wunsch (Needleman & Wunsch, 1970). Indigo also decodes sequences by comparing chromatograms, as ChromatVis does, and by querying one of several genomes. Implementing genomic references could be a possibility for ChromatVis.

Manual basecall annotation

Having more algorithms to choose from provides easier access as well as added confidence for users by verifying the results of one against another. However, once an algorithm has done most of the labor of detecting peaks and parsing sequences, it's easy for the user to see places where the algorithm made a mistake. As it stands currently, the only thing a ChromatVis user can do in that case is export the basecalls and manually correct the text file. One of the principles of ChromatVis is interactivity, therefore, there are plans to develop additional mouse event triggers and perhaps also a mode for canvas interaction that can be toggled. The implementation of those features and what implementation requires from certain objects are as yet unclear.

Where to find ChromatVis

The latest version of ChromatVis can be found at github.com/pdmcgurk.

Chapter 5: Future Directions

In this dissertation, I have shown coordinated morphogenesis of neural crest-derived tendon progenitors and cranial mesoderm-derived muscles. I have demonstrated that neural crest cells require *cyp26b1* for the aggregation of tendon progenitors and for proper muscle attachment. However, cells expressing *cyp26b1* appear to be separate from tenoblasts during the period when Cyp26 function is required. I hypothesized that *scxa* and *xirp2a*, genes expressed in tendon and in muscle tips, respectively, were critical for patterning craniofacial musculature. Therefore, I generated mutations in each gene using CRISPR/Cas9 technology. Preliminary results suggest that loss of either gene alone or both in tandem does not disrupt normal muscle pattern anywhere in the developing embryo. Similarly, muscle strain caused no significant deleterious effects on muscle attachments in single or double mutants.

Addressing the role of retinoic acid in patterning cranial muscle attachments

The zone of *cyp26b1* expression in neural crest cells abutting tenoblasts at the mandibulohyoid junction appears to promote aggregation and maturation of the tenoblasts. Though the simplest explanation of the *cyp26b1* mutant phenotype is that excess retinoic acid (RA) alters tenoblast behaviors, my work did not directly test the function of RA. Previous work has suggested RA-independent functions of Cyp26b1 (Kumar et al., 2011). However, subsequent studies demonstrate more clearly that Cyp26b1 does regulate RA-dependent sex determination (Saba et al., 2014; Bowles et al., 2016). Having determined the timing of Cyp26b1 function in mandibulohyoid junction formation, it behooves us to directly investigate the RA dependence of the *cyp26b1* mutant phenotype.

My preliminary data suggest that inhibiting RA synthesis with DEAB improves muscle phenotypes, but my results have been inconsistent between experiments. DEAB

experiments performed by Mary Swartz also failed to recapitulate rescue of skeletal phenotypes published previously (Laue et al., 2008). The reason for such inconsistent results is unclear but could relate to its inconsistent inhibitory effects on various Aldh enzymes (Morgan et al., 2015). Direct injection of RA during tenoblast condensation could test the function of RA in this process, although the substantial self-regulatory nature of the pathway (Hernandez et al., 2007) could confound these analyses. Other pharmaceutical treatments are available that might work more directly and consistently, such as retinoic acid receptor (RAR) antagonists like AGN193109 (Johnson et al., 1995). RAR agonists could hypothetically recapitulate the mandibulohyoid junction phenotype in wild-type zebrafish, and RAR-selective agonists like Am580 (RARα; Kagechika et al., 1988), CD 2019 (RARβ; Martin et al., 1992), and CD 437 (RARγ; Martin et al., 1992) might further detail a mechanism. If RA does in fact inhibit mandibulohyoid junction morphogenesis, I propose investigating two models that encompass known mechanisms of developmental regulation by RA: transcriptional regulation and migratory cell guidance.

The canonical model – transcriptional regulation (Fig 5.1A)

RA most often functions by causing retinoic acid receptor binding to DNA response elements where they regulate gene transcription. One possible transcriptional mechanism is direct regulation of Sox9 in tenoblasts adjacent to the ceratohyal cartilage condensations. In mice, dual-positive *Sox9;Scx* cells form bony eminences in mice that connect limb tendons to the bone proper (Blitz et al., 2013; Sugimoto et al., 2013). Tenoblast conditional knockout of *Sox9* via *Scx:Cre* results in a loss of bone eminences, but not a loss of tendon (Blitz et al., 2013; Sugimoto et al., 2013), demonstrating a requirement for *Sox9* in proper tendon-to-bone attachment. In zebrafish, Sox9 is duplicated (Chiang et al., 2001), and existing data suggest that at least *sox9a* is expressed at jaw muscle attachment sites at 62 hpf, after strong expression of *sox9a* in cartilage condensations is turned off (Ignatius et
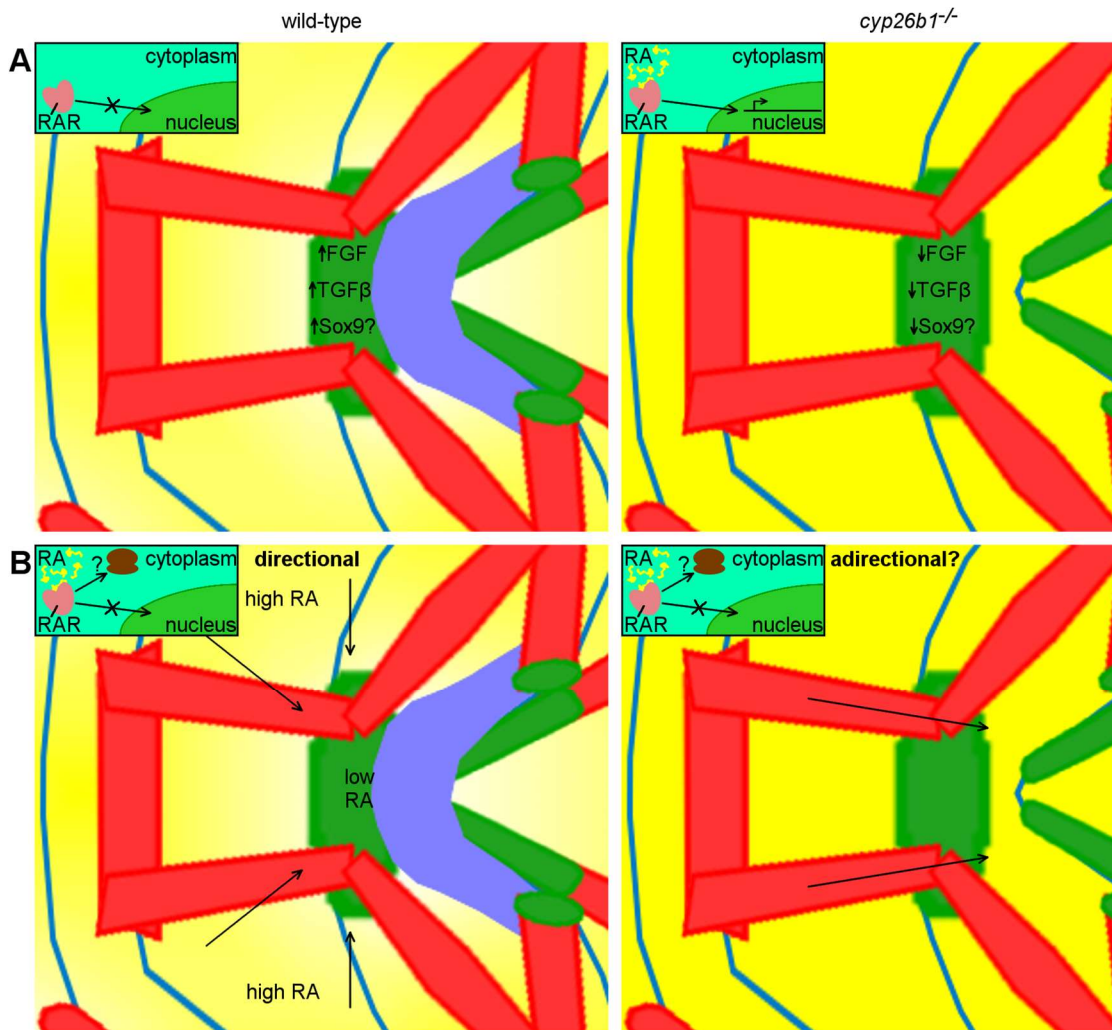
Fig 5.1. Alternate models of RA regulation of mandibulohyoid tenoblast aggregation.

(A) In the canonical model, neural crest cells expressing *cyp26b1* (blue) limit the concentration of retinoic acid (yellow) at the mandibulohyoid junction, where differentiating tenoblasts express high levels of FGF and TGFβ signaling genes, and some may express *sox9a/b*. In *cyp26b1* mutants, higher RA concentrations allow RARs to translocate to the nucleus, where they downregulate the FGF and TGFβ pathways, and possibly *sox9a/b*. (B) In the non-transcriptional model, RA binding to RARs stimulates a cytoplasmic response including protein synthesis, independent of transcription, that drives tenoblasts to migrate down the RA gradient toward *cyp26b1*-expressing neural crest. In *cyp26b1* mutants, the RA gradient around the mandibulohyoid junction is lost, disrupting directional movements of tenoblasts, which may instead migrate along with intermandibularis posterior muscle elongation.

82

al., 2013). Cranial tenoblasts expressing *scxa* are mislocalized in *sox9a;sox9b* double mutants, but appear to differentiate normally in the absence of cartilage or Sox9 protein (Chen & Galloway, 2014). It is unclear, however, if *scxa;sox9a* or *scxa;sox9b* double positive cells occupy muscle attachment sites. If so, RA-mediated regulation of *sox9a* and/or *sox9b* could result in the phenotypes we observe in *cyp26b1* mutants.

There is conflicting evidence concerning RA regulation of *Sox9* expression in cultured mouse chondrocytes. Retinoic acid up-regulates *Sox9* expression in cartilage-derived TC6 cells (Sekiya et al., 2000). However, RA represses *Sox9* in primary chondrocytes (Sekiya et al., 2001). A functional role for RA in Sox9 expression *in vivo* is unknown. However, *cyp26b1* mutants and talarozole-treated embryos provide a way to determine whether RA regulates Sox9 expression in developing cranial tendons.

In silico analysis supports the hypothesis that RA could regulate at least one of the zebrafish *Sox9* paralogs. A canonical "DR5" sequence (5'- RGKTCANNNNNRGKTCA-3', direct repeat with 5-base spacer) for retinoic acid receptor binding is located 5209 bases upstream of the *sox9a* transcription start site. No DR5 sequence is predicted within 10 kilobases upstream or downstream of *sox9b*. Nor is there a predicted DR5 sequence near chicken, mouse, or human *Sox9* orthologs. Thus, RA regulation of *sox9a* may be a zebrafish, or teleost, novelty. A cross-species analysis of *Sox9* responsiveness to RA could shed insight into this exciting possibility.

While Sox9 has a specific role in tendon-to-bone attachment, tendon differentiation more generally requires FGF and TGFβ signaling. RA could potentially interact with these pathways to regulate tendon development. Both pathways promote *Scx* expression in tenoblasts (Edom-Vovard et al., 2002; Brent et al., 2003; Brent and Tabin, 2004; Pryce et al., 2009; Havis et al., 2014). During the muscle-dependent phase of tendon development, TGFβ is sufficient to restore lost expression of *Scx, Tnmd,* and *Thbs2* in paralyzed mouse

limbs (Havis et al., 2016). FGF is sufficient to restore *Scx* expression, but not those markers of tendon differentiation (Havis et al., 2016). Thus, the abnormal behaviors of *scxa*-positive tenoblasts and delayed deposition of Tsp4b (Thbs4) in zebrafish *cyp26b1* mutants would be consistent with RA inhibiting TGFβ signaling.

In other contexts RA is known to inhibit TGFβ signaling. In RA-deficient mice, lungs fail to form due to ectopic TGFβ signaling (F. Chen et al., 2007), but no transcriptional targets of RA were identified in this study. The TGFβ pathway, like the FGF pathway, includes many ligands, receptors, and effectors that could be transcriptionally regulated by RA, making the identification of such targets difficult. Additionally, only a fraction of the total number of *scxa*-expressing cells are disrupted in cyp26b1 mutants. Therefore transcriptional profiling techniques, such as RNA-seq, may not readily identify these targets either. The availability of a TGFβ reporter line, *Tg(12xSBE:EGFP)^{ia16}* (Casari et al., 2014), would aid in determining if RA alters the TGFβ pathway. RA-dependent alterations in TGFβ signaling could be spatiotemporally localized in *12xSBE:EGFP; scxa:mCherry* double transgenics.

If TGFβ signaling is altered by RA, numerous resources exist to explore the functional consequence of TGFβ misregulation. Loss- and gain-of-function approaches could be used in a "recapitulate and rescue" strategy as described above for RARs. For example, if RA inhibits TGFβ in tenoblasts, then treating wild-type embryos with the TGFβ inhibitor SB-431542 (Inman et al., 2002) during the necessary window for Cyp26b1 activity would cause defects in mandibulohyoid tenoblast aggregation and muscle patterning. Conversely, converting the C-terminal SSVS motif of Smad3 to DDVD constitutively activates the TGFβ pathway (X. Liu et al., 1997). Transgenic expression of CA-Smad3 would improve tenoblast and muscle phenotypes in *cyp26b1*-deficient embryos if these phenotypes are due to RA inhibition of TGFβ signaling. A positive result in any of

these experiments would incentivize an investigation into the expression of TGFβ-related genes in *cyp26b1* mutants.

Based on the phenotypes in *cyp26b1* mutants, it is unlikely that FGF signaling plays a role. However, RA has been shown to antagonize FGF signaling (reviewed in Duester, 2013). In zebrafish, for example, *cyp26a1* and *cyp26c1* morpholinos reduce expression of *fgf8a*, partly accounting for second heart field defects (Rydeen & Waxman, 2016). Here again, the "recapitulate and rescue" paradigm is useful to test the involvement of FGF signaling in *cyp26b1* mutant musculoskeletal defects. SU-5402 (Sun et al., 1999) is an FGF pathway inhibitor that could be used in experiments similar to those carried out with talarozole. Replacing S217 and S221 of MEK with glutamic acid residues mimics phosphorylation of these serines and induces ERK/MAPK activity, downstream of FGF (Cowley et al., 1994), allowing for transgenic approaches to activate the FGF pathway. Additionally, time-specific activation of MAPK can be achieved by chemical activation of ERK with U44619 (Karim et al., 1997). Positive results from these experiments would be surprising, but of interest for detailing the relative roles of RA, FGF, and TGFβ signaling in cranial tendons.

A non-transcriptional model – RA-mediated tenoblast chemotaxis (Fig 5.1B)

Traditionally, retinoic acid signaling is known to act over a distance, determining cell fates according to a gradient of concentrations from its source to a "sink" of Cyp26-expressing cells that degrade RA. That tenoblast movements appear to be restricted by a population of *cyp26b1*-expressing cells suggest that perhaps a more short-range signaling interaction directs tenoblast aggregation. RA induces neurite outgrowth in newt and mollusk neurons, whose axon growth cones turn toward a source of RA, indicating that RA acts as a chemoattractant (Dmetrichuk et al., 2005, 2006). Further work demonstrated that mollusk neurites grow and turn in response to RA independently of gene transcription

(Farrar et al., 2009). Rather, protein synthesis and calcium influx in the neurite appears to drive the response to RA (Farrar et al., 2009). Since these axons migrate up a gradient of RA, we wonder if a similar mechanism might drive tenoblasts down an RA gradient toward *cyp26b1*-expressing cells.

The most direct way to assess chemotactic response to RA is *in vitro*. The *scxa:mCherry* reporter allows for sorting of dissociated cranial tenoblasts from zebrafish heads, but may also allow for assays in heterogeneous cultures, obviating the need for cell sorting. Under this model, *scxa:mCherry*-positive cells would hypothetically migrate away from a source of RA. If RA has a chemotactic effect *in vitro*, RA-coated beads could then be applied to disrupt *in vivo* migration. Cranial tenoblast morphogenesis occurs on a relatively small scale, however, perhaps precluding effective use of beads. These challenges could be overcome using localized heat-shock activation of *aldh1a2* to generate an ectopic source of RA.

If alterations to RA signaling disrupt tenoblast behaviors *in vivo*, the chemotaxis model suggest that the effects of RA on tenoblasts are transcription-independent. Unfortunately, transcriptional inhibitors like actinomycin D have only been used in live zebrafish by injection to assess requirements for zygotic transcription. Therefore, it's unclear that *in vivo* analyses would be feasible. However, *in vitro* treatment of cultured cells has been widely demonstrated for actinomycin D, which we would predict inhibits RA-mediated chemotaxis of *scxa:mCherry*-positive tenoblasts.

Determining the RA signaling environment during musculoskeletal development

For either of the above models, it is crucial also to understand the landscape of RA sources and sinks in the jaw to have an idea of what the RA gradient might look like. Around the time of tenoblast aggregation, neural crest cells expressing *aldh1a2* appear to be the source of RA in the head (Grandel et al., 2002; Liang et al., 2008), although more

detailed analyses covering the time window of musculoskeletal attachment are needed. Though *cyp26b1* is functionally involved in craniofacial musculoskeletal attachment, we have not ruled out roles for the other RA degrading enzymes, *cyp26a1* and *cyp26c1*. Because of feedback mechanisms in RA signaling, *cyp26b1* mutants may upregulate expression of *cyp26a1* or *cyp26c1*. *In situ* hybridization will provide insight into the position of RA sources and sinks around the forming mandibulohyoid junction.

Visualizing the actual RA gradient in the zebrafish jaw is a complicated task. Cutting edge technology has the potential to directly visualize autofluorescent endogenous RA *in vivo* with fluorescence lifetime microscopy (Stringari et al., 2011). Another novel technology, Genetically Encoded reporter Probes for RA (GEPRA) allow imaging of endogenous RA gradients indirectly, with some limitations (Shimozono et al., 2013). GEPRA proteins are fusions between RA-binding peptides and fluorophores that change fluorescence resonance energy transfer (FRET) states between their RA-bound and unbound states. Either of these approaches could prove useful in understanding the precise RA signaling landscape in the developing face.

Determining the role of Scx and Xirp2a

No "master regulator" of tendon has yet been discovered. Genetic investigation of tendon development has been complicated by the fact that the tissue develops independent of the transcription factor associated with its specification. *Scx* has long been associated with developing tendons across model systems (Schweitzer et al., 2001). *Scx* is duplicated in zebrafish and it is *scxa* that is expressed by tendons (Chen & Galloway, 2014). Xin actin-binding repeat proteins are expressed in cardiac and skeletal muscles at myotendinous junctions (reviewed in Wang et al., 2014). Zebrafish *xirp1* and the duplicate paralogs of *Xirp2* have semi-overlapping expression in cardiac muscle and skeletal muscle, but *xirp2a* alone is expressed strongly at all myotendinous junctions. We found that mutation of *scxa*

or *xirp2a* singly or together failed to cause overt tendon defects. *Scx* loss of function has been characterized previously in a knockout mouse model. Similar to our results, there are no overt tendon phenotypes in these mice, however, the structure of certain force-transducing tendons is anomalous (Murchison et al., 2007). *Xin* function has also been studied in mouse and chick loss of function models. *Xin/Xirp1* is required for normal heart morphogenesis in chick and mouse development (D. Z. Wang et al., 1999; Gustafson-Wagner et al., 2007), as is *Xirp2* in mouse (Q. Wang et al., 2010), but these mutants lack skeletal muscle defects like our *xirp2a* mutants do. The absence of a developmental phenotype can belie instability at myotendinous junctions, however. Knocking down the tendon extracellular matrix component *tsp4b* in zebrafish did not cause overt muscle or tendon defects (Subramanian & Schilling, 2014). However, when subjected to contractile forces generated by electrical stimulation, muscle fibers broke free of their attachments (Subramanian & Schilling, 2014). Collectively this work suggests that many tendon and muscle tip markers may be regulating the strength of the myotendinous junction.

We sought to test the stability of MTJs and to do so with an easily accessible and quantitative approach. We put force on the MTJ via a convulsant drug, PTZ. Though our work failed to demonstrate functional losses in development or stability of MTJs in *scxa* or *xirp2a* mutants, an important caveat is that we do not yet know that these mutants are true null or loss-of-function alleles. While functioning antibodies for these proteins have not been characterized, the indels present in each mutant would be predicted to result in nonsense mediated decay.Our work will continue to determine whether nonsense-mediated decay occurs, via *in situ* hybridization and PCR, and test whether any functional protein is made should antibodies become available.

If our *scxa* and *xirp2a* mutants are null, further work could delve into their apparent lack of phenotypes. First, these genes have duplicates that could be upregulated to

compensate for their loss of function (Chen & Galloway, 2014; Otten et al., 2012). This would necessitate the generation of double or triple mutant lines. However, it is possible that morpholino knockdown could avoid compensatory mechanisms (Rossi et al., 2015; El-Brolosy & Stainier, 2017), although morpholinos have come under scrutiny (Kok et al., 2015). Second, these genes could simply be dispensable for the normal formation and maintenance of musculoskeletal attachments. There are many dystrophic zebrafish mutants that have developmental phenotypes under endogenous muscle strain (reviewed in Berger & Currie, 2012), whereas the *tsp4b* knockdown model uniquely disrupts MTJ stability without an innate phenotype. Though PTZ offered a simple way to test the effects of strain on the phenotypes of my mutants, it would be useful to test electrical stimulation as a positive control for our analyses.

Interpreting any analysis of *scxa* mutants depends on an understanding of tendon structures in different anatomical contexts (e.g. mandibulohyoid junction vs. sternohyal tendon vs. somite myoseptum) in wild-type and *scxa* mutant zebrafish. How do collagen fibrils organize in these different situations? Is there a phenotype resembling any of what is seen in mouse mutants? Answering these questions would provide a framework to address questions to come, like "Does Tnmd (or some other downstream effector) restore normal tendon structure in *scxa* mutants?"

Elucidating a tendon-specific gene regulatory network

Two resources used in this work would be well applied to examining the gene regulatory function of Scx. Transcriptomic analysis of the *scxa* lineage could be done by crossing the *scxa:mCherry* transgenic line with the *scxa* mutant line. Homozygous *scxa* mutants appear viable and fertile. The large clutch sizes from zebrafish matings, make it feasible to generate mRNA libraries for several stages along the course of tendon

specification, differentiation, and maturity. This would provide a temporal understanding of how Scxa regulates the transcriptome.

Transcriptomic analyses would be well complemented by analysis of *scxa* direct binding targets. The Tol2kit (Kwan, 2007) facilitates the construction of a tagged *scxa* expression vector for immunoprecipitation. Tagged Scxa could be specifically expressed in the *scxa* lineage using transgenic constructs on hand. Data from whole genome ChIP-Seq could be compared against RNA-seq datasets to assess direct targets of Scxa. Alternatively, a candidate gene approach could be done *in vitro* with tagged Scxa protein and synthesized DNA fragments of suspected E-box *scxa*-binding sequences (Y. Liu et al., 1997) from promoters of genes highlighted by transcriptomic analyses. These experiments would reveal much that is currently unknown about the function of tendon's earliest lineage marker.

Clearly, there is much to discover regarding musculoskeletal attachment. To uncover the genes and mechanisms at work, zebrafish research may be key. As established already, forward genetic screens are one strength of the model. Transgenics for tendon and muscle facilitate the detection of defects to these tissues in forward genetic mutants. Now that reverse genetic approaches are easily accessible in zebrafish, candidate genes can be screened with methodologies for quantitative analysis of MTJ integrity.

References

Alberton, P., Popov, C., Prägert, M., Kohler, J., Shukunami, C., Schieker, M., & Docheva, D. (2012). Conversion of human bone marrow-derived mesenchymal stem cells into tendon progenitor cells by ectopic expression of scleraxis. *Stem Cells and Development*, *21*(6), 846–58. https://doi.org/10.1089/scd.2011.0150

Baraban, S. C. (2013). Forebrain Electrophysiological Recording in Larval Zebrafish. *Journal of Visualized Experiments*, *10*(71), 227–236. https://doi.org/10.3791/50104

Berger, J., & Currie, P. D. (2012). Zebrafish models flex their muscles to shed light on muscular dystrophies. *Disease Models & Mechanisms*, *5*, 726–732. https://doi.org/10.1242/dmm.010082

Berger, J., & Currie, P. D. (2013). 503unc, a small and muscle-specific zebrafish promoter. *Genesis (New York, N.Y. : 2000)*, *51*(6), 443–7. https://doi.org/10.1002/dvg.22385

Blitz, E., Sharir, A., Akiyama, H., & Zelzer, E. (2013). Tendon-bone attachment unit is formed modularly by a distinct pool of Scx- and Sox9-positive progenitors. *Development*, *140*(13).

Boglioli, E., & Richard, M. (2015). Rewriting The Book Of Life: A New Era in Precision Gene Editing.

Bowles, J., Feng, C.-W., Miles, K., Ineson, J., Spiller, C., & Koopman, P. (2016). ALDH1A1 provides a source of meiosis-inducing retinoic acid in mouse fetal ovaries. *Nature Communications*, *7*(May 2015), 10845. https://doi.org/10.1038/ncomms10845

Brent, A. E., Braun, T., & Tabin, C. J. (2005). Genetic analysis of interactions between the somitic muscle, cartilage and tendon cell lineages during mouse development. *Development (Cambridge, England)*, *132*(3), 515–28. https://doi.org/10.1242/dev.01605

Brent, A. E., Schweitzer, R., & Tabin, C. J. (2003). A somitic compartment of tendon progenitors. *Cell*, *113*(2), 235–48.

Brent, A. E., & Tabin, C. J. (2004). FGF acts directly on the somitic tendon progenitors through the Ets transcription factors Pea3 and Erm to regulate scleraxis expression. *Development*, *131*(16).

Buckingham, M., & Vincent, S. D. (2009). Distinct and dynamic myogenic populations in the vertebrate embryo. *Current Opinion in Genetics and Development*. https://doi.org/10.1016/j.gde.2009.08.001

Casari, A., Schiavone, M., Facchinello, N., Vettori, A., Meyer, D., Tiso, N., … Argenton, F. (2014). A Smad3 transgenic reporter reveals TGF-beta control of zebrafish spinal cord development. *Developmental Biology*, *396*(1), 81–93.

https://doi.org/10.1016/j.ydbio.2014.09.025

Charvet, B., Ruggiero, F., & Le Guellec, D. (2012). The development of the myotendinous junction. A review. *Muscles, Ligaments and Tendons Journal*, *2*(2), 53–63.

Chen, F., Desai, T. J., Qian, J., Niederreither, K., Lü, J., & Cardoso, W. V. (2007). Inhibition of Tgf beta signaling by endogenous retinoic acid is essential for primary lung bud induction. *Development (Cambridge, England)*, *134*(16), 2969–79. https://doi.org/10.1242/dev.006221

Chen, J. W., & Galloway, J. L. (2014). The development of zebrafish tendon and ligament progenitors. *Development*, *141*(10), 2035–2045. https://doi.org/10.1242/dev.104067

Chen, J. W., King, M. J., Tabin, C. J., & Galloway, J. L. (2017). The mevalonate pathway is a critical regulator of tendon cell speciifcation. *Development*.

Chen, X., Yin, Z., Chen, J., Shen, W., Liu, H., Tang, Q., … Ouyang, H. (2012). Force and scleraxis synergistically promote the commitment of human ES cells derived MSCs to tenocytes. *Scientific Reports*, *2*. https://doi.org/10.1038/srep00977

Cherepanova, O., Orlova, A., Galkin, V. E., Van Der Ven, P. F. M., Fürst, D. O., Jin, J. P., & Egelman, E. H. (2006). Xin-repeats and nebulin-like repeats bind to F-actin in a similar manner. *Journal of Molecular Biology*, *356*(3), 714–723. https://doi.org/10.1016/j.jmb.2005.11.082

Chevalier, B. S., Kortemme, T., Chadsey, M. S., Baker, D., Monnat, R. J., & Stoddard, B. L. (2002). Design, activity, and structure of a highly specific artificial endonuclease. *Molecular Cell*, *10*(4), 895–905. https://doi.org/10.1016/S1097-2765(02)00690-1

Chevallier, A., Kieny, M., & Mauger, A. (1977). Limb-somite relationship: origin of the limb musculature. *Journal of Embryology and Experimental Morphology*, *41*, 245–58.

Chiang, E. F., Pai, C. I., Wyatt, M., Yan, Y. L., Postlethwait, J., & Chung, B. (2001). Two sox9 genes on duplicated zebrafish chromosomes: expression of similar transcription activators in distinct sites. *Developmental Biology*, *231*(1), 149–63. https://doi.org/10.1006/dbio.2000.0129

Couly, G., Coltey, P. M., & Le Douarin, N. M. (1992). The developmental fate of the cephalic mesoderm in quail-chick chimeras. *Development (Cambridge, England)*, *114*(1), 1–15.

Couly, G., Creuzet, S., Bennaceur, S., Vincent, C., & Le Douarin, N. M. (2002). Interactions between Hox-negative cephalic neural crest cells and the foregut endoderm in patterning the facial skeleton in the vertebrate head. *Development (Cambridge, England)*, *129*(4), 1061–1073.

Cowley, S., Paterson, H., Kemp, P., & Marshall, C. J. (1994). Activation of MAP kinase

kinase is necessary and sufficient for PC12 differentiation and for transformation of NIH 3T3 cells. *Cell*, *77*(6), 841–852. https://doi.org/0092-8674(94)90133-3 [pii]

Crowe, M. L. (2005). SeqDoC: rapid SNP and mutation detection by direct comparison of DNA sequence chromatograms. *BMC Bioinformatics*, *6*, 133. https://doi.org/10.1186/1471-2105-6-133

Cserjesi, P., Brown, D., Ligon, K. L., Lyons, G. E., Copeland, N. G., Gilbert, D. J., … Olson, E. N. (1995). Scleraxis: a basic helix-loop-helix protein that prefigures skeletal formation during mouse embryogenesis. *Development*, *121*(4), 1099–1110.

Depew, M. J., Lufkin, T., & Rubenstein, J. L. R. (2002). Specification of jaw subdivisions by Dlx genes. *Science (New York, N.Y.)*, *298*(5592), 381–5. https://doi.org/10.1126/science.1075703

Depew, M. J., Simpson, C. a, Morasso, M., & Rubenstein, J. L. R. (2005). Reassessing the Dlx code: the genetic regulation of branchial arch skeletal pattern and development. *Journal of Anatomy*, *207*(5), 501–561. https://doi.org/10.1111/j.1469-7580.2005.00487.x

Dex, S., Lin, D., Shukunami, C., & Docheva, D. (2016). Tenogenic modulating insider factor: Systematic assessment on the functions of tenomodulin gene. *Gene*. https://doi.org/10.1016/j.gene.2016.04.051

Dmetrichuk, J. M., Carlone, R. L., & Spencer, G. E. (2006). Retinoic acid induces neurite outgrowth and growth cone turning in invertebrate neurons. *Developmental Biology*, *294*(1), 39–49. https://doi.org/10.1016/j.ydbio.2006.02.018

Dmetrichuk, J. M., Spencer, G. E., & Carlone, R. L. (2005). Retinoic acid-dependent attraction of adult spinal cord axons towards regenerating newt limb blastemas in vitro. *Developmental Biology*, *281*(1), 112–120. https://doi.org/10.1016/j.ydbio.2005.02.019

Docheva, D., Hunziker, E. B., Fässler, R., & Brandau, O. (2005). Tenomodulin is necessary for tenocyte proliferation and tendon maturation. *Molecular and Cellular Biology*, *25*(2), 699–705. https://doi.org/10.1128/MCB.25.2.699-705.2005

Duester, G. (2013). Retinoid signaling in control of progenitor cell differentiation during mouse development. *Seminars in Cell and Developmental Biology*. https://doi.org/10.1016/j.semcdb.2013.08.001

Eberhart, J. K., Swartz, M. E., Crump, J. G., & Kimmel, C. B. (2006). Early Hedgehog signaling from neural to oral epithelium organizes anterior craniofacial development. *Development (Cambridge, England)*, *133*, 1069–1077. https://doi.org/10.1242/dev.02281

Edom-Vovard, F., Schuler, B., Bonnin, M.-A., Teillet, M.-A., & Duprez, D. (2002). Fgf4 positively regulates scleraxis and tenascin expression in chick limb tendons. *Developmental Biology*, *247*, 351–366. https://doi.org/10.1006/dbio.2002.0707

El-Brolosy, M. A., & Stainier, D. Y. R. (2017). Genetic compensation: A phenomenon in search of mechanisms. *PLOS Genetics*, *13*(7), e1006780.

Eloy-Trinquet, S., Wang, H., Edom-Vovard, F., & Duprez, D. (2009). Fgf signaling components are associated with muscles and tendons during limb development. *Developmental Dynamics*, *238*(5), 1195–1206. https://doi.org/10.1002/dvdy.21946

Eulitz, S., Sauer, F., Pelissier, M.-C., Boisguerin, P., Molt, S., Schuld, J., … Fürst, D. O. (2013). Identification of Xin-repeat proteins as novel ligands of the SH3 domains of nebulin and nebulette and analysis of their interaction during myofibril formation and remodeling. *Molecular Biology of the Cell*, *24*(20), 3215–26. https://doi.org/10.1091/mbc.E13-04-0202

Farrar, N. R., Dmetrichuk, J. M., Carlone, R. L., & Spencer, G. E. (2009). A novel, nongenomic mechanism underlies retinoic acid-induced growth cone turning. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(45), 14136–14142. https://doi.org/10.1523/JNEUROSCI.2921-09.2009

Fonfara, I., Richter, H., Bratovič, M., Le Rhun, A., & Charpentier, E. (2016). The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature*, 1–19. https://doi.org/10.1038/nature17945

Fratzl, P. (2003). Cellulose and collagen: From fibres to tissues. *Current Opinion in Colloid and Interface Science*. https://doi.org/10.1016/S1359-0294(03)00011-6

Galis, F., & Metz, J. A. J. (2001). Testing the vulnerability of the phylotypic stage: On modularity and evolutionary conservation. *Journal of Experimental Zoology*, *291*(2), 195–204. https://doi.org/10.1002/jez.1069

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

Gaut, L., & Duprez, D. (2016). Tendon development and diseases. *Wiley Interdisciplinary Reviews: Developmental Biology*, *5*(1), 5–23. https://doi.org/10.1002/wdev.201

Grandel, H., Lun, K., Rauch, G.-J., Rhinn, M., Piotrowski, T., Houart, C., … Brand, M. (2002). Retinoic acid signalling in the zebrafish embryo is necessary during pre-segmentation stages to pattern the anterior-posterior axis of the CNS and to induce a pectoral fin bud. *Development*, *129*(12), 2851–2865.

Grenier, J., Teillet, M.-A., Grifone, R., Kelly, R. G., & Duprez, D. (2009). Relationship between neural crest cells and cranial mesoderm during head muscle development. *PloS One*, *4*(2), e4381. https://doi.org/10.1371/journal.pone.0004381

Gros, J., Serralbo, O., & Marcelle, C. (2009). WNT11 acts as a directional cue to organize the elongation of early muscle fibres. *Nature*, *457*(7229), 589–93. https://doi.org/10.1038/nature07564

Gustafson-Wagner, E. A., Sinn, H. W., Chen, Y.-L., Wang, D.-Z., Reiter, R. S., L-C Lin,

J., … J-C Lin, J. (2007). Loss of mXinα, an intercalated disk protein, results in cardiac hypertrophy and cardiomyopathy with conduction defects. *Am J Physiol Heart Circ Physiol*, *293*(5), 2680–2692. https://doi.org/10.1152/ajpheart.00806.2007

Havis, E., Bonnin, M.-A., Esteves de Lima, J., Charvet, B., Milet, C., & Duprez, D. (2016). TGFβ and FGF promote tendon progenitor fate and act downstream of muscle contraction to regulate tendon differentiation during chick limb development. *Development*, *143*(20), 3839–3851. https://doi.org/10.1242/dev.136242

Havis, E., Bonnin, M.-A., Olivera-Martinez, I., Nazaret, N., Ruggiu, M., Weibel, J., … Duprez, D. (2014). Transcriptomic analysis of mouse limb tendon cells during development. *Development*, *141*(19).

Hernandez, R. E., Putzke, A. P., Myers, J. P., Margaretha, L., & Moens, C. B. (2007). Cyp26 enzymes generate the retinoic acid response pattern necessary for hindbrain development. *Development (Cambridge, England)*, *134*(1), 177–87. https://doi.org/10.1242/dev.02706

Heude, E., Bouhali, K., Kurihara, Y., Kurihara, H., Couly, G., Janvier, P., & Levi, G. (2010). Jaw muscularization requires Dlx expression by cranial neural crest cells. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(25), 11441–6. https://doi.org/10.1073/pnas.1001582107

Hill, J. T., Demarest, B. L., Bisgrove, B. W., Su, Y. C., Smith, M., & Yost, H. J. (2014). Poly peak parser: Method and software for identification of unknown indels using sanger sequencing of polymerase chain reaction products. *Developmental Dynamics*, *243*(12), 1632–1636. https://doi.org/10.1002/dvdy.24183

Hinits, Y., Williams, V. C., Sweetman, D., Donn, T. M., Ma, T. P., Moens, C. B., & Hughes, S. M. (2011). Defective cranial skeletal development, larval lethality and haploinsufficiency in Myod mutant zebrafish. *Developmental Biology*, *358*(1), 102–112. https://doi.org/10.1016/j.ydbio.2011.07.015

Hruscha, A., Krawitz, P., Rechenberg, A., Heinrich, V., Hecht, J., Haass, C., … Doudna, J. A. (2013). Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. *Development (Cambridge, England)*, *140*(24), 4982–7. https://doi.org/10.1242/dev.099085

Huang, A. H., Riordan, T. J., Pryce, B., Weibel, J. L., Watson, S. S., Long, F., … Schweitzer, R. (2015). Musculoskeletal integration at the wrist underlies the modular development of limb tendons. *Development (Cambridge, England)*, *142*(14), 2431–41. https://doi.org/10.1242/dev.122374

Huang, A. H., Riordan, T. J., Wang, L., Eyal, S., Zelzer, E., Brigande, J. V., & Schweitzer, R. (2013). Repositioning forelimb superficialis muscles: Tendon attachment and muscle activity enable active relocation of functional myofibers. *Developmental Cell*, *26*(5), 544–551. https://doi.org/10.1016/j.devcel.2013.08.007

Huang, H. T., Brand, O. M., Mathew, M., Ignatiou, C., Ewen, E. P., McCalmon, S. A., & Naya, F. J. (2006). Myomaxin Is a novel transcriptional target of MEF2A that encodes a Xin-related α-actinin-interacting protein. *Journal of Biological Chemistry*, *281*(51), 39370–39379. https://doi.org/10.1074/jbc.M603244200

Ignatius, M. S., Unal Eroglu, A., Malireddy, S., Gallagher, G., Nambiar, R. M., & Henion, P. D. (2013). Distinct Functional and Temporal Requirements for Zebrafish Hdac1 during Neural Crest-Derived Craniofacial and Peripheral Neuron Development. *PLoS ONE*, *8*(5), e63218. https://doi.org/10.1371/journal.pone.0063218

Inman, G. J., Nicolás, F. J., Callahan, J. F., Harling, J. D., Gaster, L. M., Reith, A. D., … Hill, C. S. (2002). SB-431542 is a potent and specific inhibitor of transforming growth factor-beta superfamily type I activin receptor-like kinase (ALK) receptors ALK4, ALK5, and ALK7. *Molecular Pharmacology*, *62*(1), 65–74. https://doi.org/10.1124/mol.62.1.65

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A Programmable Dual-RNA – Guided DNA Endonuclease in Adaptice Bacterial Immunity. *Science (New York, N.Y.)*, *337*(August), 816–822. https://doi.org/10.1126/science.1225829

Johnson, A. T., Klein, E. S., Gillett, S. J., Wang, L., Song, T. K., Pino, M. E., & Chandraratna, R. A. S. (1995). Synthesis and Characterization of a Highly Potent and Effective Antagonist of Retinoic Acid Receptors. *Journal of Medicinal Chemistry*, *38*(24), 4764–4767. https://doi.org/10.1021/jm00024a003

Jozsa, L., & Kannus, P. (1997). Structure and metabolism of normal tendons. In *Human tendons: anatomy physiology and pathology* (Vol. 7, pp. 46–49).

Kagechika, H., Kawachi, E., Hashimoto, Y., Shudo, K., & Himi, T. (1988). Retinobenzoic acids. 1. Structure-activity relationships of aromatic amides with retinoidal activity. *Journal of Medicinal Chemistry*, *31*(11), 2182–2192. https://doi.org/10.1021/jm00119a021

Kahn, J., Shwartz, Y., Blitz, E., Krief, S., Sharir, A., Breitel, D. A., … Zelzer, E. (2009). Muscle Contraction Is Necessary to Maintain Joint Progenitor Cell Fate. *Developmental Cell*, *16*(5), 734–743. https://doi.org/10.1016/j.devcel.2009.04.013

Kardon, G. (1998). Muscle and tendon morphogenesis in the avian hind limb. *Development (Cambridge, England)*, *125*(20), 4019–32.

Karim, S., Berrou, E., Lévy-Toledano, S., Bryckaert, M., & MacLouf, J. (1997). Regulatory role of prostaglandin E2 in induction of cyclo-oxygenase-2 by a thromboxane A2 analogue (U46619) and basic fibroblast growth factor in porcine aortic smooth-muscle cells. *The Biochemical Journal*, *326 (Pt 2*, 593–9.

Kim, H., & Kim, J. S. (2014). A guide to genome engineering with programmable nucleases. *Nat Rev Genet*, *15*(5), 321–334. https://doi.org/10.1038/nrg3686

Kok, F. O., Shin, M., Ni, C. W., Gupta, A., Grosse, A. S., van Impel, A., … Lawson, N. D. (2015). Reverse genetic screening reveals poor correlation between morpholino-induced and mutant phenotypes in zebrafish. *Developmental Cell*, *32*(1), 97–108. https://doi.org/10.1016/j.devcel.2014.11.018

Köntges, G., & Lumsden, A. (1996). Rhombencephalic neural crest segmentation is preserved throughout craniofacial ontogeny. *Development (Cambridge, England)*, *122*(10), 3229–42.

Kozhemyakina, E., Lassar, A. B., & Zelzer, E. (2015). A pathway to bone: signaling molecules and transcription factors involved in chondrocyte development and maturation. *Development (Cambridge, England)*, *142*(5), 817–31. https://doi.org/10.1242/dev.105536

Kumar, S., Chatzi, C., Brade, T., Cunningham, T. J., Zhao, X., & Duester, G. (2011). Sex-specific timing of meiotic initiation is regulated by Cyp26b1 independent of retinoic acid signalling. *Nature Communications*, *2*(May 2010), 151. https://doi.org/10.1038/ncomms1136

Kwan, K. M., Fujimoto, E., Grabher, C., Mangum, B. D., Hardy, M. E., Campbell, D. S., … Chien, C.-B. (2007). The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Developmental Dynamics : An Official Publication of the American Association of Anatomists*, *236*(11), 3088–99. https://doi.org/10.1002/dvdy.21343

Laue, K., Janicke, M., Plaster, N., Sonntag, C., & Hammerschmidt, M. (2008). Restriction of retinoic acid activity by Cyp26b1 is required for proper timing and patterning of osteogenesis during zebrafish development. *Development*, *135*(22), 3775–3787. https://doi.org/10.1242/dev.021238

Lawson, N. D., & Weinstein, B. M. (2002). In vivo imaging of embryonic vascular development using transgenic zebrafish. *Developmental Biology*, *248*(2), 307–18.

Lessman, C. A. (2011). The developing zebrafish (Danio rerio): A vertebrate model for high-throughput screening of chemical libraries. *Birth Defects Research Part C - Embryo Today: Reviews*. https://doi.org/10.1002/bdrc.20212

Li, J., & Dong, S. (2016). The signaling pathways involved in chondrocyte differentiation and hypertrophic differentiation. *Stem Cells International*. https://doi.org/10.1155/2016/2470351

Liang, D., Zhang, M., Bao, J., Zhang, L., Xu, X., Gao, X., & Zhao, Q. (2008). Expressions of Raldh3 and Raldh4 during zebrafish early development. *Gene Expression Patterns*, *8*(4), 248–253. https://doi.org/10.1016/j.gep.2007.12.007

Life Technologies Corporation. (2009). Applied Biosystems Genetic Analysis Data File Format, 1–56.

Lin, C. Y., Chen, W. T., Lee, H. C., Yang, P. H., Yang, H. J., & Tsai, H. J. (2009). The transcription factor Six1a plays an essential role in the craniofacial myogenesis of

zebrafish. *Developmental Biology*, *331*(2), 152–166. https://doi.org/10.1016/j.ydbio.2009.04.029

Liu, X., Sun, Y., Constantinescu, S. N., Karam, E., Weinberg, R. A., & Lodish, H. F. (1997). Transforming growth factor beta-induced phosphorylation of Smad3 is required for growth inhibition and transcriptional induction in epithelial cells. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(20), 10669–74. https://doi.org/10.1073/PNAS.94.20.10669

Liu, Y., Nifuji, A., Tamura, M., Wozney, J. M., Olson, E. N., & Noda, M. (1997). Scleraxis messenger ribonucleic acid is expressed in C2C12 myoblasts and its level is down-regulated by bone morphogenetic protein-2 (BMP2). *Journal of Cellular Biochemistry*, *67*(1), 66–74. https://doi.org/10.1002/(SICI)1097-4644(19971001)67:1<66::AID-JCB7>3.0.CO;2-U

Lovely, C. Ben, Swartz, M. E., McCarthy, N., Norrie, J. L., & Eberhart, J. K. (2016). Bmp signaling mediates endoderm pouch morphogenesis by regulating Fgf signaling in zebrafish. *Development*, *143*(11).

Lu, H. H., & Thomopoulos, S. (2013). Functional Attachment of Soft Tissues to Bone: Development, Healing, and Tissue Engineering. *Annual Review of Biomedical Engineering*, *15*(1), 201–226. https://doi.org/10.1146/annurev-bioeng-071910-124656

Martin, B., Bernardon, J. M., Cavey, M. T., Bernard, B., Carlavan, I., Charpentier, B., … Reichert, U. (1992). Selective synthetic ligands for human nuclear retinoic acid receptors. *Skin Pharmacol*, *5*(1), 57–65.

Medeiros, D. M., & Crump, J. G. (2012). New perspectives on pharyngeal dorsoventral patterning in development and evolution of the vertebrate jaw. *Developmental Biology*. https://doi.org/10.1016/j.ydbio.2012.08.026

Mendias, C. L., Gumucio, J. P., Bakhurin, K. I., Lynch, E. B., & Brooks, S. V. (2012). Physiological loading of tendons induces scleraxis expression in epitenon fibroblasts. *Journal of Orthopaedic Research*, *30*(4), 606–612. https://doi.org/10.1002/jor.21550

Morgan, C. A., Parajuli, B., Buchman, C. D., Dria, K., & Hurley, T. D. (2015). N,N-diethylaminobenzaldehyde (DEAB) as a substrate and mechanism-based inhibitor for human ALDH isoenzymes. *Chemico-Biological Interactions*, *234*, 18–28. https://doi.org/10.1016/j.cbi.2014.12.008

Murchison, N. D., Price, B. A., Conner, D. A., Keene, D. R., Olson, E. N., Tabin, C. J., & Schweitzer, R. (2007). Regulation of tendon differentiation by scleraxis distinguishes force-transmitting tendons from muscle-anchoring tendons. *Development*, *134*(14), 2697–2708. https://doi.org/10.1242/dev.001933

Nilsson, M. I., Nissar, A. A., Al-Sajee, D., Tarnopolsky, M. A., Parise, G., Lach, B., … Hawke, T. J. (2013). Xin is a marker of skeletal muscle damage severity in

myopathies. *American Journal of Pathology*, *183*(6), 1703–1709. https://doi.org/10.1016/j.ajpath.2013.08.010

Nissar, A. A., Zemanek, B., Labatia, R., Atkinson, D. J., van der Ven, P. F. M., Furst, D. O., & Hawke, T. J. (2012). Skeletal muscle regeneration is delayed by reduction in Xin expression: consequence of impaired satellite cell activation? *AJP: Cell Physiology*, *302*(1), C220–C227. https://doi.org/10.1152/ajpcell.00298.2011

Noden, D. M. (1983a). The embryonic origins of avian cephalic and cervical muscles and associated connective tissues. *American Journal of Anatomy*, *168*(3), 257–276. https://doi.org/10.1002/aja.1001680302

Noden, D. M. (1983b). The role of the neural crest in patterning of avian cranial skeletal, connective, and muscle tissues. *Developmental Biology*, *96*(1), 144–65.

Noden, D. M. (1988). Interactions and fates of avian craniofacial mesenchyme. *Development (Cambridge, England)*, *103 Suppl*, 121–40.

Northcutt, R. G. (2005). The new head hypothesis revisited. In *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* (Vol. 304, pp. 274–297). https://doi.org/10.1002/jez.b.21063

O'Haver, T. C. (2014). A Pragmatic Introduction to Signal Processing. https://doi.org/ISBN: 9781533372857

Okano, J., Kimura, W., Papaionnou, V. E., Miura, N., Yamada, G., Shiota, K., & Sakai, Y. (2012). The regulation of endogenous retinoic acid level through CYP26B1 is required for elevation of palatal shelves. *Developmental Dynamics*, *241*(11), 1744–1756. https://doi.org/10.1002/dvdy.23862

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, *9*(1), 62–66. https://doi.org/10.1109/TSMC.1979.4310076

Otten, C., van der Ven, P. F., Lewrenz, I., Paul, S., Steinhagen, A., Busch-Nentwich, E., … Abdelilah-Seyfried, S. (2012). Xirp proteins mark injured skeletal muscle in zebrafish. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0031041

Pacholsky, D., Vakeel, P., Himmel, M., Löwe, T., Stradal, T., Rottner, K., … van der Ven, P. F. M. (2004). Xin repeats define a novel actin-binding motif. *Journal of Cell Science*, *117*(22).

Piotrowski, T., Schilling, T. F., Brand, M., Jiang, Y. J., Heisenberg, C. P., Beuchle, D., … Nüsslein-Volhard, C. (1996). Jaw and branchial arch mutants in zebrafish II: anterior arches and cartilage differentiation. *Development (Cambridge, England)*, *123*, 345–56.

Preibisch, S., Saalfeld, S., & Tomancak, P. (2009). Globally optimal stitching of tiled 3D microscopic image acquisitions. *Bioinformatics*, *25*(11), 1463–1465. https://doi.org/10.1093/bioinformatics/btp184

Pryce, B. A., Brent, A. E., Murchison, N. D., Tabin, C. J., & Schweitzer, R. (2007). Generation of transgenic tendon reporters, ScxGFP and ScxAP, using regulatory elements of the scleraxis gene. *Developmental Dynamics*, *236*(6), 1677–1682. https://doi.org/10.1002/dvdy.21179

Pryce, B. A., Watson, S. S., Murchison, N. D., Staverosky, J. A., Dünker, N., & Schweitzer, R. (2009). Recruitment and maintenance of tendon progenitors by TGFbeta signaling are essential for tendon formation. *Development (Cambridge, England)*, *136*(8), 1351–61. https://doi.org/10.1242/dev.027342

Rausch, T. (2017). Indigo: InDel Discovery in Chromatogram traces obtained from Sanger sequencing of PCR products. Retrieved June 9, 2017, from http://gear.embl.de/indigo

Rodriguez-Guzman, M., Montero, J. A., Santesteban, E., Gañan, Y., Macias, D., & Hurle, J. M. (2007). Tendon-muscle crosstalk controls muscle bellies morphogenesis, which is mediated by cell death and retinoic acid signaling. *Developmental Biology*, *302*(1), 267–80. https://doi.org/10.1016/j.ydbio.2006.09.034

Rossi, A., Kontarakis, Z., Gerri, C., Nolte, H., Hölper, S., Krüger, M., & Stainier, D. Y. R. (2015). Genetic compensation induced by deleterious mutations but not gene knockdowns. *Nature*, *Aug 13*(524), 230–3. https://doi.org/10.1038/nature14580

Rot-Nikcevic, I., Reddy, T., Downing, K. J., Belliveau, A. C., Hallgrímsson, B., Hall, B. K., & Kablar, B. (2006). Myf5 −/− :MyoD −/− amyogenic fetuses reveal the importance of early contraction and static loading by striated muscle in mouse skeletogenesis. *Development Genes and Evolution*, *216*(1), 1–9. https://doi.org/10.1007/s00427-005-0024-9

Ruhin, B., Creuzet, S., Vincent, C., Benouaiche, L., Le Douarin, N. M., & Couly, G. (2003). Patterning of the hyoid cartilage depends upon signals arising from the ventral foregut endoderm. *Developmental Dynamics*, *228*(2), 239–246. https://doi.org/10.1002/dvdy.10380

Rydeen, A. B., & Waxman, J. S. (2016). Cyp26 Enzymes Facilitate Second Heart Field Progenitor Addition and Maintenance of Ventricular Integrity. *PLoS Biology*, *14*(11). https://doi.org/10.1371/journal.pbio.2000504

Saba, R., Wu, Q., & Saga, Y. (2014). CYP26B1 promotes male germ cell differentiation by suppressing STRA8-dependent meiotic and STRA8-independent mitotic pathways. *Developmental Biology*, *389*(2), 173–181. https://doi.org/10.1016/j.ydbio.2014.02.013

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, *74*(12), 5463–7. https://doi.org/10.1073/pnas.74.12.5463

Schilling, T. F., & Kimmel, C. B. (1997). Musculoskeletal patterning in the pharyngeal segments of the zebrafish embryo. *Development (Cambridge, England)*, *124*(15),

2945–60.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., … A., C. (2012). Fiji: an open source platform for biological image analysis. *Nature Methods*, *9*(7), 676–682. https://doi.org/10.1038/nmeth.2019.Fiji

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, *9*(7), 671–675. https://doi.org/10.1038/nmeth.2089

Schweitzer, R., Chyung, J. H., Murtaugh, L. C., Brent, A. E., Rosen, V., Olson, E. N., … Tabin, C. J. (2001). Analysis of the tendon cell fate using Scleraxis, a specific marker for tendons and ligaments. *Development (Cambridge, England)*, *128*(19), 3855–66.

Schweitzer, R., Zelzer, E., & Volk, T. (2010). Connecting muscles to tendons: tendons and musculoskeletal development in flies and vertebrates. *Development (Cambridge, England)*, *137*(17), 2807–2817. https://doi.org/10.1242/dev.047498

Sekiya, I., Koopman, P., Tsuji, K., Mertin, S., Harley, V., Yamada, Y., … Noda, M. (2001). Transcriptional suppression ofSox9 expression in chondrocytes by retinoic acid. *Journal of Cellular Biochemistry*, *81*(S36), 71–78. https://doi.org/10.1002/jcb.1077

Sekiya, I., Tsuji, K., Koopman, P., Watanabe, H., Yamada, Y., Shinomiya, K., … Noda, M. (2000). SOX9 enhances aggrecan gene promoter/enhancer activity and is up-regulated by retinoic acid in a cartilage-derived cell line, TC6. *The Journal of Biological Chemistry*, *275*(15), 10738–44. https://doi.org/10.1074/JBC.275.15.10738

Sharir, A., Stern, T., Rot, C., Shahar, R., & Zelzer, E. (2011). Muscle force regulates bone shaping for optimal load-bearing capacity during embryogenesis. *Development*, *138*(15).

Sheehan-Rooney, K., Swartz, M. E., Zhao, F., Liu, D., & Eberhart, J. K. (2013). Ahsa1 and Hsp90 activity confers more severe craniofacial phenotypes in a zebrafish model of hypoparathyroidism, sensorineural deafness and renal dysplasia (HDR). *Disease Models & Mechanisms*, *6*(5), 1285–91. https://doi.org/10.1242/dmm.011965

Shimozono, S., Iimura, T., Kitaguchi, T., Higashijima, S.-I., & Miyawaki, A. (2013). Visualization of an endogenous retinoic acid gradient across embryonic development. *Nature*, *496*(7445), 363–6. https://doi.org/10.1038/nature12037

Shu, D.-G., Luo, H.-L., Conway Morris, S., Zhang, X.-L., Hu, S.-X., Chen, L.-Z., … Chen, L.-Z. (1999). Lower Cambrian vertebrates from south China. *Nature*, *402*(6757), 42–46. https://doi.org/10.1038/46965

Shwartz, Y., Farkas, Z., Stern, T., Aszódi, A., & Zelzer, E. (2012). Muscle contraction controls skeletal morphogenesis through regulation of chondrocyte convergent

extension. *Developmental Biology*, *370*(1), 154–163.
https://doi.org/10.1016/j.ydbio.2012.07.026

Smith, J., Grizot, S., Arnould, S., Duclert, A., Epinat, J. C., Chames, P., … Duchateau, P. (2006). A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. *Nucleic Acids Research*, *34*(22). https://doi.org/10.1093/nar/gkl720

Spoorendonk, K. M., Peterson-Maduro, J., Renn, J., Trowe, T., Kranenbarg, S., Winkler, C., & Schulte-Merker, S. (2008). Retinoic acid and Cyp26b1 are critical regulators of osteogenesis in the axial skeleton. *Development (Cambridge, England)*, *135*(22), 3765–74. https://doi.org/10.1242/dev.024034

Stringari, C., Cinquin, A., Cinquin, O., Digman, M. A., Donovan, P. J., & Gratton, E. (2011). Phasor approach to fluorescence lifetime microscopy distinguishes different metabolic states of germ cells in a live tissue. *Nat Acad Sci Proc*, *108*(33), 13582–13587. https://doi.org/10.1073/pnas.1108161108

Stucky, B. J. (2012). Seqtrace: A graphical tool for rapidly processing DNA sequencing chromatograms. *Journal of Biomolecular Techniques*, *23*(3), 90–93. https://doi.org/10.7171/jbt.12-2303-004

Subramanian, A., & Schilling, T. F. (2014). Thrombospondin-4 controls matrix assembly during development and repair of myotendinous junctions. *eLife*, *2014*(3). https://doi.org/10.7554/eLife.02372

Sugimoto, Y., Takimoto, A., Akiyama, H., Kist, R., Scherer, G., Nakamura, T., … Shukunami, C. (2013). Scx+/Sox9+ progenitors contribute to the establishment of the junction between cartilage and tendon/ligament. *Development*.

Sun, L., Tran, N., Liang, C., Tang, F., Rice, A., Schreck, R., … Tang, C. (1999). Design, synthesis, and evaluations of substituted 3-[(3- or 4- carboxyethylpyrrol-2-yl)methylidenyl]indolin-2-ones as inhibitors of VEGF, FGF, and PDGF receptor tyrosine kinases. *Journal of Medicinal Chemistry*, *42*(25), 5120–5130. https://doi.org/10.1021/jm9904295

Talbot, J. C., & Amacher, S. L. (2014). A streamlined CRISPR pipeline to reliably generate zebrafish frameshifting alleles. *Zebrafish*, *11*(6), 583–5. https://doi.org/10.1089/zeb.2014.1047

Tavares, A. L. P., Cox, T. C., Maxson, R. M., Ford, H. L., & Clouthier, D. E. (2017). Negative regulation of endothelin signaling by SIX1 is required for proper maxillary development. *Development*, *144*(11).

Thisse, B., Pfumio, S., Fürthauer, M., B., L., Heyer, V., Degrave, A., … Thisse, C. (2001). Expression of the zebrafish genome during embryogenesis. *ZFIN Online Publication*. https://doi.org/cb646

Tidball, J. G., & Lin, C. (1989). Structural changes at the myogenic cell surface during the formation of myotendinous junctions. *Cell and Tissue Research*, *257*(1), 77–84.

https://doi.org/10.1007/BF00221636

Tokita, M., & Schneider, R. A. (2009). Developmental origins of species-specific muscle pattern. *Developmental Biology*, *331*(2), 311–325. https://doi.org/10.1016/j.ydbio.2009.05.548

Trainor, P. A., Tan, S.-S., & Tam, P. P. L. (1994). Cranial paraxial mesoderm: regionalisation of cell fate and impact on craniofacial development in mouse embryos. *Development (Cambridge, England)*, *120*(9), 2397–2408.

Wang, D. Z., Reiter, R. S., Lin, J. L., Wang, Q., Williams, H. S., Krob, S. L., … Lin, J. J. (1999). Requirement of a novel gene, Xin, in cardiac morphogenesis. *Development (Cambridge, England)*, *126*, 1281–94.

Wang, Q., Lin, J. L. C., Erives, A. J., Lin, C. I., & Lin, J. J. C. (2014). New insights into the roles of xin repeat-containing proteins in cardiac development, function, and disease. *International Review of Cell and Molecular Biology*, *310*, 89–128. https://doi.org/10.1016/B978-0-12-800180-6.00003-7

Wang, Q., Lin, J. L. C., Reinking, B. E., Feng, H. Z., Chan, F. C., Lin, C. I., … Lin, J. J. C. (2010). Essential roles of an intercalated disc protein, mXinβ, in postnatal heart growth and survival. *Circulation Research*, *106*(9), 1468–1478. https://doi.org/10.1161/CIRCRESAHA.109.212787

Westerfield, M. (2007). The Zebrafish Book. A Guide for the Laboratory Use of Zebrafish (Danio rerio), 5th Edition. *University of Oregon Press, Eugene (Book)*.

Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., … Zhang, F. (2015). Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell*, *163*(3), 759–771. https://doi.org/10.1016/j.cell.2015.09.038

Vita

Patrick Dooling McGurk was born January 31, 1987, in Omaha, Nebraska. He spent his childhood overachieving in most school subjects and underachieving in most sports. He got involved with the Boy Scouts of America in the first grade and learned many things over the years, most importantly to try everything. Thus, when his love for theatre led to a scheduling conflict his senior year of high school, he confidently unenrolled from advanced chemistry and filled the empty block with a college-level biology course. There, Patrick met Iris Flournoy, the teacher who launched him on a career path.

Patrick enrolled at the University of Kansas in the fall of 2006. In his time there, he studied molecular biology, chemistry, and German, among other things. He also ritually attended basketball games at Allen Fieldhouse and reveled with thousands of other students and fans on Massachusetts Street the night our Jayhawks won the national championship in 2008. That fall, Patrick got a job in a lab for the first time, scrubbing dishes and making *C. elegans* media for Dr. Erik Lundquist. Patrick was able to work his way into a spot at the bench, under the mentorship of Rafael Demarco. Patrick was awarded departmental honors for his senior thesis on their axon guidance research. Though he graduated with a B.S. in Microbiology in May 2010, the mysteries of development intrigued him.

Patrick arrived in Austin on a hot August day, a few months after graduating. He began graduate school at the University of Texas at Austin and explored his options in microbiology and neurobiology labs before joining the lab of Dr. Johann Eberhart in May 2011 to study the development of the zebrafish head. He married the literal girl next door in October 2014. Patrick and his wife, Elizabeth, welcomed their first child, Clara Elizabeth, to the world in February 2017.

Permanent email: pdmcgurk [at] gmail [dot] com

This dissertation was typed by Patrick Dooling McGurk.