Transcriptional Regulation of Esrp1 and its Role in Craniofacial Morphogenesis

Citation

Macias Trevino, Claudio. 2022. Transcriptional Regulation of Esrp1 and its Role in Craniofacial Morphogenesis. Doctoral dissertation, Harvard University Graduate School of Arts and Sciences.

HARVARD UNIVERSITY
Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Division of Medical Sciences

in the subject of Biological and Biomedical Sciences

have examined a dissertation entitled

Transcriptional regulation of esrp1 and its role in craniofacial morphogenesis

presented by Claudio Macias Trevino

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature: Matthew Harris (Jun 13, 2022 14:10 EDT)

Typed Name: Dr. Matthew Harris

Signature: Trista E North (Jun 13, 2022 14:32 EDT)

Typed Name: Dr. Trista North

Signature: april craft
april craft (Jun 13, 2022 14:42 EDT)

Typed Name: Dr. April Craft

Signature: Shannon Fisher (Jun 13, 2022 15:58 EDT)

Typed Name: Dr. Shannon Fisher

Date: June 08, 2022

Transcriptional Regulation of Esrp1 and its Role in Craniofacial Morphogenesis

A dissertation presented by

Claudio Macias Trevino

To

The Division of Medical Sciences

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

In the subject of

Biological and Biomedical Sciences

Harvard University

Cambridge, Massachusetts

June 8th, 2022

Dissertation Advisor: Eric C. Liao Claudio Macias Trevino

Transcriptional Regulation of Esrp1 and its Role in Craniofacial Morphogenesis

Abstract

Alternative splicing creates different messenger RNAs from the same genetic locus, an essential regulatory mechanism that increases protein diversity and regulates key developmental processes. Recent studies have identified epithelial-specific alternative splicing regulatory factors ESRP1 and ESRP2 as crucial proteins that regulate craniofacial morphogenesis during embryonic development. Animal models carrying *Esrp1* and *Esrp2* loss-of-function mutations lead to a cleft palate phenotype or anterior neurocranium cleft phenotype in mice and zebrafish. However, no mechanism has been proposed to describe how disruption of epithelial-specific alternative splicing events leads to morphogenic defects during development. Additionally, there is a need for experimental models to functionally annotate gene variants in *ESRP1* and *ESRP2* found in human cohorts of cleft lip and/or palate (CL/P) congenital malformations. Using RNA-sequencing of *irf6*$^{-/-}$ embryos, RNAscope in-situ hybridization, and phenotypic analysis of optogenetic *irf6*$^{-/-}$ and *esrp1/2* DKO zebrafish embryos, we established a regulatory axis between the key transcriptional regulator of epithelial maturation and craniofacial development, Irf6, and the Esrp1/2 proteins. In order to elucidate the mechanism for the ANC cleft phenotype, we employed the Tg(sox10:kaede) photoconvertible zebrafish line to lineage trace frontonasal neural crest cells that give rise to the medial ANC in developing zebrafish. We found that cranial neural crest cells in *esrp1/2* DKO fish migrate to the medial ANC but fail to differentiate into chondrocytes. Meanwhile, we developed an *esrp2* morphant assay that phenocopies the *esrp1/2* cleft ANC phenotype to test mRNAs encoding for 18 *esrp1* or *esrp2* gene variants in a phenotypic rescue assay. We found that only 4 variants found in the RNA-Recognition Motifs 1 and 3 of *esrp2* are pathogenic and fail to rescue the cleft ANC phenotype in our morphant assay. We further validated our data by performing a molecular splicing assay for putative *Esrp1/2* target genes *Arhgef11* and *Ctnnd1* in *Esrp1/2* DKO mouse PY2T cells that confirmed our zebrafish morphant assay results. Lastly, we defined alternatively spliced patterns for *Ctnnd1* in WT and *esrp1/2* DKO zebrafish and show that *Ctnnd1* transcripts partially rescue the zebrafish cleft ANC phenotype. This is a critical finding that confirms *Ctnnd1* as a target gene of esrp1 and esrp2 and implicates the importance of intercellular junctions in craniofacial morphogenesis and development of the cleft ANC phenotype.

Table of Contents

Chapter III: An Irf6-Esrp1/2 regulatory axis controls midface morphogenesis in vertebrates 46

Acknowledgments

First, I would like to convey my sincerest gratitude to my dissertation advisor, Dr. Eric Chienwei Liao, for his support and mentorship throughout the past few years. He has encouraged me to pursue excellence in my scientific and clinical endeavors, provided a rich environment to germinate as a budding scientist, and did not hesitate to push me to reach ever-higher goals that were a testament to the potential he saw in me. As I look back on my training, I feel that I am a better scientist, clinician, and even, a better person because of his mentorship.

I would also like to thank the members of my dissertation advisory committee: Dr. Matthew Warman, Dr. Matthew Harris, and Dr. Allan Goldstein. Despite many challenges, including the rising COVID-19 pandemic during my PhD years, they were present to provide direction, advice, and collaboration in an already challenging endeavor. I would also like to thank my colleagues at MGH, the center for regenerative medicine, for providing a top-notch environment filled with engaging discussion, collaborative efforts, and a pleasant environment I was proud to be part of. Specifically, I would like to thank Casey Tsimbal, with whom I had the pleasure of working for the latter part of my PhD. I can’t wait to see your next steps and future contributions to science. I would also like to thank collaborators from other institutions, particularly Dr. Russ Carstens and Dr. Elizabeth Leslie, who taught me that research is truly a global effort.

Lastly, I thank my mother, Thelma Iliana Trevino Santoy, my brother, Salvador Macias Trevino, and my family. They taught me many valuable life lessons and went through both hardships and some of my fondest memories with me. I strive to make them proud daily. I would not be where I am without the many sacrifices made by my family to help me develop into the person I am today. As I look back to the end of my graduate training, I can’t help but think about the way Henri Poincaré described mathematics: *The mathematician does not study mathematics because it is useful; he studies it because he delights in it and he delights in it because it is beautiful.* I feel the same way about science, and I look forward to the beauty of science and biology yet to be discovered.

Chapter I: General Introduction

Craniofacial development in the early embryo

Craniofacial structures are an evolutionary development that encompasses higher cognitive and sensory needs of complex organisms in response to increasing adaptations to environmental insults and higher order social interactions. At the simplest level, the craniofacial skeleton is divided dorsoventrally to compartmentalize the neurocranium as a supportive capsule to the brain, and the viscerocranium to support the pharynx[1]. Historically, higher complexities of the craniofacial skeleton were described as the result of segmentation to house protect vital and sensory needs of increasingly complex organisms, such as occipital protrusions for visual system, middle ear ossicles for the auditory system, or accommodating and compartmentalizing alimentary and respiratory tracts[2]. Embryologically, tissues of the head and neck are determined during the setting of the antero-posterior (AP) axis in the developing embryonic hindbrain. Early segmentation of the hindbrain shapes numbered rhombomeres of neuroectoderm that will give rise to cranial nerves that innervate the face and neck. Rhombomeres vary in number across species, with up to 8 rhombomeres (r1-r8) being described in vertebrates. Molecularly, the early setting of the AP axis in vertebrates is largely driven by an early gradient of ventral sonic-hedgehog (Shh) signals that give rise to a patterned expression of the homeodomain *Hox* genes[3, 4]. The combinatorial Hox pattern gives rise to an ordered structure of the developing rhombomeres, except for r1 that is largely devoid of Hox expression. Along with rhombomeres, the neuroectoderm of the neural tube gives rise to a migratory cell population that is also segmentally derived to give rise to most of the facial cartilages, the neural crest.

Neural crest cells (NCCs), termed by some as the “fourth germ layer”, give rise to a variety of cell types that would normally be restricted to one of the three canonical germ layers: ectoderm, mesoderm, and endoderm[5, 6]. These can include neural ganglia, peripheral neurons, and pigmented melanocytes (traditionally ascribed to ectodermal origin), in addition to chondrocytes and osteocytes (traditionally ascribed to mesodermal origin). Initially developing at the border of neuroectoderm and non-neural ectoderm in the developing neural tube, NCCs undergo epithelial-mesenchymal transformation (EMT) that allows them to migrate away from the neural tube and into their intended targets[5]. Furthermore, NCCs undergo early specification determined by their relative origin in the AP axis of the developing embryo. NCCs are broadly categorized by their eventual destination, where rostral populations of NCCs that give rise to craniofacial structures in the head and neck are called cranial neural crest cells (CNCCs) while caudal populations of NCCs are subdivided into trunk neural crest cells, sacral (or vagal) neural crest cells, and cardiac neural crest cells. Migrating CNCCs originating in the hindbrain, along with paraxial mesoderm on either side of the neural tube, migrate ventrally to give rise to outpouchings of tissue that will form much of the developing facial cartilage, nerves, muscle, and vasculature[7, 8]. These structures are called the pharyngeal arches (PAs, also called branchial arches) separated by pharyngeal clefts (PCs) that, along with rhombomeres, form a highly conserved segmented facial pattern across vertebrates (**Figure 1**A). Simultaneously, outpouchings develop in the lateral walls of the developing pharynx, called pharyngeal pouches, that are comprised of endoderm that will develop into the luminal structures of the pharynx. Altogether, vertebrates generate at least four pharyngeal arches and as many as six arches in mammals, generally numbered PA1 through PA6, although the fifth arch regresses during mammalian development. Each arch is associated with a set of bones, musculature, arteries, and nerves, with the corresponding developing vascular system named the aortic arches. The first pharyngeal arch (populated by CNCCs from r1 and r2) gives rise to two structures called the maxillary and mandibular prominences, with the maxillary prominence being the dorsal portion relative to the mandibular prominence. The maxillary prominence gives rise to the maxilla (upper jaw in most vertebrates) that include the lateral aspects of the secondary palate in mammals, in addition to the incus bone in the middle ear. The mandibular prominence gives rise to the mandible via the development of Meckel's cartilage (**Figure 1**B).

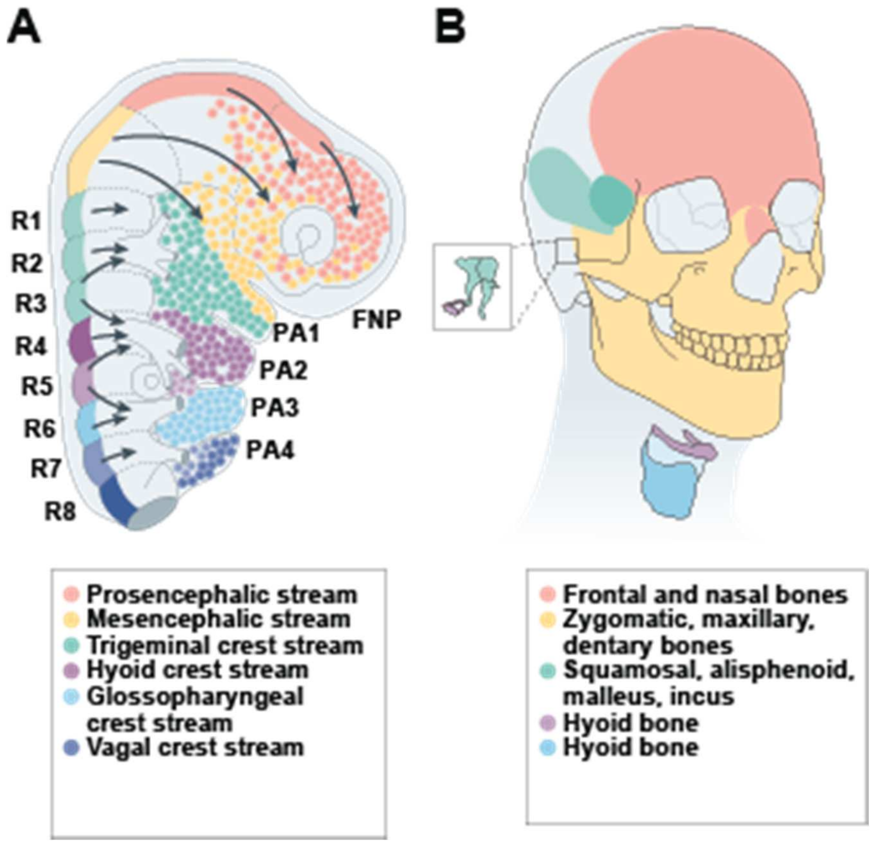


Figure 1: Cranial neural crest cell (CNCC) in craniofacial development

(A) Schematic illustrating the origin of CNCCs in the posterior neural plate and rhombomeres to populate the developing frontonasal process (FNP) and pharyngeal arches (PA1-PA4) through a series of streams. (B) Illustration of structures from the facial skeleton derived from neural crest streams. Figure adapted from Martik & Bronner, *Nature Reviews Neuroscience.*[6]

While CNCCs are restricted to their anatomic destination based on their relative origin in the A-P axis and rhombomeric origin, their homeodomain gene expression patterns approximate, but do not replicate the exact patterning observed in numbered rhombomeres. This is critically important for CNCCs populating PA1 that originate from r1 and r2 that are devoid of Hox expression patterns, despite the presence of Hoxa2 in the boundary between the first and second rhombomeres[4, 9, 10]. Furthermore, this lack of Hox expression in CNCCs of PA1 is hypothesized to be a driving factor behind the plasticity of neural crest that gives rise to large vertebrate diversity in facial forms, in addition to differences in Hox expression patterns along the AP axis between vertebrate species[11-15]. Thus, one puzzling aspect of craniofacial morphogenesis is to determine how seemingly plastic CNCCs devoid of Hox gene expression become terminally differentiated into craniofacial structures based on their end location[16, 17]. One possibility is that local positional and trophic cues in the developing face guide morphogenesis and CNCC differentiation to form the cartilaginous structure of the face.

Craniofacial Morphogenesis and Palatogenesis

Embryonic morphogenesis of facial structures occurs after the development of the pharyngeal arches through the formation of five prominences. A single frontonasal prominence (FNP) originates from CNCCs in the forebrain that gives rise to symmetric medial nasal prominences and lateral nasal prominences that flank the developing nasal placodes. In addition, outgrowths from PA1 give rise to bilateral mandibular prominences (MNP) and maxillary prominences (MXP)[18] (**Figure 2**A). After the prominences develop, three major morphogenetic events occur. First, the medial and nasal prominences of the

FNP fuse to form the nasal pits that give rise to the developing nostrils (**Figure 2**B). Additionally, the medial nasal prominence of the FNP extends ventrally and fuses to the MXPs to give rise to the upper lips and philtrum in the intermaxillary segment. Lastly, the MNPs converge and fuse at the midline to give rise to the mandibular Meckel's cartilages (**Figure 2**C). Ultimately, the fusion events between the MXPs and FNP will create a sheet of tissue to separate the nasal and oral cavities, the palate.

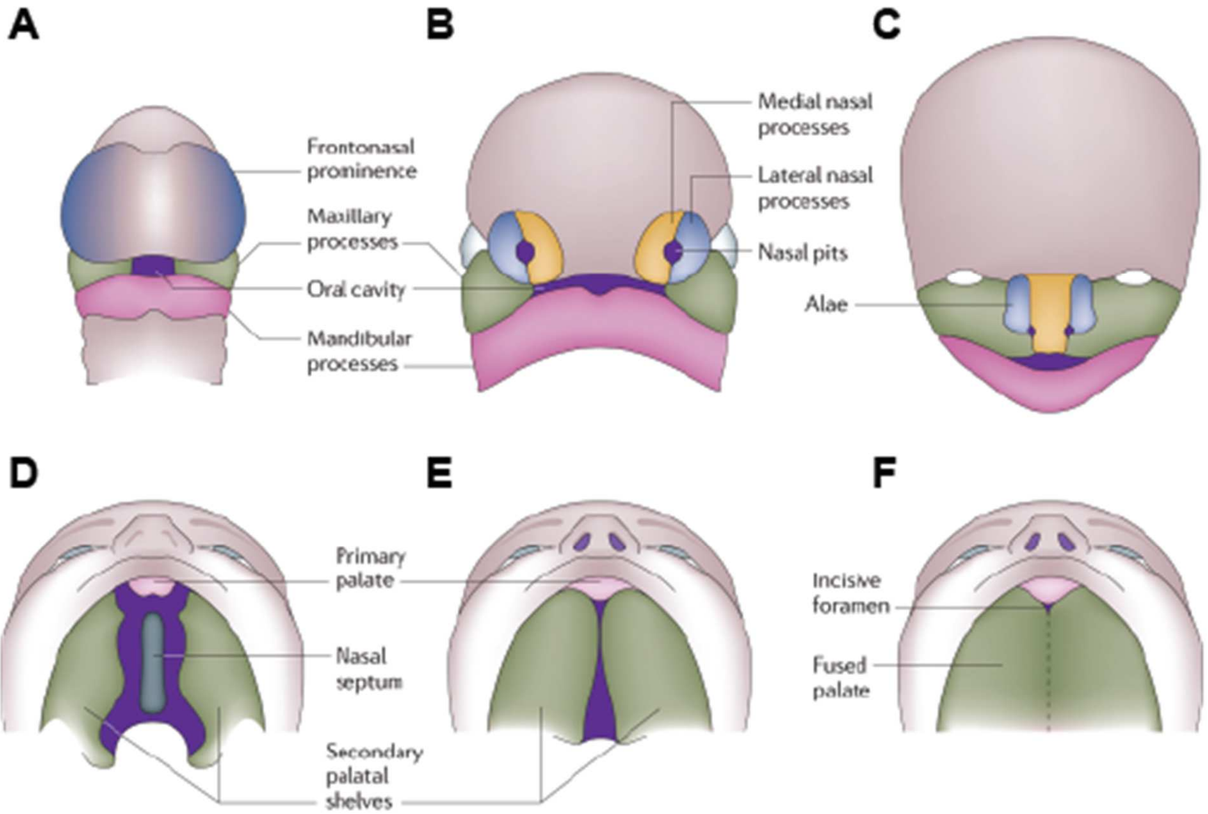


Figure 2: Development of primary and secondary palates from the early facial prominences.

(A) Illustration of the frontonasal prominence (FNP), maxillary prominences (MXPs), and mandibular prominences (MNPs) in the early embryo. (B) Formation of the nasal pits from fusion of the medial nasal prominences and lateral nasal prominences of the FNP. (C) Fusion of the FNP and MXPs gives rise to structures of the midface, while fusion of the MNPs gives rise to the lower jaw. (D) Formation of the primary palate through the development of an intermaxillary segment, and bilateral palatal shelves. (E) Fusion of the FNP-derived intermaxillary segment and secondary palatal shelves forms the hard palate. (F) Fusion between the secondary palatal shelves posterior to the primary palate forms the remainder of the hard palate and soft palate. Figure adapted from Dixon, *Nature Reviews Genetics*.[21]

The palate is subdivided into the primary palate, and the secondary palate (**Figure 2**D-F). The primary palate is derived from the inter-maxillary segment and consists of a triangular piece in the anteromedial portion of the palate. The secondary palate extends posteriorly from the primary palate and is primarily derived from a caudal portion of the MXPs. Coincidentally, the developed palate is also subdivided in two anteroposterior portions: an anterior bony portion called the hard palate, and a mucosal posterior portion called the soft palate. Palatogenesis occurs as the MXPs generate bilateral palatal shelves at the level of the tongue that first ascend to the superior edge of the oral cavity, and then extend medially to fuse with each other at the midline and the inter-maxillary segment ventrally (**Figure 2**E, F). Even though the process of craniofacial morphogenesis through contributions of the FNP, MXPs, and MNPs is highly-conserved across amniotes, only mammals and crocodilians possess a full secondary palate, while birds and lizards maintain a midline cleft[19].

The process of palatal shelf fusion has been thoroughly studied in the mouse, where the palatal shelves developing from the MXP can be appreciated as early as E11, elevate during E12-E13, reorient themselves to a horizontal position above the tongue at 14.5, and converge towards the midline and complete fusion by E15[20-22] (**Figure 3**A-C). The palatal shelves form as a bulk of mesenchymal tissue covered by an epithelial layer called the medial epithelial edge (MEE). Upon convergence, a middle epithelial seam (MES) from the converging bilateral MEEs forms, and degrades through a combination of apoptosis, cell migration, and EMT[23-28] (**Figure 3**D-F). The end result is a continuous sheet of CNCC-derived mesenchyme that forms a mesenchymal bridge connecting tissues derived from the MXPs[19].

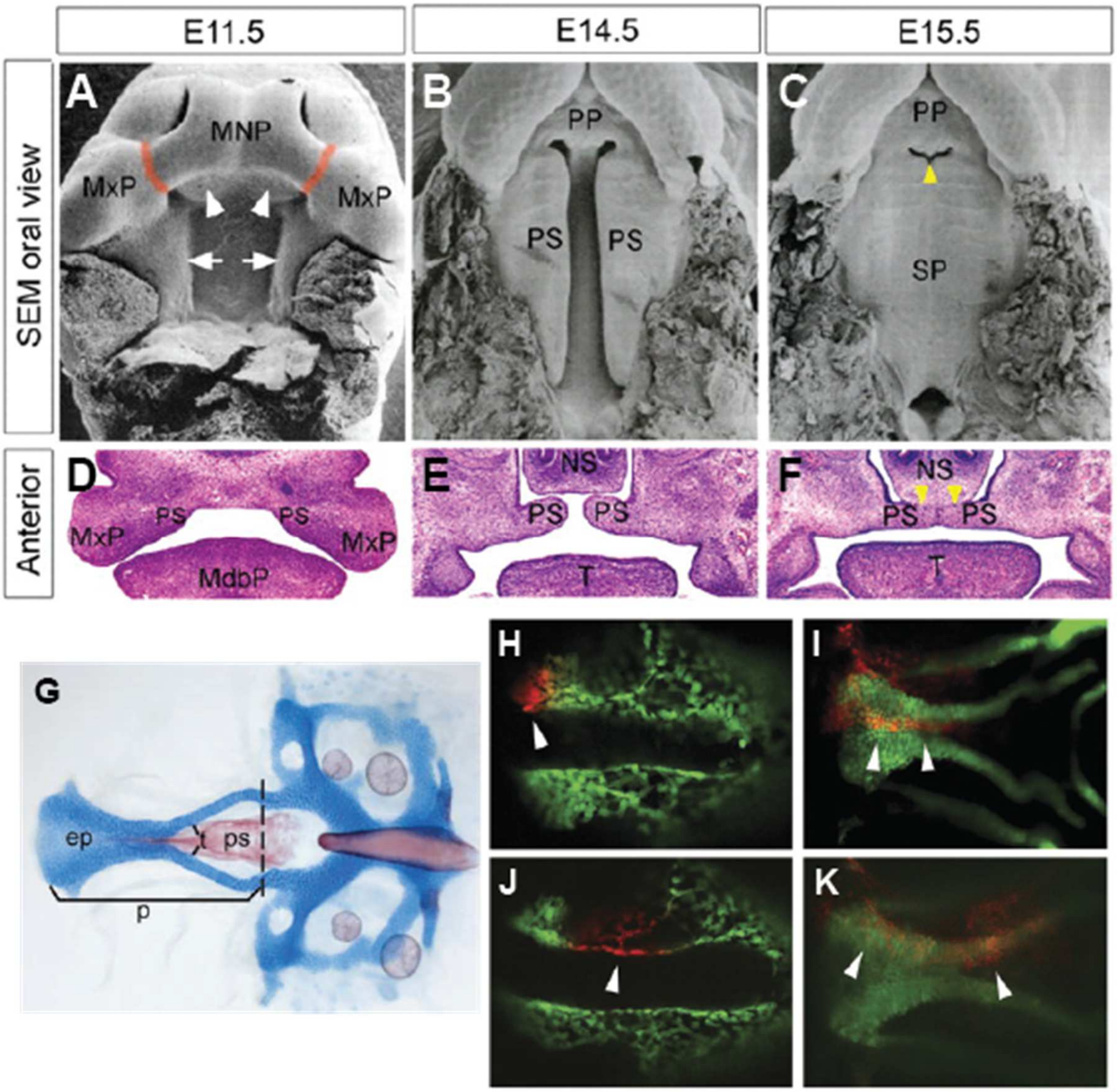


Figure 3: Palatal development in mouse and zebrafish models

(A) SEM oral view from an E11.5 mouse embryo. Arrowheads: MNP, medial nasal prominences. MxP Maxillary prominences. Arrows: developing palatal shelves (B) SEM oral view from an E14.5 embryo. The palatal shelves have formed and approximate each other at the midline. PP, primary palate. PS, palatal shelves. (C) SEM oral view from an E15.5 mouse embryo. The palatal shelves have fused to each other, and to the primary palate. PP, primary palate. SP, secondary palate (D) H&E-stained coronal sections through E11.5 mouse embryos. PS, palatal shelves. MxP, maxillary prominence. MdbP, mandibular prominence.

Figure 3 (Cont'd) (E) H&E-stained coronal sections through E14.5 mouse embryos. PS, palatal shelves. NS, nasal septum. T, tongue. (F) H&E-stained coronal sections through E15.5 mouse embryos. PS, palatal shelves. NS, nasal septum. T, tongue. (G) Acian blue and alizarin red staining of the zebrafish neurocranium. p, palate. ep, ethmoid plate. ps, parasphenoid bone (H) Lineage tracing in the Tg(sox10:kaede) zebrafish line. Photoconversion of a frontonasal neural crest subset (red, arrowhead) in 20hpf Tg(sox10:kaede) embryos. (I) At 4dpf, the photoconverted frontonasal neural crest signal localizes to the medial ethmoid plate (arrowheads). (J) Photoconversion of a maxillary neural crest subset (red, arrowhead) in 20hpf Tg(sox10:kaede) embryos. (K) At 4dpf, the photoconverted frontonasal neural crest signal localizes to the ipsilateral trabecula of the ANC (arrowheads). Images adapted from Bush, Jiang. *Development*[20] (A-F), Swartz et al. *Developmental Dynamics*[22] (G), and Dougherty et al. *Journal of Craniofacial Surgery*[29] (H-K)

In zebrafish, the roof of the mouth is formed by a set of bones that separates the oral cavity from the brain. As early as 4dpf, three structures can be appreciated at the roof of the mouth: Cartilaginous paired trabeculae (lateral) and ethmoid plate (medial), together hereafter referred to as the anterior neurocranium (ANC), and the dermal parasphenoid bone that contribute to the juvenile and adult palatal skeletons [30-32] (**Figure 3**G). Just like the mouse and other amniotes, these structures are derived from frontonasal CNCCs that originate dorsally and migrate anteriorly and inferiorly; and bilateral maxillary and mandibular CNCCs that pass through PA1 and converge to fuse at the midline. However, it appears that palatogenesis in fish species occurs directly between the FNP and MXPs without the development of a MES, although the presence of rudimentary palatal shelves has been described in salmon[22, 33]. Further evidence for the fusion of the FNP to the MXPs in the zebrafish anterior neurocranium has been shown through lineage tracing experiments using the photoconvertible Tg(sox10: kaede) line. After photoconversion, frontonasal NCCs track to the medial ANC, while the MXP NCCs track to the lateral ANC and trabeculae[29, 34, 35] (**Figure 3**H-K). Ultimately, while the anatomy of the zebrafish neurocranium is not an exact replica of the mammalian palate, the processes that lead to the ANC formation are largely conserved and involve similar developmental trajectory and molecular patterns.

Transplantation experiments in the chick have shown that the anatomic origin of CNCCs provides some early determination of tissue types, however CNCCs retain some plasticity and are reliant on positional cues and morphogenic signals at their target locations. This has led to the discovery of organizing centers such as the frontonasal ectodermal zone (FEZ)[36]. The FEZ is identified by a region of juxtaposed dorsal Fgf8 expression and ventral Shh expression to drive the outgrowth of the FNP. Studies surrounding the FEZ have shown that the overlying ectoderm provides a patterning role, given that transplantation of the FEZ can lead to the formation of ectopic beaks at different locations in the FEZ. Furthermore, the induction of an ectopic upper beak be replicated by the implantation of beads that release Shh, and ablation of Shh signals in the overlying ectoderm prevents the formation of the FNP. Still, implantation of the FEZ into other anatomic regions such as PA1 does not induce ectopic beaks. Thus, CNCCs possess an intrinsic pre-patterning based on their rhombomeric origin that is guided by morphogenic signals in their intended anatomic destinations.

Pathology associated with aberrant craniofacial morphogenesis

Orofacial clefts (OFCs) are one of the most common craniofacial anomalies, with an incidence between 1/700 to 1/1200 live births[21, 37, 38]. While OFCs don't offer a major contribution to mortality in developed countries, OFCs impose financial and social burdens to affected individuals, in addition to overall reduced quality of life[39]. Depending on the cleft severity, patients with OFCs can present with a variety of symptoms that range from difficulties with feeding, speech, or social interactions to hearing loss from recurrent ear infections or life-threatening aspiration pneumonias. Current treatment for OFCs involve a mixture of surgery, dental care, speech therapy, and psychosocial interventions, and can introduce additional post-operative complications such as infection or hemorrhage. The majority of OFCs are clinically classified under four properties: 1) Structure containing the cleft (lips/dental alveoli/hard palate/soft palate) 2) Severity of the cleft (width and involved craniofacial structures), 3) Laterality of the cleft(s) (unilateral/bilateral), and 4) Combination of clefts and presence of other structural and cognitive abnormalities (isolated cleft lip or palate/cleft lip and palate/syndromic cleft lip and/or palate/non-syndromic cleft lip and/or palate).

At the simplest level, OFCs are classified under three groups: cleft lip only (CLO), cleft palate only (CPO), or cleft lip and palate (CLP). Beyond that, they can be further described based on whether they are unilateral or bilateral (**Figure 4**). The presence of other systemic findings has led to the identification of genetic syndromes for which cleft lip and/or palate (CL/P) is a feature. Two important syndromes for this work are Van der Woude Syndrome (VWS) and Popliteal Pterygium Syndrome (PPS). Van der Woude patients have characteristic lower lip pits in addition to CL/P, and some patients will also present with hypodontia (**Figure 4**, B, e). PPS is characterized by the formation of webbing in different structures, most commonly the back of the knee (popliteal region).

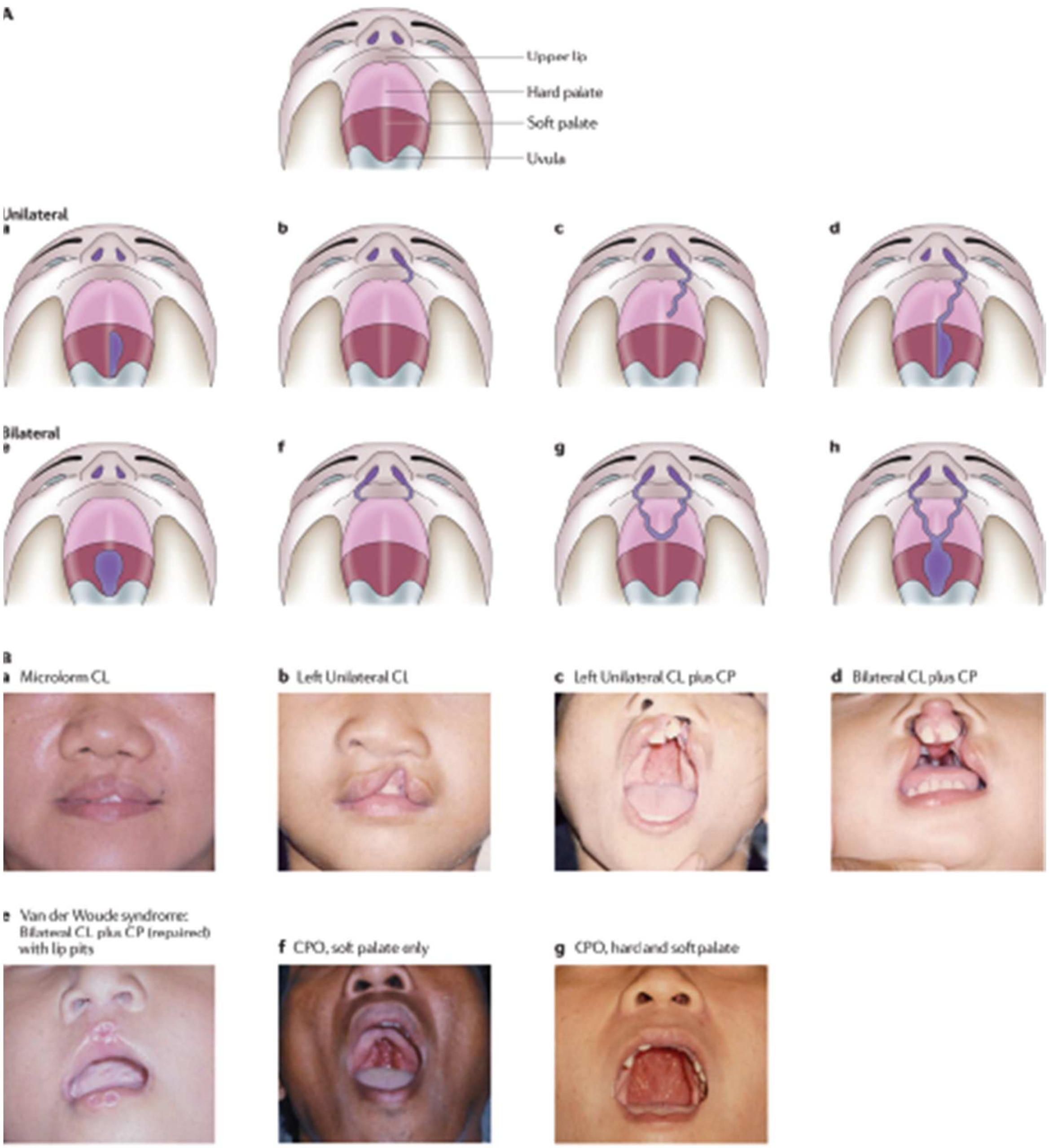


Figure 4 Classification of OFC presentations with clinical examples

(A) Cartoons depicting the common presentation of OFCs. (a) Unilateral cleft palate only (CPO) of the soft palate (b) Unilateral cleft lip only (CLO), (c) Unilateral cleft lip and cleft hard palate (CLP) (d) Unilateral CLP extending to the soft palate (e) Bilateral CPO of the soft palate (f) Bilateral CLO (g) Bilateral CLP extending to the hard palate (h) Bilateral

Figure 4 (Cont'd) CLP extending to the soft palate (B) Clinical examples of OFCs (a) microform CLO (b) Unilateral CLO (c) Unilateral CLP (d) Bilateral CLP (e) Van der Woude Syndrome (VWS) depicting lower lip pits, and bilateral CLP post-repair (f) CPO of the soft palate (g) CPO of the hard and soft palate. Image adapted from Dixon et al. *Nature Reviews Genetics* [21]

A regulatory axis involving IRF6 is critical for craniofacial development and craniofacial morphogenesis

Our current understanding of OFC cleft pathogenesis is largely driven by the discovery of genetic drivers of disease in CL/P. Among these, *IRF6* is one of the most important factors, as it is found to cause syndromic CL/P in Van der Woude syndrome (VWS) and Popliteal Pterygium Syndrome (PPS) in addition to non-syndromic forms of CL/P[40-45]. It is worth to highlight that the clinical features of both VWS and PPS reach beyond cleft lip and palate and include epithelial defects: lower lip pits and webbing of structures like the back of the knee, hands and feet, or eyelids (popliteal pterygia, syndactyly, ankyloblepharon). As a transcription factor, IRF6 is not the end effector protein for morphogenetic processes but regulates other target genes in the embryonic periderm to direct craniofacial morphogenesis. One well-characterized target of IRF6, GRHL3, is critical for proper formation and function of the embryonic periderm, regulates formation of the neural tube and craniofacial skeleton [46, 47]. Moreover, the majority of VWS patients without *IRF6* mutations have mutations in *GHL3*, and thus mutations in the *IRF6*-*GHL3* axis account for the vast majority of VWS cases[40, 45]. *IRF6* in turn is controlled by p63 and TFAP2a, which all account for important axes in neurulation and craniofacial development [47].

IRF6 expression is restricted to the epithelium, including an early epithelial covering of the embryo, the periderm (also called the enveloping layer, EVL, in zebrafish)[48]. Despite its nature as an early epithelial structure, the periderm is a complex tissue that is critical for morphogenic events in the early embryo, craniofacial morphogenesis, and palatal development[48, 49]. There are multiple mouse models of *Irf6* loss-of-function which cause embryonic lethality, and the embryos develop cleft palate, truncated limbs, and a loss of the epithelial-barrier function[50]. A reporter construct for a putative *Irf6* enhancer and Irf6 immunofluorescence have shown that *Irf6* is expressed in the oral epithelium and MEE of the developing palatal shelves in E13.5 mouse embryos, and colocalizes with p63 expression (**Figure 5**A, C, F). Moreover, expression of *Irf6* and other periderm markers at the MES is lost upon fusion of the secondary palatal shelves around E14.5 (**Figure 5** D, E, G, H).

The functional significance of both irf6, and the embryonic periderm is further highlighted in zebrafish models of *irf6* loss-of-function. *irf6* in zebrafish is maternally deposited such that irf6 protein synthesis occurs immediately in the early embryo[51]. While first-generation *irf6*$^{-/-}$ fish are viable and develop into adulthood, maternal-null *irf6*$^{-/-}$ embryos from *irf6*$^{-/-}$ female zebrafish rupture early in development during gastrulation, around 6hpf[51] (**Figure 5** I-M).

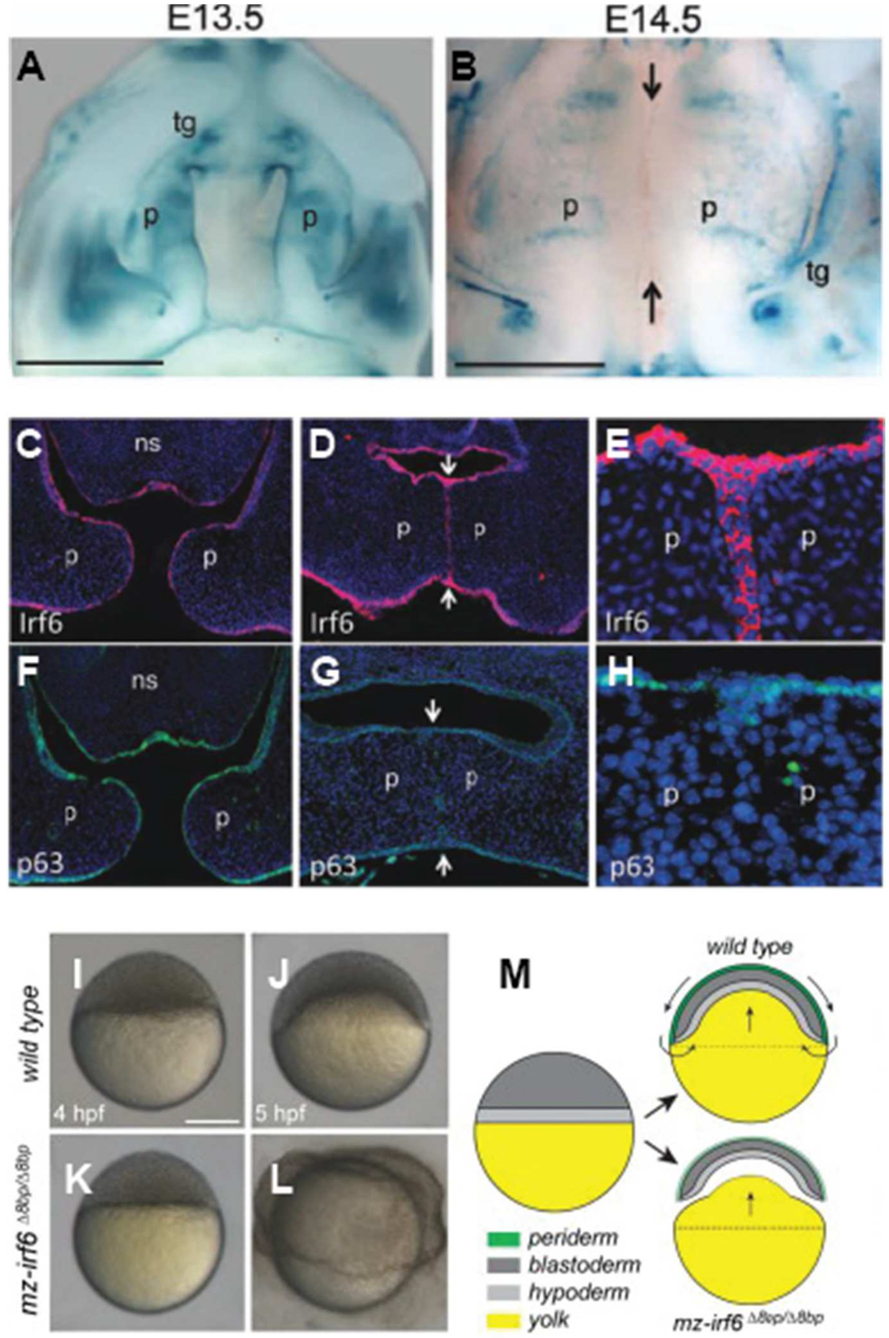


Figure 5 Irf6 and the embryonic periderm are critical for embryonic development and craniofacial morphogenesis

(A) β-galactosidase expression of the putative *Irf6* enhancer, MCS9.7. Expression is restricted to the oral epithelium overlying the palatal shelves, the MEE. (B) Expression is reduced upon palatal fusion at E14.5. (C-E) Immunofluorescence for Irf6 in the (C) MEE at E13.5 and (D-E) MES at E14.5. (F-H) Immunofluorescence for the *Irf6* inducing factor p63 (F) p63 is expressed at the MEE at E13.5 (G-H) p63 expression is lost in the fusing MES

Figure 5 (Cont'd) at E14.5. p, palatal shelf. tg, tooth germ. (I-M) Embryonic rupture of maternal zygotic-null (mz) i*rf6*$^{-/-}$ zebrafish embryos. (I-J) Wild-type zebrafish embryos at (I) 4hpf and (J) 5hpf. (K-L) mz i*rf6*$^{-/-}$ embryos rupture around 5-6hpf (M) Embryonic rupture is a consequence of abnormal periderm function during zebrafish gastrulation. Images adapted from Fakhouri et al. *Sci Rep* (A-H). and Li, et al. *PLoS Genet* (A-M).

Chapter II: Synergistic roles of Wnt Modulators R-spondin2 and R-spondin3 in Craniofacial Morphogenesis and Dental Development

Introductory Statement

This article explores the role of a wnt-pathway regulator, *Rspo3*, in craniofacial development and odontogenesis in mice and zebrafish, published in *Scientific Reports* in 2021. We followed up on another irf6 target identified by our previous RNA-seq experiments in zebrafish, *rspo3*, and set out to determine the effects of *rspo3* loss-of-function in craniofacial development, tooth development, and interactions with its ortholog *rspo2*. I assisted Nora Alhazmi in the initial stages of exploring the phenotypic effects of *rspo3* LOF in zebrafish and provided consistent instruction, mentoring, and revision of experimental design through her time in our laboratory. I assisted Nora with establishing a CRISPR-mutagenized rspo3-/- line, performing alcian blue and alizarin red staining and dissection of mutant zebrafish, and RNAscope ISH efforts for the paper. Nora Alhazmi performed most of the exploratory experiments in the article, and Shannon Carroll performed experiments for revision and acceptance of the manuscript.

Synergistic roles of Wnt Modulators *R-spondin2 and R-spondin3* in Craniofacial Morphogenesis and Dental Development

Nora Alhazmi[1,†], Shannon H. Carroll[2,3,4†], Kenta Kawasaki[2,3], Katherine C. Woronowicz[5,6], Shawn A. Hallett[2,4], Claudio Macias Trevino[2,4], Edward B. Li[2,4], Roland Baron[1,4], Francesca Gori[1], Pamela C. Yelick[7], Matthew P. Harris[5,6], and Eric C. Liao[2,3,4,8*]

[1] Harvard School of Dental Medicine, Boston, Massachusetts, United States of America.
[2] Center for Regenerative Medicine, Massachusetts General Hospital, Boston, Massachusetts, United States of America.
[3] Shriners Hospital for Children, Boston, Massachusetts, United States of America.
[4] Department of Medicine, Harvard Medical School, Boston, Massachusetts, United States of America.
[5] Department of Genetics, Harvard Medical School, Boston, Massachusetts, United States of America.
[6] Department of Orthopedics, Boston Children's Hospital, Boston, Massachusetts, United States of America.
[7] Department of Orthodontics, Division of Craniofacial and Molecular Genetics, Tufts University School of Dental Medicine, Boston, Massachusetts, United States of America.
[8] Division of Plastic and Reconstructive Surgery, Massachusetts General Hospital, Boston, Massachusetts, United States of America.
† These authors contributed equally to this work.

*Author for correspondence: cliao@partners.org

Telephone: ECL: (617) 643-5975

Key words: *wnt*, *rspo2*, *rspo3*, craniofacial, dental, bone, development

Abstract

Wnt signaling plays a critical role in craniofacial patterning, as well as tooth and bone development. *Rspo2* and *Rspo3* are key regulators of Wnt signaling. However, their coordinated function and relative requirement in craniofacial development and odontogensis are poorly understood. We showed that in zebrafish *rspo2* and *rspo3* are both expressed in osteoprogenitors in the embryonic craniofacial skeleton. This is in contrast to mouse development, where *Rspo3* is expressed in osteoprogenitors while *Rspo2* expression is not observed. In zebrafish, *rspo2* and *rspo3* are broadly expressed in the pulp, odontoblasts and epithelial crypts. However, in the developing molars of the mouse, *Rspo3* is largely expressed in the dental follicle and alveolar mesenchyme while *Rspo2* expression is restricted to the tooth germ. While *Rspo3* ablation in the mouse is embryonic lethal, zebrafish *rspo3*-/- mutants are viable with modest decrease in Meckel's cartilage rostral length. However, compound disruption of *rspo3* and *rspo2* revealed synergistic roles of these genes in cartilage morphogenesis, fin development, and pharyngeal tooth development. Adult *rspo3-/-* zebrafish mutants exhibit a dysmorphic cranial skeleton and decreased average tooth number. This study highlights the differential functions of *Rspo2* and *Rspo3* in dentocranial morphogenesis in zebrafish and in mouse.

Introduction

The Wnt signaling pathway plays a major role in skeletal patterning and differentiation during embryonic development, and in maintaining postnatal bone homeostasis.[94-96] Impairment and potentiation of Wnt signaling affects overall bone mass and density.[94-96] Canonical β-catenin mediated Wnt signaling directly regulates osteoblast differentiation and activity and likely has indirect effects on osteoclasts during bone metabolism.[97] Moreover, a study reported the direct negative influence of canonical Wnt/β-catenin signaling on osteoclast development using *in vitro* cell models and *in vivo* mouse studies.[98] During embryogenesis, canonical Wnt/β-catenin signaling is highly active in the first pharyngeal arch as well as in multiple craniofacial regions in mouse, chicken and zebrafish.[99-101] Wnt signaling is involved in regulating skeletogenic neural crest cells, such as the subdivision of each pharyngeal arch into dorsal and ventral elements in zebrafish during craniofacial development.[102] In addition, Wnt signaling plays a role in mediating regional specification in the vertebrate face.[100] The identification of modulators of Wnt signaling during development and homeostasis of adult skeletal tissues may lead to new insights into disease etiology and identify potential targets for therapeutic mediation.

Human genome wide association studies revealed many regulators of canonical Wnt signaling that are involved in regulating bone metabolism.[103, 104] The *R-spondin* (*Rspo*) family of secreted proteins includes four members (*Rspo1-4*) in the thrombospondin type 1 repeat (TSR1)- containing protein superfamily that have been shown to potentiate the canonical Wnt/β-catenin pathway.[105, 106] RSPO proteins modulate Wnt signaling through interactions with the LGR4-6 receptors, leading to stabilization of Frizzled and LRP5/6 at the cell membrane, and through regulation of the ubiquitin ligases ZNFR3 and RNF43 that degrade Frizzled receptors.[106, 107] *Rspo2* and *Rspo3* also have been shown to augment Wnt/β-catenin signaling independent of LGRs by binding to heparin sulfate proteoglycans.[108] *Rspo* genes are essential for normal development and have been shown to regulate skeletal patterning during development.[109] In particular, *Rspo2* has been shown to be essential for limb patterning.[110][17] Additionally, several GWAS conducted in humans have associated *RSPO2* and *RSPO3* with bone mineral density.[103, 111]

Rspo3 was identified as a candidate gene that contributes to cleft lip/palate and dental anomalies.[112] *Rspo3* was also reported to have a critical role in mouse placental development.[110] However, since mouse embryos lacking *Rspo3* function die at E10.5 due to placenta and vascular defects, this precluded analysis of its role during later embryonic development.[110] Conditional ablation of the *Rspo3* in limb mesenchymal cells caused modest delay in limb growth during development.[113] *Rspo3* and *Rspo2* double mutant mice however developed severe hindlimb truncations, suggesting a redundant function of these genes.[113] The function of *Rspo3* during craniofacial morphogenesis has yet to be defined.[114]

Wnt/β-catenin pathway also plays a critical role in tooth development and can affect craniofacial development more broadly.[115, 116] Tooth formation initiates from the interactions between the dental epithelial layer and the underlying mesenchyme.[117, 118] Mice have a single set of dentition (monophyodont) that consists of continuously erupting incisors, and three molars in single row on both sides of the upper and lower jaws that do not exhibit continuous growth or replacement.[119, 120] The zebrafish dentition is more numerous, unlike in the mouse, zebrafish teeth exhibit continuous replacement throughout life (polyphyodont).[121] Despite these differences, the molecular and cellular mechanisms regulating tooth development are highly conserved between zebrafish and mammals.[121] Therefore, studies in zebrafish can provide novel insights into the regulation of craniofacial structures that can complement the mouse.

Here, we focused on the roles of *Rspo2* and *Rspo3* in regulating dental and craniofacial development. We utilized RNAscope probes to gain high resolution images of *Rspo2 and Rspo3* gene expression in zebrafish and mouse. We examined the genetic requirement of *rspo2 and rspo3* in zebrafish development, using complementary CRISPR/Cas9-mediated targeted mutagenesis. Using these approaches, we revealed roles for r*spo2 and rspo3* in tooth development and in morphogenesis of the craniofacial complex.

Results

Rspo2 **and** ***Rspo3*** **are expressed in the craniofacial complex and in the perichondrium and osteoprogenitor cells during zebrafish craniofacial morphogenesis.**

Gene expression patterns of *rspo2* and *rspo3* during zebrafish embryogenesis were delineated by whole-mount RNA *in situ* hybridization (WISH). *rspo2* and *rspo3* transcripts were detected in the brain, otic vesicle, and endodermal pouches at 24 hours post-fertilization (hpf) and as well as in regions consistent with the ethmoid plate and Meckel's cartilage at 48 hpf (**Figure 6**A). Using RNAscope *in situ* hybridization, we identified diffuse *rspo3* transcript expression throughout the mesenchyme with concentrated expression in cells that circumscribe the pre-cartilage mesenchyme (48 hpf) and the paired trabeculae, ethmoid plate, and Meckel's cartilage at 5 days post-fertilization (dpf) (**Figure 6**B). We also detected *rspo3* expression within ethmoid plate chondrocytes at 5 dpf (**Figure 6**B). *rspo2* expression generally overlapped with *rspo3* at both developmental timepoints.

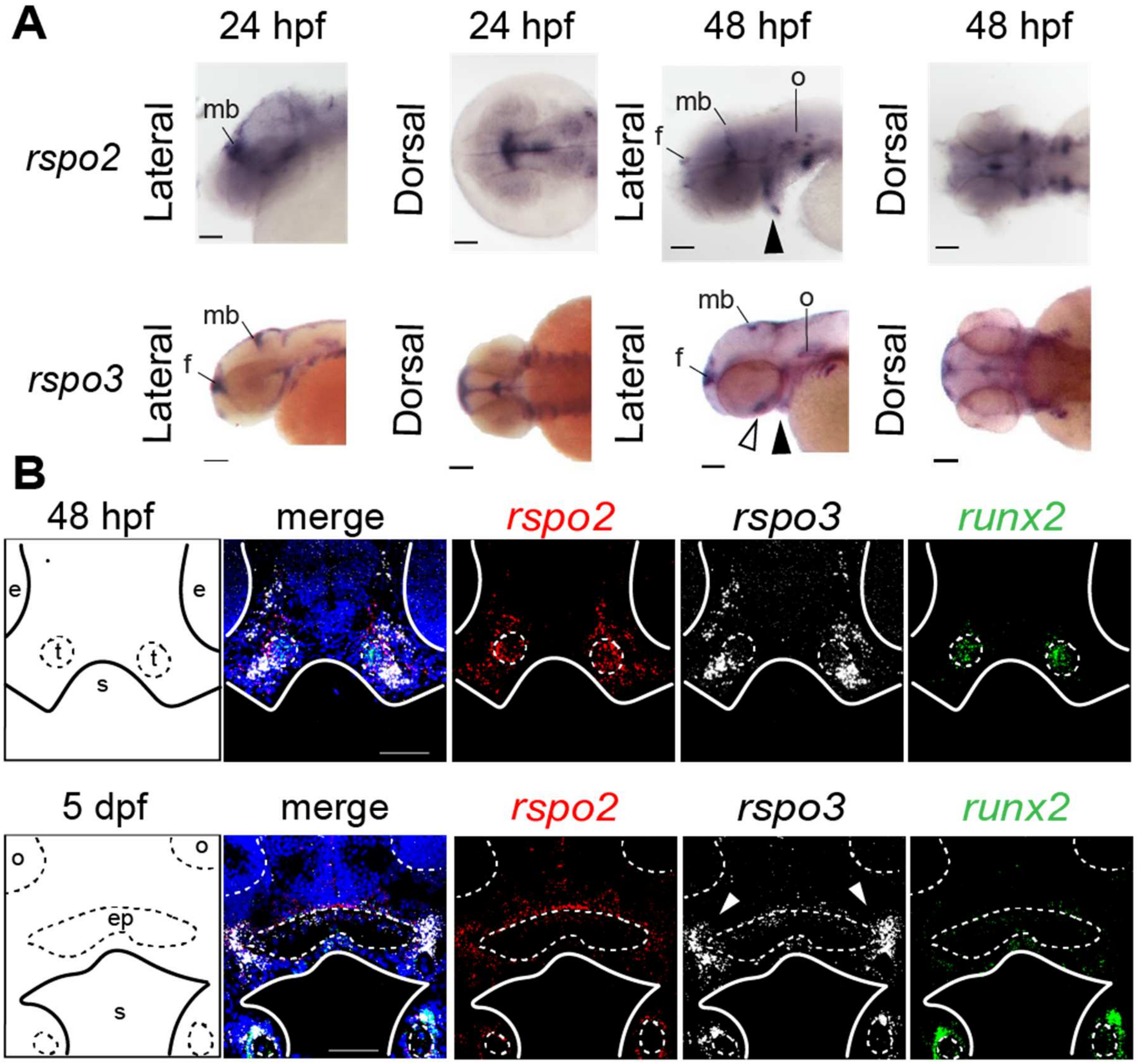


Figure 6 High resolution gene expression analysis detected dynamic spatiotemporal localization of *rspo2* and *rspo3* transcripts in zebrafish cranial mesenchyme.

Figure 6 (Cont'd) A) Whole-mount RNA *in situ* hybridization detected *rspo2* and *rspo3* transcripts in the midbrain (mb), forebrain (f), otic vesicle (o), and Meckel's cartilage (arrowhead) at 24 and 48 hpf in lateral and dorsal views. Transcript of *rspo3* was additionally detected at 24 and 48 hpf in the forebrain (f) and at 48 hpf in the ethmoid plate (open arrowhead). Scale bar: 100 um. B) Maximum projections of z-stacks of coronal sections of zebrafish embryos, section in plane with eyes. Labeled schematic at left. At 48 hpf, *rspo2* and *rspo3* transcripts were highly co-localized in the mesenchyme surrounding the condensing trabeculae cartilage mesenchyme, marked by *runx2a* expression. *rspo2* is also detected within the condensing mesenchyme at this timepoint. At 5 dpf *rspo2* and *rspo3* continue to be co- expressed in mesenchyme and perichondrium surrounding cartilage elements, specifically the ethmoid plate and Meckel's cartilage. *Rspo3* expression is particularly high in *runx2a* expressing osteogenic precursor cells associated with Meckel's cartilage, as well as where the palatoquadrate meets the ethmoid plate (white arrowhead). Abbreviations: e: eye, o: olfactory organ, s: stomodeum, t: trabeculae. Scale bar: 100 μm

Expression of *rspo3* is similar in mouse and zebrafish, however *rspo2* expression is distinct.

To test the conservation of *Rspo2* and *Rspo3* expression between vertebrates, we analyzed expression in mouse embryos with RNAscope *in situ* hybridization and immunofluorescence. At E13.5 we detected *Rspo3* expression in regions consistent with osteogenesis, including the developing mandible. We found cellular co-localization of

Rspo3, C*ol1a1* mRNA and *Runx2* protein, indicating a potential role in osteogenesis (**Figure 7**A). In the E15.5 mouse embryo, *Rspo3* transcripts were detected widely throughout the mesenchyme as well as within Meckel's cartilage (**Figure 7**B). *Rspo3* expression was also detected within *Runx2* positive, presumptive osteoprogenitor cells (**Figure 7**B). In contrast to gene expression results in zebrafish, we did not detect *Rspo2* expression in the mesenchyme of mouse embryos or associated with cartilage elements. Instead, within the developing mandible, *Rspo2* expression was restricted to developing teeth (**Figure 7**C). Unlike in zebrafish where *rspo2* and *rspo3* expression largely overlap, in the mouse transcripts of *Rspo2* and *Rspo3* appear to be anatomically distinct. (**Figure 7**)

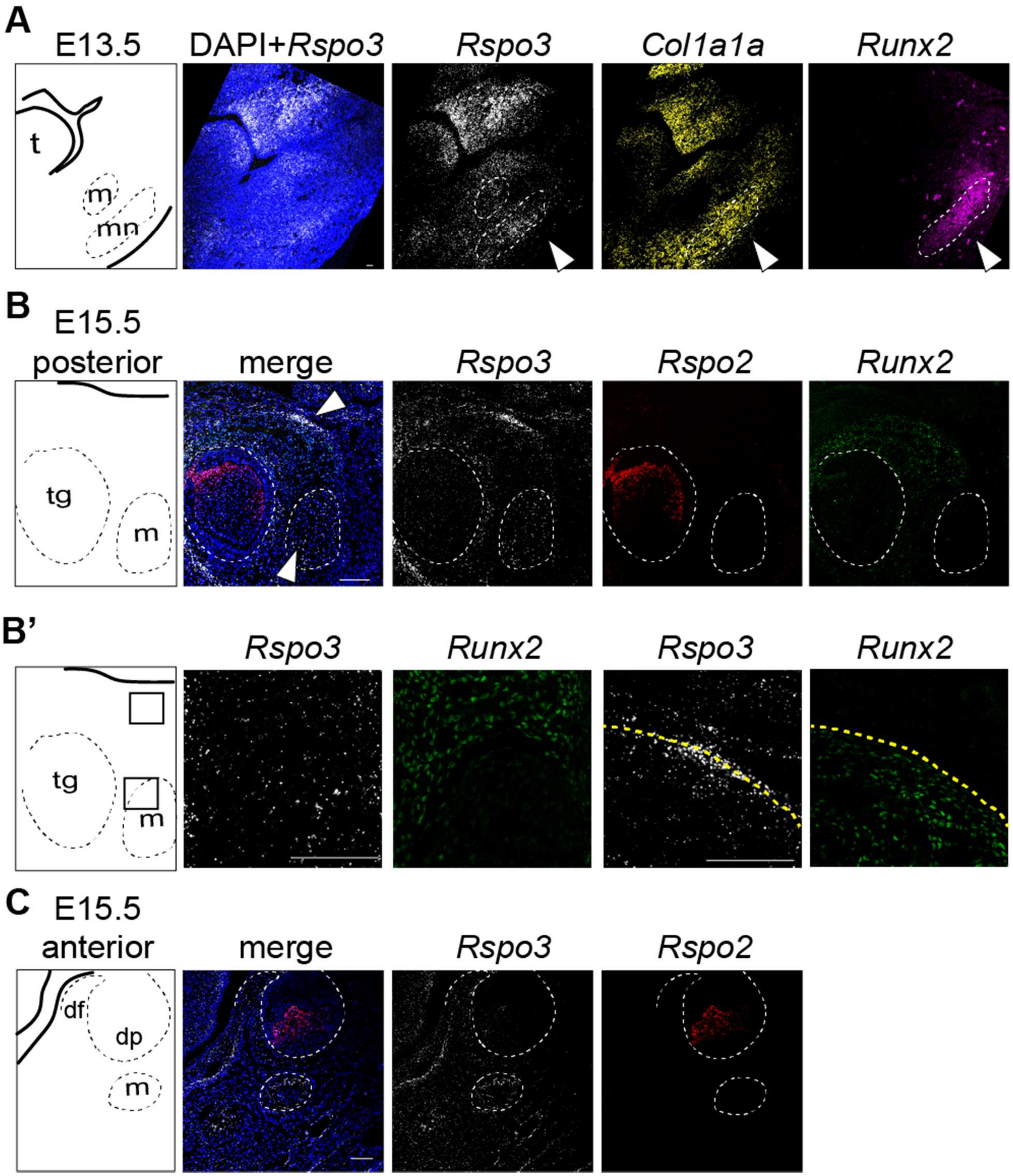


Figure 7 RNAscope gene expression analysis of *Rspo2* and *Rspo3* in mouse embryos.

Figure 7 (Cont'd) Coronal sections of wild-type mouse embryos at E13.5 and E15.5. Labeled schematic at left. A) RNAscope revealed diffuse expression of *Rspo3* throughout the cranial mesenchyme at E13.5. *Col1a1 in situ* hybridization and *Runx2* immunofluorescence were used to identify osteogenic precursor cells and *Rspo3* expression was detected within these cells. Of particular note is *Rspo3* expression within Meckel's cartilage (m) and the developing mandible (arrow; mn). B) At E15.5, *Rspo3* expression is detected diffusely thorough the mesenchyme, including in Meckel's cartilage (m) and within osteogenic precursor cells (co-expressed with *Runx2*). *Rspo2* expression is isolated to discrete cells within the developing tooth germ (tg). B') Higher magnification images (boxes depict location) of *Rspo3* and *Runx2* co-expression. C) Within the developing molar at E15.5, *Rspo2* and *Rspo3* transcripts were detected in distinct non-overlapping regions, with *Rspo3* expression in Meckel's cartilage (m) and the dental follicle (df), while *Rspo2* is expressed exclusively within the dental pulp (dp). Scale bar: 100 µm.

***rspo2* and *rspo3* are differentially expressed within zebrafish dental structures.**

Given the expression of *Rspo2* and *Rspo3* in developing mouse teeth, we examined the gene expression of *rspo2* and *rspo3* within and surrounding the tooth structure in zebrafish. *rspo3* gene expression was detected at low levels diffusely throughout the dental pulp and the surrounding mesenchyme (**Figure 8**). In contrast, high levels of *rspo2* gene expression were detected in the enamel epithelium (**Figure 8**). Furthermore, *rspo3* gene expression was highest within odontoblasts of regenerating teeth (**Figure 8**).

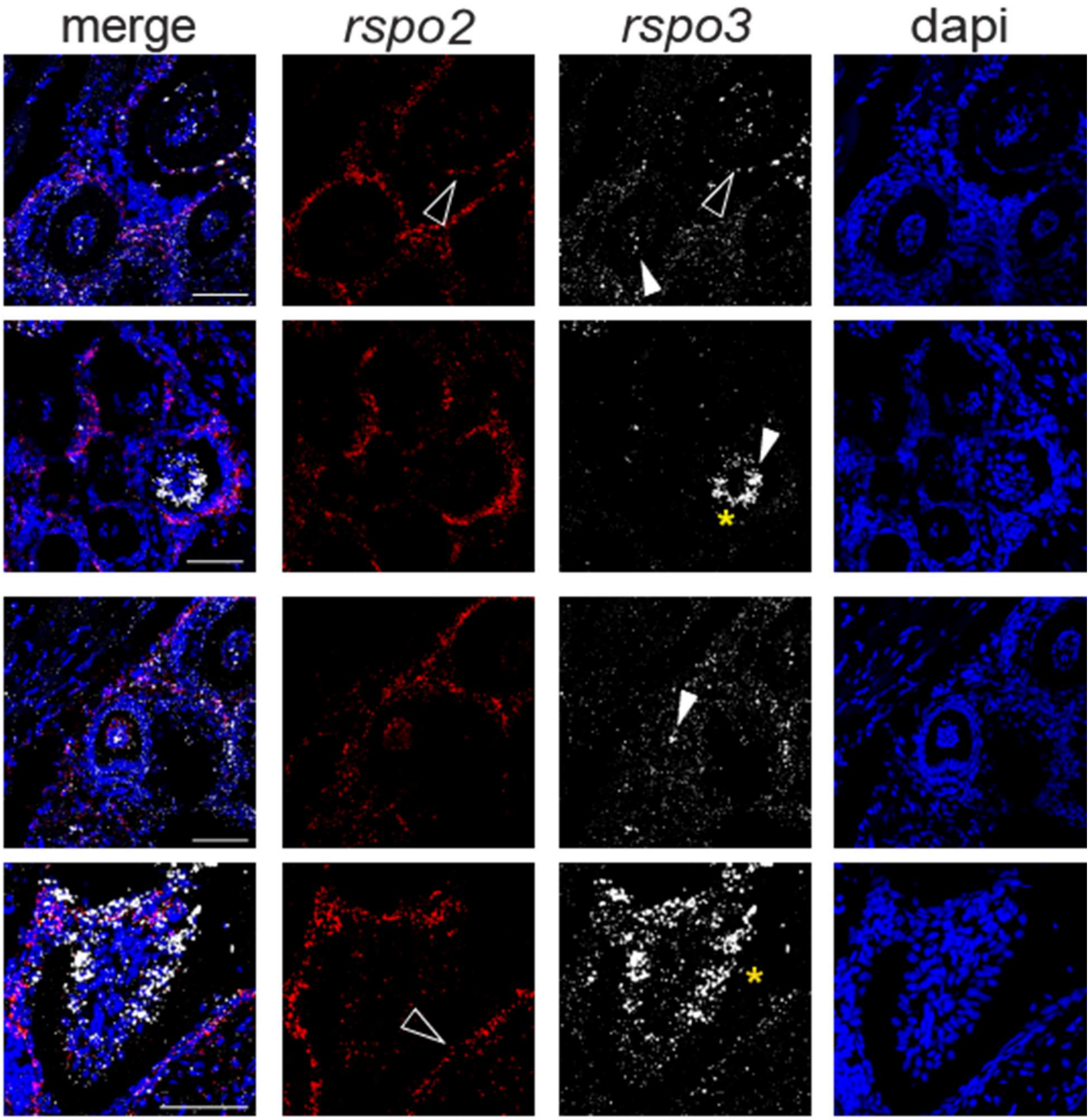


Figure 8 *rspo2* and *rspo3* are differentially expressed in zebrafish pharyngeal teeth. Representative images of maximum projections of z-stacks of sagittal sections of 180 dpf zebrafish. *rspo2* (red) and *rspo3* (white) transcripts were detected by RNAscope *in situ* hybridization. *rspo3* is diffusely expressed throughout the supporting mesenchyme and highly expressed in enamel epithelium (open arrow) and some tooth

Figure 8 (Cont'd) pulp (closed arrow). Often within the tooth pulp, *rspo3* expression is restricted to the outermost odontoblasts (*). Meanwhile, *rspo2* expression is highest in enamel epithelium (open arrow) and very low within the tooth pulp. Scale bar: 100 µm.

Combined disruption of *rspo2 and rspo3* resulted in cartilage dysmorphogenesis.

Given the specific expression of *rspo3* in early palate and in Meckel's cartilage development, we used CRISPR/Cas9-mediated genome editing to generate *rspo3* mutant alleles. Guide RNAs targeting the *rspo3* gene in exon2 were used to create a *rspo3* mutant germline allele. A frame shift mutation was generated by introducing a -20 bp gene deletion, as detected by microsatellite genotyping and confirmed by Sanger sequencing. The efficiency of this *rspo3* -20 bp deletion allele (hereafter called *rspo3*$^{-/-}$) was assessed by qRT-PCR at 6 hpf, where we observed that *rspo3* mRNA was significantly reduced by 6-fold in the mutant compared as compared to wild-type clutch-mates ($p < 0.05$).

To characterize requirements for *rspo3* during early craniofacial morphogenesis, Alcian blue cartilage staining was performed at 5 dpf. The effects of *rspo3* disruption on larval cartilage skeleton were found to be subtle. As *rspo2* is known to also function in regulating Wnt signaling and has overlapping expression in the zebrafish, we hypothesized that *rspo2* action may be compensating for *rspo3* germline disruption. Therefore, to determine the combined requirement of *rspo2* and *rspo3*, we targeted *rspo2* by injection of multiple gRNA into *rspo3* homozygous embryos (**Figure 9**A), commonly referred to as a crispant and denoted here as *rspo2*$^{\Delta}$.[29] Embryos generated from *rspo3*$^{+/-}$ in-crossed zebrafish were either raised for analysis of the single mutant or were injected at the 1-cell stage with gRNAs targeting *rspo2* (*rspo2*$^{\Delta}$). The resulting larvae were stained at 9 dpf with Alcian blue and Alizarin Red S. Following imaging and phenotyping, individual larvae were genotyped. We identified a subset of zebrafish with disrupted pectoral fin development where either the fin was partially formed or was absent (**Figure 9**B). We found that *rspo2* was required for pectoral fin development, and that haploinsufficiency of *rspo3* exacerbated the loss of pectoral fin formation (**Figure 9**B, C).

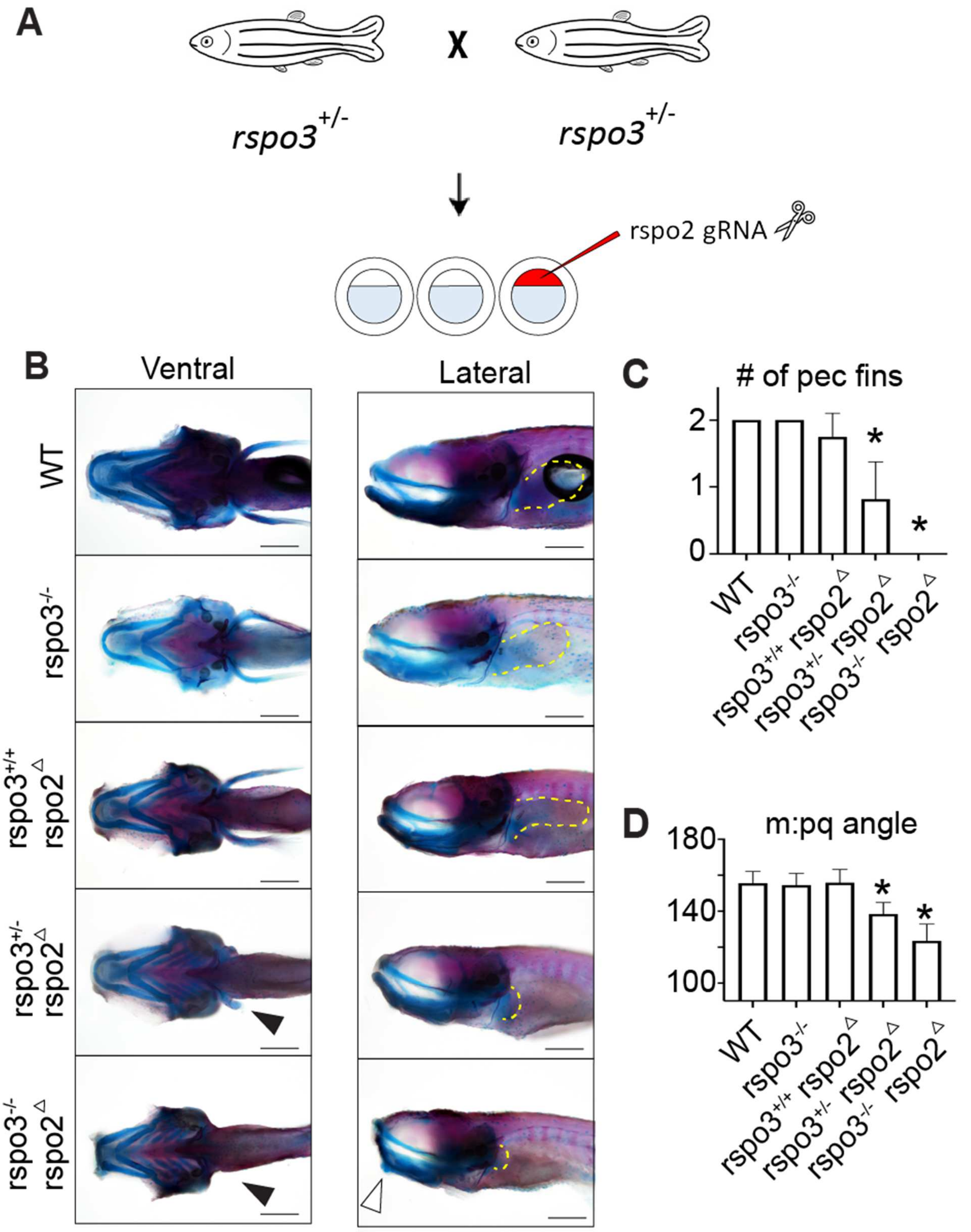


Figure 9 Synergistic effect of *rspo2* and *rspo3* ablation on zebrafish limb development and craniofacial morphology.

Figure 9 (Cont'd) A) Schematic illustrating experimental design. Targeted mutagenesis of *rspo3*$^{-/-}$ in zebrafish was carried out using CRISPR/Cas9 gene editing. A –20 bp deletion was bred to homozygosity. Intercross or *rspo3*$^{+/-}$ were injected with 4 gRNAs against *rspo2* and the resulting larvae were genotyped and analyzed for phenotype. B) Whole mount ventral and lateral images of Alcian blue/Alizarin red S stained 9 dpf larvae. *rspo3*$^{-/-}$ embryos that were *rspo2* gRNA/Cas9 injected (*rspo2*$^{\Delta}$) larvae were similar to wild-type except that *rspo2*$^{\Delta}$ larvae exhibited disrupted development of the pectoral fin. Impaired fin development was exacerbated with decreasing genetic dosage of *rspo3* (black arrows, dotted yellow lines delineate fins). While craniofacial development in *rspo3*$^{-/-}$ and *rspo2*$^{\Delta}$ larvae were largely normal, *rspo3*$^{-/-}$; *rspo2*$^{\Delta}$ double mutants exhibited a dysmorphic lower jaw (white arrow). Scale bar: 100 μm. C) Quantification of pectoral fin developmental disruption. *rspo2*$^{\Delta}$ larvae tended to have disrupted development of a single pectoral fin. This effect was significantly exacerbated with decreasing genetic dosage of *rspo3*, as *rspo3*$^{-/-}$; *rspo2*$^{\Delta}$ double mutant larvae failed to develop pectoral fins altogether. D) Quantification of angle measurements between Meckel's cartilage (m) and palatoquadrate (pq). While *rspo3*$^{-/-}$ and *rspo2*$^{\Delta}$ mutants had normal lower jaw morphology, rspo3$^{+/-}$; *rspo2*$^{\Delta}$ and rspo3$^{-/-}$; *rspo2*$^{\Delta}$ mutants displayed a significantly decreased angle at the Meckel's/palatoquadrate joint. N=10-16. $p<0.01$. *indicates significance relative to wild-type.

In addition to altered fin development, we identified a subset of zebrafish with altered craniofacial morphology affecting the lower jaw. We found that *rspo2*$^{\Delta}$, *rspo3*$^{+/-}$ and *rspo3*$^{-/-}$; *rspo2*$^{\Delta}$ larvae displayed a significantly reduced angle where the palatoquadrate meets Meckel’s cartilage (**Figure 9**B, D).

To evaluate craniofacial effects in greater detail and visualize individual cartilage elements, we dissected out the ethmoid plate and ventral cartilages, including the pharyngeal teeth (**Figure 9**). Analyses of Alcian blue/Alizarin Red S zebrafish at 9 dpf revealed that disruption of *rspo2* caused a decrease in the number of pharyngeal teeth, with an average of 2 total teeth rather than the 8 teeth observed in the control (**Figure 9**B). Although *rspo3*$^{-/-}$ larvae did not exhibit a difference in the number of teeth at 9 dpf, haploinsufficiency of *rspo3* decreased tooth number in the *rspo2*$^{\Delta}$ larvae, with the *rspo3*$^{-/-}$; *rspo2*$^{\Delta}$ double mutant having no mineralized teeth at 9 dpf (**Figure 9**A, B).

Flat-mount imaging of Alcian blue/Alizarin Red S-stained ventral cartilage revealed a significant decrease in anterior-posterior/rostral length of Meckel’s cartilage in the *rspo3*$^{-/-}$ larvae while *rspo2* disruption alone had no effect (**Figure 10**A, C). The requirement for *rspo3* on Meckel’s cartilage rostral length was significantly exacerbated by *rspo2* disruption (**Figure 10**C). The effect of *rspo3* on Meckel’s cartilage rostral length is specific, rather than due to a total anterior-posterior shortening, as ceratohyal length anterior-posterior length was not different in these zebrafish (**Figure 10**D).

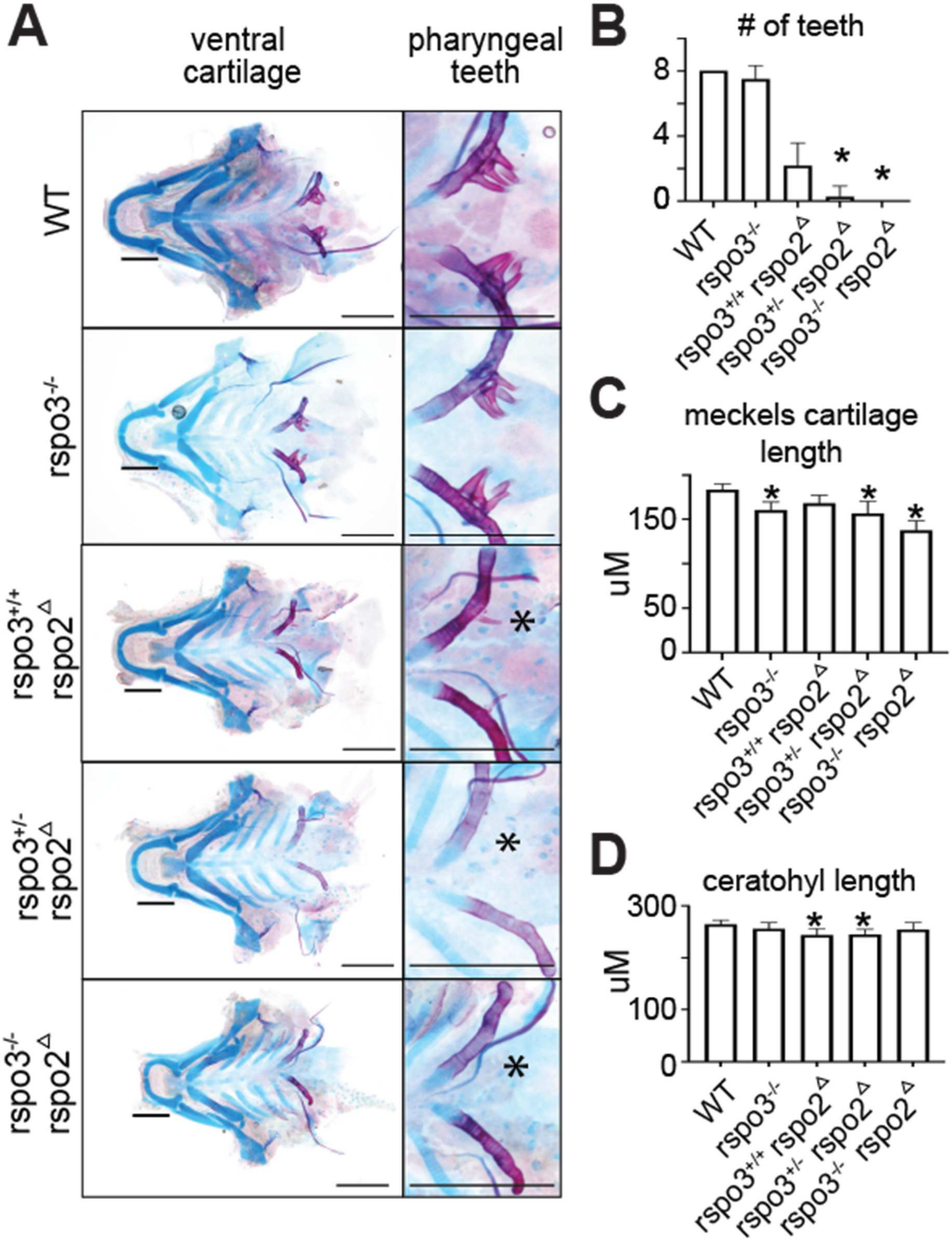


Figure 10 Synergistic effect of *rspo2* and *rspo3* ablation on zebrafish tooth development and Meckel's cartilage. A) Flat-mount images of Alcian blue/Alizarin

Figure 10 (Cont'd) red S stained 9 dpf zebrafish ventral cartilages. Zoom of pharyngeal teeth to right. *rspo3*$^{-/-}$ larvae displayed anterior shortening of Meckel's cartilage, which was exacerbated with *rspo3*$^{-/-}$; *rspo2*$^{\Delta}$ gRNA disruption (black bars) *indicate absent teeth. Scale bar: 200 µm. B) Alizarin red S staining of pharyngeal teeth shows that *rspo3*$^{-/-}$ are generally normal relative to wild-type while *rspo2*$^{\Delta}$ larvae have a reduced number of teeth (average of 2 versus 8). Tooth number in *rspo2*$^{\Delta}$ larvae decreased further with decreasing wild-type alleles of *rspo3* (zero teeth detected in *rspo3*$^{-/-}$; *rspo2*$^{\Delta}$ mutant). C) Quantification of the anterior-posterior/rostral length of Meckel's cartilage shows a primary effect in *rspo3*$^{-/-}$ larvae, which is exacerbated in *rspo3*$^{-/-}$; *rspo2*$^{\Delta}$ mutants. D) Quantification of the anterior-posterior length of ceratohyal cartilage shows no effect in rspo3$^{-/-}$ larvae, suggesting a cartilage element-specific effect of *rspo3* and *rspo2*. N=10-16. $p<0.01$. *indicates significance relative to wild-type.

***rspo3* influences osteoclast activity during zebrafish development.**

To assess the role of *rspo3* on osteogenesis in developing zebrafish we performed live Alizarin Red S staining on 10 dpf *rspo3*$^{+/+}$ and *rspo3*$^{-/-}$ larvae. The *rspo3* mutant allele was also bred onto a *sox10:kaede* background in order to visualize cartilage elements. Confocal analyses of whole mount embryos revealed no differences in Alizarin Red S intensity (**Figure 11**A). No obvious changes in cartilage morphology were observed in the *rspo3*$^{-/-}$ fish. Interestingly, we did observe increased tartrate-resistant acid phosphatase (TRAP) positive area in *rspo3*$^{-/-}$ mutants at 14 and 21 dpf, suggesting increased osteoclast number (**Figure 11**B). Therefore, these results indicate that *rspo3* may regulate aspects of bone homeostasis after larval development as the animals mature during adult life.

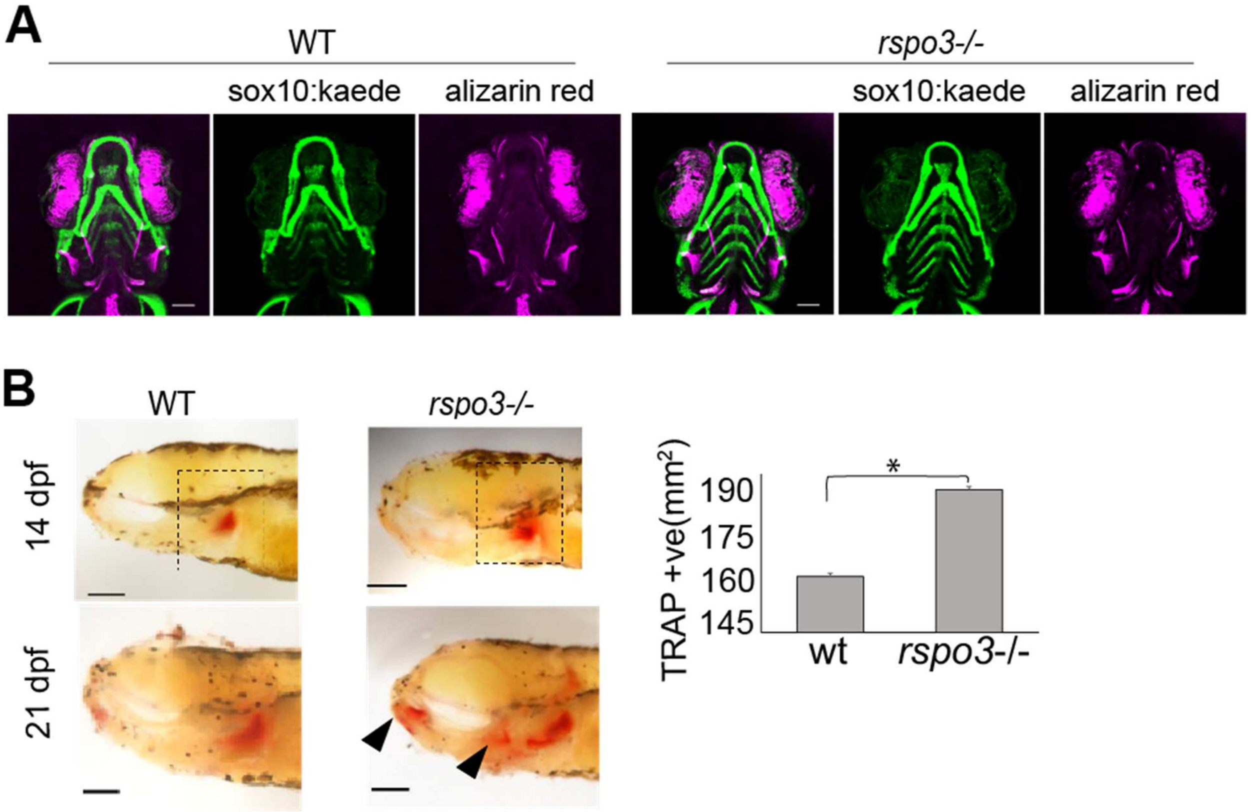


Figure 11 *rspo3* ablation does not impact initial bone mineralization but increases osteoclast area of activity. A) Representative images of maximum intensity projections of confocal z-stack of whole-mount 10 dpf zebrafish. Zebrafish express *sox10:kaede* transgene allowing fluorescent imaging of cartilage elements. Zebrafish were live-stained with Alizarin red S to visualize mineralized structures. No differences in intensity of Alizarin staining, elements stained, or cartilage architecture were noted in the *rspo3*$^{-/-}$ fish. Scale bar: 100 μm. B) Lateral view of 14 dpf showing increased osteoclast activity (red stain in the black dotted box) in *rspo3* mutant as compared to wild-type. At 21 dpf, more areas of osteoclast activity in the dentary, hyomandibular, pharyngeal teeth and jaws (solid arrow) were observed in *rspo3*$^{-/-}$ compared to wild-type. Quantification of total area of red staining. P<0.05. N=5. Scale bar: 100 μm.

Adult *rspo3* zebrafish mutants have decreased body length and exhibit a midface deficiency.

As *rspo3*$^{-/-}$ mutant zebrafish larvae matured to adult fish, we observed midface hypoplasia compared to wild-type clutch-mates (**Figure 12**A). Statistically significant differences in body length (measured from tip of mouth opening to the base of the tail, STL) were observed in *rspo3*$^{-/-}$ mutant as compared to wild-type clutch-mates (**Figure 12**B). In addition, *rspo3*$^{-/-}$ mutant exhibited significantly decreased parasphenoid and anguloarticular bone volume compared to wild-type clutch-mates (**Figure 12**C, D). The altered morphology of individual bony elements in *rspo3*$^{-/-}$ zebrafish also resulted in altered relationships between the bony elements. Cephalometric analysis revealed significant frontal bossing in *rspo3*$^{-/-}$ mutant adults, with increased parasphenoid-frontal angle (**Figure 12**E). Furthermore, we observed midface hypoplasia in adult *rspo3*$^{-/-}$ zebrafish as compared to wild-type, with significant increased distance between nasal bone and a line drawn between dentary and frontal bone landmarks (**Figure 12**F).

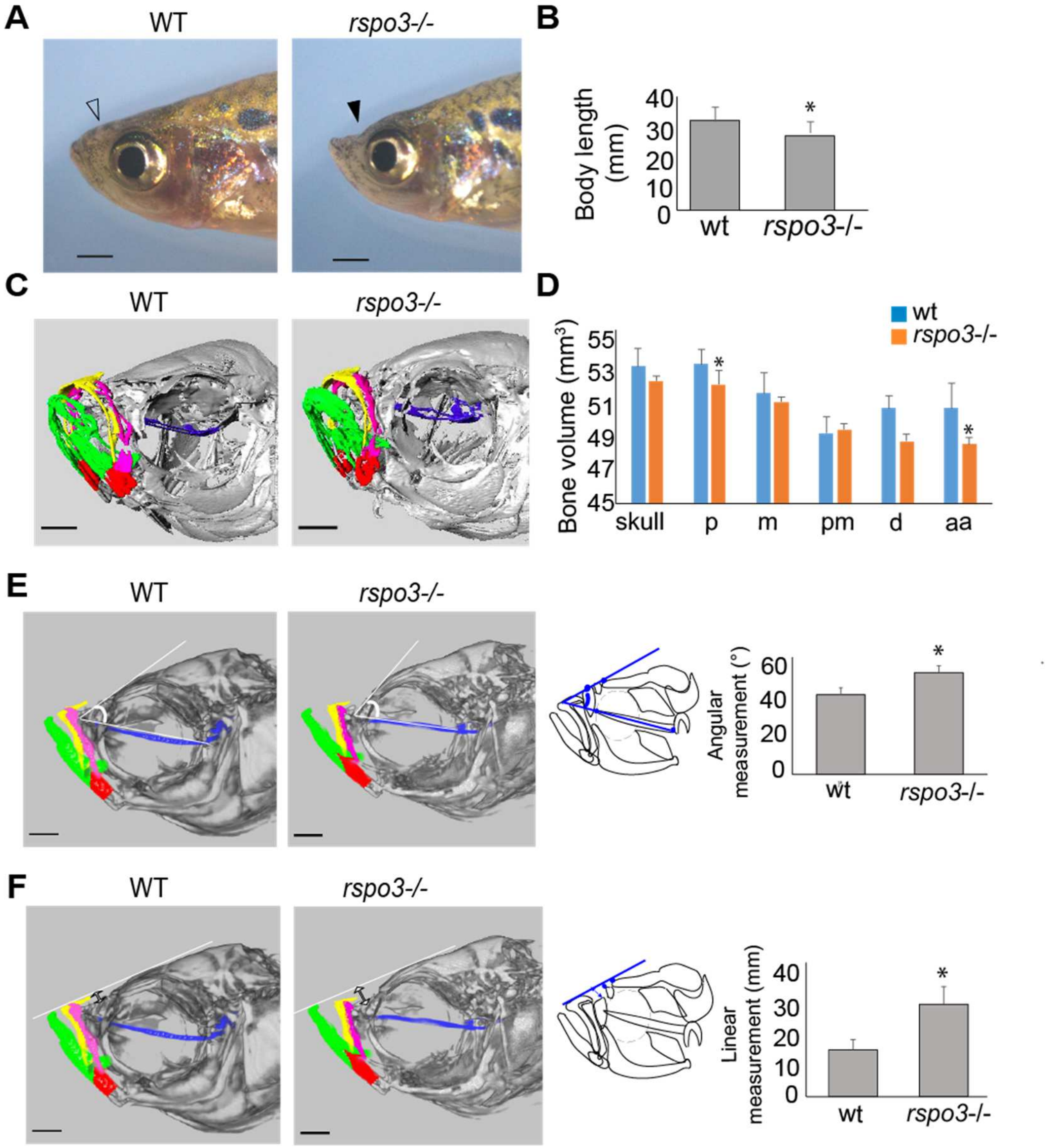


Figure 12 Figure 7. *rspo3* mutants exhibited midface deficiency, frontal bossing and decreased body length. Reduced body length, midface deficiency and frontal bossing were observed in adult *rspo3*$^{-/-}$ (180 dpf). A) Lateral image of adult zebrafish showing midface depression in *rspo3*$^{-/-}$(solid arrow) compared to wild-type (open arrow). B) Body length was significantly decreased in *rspo3*$^{-/-}$ mutants relative to wild-type. C) Oblique micro-CT image of *rspo3*$^{-/-}$ and wild-type fish at 180 dpf. Individual bone elements are color coded (blue: parashenoid, pink: maxilla, yellow: premaxilla,

Figure 12 (Cont'd) green: dentary and red: anguloarticular). Scale bar: 10 μm. D) Bone volume of the skull and of specific bones in *rspo3*$^{-/-}$ and wild-type individuals demonstrates element-specific differences in volume. Parasphenoid and anguloarticular bone volume were significantly reduced in *rspo3*$^{-/-}$ compared to wild-type fish. Abbreviations: aa: anguloarticular, d: dentary, m: maxilla, p: parasphenoid, pm: premaxilla. E) 2D cephalometric analysis obtained from micro-CT of *rspo3*$^{-/-}$ and wild-type fish. The angle formed by parasphenoid line and a line tangent to frontal bone identified frontal bossing, with increased angle in *rspo3*$^{-/-}$ compared to wild-type. Diagram of lateral view of adult zebrafish showing the angular measurement. Bar chart showing statistical differences in the angular measurement between *rspo3*$^{-/-}$ and wild-type. F) 2D cephalometric analysis of *rspo3*$^{-/-}$ and wild-type fish. The distance between nasal bone and a line drawn between dentary and frontal bone landmarks were measured. Diagram of lateral view of adult zebrafish showing the linear measurement from nasal bone to a line tangent to the frontal bone and dentary. The linear measurement value was significantly greater in *rspo3*$^{-/-}$ mutants than in wild-type indicating the presence of midface hypoplasia. $^{*}p \le 0.05$. Scale bar: 100 μm.

***rspo3* is required for normal tooth maintenance.**

Analysis of pharyngeal tooth morphology in adult zebrafish using micro-CT illustrated decreased tooth number in *rspo3*$^{-/-}$ mutant zebrafish, as compared to wild-type clutch-mates (**Figure 13**A, B). On average, *rspo3*$^{-/-}$ adult zebrafish had two fewer teeth on both the right and left sides of the jaw (**Figure 13**B). As we found no difference in the number of teeth during the larval stage in *rspo3*$^{-/-}$ animals (**Figure 10**), we suggest that *rspo3* functions in the maintenance of teeth, rather than tooth development, either by regulating tooth integrity or regulating tooth regeneration.

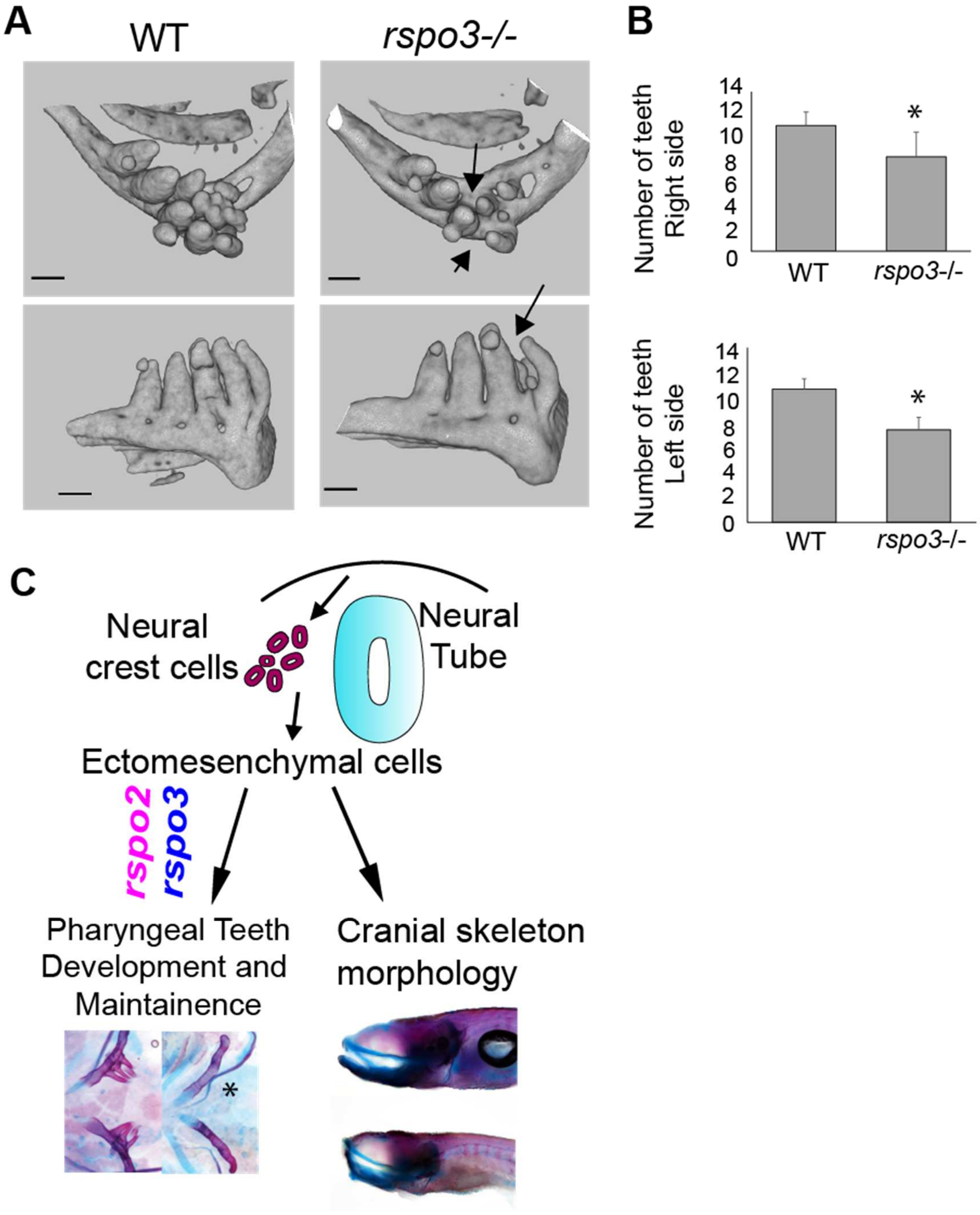


Figure 13 Adult *rspo3* mutant zebrafish have reduced teeth number. A) Micro-CT of 180 dpf zebrafish reveals that the adult *rspo3*$^{-/-}$ animals exhibit decreased tooth number with several sockets missing teeth that are

Figure 13 (Cont'd) present in the wild-type. B) *rspo3*$^{-/-}$ have significantly fewer teeth than wild-type on both right and left sides. N=9,7. *p<0.05. C) Summary diagram illustrating that *rspo2* and *rspo3* both regulate pharyngeal tooth development as well as have roles in morphogenesis of the craniofacial skeleton.

Discussion

This study reports *Rspo2* and *Rspo3* gene expressions and functions in craniofacial and dental morphogenesis, using zebrafish and mouse models. *Rspo3* is diffusely expressed through the craniofacial mesenchyme whereas *Rspo2* is expressed in distinct domains. In zebrafish, there is overlap in *rspo2* and *rspo3* gene expressions, whereas in the mouse embryo the expression domains of these paralogs are distinct. We showed that *Rspo3* is expressed in perichondral cells, and *Runx2* positive osteoprogenitors in embryonic palate and Meckel's cartilage in zebrafish, as well as in mouse osteoprogenitors. In zebrafish teeth, *rspo3* is expressed in newly formed replacement teeth, where it is broadly expressed in dental pulp, odontoblasts, and crypt epithelium. Analysis of adult *rspo3*$^{-/-}$ zebrafish suggest that *rspo3* is required for adult teeth maintenance. Loss of *rspo3* did not affect larval osteogenesis but did result in increased area of TRAP staining, midface hypoplasia, and reduced numbers of attached teeth in adult zebrafish. Importantly, *rspo2* and *rspo3* genetically interact, where haploinsufficiency of *rspo3* exacerbates defects in tooth formation and pectoral fin bud extension. We show that zebrafish *rspo2* and *rspo3* are required for limb development, analogous to the mouse function of *Rspo2/3* that was previously reported.[113] Together, these gene expression studies and genetic analyses are consistent with functions for *rspo3* in progenitor cell populations contributing to the craniofacial skeleton and teeth, and in the maintenance of craniofacial bones and teeth in zebrafish (**Figure 13**C).

We showed that high resolution gene expression analysis using RNAscope on sectioned specimens provided greater resolution as compared to whole mount (ISH) approaches (compare **Figure 6**A and **Figure 6**B). Using RNAscope, we were able to determine that *rspo2* and *rspo3* transcripts were detected in a cell layer that surrounds the chondrogenic elements in the zebrafish anterior neurocranium, trabeculae and in Meckel's cartilage. Moreover, *Rspo3* is co-expressed with *Runx2* (osteoprogenitor marker) in zebrafish and mouse (**Figure 6**, **Figure 7**). In support of our findings, a human genetic study recently reported the involvement of *RSPO3* in bone mineral density and bone fractures.[122] Furthermore, *RSPO3* was reported to regulate osteoblastic differentiation.[123] In addition, human *RSPO3* was identified as a candidate gene that contributes to cleft lip/palate and dental anomalies, consistent with its role in skeletal development and in human adipose-derived stem cells.[112, 123] Taken together, this and other studies corroborate that *Rspo3* has conserved functions in the development of craniofacial bone and tooth structures across vertebrates.

This study also identified a key requirement for *rspo3* in regulating tooth development. Zebrafish teeth are continuously replaced through its life, where the regenerative process is analogous to human adult tooth replacement of a deciduous, baby tooth and to mouse continuously growing incisors.[121, 124, 125] We described *rspo3* gene expression in dental pulp, odontoblasts and crypt and dental epithelium in zebrafish and mouse, suggesting possible roles in the regulation of tooth development, odontogenesis and ameloblast differentiation. Importantly, *rspo3* is highly detected in zebrafish replacement teeth as compared to mature teeth, indicating potential roles in dental progenitor cell populations as compared to more differentiated dental cell types. Moreover, adult *rspo3*$^{-/-}$ zebrafish exhibit reduced attached tooth numbers as compared to age-matched wild- type fish. The normal tooth formation in *rspo3*$^{-/-}$ mutant zebrafish at 9 dpf while having reduced tooth number at 180 dpf suggests a role for rspo3 in the maintenance of adult teeth rather than in their initial development. Differences in *Rspo2/3* spatiotemporal gene expression in mouse and zebrafish may reflect differences in the regenerative odontogenic potential of zebrafish. Wnt/β- catenin signaling is important for tooth morphogenesis, and consistent with the observation that *rspo3* disruption resulted in inhibited dental tissue development.[118]

Adult *rspo3*$^{-/-}$ zebrafish exhibited midface hypoplasia, frontal bossing and reduced tooth number as compared to aged-matched wild-type clutch-mates. Our data showed co-expression of *rspo3* and *col1a1a* during embryogenesis, suggesting that these two genes could be functionally associated. Consistent with this result, previous studies reported that patients with osteogenesis imperfecta have mutations in *COL1A1A* which is characterized by frontal bossing, midface hypoplasia and dentinogenesis imperfecta.[126-128] Future studies are recommended to investigate the molecular mechanisms regulated by *rspo3*, including its interactions with Wnt signaling pathway genes in regulating dental and bone development.

Chapter III: An Irf6-Esrp1/2 regulatory axis controls midface morphogenesis in vertebrates

Introductory Statement

We present our work in establishing a regulatory axis between IRF6, and the epithelial-specific splicing regulators Esrp1 and Esrp2 published in *Development* in 2020. In this article, we applied both mouse and zebrafish models to study the gene expression patterns of *Irf6* and *Esrp1/2* and characterized the phenotype of *irf6* and *esrp1/2* loss-of-function (LOF) zebrafish models. This work was a collaborative effort between a mouse geneticist, Shannon Carroll, who performed the mouse experiments, and my design and experimental approach for the zebrafish work. Edward Li performed the RNA-sequencing experiments in *irf6*$^{-/-}$ embryos and initial experiments for the dominant-negative irf6 optogenetic system. Specifically, I optimized the dominant-negative irf6 optogenetic system previously formulated by Edward Li, characterized the *esrp1*$^{-/-}$*; esrp2*$^{-/-}$ phenotype, developed a morpholino-based approach to phenocopy the *esrp1/2* LOF phenotype, and performed the lineage tracing experiments on the *esrp1*$^{-/-}$; *esrp2*$^{-/-}$ zebrafish in the Tg(sox10:kaede) reporter line. This article establishes the Esrp1/2 proteins as downstream effectors of a regulatory axis modulated by Irf6, and implicates alternative splicing of factors as a critical step in the regulation of craniofacial development by the embryonic periderm.

An *Irf6-Esrp1/2* regulatory axis controls midface morphogenesis in vertebrates

Shannon H. Carroll[1,2,#], Claudio Macias Trevino[1,3,#], Edward B-H Li[3], Kenta Kawasaki[1,2], Nikita Myers[1], Shawn A. Hallett[1], Nora Alhazmi[4], Justin Cotney[7], Russ P. Carstens[6] and Eric C. Liao[1,2,3,*]

[1]Center for Regenerative Medicine, Massachusetts General Hospital, Boston, MA 02114, USA.

[2]Shriners Hospital for Children, Boston, MA 02114, USA.

[3]Harvard Medical School, Boston, MA 02115, USA.

[4]Harvard School of Dental Medicine, Boston, MA 02115, USA.

[5]Division of Plastic and Reconstructive Surgery, Massachusetts General Hospital, Boston, MA 02114, USA.

[6]Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA.

[7]Department of Genetics and Genome Sciences, University of Connecticut Health, CT 06030, USA.

[#]Authors contributed equally to this work

[*]Author for correspondence: cliao@partners.org

Telephone: ECL: (617) 643-5975

Key words: *IRF6*, *ESRP1*, craniofacial, cleft, development

Summary statement

These studies elucidate *Irf6* and *Esrp1/2* function in periderm and embryonic epithelium during palate development, using complementary mouse and zebrafish models. Targeted disruption of *Irf6* and *Esrp1/2* in mouse and zebrafish resulted in orofacial clefts, where an aberrant mesenchymal/epithelial cell population is identified in the cleft of the *esrp1/2* zebrafish mutant.

Abstract

Irf6 and *Esrp1* are important for palate development across vertebrates. In zebrafish, we found that *irf6* regulates the expression of *esrp1*. We detailed overlapping *Irf6* and *Esrp1/2* expression in mouse orofacial epithelium. In zebrafish, *irf6* and *esrp1/2* share expression in periderm, frontonasal ectoderm, and oral epithelium. Genetic disruption of *irf6* and *esrp1/2* in zebrafish resulted in cleft of the anterior neurocranium. The *esrp1/2* mutant also developed cleft of the mouth opening. Lineage tracing of cranial neural crest cells revealed that cleft resulted not from migration defect, but from impaired chondrogenesis. Analysis of aberrant cells within the cleft revealed expression of *sox10, col1a1* and *irf6* and were adjacent to cells *krt4 and krt5* positive. Breeding of mouse *Irf6;Esrp1;Esrp2* compound mutants suggested genetic interaction, as the triple homozygote and the *Irf6*;*Esrp1* double homozygote was not observed. Further, *Irf6* heterozygosity reduced *Esrp1/2* cleft severity. These studies highlight the complementary analysis of *Irf6* and *Esrp1/2* in mouse and zebrafish and captured a unique aberrant cell population expressing *sox10, col1a1 and irf6*. Future work characterizing this cell population will yield additional insight into cleft pathogenesis

Introduction

Development of vertebrate craniofacial structures requires coordinated cellular induction, migration, proliferation, and differentiation, which allow for the positioning of adjacent epithelial-lined facial processes that ultimately merge [19, 25, 35, 52-57]. Morphogenesis of facial structures such as the midface, lip, and palate requires convergence of the medial and lateral nasal prominences and the fusion of the secondary palatal shelves at the midline [19, 25, 58]. Failure of these processes to fuse results in orofacial clefts (OFCs) of the lip, primary palate, or secondary palate [59]. Orofacial clefts are among the most common congenital structural anomalies [38, 60, 61]. From genome-wide association studies carried out over a decade ago to more recent whole-genome sequencing projects of orofacial cleft cohorts, cleft-associated genetic loci continue to be identified, where the transcription factor *IRF6* is one of the most commonly associated genes [45, 62-64]. *IRF6* disruption is causal for syndromic cleft in Van der Woude and Popliteal Pterygium syndromes and associated with non-syndromic orofacial clefts [40, 45, 65, 66].

Several IRF6 transcriptional targets such as *GRHL3*, *WDR65*, *OVOL1, KLF4* have been identified, which are also important for palate development and implicated in human cleft pathogenesis [49, 67-69]. These studies support the premise that investigation of *Irf6* and its transcriptional network will identify key genes that regulate palate development. Multiple mouse models have been generated to investigate *Irf6* function, including a total *Irf6* knockout and substitution of key functional residue *Irf6*^{R84C} in the DNA-binding domain [50, 70]. These *Irf6* mutant mouse models exhibited disrupted epithelial terminal differentiation and lack of a functional periderm, leading to pathological adhesions of epithelial embryonic tissues [46, 50, 70, 71]. The epithelial differentiation and adhesion defects are thought to prevent elevation of the palatal shelves, and ultimately these mice develop a cleft in the secondary palate. Additionally, the midfaces of these mice were hypoplastic, a phenotype that was attributed to the dysfunctional embryonic epithelium [50, 70].

Epithelial splicing regulatory proteins 1 and *2* (*Esrp1*, *Esrp2*) are also important in embryonic epithelial differentiation and palate development [72-74]. *Esrp2* and its homolog *Esrp1* are regulators of RNA splicing that are specifically expressed in the epithelium [75]. *Esrp1/2* knockout mice exhibit bilateral cleft of the lip and primary palate, as well as a secondary palate cleft [74]. *Esrp1/2* are unusual among regulators of RNA-splicing in that they are tissue-restricted and exhibit dynamic expression during embryogenesis [74, 76]. The developmental importance of *Esrp1/2* is underscored by their conservation across species, from ascidians to zebrafish, *Xenopus*, mouse, and humans [76]. Gene variant in *ESRP2* was also recently reported in human orofacial cleft cohorts [64].

The mouse has been an important experimental model to study craniofacial and palate morphogenesis [59]. Secondary palate development in the mouse is similar to humans, with the analogous stages of vertical outgrowth, elevation, horizontal growth, and fusion [59, 61]. Many genes associated with cleft lip and palate (CL/P) in humans, when disrupted in the mouse, result in cleft of the secondary palate without affecting morphogenesis of the primary palate and lip appear [59, 77]. So while the mouse model can be useful to study the secondary palate, the use of mouse models to study cleft of the lip and primary palate has been less effective as there are remarkably few mouse models where development of the lip and primary palate are perturbed [59]. Meanwhile, clinically cleft of the lip and palate (CL/P) is more common than isolated cleft of the palate only (CPO), and human genetic studies have suggested that the genetics underpinning CL/P and CPO are distinct [59, 61]. The developmental processes of outgrowth of the facial prominences followed by convergence and fusion are thought to be conserved across mammals [19]. Therefore, it is hypothesized that differences in mouse versus human phenotypic presentation are due to spatiotemporal differences in craniofacial development [59]. In this context, the phenotype of bilateral clefts affecting the lip, primary and secondary palate in the *Esrp1/2* mutant mouse is unique among mouse models and is a valuable tool to study lip and palate morphogenesis.

Zebrafish has been favored as an animal model by embryologists to study craniofacial development due to its accessibility, transparency and genetic tractability [78-80]. Although a secondary palate, which partially or entirely separates the oral and nasal cavities, is reserved to amniotes, the primary palate is appreciably conserved across vertebrates [19]. The primary palate establishes the intact upper jaw [19], which in the larval zebrafish consists of the ethmoid plate, also known as the anterior neurocranium (ANC). In all vertebrates, the most anterior cranial neural crest cells (CNCCs), that migrate rostral then turn caudal and ventral to the eye, contribute to the median frontonasal prominence, and a second CNCC stream, that migrate inferior to the eye and into the first pharyngeal arch, generate the paired maxillary prominences [22, 35, 80, 81]. The ANC of the zebrafish is formed from the convergence of the median element, that is derived from the frontonasal prominence, and paired lateral elements that are derived from the maxillary prominences [22, 53, 82]. Zebrafish homologs of human genes associated with orofacial clefts will disrupt morphology of the ANC, as have been observed for a number of genes such as *capzb*, *pitx2*, *pdgfra*, *smad5*, *tgfb2*, *fgf10a* and *wnt9a* [53, 77, 82].

Here, we carried out detailed gene expression analysis of *Irf6*, *Esrp1* and *Esrp2* in mouse and zebrafish in order to understand the comparative morphogenesis of facial structures and periderm between these important vertebrate genetic models. We analyzed and compared the *Irf6* and *Esrp1/2* mutant phenotypes to elucidate the comparative morphologies and genetic epistasis between these genes. Further, we generated zebrafish *irf6* and *esrp1/2* zebrafish mutants and examined their requirement in morphology of the stomodeum opening and ANC. Interestingly, we identified an aberrant cell population with epithelial and mesenchymal molecular signatures that localized to the region of the ANC cleft. This work highlights the relative strengths of the mouse and zebrafish models for investigating the morphogenetic mechanisms of orofacial clefts and contributes new insights into the function of *Irf6* and *Esrp1/2* during palatogenesis.

Results

irf6 null zebrafish embryos have decreased expression of *esrp1*

We previously generated a functionally null *irf6* zebrafish allele [51]. Using CRISPR/Cas9, an 8bp deletion in exon 6 of the *irf6* coding region resulting in a frameshift and premature stop codon, leading to the ablation of *irf6* function. It was observed that embryos lacking maternally expressed *irf6* exhibited epiboly arrest and periderm rupture at 4-5 hours post-fertilization (hpf) [51]. Utilizing this *irf6*-null model, we aimed to identify genes that were differentially downregulated in *irf6*-null versus wild type (WT) embryos. We performed RNA-seq on WT and maternal/zygotic *irf6*-null (mz-*irf6*$^{-8bp/-8bp}$) embryos at 4.5hpf, just before embryo rupture at the onset of gastrulation. Differential expression analysis reveals a substantial number of significantly differentially expressed genes (DEGs; n=10299, p-adjust < 0.05) (**Figure 14**A-C). To visualize changes in this large number of DEGs a heat map was generated which illustrates 1377 upregulated and 1799 downregulated genes with an absolute fold change greater than 2 in *irf6*-null relative to WT (**Figure 14**A). The patterns of gene expression amongst these strongly differentially expressed genes were highly reproducible across the two genotypes, and demonstrated relatively similar numbers of down and upregulated genes. When we visualized DEGs using significance values relative to fold change in expression, we found that the most significant and strongest effects on gene expression were biased toward those downregulated in the irf6-null embryos. (**Figure 14**B). The RNA-seq results revealed significant downregulation of genes previously known to be downregulated with disruptions in *Irf6* function (**Figure 14**B,C). Disruption of *irf6* via injecting dominant-negative *irf6* mRNA led to downregulation of many periderm-enriched genes (including *grhl1*, *krt5*, *krt18*, *tfap2a* and *klf2b*) and genes for adhesion molecules (including claudins and cadherins) [69]. Here we found a similar expression profile in the mz-*irf6*$^{-8bp/-8bp}$ embryos relative to WT (**Figure 14**B,C).

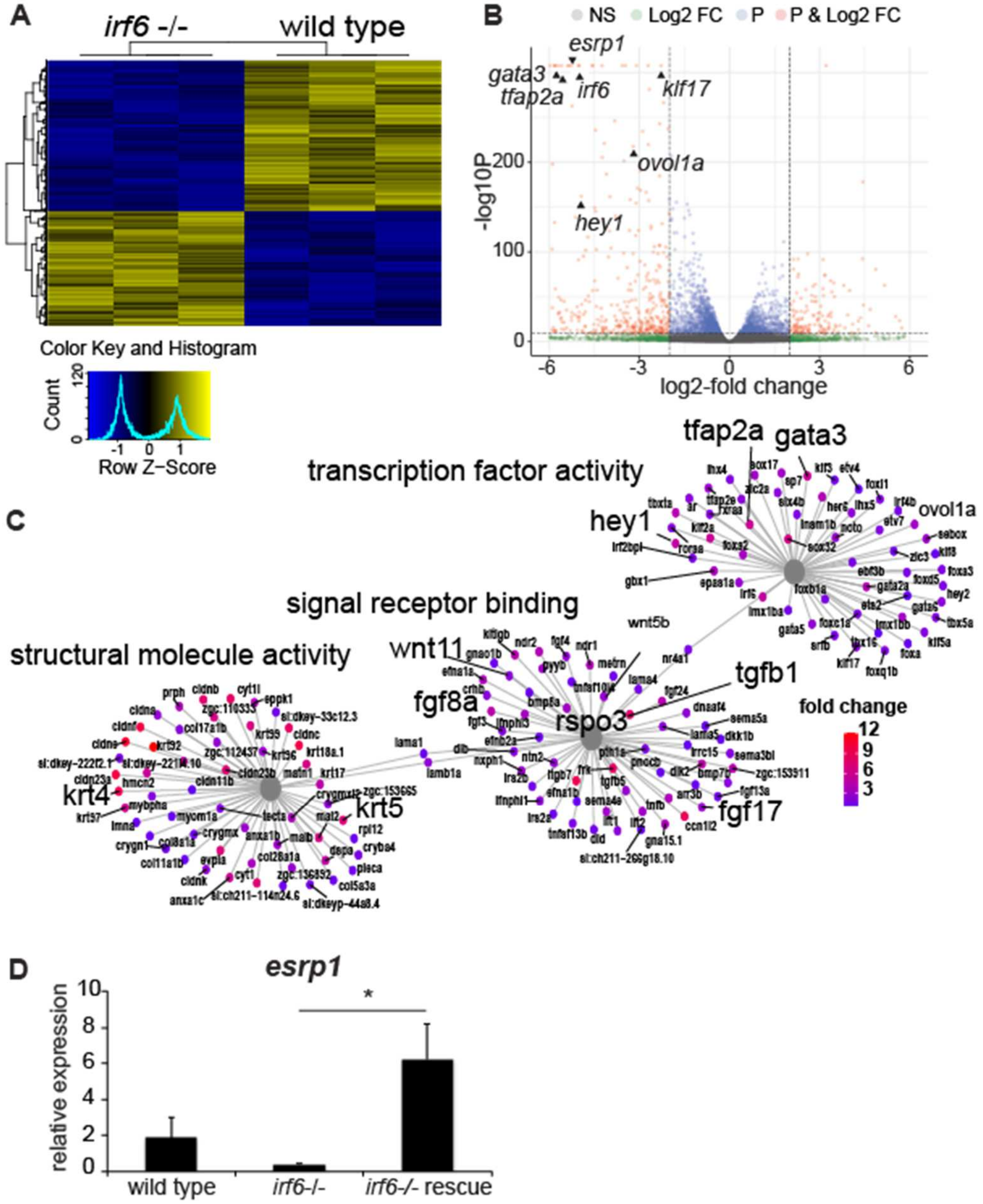


Figure 14 esrp1 expression is downregulated in irf6 null zebrafish embryos. (A) Hierarchical clustering of top differentially expressed genes (DEGs) defined by RNA-

Figure 14 (Cont'd) sequencing performed on WT vs. mz-irf6-8bp/-8bp (irf6-/-) zebrafish embryos at 4-5hpf. Top DEGs were identified by selecting genes with an adjusted p-value (Benjamini-Hochberg) <0.01 and absolute log2-fold change > 2. Data shown for 3 biological replicates. Color scale at top left represents relative levels of expression with yellow showing higher expression levels and blue showing lower expression. **(B)** Volcano plot from the RNA-seq dataset showing the distribution of DEGs based on p-values and log2-fold change. Previously published irf6-regulated genes are expressed at significantly higher level in WT relative to mz-irf6-/-, including grhl3, klf17, and wnt11. The newly identified cleft-associated gene esrp1 is also expressed significantly higher in ST relative to irf6-/-. Horizontal and vertical lines represent the p-value cutoff of 0.01 and the log2-fold change cutoff of 2, respectively. **(C)** Gene ontology (GO) gene-concept network analysis of RNA-seq data showing that irf6-/- embryos have perturbations in processes such as transcription factor activity, signal receptor binding, and structural molecule activity. Note that many of these genes, such as wnt11, fgf8, tgfb1, krt4, and krt5, are implicated in ectoderm development and cell specification. Grey nodes show GO terms, colored nodes show individual genes from the RNA-seq dataset, and edges connect genes to one or more associated GO terms. Colored nodes show relative enrichment (measured by fold-change) of genes in WT samples relative to irf6-/- embryos. Maps were generated using the enrichplot package in R. **(D)** qPCR gene expression analysis for esrp1, showing approximately 6-fold esrp1 downregulation in mz-irf6-8bp/-8bp embryos compared to WT at 4 hpf, and rescued esrp1 gene over-expression in mz-irf6-8bp/-8bp embryos injected with WT zebrafish irf6 mRNA. n = 4. Students t-test, *p<0.05.

To further understand the molecular pathways and biological functions being affected in the *irf6*-null embryo we performed gene ontology (GO) analyses on upregulated and downregulated gene sets. Of particular interest were significant changes in genes enriched for functions related to transcription factor activity, signal receptor binding and structural molecule expression (**Figure 14**C). When compared to previously published *IRF6* siRNA human keratinocyte DEG expression data [83], there were major overlaps of genes in molecular pathways responsible for epithelial regulation, including *gata3*, *krt18* and *cldn4* (Figure 14B,C). Many key developmental signaling pathways including Fgf (*fgf8a*, *fgf17* and *fgf24*) and Wnt (*wnt11*, *dact2*, *rspo3*, *frzb*, *fzd5*) pathways were also heavily represented in our dataset as genes downregulated due to *irf6* ablation (Figure 14B,C). Further, a number of genes associated with human orofacial clefts (OFCs) are also downregulated in the *irf6* null embryos, including *hey1*, *gata3*, *wnt11*, and *fgf8* (Figure 14B,C).

Interestingly, one of the most downregulated genes was *esrp1*. *Epithelial splicing regulatory protein 1* (*esrp1)* and its paralog *esrp2* are epithelial restricted RNA splicing regulators. *ESRP2* genetic variants in humans are associated with OFCs [64], and *Esrp1* and *Esrp1/2* knockout mice display a bilateral cleft of the lip, primary and secondary palate [72, 74]. To confirm the RNA-seq results, we performed qPCR on mz-*irf6*$^{-8bp/-8bp}$ and WT embryos at 4-5hpf. Relative to WT, mz-*irf6*$^{-8bp/-8bp}$ embryos had approximately 5-fold downregulation of *esrp1* expression. Additionally, injection of mz-*irf6*$^{-8bp/-8bp}$ embryos with *irf6* mRNA at the 1-cell stage rescued *esrp1* expression, resulting in increase that was approximately 3-fold higher than WT (Figure 14D). These rescued fish were phenotypically normal, as previously shown [51].

We tested *Esrp1* and *Esrp2* mRNA expression in E11.5 *Irf6* mutant mouse embryos (*Irf6*$^{R84C/R84C}$) and found expression to be significantly decreased relative to littermate WTs. Additionally, *Shh* expression was decreased in *Irf6*$^{R84C/R84C}$ embryos, consistent with a previous report of decreased *Shh* expression in *Esrp1*$^{-/-}$ mice [72]. These results in zebrafish and mouse suggest that *Esrp1* gene expression is dependent on *Irf6*, either through direct regulation or the requirement of a normal periderm.

***irf6*, *esrp1* and *esrp2* are co-expressed in the oral epithelium of zebrafish during craniofacial development**

Previous mouse studies have described *Irf6* [41] and *Esrp1*/*2* [72, 74, 75] gene expression in oral epithelium during palate development. To determine the gene expression of *irf6* and *esrp1/2* in the zebrafish during epithelial and craniofacial development, we performed whole-mount *in situ* hybridization (WISH). Maternal deposition of *irf6*, *esrp1* and *esrp2* mRNA was detectable at 8-cell stage (**Figure 15**A). The maternal transcripts were also detected in the periderm of the gastrulating embryo, although expression of *esrp2* appears lower than *irf6* and *esrp1* (**Figure 15**A). During craniofacial development WISH demonstrated specific expression of *irf6*, *esrp1,* and *esrp2* lining the embryonic oral epithelium, and circumscribing surface epithelium concentrated around the developing stomodeum (**Figure 15**B, C).

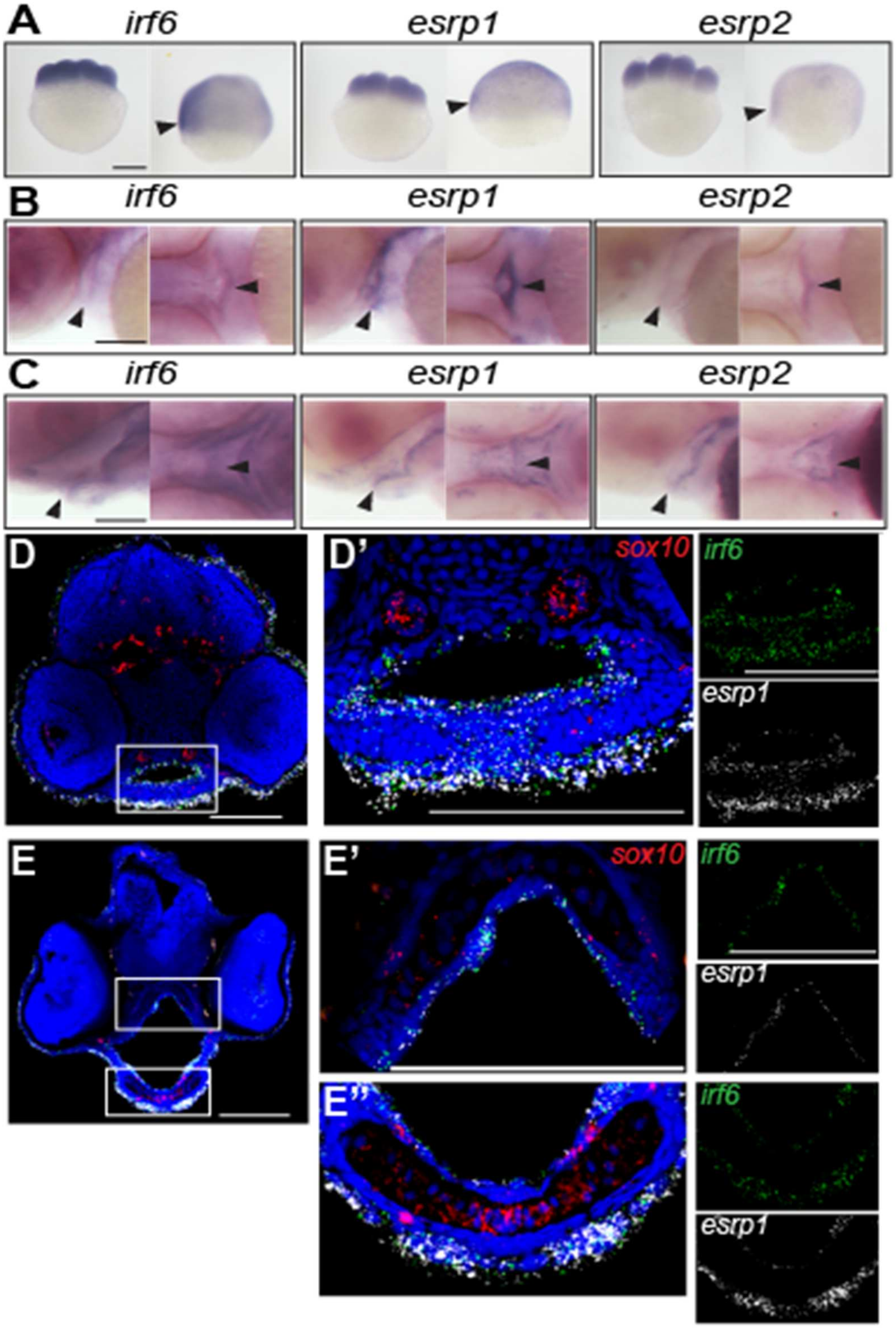


Figure 15 *irf6, esrp1* and *esrp2* are co-expressed in the oral epithelium of zebrafish embryos. (A-C) Whole mount *in situ* hybridization showing *irf6*, *esrp1* and

Figure 15 (Cont'd) *esrp2* maternal deposited transcripts are detected at the 8-cell and shield stage (A; arrow depicts periderm) circumscribes the developing stomodeum and lines the oral epithelium of zebrafish embryos at 48 (B) and 72 (C) hpf (depicted by arrow). All whole-mount embryos are oriented with anterior toward left of page, dorsal toward top of page. Coronal sections of 48 (D) and 72 (E) hpf embryos analyzed by RNAscope *in situ* hybridization, (dorsal toward top of page) showing cellular RNA co-expression of *irf6* (green) and *esrp1* (white) in surface and oral epithelial cells. *sox10* (red) staining depicts cartilage elements of the palate. Scale bars: 250µm (A) and 100µm (B-E).

To resolve the specific cell populations that express *irf6*, *esrp1,* and *esrp2*, we performed RNAscope *in-situ* hybridization of coronal cryosections taken through the developing mouth and palate at 48 and 72hpf. We found that *irf6* and *esrp1* were co-expressed within epithelial cells lining the oral cavity as well as the surface epithelium (**Figure 15**D, E). No expression of these genes was detected within the cartilage elements, identified by *sox10* expression. Further, we detected *irf6* and *esrp1* transcripts within the same cells, importantly within cells separating adjacent mesenchymal elements (**Figure 15**D'); these cells are likely in the epithelial lineage as *esrp1* is an epithelia-specific gene.

***Irf6*, *Esrp1,* and *Esrp2* are co-expressed in murine frontonasal and oral epithelium during palate and lip development**

Irf6 expression within the embryonic oral epithelium and surrounding the developing palatal shelves has been well established [41, 71, 84]. *Esrp1/2* expression was previously shown in the oral epithelium of developing mice [72, 75, 85]. Ablation of *Irf6* or *Esrp1/2* causes a cleft of the secondary palate, but the disruption of the lip and primary palate phenotypes differ between the *Irf6* and *Esrp1/2* mutants [50, 70, 72, 74]. To determine whether *Irf6, Esrp1* and *Esrp2* transcripts co-localize during mouse craniofacial development, we performed WISH for each gene at E10.5, as the frontonasal prominences and lambdoidal junction are taking shape at this time point. We found that *Irf6*, *Esrp1,* and *Esrp2* were expressed similarly with high levels of expression in areas of craniofacial development (**Figure 16**A). The mouse gene expression pattern was similar to that observed in zebrafish, with more concentrated expression to the developing head. Higher-resolution imaging with RNAscope ISH detected *Irf6*, *Esrp1,* and *Esrp2* transcripts in the periderm and the basal epithelium across all time points examined (**Figure 16**B-F). *Irf6*, *Esrp1* and *Esrp2* were co-expressed in the surface ectoderm overlying the developing frontonasal prominences (**Figure 16**B), a cell population with important signaling and inductive functions [36]. Further, co-expression included cells at critical fusion points, specifically between the medial and lateral nasal prominences (**Figure 16**C) and the palatal shelves (**Figure 16**E, F). The co-expression of *Irf6* and *Esrp1* and *2* within cells with key roles during epithelial fusion supports the existence of an *Irf6/Esrp* regulatory axis during craniofacial morphogenesis.

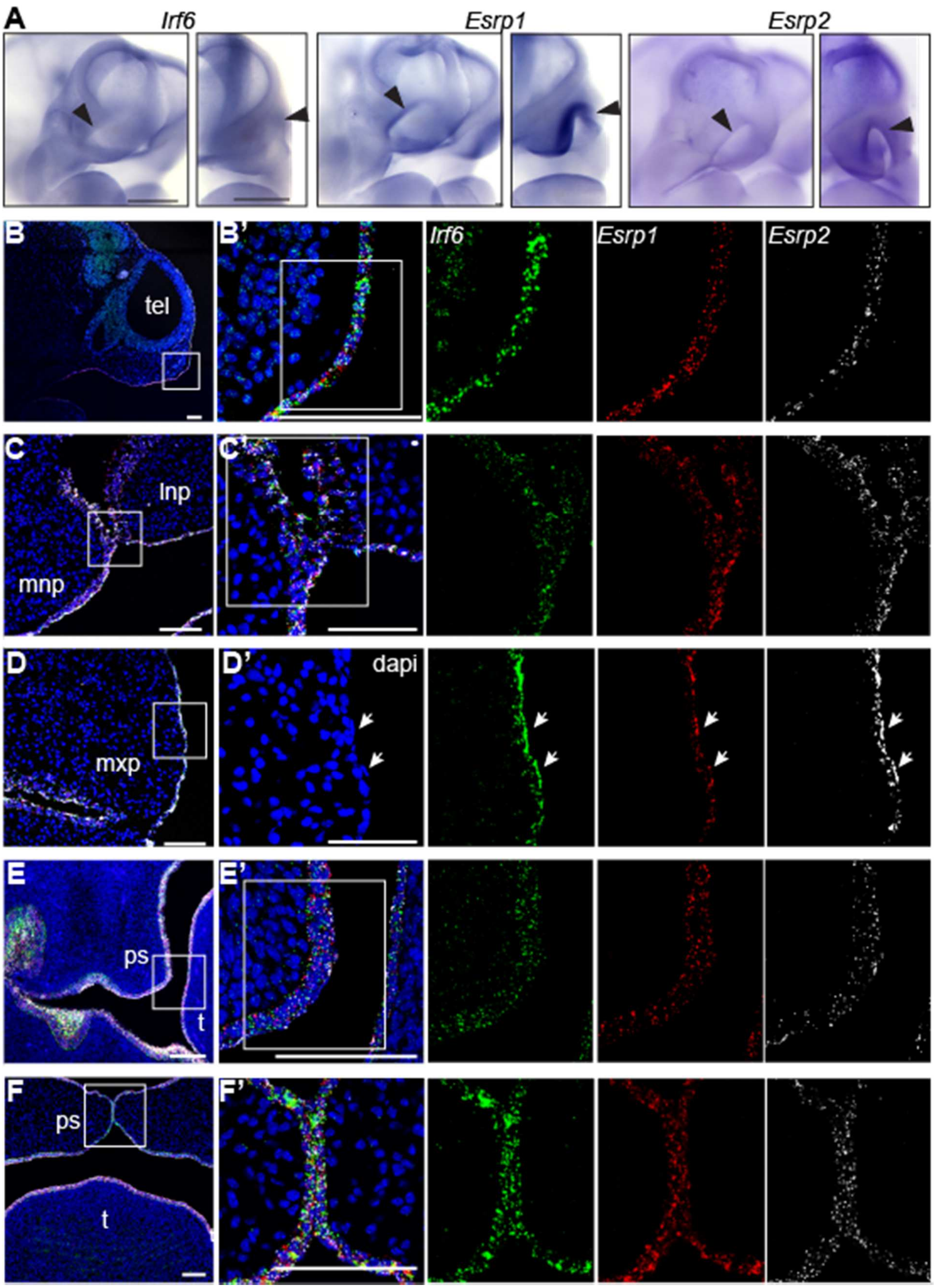


Figure 16: ***Irf6*****,** ***Esrp1*** **and** ***Esrp2*** **are co-expressed in the oral epithelium of mouse embryos.**

Figure 16 (Cont'd) (A) Whole mount *in situ* hybridization of E10.5 embryos showing *Irf6*, *Esrp1* and *Esrp2* mRNA expression in the surface epithelium and concentrated within the ectoderm of the frontonasal prominences and first brachial arch. Oblique and frontal orientation. Scale bars: 500 µm. (B-F) Sections of (B) E10, (C,D) E11.5, (E) E13.5 and (F) E15 embryos analyzed by RNAscope *in situ* hybridization showing mRNA cellular co-expression of *Irf6* (green), *Esrp1* (red) and *Esrp2* (white) in the surface ectoderm (E10), lining the frontonasal and maxillary prominences, including expression in periderm (arrows) (E11.5) and lining the palatal shelves (E13.5, E15). (B) sagittal and (C-F) coronal sections. Scale bars: 100 µm.

Interestingly, in addition to *Irf6* expression in the epithelium, RNAscope detected *Irf6* mRNA expression in the mesenchyme, particularly at E10 and E15.5 (**Figure 16**B, E). Expression of *Irf6* in this craniofacial mesenchyme has not been reported previously, where high transcript detection with RNAscope may be delineating gene expression not previously observed. Non-epithelial *Irf6* expression was detected in CNCCs of the first and second pharyngeal arches at E9 [86] and *Irf6* is expressed in cells of the developing tongue [87]. Further, we previously reported that zebrafish expressing the *irf6*^{R84C} variant under a *sox10* promoter exhibit a partial cleft of the ANC [29]. Together, these results suggest an additional role of *Irf6* in craniofacial development beyond its role in epithelial cell differentiation.

Disruption of *irf6* during neural crest cell migration results in cleft in zebrafish

Germline mutation of *irf6* results in early embryonic lethality due to periderm rupture, which precluded evaluation of palate morphogenesis [51, 88]. To circumvent embryonic lethality, we employed an optogenetic gene activation system based on the light-sensitive protein EL222, that serves to induce the expression of genes downstream of the C120 promoter [89]. To this end, a dominant-negative form of *irf6* consisting of a fusion protein of the *irf6* protein-binding domain and the engrailed repressor domain (*irf6*-ENR) was cloned downstream of the C120 promoter (C120-*irf6*-ENR; Fig. 4A) [88]. When co-injected with VP-16 mRNA, this light activated *irf6*-ENR construct enabled us to control the timing of *irf6* disruption by exposing the embryos to a 465 nm light-source later in embryogenesis (**Figure 17**A). Zebrafish embryos injected with the optogenetic system and continuously exposed to blue light from 10-96hpf were able to survive but developed with a slightly curved body axis and a dysmorphic ventral cartilage phenotype, which were not observed in control injected embryos or injected embryos that were raised in the dark (**Figure 17**B). Further analysis of the cartilage in these embryos revealed a cleft in the ANC where a population of cells in the median portion was absent (**Figure 17**C). Moreover, injecting increasing doses of EL222-VP-16 mRNA and/or C120-*irf6*-DN plasmid led to a dose-dependent effect in the proportion of zebrafish embryos with a cleft phenotype that is more pronounced for injected embryos grown in blue light starting at 10hpf compared to embryos grown in the dark. Consistent with decreased expression of *esrp1* in mz-*irf6*$^{-8bp/-8bp}$ embryos, disruption of *irf6* using this optogenetic system resulted in decreased expression of *esrp1*.

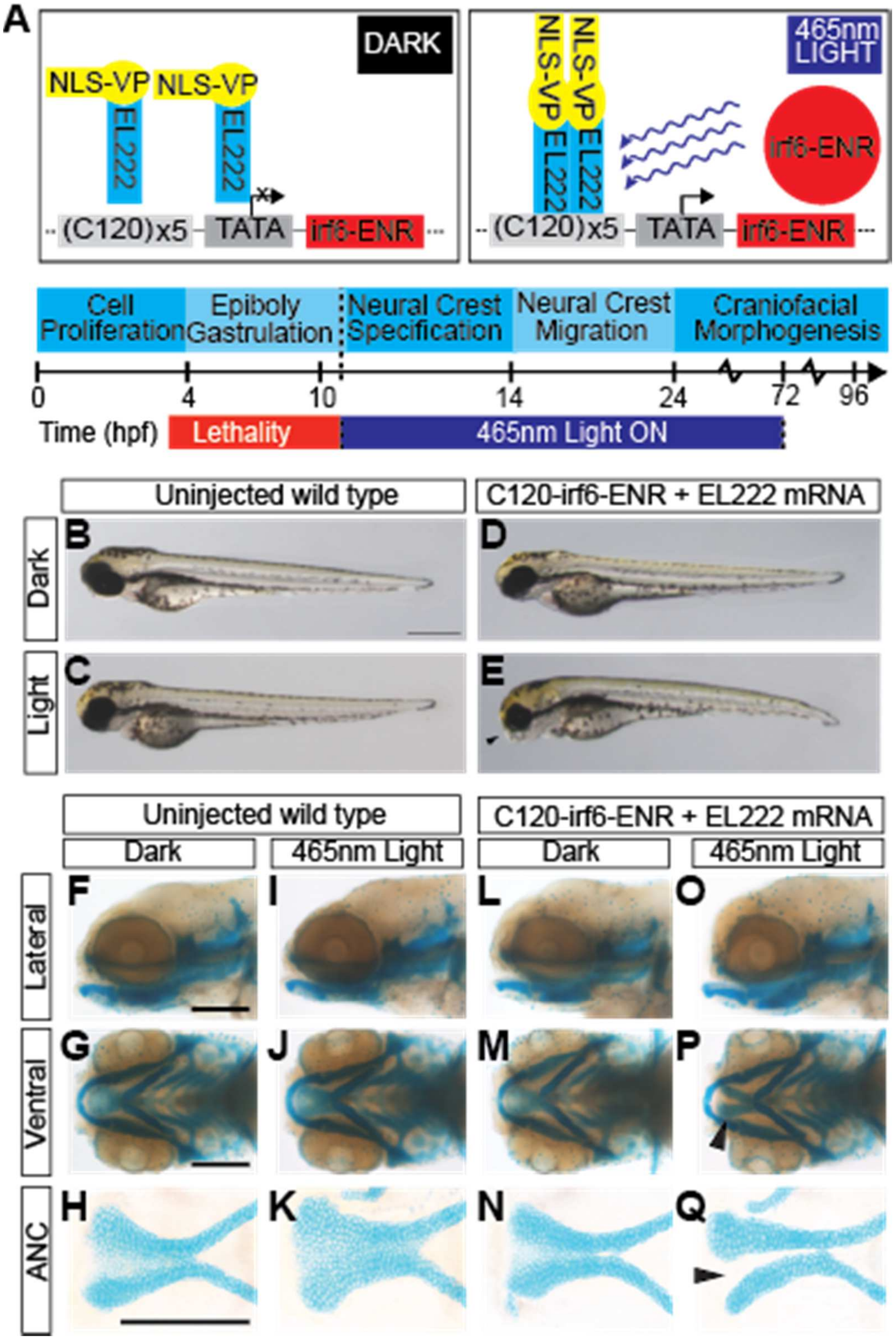


Figure 17 EL222 optogenetic disruption of *irf6* circumvents early embryonic lethality and causes a cleft palate phenotype. (A) Schematic of EL222 Optogenetic

Figure 17 (Cont'd) system. VP16-EL222 monomers are inactive under dark conditions. Upon stimulation by 465nm light, VP16-EL222 dimerizes, drives gene expression downstream of the C120 promoter, and induces the expression of a dominant negative form of irf6 (*irf6*-ENR). Embryos were exposed to blue light from 10-72hpf to circumvent embryonic lethality in mz-*irf6*$^{-8bp/-8bp}$ embryos. (B-E) Brightfield microscopy of 72hpf zebrafish embryos injected with the optogenetic system and grown in the dark (D) or exposed to blue light from 10-72hpf (E) compared to control injected embryos (B, C). Injected fish exposed to blue light exhibit retrusion of the midface (arrowhead) and curved body not observed in the other groups. (F-Q) Alcian blue staining of cartilage and microdissection of the palate of 72hpf embryos reveals a midface retrusion and cleft phenotype through the medial ethmoid plate (panel Q arrowhead) in the C120-*irf6*-ENR injected embryos grown under blue light (O-Q) which is not seen in control injected embryos (I-K) or injected embryos grown in the dark (L-N). Scale bars: 150um.

Compound homozygote of *esrp1* and *esrp2* exhibits cleft lip and ANC in zebrafish

To investigate the genetic requirement of *esrp1* and *esrp2* on zebrafish craniofacial development, CRISPR/Cas9 genome editing was utilized to generate *esrp1* and *esrp2* mutant alleles. Several alleles of *esrp1* and *esrp2* were were obtained, where alleles harboring -4bp and -14bp indels that lead to frameshift mutations and early protein truncation were selected for breeding, hereafter referred to as *esrp1*$^{-4bp/-4bp}$ and *esrp2*$^{-14bp/-14bp}$, respectively. No phenotype was observed in the *esrp1*$^{-4bp/-4bp}$ embryos, and *esrp2*$^{-14bp/-14bp}$ fish developed normally except that females were infertile, as previously published in independently derived CRISPR alleles of *esrp1/2* [76]. However, compound homozygote *esrp1*$^{-4bp/-4bp}$; *esrp2*$^{-14bp/-14bp}$ zebrafish exhibit several phenotypes, consistent with previously published mutants [76]. The *espr1*$^{-4bp/-4bp}$; *esrp2*$^{-14bp/-14bp}$ embryos also failed to inflate the swim bladder, and the pectoral fins were formed but diminutive and the margins of the fin appeared dysplastic with irregular morphology. Further, Alcian blue staining of double knockouts revealed cleft in the ANC, whereas the ventral cartilages, including the Meckel's cartilage, were formed and appeared wild-type (**Figure 18**A).

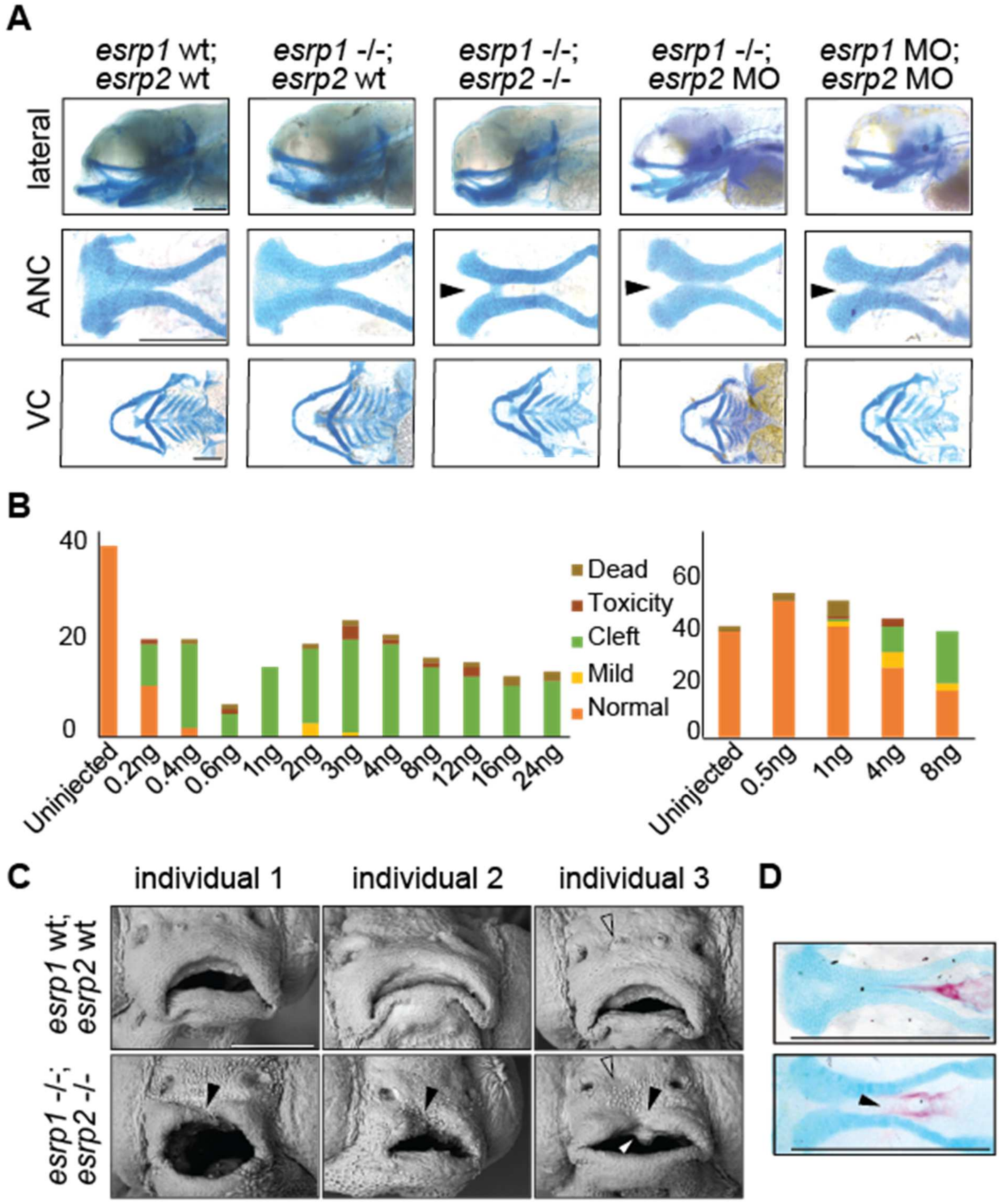


Figure 18 ***esrp1/2*** **double mutants display a cleft lip and palate.**

Figure 18 (Cont'd) (A) Alcian blue staining of 4dpf zebrafish. Representative images of WT, *esrp1* CRISPR mutant (*esrp1* -/-), and *esrp1/2* double CRISPR mutant (*esrp1* -/-; *esrp2* -/-), as well as *esrp1* CRISPR mutant treated with *esrp2* morpholino and WT treated with *esrp1* and *esrp2* morpholino (*esrp1* MO, *esrp2* MO). Flat-mount images of the anterior neurocranium (ANC) show a cleft (arrow) between the median element and lateral element of the ANC when both *esrp1* and *esrp2* function were disrupted. Lateral images and flat-mount images of the ventral cartilage (VC) show only subtle changes in morphology between WT and *esrp1/2* -/- zebrafish. (B) Morphant phenotypes observed over a range of *esrp1* and *esrp2* morpholino doses. Single *esrp2* MO injections in the *esrp1*-/- background achieves nearly 100% phenotype penetrance, even at very low MO doses. (C) SEM of 5dpf zebrafish showing discontinuous upper lip (closed arrow) in the *esrp1/2* double CRISPR mutant as well as absent preoptic cranial neuromasts (open arrow) and abnormal keratinocyte morphology. White arrow depicts aberrant cell mass (D) Representative images of alizarin red/Alcian blue staining of 9dpf *esrp1/2* double CRISPR mutant zebrafish and WT clutch-mate controls. *Esrp1/2* ablation causes abnormal morphology of the mineralizing parasphenoid bone, where the bone appears wider and with a cleft (arrow).

Inter-cross of *esrp1*$^{-4bp/-4bp}$; *esrp2*$^{wt/-14bp}$ produces predicted Mendelian ratio of 25% *esrp1*$^{-4bp/-4bp}$; *esrp2*$^{-14bp/-14bp}$ embryos for downstream phenotypic analysis, where 75% of the embryos appeared wild-type. In order to increase the percentage of embryos that can be utilized for analysis to 100%, we asked whether morpholino disruption of *esrp2* in the *esrp1*$^{-4bp/-4bp}$ background would yield consistent a cleft ANC phenotype that phenocopied *esrp1*$^{-4bp/-4bp}$; *esrp2*$^{-14bp/-14bp}$ mutant. We successfully phenocopied the cleft ANC phenotype by co-injecting *esrp1* and *esrp2* MOs into WT embryos. However, the morpholino concentrations needed were relatively high, requiring 2-8 ng of each MO to be injected for ~25-50% of embryos to develop a cleft (**Figure 18**A, B). Importantly, when *esrp1*$^{-4bp/-4bp}$ embryos were injected with *esrp2* MO, the cleft ANC phenotype was consistent and observed in nearly 100% of embryos, even when the MO concentration was reduced as low as 0.4 ng (**Figure 18**A, B). One explanation for this observation is that transcriptional compensation between *esrp1* and *esrp2* occurs when each gene is targeted, thereby requiring higher doses of each MO to ablate *esrp* activity sufficiently [90]. But when one of the *esrp* genes is already disrupted in the homozygous *esrp1*$^{-4bp/-4bp}$ mutant, the threshold for full *esrp* loss of function is lower and required a much smaller dose of MO to generate the cleft ANC phenotype.

Using scanning electron microscopy, we observed that the cleft of the upper margin of the stomodeum invaginates and extended into the cleft of the ANC. Additionally, the keratinocyte morphology of the surface epithelium appeared irregular and round with epithelial blebs in the *esrp1*$^{-4bp/-4bp}$; *esrp2*$^{-14bp/-14bp}$ embryo. In contrast, the WT surface epithelium keratinocytes appeared octagonal or hexagonal without epithelial blebs (**Figure 18**C). Alizarin red staining of the larvae at 9dpf also revealed a lack of mineralization at the midline of the parasphenoid bone (**Figure 18**D), consistent with a cleft of ANC that persisted to the ossification stage and subsequent larval development.

Zebrafish ANC morphogenesis is dependent on epithelial interactions with infiltrating cranial neural crest cells

Formation of the zebrafish ANC involves migration of anterior most CNCCs to populate the median portion (frontonasal derived) while more posterior CNCCs migrate from each side (maxillary derived). These 3 discrete embryonic elements fuse to form the ANC. Concurrent with these cellular movements, the CNCCs undergo differentiation to chondrocytes [29, 81]. We found that the ablation of *irf6* (a key periderm/epithelial gene) and *esrp1/2* (epithelial-restricted genes) both resulted in a cleft in the ANC, where chondrocytes were absent along the fusion plane between the frontonasal-derived median element and one side of the maxillary-derived lateral element (**Figure 17**C and **Figure 18**A).

To investigate the absence of these ANC chondrocytes, we performed lineage tracing of CNCCs in *esrp1/2* ablated embryos. Previously, we and others identified that the anterior most CCNC populations at 20 somites migrate to and populate the median (frontonasal) element of the ANC [29, 81]. Accordingly, we labeled the CNCCs at 20-somite stage through photo-conversion of kaede under the lineage specificity of the *sox10* promoter. CNCCs of WT or *esrp1/2* CRISPR mutants or *esrp1/2* morphants were photo-converted at 12-15hpf (**Figure 19**A, B). Embryos were imaged at 4dpf to determine the population of the ANC contributed by photo-converted cells. We found that *esrp1/2* ablation did not affect the ability of CNCCs to migrate into the ANC and reached posterior positions without clustering anteriorly (**Figure 19**A, B). These results suggest that the cleft of the ANC in the *esrp1/2* mutants is not due to a total absence of progenitor cells or defect in CNCC migration into the ANC. Nevertheless, Alcian blue staining confirmed that chondrocytes were absent from a cleft in the ANC in the *esrp1/2* mutants (**Figure 18**A).

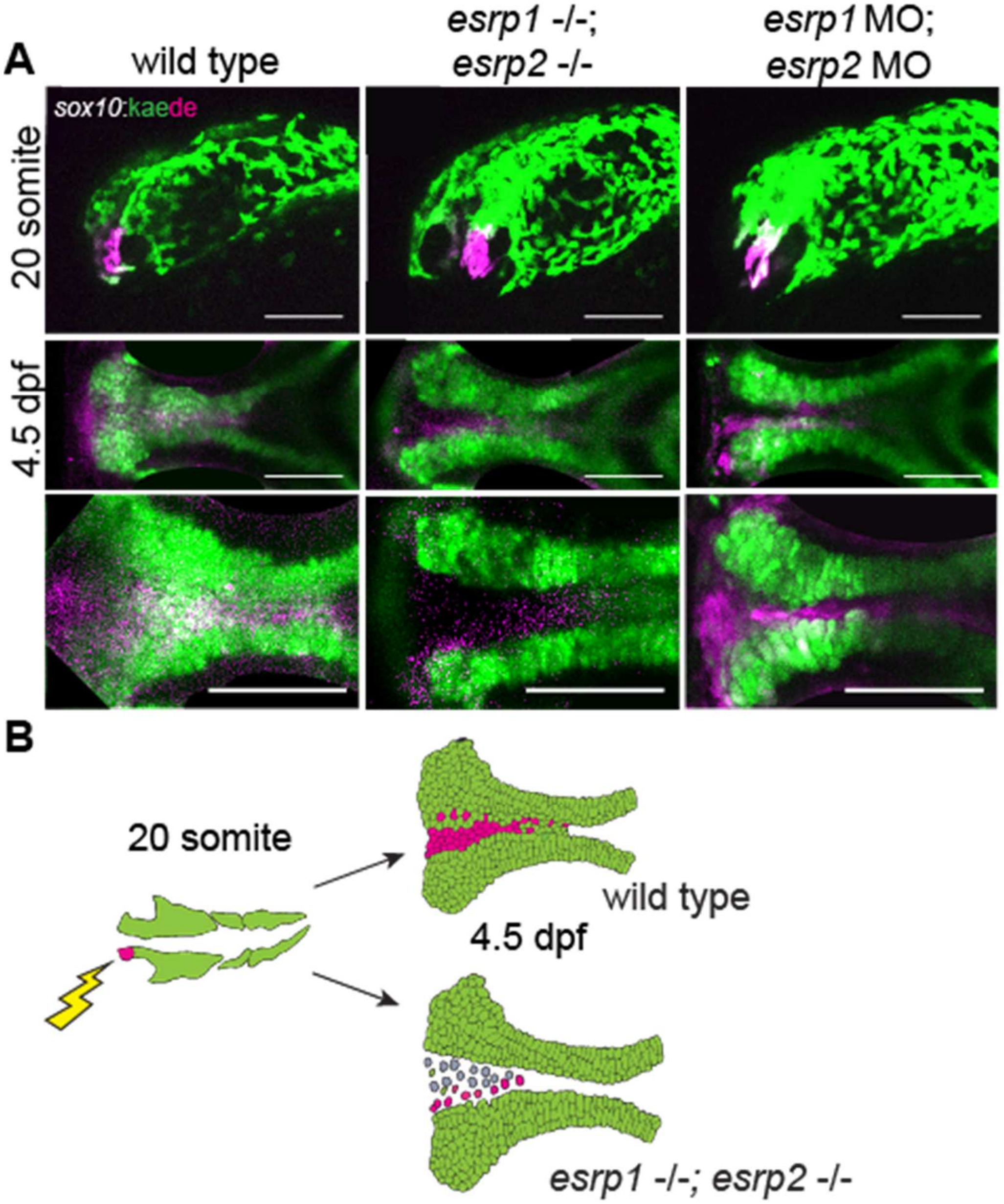


Figure 19 ***esrp1/2*-null cranial neural crest cells migrate to the ANC but do not differentiate to chondrocytes.**

Figure 19 (Cont'd) (A) Lineage tracing of WT or *esrp1/2* morphant zebrafish embryos using the Tg(*sox10*:kaede) line, native kaede fluorescence is shown in green, and photo-converted kaede is shown in magenta. Sagittal and horizontal views of zebrafish embryos at 19hpf and 4.5dpf, respectively. The anterior-most neural crest frontonasal prominence (FNP) progenitors were photoconverted at 19hpf. At 4.5dpf, the WT signal tracks to the medial portion of the ANC. Both the *esrp1*/esrp2 double CRISPR mutants and *esrp1/2* morphants exhibit a cleft in the ANC with absence of a portion of *sox10*-positive cells in the medial portion of the ANC, but the labeled CNCC representing FNP progenitors did reach and populate the entire length of the ANC. (B) Illustrative summary of lineage tracing results showing that photo-converted anterior most CNCCs contributing to FNP do migrate into the ANC in *esrp1/2* mutant embryos but a cleft forms at the juxtaposition of the FNP-derived median element and the maxillary-derived lateral element.

To investigate the cellular composition of the ANC cleft, we performed RNAscope ISH staining of WT and *esrp1/ 2* mutants at 4dpf. Sections through ANC clefts showed a dense population of cells in the location of the cleft (**Figure 20**A, B). In fact, this mass of cells can be localized in the SEM image of the *esrp1/2* mutant larvae (Fig. 5C). These cells are *col2a1* negative, consistent with absent Alcian blue staining. Instead, this aberrant cell population expresses *irf6* while *krt4* staining is restricted to the periphery, consistent with the epithelial lining of the oral cavity (**Figure 20**, **Figure 21**). Coronal and sagittal sectioning through the medial ANC of WT and *esrp1/2* mutant embryos confirmed the ectopic expression of *irf6* and revealed *sox10* expression in these aberrant, Alcian blue negative cells (**Figure 21**A, B). Like *krt4*, the expression of *krt5* outlines the oral cavity (**Figure 21**B). The expression of *sox10* suggests that at least a portion of these cells were CNC-derived, whereas *krt4* expression indicates an epithelial lineage. The presence of *irf6* expression may be indicative of epithelial/periderm cells, or indicative of expression by CNCCs as has been previously reported [29, 47]. Based on these results, we hypothesize that epithelial (and/or periderm) cells associated with frontonasal and maxillary prominence derivatives are defective in the *esrp1/esrp2* null mutants, and either disrupt or fail to promote fusion of the median and lateral elements of the ANC, causing a cleft to form (**Figure 22**). In this way, this is the first direct evidence of cleft pathogenesis in the zebrafish due to epithelial defect and suggests a model to consider how cleft pathogenesis involving the primary palate is conserved across vertebrates [50, 70, 71, 74].

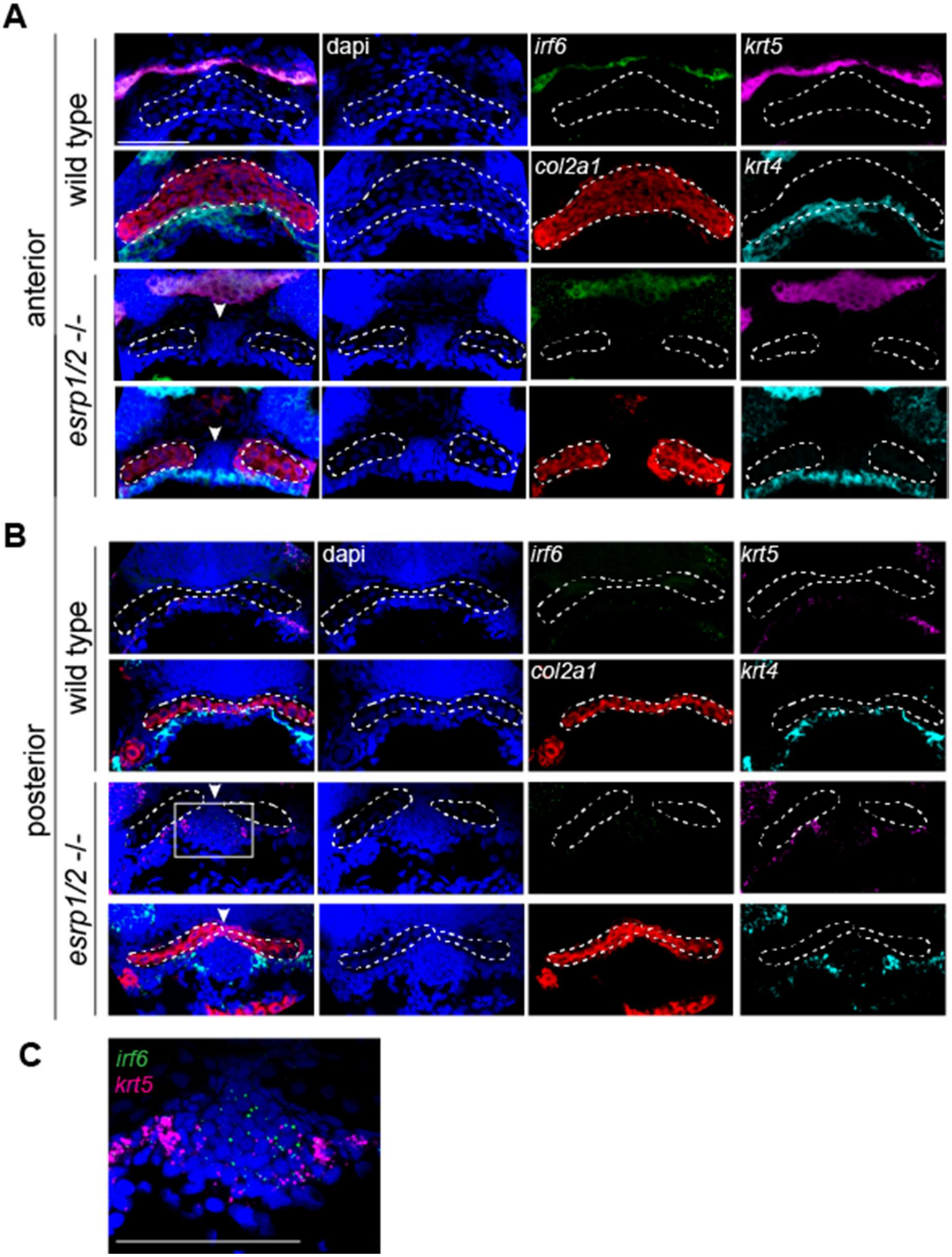


Figure 20 ANC of esrp1/2 double mutants is populated by undifferentiated cells.

Figure 20 (Cont'd) Representative z-stacks of RNAscope ISH of coronal sections of esrp1/2 double CRISPR mutants and WT clutch-mate controls at 4dpf. (A) Sections through ANC anterior to the eyes. col2a1 (red) staining depicts normal morphology of the ANC cartilage elements in WT while a cleft is apparent in the esrp1/2-/- zebrafish, with dapi (blue) stained cells between adjacent trabeculae (white arrow). These col2a1 negative cells do not express epithelial markers krt4 (cyan) or krt5 (magenta), except around the periphery. (B) Sections posterior to those in (A) show col2a1 negative cells continuing inferior to the trabeculae in the esrp1/2 mutant zebrafish and cells have low expression of irf6 (white box). (C) Zoomed image of col2a1 negative cells from (B) showing irf6 expression. Scale bars: 50 μm.

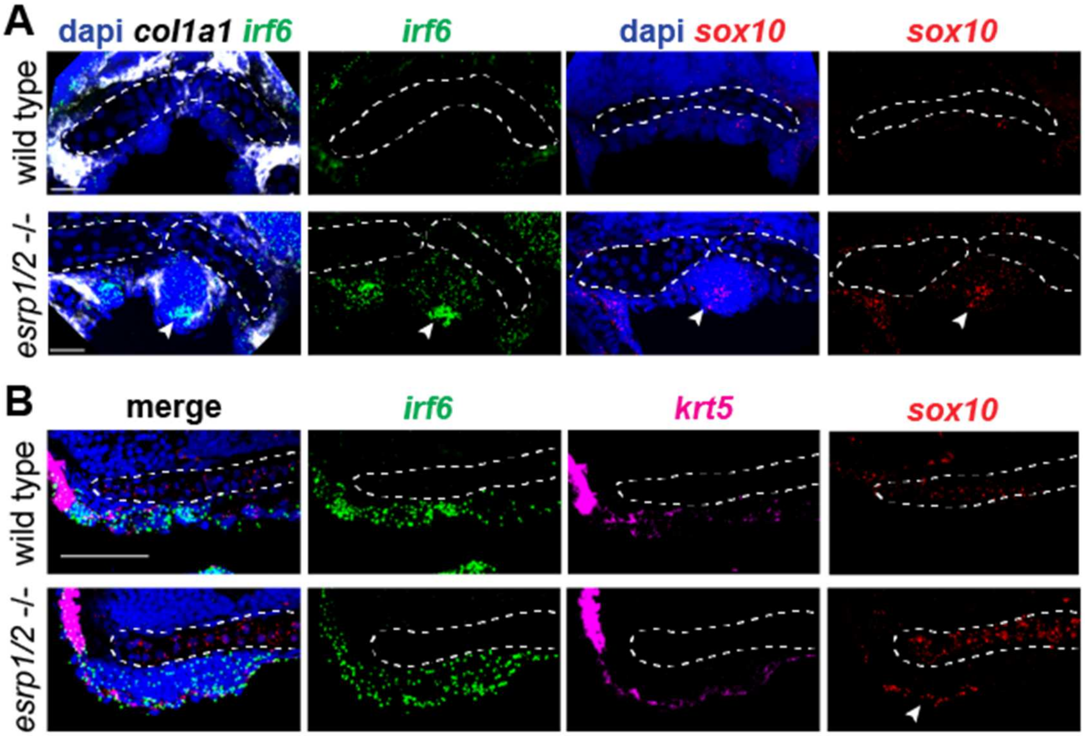


Figure 21 Aberrant ANC cells of *esrp1/2* double mutants express CNCC and epithelial cell markers.

Figure 21 (Cont'd) Representative z-stacks of RNAscope ISH of coronal sections of *esrp1/2* double CRISPR mutants and WT clutch-mate controls at 4dpf. (A) Sections through ANC anterior to the eyes. (B) Medial sagittal sections through ANC (anterior to left). ANC cartilage elements are outlined in white. *col1a1* (white) staining depicts perichondrium surrounding the aberrant mass of cells in the *esrp1/2* mutant zebrafish consistent with chondrogenic condensation (arrow). *Irf6* (green) and *sox10* (red) expression is apparent in these cells (arrowheads). Dapi (blue). Scale bars: 20 µm.

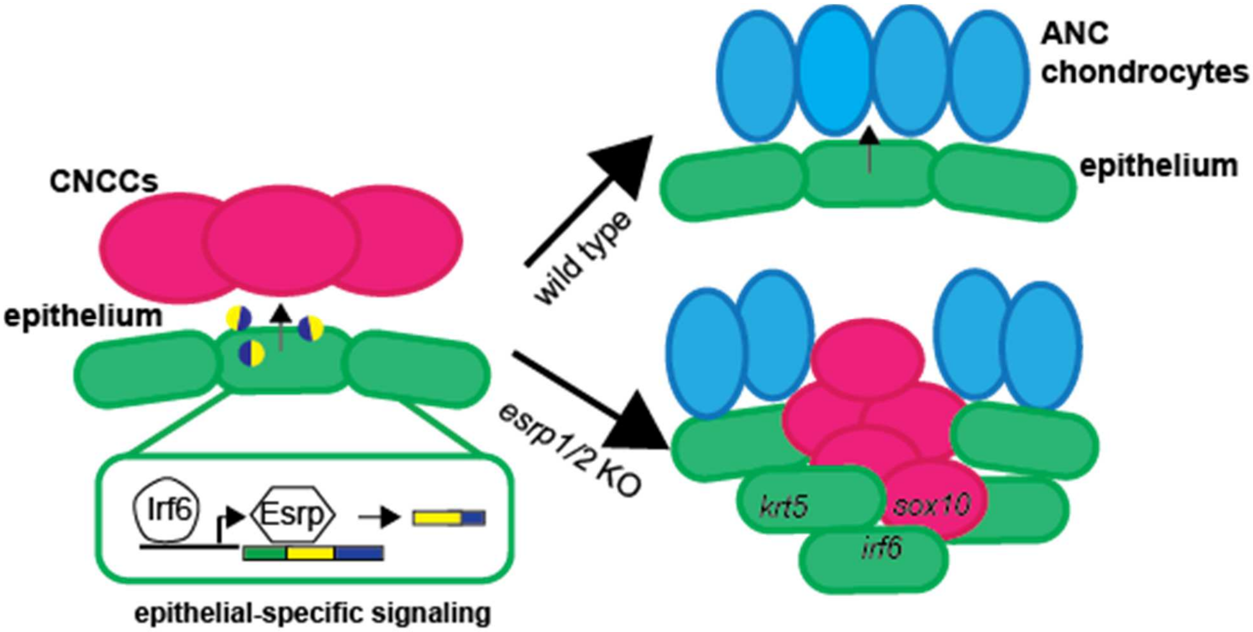


Figure 22 Illustrative summary of results. Ablation of the epithelial-restricted splicing factors *esrp1* and *2* led to the dysregulation of CNCC integration and differentiation in the medial ANC, causing a cleft between lateral ANC elements. These results suggested that epithelial-specific splice variants of yet to be determined factors are required for directing the juxtaposed mesenchymal-derived cells and promoting normal morphogenesis.

Genetic interaction of *Irf6R84C* with *Esrp1* and *Esrp2*

To test the hypothesis that *Irf6* and *Esrp1/2* genes function in the same developmental pathway, we carried out genetic epistasis analysis and generated *Irf6*; *Esrp1*; *Esrp2* compound mutants. We hypothesized that if *Irf6* and *Esrp1/2* genetically interact, then *Irf6* and *Esrp1* heterozygosity on an *Esrp2* null background may result in a cleft phenotype, when *Irf6* and *Esrp1* heterozygotes do not normally form a cleft. As expected, we observed that *Irf6*$^{R84C/+}$; *Esrp1*$^{+/-}$; *Esrp2*$^{+/-}$ mice developed and reproduced normally. To generate *Irf6*$^{R84C/+}$; *Esrp1*$^{+/-}$; *Esrp2*$^{-/-}$ embryos, we intercrossed the triple heterozygous mice. We collected a total of 79 embryos from 9 litters from E12.5-E18.5 and tabulated the resulting genotypes (Table 1). Based on Mendelian genetics, we expected approximately 5 *Irf6*$^{\mathrm{R84C/R84C}}$; *Esrp*1$^{-/-}$ double homozygous mice. However, these breedings did not produce any *Irf6*$^{R84C/R84C}$; *Esrp1*$^{-/-}$ embryos (Table 1). This result suggested that compound ablation of *Irf6* and *Esrp1* is more deleterious to development than either genotype alone, and supports a genetic interaction between *Irf6* and *Esrp1*, which may be essential early in development.

Table 1. Irf6, Esrp1 and Esrp2 genotypes interact to produce non-Mendelian embryo ratios

Irf6	*Esrp1*	*Esrp2*	probability	Expected	Observed
Het	WT	WT	1/32	2.5	5
Het	WT	Het	1/16	4.9	4
Het	WT	KO	1/32	2.5	3
Het	Het	WT	1/16	4.9	10
Het	Het	Het	1/8	9.9	8
Het	Het	KO	1/16	4.9	2
Het	KO	WT	1/32	2.5	2
Het	KO	Het	1/16	4.9	6
Het	KO	KO	1/32	2.5	3
R84C	WT	WT	1/64	1.2	1
R84C	WT	Het	1/32	2.5	2
R84C	WT	KO	1/64	1.2	0
R84C	Het	WT	1/32	2.5	5
R84C	Het	Het	1/16	4.9	4
R84C	Het	KO	1/32	2.5	5
R84C	KO	WT	1/64	1.2	0
R84C	KO	Het	1/32	2.5	0
R84C	KO	KO	1/64	1.2	0

(Table 1 Cont'd) *Irf6R84C/+*; *Esrp1*+/-; *Esrp2*+/- triple heterozygous mice were in-crossed and embryos were collected between E12.5 and E21. Table 1 is a subset of expected number of embryos based on Mendelian genetics versus the observed number of viable embryos. The *Irf6R84C/R84C*; *Esrp1*-/- genotype appears to be lethal prior to E12.5 as approximately 5 embryos were expected but zero embryos were observed. A total of 79 embryos were collected from 9 different litters.

To test for phenotypic affects in the resulting *Irf6*; *Esrp1*; *Esrp2* compound mutants we imaged embryos at E18.5. We did not observe any cleft lip or palate in any genotype collected except for the expected clefts when null for *Irf6* or *Esrp1* (Fig. 9). *Irf6*$^{R84C/+}$; *Esrp1*$^{+/-}$; *Esrp2*$^{-/-}$ embryos that we predicted to be susceptible to cleft lip and/or palate were grossly normal. We noticed some differences in the shape of the palate between heterozygous genotypes and measured the length (from philtrum to first rugae) relative to the width (space between lips). We found that *Irf6*$^{R84C/+}$ heterozygotes exhibited a shorter palate than WT (**Figure 23**A, B). A shorter snout has been previously reported in *Irf6* KO mice [50]. The length/width ratio of *Esrp1/2* double heterozygotes was similar to WT and in the *Irf6*; *Esrp1*; *Esrp2* triple heterozygote, the shorter palate phenotype of the *Irf6* heterozygote was reversed (**Figure 23**A, B). Taken together these breeding and morphologic analyses suggest an overlapping role of *Irf6* and *Esrp* in regulating midface morphogenesis.

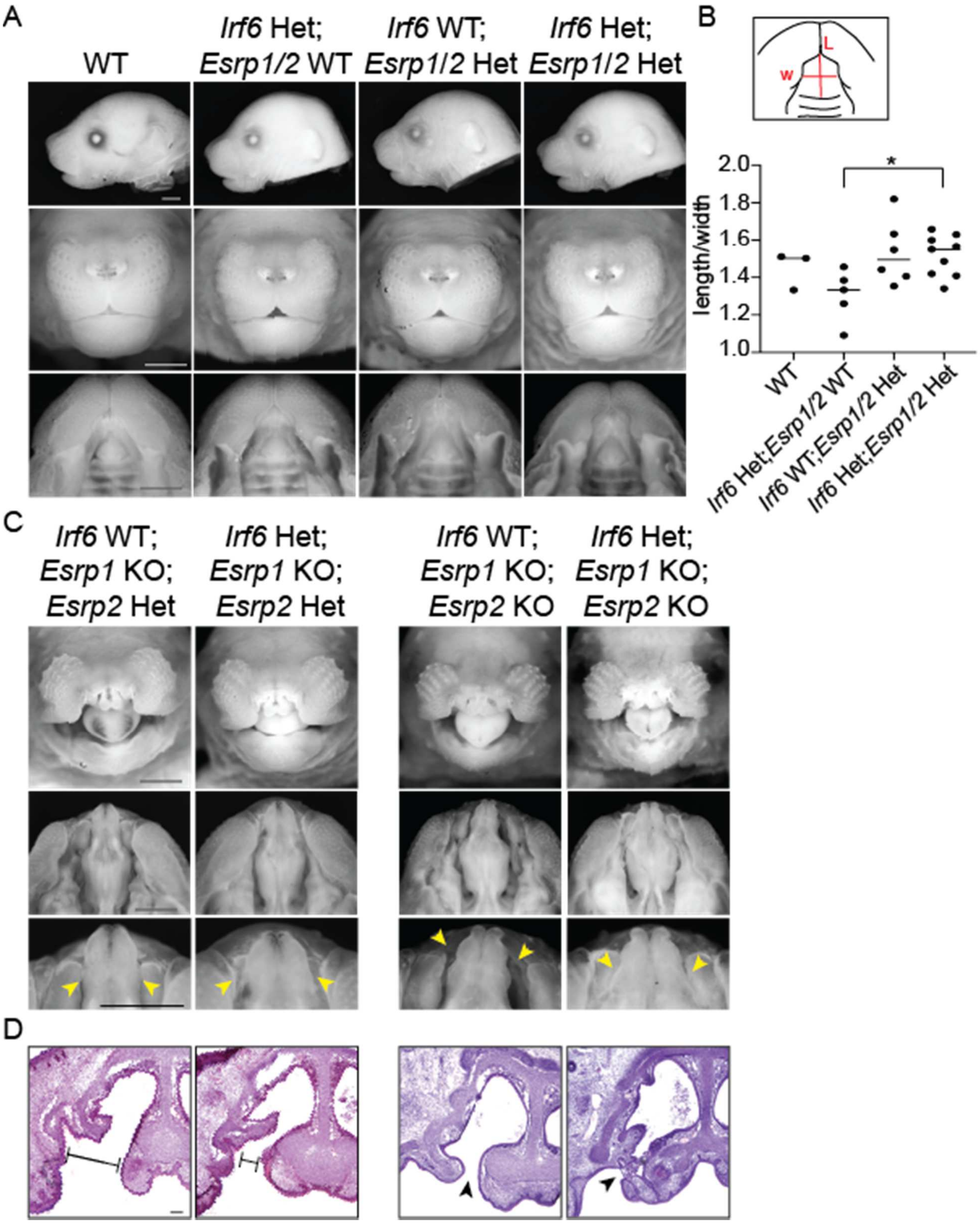


Figure 23 Irf6 and Esrp1/2 interact to modify palate phenotypes.

Figure 23 (Cont'd) Mice compound heterozygous for *Irf6*^{R84C}, *Esrp1* and *Esrp2* were generated by breeding *Irf6*$^{R84C/+}$ with *Esrp1*$^{+/-}$; *Esrp2*$^{-/-}$ mice. The triple heterozygotes were then inter-crossed and embryos were collected at E18.5. (A) Representative lateral, frontal and oral images of embryos, comparing WT (*Irf6*$^{+/+}$; *Esrp1*$^{+/+}$; *Esrp2*$^{+/+)}$, *Irf6*^{R84C} heterozygote (*Irf6*$^{R84C/+}$; *Esrp1*$^{+/+}$; *Esrp2*$^{+/+}$), *Esrp1/2* double heterozygote (*Irf6*$^{+/+}$; *Esrp1*$^{+/-}$; *Esrp2*$^{+/-}$) and triple heterozygote (*Irf6*$^{R84C/+}$; *Esrp1*$^{+/-}$; *Esrp2*$^{+/-}$). (B) Measurements of palate length relative to width. *Irf6*$^{R84C/+}$ embryos tend to have a shorter palate as compared to WT, however this genotype on an *Esrp1*$^{+/-}$;*Esrp2*$^{+/-}$ background significantly increases palate length relative to *Irf6*$^{R84C/+}$*;* *Esrp1*$^{+/+}$; *Esrp2*$^{+/+}$ (one-way ANOVA, $p<0.05$). (C) Representative frontal and oral images of embryos, comparing *Irf6*$^{+/+}$; *Esrp1*$^{-/-}$; *Esrp2*$^{+/-}$ to *Irf6*$^{R84C/+}$; *Esrp1*$^{-/-}$; *Esrp2*$^{+/-}$ and *Irf6*$^{+/+}$; *Esrp1*$^{-/-}$; *Esrp2*$^{-/-}$ to *Irf6*$^{R84C/+}$; *Esrp1*$^{-/-}$; *Esrp2*$^{-/-}$. Scale bars: 50 μm. (D) Hematoxylin and eosin staining of coronal sections through the vomeronasal cavity and primary palate of the same embryos. *Irf6*^{R84C} heterozygosity modifies the *Esrp1* KO and *Esrp1/2* dKO cleft lip and palate such that the cleft space between adjacent elements is narrower and, in some cases, we noticed epithelial adhesions that limited the cleft. Scale bar: 100 μm.

As previously reported, *Esrp1* null and *Esrp1/2* double null embryos displayed bilateral cleft lip and palate (**Figure 23**C). Interestingly, we noted a modification of this cleft phenotype when *Esrp1* and *Esrp1/2* null embryos were also heterozygous for *Irf6*. *Irf6*$^{R84C/+}$; *Esrp1*$^{-/-}$; *Esrp2*$^{+/-}$ and *Irf6*$^{R84C/+}$; *Esrp1*$^{-/-}$; *Esrp2*$^{-/-}$ embryos had less space or less wide clefts between the lateral lips and maxilla and the midline nasal capsule (**Figure 23**C). This decreased space between tissue (or cleft severity) but persistence of a cleft was confirmed in histological sections (**Figure 23**D). Further, histological sections showed presumed epithelial adhesions between lateral and medial portions of the nasal cavity in *Irf6* heterozygotes, whereas this space is open in *Esrp1* null and *Esrp1/2* double null embryos (**Figure 23**D).

Discussion

Orofacial clefts are a common birth defect, and GWASs have identified some critical genes associated with syndromic and non-syndromic cleft. Here we described mouse and zebrafish models using genes with known genetic variants in human cleft patients, *IRF6* and *ESRP1/2*. We present evidence to support that *Irf6* and *Esrp* function in the same regulatory pathway. We observed that mz-*irf6*$^{-8bp/-8bp}$ zebrafish embryos have significantly decreased expression of *esrp1,* and this is rescued upon introduction of *irf6* mRNA. This finding is consistent with *esrp1* being a transcriptional target of *irf6,* and putative *irf6* response elements [91] can be found surrounding the *esrp1* transcriptional start site. Additionally, RNA-seq identified known *irf6* targets including *grhl3* and *tfap2a*. Direct molecular experiments are needed however to test transcriptional regulation of *esrp1* by *irf6*. We found that *Irf6* and *Esrp1/2* are consistently co-expressed in the embryonic frontonasal ectoderm and oral epithelium associated with the palate, and epithelium of the mouth opening, in both mouse and zebrafish.

In zebrafish, *irf6* null embryos ruptured during gastrulation whereas *esrp1/2* null embryos survive to larval stage. However, post-gastrulation ablation of *irf6* resulted in a similar cleft morphology of the ANC as the *esrp1/2* null. Further analysis of the *esrp1/2* null showed that the cleft of the ANC correlated with a cleft in the upper margin of the mouth opening, reminiscent of a human cleft lip. Further, using a neural crest-specific photo-convertible reporter line, we were able to show that migration of CNCC to the developing ANC occurred but chondrogenesis was impaired.

The early lethality of *irf6* null zebrafish initially precluded analysis of the *irf6* zygotic requirement in craniofacial development. Here, we utilized an optogenetics strategy to disrupt *irf6* function after gastrulation when the embryonic body axis had formed, thereby revealing the zygotic requirement for *irf6*. Future studies will use the *irf6* optogenetic model to study the roles of *irf6* during ANC and lip morphogenesis. Interestingly, periderm markers identified in the mouse lambdoidal junction were found to be dysregulated in the *irf6* mutant zebrafish model, specifically *grhl3*, *tfap2a* and *perp*. Additionally, *gata3,* which was identified as a mesenchymal marker at the fusion zone of mice [92], is dysregulated in the *irf6* null zebrafish. This work highlights the utility of complementary studies of palate morphogenesis in zebrafish and mouse models. The zebrafish model affords the transgenic tractability and visualization of CNCC migration, enabling us to determine the cellular mechanism responsible for the cleft ANC. The mouse mutants provide the mammalian anatomic contexts to examine cleft malformation.

While the *esrp1*$^{-4bp/-4bp}$*; esrp2*$^{-14bp/-14bp}$ zebrafish exhibited consistent cleft lip and cleft ANC phenotype, the infertility of the *esrp2*$^{-14bp/-14bp}$ fish preclude large-scale experiments to analyze downstream mechanisms of the development of cleft palate. We generated a robust *esrp1*$^{-4bp/-4bp}$; *esrp2* morphant assay that can be applied to toward chemical screening experiments and functional testing of human *ESRP1/2* gene variants.

In humans, CPO is less common than CL/P [20, 59, 77]. Although humans and mice share approximately 99% of their genes and the early craniofacial development of the mouse embryo closely mirrors the human [22], there is a striking difference in the manifestation of orofacial cleft defects [59]. Most often when a human CL/P-associated gene has been disrupted in mice, a cleft of the palate forms but the lip appears normal. Our current understanding in humans is that CL/P and CPO are different genetic disorders [21, 59, 61]. These discrepancies between humans and mouse models hampers understanding of the etiopathogenesis of human CL/P. Here we characterize the *Esrp*1/2 null mouse, exhibiting bilateral CL/P, as an important model for studying OFC etiopathogenesis. Additionally, as we place *ESRP1* in the *IRF6* gene regulatory pathway, we hope to better understand how alternative isoforms regulated by *ESRP1* may in turn be important for palate development.

Whereas zebrafish have historically been an excellent model organism for forward genetic screens, CRISPR/Cas9 gene editing technology has permitted relatively efficient reverse genetic engineering of zebrafish [93]. This utility of the zebrafish embryo for studying developmental processes and modeling human cleft-associated genes necessitates further study into their craniofacial morphogenesis. Transplant and lineage tracing experiments have illuminated the neural crest origin of the zebrafish ANC, and how the frontonasal and paired maxillary cartilage elements converge into a continuous cartilage structure [29, 35, 81]. We show that *IRF6* and *ESRP1* are conserved in their requirement for ANC morphogenesis, where disruption result in orofacial cleft in human, mouse, and zebrafish. These findings provide evidence of conserved molecular and morphological processes occurring in the merging and fusion of the mouse and zebrafish midface.

We suspect that non-epithelial expression of *Irf6* contributes to normal craniofacial morphogenesis and may explain some differences in the *Irf6* and *Esrp* mutant phenotypes. Future research utilizing tissue-specific knockout of *Irf6* will address this hypothesis. We also suspect that the *Irf6* phenotype is more severe because *Irf6* acts upstream of *Esrp1*, along with additional targets, and ongoing experiments on the transcriptional activity of *Irf6* will be important. Recently an in-depth analysis of a lineage-specific *Esrp1* knockout mouse was completed and found that *Esrp1* regulates proliferation of the mesenchyme of the lateral nasal prominences, along with being required for fusion of the medial and lateral nasal prominences [72]. Ongoing work to identify *Esrp1/2* molecular targets and mechanistic studies of these targets will provide new insight into palate morphogenesis.

These studies highlight the utility of complementary mouse and zebrafish models to elucidate mechanisms of orofacial cleft development. Additionally, this work has expanded the scope of Irf6 gene regulation in craniofacial development.

Chapter IV: Functional analysis of *ESRP1/2* gene variants and identification of *CTNND1* as Esrp-regulated isoforms in orofacial cleft

Introductory Statement

This is a current manuscript in preparation for submission. In this article, we employed two independent models of *Esrp1/2* LOF to functionally test *ESRP1* and *ESRP2* gene variants found in cohorts of patients with CL/P or autosomal recessive deafness. This article is a follow up to our *Development* article in Chapter II, where we set to investigate the protein function of *esrp1/2* through the testing of gene variants and attempt to establish causal relationships for putative *Esrp1/2* target genes. I performed all the experimental design and execution for the experiments listed with technical assistance from Casey Tsimbal, and Jason Mitchell.

Functional analysis of ***ESRP1/2*** gene variants and identification of ***CTNND1*** as Esrp-regulated isoforms in orofacial cleft

Claudio Macias Trevino[1,2], Jason Mitchell[1,2], Casey Tsimbal[1,2], Shannon Carroll[1,2], Marie Froemmrich[1,2], Kenta Kawasaki[1,2], Russ P. Carstens[3], Elizabeth Leslie[4], and Eric C. Liao[1,2,*]

[1] Division of Plastic and Reconstructive Surgery, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

[2] Shriners Hospital for Children, Boston, MA, USA

[3] University of Pennsylvania, Philadelphia, PA, USA

[4] Emory School of Medicine, Atlanta, GA, USA

* Corresponding author

Eric C. Liao, MD, PhD

Professor of Surgery, Harvard Medical School

Vice Chair of Surgery, Massachusetts General Hospital

Chief of Staff, Shriners Children’s Boston

Director, Cleft and Craniofacial Center

Simches Research Building, CPZN4

185 Cambridge Street

Boston, MA 02114

Introduction

The study of orofacial cleft (OFC) has been foundational to genetic analysis of congenital anomalies, as facial structural malformations are amenable to detailed phenotypic classification to apply whole genomic studies to discover associated genes[42, 43, 94-100]. Since OFCs are among the most common congenital anomalies, the study of OFC genetics is not only scientifically important but also addresses a major clinical problem that has a lifelong impact on patients and families. As whole-genome sequencing (WGS) strategies and technologies advance, many genes and gene variants are associated with OFC phenotypes and subtypes [101]. These approaches have identified key regulators, such as the transcription factors *IRF6*, *GRHL3* and *TFAP2A,* associated with syndromic OFCs [40, 45]. Non-syndromic OFCs account for the majority of cleft cases, and the role of epithelial cell adhesion is increasingly recognized to play a key role in cleft pathogenesis, where *CTNND1* and *ESRP2* have been recently implicated in human cohorts and studies in murine and Xenopus models [64, 102]

As the field continues to identify genes associated with syndromic and non-syndromic OFC, there is an ever-increasing need to impute the pathogenicity of associated gene variants. As most cases of non-syndromic OFC are *de novo* cases, imputing the functional significance of gene variants detected from WGS is challenging. The American College of Medical Genetics and Genomics, in conjunction with the Association for Molecular Pathology (ACMG-AWP), recommends a set of criteria to identify gene variants with a high likelihood of pathogenicity. Among these criteria, in vitro and in vivo functional studies provide the most substantial supporting evidence for asserting pathogenic potential in gene variants for genes not previously established as causal for a particular disease [103, 104] . Multiple *in silico* predictive algorithms such as SIFT, PolyPhen, and PROVEAN offer functional imputations for gene variants that utilize amino acid sequence information, including sequence conservation, biophysical properties, or homolog alignment [105-108]. However, when given the same gene variants, different predictive tools can often provide null values or contradicting results. Thus, no primary predictive method can convert genetic data into clinically actionable information and determine the pathogenicity of gene variants in a given patient cohort, despite advances in sequencing and statistical approaches to genetics [42, 95, 97, 100, 109-111]. Predicting risks from genomic data is even more challenging with the proliferation of polygenic risk scores, adding more confusion than clinically actionable information for prenatal diagnosis and counseling [104, 106, 112].

Despite continued advances in genomics and molecular diagnostics, there remains an unmet need for functional assays that can impute the pathogenicity of gene variants from sequencing data [113, 114]. In fact, although many *IRF6* gene variants had been previously reported from GWAS and whole exome studies to be associated with cleft lip and palate, none of the variants were functionally validated. When we applied a zebrafish *irf6* mutant model to impute pathogenicity of 32 *IRF6* gene variants through an embryonic rescue assay, we found that a number of gene variants previously imputed to be pathogenic based on *in silico* prediction were functionally benign, corroborated by more recent population sequencing data from gnomAD [51]. This and other studies underscore the essential role of functional validation of gene variants as current prediction and association analysis are often insufficient [115-120]

Since *IRF6* is one of the most commonly associated genes in syndromic and non-syndromic OFC and plays a central role in embryonic epithelial differentiation, we focused on delineating the IRF6 regulated pathway to identify additional genes that will be important in periderm biology and OFC pathogenesis[40, 45, 50, 70]. We conducted a comparative RNAseq analysis between wild-type and irf6 mutant and found that the epithelial-specific splicing regulator *esrp1* was differentially expressed[121]. We found that *esrp1* and *esrp2* are colocalized in the periderm and oral epithelium and are required for the formation of the anterior neurocranium (ANC), a teleost embryonic structure developmentally analogous to the mammalian primary palate in the manner that it is formed from the convergence of frontonasal derived midline prominence and paired maxillary projections [22, 29, 30, 32, 35, 122] In the *esrp1/2* double homozygote, cleft formed in the ANC and extended to the upper edge of the mouth opening, analogous to the cleft lip and palate phenotype observed in the *Esrp1/2* mutant mice [121].

Given previous work and emerging gene variants associating *ESRP1* and *ESRP2* to orofacial cleft, we exploited the zebrafish *esrp1/2* mutant to functionally investigate human gene variants in a phenotype rescue paradigm [64]. To increase the rigor of this *in vivo* assay, we also employed murine *Esrp1/2* double knockout PY2T cells as an independent functional assay to investigate the *ESRP1* and *ESRP2* gene variants[72, 73]. Indeed, we found that the *in vivo* and *in vitro* assays were complementary approaches that provided functional evidence that either confirmed *in silico* predictions, provided information when predictions were inconclusive, or, more importantly, demonstrated some variants to be functionally benign when predicted to be pathogenic. We also showed that *Ctnnd1* is a key gene co-localized with *Esrp1* and *Esrp2* in the mouse and zebrafish, where Esrp1 and Esrp2 regulate the lineage restriction of the epithelial isoform.

Results

Our zebrafish model of cleft ANC requires double mutants for *esrp1* and *esrp2*, and the *esrp2*$^{-/-}$ fish are infertile, precluding crosses that would yield progeny with 100% penetrance of the cleft ANC phenotype[76]. However, the cleft ANC phenotype was phenocopied by injecting anti-*esrp2* morpholinos into *esrp1*$^{-/-}$ embryos, yielding a method to generate fully penetrant *esrp1/2* loss-of-function clutches[121]. In addition, PY2T mouse breast cancer cells have been edited using CRISPR to harbor loss-of-function mutations in *Esrp1* and *Esrp2* and exhibit splicing deficiencies in potential Esrp target genes [123]. Moreover, differential splicing events between the epithelium and mesenchyme in *Esrp1* mutant mice have been characterized to identify potential Esrp1 target genes [72]. However, further work is needed to define which Esrp1 target genes are essential in CL/P pathogenesis.

We have adapted our *esrp1/2* morphant assay as a platform to test for the functional rescue of the cleft ANC phenotype with gene variants for *esrp1* and *esrp2* and impute phenotypic pathogenicity from variants of otherwise uncertain functional significance. Furthermore, we adapted the *Esrp1/2* double KO PY2T cell line as a rescue assay to obtain a molecular readout and enhance our understanding of pathogenic potential through cross-species conservation and mechanistic insight. We have tested 19 variants identified from ClinVar and the GMKF children's dataset across *ESRP1* and *ESRP2* in our zebrafish and cellular assays for ESRP protein function. We provide functional and molecular evidence for the tissue-specific regulation of *CTNND1* and *ARHGEF11* by Esrp1/2 in mice and zebrafish.

Establishing complementary *in vitro* and *in vivo* assays for *ESRP1* and *ESRP2* gene variant testing

Alcian blue staining of 4dpf zebrafish embryos with the *esrp1*$^{-/-}$; *esrp2*$^{-/-}$ genotype reveals a cleft ANC phenotype where a population of chondrocytes in the medial ANC is absent, compared to the ANC of wild-type and esrp1$^{-/-}$; esrp2$^{+/+}$ embryos. A similar phenotype is observed when 8ng of translation-blocking *esrp2* morpholinos were injected into *esrp1*$^{-/-}$ embryos at the 1-cell stage and grown to 4dpf (**Figure 24**A). To functionally test *esrp1* or *esrp2* gene variants, we devised a strategy to introduce point mutations into zebrafish *esrp1* or *esrp2* coding sequences and co-injecting 8ng of anti-esrp2 MO with either: (1) capped *esrp1* mRNA, (2) capped *esrp2* mRNA mutagenized with synonymous mutations at the MO binding site, or (3) either *esrp1* or MO-resistant *esrp2* mRNA encoding for human *ESRP1* or *ESRP2* gene variants of unknown significance. We hypothesized that benign variants that preserve protein function would robustly rescue the cleft ANC phenotype like native *esrp1* or *esrp2* mRNA. Conversely, pathogenic gene variants that result in protein loss-of-function would fail to rescue the cleft ANC phenotype (**Figure 24**B).

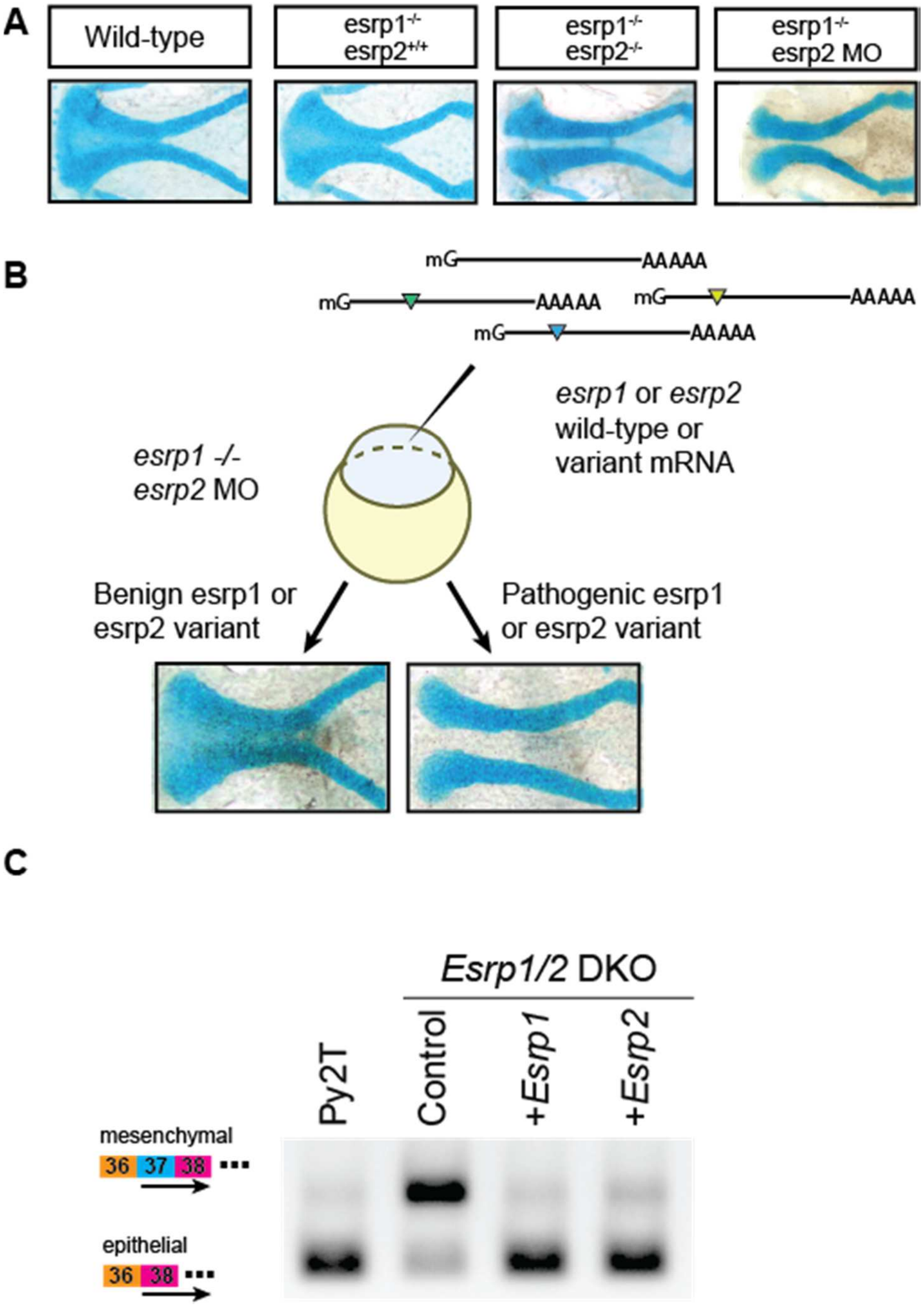


Figure 24 Complementary *in vivo* and *in vitro* models of *Esrp1/2* function for gene variant testing

Figure 24 (Cont'd) (A) Microdissected ANC of Alcian-blue stained embryos at 4dpf for wild-type, *esrp1*$^{-/-}$; *esrp2*$^{+/+}$, *esrp1*$^{-/-}$; *esrp2*$^{-/-}$, and *esrp1*$^{-/-}$; *esrp2 MO* embryos. (B) Schematic for the *esrp* morphant variant assay in zebrafish. *esrp1* or *esrp2* gene variants are co-injected with *esrp2* MO into one-cell stage *esrp1*$^{-/-}$ embryos. Variants that robustly rescue the cleft ANC phenotype are scored as benign, while variants that fail to rescue the cleft ANC phenotype are scored as pathogenic. (C) RT-PCR using primers spanning exons 36-38 of *Arhgef11* on cDNA isolated from wild-type mouse PY2T cells, *Esrp1/2* double-knockout PY2T cells, or *Esrp1/2* double-knockout PY2T cells electroporated with plasmids encoding for either *Esrp1* or *Esrp2*. Arrow markers point to the epithelial (short) isoform and mesenchymal (long) isoform skipping exon 37, or including exon 37, respectively.

To further validate our approach, we adapted splicing assays in mouse PY2T cells for *Esrp1* and *Esrp2* gene variant testing. RT-PCR performed on wild-type mouse PY2T cell cDNA using primers spanning splice junctions for two putative *Esrp* target genes demonstrates the presence of two major isoforms for *Arhgef11* transcripts. The difference between these two isoforms involves exon 37, which is included in mesenchymal cells, but skipped in epithelial cells such as PY2Ts[73, 124, 125]. PY2T cells carrying CRISPR-edited *Esrp1* and *Esrp2* loss-of-function alleles have reduced expression for the epithelial (short) isoform of *Arhgef11*, and the long isoform, previously absent, is expressed. To test whether *Esrp1* or *Esrp2* expression rescued the *Arhgef11* splicing patterns in *Esrp1/2* DKO PY2T cells, we cloned mouse mCherry-tagged constructs of *Esrp1* and *Esrp2* into a mammalian expression pcDNA3.1 vector, electroporated them individually into *Esrp1/2* DKO PY2T cells and collected RNA 24hr post-transfection for RT-PCR using primers spanning exon 37 of *Arhgef11*. Restoration of *Esrp1* or *Esrp2* in the Esrp1/2 DKO PY2T cells rescued the native splicing patterns in wild-type PY2T cells (**Figure 24**C).

Identifying ESRP1 and ESRP2 variants for testing in rescue assays

The Gabriella Miller Kids First (GMKF) Pediatric Research Program was developed to study structural birth defects and childhood cancer through whole-genome sequencing. We filtered sequencing data from the GMKF data resource for patients with orofacial clefts and found 17 gene variants for either *ESRP1* or *ESRP2* in pediatric orofacial cleft patients. We further supplemented the list with *ESRP1* or *ESRP2* variants reported in ClinVar in patients with CL/P or autosomal recessive deafness to generate a list containing 32 disease-associated gene variants in *ESRP1* or *ESRP2*. We performed cross-species alignments between human, mouse, and zebrafish amino acid sequences for *ESRP1* and *ESRP2* to determine which gene variant residues were highly conserved across species and, thus, likely critical for ESRP1 and ESRP2 function. For ESRP1, the overall amino acid sequence identity was 97% and 64.68% similar between humans and mice, or humans and zebrafish, respectively. However, when focusing on the RNA-recognition motif (RRM) domains of ESRP1, the similarity of the sequences between humans and mice and humans and zebrafish increased to 98.82% and 94.12% for RRM1, 99.08% and 79.82% for RRM2, and 95.06% and 77.78% for RRM3. Similarly, for ESRP2, the overall amino acid sequence similarity was 98.67% between humans and mice and 85.33% between humans and zebrafish. The domain-specific amino acid sequence similarities were 98.67% and 85.33% for RRM1, 98.13% and 81.31% for

RRM2, and 96.3% and 77.78% for RRM3 between humans and mice, and humans and zebrafish, respectively. Altogether, we found only 18 out of the 32 identified gene variants in residues fully conserved between human, mouse, and zebrafish *ESRP1* and *ESRP2*. Moreover, the 18 identified gene variants were part of fully conserved amino acid stretches surrounding each variant. The gene variants were evenly spread throughout both proteins and included two gene variants in the RRM1 domain of *ESRP1* and two variants each in the RRM1, RRM2, and RRM3 domains of *ESRP2* (Figure 25).

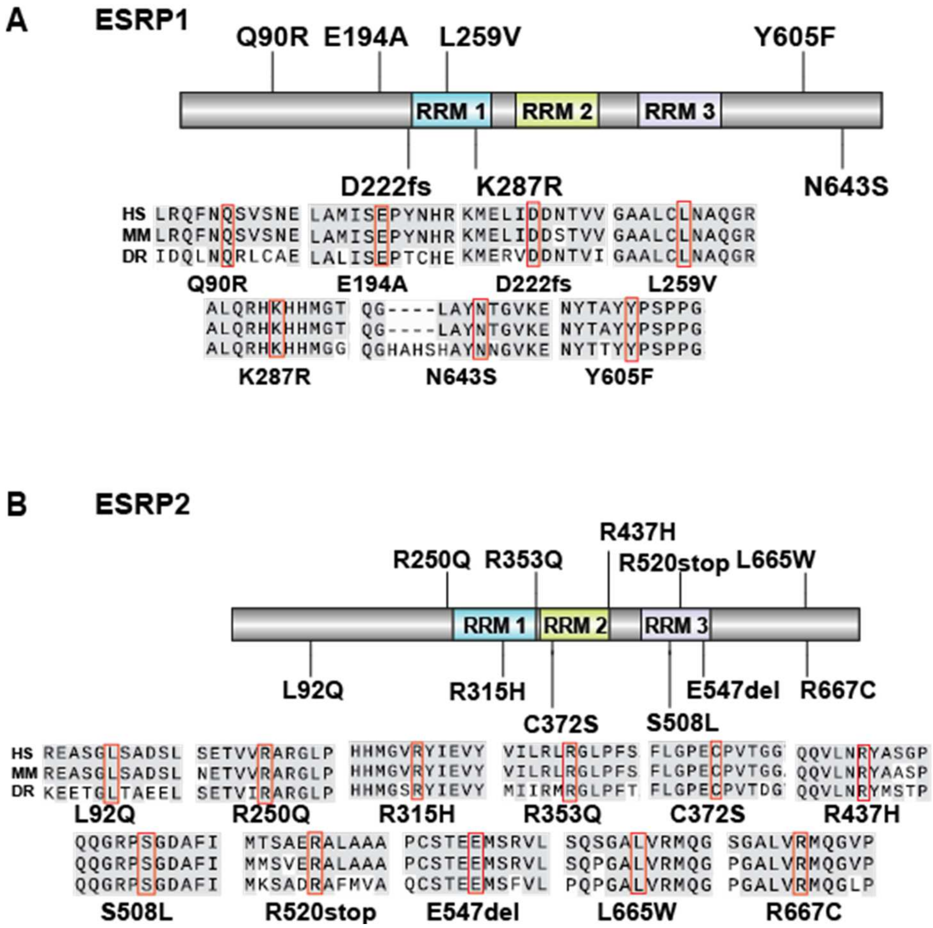


Figure 25 *ESRP1* and *ESRP2* gene variants are found in regions of high cross-species conservation

Figure 25 (Cont'd) 32 *ESRP1* and *ESRP2* gene variants were obtained by combining variants from the OFC cases in the GMFK Children's dataset and ClinVar variants associated with cleft lip and/or palate or autosomal recessive deafness. 18 variants were found to lie in amino acids fully conserved between humans, mice, and zebrafish (A) Schematic of the ESRP1 protein labeled with 7 identified gene variants, including two variants in RRM1. Truncated alignments surrounding the gene variants are provided below the diagram. (B) Schematic of the ESRP2 protein labeled with 11 identified gene variants, including two variants in each of RRM1 and RRM2, three variants are in RRM3. Truncated alignments surrounding the gene variants are provided below the diagram. Fully conserved residues between human, zebrafish, and mouse amino acid sequences surrounding the gene variants are highlighted in grey.

After identifying 18 promising gene variants based on amino acid conservation across humans, zebrafish, and mice, we performed analyses through SIFT and PolyPhen to obtain *in silico* predictions for the gene variant effects on protein function. We found that the gene variant predictions using both tools followed one of four patterns: (1) Concordant predictions from both tools annotating the variant as benign, (2) Concordant predictions from both tools annotating the variant as damaging, (3) Discordant predictions from both tools, (4) Tools unable to predict the effect of the variant on protein function. Altogether, one variant from *Esrp1* (N643S) and two variants from *Esrp2* (C372S and T475T) were predicted by both SIFT and PolyPhen to have a benign effect on protein function. Two variants from *Esrp1* (Q90R and L259V) and four from *Esrp2* (R250Q, R315H, R353Q, and R667C) were predicted by both SIFT and

PolyPhen to have a deleterious effect on protein function. SIFT and PolyPhen could not offer predictions for one variant from *Esrp1* (D222fs) and two variants from *Esrp2* (R520*and E547del). However, the remaining two *Esrp1* variants (K287R and Y605F) and four *Esrp2* variants (L92Q, R437H, S508L, and L665W) had discordant predictions between both SIFT and PolyPhen algorithms. Thus, *in silico* predictions were insufficient to annotate roughly half of our selected gene variants and required an alternate approach to predict their effects on protein function.

Imputing pathogenicity of *ESRP1* and *ESRP2* variants in our zebrafish cleft ANC rescue assay and in vitro *Arhgef11* splicing assay

After identifying 18 promising gene variants, we subjected them to our phenotypic assay in the *esrp1/2* loss-of-function zebrafish model to test for robust preservation of protein function. Both *esrp1* and *esrp2* cDNA sequences were cloned into the pCS2+8 vector backbone to generate capped mRNA for microinjection into zebrafish embryos. The *esrp1/2* morphant assay requires the microinjection of *esrp2* translation-blocking MOs into *esrp1*$^{-/-}$ embryos to maximize the number of embryos with cleft ANC phenotype in the clutch. However, since the *esrp2* MO would also neutralize exogenous *esrp2* mRNA upon co-injection into zebrafish embryos, we introduced synonymous mutations in the translational start site of the pCS2+8-*esrp2* plasmid via site-directed mutagenesis to generate *esrp2* MO-resistant *esrp2* mRNA transcripts. Co-injection of 8ng of *esrp2* MO with 200pg of either *esrp1* mRNA or MO-resistant *esrp2* mRNA fully rescued the ANC phenotype in over 75% of the injected clutches at 4dpf (**Figure 26** A-C). To test for the ability of human gene variants to rescue the cleft ANC phenotype in zebrafish, the variants were introduced into the pCS2+8-*esrp1* and pCS2+8-*esrp2* plasmids through site-directed mutagenesis. 200pg of mRNA encoding for *esrp1* or *esrp2* carrying the gene variants was co-injected with *esrp2*-MO into *esrp1*$^{-/-}$ zebrafish embryos at the 1-cell stage and stained with alcian blue at 4dpf. We found that for *esrp1,* 6 out of 7 tested gene variants rescued the ANC phenotype in a ratio like that of the *esrp1* control and were scored as benign variants. Only one variant, a frameshift mutation at the 222 aspartate residue (D222fs), had a significantly higher proportion of cleft ANC in the injected clutch compared to embryos injected with wild-type *esrp1* mRNA and was scored as a pathogenic variant (**Figure 26**B). For *esrp2*, 6 out of 12 tested gene variants rescued the ANC phenotype in a ratio similar to the esrp2 mRNA control and were scored benign. We included a silent mutation at threonine 475 (T475T) that served as an internal negative control and scored as benign. The remaining six gene variants failed to rescue the ANC phenotype at a proportion similar to the *esrp2* mRNA and were scored pathogenic (**Figure 26**C).

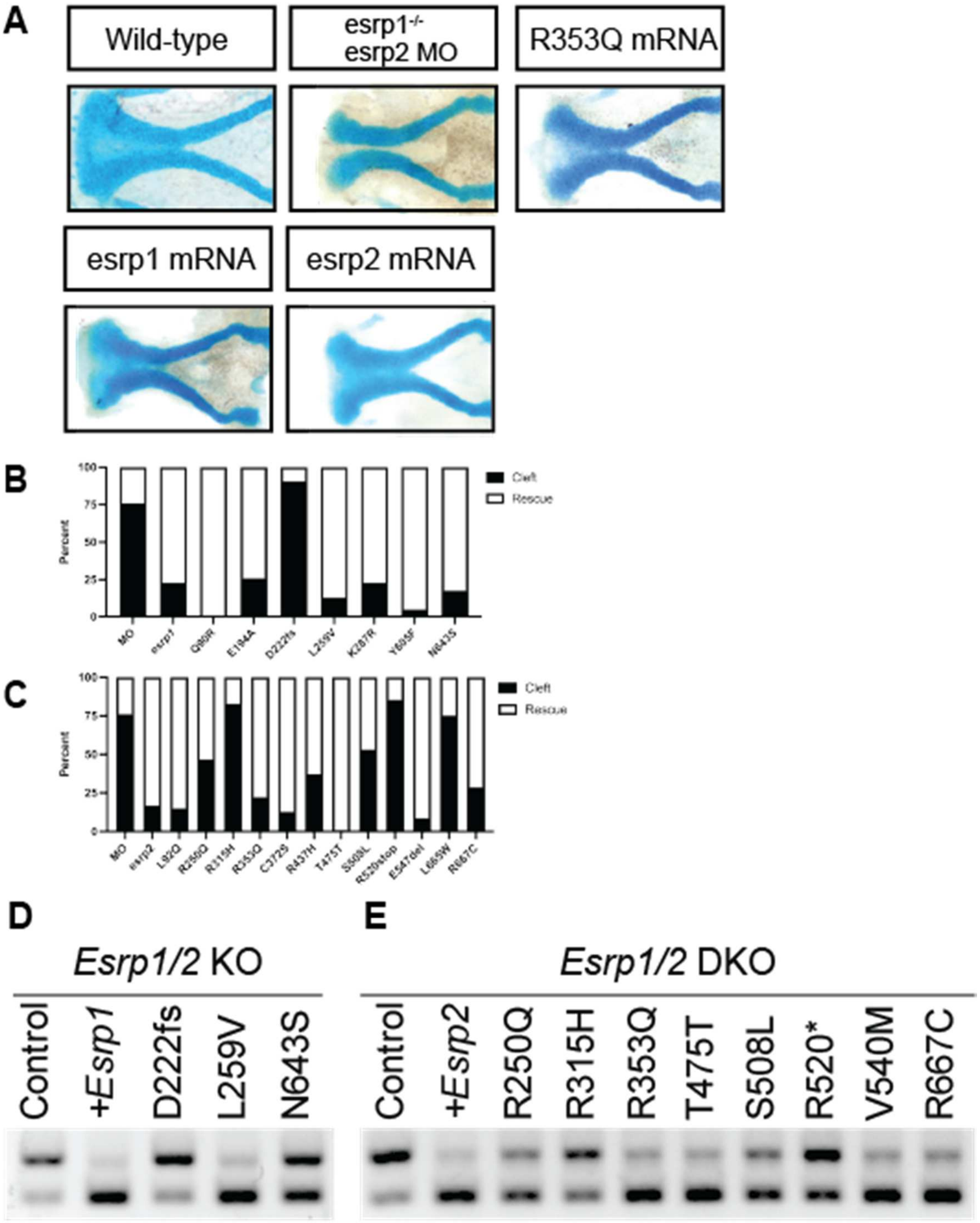


Figure 26 ***in vivo*** **and** ***in vitro*** **assays for** ***Esrp1/2*** **variant function reveal domains of functional importance**

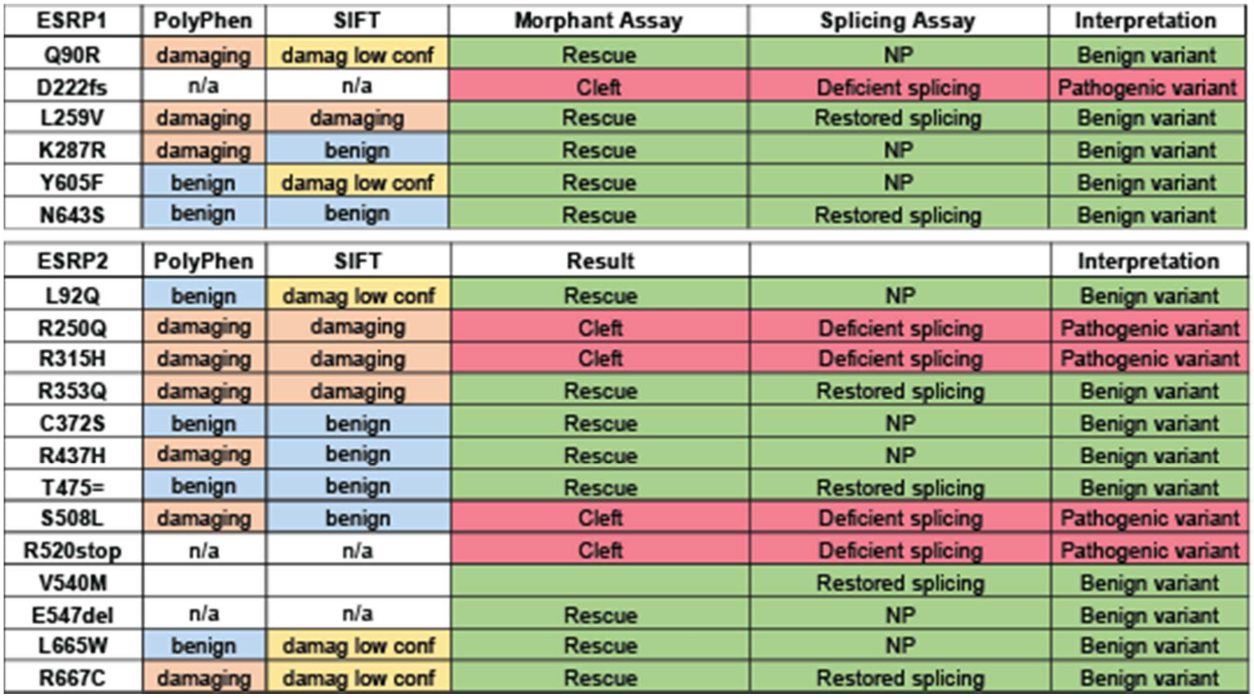
Figure 26 (Cont'd) (A) Representative images of the ANC from alcian-blue stained embryos at 4dpf after injection with *esrp2* MO and 200pg of: *esrp1* mRNA, *esrp2* R353Q mRNA. ANC was scored as a rescued ANC or cleft ANC (B) Variant rescue assay results for clutches of 30+ embryos injected with *esrp2* MO and 200pg of *esrp1* variant mRNA. (C) Variant rescue assay results for clutches of 30+ embryos injected with *esrp2* MO and 200pg of *esrp2* variant mRNA. (D) RT-PCR spanning exons 36-38 of Arhgef11 in wild type PY2T cells, *Esrp1/2* DKO PY2T cells, and *Esrp1/2* DKO PY2T cells electroporated with plasmids encoding for *Esrp1* or *Esrp1* gene variants. (E) RT-PCR spanning exons 36-38 of Arhgef11 in wild type PY2T cells, *Esrp1/2* DKO PY2T cells, and *Esrp1/2* DKO PY2T cells electroporated with plasmids encoding for *Esrp2* or *Esrp2* gene variants

Interestingly, four variants with the most potent phenotypic effect were distributed along RRM1 and RRM3 of *esrp2*. These comprised changes from arginine 250 to glutamine (R250Q) and arginine 315 to histidine (R315H) in RRM1, a change from serine 508 to leucine (S508L) in RRM3, and the introduction of a premature stop codon at arginine 520 (R520*) in RRM3. The last variant with a large effect size was found close to the C-terminus of the protein, with a change from leucine 665 to tryptophan (L665W). The remaining two variants had milder pathogenic effects. They consisted of a change from arginine 437 into histidine (R437H) at the c-terminal end of RRM2 and arginine 667 to cysteine (R667C) close to the C-terminus of the protein. These experiments underscore the functional importance of RRM1 and RRM3 for *in vivo esrp2* function. Additionally, there appears to be additional functional significance to the C-terminus of the protein, as the L665W and R667C variants disrupt *esrp2* function *in vivo*.

Since introducing *ESRP1* and *ESRP2* gene variants into zebrafish transcripts disrupted *in vivo* function of variant mRNA in our zebrafish ANC rescue, we tested a subset of 3 *Esrp1* and 8 *Esrp2* variants from our list in our PY2T cell assay to validate further their pathogenic potential and test for disruption of splicing. We performed site-directed mutagenesis on the pcDNA3.1-*Esrp1*-mCherry and pcDNA3.1-*Esrp2*-mCherry plasmids to introduce the 11 gene variants, individually electroporated them into *Esrp1/2* DKO PY2T cells and performed the RT-PCR assay 24hrs post-electroporation. We found that for *Esrp1,* only one of the tested variants (L259V) could restore *Arhgef11* splicing to the wild-type *Esrp1* mRNA levels. In contrast, the frameshift variant found to be pathogenic in our *in vivo* assay, D222fs, could not rescue splicing patterns. However, the last tested variant, N643S, partially restored some of the splicing function of *Esrp1*, expressing both *Arhgef11* isoforms in a 1:1 ratio (**Figure 26**D). Similarly, for *Esrp2,* the splicing assay results largely reproduced our *in vivo* assay results. Four out of the eight tested variants rescued the molecular splicing of *Arhgef11* to a similar degree to wild-type *Esrp2* transcript and wild-type PY2T cells. All four variants were scored benign in our *in vivo* assay, including the T475T mutation. Additionally, the remaining four variants (R250Q, R315H, S508L, and R520*) that exhibited deficient *Arhgef11* splicing in our PY2T rescue assay were all pathogenic.

The *in vivo* zebrafish ANC rescue assay and the *in vitro* PY2T splicing assays were largely concordant in imputing pathogenicity from the 18 tested gene variants (**Table 2**). Individually, the computational algorithms were poor predictors of protein function.

PolyPhen correctly predicted the effect of 8/18 (44.4%) tested gene variants, while SIFT correctly predicted the effect of 7/18 (38.8%) gene variants. However, when the predictions of both algorithms were in agreement, they could correctly predict the effect of 5/7 (71.4%) gene variants on protein function. The performance of concordant predictions was better for annotating benign variants where the algorithms correctly identified all three concordant benign variants with benign effects in both of our assays. Strikingly, the computational agreement incorrectly annotated 2/4 (50%) variants as pathogenic that had benign effects in our rescue assays. Ultimately, computational predictions of these algorithms were unable to annotate half of the identified variants and greatly overestimated the prevalence of pathogenic variants (**Table 2**).



ESRP1	PolyPhen	SIFT	Morphant Assay	Splicing Assay	Interpretation
Q90R	damaging	damag low conf	Rescue	NP	Benign variant
D222fs	n/a	n/a	Cleft	Deficient splicing	Pathogenic variant
L259V	damaging	damaging	Rescue	Restored splicing	Benign variant
K287R	damaging	benign	Rescue	NP	Benign variant
Y605F	benign	damag low conf	Rescue	NP	Benign variant
N643S	benign	benign	Rescue	Restored splicing	Benign variant

ESRP2	PolyPhen	SIFT	Result		Interpretation
L92Q	benign	damag low conf	Rescue	NP	Benign variant
R250Q	damaging	damaging	Cleft	Deficient splicing	Pathogenic variant
R315H	damaging	damaging	Cleft	Deficient splicing	Pathogenic variant
R353Q	damaging	damaging	Rescue	Restored splicing	Benign variant
C372S	benign	benign	Rescue	NP	Benign variant
R437H	damaging	benign	Rescue	NP	Benign variant
T475=	benign	benign	Rescue	Restored splicing	Benign variant
S508L	damaging	benign	Cleft	Deficient splicing	Pathogenic variant
R520stop	n/a	n/a	Cleft	Deficient splicing	Pathogenic variant
V540M				Restored splicing	Benign variant
E547del	n/a	n/a	Rescue	NP	Benign variant
L665W	benign	damag low conf	Rescue	NP	Benign variant
R667C	damaging	damag low conf	Rescue	Restored splicing	Benign variant

Table 2. Summary of results from *in silico* prediction and assays for *ESRP1/2* gene variants.

Ctnnd1* splicing and expression patterns are dependent on *Esrp1/2

CTNND1 (also known as p120-catenin) is a multi-functional protein in the catenin superfamily of proteins. Like other catenins, it possesses dual functions as a stabilizer for cell-adhesion molecules in adherens junctions and a transcriptional regulator[102, 126-130]. In addition, functional differences between epithelial and mesenchymal forms of *CTNND1* have been described[131-133]. Four major isoforms for *CTNND1* have been characterized in humans[134]. The full-length isoform, isoform 1, has a translational start site at the first methionine in the sequence (1 Met), while isoforms 2, 3, and 4 undergo splicing events that cause a 5’ truncation of the transcript and change the translational start site to methionines 55, 102, and 324, respectively. Isoform 1 of *CTNND1* is predominantly expressed in the mesenchyme, while isoform 3 is restricted to the epithelium in its expression domain. The remaining isoforms, 2 and 4, are less abundant and have not been thoroughly characterized[134].

When we aligned the amino acid sequences between human, mouse, and zebrafish CTNND1 paralogs, we found that methionines 1 and 102 are conserved in all three species. In contrast, methionine 55 is part of a 14 aa stretch missing in zebrafish (**Figure 27**A). Moreover, the splicing patterns are well-conserved, given that transcripts for the long (mesenchymal) isoform of *CTNND1* and short (epithelial) isoform after splicing of exon two that shifts the translational start site from the first methionine to methionine 102 have all been annotated in NCBI for human, mouse, and zebrafish transcripts. Previous studies have shown that both *ESRP2* and a truncated form of the full-length CTNND1 protein are colocalized in the periderm of human embryos[64], and splicing for *Ctnnd1* transcripts is deficient in the embryonic epithelium of *Esrp1*$^{-/-}$ mice

[72]. We detected the two isoforms for *Ctnnd1* in the mouse PY2T cells by performing RT-PCR using primers spanning exon 2, which is skipped in the short isoform for *Ctnnd1*. In the *Esrp1/2* DKO cell line, the splicing pattern shifts and is biased towards the longer isoform (**Figure 27**B). However, expressing Esrp1 or Esrp2 in the Esrp1/2 DKO cells restores the native splicing pattern in wild-type PY2Ts (**Figure** 27C).

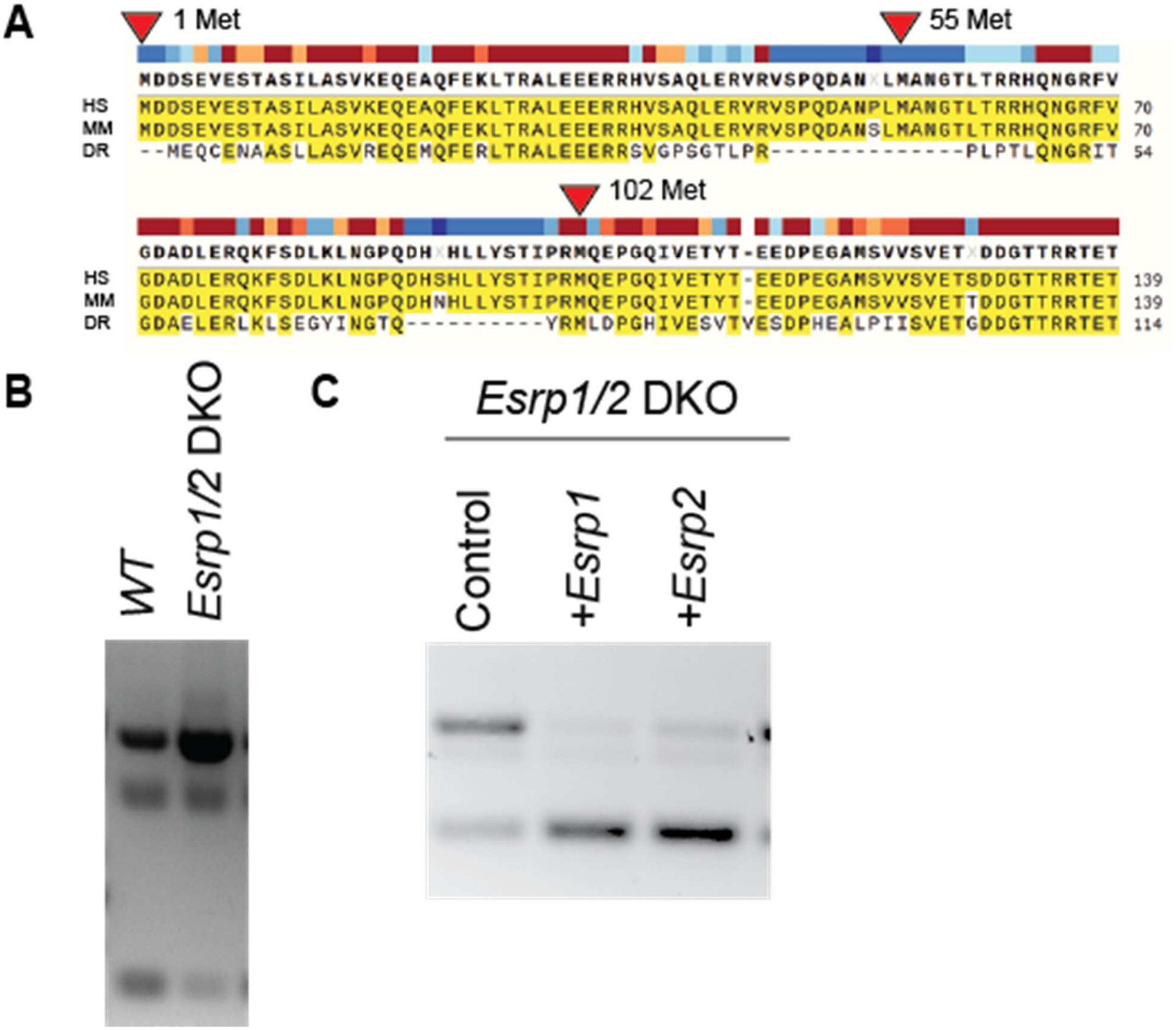


Figure 27 ***Ctnnd1*** **splicing patterns are dependent on** ***Esrp1*****/2 expression in mouse PY2T cells.**

Figure 27 (Cont'd) (A) Protein sequence alignment of the first 140 amino acids of human, mouse, and zebrafish CTNND1. Translation for isoform 1 of CTNND1 begins at the 1 methionine, while isoform 3 encodes a truncated form that starts translation at methionine 102. Methionines 55 and 324 are not conserved across all three species. (B) RT-PCR for *Ctnnd1* using primers spanning exon 2. Expression of the short isoform is reduced in *Esrp1/2* DKO cells, while expression of the long isoform is increased. (C) Expression of *Esrp1* or *Esrp2* increases the expression of the short isoform of *Ctnnd1* while decreasing the expression of the long isoform.

Discussion

Alternative splicing proteins play vital roles in craniofacial morphogenesis. Clinically, spliceosomopathies are often associated with syndromic craniofacial abnormalities due to disruption of splicing factors such as PUF60, ETUD2, SF3B4, RBM10, and ESRP2[135]. Experimentally, a growing body of animal models exhibit craniofacial phenotypes that include *Esrp1/2*-dependent models of cleft palate in mice and zebrafish in addition to other splicing regulators such as Rbfox2, srsf3, and sf3b2[74, 76, 121, 136, 137]. The ESRP proteins are uniquely expressed in epithelial structures and direct post-transcriptional modifications that distinguish protein isoforms between epithelium and mesenchyme. We applied complementary phenotypic and molecular assays to interrogate the functional consequence of identified *ESRP1/2* gene variants in cohorts of autosomal recessive deafness, cleft lip, and cleft palate.

As the magnitude of available WGS data increases, there remains a need for assigning clinically actionable information to an overabundance of genetic features of unknown significance. The ACMG-AMP criteria seek to provide a framework to evaluate clinical data and assign a predictive classification of variants based on the strength of evidence suggesting pathogenicity from variants. The SVI working group from ACMG-AMP frequently reconvenes to update, revise, and refine the ACMG criteria to provide the clearest guidance possible [103, 104]. Most recently, the working group provided further guidance regarding functional assays and experimental model systems. Among these, they highlighted the need to ascertain the gene variants' physiologic context and molecular consequence. Here, we apply complementary phenotypic assays in the zebrafish ANC rescue, in addition to our PY2T splicing assay, to assess the physiologic and molecular consequences of *ESRP1/2* gene variants observed in clinical cohorts. Our methods allowed us to discern 7 pathogenic variants out of 18 tested *ESRP1/*2 variants. Moreover, we provide functional readouts of orthologous systems across species that attest to the strongly conserved nature of epithelial splicing by the *ESRPs* in craniofacial morphogenesis.

Our methods highlight the need for experimental models to enhance the validity of *in silico* predictions of protein function. We found that while the SIFT and PolyPhen algorithms have a remarkable positive predictive value when they agree to predict benign variants, they tend to overestimate the prevalence of pathogenic variants. Furthermore, the computational agreements offer contradicting predictions in half of our tested variants and cannot accurately attribute pathogenicity without the aid of experimental models. An abundance of more computational tools further compounds this problem. However, new approaches incorporating machine learning and a meta-synthesis utilizing larger numbers of independent algorithms may provide a comprehensive tool to screen large numbers of variants for further validation in experimental systems. Still, we underscore that even though the computational predictions improved upon implementing two independent algorithms, their predictions also disagreed with half of the selected gene variants.

Our gene variant assays revealed novel insights into *ESRP1/2* protein function and downstream targets spliced by the ESRPs. We found that the gene variants with the largest effect size for our zebrafish ANC rescue assay lie in RRM1 and RRM3 of *ESRP2*. Furthermore, we have provided molecular evidence that *Esrp* transcripts rescue molecular splicing patterns of putative *Esrp*-target genes *Arhgef11* and *Ctnnd1*. Moreover, gene variants with pathogenic potential do not restore splicing patterns of *Arhgef11* and *Ctnnd1*, providing evidence that the gene variants impair *Esrp* function and likely contribute to disease pathogenicity. Our assays also answer one key desirable feature suggested by the ACMG-AMP standards, the ability of molecular assays to contribute to our understanding of mechanisms for disease.

Mutations in *CTNND1* and *CDH1* (E-cadherin) are known to cause blepharocheilodontic syndrome, including abnormal eyelids, upper lip, palate, and teeth development[64, 134, 138]. In epithelial cells, CTNND1 binds to E-cadherin to stabilize adherens junctions and desmosomes, and displacement of CTNND1 causes endocytosis of CDH1 and loss of the junction. Additionally, CTNND1 is known to modulate transcription by binding to transcription factors such as Kaiso in the canonical WNT pathway[139, 140]. It is known that alternatively spliced isoforms of *CTNND1* are differentially expressed in the epithelium and mesenchyme, and we show that the splicing patterns are dependent on *Esrp1/2* activity. However, it is not known how the alternatively spliced isoforms differ in function, alter embryonic and craniofacial morphogenesis, or contribute to disease. Thus, further studies into the functional differences between *CTNND1* isoforms are warranted and would provide insight into the disease etiology of BCD or the mechanism of the cleft palate from *ESRP* loss-of-function. Lastly, as shown in this article, many CTNND1 gene variants have been identified and would benefit from a similar experimental approach. Specifically, it would be interesting to explore the pathogenic potential and functional consequences of gene variants located within alternatively spliced exons of *CTNND1.* Such analyses would likely yield critical information into the functional differences between *CTNND1* isoforms and provide insight into the disease etiology of BCD or the mechanism of the cleft palate from *ESRP* loss-of-function.

Chapter V: Conclusions and Future Directions

Embryonic periderm is critical for craniofacial development through an *IRF6* dependent axis

OFC pathogenesis from VWS, PPS, and non-syndromic CL/P defined IRF6 as a critical factor in embryonic periderm for craniofacial morphogenesis. A growing body of work has defined the critical role of a regulatory axis involving TP63, TFAP2A, IRF6, and GRHL3 in neurulation and the formation of craniofacial structures [46, 47, 50, 141]. In Chapter II, we utilized RNA-sequencing, RNAscope ISH, and phenotypic analysis to add another target gene, *ESRP1*, to the growing list of factors regulated by IRF6. Both proteins are restricted in their expression to the periderm and epithelium of developing embryos. Unsurprisingly, loss-of-function for either *Irf6* or *Esrp1/2* leads to epithelial phenotypes: a stiff periderm in mice, periderm rupture during zebrafish gastrulation, and cleft lip in mice and zebrafish. However, even though loss-of-function of *Irf6* and *Esrp1/2* should be restricted to the epithelium, we also find mesenchymal phenotypes: mouse embryos with cleft palate and truncated limbs; and zebrafish exhibiting cleft ANC and abnormal fin development.

Interestingly, there are parallels between morphogenic signals that guide craniofacial development and limb bud development. In the developing face, CNCCs are subjected to overlapping expression domains of Shh and Fgf8 originating from both the forebrain and the embryonic epithelium in the FEZ to drive cell fate and differentiation. These same factors are responsible for the formation of the limb bud, with an apical ectodermal ridge (AER) that generates a source of Fgf8 to the developing limb, and an autocrine Shh gradient that originates from the zone of polarizing activity (ZPA). In both systems, the presence of these morphogens drive signaling cascades that direct tissue outgrowth and differentiation through Fgfr2 and wnt-pathway intermediates. Our lineage tracing experiments show that CNCC migration is not impaired in our *esrp1/2* LOF zebrafish cleft ANC model, yet FNP CNCCs fail to differentiate into chondrocytes in the medial ANC. Thus, there must be a signaling pathway dependent on *Irf6* and *Esrp* between the embryonic periderm and differentiating CNCCs. It remains to be shown whether this interaction is mediated through paracrine signaling or other forms of cell-cell communication, and if so, through which pathways. In addition, it is unclear whether the epithelial phenotypes and mesenchymal phenotypes share and etiology or represent different aspects of epithelial dysfunction. Similarly, it remains to be shown whether the limb and craniofacial defects seen in *Irf6* and *Esrp1/2* LOF share a common etiology.

As a transcription factor, *IRF6* regulates the expression of many genes. Consequently, our approach to find *irf6* targets whose transcripts are downregulated in *irf6*$^{-/-}$ zebrafish embryos revealed other genes dependent on *irf6* expression. Among these we found genes such as *rspo3*, a modulator of the wnt-pathway that is important for tooth development and bone remodeling as described in Chapter III. We also identified an additional gene, *dact1*, that binds to dishevelled and serves as an antagonist to wnt/B-catenin signaling. *dact1* and its paralog *dact2* are important for convergent extension in the early embryo. Ongoing work in the lab is further exploring a craniofacial phenotype in *dact1/2* compound mutant zebrafish and looking at epistatic relationships between *dact1/2* and other wnt-pathway genes.

Alternative splicing events in the embryonic periderm and epithelium are critical for craniofacial development

Alternative splicing proteins play pivotal roles in creating protein diversity to regulate key developmental processes. There is a growing list of disorders for patients with loss-of-function of splicing proteins called the spliceosomopathies[135]. Clinical features of these spliceosomopathies vary but can be categorized into diseases of the retina (retinitis pigmentosa), myelodysplastic syndromes, and craniofacial and developmental abnormalities. Several developmental syndromes from splicing proteins cause facial dysostoses summarized in Table 3.

Syndrome	Causal Gene	Features
Verheij Syndrome	*PUF60*	coloboma, ventricular septal defect, hip/digit abnormalities, facial dysmorphisms
Mandibulofacial dysostosis, Guion-Almeida type (MFDGA)	*EFTUD2*	microcephaly, developmental delay, craniofacial malformations
Nager Syndrome	*SF3B4*	midface retrusion, micrognathia, absent thumbs, radial hypoplasia
cerebro-costo-mandibular syndrome	*SNRPB*	rib defects, intellectual disability, cleft palate, glossoptosis, microretrognathia
N/A	*SNRPA*	craniofacial defects, intellectual defects, short stature, hand anomalies
Acrofacial dysostosis, Richieri-Costa-Pereira syndrome	*EIF4AA3*	Robin sequence (including cleft palate), laryngeal abnormalities, club feet, midline cleft mandible

Table 3 (Cont'd)		
Burn-McKeown Syndrome	*TXNL4A*	choanal atresia, sensorineural deafness, cardiac defects, and craniofacial dysmorphism
thrombocytopenia-absent radius syndrome	*RBM8A*	skeletal anomalies, craniofacial defects, microcephaly
Au-kline Bain type mental retardation and early infantile epileptic encephalopathy-54	*HNRNPR*	intellectual disability, seizures, abnormalities of the skeleton and face
TARP syndrome	*RBM10*	Talipes equinovarus (clubfoot), Atrial septal defect, Robin sequence, Persistence left superior vena cava

In chapters II and IV, we explored the phenotypes observed upon loss of epithelial-specific alternative splicing from *esrp1* and *esrp2* in both zebrafish and mouse models. We showed that esrp1/2 LOF (restricted to the embryonic epithelium) leads to cleft palate in mice and cleft ANC in zebrafish, and that splicing patterns for putative esrp1/2 target genes are disrupted. It is critical to reiterate that since esrp1/2 are specific to the epithelium, aberrant splicing within epithelial cells drives the pathology behind the phenotypes. The ESRP proteins are widely understudied, and research efforts have only just begun to identify genes whose transcripts are spliced by ESRP1/2. The FGFR2 is a well-known target of ESRP1 and undergoes an isoform switch that exchanges exon8(included in the IIIc isoform) for exon9 (included in the IIIb isoform) in epithelial cells [74, 75]. Consequently, the FGFR2IIIb isoform has modified affinity for FGF-family ligands and thus responds to a different set of factors and induces different signals to the mesenchymal FGFR2IIc isoform. While the FGFR2 isoform switch does not change ligand affinity for FGF8, a key signaling molecule in the FEZ and developing limb buds, ablation of FGFR2IIIb prevents expression of Shh in the AER and is essential for limb bud maintenance and growth[142]. Given that there are feedback mechanisms between FGFR2IIIb expressing epithelial cells and FGFR2IIIc-expressing mesenchymal cells, there is potential to uncover mechanisms of crosstalk between the developing craniofacial mesenchyme.

Studies in embryonic epithelium vs. mesenchyme in WT and *Esrp1*$^{-/-}$ mice have revealed a set of putative Esrp1 targets. Among these targets several cell-junction associated genes were identified such as *Arhgef10L, Arhgef11, Arhgap17, Enah* and *Ctnnd1.* Interestingly, *Arhgef10L, Arhgef11,* and *Arhgap17,* all belong to the Rho small family GTPase family that consists of two counteractive elements that transduce extracellular signaling and regulate intracellular actin dynamics. This system involves the phosphorylation of Rho-GDP by Rho guanine nucleotide exchange factors (RhoGEFs) like *Arhgef10L* and *Arhgef11*, that is counteracted by removal of phosphate groups from Rho-GTP by Rho GTPase activating proteins (RhoGAPs) like *Arhgap17*. This signaling pathway is particularly important for the development of tight junctions, and *Esrp1*$^{-/-}$ mice have abrogated splicing of *Arhgef11* and deficient skin integrity and barrier function. Knowing that multiple isoforms of these factors exist and are regulated by Esrp suggests the interesting hypothesis that alternative splicing of the RhoGEFs and RhoGAPs shifts the balance of RhoGTP phosphorylation states, and thus alters epithelial function.

As discussed in chapter IV, *CTNND1* is another putative target gene for Esrp1 splicing and is spliced into multiple isoforms, one of which is exclusively expressed in epithelial structures. Like other catenins, *CTNND1* has been shown to have dual functions in stabilizing cytoskeletal components, and binding to the transcription factor Kaiso, an activator of the canonical wnt pathway. Mutations in *CTNND1* cause blepharocheilodontic (BCD) syndrome, that presents with eyelid defects, bilateral cleft lip and/or cleft palate, and fewer teeth than normal that are misshapen. CTNND1 contains 3 main structural components, an N-terminal DIPA-family coiled coil that interacts with CCDC85B, 10 armadillo repeats that are characteristic of the catenin superfamily proteins, and multiple nuclear localization signal (NLS) motifs scattered throughout the protein.

Several isoforms have been identified in humans but only two major isoforms, 1 and 3, have been studied in any significant detail. Isoform 3, the shorter isoform, has two major differences from the longest transcript, isoform 1. The first difference involves the removal of exon 2 and part of exon 3 in the shorter isoform, changing the translational start site from Methionine 1 to Methionine 102 (**Figure 27**A). This splicing event results in a truncated form of *CTNND1* that skips the first hundred amino acids, about one-third of the N-terminal portion of the CCDC85B-binding domain. CCDC85B has been shown to serve as a transcriptional repressor of β-catenins, and potentially has a similar effect in *CTNND1*. If so, truncation of the CCDC85B-binding domain of *CTNND1* could serve as a regulatory step to remove repression of the transcriptive functions of *CTNND1*. Human isoforms have further spliced complexities (**Figure 28**A). At least three additional splicing events have been observed in exons termed as exon A, exon B, and exon C. Isoform 3, the epithelial isoform, consists of both the early truncation event, and the skipping of exon C in the 6th armadillo domain. Surprisingly, exon C carries an NLS motif, and could potentially offer a regulatory mechanism to compartmentalize different isoforms of *CTNND1* to the nucleus and enable its transcriptional function or the cytoplasm to activate the cytoskeletal function. Lastly, both exon A and B are located at the C-terminal end of *CTNND1* although, they are only differentially spliced in less abundant isoforms than isoforms 1 and 3. A nuclear export signal (NES) is localized in exon B, while no known features are contained in exon A. Thus, it is hypothesized that alternative splicing of *CTNND1* transcripts in epithelium vs. mesenchyme biases the resulting proteins towards one of its dual functions (**Figure 28**B). Our understanding of CTNND1 function would benefit from experimental approaches that measure subcellular localization and protein binding partners of different CTNND1 isoforms.

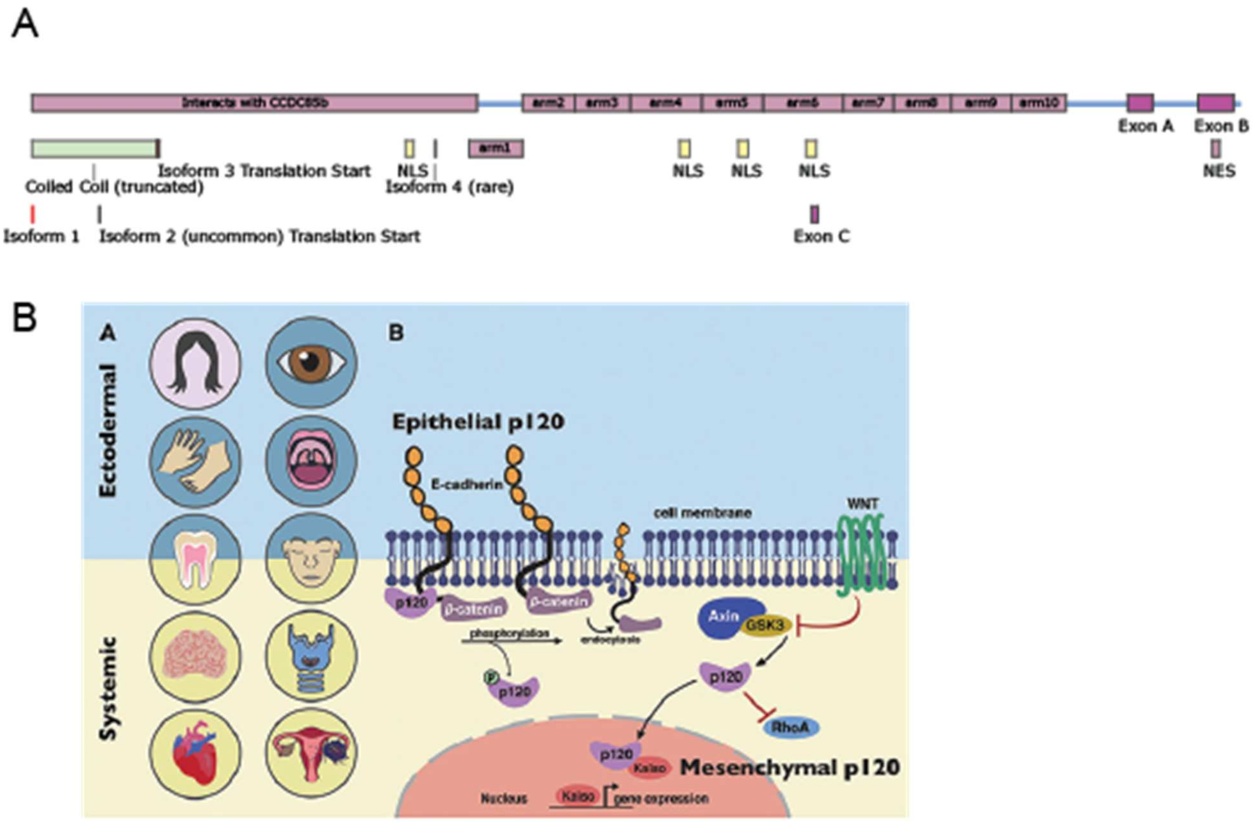


Figure 28 Alternative splicing differences in CTNND1 may bias protein function

Figure 28 (Cont'd) (A) Schematic showing several differences between known CTNND1 isoforms in humans. There are four potential translational start sites for isoforms 1, 2, 3, and 4. Isoforms 3 has a truncated CCDC85B-binding domain and skips exon C that contains an NLS compared to isoform 1. Exons A and B at the C-terminal end of CTNND1 are also spliced for other *CTNND1* transcripts. (B) Hypothesized epithelial vs systemic functions of CTNND1. Epithelial CTNND1 (p120) stabilizes cytoskeletal proteins such as E-cadherin, while mesenchymal CTNND1 also functions to activate transcription in mesenchymal-specific pathways. Cartoon adapted from Alharatani et al. *Human Molecular Genetics* [102].

Use of complementary *in vivo* and *in vitro* models across species to understand disease mechanisms and pathogenesis

Mouse models are by large the preferred mammalian *in vivo* model system for preclinical studies in biomedical research. The mouse genome has a high degree of homology to the human genome (99%) and has a robust genetic and molecular toolbox for genetic manipulation. In addition, mice have a comparatively small size compared to other mammals, making them cost-efficient for maintaining large numbers of strains and individual mice. The embryonic development of mice closely mimics that of humans and is unparalleled for high-throughput comparisons in anatomic structures throughout development. However, mouse embryos develop *in utero* which severely limits the ability to observe or experimentally manipulate embryos during development. Gestation in mice takes between 19-21 days, and a typical mouse litter consists of 6-8 mice which makes it challenging to perform experiments requiring large numbers of embryos or embryonic stages prior to parturition. The overall process of palatogenesis is grossly symmetrical between mice and humans, and both follow the development of the primary palate through fusion processes of the FNP and MXPs, and the secondary palate through palatal shelf intermediates. However, the granular detail of molecular pathways may diverge due to species' specific developmental patterns. OFCs in humans commonly involve both the lips and the palate (CLP), while most mouse models exhibit isolated cleft palate (CPO), and cleft lips (CLOs) or CLPs are rare [143, 144]. The common presentation for human patients carrying *IRF6* mutations is CLP with the presence of lower lip pits, while *Irf6* LOF mouse models have a CPO phenotype.

Zebrafish, on the other hand, complete embryonic development, and hatch within 3-4 days, in clutch sizes that can number in the hundreds for a single breeding pair. Zebrafish embryos develop within a translucent chorion that allows for the direct visualization of developmental processes and is enhanced by the availability of a wide variety of reporter lines. It is also easy to access and manipulate zebrafish embryos that can be microinjected with DNA, RNA, protein, or other molecules and constructs to perform knockdown, overexpression, or other *in vivo* assays. The accessibility of embryos allowed us to employ the optogenetic dominant-negative zebrafish system to circumvent embryonic lethality from maternal-null *irf6*$^{-/-}$ embryos and show that *irf6* loss-of-function in zebrafish causes a cleft ANC phenotype. While palatogenesis in zebrafish lacks the formation of palatal shelves and a secondary palate, the fusion processes between the FNP and MXPs in the formation of the primary palate are still present.

In chapter II, we employed both mouse and zebrafish models to characterize the expression patterns of *Esrp1/2*, and the *Esrp1/2* LOF phenotypes. SEM of *esrp1*$^{-/-}$*;* *esrp2*$^{-/-}$ zebrafish demonstrated that the cleft ANC phenotype extended beyond the ethmoid plate and included a cleft lip, demonstrating the potential for zebrafish as a model for CLO. It is important to note that most mutations found in human cleft palate cohorts are found in *ESRP2* and is largely confirmed by our gene variant testing for variants in either *ESRP1* or *ESRP*2 [64]. In contrast, *Esrp1*$^{-/-}$ mice carry a bilateral CLP phenotype, while *Esrp2*$^{-/-}$ mice do not exhibit phenotype and are viable. Zebrafish, on the other hand, require ablation of both *esrp1* and *esrp2* to exhibit phenotype. Lineage tracing experiments refined our understanding of the cleft ANC pathology by revealing that CNCCs migrate, but do not differentiate into chondrocytes at the medial ANC. Despite evolutionary differences where individual contribution of ESRP1 or ESRP2 to cleft palate pathogenesis is not fully equivalent between all three species, we were able to show that expression patterns are largely conserved, and loss of function of these factors yields similar phenotypic consequences in multiple animal models.

During these efforts, I developed a morpholino-based assay to phenocopy the *esrp1/2* zebrafish cleft ANC phenotype that proved to be a useful model as a phenotypic readout for gene variant testing in Chapter IV. However, this approach largely benefitted from complementary molecular testing through an RT-PCR based splicing readout in mouse PY2T cells. Strikingly, we were able to uncover regions of functional importance for *ESRP2,* two lying in the annotated RRM1 and RRM3 domains, and additional function at the C-terminal end of the protein. Ultimately, using a multifaceted approach that incorporates both mouse and zebrafish experimental models offers a broad understanding of conserved processes in craniofacial and palatal development. Moreover, the additional incorporation of molecular models that offer insight into disease mechanisms offer the highest level of evidence as convened by the ACMG standards.

Functional testing of gene variants reveals novel protein characteristics and expands diagnostic repertoire

As gene sequencing technologies and capabilities improve with every generation, there is an ever-increasing amount of gene sequencing data of uncertain significance being generated. In addition, due to advances in our understanding of disease pathology in combination with the development of new diagnostic techniques of increasing sensitivity we can harvest measurements for a wealth of clinical biomarkers for patients. In chapter IV, we explored an approach to gene variant testing using complementary *in vivo* and *in vitro* models of phenotype and molecular rescue in *esrp1/2* loss of function zebrafish or PY2T cells. These efforts to functionally annotate known variants of uncertain significance performed better than *in silico* predictors of protein consequences but were also able to highlight the importance of the RRM1 and RRM3 domains of ESRP2. However, ESRP1/2 function is somewhat limited in scope as it is a specialized regulator for splicing of RNA substrates in epithelial cells. Applying a similar complementary approach to test gene variants in more complicated systems may prove a lucrative approach to reveal novel interactions or mechanisms of protein function.

As discussed previously, CTNND1 is a complex protein with multiple functions as a stabilizer of cytoskeletal components of epithelial cells, but also as a transcriptional regulator for factors in the Wnt pathway. While it is known that the splicing of *CTNND1* into different isoforms seems to be important for epithelial and mesenchymal cells, the exact mechanism behind the functional differences is not known. There have been some efforts to establish animal models for *CTNND1* loss of function [102, 145]. A zebrafish CRISPR mutant for *ctdnn1* has been made that exhibits epithelial phenotypes [145]. Heterozygous *ctnnd1* mutant fish exhibit epithelial blebbing around the 6-somite stage of development, which resolves by the 20 somite-stage (**Figure 29**A, A'). However, homozygous *ctnnd1* mutations are embryonic lethal, and embryos undergo rupture after gastrulation (**Figure 29**B, B'). While this phenotype is reminiscent of the maternal-null *irf6*$^{-/-}$ rupture phenotype, one striking difference is that i*rf6*$^{-/-}$ embryos rupture during gastrulation. This rupture phenotype in *ctnnd1* embryos provides a useful background to perform gene variant testing, as we have done previously with *irf6*[51]. We have identified several gene variants in the coding sequence of *CTNND1* and testing their pathogenicity could tease out the contributing functions of spliced exons in its many domains (**Figure 29**C, D).

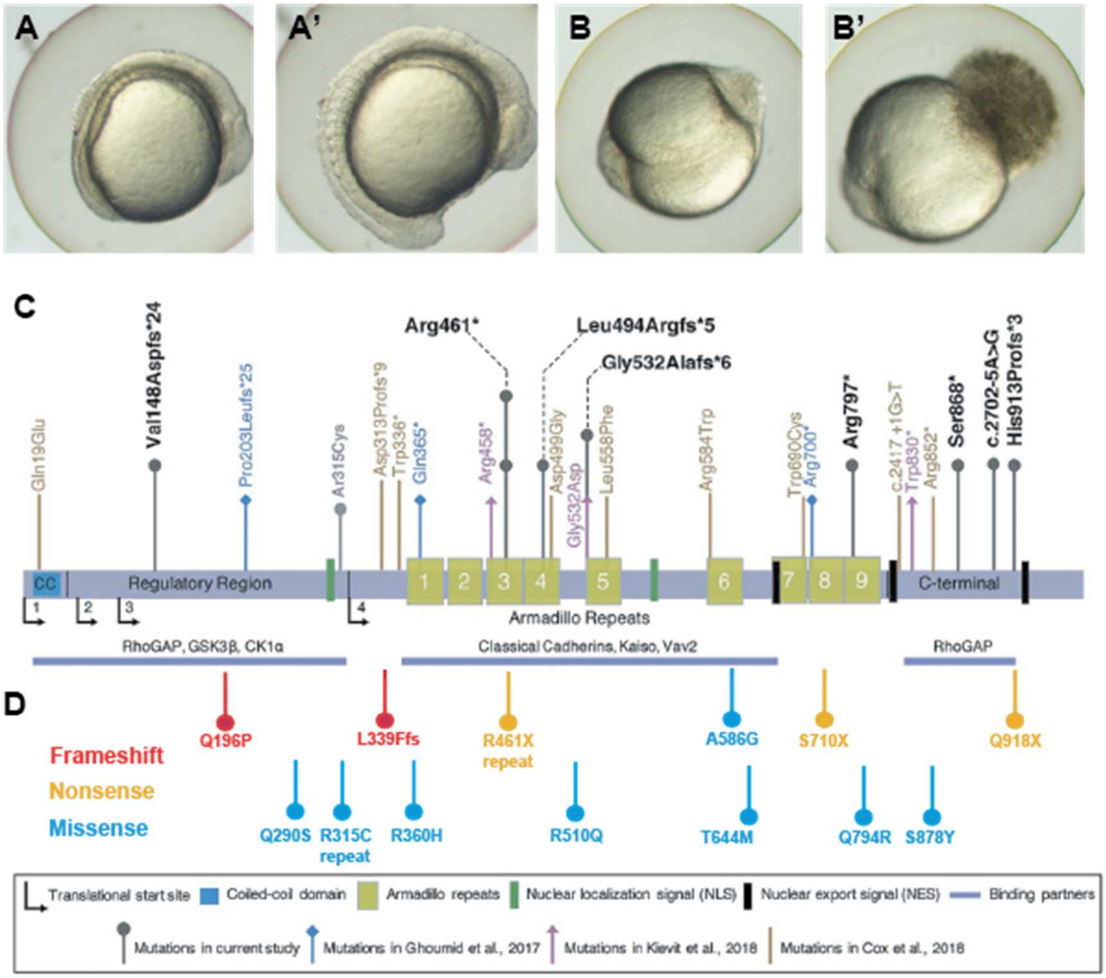


Figure 29 *ctnnd1* zebrafish embryonic rupture model for gene variant testing

(A, A') Heterozygous *ctnnd1* LOF embryos exhibit epithelial blebbing around the 6-somite stage that regenerates by the 20 somite-stage. (B, B') Homozygous *ctnnd1* LOF embryos undergo blebbing after gastrulation that progresses to embryonic rupture around 12hpf. (C) Previously identified gene variants for *CTNND1* covering the entire length of the protein (D) Newly identified frameshift, nonsense, and missense gene variants of unknown significance.

While disturbed function of *IRF6, ESRP1/2, CTNND1*, and other proteins discussed in this work contributes to malignancy, much of their detailed discussion lies beyond the scope of this dissertation [46, 70, 83, 125, 131-133]. However, it should be noted that there is a dichotomy between EMT events during development and events that contribute to malignant metastatic processes. As such, our complementary experimental models have allowed us to interrogate fundamental mechanistic questions about key genes critical for craniofacial morphogenesis but may also provide ways to characterize malignancies and guide diagnostic decisions, treatment modalities, or prognostic assessments in the management of patients. Ultimately, we envision a future where experimental approaches and disease models can guide clinical decisions that range from genetic counseling and maternal-fetal diagnostics, to understanding structural protein elements critical for disease that provide new therapeutic targets in development and cancer.

Materials and Methods

Animal breeding and gene editing

All animal experiments were performed in accordance with protocols approved by Massachusetts General Hospital Animal Care and Usage Committee. C57Bl/6J (WT) animals were obtained from the Jackson Laboratory. *Irf6*$^{R84C/+}$ mice were received as a gift from Dr. Yang Chai. *Esrp1*$^{+/-}$*; Esrp2*$^{-/-}$ mice were received from Dr. Russ Carstens. *Rspo3* mutant mice were kindly provided to Dr. Baron by Dr. Christof Nierhs (German Cancer Research Center, Heidelberg, Germany). Embryonic day 0.5 was considered to be noon on the day of the copulatory plug.

Zebrafish (*Danio rerio*) adults and embryos were maintained in accordance with approved institutional protocols at Massachusetts General Hospital. Embryos were raised at 28.5°C in E3 medium (5.0 mM NaCl, 0.17 mM KCl, 0.33 mM $CaCl_2$, 0.33 mM $MgSO_4$) with 0.0001% methylene blue. Embryos were staged according to standardized developmental timepoints by hours or days post fertilization (hpf or dpf, respectively) [146]. All zebrafish lines used for experimentation were generated from the Tübingen strain.

CRISPR sgRNA target sites were identified by a variety of online CRISPR computational programs such as zifit.partners.org/ZiFiT [147], crispr.mit.edu [148], and chopchop.rc.fas.harvard.edu [149]. sgRNAs were designed with the traditional sequence constraint of a 3' PAM sequence containing NGG and an additional sequence constraint of a 5' NG for *in vitro* RNA synthesis.

The *esrp1, esrp2* and *irf6* CRISPR sgRNAs were generated by *in vitro* transcription from a SP6 promoter as described [150]. Lyophilized Cas9 protein (PNA Bio) was resuspended in ddH_2O to a stock concentration of 1 µg/µl and stored in single-use aliquots in -80°C and kept for 6 months. One-cell staged zebrafish embryos were microinjected directly in the cytoplasm with 2 nl of a solution containing 15 ng/µl of sgRNA and 100 ng/µl of Cas9 protein pre-complexed for 5-10 minutes at room temperature. A subset of embryos injected with the sgRNA and Cas9 protein mixture were harvested for genomic DNA to confirm the presence of indels, and the rest were grown into adulthood as F0 mosaic fish. F0 adult fish were subsequently outcrossed with WT fish to generate F1 founders with germline transmission of indel alleles. F1 founders were further outcrossed with WT fish to yield a large number of heterozygotes and minimize the presence of off-target edits. Lastly, F2 heterozygotes were in-crossed to generate homozygote embryos for phenotypic analysis.

DNA for genotyping was isolated from either whole 24hpf embryos or tail fin clips using the HotSHOT method as described [151]. Genotyping primers flanking the CRISPR sgRNA site were designed using a combination of ChopChop (chopchop.rc.fas.harvard.edu) and NCBI primer BLAST (ncbi.nlm.nih.gov/tools/primer-blast/). Forward primers were synthesized by Invitrogen with 5'-FAM modifications. Microsatellite sequencing analyses were used to determine indel mutation sizes and frequencies (MGH DNA Core), and Sanger sequencing was performed on PCR amplicons of CRISPR sgRNA to confirm the exact sequence changes resulting from CRISPR mutagenesis.

We used targeted genome editing via CRISPR-cas9 mutagenesis in zebrafish to perform functional analysis of *rspo3*. A *rspo3* mutant zebrafish line was created using the cas9 RNA CCTGGCAGCCCTGGGAGCTC, which resulted in a 20 bp deletion. Genotyping primers for the *rspo3* mutant line are 5'-AAGCAGCAAAAATAAGTTCCCA- 3' and 5'-CCACTCCCCATTGCTTTATTAC-3' with FAM modification on the reverse primer for microsatellite analysis. The mutant peak was observed at 337bp and wild-type peak observed at 357bp.

CRISPR gRNA were designed using CRISPOR (http://crispor.tefor.net/) to target *rspo2* translational start sites as previously described.[29] Due to the presence of two *rspo2* transcript variants with unique translational start sites (TSS), specific pairs of gRNAs (4 total) were designed to flank each TSS. Guides ordered from Synthego were the following: AGCTCATATACGGACCCTGAAGG, AGACGCAGCAGTCCCACCGCTGG, ATGTCTTTGTACCAAACGATTGG, TCCTCTCCCTCCTCAGGAACAGG.

All four gRNAs were co-injected into *rspo3*$^{+/-}$ in-crossed single cell zebrafish embryos. Each guide was prepared at a final concentration of 1.25 μM and 2 nL were injected into each embryo. Injected embryos were raised to 9 days post fertilization, where they were subsequently fixed and stained for detailed phenotypic analysis. Stained fish were imaged using a Nikon Eclipse 80i compound microscope with a Nikon DS Ri1 camera. Measurements were taken in ImageJ. Transgenic line Tg(*sox10*:kaede)[39] was also used in this study.

mRNA sequencing and qPCR

Total RNA was isolated from 4hpf WT and maternal-null *irf6*$^{-/-}$ embryos by TRIzol and phenol-chloroform ethanol precipitation. Total RNA was quantified with the Nanodrop 2500 and assessed for quality with Bioanalyzer 2100 RNA chips (Agilent). Samples with

RNA integrity numbers (RIN) over 9 were selected to proceed with sequencing library preparation. mRNA-seq libraries were prepared with the NEBNext Ultra RNA library preparation kit with poly(A) mRNA magnetic isolation module (NEB) essentially according to manufacturer protocols. Resulting cDNA libraries were quantified by a Qubit fluorometer and assessed for quality with a Bioanalyzer. The sequencing-ready cDNA libraries were quantified with the NEBNext library quantification kit for Illumina (NEB). mRNA-seq libraries were sequenced with single-end 50 at ≈20 million reads per sample with biological triplicates. Sequencing data is available at the Gene Expression Ombibus (accession number GSE153828).

For qPCR, ~ 30 zebrafish embryos per sample were flash-frozen in liquid nitrogen. Mouse embryos from E11.5 timed pregnancies were isolated and dissected so that the head portion was flash frozen for RNA isolation and a posterior portion was frozen for genotyping. Samples were homogenized using a rotor-stator homogenizer, and RNA was isolated using a RNeasy Mini Kit (Qiagen). Total mRNA was quantified using a Nanodrop spectrophotometer and used for cDNA synthesis (Thermo). qPCR was performed with Taqman probes and reagents (Thermo), and expression was normalized to 18s rRNA or *TBP* expression.

Zebrafish embryo microinjection of mRNA and morpholinos

Microinjection of mRNA was performed by injecting 2 nl of mRNA solution with 0.05% phenol red directly into the cytoplasm of one-cell staged embryos. Lyophilized morpholinos were resuspended with ddH2O to a stock concentration of 20 ng/µl and stored at RT in aliquots. Individual aliquots were heated to 70°C and briefly vortexed before preparation of the injection mix to ensure full dissolution. Mismatch control morpholinos were injected under identical conditions to control for potential toxicities. Embryos from all methods of microinjection were examined at 3hpf to remove unfertilized embryos, which were quantified against the total number of microinjected embryos to ensure no fertilization defects were observed.

Whole-mount *in situ* hybridization

Embryos were isolated at various time points and fixed in 4% formaldehyde at 4°C for 12-16 hrs. Subsequently, embryos were washed and stored in methanol. WISH and DIG-labeled riboprobes were synthesized as described [152]. Briefly, for riboprobe synthesis, PCR was performed using embryonic cDNA as templates and T7 promoter sequence-linked reverse primers to generate cDNA templates for *in vitro* transcription. PCR reactions were purified using the NucleoSpin gel and PCR clean-up kit (Machery-Nagel). *In vitro* transcription was performed using a T7 polymerase (Roche) and DIG labeling mix (Roche). DIG-labeled riboprobes were isolated with ethanol-NaOAc precipitation, resuspended in DEPC-treated ddH_2O, and stored at -20°C. All PCR products were TOPO cloned into pGEM-T Easy vectors (Promega) and sequence verified by Sanger sequencing. WISH colorimetric signal detection was performed using an alkaline phosphatase-conjugated anti-DIG antibody (Roche) and BM Purple AP substrate (Roche).

The primers used to generate the *rspo3* RNA probe were the forward primer 5'-AACCTGTGGCTTCAAATGG-3' and reverse primer 5'-TTGTTGTCGCTCATCCAGTA-3'.[40]

The T7 promotor (gaaattaatacgactcactatagg) was added to all reverse primers. The RNA products were confirmed by gel-electrophoresis. WISH in zebrafish was performed as previously described.[41]

RNAscope *in situ* hybridization, immunofluorescence, and confocal imaging

Zebrafish and mouse embryos were fixed in 4% formaldehyde, taken through a sucrose gradient and cryo-embedded and sectioned. Probes were designed and purchased from ACD Bio, and hybridization and staining were performed according to the manufacturer's protocol. Stained sections were imaged using either a confocal microscope, where a z-stack was obtained and analyzed on ImageJ for z-stack maximum intensity projections or a standard fluorescent microscope.

For sample preparation, 48 hpf and 5 dpf zebrafish embryos were fixed using 4% formaldehyde overnight (ON) at 4°C. Adult zebrafish (6 months old) were fixed using 4% formaldehyde ON at 4°C and then decalcified ON using 0.35 M EDTA as previously published.[45] The E13.5 and E15.5 mouse embryos were fixed with 4% formaldehyde ON. n= 3 zebrafish embryos and n=3 mouse embryos were analyzed.

Subsequently, all samples were placed in 15% sucrose in PBS until the tissue sank, and then placed in 30% sucrose in PBS ON. Samples were then embedded in OCT (Tissue-Tek) and serially sectioned (10 um) in coronal orientation using a Leica CM1850 cryostat.

RNAscope probes included: Dr-rspo3-C2 (catalog number: 555121-C2), Dr-runx2a-C1 (catalog number: 409521), Dr-rspo2-C3 (catalog number: 899271-C3) Mm-Rspo3-C3 (catalog number: 402011-C3), Mm-rspo2-C2 (catalog number: 402008-C2). All probes were manufactured by Advanced Cell Diagnostics in Newark, NJ, USA. Sample pre-treatment and RNAscope were performed according to the manufacturer's instructions (Advanced Cell Diagnostics, Newark, NJ, USA). Stained slides were imaged using a

Leica SP8 inverted confocal laser scanning microscope and image processing was performed using ImageJ version 2.0 (2018). Immunofluorescence detection of mouse Runx2 (Abcam primary antibody, catalog number: ab192256; Invitrogen Alexa Fluor 488 goat anti rabbit secondary antibody) was performed following RNAscope *in situ* hybridization as described by Advanced Cell Diagnostics.

Skeletal staining and brightfield imaging

Zebrafish embryos were fixed at 96hpf or 120hpf in 4% formaldehyde and stored at 4°C overnight, washed with PBS, dehydrated in 50% ethanol, and stained with acid-free Alcian blue overnight on a rotating platform at RT as described [152]. Stained embryos were washed with ddH2O and subsequently bleached (0.8% W/V KOH, 0.1% Tween 20, 0.9% H2O2) until cell pigmentation was no longer present. For double-stained embryos with Alcian blue and alizarin red, embryos were stained with a 0.05% alizarin red solution in ddH2O for 30 minutes on a rotating platform at RT following bleaching with KOH and H2O2. Afterward, double-stained embryos were placed in three changes of a tissue-clearing solution consisting of 25% glycerol and 0.1% KOH, each for 25 min. Whole and dissected stained embryos were mounted in 3% methylcellulose on a depression slide and imaged using a Nikon Eclipse 80i compound microscope with a Nikon DS Ri1 camera. Z-stacked images were taken to increase the depth-of-field with the NIS Element BR 3.2 software. Stacked images were processed by ImageJ to generate maximum intensity projection images.

Tartrate- resistant acid phosphatase (TRAP) staining for osteoclast activity was performed (n= 5 wild-type and 5 *rspo3*-/-) as adapted from previous study.[43] Imaging was performed using Nikon Eclipse 80i microscope (Melville, NY, USA) and NIS-

Elements Br imaging software version 4.40 (2015). Measurements were taken in ImageJ. *In vivo* Alizarin red S staining of 9 dpf zebrafish was performed as previously described.[44] Alizarin red S and *sox10:kaede* fluorescence was imaged using a Leica SP8 inverted confocal laser scanning microscope. Maximum intensity projections of z-stacks were generated using ImageJ version 2.0.

For scanning electron microscopy (SEM), 4dpf embryos were fixed in half-strength Karnovsky fixative. Samples were processed, and images were obtained by CBSET, Inc. Lexington, MA. Mouse embryos from *Irf6*$^{R84C/+}$; *Esrp1*$^{+/-}$; *Esrp2*$^{+/-}$ crosses were collected into PBS. Tail clips were saved for genotyping and embryos were fixed in 10% formalin for bright field imaging. After imaging, skulls (excluding the lower jaw) were cryosectioned and sections were stained with Hematoxylin and Eosin.

Optogenetic expression of *irf6* in zebrafish

Genes *irf6*, *irf6*-ENR, i*rf6*^{R84C}, and mCherry were isolated by PCR from various templates and inserted into the pGL4.23-(C120x5)-TATA vector with In-Fusion cloning (Clontech) according to manufacturer instructions using a 1:2 vector-to-insert ratio to generate optogenetic response plasmids. The constructs were transformed in Stellar chemically competent cells (Clontech), and colonies were screened by PCR, restriction digests, Sanger sequencing, and whole plasmid sequencing to verify the sequence identities and accuracy of the constructs. Light-sensitive response protein VP16-EL222 was subcloned into pCS2+8 and in vitro transcribed from the SP6 promoter as described above to generate capped mRNA for embryo microinjections. The optogenetics injection mix was comprised of 25 ng/µl EL222 and 10 ng/µl pGL4.23 response plasmid with 0.05% phenol red. Each embryo was microinjected with 2 nl of the optogenetics injection mix directly in the cytoplasm at the one-cell stage, immediately wrapped in aluminum foil, and placed into a dark incubator. Unfertilized and abnormal embryos were removed at 3hpf in the dark room with limited exposure to ambient light. Injected embryos were divided into two groups (dark and light) at the desired developmental stage in E3 media without methylene blue and placed under 465nm blue light (LED panel, HQRP) at 0.3 mW/cm2 (measured by a PM100D digital power meter with an SV120VC photodiode power sensor, ThorLabs) with constant illumination. Control embryos containers were wrapped in aluminum foil.

Lineage Tracing

Embryos originating from an *espr1*$^{-4bp/-4bp}$;*esrp2*$^{+/-14bp}$ in-cross were injected with 8 ng *esrp1* MO and 4 ng *esrp2* MO at the one-cell stage, or uninjected WT embryos, all in a Tg(*sox10:kaede*) background were grown until 20 somites, oriented for imaging in the sagittal position, and encased in 1% low-melt agarose. Using the 405nm UV laser and ROI setting in a Leica SP8 confocal microscope, the anterior-most portion of NCCs that contribute to the FNP were unilaterally photoconverted, keeping the alternate side as an internal control, as previously described [35]. Photoconverted embryos were carefully micro-dissected out of the agar and grown in E3 at 28.5^{o}C until 4dpf and imaged again to track the photoconverted cells. Maximum projections of the photoconverted half of the embryo, or the planes consisting of the palate, in 14hpf or 4dpf embryos, respectively, were generated using Fiji/ImageJ.

Micro-computed tomography

Wild-type and *rspo3* mutant adult zebrafish were sacrificed at 6 months of age, n= 9 wild-type and 7 mutant zebrafish. All zebrafish were scanned as previously described.[46]

The voxel size of Micro-CT analysis is 10.5 μm. The examiner (K.W.) was blinded to the genotype of the zebrafish. Images were reconstructed, analyzed and viewed using Amira software version 6.

Measurement of bone volume

The reconstructed bitmap image (BMP) files were converted to NIfTI format for simplification, using Amira software. The threshold tool values were consistent between the samples (32-72 threshold logic unit). Each zebrafish skull was segmented into bone elements (dentary, anguloarticular, premaxilla, maxilla and parasphenoid) using Amira manufacture's instruction. n= 9 wild type and 7 mutant zebrafish were analyzed at 6 months of age.

Quantitative RT-PCR

Three independent samples of wild-type and *rspo3* CRISPR/Cas9 (-20 base pairs micro- deletion mutants) at 6 hpf were collected and measured in triplicate in order to characterize the *rspo3* mutant. We decided to collect embryos at 6 hpf, because it has been reported that *rspo3* mRNA is highly expressed in zebrafish embryos at this time point.[34] In addition, three independent 1-cell stage and 24 hpf wild type embryo samples were collected and measured to define the expression of *rspo3* mRNAs. RNA extractions were performed using RNeasy Mini Kit (Qiagen). SuperScript First-Strand Synthesis System IV (Thermo Fisher Scientific) was used to synthesize first-strand cDNA. Quantitative reverse-transcription PCR (qRT-PCR) was performed using *rspo3* Taqman assay (Dr03109282_m1) Taqman Fast Advanced master mix (Thermo Fisher Scientific) and normalized to 18S rRNA expression (Hs03003631_g1). qPCR was performed on a StepOnePlus Real-Time PCR system (Applied Biosystems).

Statistical analysis

IBM SPSS statistics version 26 was used for all Student's t-test statistical analyses. Student's t-test was used to compare between the two groups. Prism 9 software was used to perform Kruskal-Wallis statistical test with multiple comparisons when more than two groups were compared. Statistical significance was set at p-value ≤ 0.05. Asterisks in the figures indicate p-value ≤ 0.05. Data presented as means ± SEM.

Gene variant identification, sequence alignment, and variant effect prediction

The Gabriella Miller Kids First (GMKF) Research Program's WGS dataset was filtered for OFC cases carrying *ESRP1* or *ESRP2* variants that were (1) heterozygous in the affected patient, (2) had a minor allele frequency no greater than 0.001 in any population in gnomAD or 1000 Genomes, and (3) had a variant consequence of missense, frameshift, stop-gain, splicing, or inframe insertion/deletion. We further supplemented the resulting list with additional gene variants from ClinVar associated with an orofacial cleft or autosomal recessive deafness to generate a list of 12 *ESRP1* and 20 *ESRP2* variants of uncertain significance.

To refine the gene variant list to amino acids for use in our mouse and zebrafish assays, the *ESRP1* and *ESRP2* homolog amino acid sequences were aligned between human, mouse, and zebrafish using Clustal Omega[Ref]. The resulting 7 *ESRP1* and 12 *ESRP2* variants were introduced into both SIFT and PolyPhen to obtain the predicted change in protein function and ranked as benign, pathogenic, or of unknown significance.

Plasmid generation, site-directed mutagenesis, and mRNA synthesis

mRNA from zebrafish embryos was collected at multiple time points from 6hpf-4dpf, reverse transcribed, and combined to make pooled cDNA to clone the *esrp1* CDS. A plasmid containing the *esrp2* CDS was purchased as a plasmid from GenomicsOnline. *esrp1* and *esrp2* were each cloned into a pCS2+8 plasmid backbone using the In-Fusion system (Clontech). The resulting pCS2+8-*esrp2* plasmid was mutagenized with synonymous mutations surrounding the translational start-site using the GeneArt site-directed mutagenesis (SDM) system (ThermoFisher) to generate *esrp2* transcripts resistant to *esrp2* morpholino binding. The 19 *ESRP1* and *ESRP2* variants were each individually introduced to the pCS2+8-*esrp1* or MO-resistant pCS2+8-*esrp2* plasmids through the GeneArt SDM system. All generated pCS2+8 plasmids were digested with NotI at 37^{o}C for 1hr, and capped mRNA was synthesized using the SP6 mMessage mMachine kit (ThermoFisher).

The pIBX-C-FF(B)-mCherry-Esrp1(2A)-+CKLP plasmid containing the mouse *Esrp1* cDNA sequence fused to a mCherry tag was a gift from Russ Carstens at the University of Pennsylvania. Mouse *Esrp2* cDNA was purchased from Genomics Online. *Esrp1* cDNA was cloned into the pcDNA3.1 backbone containing a CMV promoter and SV40 polyA tailing sequence for expression in mammalian cells using the In-Fusion cloning (Clontech) to generate the *pcDNA3.1-Esrp1-mCherry plasmid*. A mCherry tag was fused in-frame onto the *Esrp2* cDNA and introduced into the pcDNA3.1 backbone through a multi-insert in-Fusion cloning strategy, using the pIBX-C-FF(B)-mCherry-Esrp1(2A)-+CKLP as the template for the 2A-mCherry sequence to generate the pcDNA3.1-*Esrp2-mCherry plasmid*. Gene variants for *Esrp1* and *Esrp2* were introduced through the GeneArt SDM system.

Zebrafish cleft ANC mRNA rescue assay

We previously generated a zebrafish line carrying homozygous loss-of-function alleles in *esrp1* through CRISPR[121]. *esrp2* morpholinos (GeneTools) were reconstituted to a concentration of 8ug/uL in water and stored in single-use aliquots at RT. 2nL droplets containing (1) 8ng *esrp2* morpholino, (2) 0.05% phenol red and (3) 200pg of *esrp1, esrp2,* or *esrp* gene-variant mRNA were microinjected directly into the cytoplasm of one-cell stage *esrp1*$^{-/-}$ zebrafish embryos and grown until 4dpf. At 4dpf, embryos were fixed in 4% formaldehyde, stained with acid-free Alcian blue as previously described [121, 153], and microdissected to inspect the anterior neurocranium (ANC). The ANC was scored as (1) Cleft ANC, (2) Partially Rescued ANC, or (3) Fully Rescued ANC.

PY2T cell maintenance and transfection

Mouse Py2T cells and *Esrp1/2* DKO Py2T were a gift from Russ Carstens from the University of Pennsylvania [73]. Cells were maintained in DMEM supplemented with 10% FBS and penicillin/streptomycin and were not subcultured past passage 30. 10.8ug of plasmid was transfected onto 10^6 cells using the 100uL Neon system (ThermoFisher) with a single, 30second pulse at 1400V and plated onto 6-well plates. Cells were harvested for RNA after 24hr, reverse transcribed, and the cDNA was used for RT-PCR using primers spanning the splice junctions for exons 1 and 3 of *Ctnnd1* and exons 36 and 38 of *Arhgef11*.

References

1. Kuratani, S., *Craniofacial development and the evolution of the vertebrates: the old problems on a new background.* Zoolog Sci, 2005. **22**(1): p. 1-19.
2. Kuratani, S., *Developmental studies on the vertebrate head evolution.* Zoolog Sci, 2005. **22**(12): p. 1361-6.
3. Sham, M.H., et al., *The zinc finger gene Krox20 regulates HoxB2 (Hox2.8) during hindbrain segmentation.* Cell, 1993. **72**(2): p. 183-96.
4. Krumlauf, R., et al., *Hox homeobox genes and regionalisation of the nervous system.* J Neurobiol, 1993. **24**(10): p. 1328-40.
5. Shyamala, K., et al., *Neural crest: The fourth germ layer.* J Oral Maxillofac Pathol, 2015. **19**(2): p. 221-9.
6. Martik, M.L. and M.E. Bronner, *Riding the crest to get a head: neural crest evolution in vertebrates.* Nat Rev Neurosci, 2021. **22**(10): p. 616-626.
7. Santagati, F. and F.M. Rijli, *Cranial neural crest and the building of the vertebrate head.* Nat Rev Neurosci, 2003. **4**(10): p. 806-18.
8. Lumsden, A., N. Sprawson, and A. Graham, *Segmental origin and migration of neural crest cells in the hindbrain region of the chick embryo.* Development, 1991. **113**(4): p. 1281-91.
9. Prince, V. and A. Lumsden, *Hoxa-2 expression in normal and transposed rhombomeres: independent regulation in the neural tube and neural crest.* Development, 1994. **120**(4): p. 911-23.
10. Krumlauf, R., *Hox genes and pattern formation in the branchial region of the vertebrate head.* Trends Genet, 1993. **9**(4): p. 106-12.
11. Creuzet, S., et al., *Negative effect of Hox gene expression on the development of the neural crest-derived facial skeleton.* Development, 2002. **129**(18): p. 4301-13.
12. Trainor, P.A., K.R. Melton, and M. Manzanares, *Origins and plasticity of neural crest cells and their roles in jaw and craniofacial evolution.* Int J Dev Biol, 2003. **47**(7-8): p. 541-53.
13. Minoux, M., et al., *Rostral and caudal pharyngeal arches share a common neural crest ground pattern.* Development, 2009. **136**(4): p. 637-45.
14. Hunt, P., et al., *A distinct Hox code for the branchial region of the vertebrate head.* Nature, 1991. **353**(6347): p. 861-4.
15. Hunter, M.P. and V.E. Prince, *Zebrafish hox paralogue group 2 genes function redundantly as selector genes to pattern the second pharyngeal arch.* Dev Biol, 2002. **247**(2): p. 367-89.

16. Rijli, F.M., et al., *A homeotic transformation is generated in the rostral branchial region of the head by disruption of Hoxa-2, which acts as a selector gene.* Cell, 1993. **75**(7): p. 1333-49.
17. Gendron-Maguire, M., et al., *Hoxa-2 mutant mice exhibit homeotic transformation of skeletal elements derived from cranial neural crest.* Cell, 1993. **75**(7): p. 1317-31.
18. Som, P.M. and T.P. Naidich, *Illustrated review of the embryology and development of the facial region, part 1: Early face and lateral nasal cavities.* AJNR Am J Neuroradiol, 2013. **34**(12): p. 2233-40.
19. Abramyan, J. and J.M. Richman, *Recent insights into the morphological diversity in the amniote primary and secondary palates.* Dev Dyn, 2015. **244**(12): p. 1457-68.
20. Bush, J.O. and R. Jiang, *Palatogenesis: morphogenetic and molecular mechanisms of secondary palate development.* Development, 2012. **139**(2): p. 231-43.
21. Dixon, M.J., et al., *Cleft lip and palate: understanding genetic and environmental influences.* Nat Rev Genet, 2011. **12**(3): p. 167-78.
22. Swartz, M.E., et al., *Examination of a palatogenic gene program in zebrafish.* Dev Dyn, 2011. **240**(9): p. 2204-20.
23. Martinez-Alvarez, C., et al., *Medial edge epithelial cell fate during palatal fusion.* Dev Biol, 2000. **220**(2): p. 343-57.
24. Cuervo, R. and L. Covarrubias, *Death is the major fate of medial edge epithelial cells and the cause of basal lamina degradation during palatogenesis.* Development, 2004. **131**(1): p. 15-24.
25. Jiang, R., J.O. Bush, and A.C. Lidral, *Development of the upper lip: morphogenetic and molecular mechanisms.* Dev Dyn, 2006. **235**(5): p. 1152-66.
26. Carette, M.J. and M.W. Ferguson, *The fate of medial edge epithelial cells during palatal fusion in vitro: an analysis by DiI labelling and confocal microscopy.* Development, 1992. **114**(2): p. 379-88.
27. Jin, J.Z. and J. Ding, *Analysis of cell migration, transdifferentiation and apoptosis during mouse secondary palate fusion.* Development, 2006. **133**(17): p. 3341-7.
28. Nawshad, A., *Palatal seam disintegration: to die or not to die? that is no longer the question.* Dev Dyn, 2008. **237**(10): p. 2643-56.
29. Dougherty, M., et al., *Distinct requirements for wnt9a and irf6 in extension and integration mechanisms during zebrafish palate morphogenesis.* Development, 2013. **140**(1): p. 76-81.

30. Schilling, T.F. and C.B. Kimmel, *Musculoskeletal patterning in the pharyngeal segments of the zebrafish embryo.* Development, 1997. **124**(15): p. 2945-60.
31. Cubbage, C.C. and P.M. Mabee, *Development of the cranium and paired fins in the zebrafish Danio rerio (Ostariophysi, Cyprinidae).* J Morphol, 1996. **229**(2): p. 121-160.
32. Kimmel, C.B., et al., *The shaping of pharyngeal cartilages during early development of the zebrafish.* Dev Biol, 1998. **203**(2): p. 245-63.
33. Shah, R.M., E.M. Donaldson, and G.G. Scudder, *Toward the origin of the secondary palate. A possible homologue in the embryo of fish, Onchorhynchus kisutch, with description of changes in the basement membrane area.* Am J Anat, 1990. **189**(4): p. 329-38.
34. Gfrerer, L., M. Dougherty, and E.C. Liao, *Visualization of craniofacial development in the sox10: kaede transgenic zebrafish line using time-lapse confocal microscopy.* J Vis Exp, 2013(79): p. e50525.
35. Dougherty, M., et al., *Embryonic fate map of first pharyngeal arch structures in the sox10: kaede zebrafish transgenic model.* J Craniofac Surg, 2012. **23**(5): p. 1333-7.
36. Hu, D., R.S. Marcucio, and J.A. Helms, *A zone of frontonasal ectoderm regulates patterning and growth in the face.* Development, 2003. **130**(9): p. 1749-58.
37. Mossey, P., *Global strategies to reduce the healthcare burden of craniofacial anomalies.* Br Dent J, 2003. **195**(10): p. 613.
38. Goodwin, A.F., et al., *From Bench to Bedside and Back: Improving Diagnosis and Treatment of Craniofacial Malformations Utilizing Animal Models.* Curr Top Dev Biol, 2015. **115**: p. 459-92.
39. Wehby, G.L. and C.H. Cassell, *The impact of orofacial clefts on quality of life and healthcare use and costs.* Oral Dis, 2010. **16**(1): p. 3-10.
40. Kondo, S., et al., *Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes.* Nat Genet, 2002. **32**(2): p. 285-9.
41. Knight, A.S., et al., *Developmental expression analysis of the mouse and chick orthologues of IRF6: the gene mutated in Van der Woude syndrome.* Dev Dyn, 2006. **235**(5): p. 1441-7.
42. Mukhopadhyay, N., et al., *Genome-Wide Association Study of Non-syndromic Orofacial Clefts in a Multiethnic Sample of Families and Controls Identifies Novel Regions.* Front Cell Dev Biol, 2021. **9**: p. 621482.
43. Mukhopadhyay, N., et al., *Genome-wide association study of multiethnic nonsyndromic orofacial cleft families identifies novel loci*

specific to family and phenotypic subtypes. Genet Epidemiol, 2022. **46**(3-4): p. 182-198.

44. Parada-Sanchez, M.T., et al., *Disrupted IRF6-NME1/2 Complexes as a Cause of Cleft Lip/Palate.* J Dent Res, 2017. **96**(11): p. 1330-1338.
45. Zucchero, T.M., et al., *Interferon regulatory factor 6 (IRF6) gene variants and the risk of isolated cleft lip or palate.* N Engl J Med, 2004. **351**(8): p. 769-80.
46. Ferretti, E., et al., *A conserved Pbx-Wnt-p63-Irf6 regulatory module controls face morphogenesis by promoting epithelial apoptosis.* Dev Cell, 2011. **21**(4): p. 627-41.
47. Kousa, Y.A., et al., *The TFAP2A-IRF6-GRHL3 genetic pathway is conserved in neurulation.* Hum Mol Genet, 2019. **28**(10): p. 1726-1737.
48. Hammond, N.L., J. Dixon, and M.J. Dixon, *Periderm: Life-cycle and function during orofacial and epidermal development.* Semin Cell Dev Biol, 2019. **91**: p. 75-83.
49. Liu, H., et al., *Irf6 directly regulates Klf17 in zebrafish periderm and Klf4 in murine oral epithelium, and dominant-negative KLF4 variants are present in patients with cleft lip and palate.* Hum Mol Genet, 2016. **25**(4): p. 766-76.
50. Ingraham, C.R., et al., *Abnormal skin, limb and craniofacial morphogenesis in mice deficient for interferon regulatory factor 6 (Irf6).* Nat Genet, 2006. **38**(11): p. 1335-40.
51. Li, E.B., et al., *Rapid functional analysis of computationally complex rare human IRF6 gene variants using a novel zebrafish model.* PLoS Genet, 2017. **13**(9): p. e1007009.
52. Reid, B.S., et al., *Ectodermal Wnt/beta-catenin signaling shapes the mouse face.* Dev Biol, 2011. **349**(2): p. 261-9.
53. O'Donoghue, S., et al., *Brain Development in School-Age and Adolescent Girls: Effects of Turner Syndrome, Estrogen Therapy, and Genomic Imprinting.* Biol Psychiatry, 2020. **87**(2): p. 113-122.
54. Knight, R.D. and T.F. Schilling, *Cranial neural crest and development of the head skeleton.* Adv Exp Med Biol, 2006. **589**: p. 120-33.
55. Helms, J.A., D. Cordero, and M.D. Tapadia, *New insights into craniofacial morphogenesis.* Development, 2005. **132**(5): p. 851-61.
56. Creuzet, S., G. Couly, and N.M. Le Douarin, *Patterning the neural crest derivatives during development of the vertebrate head: insights from avian studies.* J Anat, 2005. **207**(5): p. 447-59.

57. Cordero, D.R., et al., *Cranial neural crest cells on the move: their roles in craniofacial development.* Am J Med Genet A, 2011. **155A**(2): p. 270-9.
58. Losa, M., et al., *Face morphogenesis is promoted by Pbx-dependent EMT via regulation of Snail1 during frontonasal prominence fusion.* Development, 2018. **145**(5).
59. Gritli-Linde, A., *The etiopathogenesis of cleft lip and cleft palate: usefulness and caveats of mouse models.* Curr Top Dev Biol, 2008. **84**: p. 37-138.
60. Yuan, Q., S.H. Blanton, and J.T. Hecht, *Genetic causes of nonsyndromic cleft lip with or without cleft palate.* Adv Otorhinolaryngol, 2011. **70**: p. 107-13.
61. Juriloff, D.M. and M.J. Harris, *Mouse genetic models of cleft lip with or without cleft palate.* Birth Defects Res A Clin Mol Teratol, 2008. **82**(2): p. 63-77.
62. Marazita, M.L., *The evolution of human genetic studies of cleft lip and cleft palate.* Annu Rev Genomics Hum Genet, 2012. **13**: p. 263-83.
63. Yu, Y., et al., *Genome-wide analyses of non-syndromic cleft lip with palate identify 14 novel loci and genetic heterogeneity.* Nat Commun, 2017. **8**: p. 14364.
64. Cox, L.L., et al., *Mutations in the Epithelial Cadherin-p120-Catenin Complex Cause Mendelian Non-Syndromic Cleft Lip with or without Cleft Palate.* Am J Hum Genet, 2018. **102**(6): p. 1143-1157.
65. Leslie, E.J., et al., *Search for genetic modifiers of IRF6 and genotype-phenotype correlations in Van der Woude and popliteal pterygium syndromes.* Am J Med Genet A, 2013. **161A**(10): p. 2535-2544.
66. Beaty, T.H., M.L. Marazita, and E.J. Leslie, *Genetic factors influencing risk to orofacial clefts: today's challenges and tomorrow's opportunities.* F1000Res, 2016. **5**: p. 2800.
67. Rorick, N.K., et al., *Genomic strategy identifies a missense mutation in WD-repeat domain 65 (WDR65) in an individual with Van der Woude syndrome.* Am J Med Genet A, 2011. **155A**(6): p. 1314-21.
68. Kousa, Y.A. and B.C. Schutte, *Toward an orofacial gene regulatory network.* Dev Dyn, 2016. **245**(3): p. 220-32.
69. de la Garza, G., et al., *Interferon regulatory factor 6 promotes differentiation of the periderm by activating expression of Grainyhead-like 3.* J Invest Dermatol, 2013. **133**(1): p. 68-77.
70. Richardson, R.J., et al., *Irf6 is a key determinant of the keratinocyte proliferation-differentiation switch.* Nat Genet, 2006. **38**(11): p. 1329-34.

71. Iwata, J., et al., *Smad4-Irf6 genetic interaction and TGFbeta-mediated IRF6 signaling cascade are crucial for palatal fusion in mice.* Development, 2013. **140**(6): p. 1220-30.
72. Lee, S., et al., *Cleft lip and cleft palate in Esrp1 knockout mice is associated with alterations in epithelial-mesenchymal crosstalk.* Development, 2020. **147**(21).
73. Lee, S., et al., *Esrp1-Regulated Splicing of Arhgef11 Isoforms Is Required for Epithelial Tight Junction Integrity.* Cell Rep, 2018. **25**(9): p. 2417-2430 e5.
74. Bebee, T.W., et al., *The splicing regulators Esrp1 and Esrp2 direct an epithelial splicing program essential for mammalian development.* Elife, 2015. **4**.
75. Warzecha, C.C., et al., *ESRP1 and ESRP2 are epithelial cell-type-specific regulators of FGFR2 splicing.* Mol Cell, 2009. **33**(5): p. 591-601.
76. Burguera, D., et al., *Evolutionary recruitment of flexible Esrp-dependent splicing programs into diverse embryonic morphogenetic processes.* Nat Commun, 2017. **8**(1): p. 1799.
77. Van Otterloo, E., T. Williams, and K.B. Artinger, *The old and new face of craniofacial research: How animal models inform human craniofacial genetic and clinical data.* Dev Biol, 2016. **415**(2): p. 171-187.
78. Schilling, T.F. and P. Le Pabic, *Fishing for the signals that pattern the face.* J Biol, 2009. **8**(11): p. 101.
79. Lieschke, G.J. and P.D. Currie, *Animal models of human disease: zebrafish swim into view.* Nat Rev Genet, 2007. **8**(5): p. 353-67.
80. Kimmel, C.B., *Genetics and early development of zebrafish.* Trends Genet, 1989. **5**(8): p. 283-8.
81. Reid, J.S., et al., *Ectodermal dysplasia: case with an unusual combination of dental features.* J Oral Med, 1986. **41**(4): p. 259-61.
82. Duncan, K.M., et al., *Zebrafish models of orofacial clefts.* Dev Dyn, 2017. **246**(11): p. 897-914.
83. Botti, E., et al., *Developmental factor IRF6 exhibits tumor suppressor activity in squamous cell carcinomas.* Proc Natl Acad Sci U S A, 2011. **108**(33): p. 13710-5.
84. Kousa, Y.A., D. Moussa, and B.C. Schutte, *IRF6 expression in basal epithelium partially rescues Irf6 knockout mice.* Dev Dyn, 2017. **246**(9): p. 670-681.

85. Bebee, T.W., et al., *Ablation of the epithelial-specific splicing factor Esrp1 results in ureteric branching defects and reduced nephron number.* Dev Dyn, 2016. **245**(10): p. 991-1000.
86. Fakhouri, W.D., et al., *Intercellular Genetic Interaction Between Irf6 and Twist1 during Craniofacial Development.* Sci Rep, 2017. **7**(1): p. 7129.
87. Goudy, S., et al., *Cell-autonomous and non-cell-autonomous roles for IRF6 during development of the tongue.* PLoS One, 2013. **8**(2): p. e56270.
88. Sabel, J.L., et al., *Maternal Interferon Regulatory Factor 6 is required for the differentiation of primary superficial epithelia in Danio and Xenopus embryos.* Dev Biol, 2009. **325**(1): p. 249-62.
89. Motta-Mena, L.B., et al., *An optogenetic gene expression system with rapid activation and deactivation kinetics.* Nat Chem Biol, 2014. **10**(3): p. 196-202.
90. Rossi, A., et al., *Genetic compensation induced by deleterious mutations but not gene knockdowns.* Nature, 2015. **524**(7564): p. 230-3.
91. Khan, A., et al., *JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework.* Nucleic Acids Res, 2018. **46**(D1): p. D1284.
92. Li, H., et al., *The molecular anatomy of mammalian upper lip and primary palate fusion at single cell resolution.* Development, 2019. **146**(12).
93. Liu, K., et al., *Expanding the CRISPR Toolbox in Zebrafish for Studying Development and Disease.* Front Cell Dev Biol, 2019. **7**: p. 13.
94. Bishop, M.R., et al., *Genome-wide Enrichment of De Novo Coding Mutations in Orofacial Cleft Trios.* Am J Hum Genet, 2020. **107**(1): p. 124-136.
95. Bureau, A., et al., *Whole exome sequencing of distant relatives in multiplex families implicates rare variants in candidate genes for oral clefts.* Genetics, 2014. **197**(3): p. 1039-44.
96. Carlson, J.C., et al., *A systematic genetic analysis and visualization of phenotypic heterogeneity among orofacial cleft GWAS signals.* Genet Epidemiol, 2019. **43**(6): p. 704-716.
97. Carlson, J.C., et al., *Identifying Genetic Sources of Phenotypic Heterogeneity in Orofacial Clefts by Targeted Sequencing.* Birth Defects Res, 2017. **109**(13): p. 1030-1038.

98. Leslie, E.J., et al., *Association studies of low-frequency coding variants in nonsyndromic cleft lip with or without cleft palate.* Am J Med Genet A, 2017. **173**(6): p. 1531-1538.
99. Mukhopadhyay, N., et al., *Whole genome sequencing of orofacial cleft trios from the Gabriella Miller Kids First Pediatric Research Consortium identifies a new locus on chromosome 21.* Hum Genet, 2020. **139**(2): p. 215-226.
100. Ray, D., et al., *Pleiotropy method reveals genetic overlap between orofacial clefts at multiple novel loci from GWAS of multi-ethnic trios.* PLoS Genet, 2021. **17**(7): p. e1009584.
101. Leslie, E.J. and M.L. Marazita, *Genetics of cleft lip and cleft palate.* Am J Med Genet C Semin Med Genet, 2013. **163C**(4): p. 246-58.
102. Alharatani, R., et al., *Novel truncating mutations in CTNND1 cause a dominant craniofacial and cardiac syndrome.* Hum Mol Genet, 2020. **29**(11): p. 1900-1921.
103. Bean, L.J.H., et al., *Diagnostic gene sequencing panels: from design to report-a technical standard of the American College of Medical Genetics and Genomics (ACMG).* Genet Med, 2020. **22**(3): p. 453-461.
104. Richards, S., et al., *Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology.* Genet Med, 2015. **17**(5): p. 405-24.
105. Adzhubei, I.A., et al., *A method and server for predicting damaging missense mutations.* Nat Methods, 2010. **7**(4): p. 248-9.
106. Choi, Y., et al., *Predicting the functional effect of amino acid substitutions and indels.* PLoS One, 2012. **7**(10): p. e46688.
107. Kumar, P., S. Henikoff, and P.C. Ng, *Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm.* Nat Protoc, 2009. **4**(7): p. 1073-81.
108. Reva, B., Y. Antipin, and C. Sander, *Predicting the functional impact of protein mutations: application to cancer genomics.* Nucleic Acids Res, 2011. **39**(17): p. e118.
109. Leslie, E.J., *Genetic models and approaches to study orofacial clefts.* Oral Dis, 2021.
110. Gowans, L.J., et al., *Association Studies and Direct DNA Sequencing Implicate Genetic Susceptibility Loci in the Etiology of Nonsyndromic Orofacial Clefts in Sub-Saharan African Populations.* J Dent Res, 2016. **95**(11): p. 1245-56.

111. Leslie, E.J., et al., *Identification of functional variants for cleft lip with or without cleft palate in or near PAX7, FGFR2, and NOG by targeted sequencing of GWAS loci.* Am J Hum Genet, 2015. **96**(3): p. 397-411.
112. Thusberg, J., A. Olatubosun, and M. Vihinen, *Performance of mutation pathogenicity prediction methods on missense variants.* Hum Mutat, 2011. **32**(4): p. 358-68.
113. Oliver, J.D., et al., *Molecular Diagnostics and In Utero Therapeutics for Orofacial Clefts.* J Dent Res, 2020. **99**(11): p. 1221-1227.
114. Frazer, J., et al., *Disease variant prediction with deep generative models of evolutionary data.* Nature, 2021. **599**(7883): p. 91-95.
115. Griesemer, D., et al., *Genome-wide functional screen of 3'UTR variants uncovers causal variants for human disease and evolution.* Cell, 2021. **184**(20): p. 5247-5260 e19.
116. Findlay, G.M., et al., *Accurate classification of BRCA1 variants with saturation genome editing.* Nature, 2018. **562**(7726): p. 217-222.
117. Glazer, A.M., et al., *High-Throughput Reclassification of SCN5A Variants.* Am J Hum Genet, 2020. **107**(1): p. 111-123.
118. Giacomelli, A.O., et al., *Mutational processes shape the landscape of TP53 mutations in human cancer.* Nat Genet, 2018. **50**(10): p. 1381-1387.
119. Mighell, T.L., S. Evans-Dutson, and B.J. O'Roak, *A Saturation Mutagenesis Approach to Understanding PTEN Lipid Phosphatase Activity and Genotype-Phenotype Relationships.* Am J Hum Genet, 2018. **102**(5): p. 943-955.
120. Jia, X., et al., *Massively parallel functional testing of MSH2 missense variants conferring Lynch syndrome risk.* Am J Hum Genet, 2021. **108**(1): p. 163-175.
121. Carroll, S.H., et al., *An Irf6-Esrp1/2 regulatory axis controls midface morphogenesis in vertebrates.* Development, 2020. **147**(24).
122. Wada, N., et al., *Hedgehog signaling is required for cranial neural crest morphogenesis and chondrogenesis at the midline in the zebrafish skull.* Development, 2005. **132**(17): p. 3977-88.
123. Warzecha, C.C., et al., *The epithelial splicing factors ESRP1 and ESRP2 positively and negatively regulate diverse types of alternative splicing events.* RNA Biol, 2009. **6**(5): p. 546-62.
124. Itoh, M., et al., *The exon 38-containing ARHGEF11 splice isoform is differentially expressed and is required for migration and growth in invasive breast cancer cells.* Oncotarget, 2017. **8**(54): p. 92157-92170.

125. Shapiro, I.M., et al., *An EMT-driven alternative splicing program occurs in human breast cancer and modulates cellular phenotype.* PLoS Genet, 2011. **7**(8): p. e1002218.
126. Davis, M.A., R.C. Ireton, and A.B. Reynolds, *A core function for p120-catenin in cadherin turnover.* J Cell Biol, 2003. **163**(3): p. 525-34.
127. Fukumoto, Y., et al., *The regulatory or phosphorylation domain of p120 catenin controls E-cadherin dynamics at the plasma membrane.* Exp Cell Res, 2008. **314**(1): p. 52-67.
128. Ireton, R.C., et al., *A novel role for p120 catenin in E-cadherin function.* J Cell Biol, 2002. **159**(3): p. 465-76.
129. Ishiyama, N., et al., *Dynamic and static interactions between p120 catenin and E-cadherin regulate the stability of cell-cell adhesion.* Cell, 2010. **141**(1): p. 117-28.
130. Reynolds, A.B., et al., *Identification of a new catenin: the tyrosine kinase substrate p120cas associates with E-cadherin complexes.* Mol Cell Biol, 1994. **14**(12): p. 8333-42.
131. Stairs, D.B., et al., *Deletion of p120-catenin results in a tumor microenvironment with inflammation and cancer that establishes it as a tumor suppressor gene.* Cancer Cell, 2011. **19**(4): p. 470-83.
132. Schackmann, R.C., et al., *p120-catenin in cancer - mechanisms, models and opportunities for intervention.* J Cell Sci, 2013. **126**(Pt 16): p. 3515-25.
133. Reynolds, A.B. and A. Roczniak-Ferguson, *Emerging roles for p120-catenin in cell adhesion and cancer.* Oncogene, 2004. **23**(48): p. 7947-56.
134. Kievit, A., et al., *Variants in members of the cadherin-catenin complex, CDH1 and CTNND1, cause blepharocheilodontic syndrome.* Eur J Hum Genet, 2018. **26**(2): p. 210-219.
135. Griffin, C. and J.P. Saint-Jeannet, *Spliceosomopathies: Diseases and mechanisms.* Dev Dyn, 2020. **249**(9): p. 1038-1046.
136. Dennison, B.J.C., et al., *Srsf3 mediates alternative RNA splicing downstream of PDGFRalpha signaling in the facial mesenchyme.* Development, 2021. **148**(14).
137. Timberlake, A.T., et al., *Haploinsufficiency of SF3B2 causes craniofacial microsomia.* Nat Commun, 2021. **12**(1): p. 4680.
138. Ghoumid, J., et al., *Blepharocheilodontic syndrome is a CDH1 pathway-related disorder due to mutations in CDH1 and CTNND1.* Genet Med, 2017. **19**(9): p. 1013-1021.

139. Del Valle-Perez, B., et al., *Wnt controls the transcriptional activity of Kaiso through CK1epsilon-dependent phosphorylation of p120-catenin.* J Cell Sci, 2011. **124**(Pt 13): p. 2298-309.
140. Park, J.I., et al., *Kaiso/p120-catenin and TCF/beta-catenin complexes coordinately regulate canonical Wnt gene targets.* Dev Cell, 2005. **8**(6): p. 843-54.
141. Letra, A., et al., *Interaction between IRF6 and TGFA genes contribute to the risk of nonsyndromic cleft lip/palate.* PLoS One, 2012. **7**(9): p. e45441.
142. Revest, J.M., et al., *Fibroblast growth factor receptor 2-IIIb acts upstream of Shh and Fgf4 and is required for limb bud maintenance but not for the induction of Fgf8, Fgf10, Msx1, or Bmp4.* Dev Biol, 2001. **231**(1): p. 47-62.
143. Cox, T.C., *Taking it to the max: the genetic and developmental mechanisms coordinating midfacial morphogenesis and dysmorphology.* Clin Genet, 2004. **65**(3): p. 163-76.
144. Tamarin, A. and A. Boyde, *Facial and visceral arch development in the mouse embryo: a study by scanning electron microscopy.* J Anat, 1977. **124**(Pt 3): p. 563-80.
145. DeLaurier, A., et al., *ZebraShare: a new venue for rapid dissemination of zebrafish mutant data.* PeerJ, 2021. **9**: p. e11007.
146. Liu, W.S., et al., *Down-regulation of dendritic spine and glutamic acid decarboxylase 67 expressions in the reelin haploinsufficient heterozygous reeler mouse.* Proc Natl Acad Sci U S A, 2001. **98**(6): p. 3477-82.
147. Sander, J.D., et al., *Zinc Finger Targeter (ZiFiT): an engineered zinc finger/target site design tool.* Nucleic Acids Res, 2007. **35**(Web Server issue): p. W599-605.
148. Ran, F.A., et al., *Genome engineering using the CRISPR-Cas9 system.* Nat Protoc, 2013. **8**(11): p. 2281-2308.
149. Montague, T.G., et al., *CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing.* Nucleic Acids Res, 2014. **42**(Web Server issue): p. W401-7.
150. Gagnon, J.A., et al., *Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs.* PLoS One, 2014. **9**(5): p. e98186.
151. Meeker, N.D., et al., *Method for isolation of PCR-ready genomic DNA from zebrafish tissues.* Biotechniques, 2007. **43**(5): p. 610, 612, 614.
152. Thisse, C. and B. Thisse, *High-resolution in situ hybridization to whole-mount zebrafish embryos.* Nat Protoc, 2008. **3**(1): p. 59-69.

153. Walker, M.B. and C.B. Kimmel, *A two-color acid-free cartilage and bone stain for zebrafish larvae.* Biotech Histochem, 2007. **82**(1): p. 23-8.