Title

Cell Cycle Arrest of the Midface Epithelium Promotes Face Morphogenesis in Mice and Humans and is Disrupted in Craniofacial Disorder Models

Author

Chacon, Brandon Hugh

Publication Date

2024

Cell Cycle Arrest of the Midface Epithelium Promotes Face Morphogenesis in Mice and Humans and is Disrupted in Craniofacial Disorder Models

by
Brandon Chacon

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Developmental and Stem Cell Biology

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
Daniel Lim
FC7A2B6D45EF428...
Daniel Lim
Chair

DocuSigned by:
Ralph Marcucio
[illegible]495...
Ralph Marcucio

DocuSigned by:
Licia Selleri
[illegible]4CB...
Licia Selleri

DocuSigned by:
Daniel Wagner
683F6055047E46A...
Daniel Wagner

Committee Members

By

Brandon Hugh Chacon

This dissertation is dedicated to all affected by craniofacial disease and the hope for improved medicine for all.

Acknowledgements

I wish to acknowledge all in my life who helped get me to where I am today. These acknowledgements go to the following:

Thank you to the long line of women in science that inspired and motivated me all through my life: To Jaya Ayer, for giving me Darwin and a new world perspective. To Wendy Washington who let my voice be acknowledged. To Kathleen Sigafoos for inspiring scholarliness. To Cecily Kipe for a masterful preparation and foundation in biology yet unsurpassed. To Kathy Foltz, a passionate developmental biologist, and the essence of a mentor. To Denise Montell, whose developmental biology lectures completed the puzzle. To Carmen Domingo, who gave me a fighting chance. To Diana Laird, the symbol of professionalism and a friend above all else, thank you for believing in me. To Marta Losa Llabata, without you, I'd have given up all too soon. You are a light through the tunnel and this work is equally yours. Thank you for your being my mentors.

To all who helped me chase a scientific dream at UCSB, SFSU, and UCI, especially Alex Boiko, who led a volunteer to a microscope, and the rest is history. To all at UCSF who made my progression to candidacy and graduation a success, including my thesis committee members, Demian Sainz, and Todd Nystul, I thank you from the bottom of my heart. To my DSCB friends, Thi, Nick, Matt, Emily, Jean Luke, Michael, and Gugene, what a ride it has been. May you all find success and happiness in life. To my senior DSCB mentors Karishma, Ryan, Bryan, and Arpana, but especially Nate Meyer whose help through the qualification exam led to success, thank you.

To members of the Floor, Goga, L'Etoille, Tward, Barber, Laird, and Panagiotakos labs, who were always willing to lend an ear, a hand, or a reagent, you were excellent friends and colleagues. To our collaborators in the Adameyko and Leslie labs, notably Louis and Kelsey, thank you for your dedication and hard work. To the Selleri laboratory, you crazy people were a delight to have as friends, co-workers, and fellow scientists. Viviana and Maurizio, you have helped me navigate the toughest of times and helped me more times than I can count. Tony, I could not ask for a greater co-pilot to go through the graduate school experience with. May your ambition never dim, and your experiments always succeed. Thank you to all the Selleri lab members past and present, Nuo, Jaiden, Karissa, Danny, Tatiana, Rocío, and Gaga. You all rock!

To my advisor and mentor Licia Selleri, thank you for accepting me into your lab and teaching the world of craniofacial biology to me. You are often a voice in a world that expects submissiveness, and all too often the only correct one. If you love what you do, why ever stop?

I would like to thank my friend and roommate of so many years, Eduard Puchin. You kept me grounded and helped maintain a home that was my refuge. Thank you for anchoring my daily life outside of science and for all the laughs, improv, and endless patience with me.

To my friend in science and in life, Anup Sarakki, whatever you put your mind to, you can achieve. Thank you for the love and for being my most trustworthy advisor. If I knew what I was in for when we hired you at UCI, I'd do it all over again.

To my partner, Joshua Pardo, the beacon of sunshine that has cleared the deepest of SF fogs with the warmest of smiles, even across years and miles of long-distance love. May our adventure in life and love never end. Thank you for everything. Your love got me to this point and will get me through any obstacle. I'm coming home soon.

To all of my family, but especially to all my grandparents, you were beacons of bravery and boldness that I only can aspire to. To my grandpa Paul Chacón, your efforts to make a better life for your family paid off. To my grandpa Victoriano Valdez, when you declared that the curious boy you saw before you would be a doctor one day, you gave inertia to a fulfilled prophesy. I miss you terribly and thank you for seeing me.

To my brother Adam, the one I wish I had forever to guide but who has guided me in more ways than he knows, I love you and thank you for everything. It has been a delight to see you grow into yourself. I am always proud of your direction in life wherever it may lead.

To my dad. You have the heart of an angel, the patience of a saint, and a work-ethic beyond description. Not only have you taught me the importance of loving what you do in life, but you have dedicated a life to serving others and generously sacrificing for the greater good. Somehow, even after that, you have given your life to your family. You are a success in every standard and metric I can imagine but especially as my father. What fantastic joys and memories in life I can rely on thanks to you. My accomplishments are yours to have as much credit as you deserve. You make me so incredibly proud to be your son, and I owe so much to you. Thank you for being my hero and the superhero that has never failed our family.

To my mom. Where do I even begin? The home, life, and love you helped cultivate for me is what brought me here today. In my life, no woman and no person has done more to keep me motivated, inspire excellence, and establish a love for knowledge, inquiry, and education than you. So many of my invaluable gifts are thanks to you and the sacrifices you have made. Whether through nature or nurture, my literacy, my oratory skills, my intolerance of nonsense, and my persistence are all your traits you shared in abundance. You are a fighter, a role model, my rock, and my predecessor in all ways good. You are and always have been, my best friend and there is nobody more badass than you. Thank you for everything.

CONTRIBUTIONS

This thesis work was performed under the supervision of Dr. Licia Selleri, MD, PhD. Dissertation committee members Drs. Daniel Lim MD, PhD, Ralph Marcuccio, PhD, and Daniel Wagner, PhD each contributed invaluable input and expertise in evaluating and guiding this work.

Chapter 2 is a manuscript in preparation which I wrote with guidance from Licia Selleri MD, PhD, and Marta Losa Llabata, PhD. Our collaborators Kelsey Robinson, Elizabeth Leslie, Louis Faure, and Igor Adameyko PhD each contributed to the formation of the manuscript. Experiments performed were a joint effort between the Selleri Lab at UCSF: myself, Marta Losa Llabata, PhD, Tiange (Tony) Qu, and Rocío Hernández Martínez, PhD; the Adameyko Lab of the Karolinska Institute (Stockholm, Sweden): Louis Faure, PhD and Igor Adameyko, PhD; the MRC-Wellcome Trust Human Developmental Biology Resource at the Institute of Genetic Medicine (Newcastle, UK): Aragorn Jones, PhD and Steven Lisgo, PhD; and the lab of Elizabeth Leslie, PhD at Emory University (Atlanta, Georgia): Kelsey Robinson and Elizabeth Leslie, PhD. Figures generated were efforts between all listed authors. This manuscript describes research aimed at characterizing isolated midface epithelium during fusion of the facial prominences at the lambdoidal junction in both mouse and human development. It describes a state of cell cycle arrest for a population of cells located at the site of fusion and investigates mechanisms relating to the appearance of cleft lip when cell cycle arrest of the cephalic epithelium is affected.

Chapter 3 is preliminary work relating to the characterization of a conditional deletion of *Bmpr1a* in the ectoderm of mouse. This novel deletion sought to investigate the occurrence of cleft lip and other gross anatomical defects resulting from the loss of *Bmpr1a*. I characterize the anatomy of these mutant mice and examine histological evidence for gene expression changes in the midfacial prominences during lambdoidal junction fusion.

“The cosmos is within us. We are made of star-stuff. We are a way for the universe to know itself.”

— Carl Sagan

“Beauty fades, dumb is forever.”

– Judy Scheindlin

“Life is like a box of chocolates. You never know what you’ll get.”

-Forrest Gump

Cell Cycle Arrest of the Midface Epithelium Promotes Face Morphogenesis in Mice and Humans and is Disrupted in Craniofacial Disorder Models

Brandon Hugh Chacón

Abstract

To orchestrate the complex events of craniofacial morphogenesis requires a concerted effort between gene expression and tissue specific events. In the embryonic midface, facial prominences must execute a coordinated fusion between maxillary, medial and lateral nasal prominences at the lambdoidal junction, requiring breakdown and removal of epithelium to permit coalescence of underlying mesenchyme. When this process is disrupted by genetic mutations, cleft lip, the most common craniofacial birth defect in humans, is created. Key to understanding cleft lip pathogenesis is the reliance on mouse models. Models of cleft lip and studies of the role of epithelium are rare in comparison to cleft palate and research on underlying mesenchyme. The epithelium possesses discrete roles that facilitate fusion at the lambdoidal junction, including apoptosis and epithelial to mesenchymal transitions. To better understand its heterogeneity, we sought to enrich for the midface epithelium and isolate it from mesenchyme to gain fuller resolution of the lambdoidal junction transcriptome. This dissertation contains the product of that endeavor to focus on the role of the epithelium alone both in normal conditions and in cleft lip pathogenesis. In doing so, an atlas of lambdoidal junction epithelium at timepoints before, during and after fusion of the midface was created and validated. Through these methods, a deep analysis of each cluster resulted in the localization of a cluster undergoing cell cycle arrest and located at the fusion site of each prominence. A series of experiments were conducted which demonstrated that cell cycle arrest is characteristic of this population that is altered in cleft lip models. Additional research yielded ties to normal midface morphogenesis as well as novel orofacial cleft risk candidates, including *Zfhx3* a gene with ties to cell cycle arrest. This thesis also expands upon the phenotypic characterization of a novel mouse cross which conditionally removes *Bmpr1a* from the ectoderm of the embryo, creating a model of cleft lip and palate. This work collectively produces knowledge of the midface epithelium at a scale of resolution not seen before as well as describes in detail how cell cycle arrest is a characteristic of primary palate fusion.

TABLE OF CONTENTS

TABLE OF FIGURES

LIST OF ABBREVIATIONS

λ -	lambdoidal
AER -	apical ectodermal ridge
BA -	branchial arch
BAMBI -	BMP and activin membrane bound inhibitor
BMP -	bone morphogenic protein
BMPR1A -	bone morphogenic protein receptor type 1A
BrdU -	bromodeoxyuridine
CDKN1A -	cyclin dependent kinase inhibitor 1A
ChIPSeq -	chromatin immunoprecipitation sequencing
CL -	cleft lip
CL/P -	cleft lip with or without cleft palate
CLO –	cleft lip only
CP -	cleft palate
CNC –	cranial neural crest
Cre -	causes recombination
DEG -	differentially expressed gene
E -	embryonic day
EMT –	epithelial to mesenchymal transition
EPCAM -	epithelial cell adhesion molecule
FACS -	fluorescence activated cell sorting
FISH -	fluorescent *in-situ* hybridization
FOX -	forkhead box protein

GADD45G -	growth arrest and DNA damage inducible protein gamma
ISH -	*in-situ* hybridization
LNP -	lateral nasal prominence
MdP -	mandibular prominence
MNP -	medial nasal prominence
MxP -	maxillary prominence
NCC -	neural crest cell
OFC -	orofacial cleft
PBX -	pre-B-cell leukemia homeobox 1
scRNAseq -	single cell ribonucleic acid sequencing
TGFβ -	transforming growth factor beta
TP63 -	transformation related protein 63
TUNEL -	terminal deoxynucleotidyl transferase dUTP nick-end labeling
UMAP -	uniform manifold approximation and projection
WISH -	whole mount *in-situ* hybridization
WNT -	wingless/integrated
ZFHX3 -	zinc-finger homeobox 3
ZL -	zippering lambda epithelium

CHAPTER 1: INTRODUCTION

The head is the most complicated part of the vertebrate body, home to not only our brain and sensory organs, but also our identity in the form of a recognizable face. Facial development begins early in embryonic development. This allows time for the complex patterns of tissue morphogenesis to sculpt a variation that is uniquely me or you. Remarkably, the infinite variation of faces among humans as well as across vertebrate taxa employs many conserved cellular, developmental, and genetic mechanisms around which this variation emerges. This conservation allows researchers to better understand our own facial morphogenesis while using laboratory models such as the mouse. The laboratory mouse, *Mus musculus*, allows for genetic manipulation which can mimic disorders of craniofacial development, including the most common craniofacial birth defect in humans: cleft lip and palate.

1.1 Midface Development in Mouse and Human Embryos and Cleft Lip Pathogenesis

Among the variety of orofacial clefts, cleft lip with or without cleft palate is the product of midface morphogenesis gone awry. The midface is the center of the face, defined as the region between the eyes and the upper lip, including the nose, nostrils, philtrum, primary palate, and the maxillary component of the upper jaw. It is in this region that a structural developmental disorder can arise, cleft lip with or without cleft palate (CL/P), beginning in early embryonic development. Cleft lip (CL) and cleft palate (CP) are related yet distinct in spatial-temporal development. CL affects the primary palate, which forms initially in development and consists in humans of structures anterior of the incisive foramen, including the alveolar arch, pre-maxilla, and upper lip, which form during the 4th to 5th week of gestation through week 7. Homologous structures develop in mice from stages E9.5 through E12.5 CP affects the secondary palate, which forms later in development and consists of hard and soft palate structures posterior of the incisive foramen that derive from the maxillary prominences at the 7th week of gestation in humans and E11.5 in mice.

In both mouse and humans, the emergence of two bulges in the frontonasal prominence around the olfactory placode heralds the formation of the lateral and medial nasal prominences (LNP and MNP respectively) between gestational day (E) E9.5 and E10.5 in mouse and weeks 4-5 of human embryonic development. In humans the LNP will give rise to the region of the alae and sides of the nose, and the MNP the midline of the face, nose, and philtrum. The first branchial arch (BA1), which formed at E9.5 in mouse (week 4 human), before the nasal placode expansion, has split to become the maxillary prominence (MxP) and the mandibular prominence (MdP), connected by a hinge of tissue which will later become the corners of the mouth. The maxillary prominence will contribute to both maxillae, the cheeks, and parts of the upper lip. When the lateral and medial nasal prominences swell and meet with the maxillary prominences on both sides of the face, they create a three-way juncture called the lambdoidal junction due to its resemblance to the Greek letter λ.

Within the span of one developmental stage, these prominences should merge, leaving no trace of the seam that once separated the three. Eventually, in the fully developed human face, only the edges of the philtrum leading anteriorly to the nostrils demarcate the once separate facial anlagen. When this fusion of the prominences fails, the prominences continue to grow with a gap between them that persists until birth. This gap, or cleft, can occur on one or both sides of the upper lip (unilateral or bilateral cleft lip) where, in humans, the philtral columns/ridges should be normally present. The gap(s) prevents a newborn from suckling properly without a proper mouth seal. In mice litters, pups with cleft lip, with or without cleft palate, typically cannot feed and die perinatally.

The prevalence of CL/P is high in the United States, where a recent analysis of 26% of all live births between 2010-2014 found that it ranked third in birth defect prevalence (10.00 in 10,000 live births), behind only Down syndrome (14.14 in 10,000), and clubfoot (16.87 in 10,000) (Mai et al., 2019). Of all craniofacial birth defects, it is the most common worldwide with an incidence of 1 in 700 births (Watkins et al., 2014). Both genetic and environmental factors can influence the occurrence of CL/P, such as maternal smoking or other teratogenic exposures, however the majority (70%) of CL/P cases are non-syndromic, wherein no other anomalies are detected (Murthy & Bhaskar, 2009). Without surgical intervention, CL/P can lead to malnutrition, dental anomalies, and speech impediment, severe respiratory infections, as well as social and psychological impacts due to a highly visible physical disfigurement. Although medical intervention is often quickly performed in the form of surgery, continued orthodontics, speech therapy, and more are needed, at great cost to families and individuals (Rochlin et al., 2022). Even with these interventions, visible scars and speech difficulties cannot be fully corrected with current methods and socio-psychological impacts can remain (Lorot-Marchand et al., 2015). It is worthwhile for the field of medicine to be able to prenatally diagnose and one day intervene in the onset of orofacial clefts *in utero*. A developmental biology perspective would be of great aid in this effort, while expanding upon the foundations of midface development.

1.2 Historical Context of Facial Morphogenesis and Cleft Lip

Knowledge of both facial morphogenesis and of what can happen when it goes awry dates to ancient times. Attempts to explain why some babies were born with clefts in the upper lip range from evil spirits (Ancient Greece and Rome) to the superstition of jumping over a hare while pregnant (16th century Sweden)(Bhattacharya et al., 2009). Treatment of people with CL/P likewise ranged from the barbaric, such as leaving newborns to die or drown in rivers (largely justified by notable scholar Plato), to the sensible, such as in the documented successful surgical repair of CL in 3rd century BC China. The patient, a young Wey Young-Chi, placed the credit of his later career and military successes on his corrective surgery (Wong K. Chiminand Wu Lien-teh, 1932).

Embryological perspectives of lambdoidal junction fusion were established before the age of molecular biology. Wilhelm His Sr. provided early close studies of facial morphogenesis, observing early stages of development that include the changing shapes of the midface and branchial arches. His noted the formation of several (five) processes [prominences] of the face which when improperly fused, he believed contributed to orofacial clefts (Hopwood, 2000). In addition to contributing theory to cleft lip formation, Wilhelm His theorized that differential growth of regions in the embryo contributed to the formation of complex shapes and patterns he observed, a theory that predates knowledge of the molecular events needed to alter the proliferation rates of cells (Richardson & Keuck, 2022). His' observations and theory nevertheless pinpoint a tie between cellular proliferation rates and morphogenesis which, when coupled with modern molecular biology, begins to shed light on how structures as complex as the face can develop. Neural crest cells invade the prominences and contribute to the bulk of mesenchymal tissue during the prominence formation. It was shown through multiple studies that defects in cranial neural crest cells (CNCs) can manifest in orofacial clefts (OFCs). Some of these defects are due to proliferation defects in the mesenchyme, in which the volume of the prominences doesn't expand to create a proper lambdoidal junction in the first place. Proliferation of neural crest derived mesenchyme alone, as with His' theory, cannot fully explain the formation or not of OFCs. Further analysis of the mesenchyme and the tissue surrounding the prominences, the epithelium, unveils a suite of cellular behaviors with genetic regulatory networks (GRNs) underpinning their dynamics. Although His did not have knowledge of the role of DNA, mRNA, proteins, or how they come together to execute a cellular behavior, I believe he was prescient in discussing a basic mechanism of development that we now know requires as much genetic regulation as any other: cell cycle dynamics.

1.3 Cell Cycle Arrest in the Context of Embryonic Development

Cell cycle regulation is part of everyday growth and regulation for all forms of life. Vertebrate animal cell cycles have layers of redundancy to ensure that careful coordination of growth and avoidance of unchecked proliferation are the norm (Casimiro et al., 2012). In the developing embryo, cell growth and proliferation are more frequent than rest and quiescence, such as in the face, where the embryo must rely on proliferation to expand the volume of the underlying mesenchyme of the midfacial prominences. Specific moments of cell cycle arrest and quiescence are rarer in embryonic development but are needed to complete the body plan before birth (Brantley & Di Talia, 2021; O'Farrell, 2011). A growing understanding of the role of both cell cycle pausing and modulation, and the loss of cells through programmed cell death in normal embryonic development has emerged.

There are several instances of reported cell cycle dynamics at play that include slowed cell cycle progression in the shaping of the developing embryo. In the apical ectodermal ridge (AER) of the limb bud, a specialized area of epithelium known as the progress zone undergoes cell cycle arrest in the distal epithelium, a trait believed to involve cellular senescence (Kumari & Jat, 2021). Similarly, epithelial cells in the developing tooth bud undergo a G1 arrest in the enamel knots, both primary and secondary (in the case of molars) (Mogollón et al., 2021). Disruption of this arrest permits continued proliferation of these regions that interferes with their role as signaling centers in their respective tissues and leads to malformation of the limb bud and tooth respectively. In both examples a relatively thin layer of epithelium performs the role of cell cycle modulation in the epithelium as opposed to a more abundant neighboring mesenchyme. Both the AER and enamel knot exert an influence on their surrounding tissues that includes epithelial-mesenchymal cross talk and a temporary status as a signaling node that ends in apoptosis for the arrested epithelial cells. These roles tend to be associated with "temporary" structures in embryonic development in a way that is evolutionarily conserved (Villiard et al., 2017). What role cell cycle arrest has been known to play in development in the primary palate is not understood and has not been noted in published literature aside from evidence of lowered proliferation in the epithelium at the fusion site (Jin et al., 2012; Song et al., 2009). Greater evidence is needed to ascertain this phenomenon, its ties to cleft lip pathogenesis, and any molecular mechanisms responsible for its onset.

1.4 Epithelial Specific Roles in Midface Development

The role of the epithelium in the midface is one that has been highlighted recently. The epithelium represents approximately 10% of the cells of the midface at E10.5 of mouse development. This layer is derived from the surface cephalic ectoderm and creates a single cell thick sheath around the prominences of the head. The role of epithelium is one that is both independent and context-dependent, in which signaling to the underlying mesenchyme is critical. The epithelium of both forebrain and facial ectoderm is required to properly direct CNC growth in the developing frontonasal prominence (FNP) (D. Hu & Helms, 1999) and to ensure specification of CNC fate (Kimmel et al., 2001). In contrast, mesenchymal signals can also influence the overlying epithelium as in the case of *Meis2* loss in the mandibular CNC cells abolishing *Shh* expression in the oropharyngeal epithelium and causing improper patterning of BA1 (Fabik et al., 2020). Though the field has largely focused on the role of the more abundant mesenchyme, both independently and through its mesenchymal interactions, the epithelium possesses its own suite of cellular behaviors that shape both midface fusion and the occurrence of cleft lip.

Even without considering genetics or subcellular mechanisms, close analysis of the fusion site using electron microscopy reveals changes in the outer layers of the midfacial prominences as fusion occurs in mouse at E10.5. Filopodial extensions of the epithelium as well as cellular blebbing indicate an active process taking place between epithelium on either side of the fusion seams, rather than one passively occurring due to the pressure of prominence expansion or other mechanical means. This blebbing and formation of epithelial bridges in the lambdoidal junction epithelia is found to be evolutionarily conserved in human face development during fusion (Hinrichsen, 1985).

Behavior of the epithelium during primary palate fusion is complex but has the ultimate goal of being removed from the distal tips of the prominences to make way for invading mesenchyme to bridge the gap and unite the tissues. Various mechanisms of epithelium removal are both known and proposed including apoptosis, epithelial to mesenchymal transition (EMT), collective cell migration, and cell migration (Ferretti et al., 2011; Losa et al., 2018; Teng et al., 2022). Some of these are proposed based on evidence from related yet distinct mechanisms of epithelial removal during secondary palate formation.

Specific genetic disturbances in the epithelium as opposed to the mesenchyme promote the creation of cleft lip. As highlighted by *Pbx1* conditional loss in the mesenchyme using *Wnt1-Cre* on a *Pbx2* deficient background creates a cleft palate, whereas use of *Crect-Cre* removes *Pbx1* in the surface cephalic epithelium and produces cleft lip along with cleft palate (Welsh et al., 2018). The loss of *Pbx* genes is accompanied by defective epithelial cell behaviors at the fusion site, normally needed for its removal, with both apoptosis and EMT programs becoming abrogated from the loss of *Pbx* upstream of two parallel pathways (Losa et al., 2018). In the case of apoptosis, cells at the fusion site that normally would bleb at E10.5 and begin to self-destruct would instead retain a flattened morphology (Ferretti et al., 2011) and fail to undergo programmed cell death. These cells could begin to form adhesions between each other, blocking the progression of mesenchyme invasion. EMT was characterized as a normal part of epithelial cell removal at the prominence fusion site in mouse embryos at E11.5 (Losa et al., 2018). *Pbx* genes, whose expression in the midface is more abundant in the epithelium was shown to be a critical upstream regulator of apoptosis through Wnt3/9b*,* establishing a PBX-WNT-TP63-IRF6 axis which has been linked to orofacial cleft risk in humans (Maili et al., 2020). EMT regulation, meanwhile, was seen to be mediated by *Pbx* genes through the downstream mediators SNAI1 and SMAD4. *Tgfβ,* a known OFC risk gene in its own right, has a role upstream of *Pbx1* and thus also upstream of EMT and apoptosis pathways linked to *Pbx* loss, suggesting further that *Pbx* links a variety of developmental signaling pathways enriched at the lambdoidal junction.

1.5 Mouse Models of Cleft Lip

Orofacial clefts are categorized in epidemiology by the presence of cleft palate, cleft lip, or the combination of both. In one study of European and US populations (OMIM database 2000), unilateral CL with CP was ranked highest at 30-35% of cases, CL on its own was 20%, CP 25%(Mossey & Modell, 2012). Of all oral clefts, 15% were deemed syndromic. Additional classifications include breakdowns of unilateral versus bilateral occurrences. For CLP the majority were unilateral at 80%, and 20% were bilateral. The rarity of bilateral clefts was echoed in another report finding an occurrence of CL (10.3%) and CLP (30.2%) cases (IPDTOC 2011). The relative severity of CP versus CL, failures to distinguish between the two in reporting, and its greater ties to associated defects may in turn influence the lack of research dedicated solely to CL, and abundance of mouse models focused on the secondary palate. Nevertheless, the diversity of orofacial clefts and rates of occurrence is matched by the varied types of CL/P mouse models.

There are numerous mouse models of orofacial clefts. Some exhibit cleft lip, others cleft palate only (CPO), and some both with varying degrees of penetrance. The vast range of genetic models highlights the diversity and complexity of molecular and cellular events needed to perform fusion of the facial prominences. Distinguishing between genetic models that affect the secondary palate and the primary (lip) can be difficult, especially when non-specific tissue deletions, as in some *Cre* mouse lines, is involved. The number of mouse models with CL/P is low in comparison to CPO. Likewise, the effects of a constitutive null allele can produce a more severe phenotype than a tissue specific one, including to the point of early embryonic lethality, precluding study of later stage morphogenesis in the face.

In the early 20th century, a genetic basis for CL in mice was found in the newly established A strain mouse lines. These mice and the subsequent A/J lines were found to be susceptible to spontaneous CL/P (Reed & Snell, 1931) Similarly, the development of the CL/Fr line through repeated inbreeding and selection for the CL phenotype resulted in a 26% frequency of CL in viable mouse pups compared to the 14% in A/J (Bornstein et al., 1970). In these predisposed mice, hypoxia and hyperoxia during pregnancy was found to influence the incidence rate of CL/P, highlighting an interplay between environmental effects and genetic background (Millicovsky & Johnston, 1981). Future studies reexamined the A-line strain with highest incidence of CL/P, the A/WySyn (20-30% incidence rate) mouse line. Alterations to the *clf1* (found to be a retrotransposon) locus were found responsible for epigenetic dysregulation of nearby *Wnt9b* (Green et al., 2019; Juriloff et al., 2005). In the case of CL/Fr, mRNA profiling of palatal shelves indicated alterations in several genes related to cell cycle and proliferation (*Ywhab, Nek2, Tacc1*)(Sasaki et al., 2014). Through artificial selection both the A/J and CL/Fr lines attained relatively high rates of CL/P, however in contrast to single or multiples gene targeted knockout mice, the etiopathogenesis of these models is complicated by the incomplete penetrance and ties to environmental influence.

Currently, genetic mouse models that exhibit CL/P number 20 in total. In contrast, over 600 mouse lines are reported to possess a cleft palate phenotype (JAX informatics) (Gritli-Linde, 2008; S. K. Lee et al., 2020). Single gene disturbance genetic mouse models that include a CL phenotype largely were created in the early 2000's, including *Bmp4, Bmpr1a, Wnt9b, Mkx, Lrp6, Sox11, Tgfbr1,Trp63, Tfap2a, Folr1, Spry2, Sumol,* and *Tbx10* (Alkuraya et al., 2006; Bush et al., 2004; S. G. Gong et al., 2000; S.-G. Gong & Eulenberg, 2001; Y. X. Li et al., 2007; Liu et al., 2005; Nottoli et al., 1998; Sock et al., 2004; Song et al., 2009; Thomason et al., 2008; Welsh et al., 2007). Only the recent addition of *Esrp* conditional and constitutive null mice has supplemented this list (S. K. Lee et al., 2020). Compound deletions of multiple genes also expands the list of CL mouse models further including *Boc* and *Cdo* (W. Zhang et al., 2011), *Hhat* and *Ptch1* (Kurosaka et al., 2014), *Tfap2a* and *Fgf8* (Green et al., 2015), *Pbx1* and *Pbx2* or *Pbx3* (Ferretti et al., 2011), and *Pbx1* and *Wnt9b* (Ferretti et al., 2011).

This striking imbalance of mouse model types in the face of CL prevalence in human patients highlights an area in need of further research. Disentangling the complexity of face morphogenesis and of CL/P by necessity will require a need for more precise models in the future. Separating CL pathogenesis from CP may necessitate usage of tissue-specific deletions and the creation of more specific *Cre* recombinase mouse lines active in the face. A growing number of population studies in CL/P patients may uncover novel risk candidate genes which can provide guidance on mouse model creation. Altogether these improvements will aid in the phenocopying of human patients whose condition of CL does not align with CP studies.

1.6 Single Cell RNA Sequencing as a Method of Uncovering Heterogeneity in Facial Tissues

There have been several attempts to analyze the transcriptome of the developing face during midface prominence fusion beginning with early microarray analysis and continuing into the present with single cell RNA sequencing (scRNAseq). Since its inception scRNAseq provided a means to capture moments in transcriptomic time for cells too few to be gathered in bulk (Tang et al., 2009). Since then, advances in barcoding and cell capture efficiency have brought the technique into the mainstream (Islam et al., 2012; Klein et al., 2015; Zheng et al., 2017). Single cell analysis has greatly aided the unveiling of heterogeneity in cell populations once thought homogeneous (Soldatov et al., 2019) (Choi & Kim, 2019).

Microarrays provided an early means of investigating the heterogeneity of the developing midface. Sasaki et al 2014 employed CL/Fr mouse lines to investigate the gene expression differences between mutants with CL/P versus controls yet focused entirely on gene expression patterns in the secondary palate shelves at E13.5, excluding primary palate tissue (Sasaki et al., 2014). Brunskill et al 2014 focused on earlier timepoints E8.5-E10.5 of the developing mouse head (Brunskill et al., 2014). Using laser capture microdissections they sampled regions of mostly cranial NCC and neuroepithelial cells, with the exception of olfactory placode cells at E10.5 and created microarray profiles of each region. Additionally, by physically picking 11 single cells from regions of interest at E8.5 and using microarray, the study created an early semblance of modern single cell profiling for three tissue compartments, one of which was epidermal ectoderm.

Recent publications have taken the steps towards unveiling the heterogeneity of the midface through scRNAseq, with most taking the approach of whole dissection of the face or midfacial prominences prior to sequencing. Li et al 2019 investigated and validated the transcriptomics of E11.5 whole lambdoidal junctions creating a cell atlas of unsorted midfacial tissues (H. Li et al., 2019). Xu et al 2019 similarly performed an analysis of the whole mandibular prominences at E10.5 (Xu et al., 2019). Siewert et al 2023 performed a retrospective analysis of the Li et al 2019 dataset alongside whole embryo data to search for novel OFC candidate genes (Siewert et al., 2023). Recently, Cai et al 2023 performed scRNAseq on whole lambdoidal junctions on stages E10.5-E12.5 with some validation approaches, and an emerging study (Rajderkar et al., 2023) has taken the approach of combining scRNAseq and single nucleus ATACseq of E11.5 -E15.5 whole faces (Cai et al., 2023). Despite the sudden influx of studies using this technology to explore the midface transcriptome, no publications stand out in their ability to highlight novel roles of the epithelium in the development of the midface or to attribute specific biological roles to the identified cellular clusters. Critically, these studies done in wild-type tissues also lack comparative analysis of genetic mutants.

Nguyen et al 2024 was the most recent publication to investigate the transcriptomes of the developing midface and took the novel approach of enriching for a single cell type, CNCs as well as investigating changes in transcriptomes in a model of midfacial clefts, the *Tfap2a/2b* mutants (Nguyen et al., 2024). It should be noted that these mutants do not possess CL/P. This approach allowed for greater read depth, and thus more precise cluster and trajectory analysis.

scRNAseq has limitations in both the ability to capture all cell types and to perform an adequate read depth for large amounts of cell input. At the time of writing, there have been no attempts to perform scRNAseq on the epithelium of the midface alone, curtailing transcriptomic resolution of this distinct tissue. Likewise, no analysis of CL/P mutant mouse models using scRNAseq has been performed as of this date.

1.7 Known Cell Cycle Interactors and their Relation to the Midface

Changes in Developmental Signaling Pathways Leading to Altered Proliferation Rates

Gene expression analysis of the midface has yielded considerable data demonstrating key signaling pathways enriched in the developing prominences and the epithelium. Upstream interactors of the cell cycle have been known to be enriched at the distal tips of the prominences, including *Shh, Tgfb2, Bmp4,* and *Fgf8.* These genes have been investigated for their role in guiding proliferation, with most attention on the proliferation of the CNC-derived mesenchyme, despite the expression of these genes being restricted to the ectodermal epithelium of the prominences. Various effects on proliferation, innately tied to cell cycle progression, have been documented through disruptions of these pathways.

Exogenous FGF8 was found to be "substitute" for missing ectoderm and preserve the survival and proliferation of facial mesenchyme (Firnberg & Neubüser, 2002; Neubüser et al., 1997). *Shh* expression in the ectoderm of the frontonasal process and midfacial prominences is critical to mesenchymal proliferation and facial outgrowth (D. Hu & Helms, 1999; D. Hu & Marcucio, 2009; Jeong et al., 2004). Loss of *Shh* signaling by inactivation of *Smo* in the mesenchyme both increased CNC apoptosis and later inhibited cell proliferation, whereas increased *Shh* signaling yielded overgrowth phenotypes in the facial primordia. *Bmp4* was documented to have expression in the distal ectodermal epithelium of nearly all facial prominences including MNP, LNP, MxP, and MdP using chick (Francis-West et al., 1994). Similar to *Shh*, an overgrowth phenotype was found when ectopically expressing *Bmp4* (or *Bmp2*) while reduced outgrowth and proliferation was characteristic of NOGGIN-induced BMP signal reduction (Ashique et al., 2002; P. Wu et al., 2006). TGFβ family members and their associated receptors possess complex overlapping and distinct roles in communication between the mesenchyme and epithelium. Elimination of *Tgfbr2* in the mesenchyme appears to lead to increases in *Tgfb2* and *Tgfbr3* with proliferation alterations in the secondary palate, shortened faces, and CPO being the result (J. I. Iwata et al., 2012). *Tgfbr1/Alk5* mesenchymal specific deletions notably cause drastic effects to growth of the mandible, tongue, palate, snout, cranium, and show CL/P (Dudas et al., 2006; W. Y. Li et al., 2008; H. Zhao et al., 2008). Using *Nestin-Cre* driven removal of *Tgfbr1/Alk5* conditionally in the ectoderm (though *Nestin-Cre* is notoriously nonspecific and has forebrain and mesenchyme expression) not only produces either unilateral or bilateral CL but also leads to impaired palatal shelf growth and increased cell death (W. Y. Li et al., 2008). These discrepancies likewise also highlight the ongoing difficulties in defining the independent and interdependent relationship and cross-talk of mesenchyme and epithelium. Though study of these pathway has yielded an array of proliferation related defects, the mesenchyme has been the prevailing object of interest and the failure of primary palate fusion in some cases, to be deemed incidental. Growth of the epithelium, in contrast, has lacked study and analysis.

Cell Cycle Arrest Distinction from Decreased Proliferation

Some early evidence that cell cycle inhibitor proteins played a role in palatogenesis could be traced to the occasional incidence of cleft palate in patients with retinoblastoma (Bonaïti-Pellié et al., 1975). Retinoblastoma 1 (RB1) protein is a potent oncogene normally responsible for inhibition of cyclin-dependent kinases through sequestration of transcription factor E2F. Only through continuous maintenance of multiple phosphorylation sites on RB by Cyclin/Cyclin Dependent Kinase (CDK) complexes can release E2F and permit cell cycle progression .

Additional levels of control over CDK complexes are found through CDK-inhibitor proteins. Historically these inhibitors were grouped into either "inhibitors of CDK4" (INK4), p15, p16, p18, p19, or other "CDK-inhibitor proteins" (Kip/Cip), p21, p27, p57. Genetic knockout of *p57/Cdkn1c* in mouse causes perinatal lethality and results in a severe cleft secondary palate phenotype, inflated GI tracts with missing jejunal and ileal sections, muscular, and skeletal defects (Yan et al., 1997; P. Zhang et al., 1997). Further investigation demonstrated a link between *Lhx6/8* regulation and expression of *p57/Cdkn1c* and *Gadd45g* using *Lhx6/8* null mutant mice with CPO (Cesario et al 2015). *P57/Cdkn1c* and *p14/p16/Cdkn2a* were two genes noted to be expressed in the "basal cells at fusion zone cluster" using scRNAseq of E11.5 mouse embryos (H. Li et al., 2019)

Growth arrest and DNA-damage inducible 45 gamma (GADD45G) is one of three family members centrally located in cell cycle, growth, and repair sensing and response using diverse input and output mechanisms. Gadd45g can act as both a pro and anti-apoptotic mediator, inhibit the CDK complex at G2-M phase, activate MAPK/p38 signaling, induce p21-mediated G1 arrest, initiate DNA repair mechanisms, alter histone methylation, and is linked to cell senescence. Loss of *Gadd45g* in mice is not associated with any major phenotypes. Mice with null alleles of *Gadd45g* develop normally and healthily (Hoffmeyer et al., 2001). Curiously, genetic ablation of *Gadd45g* in domestic New Zealand rabbit (*Oryctolagus cuniculus*) leads to unilateral or bilateral CL and embryonically shows disordered proliferation, apoptosis, and EMT in the medial and lateral nasal prominences (Lu et al., 2019).

P21$^{Cip1/Waf1}$*/Cdkn1a* is largely dispensable in development, with null mice not showing craniofacial or any clear developmental phenotypes, albeit with defective G1 arrest reported in cultured cells (Deng et al., 1995). Nonetheless, it has been reported to play a role in secondary palatogenesis. *P21*$^{Cip1/Waf1}$*/Cdkn1a* is normally present in the midline epithelial seam (MES) and midline epithelial edges (MEE) that must be eliminated before fusion of the palatal shelves is complete. It was reported that all trans-retinoic acid induced CPO led to loss of *p21* expression at the MEE/MES (Li HY et al 2014. Article in Chinese). *P21*$^{Cip1/Waf1}$*/Cdkn1a* is a target of ΔNp63 (ectoderm isoform of p63) which in turn is in the same pathway as IRF6 and TGFβ, of which all three are linked to CL/P and PBX1/2 (Ferretti et al., 2011; Losa et al., 2018). Alterations to the TGFβ-IRF6-ΔNp63-p21 pathway activity results in the persistence of the MEE in the secondary palate (*Irf6* mutations also showed a low penetrance of CL). Overexpression of *p21* rescues this defect and allows for dissolution of the MEE in the secondary palate and proper palatal fusion (J. ichi Iwata et al., 2013). Though p21 is dispensable in development it is sufficient to trigger secondary palate fusion in a pro-CP setting.

Ras family member proteins, previously referred to as p21^{ras}, but not to be confused with *p21/Cdkn1a,* were found to be differentially expressed in the A/WySn CL prone mouse versus non-cleft mutant BALB/cBy (K. Y. Wang et al., 1995). A/WySn non-mutants had reduced staining against pan-p21^{ras} proteins (Hras, Kras, and Nras) in both mesenchyme and ectodermal epithelial cells in the primary palate compared to BALB/cBy. A/WySn CL mutants had further reduced/absent p21^{ras} expression, suggesting a connection between Ras and primary palatogenesis. This connection was more recently substantiated on in the secondary palate where Ras signaling was demonstrated to collaborate with TGFβ signaling to achieve clearance of the MEE seam (Inubushi et al., 2022). Indeed, when considering the other cell cycle modulators believed to be at work in the secondary and primary palate, it stands to reason that the Ras family and its ties to the likes of Gadd45 could be included in the puzzle.

1.8 Outstanding Questions and Experimental Approach

If disentanglement is the goal, the knot of signaling pathways active in the face during midface fusion is gordian. While key players have been identified, and cellular mechanisms observed, the field has not yet disentangled the role of epithelium versus mesenchyme, the crosstalk mechanisms between the two, the need for proliferation versus removal of epithelial cells, CL versus normal fusion signaling, gene regulatory crosstalk, or signal transduction pathway crosstalk. To add another layer of complexity, when integrating the question of cell cycle and cell cycle arrest mechanisms into the midst of these pathways, a new field and all its remaining mysteries is added to the equation. I pose the questions of what can be revealed by investigating midface epithelium in isolation? What do models of CL have to show at a transcriptomic level? How do observations of cell cycle arrest phenomena intersect with normal primary palate fusion and cleft lip pathogenesis?

These studies follow the previous work of the Selleri Lab that first revealed independent roles of the mesenchyme versus epithelium in CL/P pathogenesis (Ferretti et al., 2011) and described discrete mechanisms of epithelial removal at the fusion site through apoptosis and EMT (Losa et al., 2018). As only approximately 50% of epithelial cell removal could be attributed to EMT or apoptosis, questions remained about the remaining cells and the overall transcriptomic state of the epithelium during midface prominence fusion.

When setting out to design this project, the initial goal was to utilize a combination of scRNAseq that focuses mainly on the epithelium of the developing midface as well as to compare the results to a model of cleft lip. In turn, a larger goal was to see if models of CL/P that possess the same phenotype at a gross anatomical scale have convergent or divergent transcriptomic signatures. In investigating the wild-type epithelium of the midface at timepoints corresponding to fusion, we set out to subtract the mesenchyme and its overwhelming cell numbers from the picture to be able to focus on the unique roles and subtle heterogeneities that would otherwise go unresolved. I set out to confirm and validate signatures that were shown through this data, and in doing so quickly found a curious pattern emerge for epithelium at the prominence fusion site. This pattern revealed a yet unreported mechanism of cell cycle arrest in the primary palate similar in some ways to secondary palatogenesis but unique in its own ways.

In utilizing the *Pbx1/2* and later *Bmpr1a* mutant models of CL/P I uncovered novel roles through transcriptomic analysis. By focusing on these mouse models, I addressed the underrepresented primary palatogenesis to better understand the human condition of CL. I combined my knowledge of cell cycle changes at the fusion site with these models, hypothesizing a connection to CL/P pathogenesis. These data enabled collaboration with other experts in the fields of human embryo anatomy, human patient genomics, and bioinformatics. This yielded a trove of new avenues that both validated in humans findings in the mouse model and revealed new OFC risk candidates for further study. The study of one of these candidates marks the end of my studies and the beginning of exploring the still vast sea of data. In taking careful consideration to appreciate a tissue type that is miniscule in cell number but quite literally the outward face of the embryo, I believe we have advanced the understanding of development, morphogenesis, and human health.

CHAPTER 2: SINGLE CELL TRANSCRIPTOMICS OF ISOLATED MIDFACE EPITHELIUM REVEALS CELL CYCLE ARREST AND NOVEL CLEFT RISK CANDIDATES

2.1 Abstract

Development of the midface requires fusion of facial prominences and removal of surface cephalic ectoderm-derived epithelium at the lambdoidal junction. We employ single cell RNA sequencing of sorted murine lambdoidal junction epithelium to disentangle its heterogeneity during stages of midface prominence fusion. We validate and describe 14 populations between stages E9.5-11.5 and unveil a transcriptomic state associated with cell cycle arrest restricted to the fusion site. This cell cluster, that we call the "zippering lambda epithelium", was found and validated to be perturbed in genetic models of cleft lip and palate. Through analysis of human embryos and genomic analyses of human orofacial cleft probands, we confirm the presence of this cell cycle arrested population in *Homo sapiens* and identify associated rare orofacial cleft risk gene variants. We examine the top risk candidate, *ZFHX3/ATBF1,* and validate ties to cell cycle gene expression and cleft lip risk.

2.2 Introduction

Cleft lip with or without cleft palate (CL/P) remain debilitating and common birth defects despite continued efforts to understand their etiology and uncover novel preventative treatments (X. Dong et al., 2017; Jiang et al., 2006; Mai et al., 2019). Development of the craniofacial midface region involves fusion of the maxillary prominences (MxP), along with the medial and lateral nasal prominences (MNP, LNP) at 4-5 weeks gestation in human embryos and in mouse at embryonic day (E)10.5. In the case of CL with or without accompanying CP, the fusion of these prominences at the lambdoidal (λ) junction is disrupted, leading to a persistent unresolved gap in the primary palate that must be surgically corrected in human patients. CL/P cases are largely non-Mendelian and non-syndromic in nature, with the underlying genetic cellular etiologies both diverse and not fully characterized. Investigations of the fine dynamics of tissues during midface morphogenesis and their genetic programs can provide further insight into the root causes of CL/P in humans.

Although the developing midface comprises several tissue types, the prominences mainly consist of cranial neural crest cell (CNC) derived mesenchyme, surrounded by a much thinner single layer of surface cephalic epithelium. In previous studies, much attention has been paid to the roles of the mesenchyme in the causation of CL/P, and recently there have been greater attempts to utilize next generation sequencing (NGS) to capture a full picture of cell types in the developing midface(H. Li et al., 2019). Yet, there remains a lack of information strictly dedicated to the less abundant epithelium that lines each of the prominences during fusion. The epithelium has a known instructional role in promoting growth, morphogenesis, and fusion of the facial prominences, with known epithelial-specific defects producing a CL/P phenotype due to loss of cellular events such as epithelial to mesenchymal transition (EMT) and apoptosis(Ferretti et al., 2011; Losa et al., 2018). Without complete removal of the λ epithelium, the prominences cannot fuse, creating the cleft phenotype. Efforts have documented active removal of epithelium from the fusion site through EMT and apoptosis, yet only were able to match these cellular processes to about 50% of epithelial cells(Losa et al., 2018). The leveraging of new advances in single cell RNA sequencing (scRNAseq) creates opportunities to decipher underlying heterogeneity and cell type transitions among cell tissue layers once thought homogenous. To unveil yet unknown cellular and molecular behaviors that drive the removal of the remaining epithelium and what underlying heterogeneity exists among this cell type in the developing midface we employ scRNAseq of purified epithelial tissue of the developing mouse midface and uncover a previously non-reported relationship between facial prominence fusion at the λ and cell cycle arrest.

Modulation of the cell cycle has morphologic roles in the developing embryo. In the developing brain, lengthening and shortening of the cell cycle is critical to the proper formation of cortical layers(Siegenthaler & Miller, 2005; S. Wu et al., 2021). In developing limbs, epithelial signaling centers of the apical ectodermal ridges (AER) undergo cell cycle arrest, and in the developing enamel knots (EK) epithelial cells cease proliferating to shape the formation of cusps before later disintegrating and reforming at secondary knots in molars(Ahtiainen et al., 2016; Jernvall et al., 1998; S. Y. Jung et al., 2018; Y. Li et al., 2018). Previous studies have noted the appearance of low proliferation among cells at the λ-junction(Song et al., 2009) but have not expanded upon mechanisms by which cell cycle may be regulated or influence the fusion event. Control of cell cycle dynamics and the maintenance of an arrested state is often transcriptional, with transcription factor (TF) control of gene regulatory networks (GRNs) at the helm.

A link between the PBX TF family and CL/P formation was found through the development of fully penetrant mouse models with constitutive and conditional loss combinations of PBX family members(Ferretti et al., 2011). These TALE homeodomain-containing TFs which are highly expressed in the developing midface(Ferretti et al., 2011; Moens & Selleri, 2006; Selleri et al., 2019), exhibit control of mechanisms of epithelial removal at the λ-junction including epithelial-to-mesenchymal transition (EMT) and apoptosis(Ferretti et al., 2011; Losa et al., 2018). Both in mouse and human patients, a pathway of Pbx-Wnt-P63-Irf6 signaling is linked to NSCLP(Ferretti et al., 2011; Losa et al., 2018; Maili et al., 2020; Selleri et al., 2019). The loss of *Pbx* in mouse models leads directly to loss of the EMT and apoptotic programs in the epithelium at the site of fusion(Ferretti et al., 2011; Losa et al., 2018). The full extent of epithelial genetic dysfunction in *Pbx* CL/P models has not been assessed at the single cell level. We reveal a state of epithelial alterations in both *Pbx* and *p63* models of CL/P that includes disruption of normal cell cycle dynamics at the fusion site.

Zinc finger homeobox 3, *Zfhx3/Atbf1,* is a DNA binding encoding a large transcriptional regulator with known roles in brain development in addition to its roles as a tumor suppressor(G. Dong et al., 2020; Q. Hu et al., 2019), metabolism(Nolan et al., 2023), and cardiovascular regulation (Zaw et al., 2017). In its capacity as a tumor suppressor, it functions to inhibit growth via interface with the cell cycle and has a high rate of mutation among several cancer types(Panaccione et al., 2017). ZFHX3 can induce cell cycle arrest in neurons as well as promote $p21^{(Waf1/Cip1)}/Cdkn1a$ expression(C. G. Jung et al., 2005; Miura et al., 2004). *ZFHX3,* along with other zinc finger homeobox family members, has been implicated in craniofacial morphogenesis and is considered an orofacial cleft risk-associated gene(Bishop et al., 2020). Via the intersection of NGS transcriptomics of the mouse midface epithelium and rare variant analysis of over 700 human trios we sought to investigate the links between differential gene expression of developing epithelial clusters and *de novo* gene variants in non-syndromic cleft lip and palate patients. We find *Zfhx3,* a top gene with *de novo* mutations in rare-variant analysis of human patients, is most highly associated with epithelial cells at the fusion site undergoing cell cycle arrest.

Collectively, our investigations into the heterogeneity of the midface epithelium create a detailed reference of this cell type during the critical moments of midface morphogenesis and λ-facial prominence fusion. We leverage validation experiments to identify the spatial-temporal dynamics of epithelial cell sub-types and uncover a novel population of cells participating in fusion while undergoing cell cycle arrest. Further investigation of this population highlights a transcriptomic state that is regulated by known CL/P risk transcription factors and disrupted by the onset of a cleft lip phenotype. Our study pinpoints *Zfhx3* as a novel CL/P risk candidate in humans and furthers our knowledge of midface development mechanics, expanding upon the role of cell cycle regulation in embryonic morphogenesis.

2.3 Results:

Single cell RNA sequencing and cluster identification of midface Epcam positive cells

Epithelium was isolated from Swiss Webster mouse embryos after dissection of facial prominences. At E9.5, the frontonasal prominence (FNP) was dissected, avoiding the forebrain, along with the MxPs. E10.5 and E11.5 midfaces were dissected first, keeping the LNP, MNP, and MxP intact, then further dissected using transverse cuts midway through the LNP and MNP, and halfway through each MxP **(Fig. 2.6.1: A-A')**. After disassociation of tissue, fluorescent activated cell sorting (FACS) targeting Epcam positive cells yielded epithelial cell enrichment **(Fig. 2.6.1A''')**. Epithelial cells represented x, y, and z percentages of total cells at E9.5, E10.5, and E11.5 stages respectively. 10X genomics based single cell sequencing of captured epithelial enriched cells revealed a total of 14 populations across E9.5, E10.5 and E11.5 mouse embryos. Visualizing of the transcriptomic data is shown through uniform manifold approximation and projection (UMAP) plots. **(Fig. 2.6.1B).** The proportional individual cell transcriptomes are represented as dots separated by distance, wherein closer distance is indicative of transcriptional similarity in UMAP space(McInnes et al., 2018). Each cluster was analyzed and differentially expressed genes (DEGs) examined to categorize and identify putative cell types.

Our E9.5 dataset corresponds to the timepoint prior to midface development and comprises 25,000 cells. Eleven clusters were defined as separate populations at E9.5 **(Fig 2.7.1C, Supplementary Files)**. Three main epithelial populations were associated with sensory and neuroepithelial organs including *Hmx1* and *Sfrp2* high early eye surface ectoderm (EE), forebrain (FB) defined by telencephalon genes responsible for forebrain development *Zic1* and *Fez1*, and olfactory epithelium (OE) defined by *Hey1* and *Dlk1*, implicated in olfactory neurogenesis and nasal pit formation respectively (Miller & Cole, 2014). Oral cavity epithelium (OR) is found in this grouping as well, defined by high *Pitx1* and *Pitx2*. These four clusters of cell transcriptomes that were close in UMAP space fell into a larger category we describe as “specialized epithelium,” high in *Epcam* expression, yet clustered away in UMAP from other *Epcam* high cells **(Fig. 2.6.1B, 2.7.1B)**. By and large these four clusters share higher expression of neuroectodermal determiner *Sox2* and lower expression of *Tfap2a, Sostdc1, Wnt6,* and *Krt7,* genes more indicative of developing protective surface and epidermal cell fates. A described PITX2-SOX2-LEF1 axis of specifying oral and dental progenitors aligns with *Sox2* high expression found in our putative OR cluster and its inclusion in this grouping (Yu et al., 2020).

At E9.5, two remaining *Epcam* high main clusters correspond to the ectodermal covering of the FNP, prominence surface 1, and prominence surface 2 (PS1, PS2). Gene expression of these clusters was similar, both high in *Wnt6, Tfap2b* and *Wnt4*, distinguished only by cell cycle genes relating to proliferation, e.g. *Top2a*. It was noteworthy that despite nasal pit ingression not yet occurring at E9.5, this timepoint has clear transcriptomic separation of the OE and presumptive olfactory placode away from the PS1 and PS2 clusters.

Two populations of cells separated away from others while remaining near FB transcriptomes suggesting a connection to neural cell-types. They were defined as neural progenitor cells (NPC) (*Stmn1, Neurod1* and *Sst)* as well as olfactory neurons (ON) based on strong differential expression of alpha and beta tubulin microtubule assembly genes *Tub1a* and *Tubb3* with known roles in neurogenesis and axonal guidance. NSC and ON transcriptomes formed a third class of epithelial cells grouped under the term “neuronal.”

Using EPCAM-based FACS sorting enriched for epithelium but to maximally capture epithelium, cells with moderate EPCAM-APC positivity were included in gating strategies **(Fig. 2.7.1A)**. Two clusters with relatively low EPCAM expression were captured and separated distantly away from other clusters **(Fig. 2.7.1B)**. Based on gene expression profiles, these clusters were given the names Mesenchyme-like (ME) based on *Prrx1* and *Twist1* gene expression, and neural crest cell (NCC)-like based on *Sox10* and *Vim* expression. An additional minor cluster expressing markers of blood/erythrocytes (B) including ferritin heavy chain (*Fth1)* and hemoglobin genes (*Hba-x, Hbb-y, Hbb-a1*) was not considered for further analysis. Clusters B, NCC and ME were excluded from further analysis and validation.

E10.5 midfaces correspond to a timepoint of midface development in which MNP and LNP have emerged and the lambdoidal junction formed with the MxPs. The initial fusion of the lambdoidal junction begins at E10.5. scRNAseq and subsequent clustering of E10.5 epithelium gives rise to 14 distinct populations, with the addition of two clusters not detected at E9.5, and major changes in the proportions of several others **(Fig.2.6.1B, 2.7.1C)**. NPC and EE clusters were most reduced by E10.5. The size and distance of the λ junction tissue from the developing eye at this timepoint may contribute to reduction of the EE cluster. Two additional populations characteristic of later progression in development. This timepoint marks the emergence of periderm-like cells (PD) defined by *Grhl3* and *Tacstd2* and heralds differentiation of a subset of prominence surface cells. This cluster was found in numbers too small to be called in E9.5. Secondly, a cluster emerged centrally between OF, OR and PS1/2 clusters that possessed strong expression of genes known to play critical roles in orofacial development such as *Fgf8, Wnt5a,* and *Msx1*. The initial identity of this cluster was believed to be the distal tips of the prominences where cells participating in the closing of the λ junction seam would be present. The timing of this cluster's appearance, coinciding with the beginning of fusion helped define this cluster as a putative "zippering lambda epithelium," (ZL).

The midface and λ junction continue to fuse through E11.5 and reaches completion shortly after at E12.0. We sampled the E11.5 timepoint in two batches of separately FACS sorted lambdoidal junctions **(Fig. 2.7.1D).** UMAP clustering of E11.5 timepoint transcriptomes reduced the number of distinct clusters to 11, with the loss of two previously called populations, NPC, FB and near complete loss of EE (not visible in UMAP) **(Fig. 2.7.1C).** The emergence of an E11.5-specific cluster enriched for *Sox9* and *Aldh1a3* suggested a connection to eye and lacrimal duct development. Both genes matched literature references to a nasolacrimal groove development around the E11.5 timepoint. The population was named nasolacrimal duct (ND) in reference to this signature.

Together, across the three sampled timepoints, ZL, PD, PS1/2 (hereafter referred jointly as PS), and ND are collectively "epithelium" of the midface separated from the specialized epithelium, neuronal, and erythrocyte, and mesenchyme-like lineages **(Fig 2.7.1E).** Mapping of midface epithelial clusters using differentially expressed genes and *in situ* hybridization.

E9.5, E10.5 and both E11.5 batches were combined into a single UMAP encompassing the spread of pre-midface fusion, initial fusion, and nearly complete fusion of the lambdoidal junction **(Fig. 2.6.1C).** The merged dataset yielded no additional clusters, for a combined total of 14. To validate the clustering and separation of these cell populations, we employed fluorescent *in-situ* hybridization (FISH) probes designed via RNAscope (Advanced Cell Diagnostics) against DEGs for each **(Fig. 2.6.1D).** Ribosomal, non-coding RNA, and non-annotated genes were excluded as candidates for validating the localization of the clusters. Probes were used against tissue sections of wild-type embryos collected at the appropriate timepoints **(Fig. 2.6.1A"-A"').** Clusters that were found predominantly in one timepoint were assessed using embryos of the corresponding stage.

In **figure 2.6.1D** the DEG for each cluster is superimposed on a heatmap of the combined E9.5-E11.5 timepoint (left). **Figure 2.6.1D** shows a representative tissue section image of the validated DEG expression (right). Olfactory epithelium (OE) has the highest *Sox2* expression of all clusters. *Sox2* expression localizes to the deep dorsal nasal pit with slight medial bias. The PS clusters share high *Wnt* gene expression, of which *Wnt6* is the most differentially expressed. *Wnt6* expression encompasses the anterior outer surface of MNP, LNP, and the MxP. These clusters have the largest surface area coverage and are among the highest in total cell number for the midface. *Wnt6* expression is not continuous with the nasal pit or the oral cavity, stopping short near the midface prominence tips and medial edge of the MNP. TUBB3 protein, using immunofluorescence as opposed to FISH, was indicative of the olfactory neurons (ON) found interlace among *Sox2* high OE cells in the dorsal nasal pit. Their localization extends past the basal lamina border in some cases and can be found anterio-medially in lower abundance. Oral cavity (OR) epithelium is defined by *Pitx2* expression on the oral side of the medial nasal prominences flanking the oral cavity. The zippering lambda epithelium (ZL) has high differential expression of BMP-inhibitor *Bambi.* FISH against *Bambi* confirmed our intuition for the cluster localization at the site of fusion where the epithelium of the prominences begins to coalesce and initiate fusion. For reference, we superimpose the ZL, OE, OR, PS, and ON cluster relative positions on a E10.5 model of mouse embryonic face generated through micro computerized tomography μCT **(Fig. 2.6.1E).**

PD and EE clusters were investigated through targeted FISH against *Grhl3* and *Hmx1* respectively. For EE, the expression of *Hmx1* and the cluster size is greatest at E9.5 and was validated using E9.5 embryonic sections. *Grhl3* expression is found on the apical surface of the *Wnt6* PS cluster, with close analysis demonstrating a second layer of cells developing by E10.5

and through E11.5. By E11.5 PD cells can also be found on the interior of the nasal pit on the anterior edge of the cavity near the site of completed lambdoidal junction fusion. ND cells were found only in sections of the deep mandibular prominences at E11.5. Using *Aldh1a3* as a target we witnessed high levels of expression in the developing eye, which was excluded from the tissues dissected for sequencing. Epithelial cells in a groove involuting near the eye also expressed *Aldh1a3* and in conjunction with the role of *Sox9* in ductal development were deemed nasolacrimal duct epithelial cells **(Fig. 2.7.1F).**

Molecular characterization of the zippering λ epithelial cluster and in vivo cell fate mapping

The ZL cluster became the focus of interest due to its prominent location at the point of λ junction prominence fusion. The emergence of this cluster at E10.5 between other clusters that existed in the E9.5 dataset led to the question of whether this cluster's appearance could be pre-emptively determined at E9.5. A gene that was consistently high in this population in both the E10.5 and E11.5 timepoints **(Fig 2.6.2B, 2.7.2B)**, *Bmp4,* has a known role in the proper development of the midface, with BMP and TGFβ signaling both linked to defects in craniofacial development. Using the E9.5 dataset, we can see that *Bmp4* expression is localized to a region between the OE, OR, and PS clusters, which also are the clusters flanking the ZL cluster at E10.5 **(Fig. 2.6.2A)**. To this end, we sought to validate *Bmp4* expression in E9.5 embryos. Sections using FISH probes revealed a prominent region of high *Bmp4* expression in the posterior side of the outgrowing FNP. Using a whole mount traditional chromogenic *in-situ* hybridization (WISH), a greater spread of *Bmp4* expression can be seen including at the midline **(Fig. 2.6.2A')**. In Figure 2B, sections of E10.5 embryos focused on the fusing lambdoidal junction confirm and validate the strong expression of *Bmp4* of the ZL cluster and revalidating the localization of this cluster to the fusion site. WISH at E10.5 likewise demonstrates a clear bias of *Bmp4* expression to the lambdoidal junction **(Fig. 2.6.2B')**.

Other transcripts that were enriched in the ZL included genes with known implications in CL/P pathogenesis and midface development. *Tgfβ*, *Fgf9*, *Wnt5a*, and *Trp63*, each linked to models of CL/P in mouse have high enrichment in the ZL cluster **(Fig. 2.6.2C)**. This strongly implicated the ZL cluster transcriptomic state in the pathogenesis of CL/P. Using several bioinformatic tools we sought to interrogate the transcriptome of the ZL cluster. RNA velocity comparisons of spliced versus un-spliced mRNA transcripts can predict lineages and relationships between scRNAseq clusters (La Manno et al., 2018). Examining the transcriptomes at E10.5 and E11.5, the vector directions are seen to generally point to three areas of convergence. The E10.5 timepoint endpoints are mapped accordingly to where first the PS clusters cycling between proliferation and non-proliferation, second a tract of neurogenesis involving NPC and ON cells, and lastly a convergence point on the ZL between PS, OR and OE clusters is seen **(Fig. 2.6.2D).** This pattern is echoed at E11.5, where the two sources of vectors to the ZL are predicted to be the OR and OE clusters. At E9.5 RNA velocity does indicate vectors between the OE, PS and OR clusters converging at a point between the three, suggesting that the determination of the ZL cluster may begin a day before the formation of the lambdoidal junction **(Fig 2.7.2E).**

We used an alternative method to predict cell lineage trajectories, CytoTRACE, which is based on a principle of associating decreased transcriptional diversity with further cell differentiation (Gulati et al., 2020). We applied the scRNAseq data of each timepoint to CytoTRACE and revealed several clusters to predict lineage trajectories. At E9.5 the ON and NPC clusters were predominantly the most differentiated, in line with the terminally differentiated state of neurons **(Fig. 2.7.2F)** and to a lesser degree, a subset of OE. At E10.5 ON and NPC clusters retained their differentiated status, but only partially. The ZL cluster in its entirety was classified as the most differentiated subset of the transcriptomic data **(Fig. 2.6.2D').** At E11.5 the ZL cluster retains a mostly differentiated state, along with portions of other clusters: PS, OR, OE, and ON **(Fig. 2.7.2F)**.

Further characterization of the ZL cluster was done using gene ontology (GO analysis) using DAVID functional annotation tools. GO Biological Process matches of the top 100 differentially expressed genes for the ZL cluster gave rise to results that suggested an association of the cluster with cell cycle, apoptosis, proliferation, and division, altogether suggesting that genes involving regulation of the cell cycle are enriched in this cluster **(Fig. 2.6.2E).** We correlated this finding with the high DE of *Cdkn1a/p21* for this cluster observed upon its initial analysis **(Fig. 2.7.1E)**.

As we previously have reported on several cellular mechanisms and states taking place in the λ junction during fusion (Ferretti et al., 2011; Losa et al., 2018), we sought to unveil further levels of heterogeneity in the ZL cluster. To accomplish this, we reclustered the ZL cluster from E10.5 and E11.5, and examined the subclusters using RNA velocity, differential gene expression, and gene ontology. Reclustering of the E10.5 ZL cluster resulted in 5 cell subclusters **(Fig. 2.6.2F center)**. RNA velocity was applied to this subset and two endpoints revealed as terminal states in subclusters 1 and 3 **(Fig. 2.6.2F right)**. Highly DE genes that appeared to coincide with these endpoints included *Cdkn1a* expression, closely overlapping with *Tgfb2* expression. Additionally, *Sostdc1,* a BMP inhibitor and *Lef1* a downstream target of *Wnt* signaling appeared to have strong correlation with the two endpoints as well.

We plotted the highly DE genes for each subcluster and performed gene ontology of the top 40 genes from subclusters 1 and 3, the locations of the terminal endpoints **(Fig. 2.6.2H)**. Our analysis of subcluster 1 revealed top GO categories matching apoptosis, EMT, and MYC targets. It should be noted that MYC is known to target cell cycle arrest genes for de-repression at the transcription level, including *Cdkn1a, Cdkn1b,* and *Cdkn2b* (García-Gutiérrez et al., 2019). E10.5 subcluster 3 had similar GO term enrichment for apoptosis and EMT, as well as estrogen and WNT-beta catenin signaling. Together these terms support previous work that identifies apoptosis and EMT mechanisms of fusion at the λ junction driven in part by upstream WNT signaling. We performed the same analysis on the ZL of the E11.5 datasets and found a similar dual endpoint corresponding to subclusters 3 and 4 **(Fig. 2.7.2G)**. E11.5 endpoints were similarly enriched in *Tgfb2, Cdkn1a,* and to a lesser degree *Lef1* and *Sostdc1* **(Fig. 2.7.2H)**. One endpoint was notable for its strong positive DE of *Isl1* and negative DE of *Pax3. Pax3* has recently been implicated in craniofacial morphogenesis (Zhou & Conway, 2022). *Isl1* is implicated in both hindlimb mesenchyme and facial epithelium of the MdP and the midface (Akiyama et al., 2014) Many of these genes are implicated in limb morphogenesis, which has a curious relationship to facial morphogenesis, including through PBX signaling pathways (Losa et al., 2023; B. T. Truong & Artinger, 2021).

RNA velocity predicted the ZL cluster to be derived from neighboring cell clusters. We employed genetic lineage tracing strategies to validate this finding using a combination of tissue specific *Cre*-recombinase and fluorescent reporter construct mouse lines **(Fig. 2.6.2J)**. OR cells were defined by *Pitx2* high DE which extended into the ZL cluster in part by E10.5. *Pitx2-Cre* mice were crossed to *R26R-RFP* and embryos collected at the E10.5 timepoint. The RFP fluorescent reporter was found overlapping two markers of the ZL cluster, *Bmp4* and *Cdkn1a.*

The PS cluster is similarly defined by strong DE of *Wnt* genes such as *Wnt6* and *Wnt4*. Using *Wnt4-Cre* as a driver of recombination yielded similar results, with RFP protein found overlapping with *Bmp4* and *Cdkn1a* at E10.5. We confirmed that both the OR and PS cells are precursors to the ZL lineage.

Non-proliferation and absence of cell cycle progression markers in the zippering lambda epithelium

The knowledge of lowered proliferation at the fusion site of the λ junction was previously shown but not reported on (Song et al., 2009). We re-examined this in light of the strong expression of *Cdkn1a* found in the ZL cluster at the fusion timepoints. We employed scVelo analysis using the 'score genes cell cycle' function, which calculates scores and assigns cell cycle phases to the transcriptomic data based on a predetermined list of cell cycle genes (Tirosh et al., 2016). When applying this function to the E9.5, E10.5 and E11.5 datasets, a mixture of scoring is found among most major clusters, with G2M predominating at E9.5. At E10.5, an absence of G2M/S scoring prevails in the ZL cluster as with the ON cluster (differentiated and non-dividing neurons). By E11.5, more cells have become non-dividing, including subsets of PS, ND, and OR clusters, but the ON and importantly the ZL cluster remain devoid of G2M/S scoring **(Fig. 2.6.3A).** The ZL cluster had quantifiably the lowest S score of all populations at E10.5

Multiple phases of the cell cycle are associated with corresponding expressed gene markers, such as phosphorylated histone H3 (pHH3) with G2M. Additionally, exogenous labeling strategies, such as thymidine analogue integration via BrdU/EdU has the potential to label cells passing through the S phase. We examined several of these markers *in-vivo* and found a marked absence of cell cycle phase-related transcripts and proteins in the ZL cluster region. pHH3 positive cells could be found in the developing olfactory epithelium, and were abundant in mesenchyme, but were rarely present in the fusing junction between MNP and LNP **(Fig. 2.6.3B-B')**. BrdU thymidine analog was injected into time-pregnancies at E10.5, and followed to several timepoints, 1-, 7-, and 24-hours post-injection. Little to no BrdU integration could be detected at the lambdoidal junction at each timepoint **(Fig 2.6.3C-C')**. We examined the scRNAseq data and found low DE of *Aurkb, Cdk1,* and *Mki67* in the ZL cluster. *Aurkb,* another G2M phase marker, was validated using sections of the midface to be absent in the ZL region denoted by high *Bambi* expression. Similarly, in both transverse and coronal sections, *Cdk1* expression, which peaks at G2M phase transition is completely absent from the lambdoidal junction during fusion. *Mki67* expression, which builds gradually from G1 through M phase, was lowly DE in the ZL, and validated to be low at the tips of the MNP and LNP, even prior to physical contact and fusion.

Genes of the pre-initiator complex were largely negatively differentially expressed in the ZL, including *Mcm7, Gins2, Cdc6* and *Cdc20* **(Fig. 2.7.3A).** Cyclin genes, *Ccna2, Ccnb1, Ccnd1,* and *Ccne2* were among genes similarly depleted. E2 promoter binding factor family member *E2f1*, *Anln,* and *Pcna* factors associated with G1/S, M (cytokinesis), and S phases respectively are similarly DE negatively as well. Depletion of these genes, as well as loss of KI67 protein would suggest G0/G1 checkpoint arrest is occurring in the ZL cluster. The combination of all these factors, allude to the ZL cluster being non-proliferative at the time of fusion, E10.5.

The zippering lambda epithelium is enriched for markers of cell cycle arrest.

High *Cdkn1a/p21* expression in the ZL cluster was out first indication that the cell cycle may be playing a role in its formation. We sought to investigate spatio-temporally the presence of *Cdkn1a* throughout midface fusion timepoints, along with other key positive markers of arrest.

In **figure 2.6.4A** we explore through WISH the evolving landscape of *Cdkn1a/p21* expression. There is a minor presence of *Cdkn1a* transcript found in E9.5 embryos in the posterior side of the FNP, similar to the location of high *Bmp4* expression detected previously. At E9.5 *Cdkn1a* transcript is highest in the mandibular prominence tips which undergo their own type of fusion at the midline at an earlier stage than the lambdoidal junction but are outside the scope of our investigation. E10.25 embryos, in which the nasal prominences are barely starting to form, already begins to show highly specific localization of *Cdkn1a/p21* in the epithelium between the presumptive LNP, MNP, and MxP, outlining the λ in its entirety. The specificity of this expression to the epithelium, and not underlying mesenchyme can be seen in section. By E10.5, this expression continues to be strongly expressed at the fusion site between the MNP/LNP, and the MNP/MxP, but has regressed in the LNP/MxP junction region. A slight gap can be seen forming in expression where the lambdoidal junction has begun to seal. By E11.25, this gap has enlarged where the MNP and MxP have joined, and only anterior portions of the MNP/LNP junction as well as an oral edge of MxP retain high expression. Altogether we correlate *Cdkn1a/p21* expression with ongoing face prominence fusion, and its loss with the completion of the fusion event.

We employ FISH probes to further validate the bioinformatically determined high DE of *Cdkn1a/p21* in both E10.5 and E11.5 sections of the midface. The presence of *Cdkn1a/p21* is found in both timepoints at the tips of the LNP and MNP prior to fusion and contact and is transcribed to protein P21 which remains localized to the ZL. **(Fig. 2.6.4B, C, E)**. *Gadd45g* and *Rb1* are both associated with cells in a state of arrest. *Gadd45g* also possesses a link towards CL pathogenesis recently unveiled (Lu et al., 2019). We validated the presence of *Gadd45g* and *Rb1* transcript at E10.5 and E11.5 and found high overlapping in the region of high *Cdkn1a/p21* expression in the ZL. *Gadd45g* expression could, like *Cdkn1a/p21* be mapped to the presumptive area of the lambdoidal junction early in its formation where the olfactory placode is specified **(Fig. 2.6.4F).** As with *Cdkn1a/p21*, *Gadd45g* expression faded from the λ junction with the zippering of the epithelium and was found to be localized to the epithelium before contact of the prominences, without clear mesenchymal expression.

Although the ZL cluster subcluster analysis found transcriptomic signatures that matched gene ontology for apoptosis, only the endpoint subclusters matched this term. Additionally, our previous studies indicated only a portion of cells underwent apoptosis and a mutually exclusive group underwent EMT (Losa et al., 2018). To this end, we investigated the overlap of apoptotic genes *in-vivo* in relation to *Cdkn1a/p21*. We found that only a small portion of *Cdkn1a/p21* high cells in the ZL were positive for cleaved-Caspase3 (cCasp3) and terminal deoxynucleotidyl transferase dUTP nick end labeling (TUNEL), suggesting that cell cycle inhibition may not directly lead to cell apoptosis in the ZL **(Fig. 2.6.4D)**. Though *Cdkn1a/p21, Rb1,* and *Gadd45g* are not directly implicated in CL/P pathogenesis in mouse, we sought to investigate whether the associated gene patterning that is indicative of cell cycle arrest in the ZL was altered in CL/P models.

Models of CL/P display cell cycle perturbations in the zippering lambda epithelium and absent or reduced cell cycle arrest.

There is a known connection between proliferation in the midfacial prominences and the occurrence of CL/P. However, the epithelium has at times been relegated as a mediator of mesenchyme proliferation, that when perturbed, leads to CL/P via prominence shortening (Liu et al., 2005). Specific roles of cell cycle and cell cycle arrest genes have been document in secondary palatogenesis in the midline epithelial edges of the palatal shelves, but they have not been correlated with primary palatogenesis or its disruption. The *Pbx1/2* model of CL/P, involving epithelial specific loss of *Pbx1* on a *Pbx2* globally deficient background produces a bilateral fully penetrant CL/P. This model serves as our initial investigation into epithelial specific disturbances in CL/P through scRNAseq.

Using E10.5 and E11.5 mouse midfaces, we performed scRNAseq of *Pbx1/2* mouse mutants, *Crect*$^{Cre/+}$*; Pbx1*$^{f/f}$*; Pbx2*$^{+/-}$, versus littermate *Cre* negative controls, *Crect*$^{+/+}$*; Pbx1*$^{f/f}$*; Pbx2*$^{+/-}$ or *Crec*$^{+/+}$*; Pbx1*$^{f/+}$*; Pbx2*$^{+/-}$. The midfaces were collected through the same dissection scheme as with Swiss Webster wild-type tissue but were sequenced without FACS sorting for EPCAM-positive tissue to avoid loss of mutant material. Bioinformatically determined EPCAM-high epithelial cells were selected and isolated from mesenchymal populations and reclustered in UMAP. This yielded 10 clusters for the E10.5 and 11 clusters for E11.5. Using label transfer function, we annotated the epithelium transcriptomic signatures with pre-determined labels from Swiss Webster scRNAseq and identified commonalities to wild-type data. Importantly, the close associations of the ZL cluster with neighboring OR, OE, and PS cells was replicated in the *Pbx1/2* scRNAseq UMAP.

We examined the scRNAseq of the E10.5 mutant versus control dataset through RNA velocity and compared the trajectories of predicted lineage vectors **(Fig. 2.6.5A)**. At E10.5 mutant *Pbx1/2* ZL cluster contains a noticeable shift in trajectories compared to controls. The shift was most pronounced across the border between the ZL and PS clusters. We determined empirically that the vectors between PS1 and ZL clusters was significantly perturbed in *Pbx1/2* mutants suggesting a malfunction of the normal lineage contributions **(Fig. 2.6.5B)**. Our E11.5 dataset possessed more compact clustering than the E10.5 dataset **(Fig. 2.6.5C)**. We investigated whether there were changes in the relative proportion of each population between *Pbx1/2* mutants versus controls. While there were significantly more ON cells in the mutants, the control dataset proportionally had more PS, ZL and PD cells than mutants **(Fig. 2.6.5D)**. We interpret this to be related to improper formation of the ZL in *Pbx1/2* mutants. We cannot exclude the possibility that other cell populations significantly affected in the *Pbx1/2* mutants contribute to the onset of CL/P, as suggested by similar cell populations in Siewert et al 2023, where periderm and other ectodermal cell clusters analyzed from mouse embryonic transcriptomic data had DEGs that were known CL/P risk candidates (Siewert et al., 2023)

We used an alternative method to quantify transcriptional changes occurring in the midface epithelium of E11.5 *Pbx1/2* embryos versus littermate controls. Bulk-RNA sequencing (bulk RNAseq) of the entire epithelium from several mutant embryos versus controls demonstrated a clear transcriptomic separation between the two groups, with 718 genes significantly altered in expression (adjusted p-value <0.05) **(Fig. 2.6.5E)**. Among the genes differentially expressed, cyclin dependent kinase inhibitor genes, *Cdkn1b, Cdkn1c, Cdkn2a, Cdkn2b,* and the prominent marker of the ZL, *Gadd45g* were significantly decreased in *Pbx1/2* mutants versus controls. RNAsieve analysis allowed us to interpolate the results of the bulk RNAseq and out scRNAseq data and examine which clusters of cells were most correlated with DE genes in the bulk RNAseq. We found that genes significantly downregulated and upregulated in our bulk RNAseq were most closely associated with the ZL cluster, strongly suggesting the disruption of this cell type in CL/P mutants **(Fig. 2.6.5F)**.

To validate the expression changes seen in *Pbx1/2* mutant embryos for key cell cycle arrest related genes, we examined transverse sections of embryos from both mutant and control genetic backgrounds for the presence of the transcripts. We introduce the use of constitutive null mutants and controls alongside their epithelial-specific counterparts. *Pbx1*$^{+/-}$*; Pbx2*$^{+/-}$ embryos do not have CL/P and have no noticeable deleterious phenotypes. In contrast, *Pbx1*$^{-/-}$*; Pbx2*$^{+/-}$ produces a severe and fully penetrant bilateral CL/P phenotype. We detect low levels of *Cdkn2b, Cdkn2a,* and *Cdkn1c* transcript at the prominence fusion site among the ZL cluster cells in control *Pbx1*$^{-/-}$*; Pbx2*$^{+/-}$ embryos **(Fig. 2.6.5G top row)**. This expression, unresolved in our scRNAseq datasets, was found to have remarkable specificity to the fusion site, albeit with sparse transcript levels. These ZL cells exhibited the same high levels of *Rb1* and *Gadd45g* transcript seen in Swiss Webster wild-type embryos. *Pbx1*$^{-/-}$*; Pbx2*$^{+/-}$ have a noticeable large separation of the LNP and MNP which does not resolve at any plane of midface sectioning due to the CL/P phenotype. Among all midface section, little to no transcript can be seen at the prominence tips for *Cdkn2b, Cdkn2a, Cdkn1c, Rb1,* or *Gadd45g,* implying the cell cycle arrest state is missing in the CL/P phenotype **(Fig. 2.6.5G middle row)**. *Crect*$^{Cre/+}$*; Pbx1*$^{f/f}$*; Pbx2*$^{+/-}$ CL/P mutants have the same large separation of the MNP and LNP and similarly we observe an absence of the cyclin inhibitor kinase transcripts as in constitutive mutants **(Fig. 2.6.5G bottom row)**. We see strong reductions in *Rb1* and *Gadd45g* transcript as well but with slight increases compared to constitutive mutants. *Cdkn1a/p21* was not found significantly altered in expression in bulkRNAseq of *Pbx1/2* embryos despite being associated with cell cycle arrest and having the highest expression of all cyclin dependent kinase inhibitors in the ZL cluster. *Cdkn1a/p21* was confirmed to be present in abundance in both global and conditional *Pbx1/2* CL/P mutants **(Fig 2.7.4A)**. *Cdkn1a/p21* does not have its own CL/P phenotype in knockout models but leads to severe embryonic defects in conjunction with loss of other family members (Campbell et al., 2020). While we conclude that *Cdkn1a/p21* expression is not dependent on *Pbx1/2* signaling, we cannot exclude the possibility that it plays a role in CL/P risk.

Trp63 executes a wide array of tissue and context specific roles, some of which are isoform dependent. This is reflected in the diverse range of disease-causing mutations found in specific DNA binding domains. In humans, *TRP63* mutations that are tied to non-syndromic OFC risk are found throughout the gene as opposed to domain centric mutations such as in ankyloblepharon-ectodermal defects-cleft lip/palate syndrome (AEC) (Osterburg et al., 2023). We employ the use of a the *Trp63*^{tm1Fmc} mutant mouse line affecting all isoforms of the gene, including those specifically linked to ectodermal and orofacial defects such as *ΔNp63* (Yang et al., 1999). The *Trp63*^{tm1Fmc} mouse line phenocopies the *Pbx1/2* global and conditional mutant CL/P with severe and fully penetrant bilateral CL/P. This model of CL/P is related to the *Pbx1/2* models via a shared pathway, wherein *Trp63* is believed to act downstream of PBX transcription factors in an established PBX-WNT-TP63-IRF6 axis of control also implicated in human non-syndromic CL/P pathogenesis (Ferretti et al., 2011; Losa et al., 2018; Maili et al., 2020; Thomason et al., 2008). We collected homozygous *Trp63*^{tm1Fmc} (*p63*$^{-/-}$) embryos for FISH validation of transcripts related to cell cycle arrest. As with *Pbx1/2* global and conditional CL/P mutants, we could not detect transcripts of the predicted downregulated cyclin-dependent kinase inhibitor genes from our *Pbx1/2* bulk RNAseq in homozygous *Trp63*^{tm1Fmc} mutant embryos **(Fig. 2.6.5H)**. To an extent, *Rb1* and *Gadd45g* transcripts were found in even less abundance in homozygous *Trp63*^{tm1Fmc} mutant embryos than the *Pbx1/2* CL/P mutants.

There are conflicting reports of *Cdkn1a/p21* being a known target of ΔNp63, TAp63, or both classes of *Trp63* isoforms (Dohn et al., 2001; Ghioni et al., 2002; Helton et al., 2007). The role of *Trp63* in repression adds a layer of complexity, with ΔNP63 also stated to repress *Cdkn1a/p21*(A. B. Truong et al., 2006; Westfall et al., 2003). A frameshift mutation of either TA or ΔN isoforms in exon 13 was reported to increase *Cdkn1a/p21* transcription yet also the proline rich N-terminus residues 124-127 were shown to be most critical to *P21* transactivation (Ghioni et al., 2002; Helton et al., 2007). Curiously, we detect high amounts of *Cdkn1a/p21* transcript in homozygous *Trp63*^{tm1Fmc} mutant embryos **(Fig. 2.7.4A)**. *Trp63*^{tm1Fmc} bears a disrupting mutation in exons 6-8, therefore we cannot exclude the possibility that intact N-terminus exons are translated into a protein with functional active domains capable of inducing *Cdkn1a/p21* (Bergholz & Xiao, 2012; Helton et al., 2006). We also cannot exclude the possibility that side pathway interactions are implicated in *p63* null loss of these markers, but the evidence suggests PBX-WNT-P63-IRF signaling is upstream of cell cycle arrest outcomes in the zippering lambda epithelium during fusion of the midface prominences.

Midface prominence fusion in human embryos is associated with cell cycle arrest in the epithelium.

Midface prominence fusion is believed to proceed in a phenotypically similar manner among many animals, including humans. In human development fusion of the prominences is proposed to involve similar mechanisms to those in mice, wherein confluence of the MNP, LNP, and MxP forms a lambdoidal junction prior to merging of the prominences into a single unit. The timing of these events is believed to be around the 5th through 7th week of human development, corresponding to roughly Carnegie stage (CS) 15-17 (Selleri & Rijli, 2023). We lack understanding of the cellular mechanisms at work in human midface morphogenesis. There is some evidence of cell death at the λ junction of human embryos before fusion, but knowledge of molecular mechanisms is absent (Jiang et al., 2006). We investigated the presence of cell cycle arrest at the λ junction in human embryos during midface fusion.

CS16 human embryonic head sections were prepared using a transverse approach through the midfacial prominences. At CS16, the MNP and LNP have started to fuse, and the epithelium at the distal tips of the prominences coalesces, in similar fashion to E10.5 mouse embryos **(Fig. 2.6.6A)**. RNAscope FISH against *GADD45G* revealed a region of high expression where the epithelium bridging the fusion site is located. Immunofluorescence against KI67 indicated that the *GADD45G* high region is similarly depleted for the marker of cell proliferation. Using a slight coronal section of the midface, we capture the λ junction in full including the MxP, MNP and LNP junctions. Like in mouse, the epithelium within the human three-way junction is depleted for *CDK1* transcript and KI67 protein **(Fig. 2.6.6B)**. CS17 embryonic sections demonstrate midface prominence fusion at a more advanced stage similar in appearance to mouse E11.5 embryos. As in CS16, the λ epithelium is enriched with *GADD45G* transcript and lacks KI67 and *CDK1*, but notably possesses enrichment of *CDKN1A/P21* transcript as well **(Fig. 2.6.6C, 2.7.5B)**. The decreased cellular proliferation and increase in expression of cell cycle arrest/DNA damage markers, indicates an evolutionarily conserved phenomenon between mouse and human midface development through at least the 7th week of human development, and suggests our ZL-cluster transcriptomic signature may have an equivalent in human embryos.

An array of datasets has been generated by collecting genomic information from CL/P human patients worldwide. The study of CL/P patient trios, in which whole genome sequencing of OFC case patients and their non-affected parents, is suitable for the discovery of *de-novo* mutations. Using a database of 759 orofacial cleft (OFC) trios from a diverse ethnic pool, the top 150 differentially expressed genes for each main cell cluster, identified by our scRNAseq of mouse embryonic midface epithelium, were used to identify OFC risk variations **(Fig 2.6.6D)**. Both *de novo* variants, as well as rare inherited variants, were examined in the patient trio data. As an outgroup, patient trio data for autism spectrum disorder (ASD) was included in the analysis for reference. *De novo* variant enrichment of all cluster marker genes from the mouse scRNAseq data sets demonstrated a significant correlation to OFC probands variations, notably protein loss of function (pLOF) and protein altering gene variants **(Fig. 2.6.6D)**. In contrast, ASD probands and unaffected siblings showed a significant but far lower enrichment score.

Rare variant analysis of unfiltered top genes from each cluster was performed using OFC and ASD outgroup probands. For comparison, additional enrichment testing using genes with known OFC association (Gene Panel), was performed **(Fig. 2.6.6E)**. Across all clusters, the highest enrichment (p-value) of genes with predicted damaging variants in OFC probands was found in the PD and PS clusters. ASD probands, in contrast, largely matched to ON enriched genes. After multiple test correction for 6 clusters ($p<0.0083$), only the PD cluster overlap for OFC probands, and ZL cluster overlap for OFC-associated genes were significant. Conducting rare variant analysis of the genes enriched in the ZL cluster, we found that the top 10 genes bearing variations in OFC probands, ordered by total number of variants, include *ZFHX3. ZFHX3* encodes a zinc finger homeobox transcription factor containing 4 homeodomains and multiple zinc finger motifs known to be involved in both neuronal and myogenic differentiation and cell cycle regulation **(Fig. 2.6.6G)** (C. G. Jung et al., 2005; Z. Zhang et al., 2019; D. Zhao et al., 2016). It was previously reported that *ZFHX3* is a *de novo* OFC risk candidate (Bishop et al., 2020). *ZFHX3* accounted for 203 variants in the OFC probands, 127 of which were unique. Following filtering steps, 42 total rare variants were identified within *ZFHX3*. The predicted damaging variants for *ZFHX3* found in probands are illustrated in **Fig. 2.6.6G**.

Notably, *Zfhx3* is enriched in the ZL cluster at E10.5 and E11.5 in mouse wild-type epithelium **(Fig.2.6.6H, 2.7.5E)**. We validated *Zfhx3* high expression in the epithelium near the ZL cluster using *Zfhx3*-targeted FISH in mouse midface sections **(Fig. 2.6.6I, 2.7.5D)**. Likewise, both CS16 and CS17 human embryos show high expression at the lambdoidal junction, both in the epithelial layer and in the prominence mesenchyme, although to a lesser extent in the latter **(Fig.2.6.6J, 2.7.5C)**. The focused ZL reclustering of both E10.5 and E11.5 scRNASeq data sets demonstrates enrichment for *Zfhx3* at all predicted RNA velocity endpoints, showing higher expression in subclusters 3 for E10.5 and 4 in E11.5 **(Fig. 2.7.5F, G)**. These expression patterns overlap strongly with *Tgfb2* expression. SMAD-dependent TGFβ signaling is known to collaborate with ZFHX3 to regulate transcriptional activity involving cell proliferation, with knockdown of *ZFHX3* preventing TGFβ-induced inhibition of cell proliferation (M. Li et al., 2020). *Zfhx3/ZFHX3* enrichment in the ZL, association with OFC risk, and previously identified role in cell cycle gene transcriptional regulation, suggested that the cell cycle arrest observed during midface prominence fusion may be at least in part controlled by this transcription factor.

2.4 Discussion

A wealth of data has been generated in the field of orofacial cleft research both historically and in recent years. Older observations have been linked to specific genes and cellular behaviors through mouse modeling of craniofacial development and observation of genetic defects. With regards to OFC occurrence, the field has increasingly begun to differentiate between the etiology of cleft palate and cleft lip, as well as the roles of specific tissues in their development.

The role of the facial prominence epithelium, as in other systems, cannot be overlooked despite its small cell numbers. Here we have generated datasets of transcriptomic signatures for the epithelium of the lambdoidal junction from the E9.5 through E11.5 timepoints. Our analysis and validation of these clusters identifies 14 unique epithelial populations spread throughout the midface, with stage specific appearances of cells implicated in development of the eye ectoderm, periderm, nasolacrimal ducts, and olfactory tissues. We home in on the appearance of a novel population of cells that can be detected only starting at E10.5 and is associated with the facial prominence tips fusing at the lambdoidal junction. We call this population the "zippering lambda epithelium" based on its association with the fusion process. Through isolated analysis of the transcriptomic signatures, we show a contribution of cells from the midface prominence surfaces, olfactory epithelium, and oral cavity in the formation of this cluster. At both E10.5 and E11.5 timepoints we find heterogeneity within this population and identify cells implicated in previously known behaviors of the midface epithelium, such as EMT and apoptosis, as well as cell cycle arrest, a new and as yet uncharacterized behavior at the lambdoidal junction.

We characterize the zippering lambda epithelium and validate the expression of differentially expressed genes related to the cell cycle, finding a major association with cell cycle inhibitors such as *Gadd45g, Cdkn1a, Cdkn2b, Cdkn2a,* and *Rb1,* that is linked to the behavior of this epithelial cell population. We describe the cell cluster as non-proliferative through a series of *BrdU* chases and conclude that these cells largely do not re-enter the cell cycle, but instead maintain a cell cycle arrested state through midface prominence fusion at least until reformation of the epithelium around a completely fused primary palate.

We employ a combination of bulk and scRNAseq strategies to interrogate differences in the lambdoidal junction epithelium of CL/P models and find significant changes to the zippering lambda cell cluster formation and transcriptomic state. We uncover that the ZL cluster is most affected by the changes brought by the loss of *Pbx1* in the surface cephalic ectoderm on a *Pbx2* deficient background and find that transcripts related to proliferation and cell cycle arrest are among the most differentially expressed genes. In two models of CL/P involving loss of *Pbx1/2* and loss of *Trp63*, respectively, we uncover a complete loss or downregulation of cell cycle arrest genes within the lambdoidal ZL cluster. This discovery links control of cell cycle arrest to existing upstream regulators of midface fusion and their corresponding regulatory networks.

Our data sets enabled a targeted investigation into human midface development and OFC pathogenesis. We find that cell cycle arrest genes enriched in the ZL of the mouse are similarly expressed in lambdoidal junction ZL-epithelium in early human development, and therefore implicate cell cycle arrest across evolutionary time as a mechanism in mammalian primary palatogenesis. Our data sets also allowed us to cross-reference genes from each population of the epithelium identified in our scRNAseq analysis with OFC trio proband genome analysis and variant data. Through *de novo* and rare variant analysis, we implicate epithelial-enriched genes in OFC risk. More precisely, we uncover a strong link between genes enriched in the ZL cluster and established OFC risk variants by identifying novel OFC-associated variations tied to the ZL cluster gene expression signatures. We find that *ZFHX3,* a previously identified OFC risk variant, contains the highest number of OFC-associated variants among the top 10 genes enriched in the ZL cell cluster that bear variations in OFC probands. *ZFHX3* is implicated in the expression of cell cycle and cell cycle arrest genes, including those enriched at the ZL cluster. A combination of tissue section analysis and subsequent ongoing methods link *Zfhx3* in mouse to control of epithelial specific expression of cell cycle arrest genes enriched in the lambdoidal junction during prominence fusion. We suggest that *Zfhx3,* in collaboration with, or upstream of, *Pbx1/2* is part of a novel genetic pathway that drives a cell cycle arrested state during the prominence fusion process, enabling normal midface morphogenesis and the ultimate removal of the epithelium from the fusion site.

Epithelium specific roles are known to guide craniofacial outgrowth but are now shown to direct and orchestrate the intricate fusion event at the lambdoidal junction. As the field has turned towards identifying novel cleft risk candidates in human patients, likewise, a growing list of potential gene pathways and their molecular process outputs have been implicated in OFCs. As the field continues to unravel the relation between these pathways, we uncover novel ways in which canonical roles are co-opted to unique events. Through our targeted specific approach to scRNAseq of the midface epithelium, we achieve a fine dissection of this embryonic tissue, its heterogeneity, and changing transcriptome throughout fusion. We uncover the identity of a population of cells at the prominence fusion site we dub the “zippering epithelium” for its association with lambdoidal junction fusion and describe a defining characteristic of cell cycle arrest. It was previously reported that cell cycle arrest can mediate developmental processes, but it was never shown that this cell behavior is critical for primary palate fusion. Through analysis of CL/P models we implicate dysregulation of ZL cluster cell cycle arrest in failure of midface prominence fusion. We employ a combination of various developmental biology and genomics methods to identify cell cycle arrest in human midface morphogenesis and OFC risk. A novel risk candidate, *ZFHX3,* is demonstrated to be associated with the ZL cell cluster and provides novel insight into an as yet unknown mechanism of cell cycle arrest during face prominence fusion.

2.5 Materials and Methods:

Cell Isolation and Fluorescence Activated Cell Sorting

Swiss Webster mice timed matings were used to determine the gestational day of embryonic development where evidence of a copulatory plug was considered day (E) 0.5. Embryonic litters were harvested and staged using somite count and limb morphology at the respective days of development and dissected into ice-cold PBS. E9.5 embryonic faces were dissected from the heads by dissection of the frontonasal prominence and maxillary prominence, taking care to avoid forebrain and developing eye placode. E10.5 and E11.5 embryonic faces were dissected away from the head as intact midfaces encompassing the MNP, LNP and MxP avoiding the eyes and brain tissues. The lambdoidal junction was isolated using surgical scalpels to dissect away the upper halves of the LNP and MNP at the horizontal midline of the nasal pits, and the lower half of the MxP at the horizontal midpoint, taking care to keep the lambdoidal junction intact without separating the individual prominences. For wild-type scRNAseq individuals were pooled together for dissociation to single cell suspensions. A 1X Liberase TL (Roche #05401020001) and 1X DNAseI Grade II (Roche #10104159001) cocktail was prepared fresh. Tissue samples were incubated in the enzymatic cocktail at 37°C without agitation for 10-13 minutes, with gentle pipetting at 5-minute intervals to break apart tissues. Enzymatic dissociation was halted using ice cold PBS. Cells were spun at 300 RCF for 10 minutes and resuspended in ice cold cell staining buffer (0.5% Bovine Serum Albumin, 2mM EDTA in PBS). 1.25µL of anti-Epcam-APC (Invitrogen) antibody was incubated with dissociated cells for 15 minutes at 4°C in darkness, before spinning down and resuspending in FACS buffer (5% FBS, 5mM EDTA in PBS). DAPI was added to cell suspension at a 1:1000 ratio.

Cells were kept chilled on ice and sorted using a Sony SH800S sorter. Gating was performed to eliminate doublets and dead cells. APC positive cells were collected at over 10^4 fluorescent units into chilled FACS buffer.

Single cell RNA sequencing

Tissues were dissociated to single cells using an enzymatic cocktail of Liberase and DNase I for 10 min at 37 degrees. Cells were passed through a 45 µm strainer to remove aggregates from the single-cell suspension. Dead cells (≤20%) were eliminated with a ‘Dead Cell Removal Kit’ with magnetic beads (MACS, Milteny Biotech). Live cells from pooled lambdoidal junctions were loaded into one well for single-cell capture using the Chromium Single-Cell 3′ Reagent Kit V2 (10X Genomics). Libraries were prepared using the Chromium Single-Cell 3′ Reagent Kit V2, and each sample was barcoded with a unique i7 index. Libraries were pooled and sequenced using an Illumina NovaSeq sequencer.

Data Analysis

The Cell Ranger v2.2.0 pipeline from 10X Genomics was used for initial processing of raw sequencing reads. Briefly, raw sequencing reads were demultiplexed, aligned to the mouse genome (mm10), filtered for quality using default parameters, and UMI counts were calculated for each gene per cell. Reads were filtered for uniquely mapped and saved in BAM file format with count matrices produced. Main analysis of the count matrix was mostly performed using the pagoda2 R package. PCA was performed on the overdispersed genes, and XX PC were retained using elbow curve selection. UMAP visualization was generated in PCA space. Clusters were identified on KNN graph using leiden algorithm (pagoda2), and pathway overdispersion analysis performed to identify relevant biological aspects. Differential gene expression was performed for each detected cluster using Wilcoxon rank testing. RNA velocity analysis used BAM files from each dataset and were processed using python command line velocity tool. Loom files for each dataset contained spliced and unspliced transcript counts which were combined into a single loom dataset and filtered according to cells kept in final dataset. The scvelo python tool was used to exclude genes with <20 spliced counts or genes with <10 unspliced counts and the 4000 top highly variable genes were kept. PCA was performed on the spliced matrix with the 30 PC kept and KNN neighbor graph produced. Moments of spliced/unspliced abundances, velocity vectors and velocity graphing were computed using default parameters. Velocity states were projected on the UMAP embedding produced during initial analysis step.

Bulk RNA sequencing

Both epithelium and mesenchyme were sorted using FACS (Sony SH800S) as described above. Epcam positive and negative cells were collected directly into separate tubes of RLT RNA extraction buffer. 6 separate litters of *Crect*$^{Cre/+}$*; Pbx1*$^{f/+}$ to *Pbx1*$^{f/f}$*; Pbx2*$^{-/-}$ crosses were used to obtain material from E11.5 midfaces. After genotyping of individual embryos, RNA from 4 embryos of the respective groups was pooled together: mutant (*Crect*$^{Cre/+}$*; Pbx1*$^{f/f}$*; Pbx2*$^{+/-}$), heterozygous controls (*Crect*$^{Cre/+}$*; Pbx1*$^{f/+}$*; Pbx2*$^{+/-}$, and *Cre*-negative controls (*Crect*$^{+/+}$*; Pbx1*$^{f/f}$*; Pbx2*$^{+/-}$ or *Crect*$^{+/+}$*; Pbx1*$^{f/+}$*; Pbx2*$^{+/-}$). RNA was extracted using the RNeasy Plus Micro Kit (Qiagen, #74034), and total RNA quantifed using Qubit RNA HD Assay Kit (Invitrogen, Q32852). RNA quality was determined with an RNA 6000 Pico kit (Agilent, Cat: 5067-1513) on a 2100 Bioanalyzer (Agilent). All RNA samples for library preparation had an RIN>9. PolyA RNAs were captured using an NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB, #E7490). RNA sequencing libraries were prepared from 100ng of RNA using NEBNext Ultra™ II RNA Library Prep Kit for Illumina (NEB, #E7775). Library size and quality was checked using an Agilent 2100 Bioanalyzer with the High Sensitivity DNA kit (Agilent, Cat: 5067-4626). Library DNA concentrations were determined with the Qubit dsDNA HS Assay Kit (Invitrogen, Q32854). Libraries were sequenced with the Illumina HiSeq 4000 to generate 50 base pair (bp) single-end reads. Reads for each tissue were mapped against the mouse genome (mm10) using the Tophat2 aligner (version 2.0.13) with default parameter settings except for setting the flag –no-coverage-search. Expression levels for each tissue were initially quantified using htseq-count. Differential expression analyses of mutant versus *Cre*-negative controls was performed using Deseq2. Genes with a fold change ≥1.2 or ≤-1.2 and an FDR≤0.05 were defined as differentially expressed genes (DEGs).

Whole Mount In Situ Hybridization

Embryos were dissected whole from timed pregnancies and fixed O/N at 4°C in 4% PFA. Embryos were dehydrated using a methanol gradient series to 100% pure MeOH. Embryos were stored immersed in 100% MeOH -20°C for 6 months to a year before rehydration through a reverse methanol series to 100% PBS. Holes were punched through the back of the cranium to the hindbrain before *in-situ* hybridization to prevent trapping of probe in head regions in embryos older than E10.5. Whole-mount in situ hybridization was performed as described(Capellini et al., 2006, 2008). In brief, embryos were rehydrated and pretreated with Proteinase K, and then hybridized O/N at 70 °C with either sense or antisense riboprobes at a final concentration of 1 µg/ml in incubation buffer containing 50% formamide, 5× SSC, 50 µg/ml yeast RNA, 1% SDS, 50 µg/ml heparin, and 0.1% CHAPS detergent (ThermoFisher Scientific). In situ hybridization probes were those used by Capellini et al. (Capellini et al., 2006). Embryos were then washed through a series of SSC solutions (5× SSC and 2× SSC, three times each for 30 min, and one time each in 0.2× SSC and 0.1× SSC for 30 min, respectively) at 70 °C. After a brief rinse in Tris-buffered saline/0.1% Tween (TBST), embryos were incubated in 10% blocking reagent (ThermoFisher Scientific, # R37620) as described and the positive signals were detected by AP-conjugated anti-digoxigenin antibody (Roche Diagnostics GmbH, Mannheim, Germany) at 1:5000 dilution. Following washing in TBST, embryos were incubated in NBT/BCIP in NTMT buffer (Roche Diagnostics GmbH, Mannheim, Germany) following the manufacturer instructions, until color fully developed. Positive hybridization was visualized by purple (NBT/BCIP) signal. At least 3 embryos for each genotype and developmental stage were analyzed to establish reproducibility.

Thymidine Analogue Incorporation and Staining

BrdU was diluted in saline solution (0.9%). E10.5 pregnant Swiss Webster dams were anesthetized using isoflurane gas inhalation. Intraperitoneal injection of BrdU solution was administered under anesthesia followed by observation and recovery. Pregnant dams were sacrificed at 1, 7, and 24 hours after BrdU injection. Embryos were collected and fixed in whole in 4% PFA solution overnight at 4°C with gentle agitation. Fixed embryos were washed with PBS and allowed to sink in 30% sucrose solution before embedding in 100% Neg-50 frozen cryoblocks and storing at -80°C. Cryosectioning of embryos was performed as with IF and RNAscope procedures. Slides were blocked using blocking buffer (10% Normal Donkey Serum, 0.3% Triton X-100) 1hr room temperature (RT). Non-BrdU antibodies were incubated overnight 4°C in antibody buffer (2% NDS, 0.1% Triton X-100 in PBS). Secondary antibodies were applied at 1:500 the next day following washes in using PBS (0.1% BSA) and incubated 2 hours at RT before washing again. 2M HCl solution was applied to slides for 30 min at 37°C. Sides were washed and reblocked with blocking buffer for 1 hour at RT. Rat anti-BrdU (abcam ab6326) antibody was diluted (1:500) in antibody buffer and incubated overnight. The following day, washes and anti-Rat secondary antibodies (1:500) were incubated along with DAPI (1:1000) counterstain before final washes and mounting with Prolong Gold antifade.

Mice

All animals were maintained according to UCSF IACUC guidelines. Breeding to generate mutant mice alleles was done using mating of stud males to female dams. Evidence of a copulatory plug was used to determine the staging. Pregnant females were euthanized using IACUC approved humane methods, and embryos removed from dissected uteri for further sample collection.

Pbx1/2 conditional mutant embryos were generated using a cross between males bearing *Crect*$^{cre/+}$ and *Pbx1*$^{f/+}$ alleles in heterozygosity and females homozygous for both conditional *Pbx1*$^{f/f}$ and global null *Pbx2*$^{-/-}$ alleles. *Pbx1/2gKO* mutants were generated using crosses between males and females that bear *Pbx1*$^{+/-}$ and *Pbx*$^{+/-}$ constitutive null deletions in heterozygosity. *Crect*$^{Cre/+}$ mice were maintained in heterozygosity on an FVB background via backcrosses to FVB wild-type mice. *Trp63*^{tm1Fmc} mice are maintained in heterozygosity. Heterozygous individuals are mated to produce homozygous "null" mutant embryos (*Trp63*$^{-/-}$*)*.

RNAscope Fluorescent In Situ Hybridization and Immunofluorescence

Frozen blocks were at 14-µm-thick cryosections that were air dried at −20 °C for 1 h and stored at −80 °C. Slides carrying hindlimb bud sections were thawed and washed with 1X PBS to remove excess freezing medium before use. Slides were assayed using an RNAscope™ Multiplex Fluorescent Reagent Kit V2 following modified manufacturer's instructions protocols. Antigen target retrieval steps were skipped, and protease treatment limited to treating slides with Protease Plus for no longer than 10 min to avoid fragile embryonic tissues damage. Probe mixes were hybridized for 2 h at 40 °C in a HybEZ™ II Oven (Advanced Cell Diagnostics, Newark, CA). The appropriate HRP channels were developed with Opal™ 520, TSA™ Cy3 Plus, and Cy5 Plus (PerkinElmer) dyes. Following DAPI staining and mounting with ProLong™ Gold (Invitrogen).

For sections involving immunofluorescence in conjunction with RNAscope, immunofluorescence assays were carried out after completion of the last RNAscope step involving blocking of the final HRP channel. A blocking solution was applied to slides containing 5% Normal Donkey Serum, 0.1% Tween-20 in 1X PBS for 45 minutes. Following this, antibodies raised against KI67, EPCAM, and P21 were diluted in an antibody buffer (1:5

dilution of blocking solution to PBS-0.1% Tween-20 solution) and applied to the slides O/N at 4°C. The following day, primary antibodies were washed away and the slides washed 3 times 15 minutes each using antibody buffer. Secondary antibodies were diluted in antibody buffer (1:500) along with DAPI (1:1000) and applied to the slides for 2 hours at room temperature. Following washes to remove secondary antibody, slides were flicked dry and mounted using Prolong Gold antifade and coverslips.

Gene Ontology Analysis

The top 100 differentially expressed genes ordered by absolute Z-score from E10.5 epithelium datasets was used as input into the NIH Database for Annotation, Visualization, and Integrated Discovery (DAVID) functional annotation tool. 100 gene DAVID ID matches to *M. musculus* were charted using the GOTERM_BP_DIRECT category and ranked by P-value.

Human Embryo Samples

Human embryonic tissues were acquired by the Human Developmental Biology Resource (HDBR) after voluntary pregnancy terminations with informed donor consent (a written consent was obtained for each case). Ethical approval from the Newcastle and North Tyneside 1 National Health Service (NHS) Health Authority Joint Ethics Committee (08/H0906/21+5).

Human embryo samples at CS16 and CS17 were paraffin-embedded sections fixed with 4% paraformaldehyde. Antigen retrieval methods were carried out according to manufacturer guidelines for RNAscope Multiplex Fluorescent Reagent Kit v2 (Advanced Cell Diagnostics), followed by RNAscope probe manufacturer’s manual for incubation and detection. Immunofluorescence was performed immediately following RNAscope Multiplex assay utilizing ACD biotechne recommended IHC methods.

Microscopy

Whole embryos were imaged using a Leica M20F5A dissecting microscope with a 1x/0.03 lens and variable zoom settings. Images were processed and exported using LAS-X software package for all Leica microscope images. Sections were imaged using a Zeiss Axio Observer.Z1 or a Zeiss Axiovert 200M with a Plan-Apochromat 20x/0.8 objective. Several Z-slices of sections were acquired. Sections with visible focused expression of RNAscope and immunofluorescence signal, were individually chosen, as the two were often at differing focal ranges. Images were processed using ZEN 3.6 or ZEN 2.3 Pro. Adjustments were made to reduce the background autofluorescence of the tissues by adjusting the absolute black and absolute white levels using a Best Fit adjustment where necessary.

Table of Primers Used

Primers to generate WISH riboprobes:

Cdkn1a/p21 F: CTCTTCCCCATCTTCGGCC

Cdkn1a/p21 R: GAGACGCTTACAATCTGAGTGG

Gadd45g F: CCGATGAAGAAGATGAGGGCG

Gadd45g R: TGAAAGAGCAGTGCAGTCGG

Bmp4 WISH probe was provided by Ian C. Welsh (see Welsh and O'Brian 2009)

List of Probes and Antibodies Used

RNAscope catalog probes (Advanced Cell Diagnostics): Gadd45g (803431-C2), Cdkn2a (411011-C3), Cdkn2b (458341-C2), Bambi (523071), Pbx1 (435171), Foxn3 (586011), Bmp4 (401301-C2), Cdkn1c (458331-C2), Pitx2 (412841-C2), Aurkb (461761), Zfhx3 (803471), Rb1

(486191-C3), Bambi (523071-C3), Igfbp5 (425731-C2), Cdk1 (476081-C2), Cdkn1b (499991), Itm2b (491791), Tfap2b (536371-C3), Sox2 (401041-C3), Hey1 (319021-C3), Dlk1 (405971-C2), Fgf9 (499811), Wnt6 (401111).

TUNEL-TMR Red Kit (Roche #12156792910).

Antibodies used: Anti Active Caspase3 (Promega G748A), Mki67 (abcam ab15580), Tubb3 (abcam ab18207), Epcam-APC (Invitrogen 17-5791-82), Anti P21/Waf1/Cip1 HUGO291 (MABE1816-100UG), Epcam (DSHB G8.8).

OFC trio samples

Case-parent trios with whole genome sequencing data originating from three ancestral groups were used for this study. The first set consists of 376 trios of European ancestry that were recruited from sites in the United States, Argentina, Turkey, Hungary, and Spain; the second is a set of 267 trios from Medellin, Colombia; and the third is a set of 116 trios from Taiwan. In 93.8% of European, 96.5% of Taiwanese, and 100% of the Colombian parents were unaffected. Because OFC etiology is likely to be multifactorial, parent phenotype status was not considered in this analysis. By proband OFC type, there were 88 cleft lip only (CL; 8 Colombian, 80 European), 613 cleft lip and palate (CLP; 259 Colombian, 238 European, 116 Taiwanese), and 58 cleft palate (CP; all European).

Subject recruitment and phenotypic assessment occurred at regional treatment centers. Each site's institutional review board (IRB) and the IRBs of the affiliated US coordinating institutions (HRPO #03-0871, IRB#HSC-MS-03-090, IRB#970405, IRB#200109094, and IRB#200109094) provided ethical approval and oversight.

Autism spectrum disorder samples

De novo variants called from 6,430 probands with autism spectrum disorder (ASD) and 2,179 unaffected siblings were obtained from a publicly available dataset. (Satterstrom et al., 2020)

Whole genome sequencing and variant calling

The case-parent OFC trios used in this study were sequenced as part of the Gabriella Miller Kids First (GMKF) Research Consortium. Sequencing was performed on blood samples when available (the majority of samples) and saliva when blood was not obtainable. Whole genome sequencing for European samples was carried out by the McDonnell Genome Institute (MGI) the Washington University School of Medicine (St. Louis, MO) followed by alignment to hg38 and variant calling at the GMKF Data Resource Center at the Children's Hospital of Philadelphia. Sequencing for Colombian and Taiwanese samples was carried out by the Broad Institute, with alignment to hg38 and variant calling by GATK pipelines. Additional details on alignment and workflow used to harmonize these datasets have been published. (Mukhopadhyay et al., 2020)

Quality control

The WGS data for all case-parent trios was subjected to several quality metrics. Individual samples were evaluated for missingness, Mendelian error rate, and average read depth, and were removed if these values were greater than three standard deviations from the mean. Additionally, samples with transition/transversion (Ts/Tv), exonic Ts/Tv, silent/replacement, or heterozygotes/homozygotes ratios outside of the expected values were removed. Due to trios within the Colombian cohort having higher rates of consanguinity than other groups, a lower ratio for heterozygotes/homozygotes ratios was allowed. Identity by descent as tested in PLINK (version 1.90b53) was used to confirm familial relationships, and sample sex was confirmed by X chromosome heterozygosity.

Identification of de novo variants

Mendelian errors were called in trios using PLINK (version 1.90b53) and bcftools (v1.9). These mendelian errors were then underwent further filtering to yield high quality *de novo* mutation calls including filtering for passing variants, bi-allelic variants only, a minor allele count (MAC) = 1, genotype quality (GQ) $\geq$ 20, and depth (DP) $\geq$ 10 using VCFtools (version 0.1.13). Furthermore, we filtered variants on the basis of allele balance (AB), with *de novo* calls requiring an AB ratio $\geq$0.30 and $\leq$0.70 in the proband and an AB ratio $<$ 0.05 in the parents. Annotation of high confidence DNMs was completed with ANNOVAR (version 201707). Following annotation, *de novo* variants in coding regions were selected based on functional classification ("exonic" and/or "splicing") and frequency (MAF < 0.3% across all of gnomAD v3.1.2)

De novo variant enrichment

Enrichment of *de novo* variants (DNs) in 759 OFC trios, 6,430 ASD probands, and 2,179 unaffected siblings was statistically analyzed using the 'DenovolyzeR' package (version 0.2.0) in R. Enrichment is tested by determining if the number of observed variants in a dataset is greater than what is expected based on mutational models described by Samocha, et al. (6). The functions 'DenovolyzeByClass' and 'DenovolyzeByGene' were used to test for an excess of DNMs both globally and per gene, respectively; however, use of the function 'includeGenes' was implemented to restrict testing to the gene sets derived from single cell RNA sequencing only. When testing enrichment per cluster, results were considered significant at $p<0.0083$ (Bonferroni correction 0.05/6 clusters).

Rare variants in top scRNAseq genes

The top 150 genes ordered by absolute Z-score of each scRNAseq cluster were converted from *Mus musculus* to *Homo sapiens* orthologues using g:Profiler, then coding variants in orthologous genes were extracted for 876 families with orofacial clefts (OFCs). First, synonymous variants were removed so that only protein-altering variants remained, then variants were filtered for minor allele frequency (MAF) of <0.5% in any population in both gnomAD v2.1.1 and v3.1.2. We used ASD probands in comparison, for which the obtained DNs underwent the same filtering steps. Qualitatively, we evaluated variants based on multiple pathogenicity predictors. We excluded variants with CADD scores <20 and SIFT scores >0.05 and compared overlaps of variants in OFC cases within both constrained (LOEUF <0.35 and pLI >0.9) and non-constrained marker genes. We then tested OFC and ASD gene overlaps with the top marker genes using GeneOverlap (version 1.26.0). Given a relatively limited sample size, we also tested overlap of marker genes with 150 genes with known associations with OFCs. For *ZFHX3*, we visualized predicted damaging variants in OFC probands with ProteinPaint.

2.6 Figures

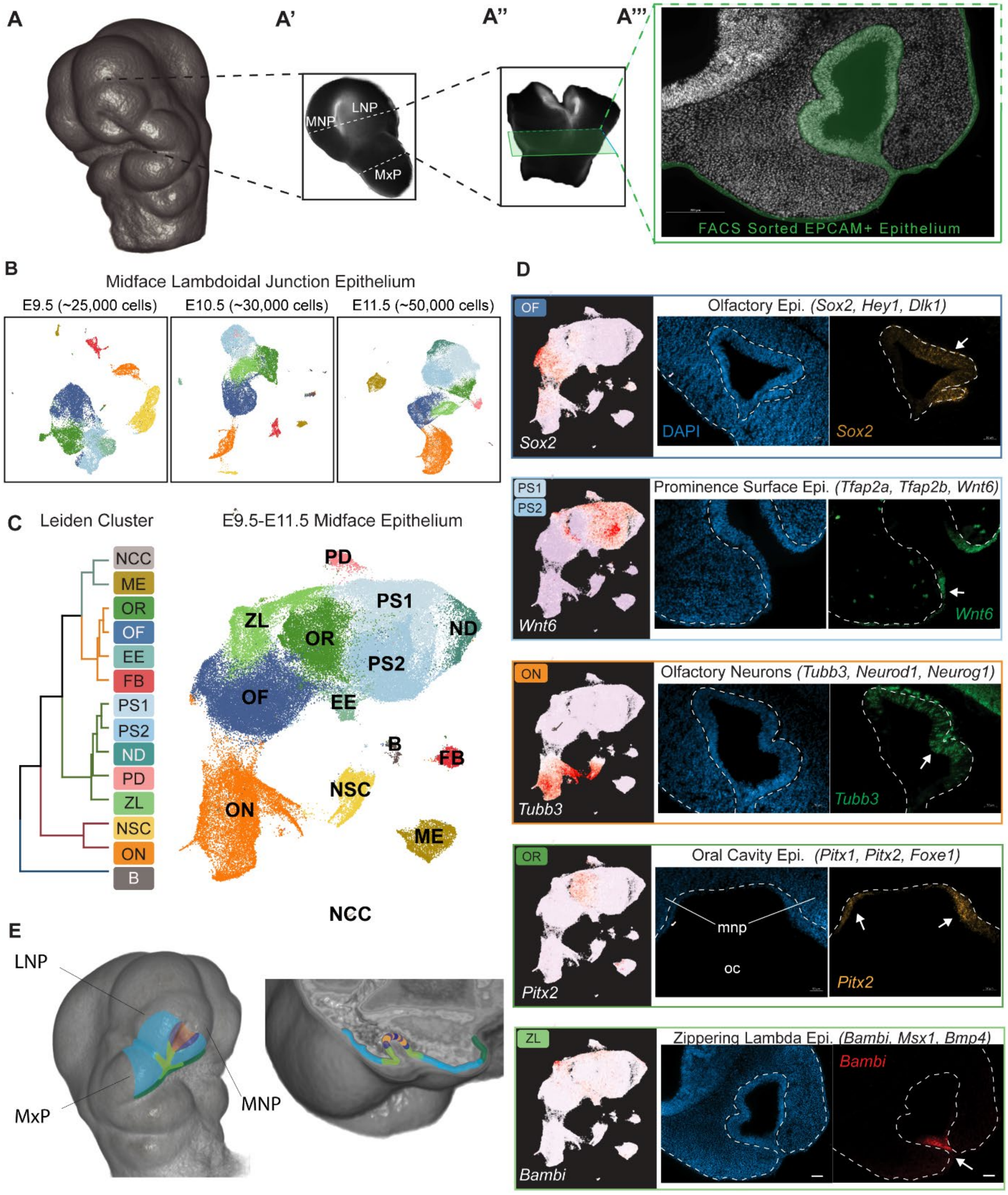


Figure 2.6.1: scRNAseq of Midface λ Epithelium Highlights Cells Heterogeneity and Spatial Localization of Select Cell Clusters
A: Experimental design and schematic of epithelial cell isolation from mouse midfacial embryonic prominences. A': Representative (Figure caption continued on the next page.)

(Figure caption continued from the previous page.) microdissection of E10.5 half-midface containing medial nasal prominence (MNP), lateral nasal prominence (LNP), and maxillary prominence (MxP), with dashed lines indicating further dissection of midface to lambdoidal junction-centric region. A”: Dissected lambdoidal junction used as basis for FACS sorting of fusion site epithelium. Green square indicates cross section of lambdoidal junction. A’”: DAPI stained section of midfacial MNP and LNP prominences (14um) with green highlight indicating epithelial cell population targeted for isolated by FACS in scRNAseq experiments. B: scRNAseq clustering uniform manifold approximation projection (UMAP) of corresponding wild-type (SW) epithelium of E9.5, E10.5, and E11.5 lambdoidal junction with captured cell numbers indicated. Clusters assigned across timepoints are consistent using Seurat label transfer C: Dendrogram and UMAP of all Leiden algorithm defined clusters across three timepoints E9.5-E11.5. A total of 14 clusters were called across the three stages. Clusters that had *Epcam* positivity were investigated for defining markers and named according to cellular phenotype and location. D: Validation of 5 largest clusters across datasets was performed using RNAscope fluorescent *in-situ* hybridization (FISH) and immunofluorescence. These 5 main clusters are shown with 3 representative gene markers. Normalized gene expression (NGE) per cell of a single cluster defining marker is superimposed on UMAP (left). A representative validation at the E11.5 timepoint demonstrates gene expression *in-vivo* with (center) DAPI stain and with (right) RNAscope FISH probes against marker of interest. Dashed lines indicate boundary of epithelium and mesenchyme. Arrows indicate areas of interest, (oc) oral cavity. E: Color-coded schemata of midfacial prominences in whole and cross section using µCT generated model (Facebase) with indicated boundaries of validated epithelium specific clusters at E10.5 (left). Transverse cutaway of model to show nasal pit and midface detail (right). Pale blue: prominence surface epithelium (PS1 and PS2). Dark blue: olfactory epithelium. Orange: olfactory neurons. Dark green: oral cavity epithelium. Light green: zippering lambda epithelium.

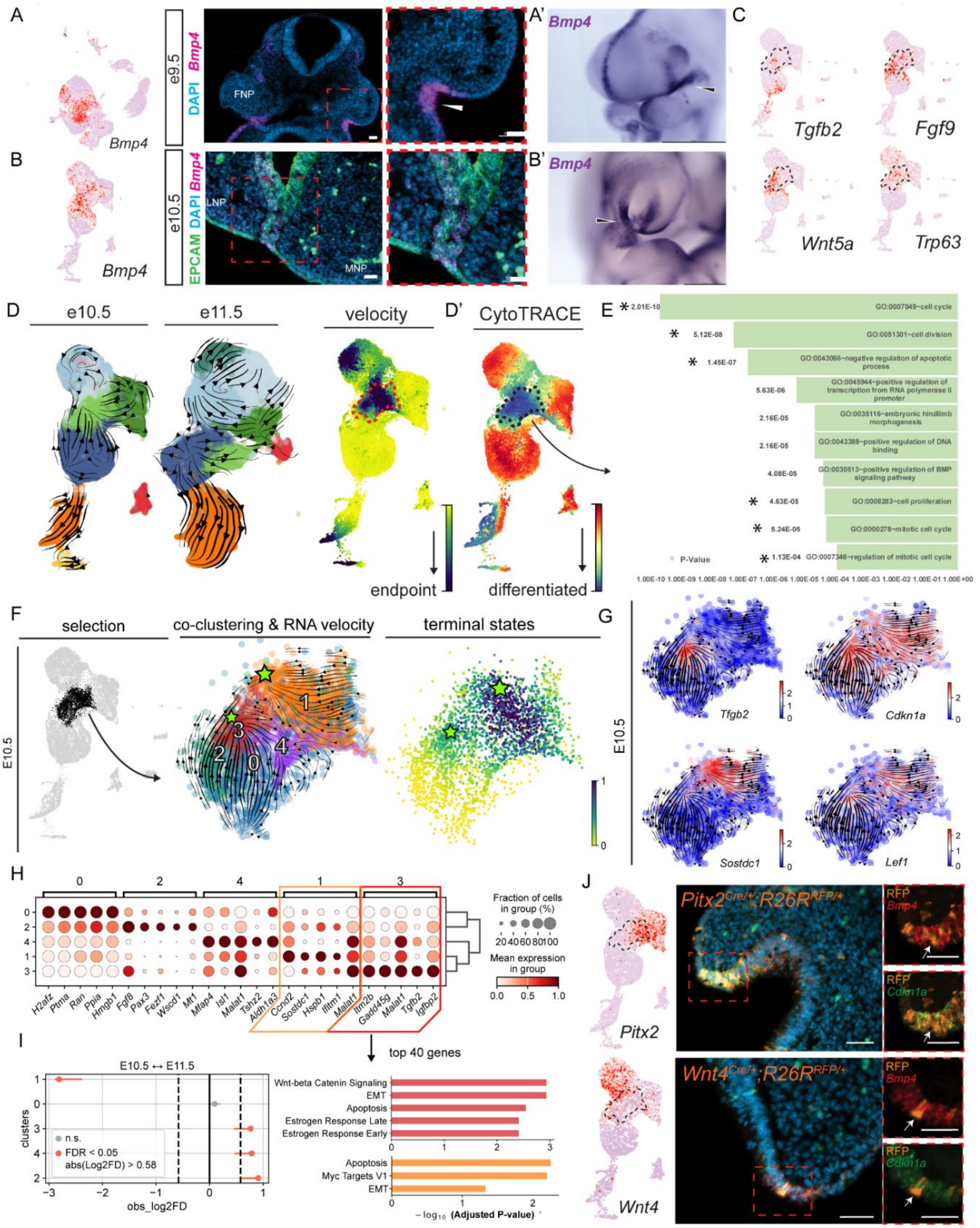


Figure 2.6.2: Molecular Characterization of the λ Zippering Epithelial Cell Cluster and In-vivo Cell Fate Mapping

A: *Bmp4* normalized gene expression (NGE) on E9.5 UMAP indicating presence of a localized *Bmp4* high region without cluster specificity (Figure caption continued on the next page.)

(Figure caption continued from the previous page.) and prior to ZL cluster formation (left). Transverse section of E9.5 head with localization of *Bmp4* expression shown using RNAscope validation of *Bmp4* expressing cells under FNP in wild-type e9.5 embryos (center). Magnified region of interest (right). Scale bars: 50 µm. A': traditional chromogenic whole-mount *in-situ* validation of predicted *Bmp4* expression at E9.5. Scale bar: 500 µm. B: One day of development later, at E10.5, *Bmp4* expression acts as marker of midface prominence tips and zippering lambda epithelium at the site of fusion. Superimposed NGE of known developmental regulators of midfacial growth and morphogenesis on E10.5 UMAP (left). Dashed black border indicates zippering epithelial cell cluster boundaries. RNAscope FISH validation of *Bmp4* expression in transverse sections of lambdoidal junction at E10.5 (center). Inset to show detail of fusion site with high *Bmp4* expression limited to epithelium (marked by Epcam expression) and fusing junction. Scale bars: 50 µm. B' Whole mount *in-situ* hybridization of *Bmp4* at E10.5 timepoint using chromogenic assay. Scale bar: 500 µm. C: Normalized gene expression heatmaps overlaid on UMAP highlight enrichment of genes known to be associated with both midface development and CL/P risk in the zippering epithelial cluster. TGFB, FGF, and WNT signaling create a known pathway implicated in fusion and apoptosis of midface epithelium along with downstream effector Trp63. Dashed black border indicates zippering epithelial cell cluster boundaries. D: E10.5 RNA velocity analysis using *scvelo* of spliced versus unspliced transcripts in E10.5 (left) and E11.5 (right) scRNAseq datasets. A distinct pattern of converging vectors from clusters neighboring zippering lambda epithelium (light green) suggest a lineage of epithelial cells from oral cavity epithelium, prominence surface epithelium, and olfactory epithelium contribute to a transcriptomic state localized at the fusion site. A weakened contribution from PS cells to this state is predicted by E11.5 while OE and OR retain strong vectors. E10.5 velocity endpoints are found to be strongest in three clusters, PS2, ZL and ON. PS2 cells appear to be in a cycling state with partner cluster PS1. ON cell endpoints highlight olfactory neurogenesis reaching terminal differentiation. ZL endpoint lies in convergence zone of neighboring three population trajectories. D' CytoTRACE (Cellular (Cyto) Trajectory Reconstruction Analysis using gene Counts and Expression) computation of E10.5 main clusters labels cells at the end of ON neurogenesis tract and ZL cluster (at the lambdoidal fusion site) as most differentiated based on predicted order of gene counts per single cell. E: Gene ontology (biological processes) terms of differentially expressed gene ZL cluster with false discovery rate correction. Asterisks indicate notable terms of interest. F: E10.5 ZL cluster was re-clustered in isolation to reveal 5 subclusters. RNA Velocity analysis performed on this isolated cluster highlights transitions from surrounding transcripts into terminal states enriched for biological processes known to be involved in midface fusion. Green stars indicate two terminal endpoints predicted for isolated ZL RNA velocity vectors located in subclusters 1 and 3. G: Normalized gene expression and RNA velocity overlay on isolated ZL UMAPs at E10.5. Genes enriched at the terminal endpoints are highlighted including *Sostdc1,* an agonist of *Bmp4* signaling, *Tgfb* and *Lef1,* previously known to be implicated in *Pbx*-mediated apoptosis signaling pathway in the lambdoidal junction, and *Cdkn1a*/*p21* a marker of cell cycle arrest. H: Mean expression of top 5 differentially expressed genes in isolated ZL subclusters. Subclusters 1 and 3, which contained RNA velocity predicted endpoints, are highlighted. Top 40 DEGs of subclusters 1 and 3 used for GO term analysis (GO biological process complete) plotted by adjusted P-value. In both cases, EMT and apoptosis, hallmarks of fusing lambdoidal junction epithelium, are top terms. (Figure caption continued on the next page.)

(Figure caption continued from the previous page.) I: For subcluster of ZL epithelium, changes in proportions were assessed between the E10.5 and E11.5 timepoints in which fusion of the lambdoidal junction is primarily occurring. Notably reduction in subcluster 1 and increases in 3 suggests a transition to increased removal of epithelial cells in the ZL via defined mechanisms of Tgfβ driven apoptosis and EMT. J: Validation of RNA Velocity predictions through genetic lineage tracing experiments using *cre* recombinase mouse lines corresponding to cluster defining markers of predicted prominence surface *(Wnt4)* and oral cavity *(Pitx2)* epithelium. (top) NGE overlap of *Wnt4* and *Pitx2* at E10.5 with dashed border indicating ZL cluster boundaries. Both *Wnt4-cre* and *Pitx2-cre* lineage tracing with an *RFP* reporter demonstrate considerable overlap of ZL cluster markers *(Bmp4* or *Bambi)*, indicating contributing lineages at E10.5.

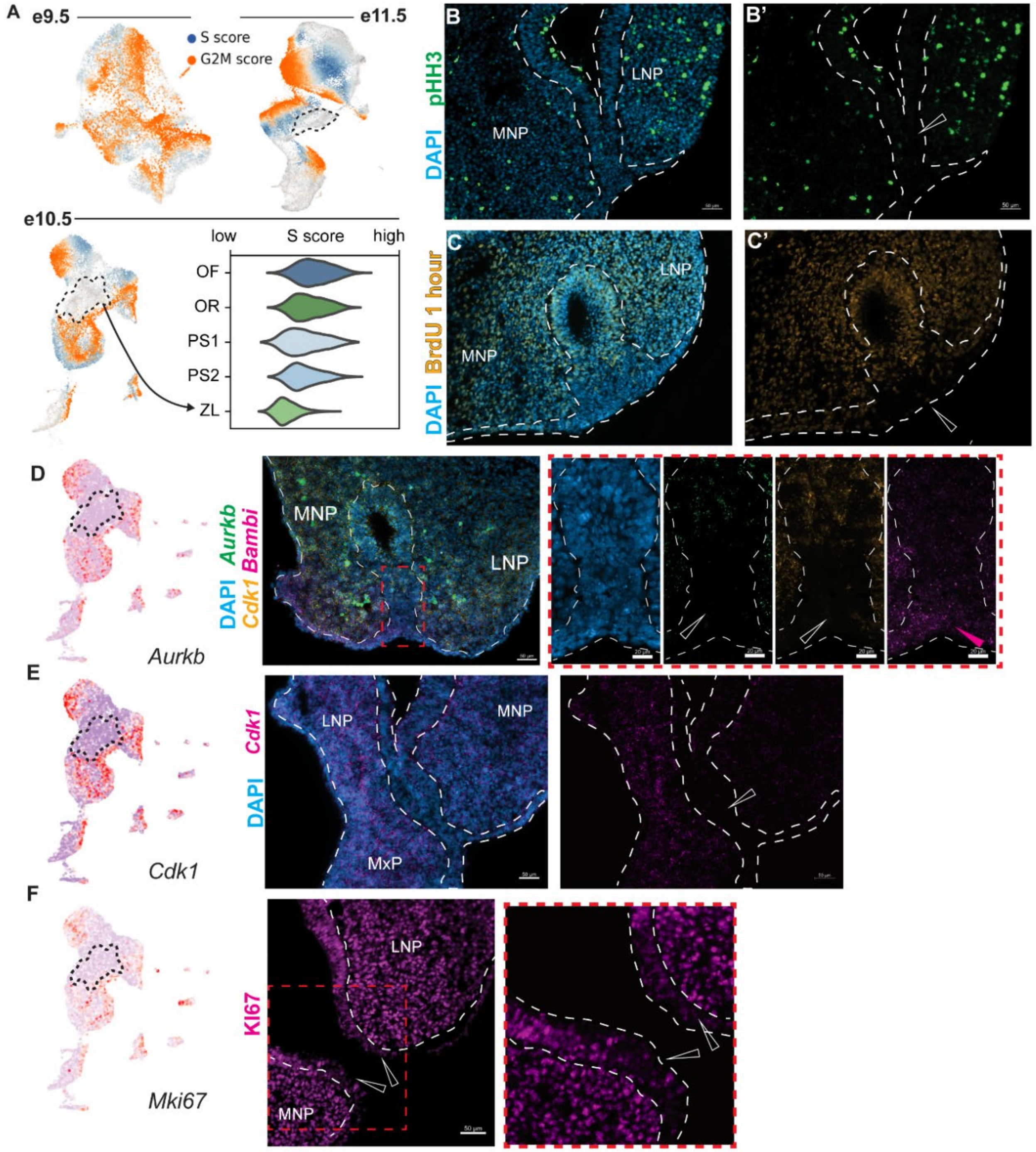


Figure 2.6.3: Cells Comprising the λ Epithelium Zippering Cluster Do Not Proliferate and are Devoid of Cell Cycle Progression Markers

A: Scoring and cell cycle phase assignment of e9.5-E11.5 timepoints using the scvelo cell cycle scoring function demonstrates mixed scores for S and G2M in all clusters at E9.5, and the appearance of a persistently cell cycle low score in the ZL and ON clusters beginning at E10.5 through E11.5. Violin plot of scvelo S-phase scoring at E10.5 for ZL and neighboring clusters. ZL epithelium ranks last in S-score among main clusters suggesting a low rate of proliferation and cell cycle progression. B-B': phosphorylated histone H3 (pHH3) immunofluorescence assay indicates a reduction of mitotic cells at (Figure caption continued on the next page.)

(Figure caption continued from the previous page.) the fusion site in contrast to nearby nasal pit epithelium and underlying mesenchyme. C-C': Following intraperitoneal injection of BrdU thymidine analogue, embryos at E10.5 were collected and anti-BrdU immunofluorescence performed in transverse sections. Minimal integration of BrdU at the ZL epithelium was found compared to surrounding cells. D-F: Other markers of proliferation corresponding to different stages of the cell cycle are shown to be absent from ZL epithelium and reduced in E10.5 scRNAseq datasets in the ZL cluster (left). Black dashed line defines ZL cluster boundary. D: RNAscope FISH in transverse section against *Cdk1* and *Aurkb* overlap region enriched in region of ZL DEG *Bambi.* Red box inset to show detail; scale bars 20µm. E: *Cdk1* RNAscope in coronal section demonstrates absence of gene involved in cell cycle progression from fusing lambdoidal seam. F: Immunofluorescence targeting Mki67 reveals lowered expression at the fusion site between LNP and MNP. Red box inset to show detail. Scale bars set to 50µm. White dashed lines denote edges and border of epithelium.

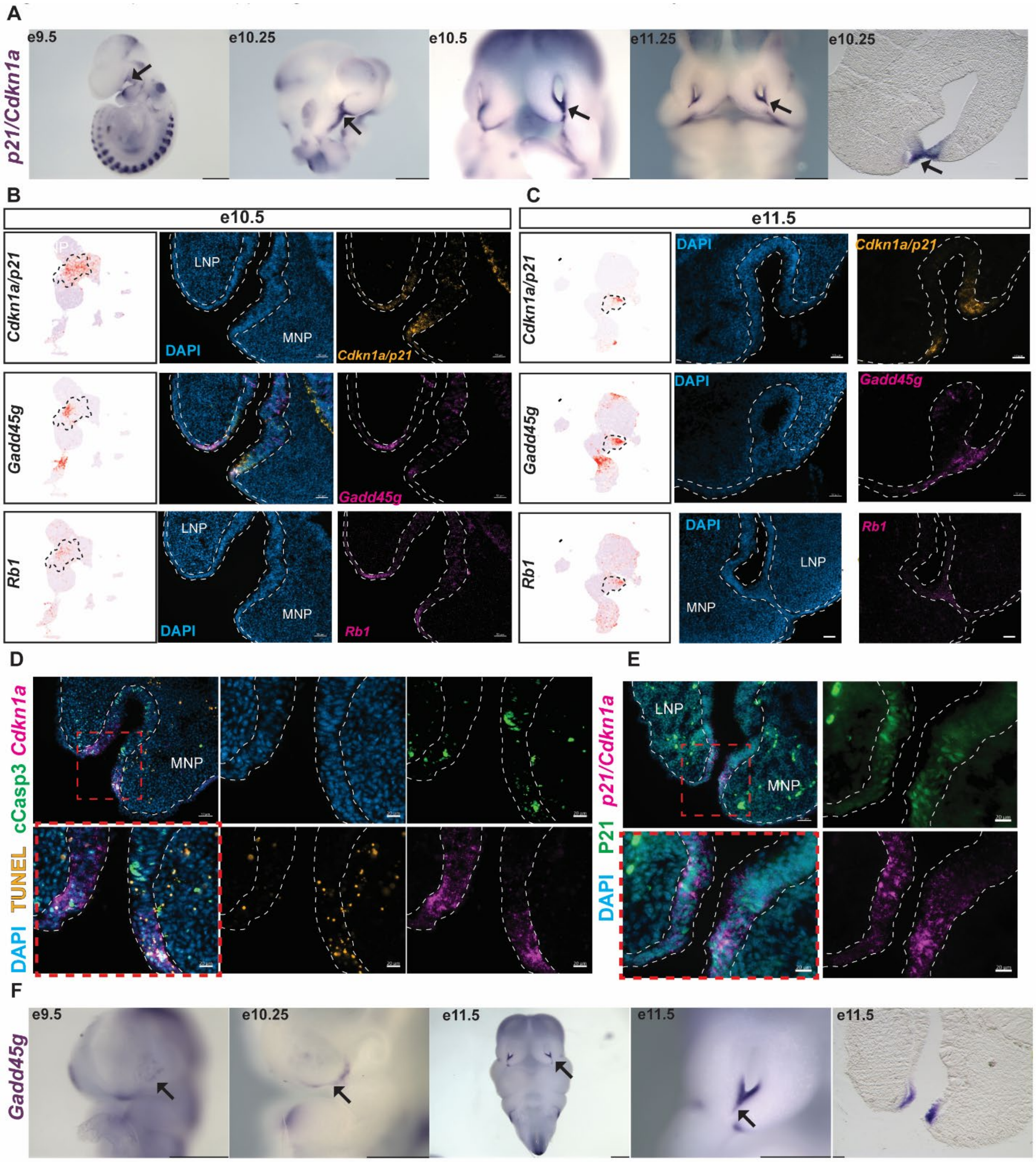


Figure 2.6.4: The λ Epithelium Zippering Cluster is Enriched for Markers of Cell Cycle Arrest
A: Stage course of chromogenic whole mount *in situ* hybridization against cell cycle arrest marker *Cdkn1a/p21* in wild-type mouse embryos stages E9.5-E11.25 as determined by somite count. Arrowheads highlight region of high expression within λ. Expression of *Cdkn1a/p21* fades as λ fusion progresses. Scale bars set to 200µm. A': 40µm section through λ of e10.25 WISH embryo demonstrates specificity to epithelium at fusion site. Scale bar: 50µm. B: Heatmap overlays of *Cdkn1a, Gadd45g,* and *Rb1* gene expression on corresponding E10.5 scRNAseq UMAP (left). RNAscope FISH against *Cdkn1a, Gadd45g,* and *Rb1* highlight strong differential expression of cell-cycle arrest markers (Figure caption continued on the next page.)

(Figure caption continued from the previous page.) at the tips of the fusing prominence epithelium at E10.5 (center and right). C: Heatmap of *Cdkn1a, Gadd45g,* and *Rb1* gene expression on corresponding E11.5 scRNAseq UMAP (left). RNAscope FISH against *Cdkn1a, Gadd45g,* and *Rb1* in sections of E11.5 wild-type midfaces (center and right). D: cCASP3 immunofluorescence, TUNEL staining, and *Cdkn1a* RNAscope probes in E10.5 transverse midface sections demonstrate partial overlap of apoptosis markers with cell cycle arrest marker. D': Inset for detail of D red dashed box. Scale bars: 50µm. E: Both protein and transcript are expressed for P21 IF with *Cdkn1a/p21* RNAscope FISH highlights protein expression correlation with transcript. Scale bars: 50µm. F: Stage course of chromogenic whole mount in situ hybridization against *Gadd45g* in wild-type mouse embryos stages e9.5-E11.5. Arrows indicate λ expression of *Gadd45g* which fades as fusion of the prominences progresses. Scale bars: 200µm.

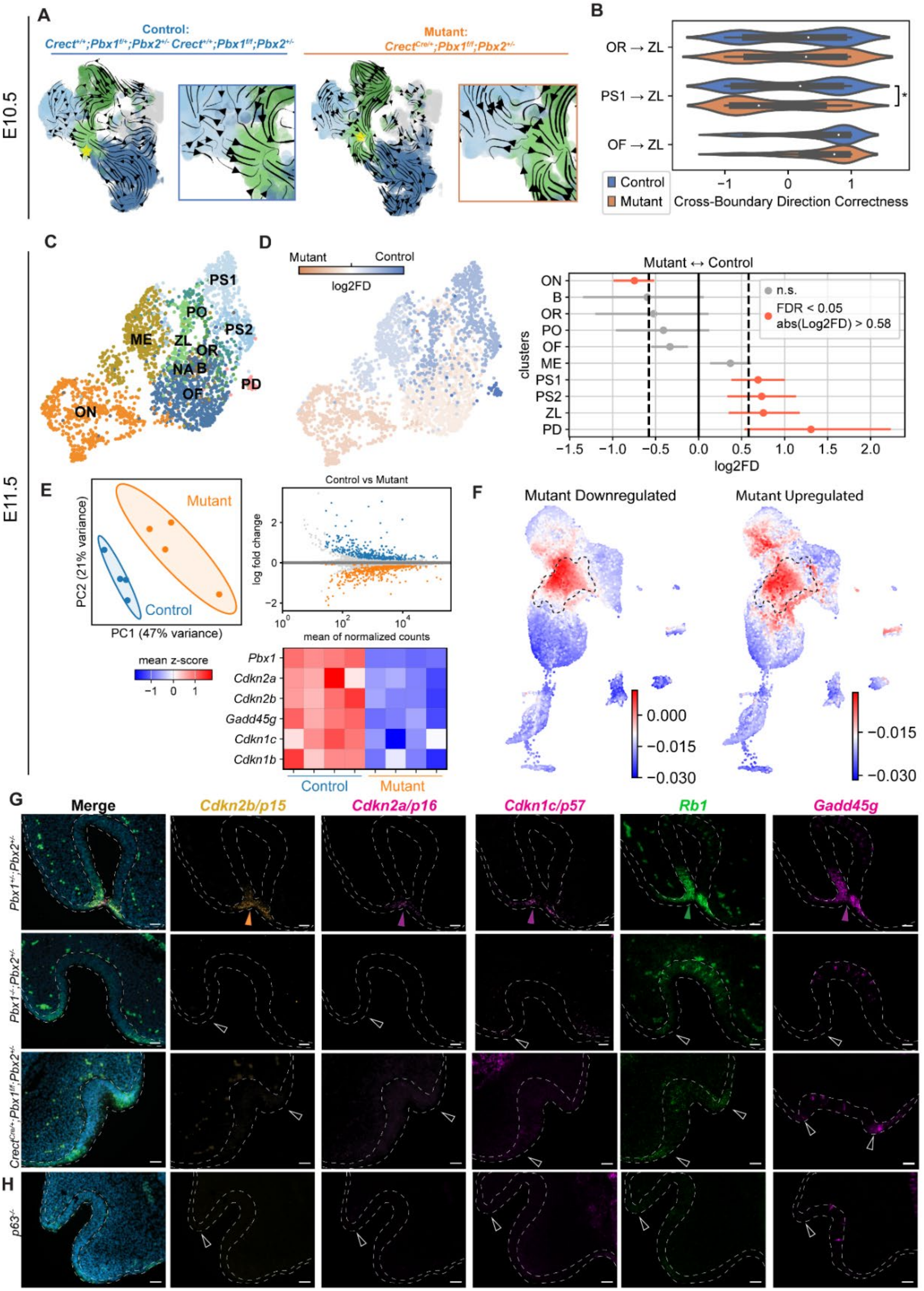
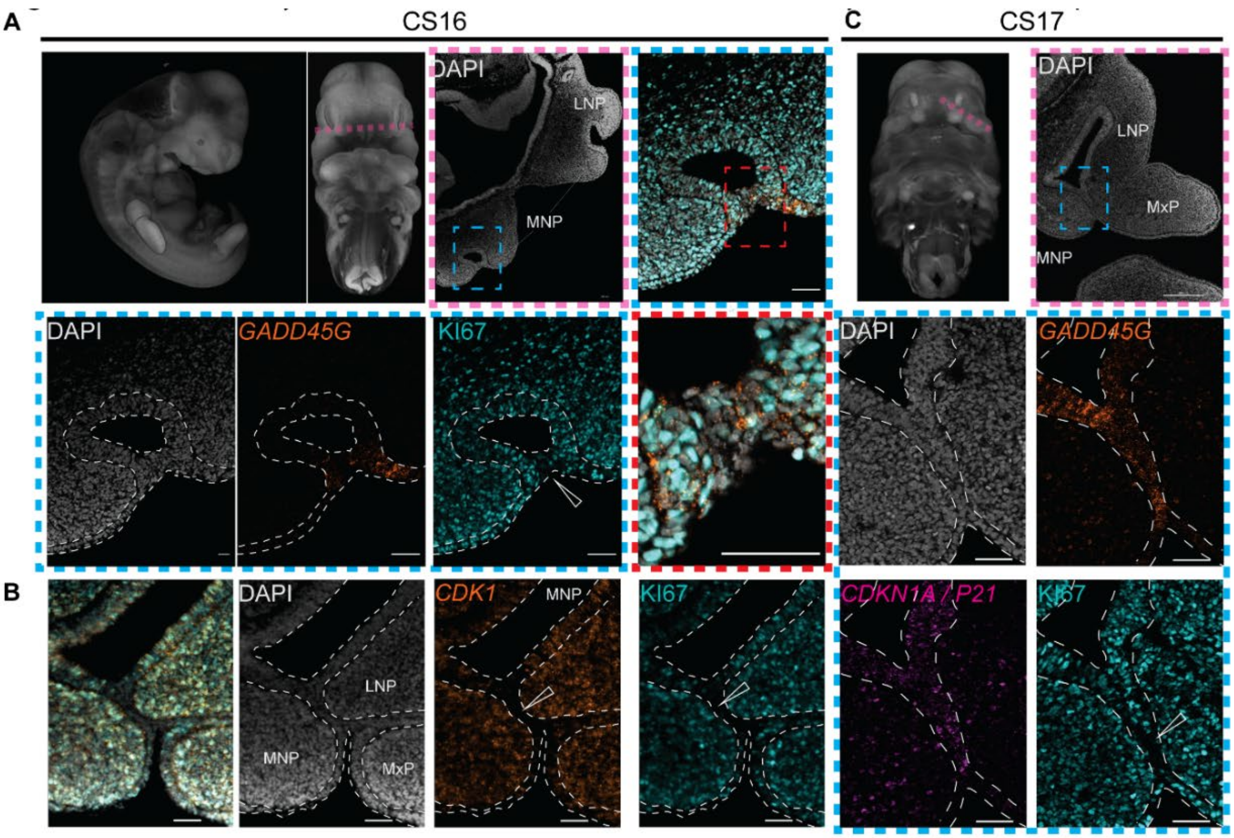

A
Control:
Crect+/+;Pbx1f/+;Pbx2+/- Crect+/+;Pbx1f/f;Pbx2+/-
Mutant:
CrectCre/+;Pbx1f/f;Pbx2+/-
E10.5
B
OR → ZL
PS1 → ZL
OF → ZL
*
Control
Mutant
−1
0
1
Cross-Boundary Direction Correctness
C
PS1
PO
PS2
ME
ZL
OR
NA
B
PD
OF
ON
D
Mutant
Control
log2FD
Mutant ↔ Control
ON
B
OR
PO
OF
ME
PS1
PS2
ZL
PD
clusters
n.s.
FDR < 0.05
abs(Log2FD) > 0.58
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
log2FD
E11.5
E
Mutant
Control
PC2 (21% variance)
PC1 (47% variance)
Control vs Mutant
log fold change
2
0
−2
10^0
10^2
10^4
mean of normalized counts
mean z-score
−1
0
1
Pbx1
Cdkn2a
Cdkn2b
Gadd45g
Cdkn1c
Cdkn1b
Control
Mutant
F
Mutant Downregulated
Mutant Upregulated
0.000
−0.015
−0.030
−0.015
−0.030
G
Merge
Cdkn2b/p15
Cdkn2a/p16
Cdkn1c/p57
Rb1
Gadd45g
Pbx1+/-;Pbx2+/-
Pbx1-/-;Pbx2+/-
CrectCre/+;Pbx1f/f;Pbx2+/-
H
p63-/-

Figure 2.6.5: Pbx1/2 Mutant Embryos Display Perturbed Velocity Trajectories into the λ Zippering Epithelial cluster. Both Pbx1/2 and p63 Mutants Exhibit Dysregulation of Cell Cycle Genes in λ Epithelium
A: RNA velocity of E10.5 control (*Pbx1*$^{f/}$*f; Pbx2*$^{+/-}$*; Crect*$^{+/+}$ and *Pbx1*$^{f/+}$*; Pbx2*$^{+/-}$*; Crect*$^{+/+}$) scRNAseq (left) versus mutants (*Pbx1*$^{f/f}$*Pbx2*$^{+/-}$*Crect*$^{Cre/+}$). Insets to show detail of boundary between ZL and PS cluster transitions. B: Cross boundary direction correctness measured between the ZL cluster and three neighboring clusters that have predicted contributions to ZL formation in control RNAvelocity versus mutants. Only cluster PS1 shows a significant change in correctness with the ZL cluster in *Pbx1*$^{f/f}$*Pbx2*$^{+/-}$*Crect*$^{Cre/+}$ mutants. C. UMAP diagram of E11.5 scRNAseq of both labeled clusters of control (*Pbx1*$^{f/}$*f; Pbx2*$^{+/-}$*; Crect*$^{+/+}$ and *Pbx1*$^{f/+}$*; Pbx2*$^{+/-}$*; Crect*$^{+/+}$) scRNAseq and mutants (*Pbx1*$^{f/f}$*Pbx2*$^{+/-}$*Crect*$^{Cre/+}$) midface epithelium with cluster abbreviations. D. Heatmap overlay on E11.5 *Pbx1/2* mutant and control midface epithelium scRNAseq UMAP showing normalized proportion of cells per cluster found in mutants versus controls (left). Plot of normalized proportions of cells by cluster in mutants versus controls showing statistically significant increases in PS1/2, ZL, PD populations in controls and a decrease in ON cells in *Pbx1/2* mutants (right). E: Four individual replicates of E11.5 control λ epithelium versus 4 *Pbx1*$^{f/f}$*Pbx2*$^{+/-}$*Crect*$^{Cre/+}$ mutant λ epithelium RNAseq group separately using principal components 1 and 2 (47% and 21% of variance, respectively) (left). Deseq2 volcano plot of differential gene expression analysis between control epithelium versus mutants (top right). *Pbx1*$^{f/f}$*Pbx2*$^{+/-}$*Crect*$^{Cre/+}$ mutant λ epithelium shows significantly downregulated transcripts of cell-cycle arrest related genes found to be enriched in the ZL region including cyclin dependent kinase inhibitors *2a, 2b, 1b*, and *1c* as well as *Gadd45g* (bottom right). F: Using RNAsieve to compare the enrichment of top differentially expressed genes in RNAseq of *Pbx1*$^{f/f}$*Pbx2*$^{+/-}$*Crect*$^{Cre/+}$ mutant λ epithelium versus controls, the top upregulated (left) and downregulated (right) genes are found to be enriched in the ZL cluster at E10.5. Heatmap of enrichment score is overlaid on wild-type scRNAseq of E10.5 epithelium. G: Transverse sections of λ-junction in both wild-type and mutant embryos with cleft lip validates downregulation of cell-cycle arrest transcripts at the fusion site and prominence tips. In *Pbx1*$^{f/f}$*Pbx2*$^{+/-}$*Crect*$^{Cre/+}$ conditional, *Pbx1*$^{-/-}$*Pbx2*$^{+/-}$ global knockout, and *Trp63*$^{-/-}$ global knockout mouse mutant models of cleft lip, the gene expression of *Cdkn2b, 2a, 1c, Rb1,* and *Gadd45g* which normally localizes at the site of fusion (corresponding to the ZL cluster), is reduced or absent. All scale bars: 50µm.


A
CS16
DAPI
LNP
MNP
C
CS17
DAPI
LNP
MxP
MNP
DAPI
GADD45G
KI67
DAPI
GADD45G
B
DAPI
CDK1
MNP
KI67
LNP
MNP
MxP
CDKN1A / P21
KI67

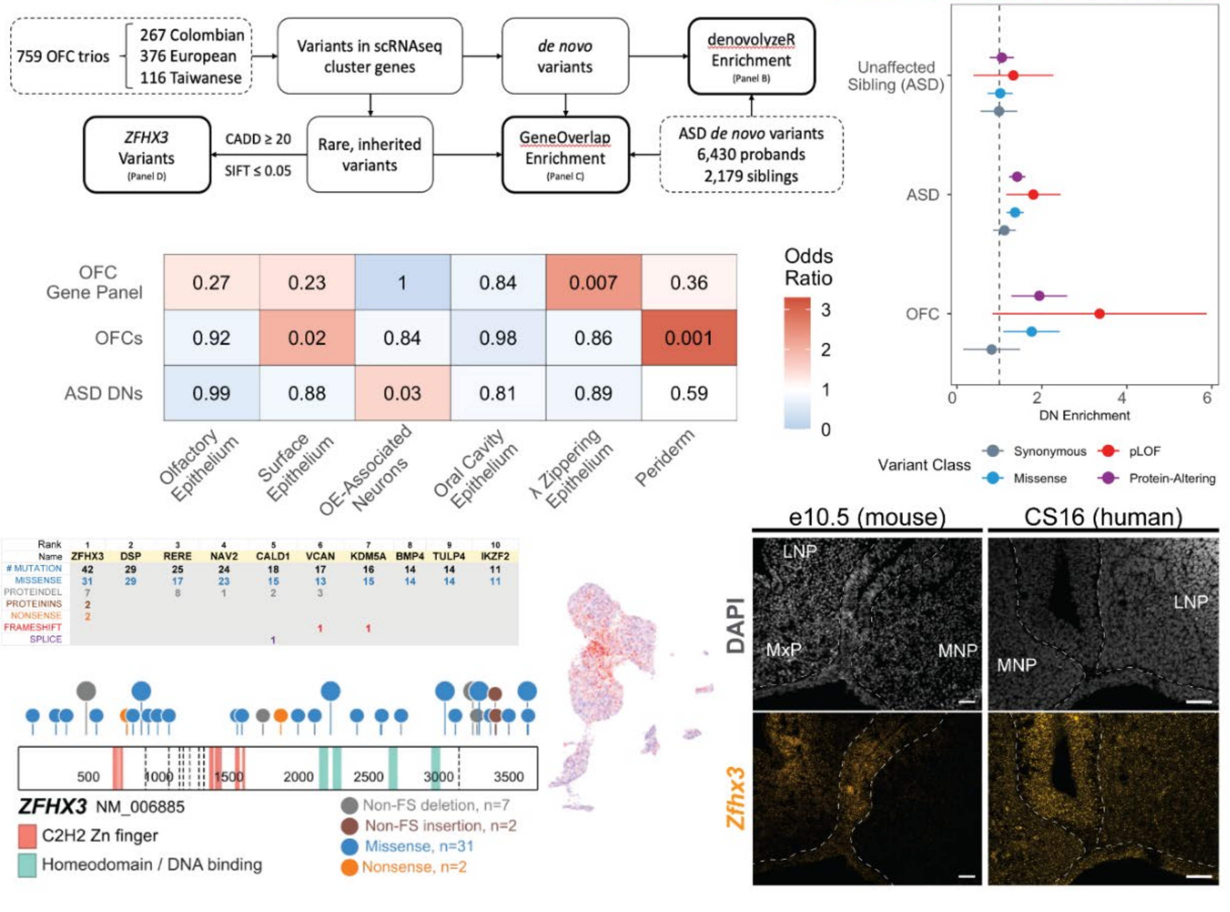

759 OFC trios
267 Colombian
376 European
116 Taiwanese
Variants in scRNAseq cluster genes
de novo variants
denovolyzeR Enrichment (Panel B)
ZFHX3 Variants (Panel D)
CADD ≥ 20
SIFT ≤ 0.05
Rare, inherited variants
GeneOverlap Enrichment (Panel C)
ASD de novo variants
6,430 probands
2,179 siblings
OFC Gene Panel
OFCs
ASD DNs
0.27
0.23
1
0.84
0.007
0.36
0.92
0.02
0.84
0.98
0.86
0.001
0.99
0.88
0.03
0.81
0.89
0.59
Odds Ratio
3
2
1
0
Olfactory Epithelium
Surface Epithelium
OE-Associated Neurons
Oral Cavity Epithelium
λ Zippering Epithelium
Periderm
Unaffected Sibling (ASD)
ASD
OFC
0
2
4
6
DN Enrichment
Variant Class
Synonymous
Missense
pLOF
Protein-Altering
Rank
1
2
3
4
5
6
7
8
9
10
Name
ZFHX3
DSP
RERE
NAV2
CALD1
VCAN
KDM5A
BMP4
TULP4
IKZF2
MUTATION
42
29
25
24
18
17
16
14
14
11
MISSENSE
31
29
17
23
15
13
15
14
14
11
PROTEINDEL
7
8
1
2
3
PROTEININS
2
NONSENSE
2
FRAMESHIFT
1
1
SPLICE
1
500
1000
1500
2000
2500
3000
3500
ZFHX3 NM_006885
C2H2 Zn finger
Homeodomain / DNA binding
Non-FS deletion, n=7
Non-FS insertion, n=2
Missense, n=31
Nonsense, n=2
e10.5 (mouse)
CS16 (human)
DAPI
LNP
MxP
MNP
LNP
MNP
Zfhx3

Figure 2.6.6: Human Embryonic Midface Fusion is Associated with Cell Cycle Arrest in the Epithelium

A: Left: High resolution episcopic microscopy image of CS16 human embryo lateral and frontal view. Magenta dashed line represents plane of sectioning for CS16 midface RNAscope and immunofluorescence. Blue dashed box: inset on right of single region of fusing lateral and medial nasal processes. Red dashed box: inset of epithelial bridge and region of λ fusion. CS16 sections with RNAscope against *GADD45G* and immunofluorescence against KI67. Empty arrowhead indicates region of λ fusion lacking KI67 expression indicating lowered proliferation. B: CS16 embryonic sections demonstrate through IF and RNAscope reduced KI67 at fusion site is accompanied by lack of *CDK1* expression. C: Optical projection tomography image of CS17 human embryo lateral and frontal view. Magenta dashed line represents plane of sections. Blue dashed box: inset on right of single region of fusing lateral and medial nasal processes. RNAscope probes against *GADD45G* and *CDKN1A/P21* and IF against KI67. Empty arrowhead indicates the area absent in positive signal for KI67. All scale bars are set to 50µm. Whole embryo images provided by HDBR Atlas. D: Visualization of workflow for data analysis using human patient trio data. E: *De novo* variant enrichment in all top scRNAseq DE genes for all clusters and all OFC cases, ASD probands, and unaffected siblings of ASD probands. F: GeneOverlap results for comparison of genes with rare, predicted damaging variants in constrained genes in OFC probands and all constrained genes in the top marker genes. The color represents odds ratios (blue = reduced OR, red = increased OR) with p-values as the overlay. G: ProteinPaint table showing the top 10 genes with the most predicted damaging variants in OFC probands for the ZL population. H: Lollipop plot illustrating all rare, predicted damaging *ZFHX3* variants in any OFC proband. I: Heatmap overlay of *Zfhx3* expression at E10.5 on wildtype epithelium UMAP demonstrating enrichment in ZL cluster (dashed border). J: RNAscope FISH showing expression of *Zfhx3/ZFHX3* in E10.5 mouse and equivalent CS16 midface sections with strong expression in the λ region of fusing epithelial cells between MNP and LNP.

2.7 Supplemental Figures

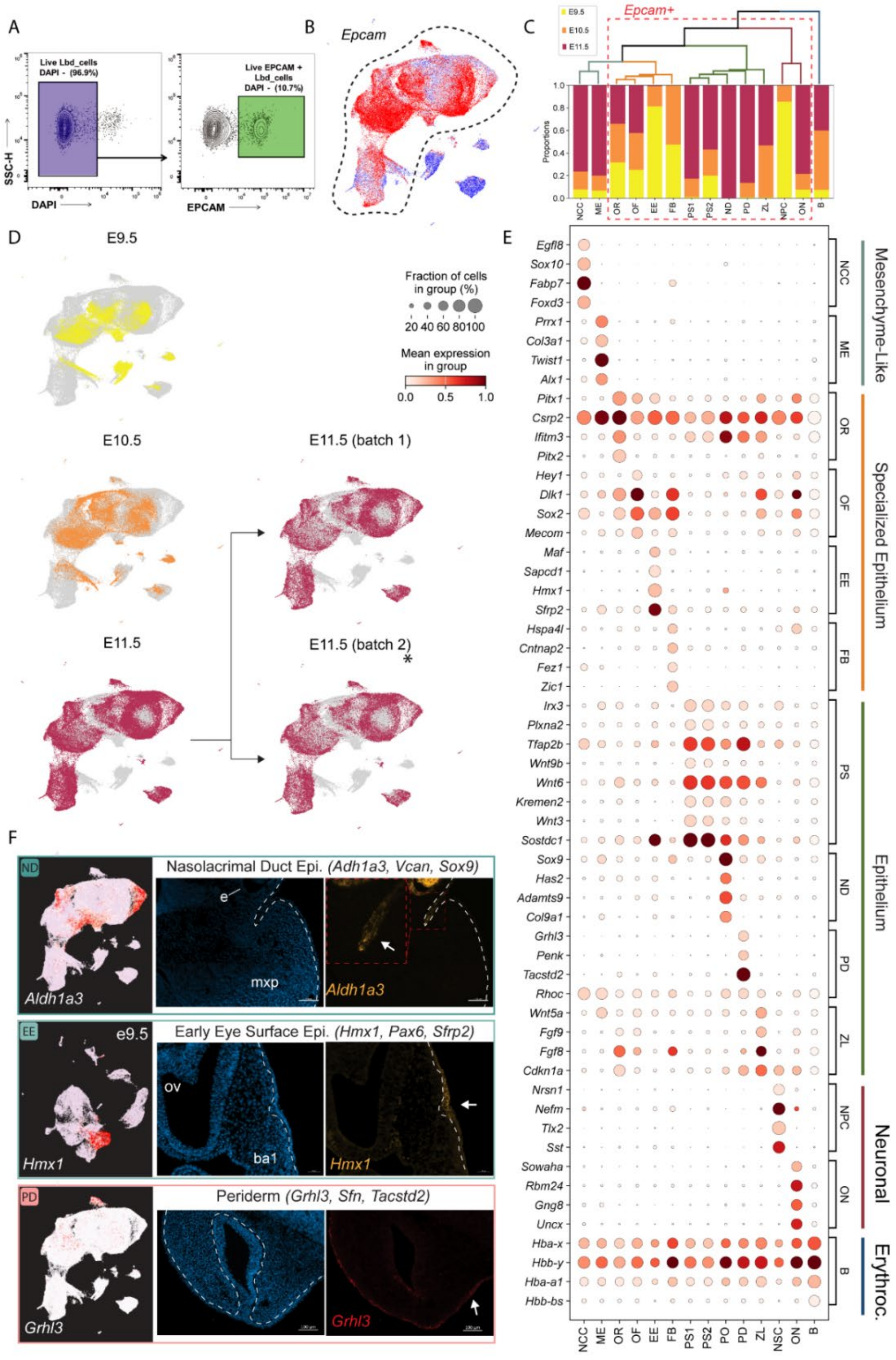


Figure 2.7.1: (Supplementary Figure 1)
A: Fluorescence activated cell sorting of midface epithelium utilized Epcam positivity as basis of isolation. Cells in the underlying mesenchyme are considered mainly negative for Epcam. We recovered cells following a DAPI viability gate followed by Epcam-APC. Cells were defined as Epcam positive if fluorescence was greater than 10^4 units. Depending on the gestational day, the range of Epcam positive cells averaged 6%-10% of sorted midface tissue. B: *Epcam* gene expression in combined dataset UMAP E9.5-E11.5 shows that post-FACS there is heterogeneity in the levels of *Epcam* in recovered cells. Dashed border indicates main clusters for study. Clusters outside dashed line were excluded from validation and, based on low *Epcam* and DEGs, were mainly considered mesenchymal or mesenchyme-like. C: Proportional representation of Leiden clusters at each of the three gestational day datasets (E11.5 batches are considered together). The cells that are considered *Epcam* high comprise the boxed red region, apart from NSC, which are believed to give rise to the *Epcam* high ON population. Combined, the NSC, ON, and FB clusters comprise a region of neuronal transcriptomes that separates away from others in UMAP. D: The individual clusters of E9.5-E11.5 are separated from the UMAP overlay into their individual timepoints. The E11.5 datasets are further divided between two separate batches. E: Mean expression of top differentially expressed genes in the individual clusters of scRNAseq. The top four genes are displayed for each cluster. The PS1 and PS2 cluster are grouped based on a shared identity, separated by cell cycle phases. Five broad categories of clusters are defined based on the validated localization or known identities based on gene expression. F: The nasolacrimal duct (PD), early eye (EE), and periderm (PD), clusters were validated using RNAscope FISH. Heatmap overlays of marker gene expression (left) are overlaid on combined E9.5-E11.5 UMAPs. The EE cluster is primarily found at E9.5 and is best represented by accompanying E9.5 UMAP with heatmap overlay. (right) PO validation targeting *Aldh1a3* near the junction of the maxillary prominence (mxp) and the eye (e), with inset to show detail (red dashed box). Scale bar 100µm. EE cluster validation with *Hmx1* RNAscope probe using transverse section of E9.5 embryo. *Hmx1* expression lies superficial to developing optical vesicle (ov) of forebrain and dorsal to the forming branchial arch 1 (ba1). Scale bar 50µm. Validation of PD cluster using *Grhl3* probes. *Grhl3* positive periderm cells are seen to develop on the prominence surfaces as well as inside the nasal pit, superficial to the monolayer of epithelium above the basal membrane (dashed white line). Scale bar 100µm.

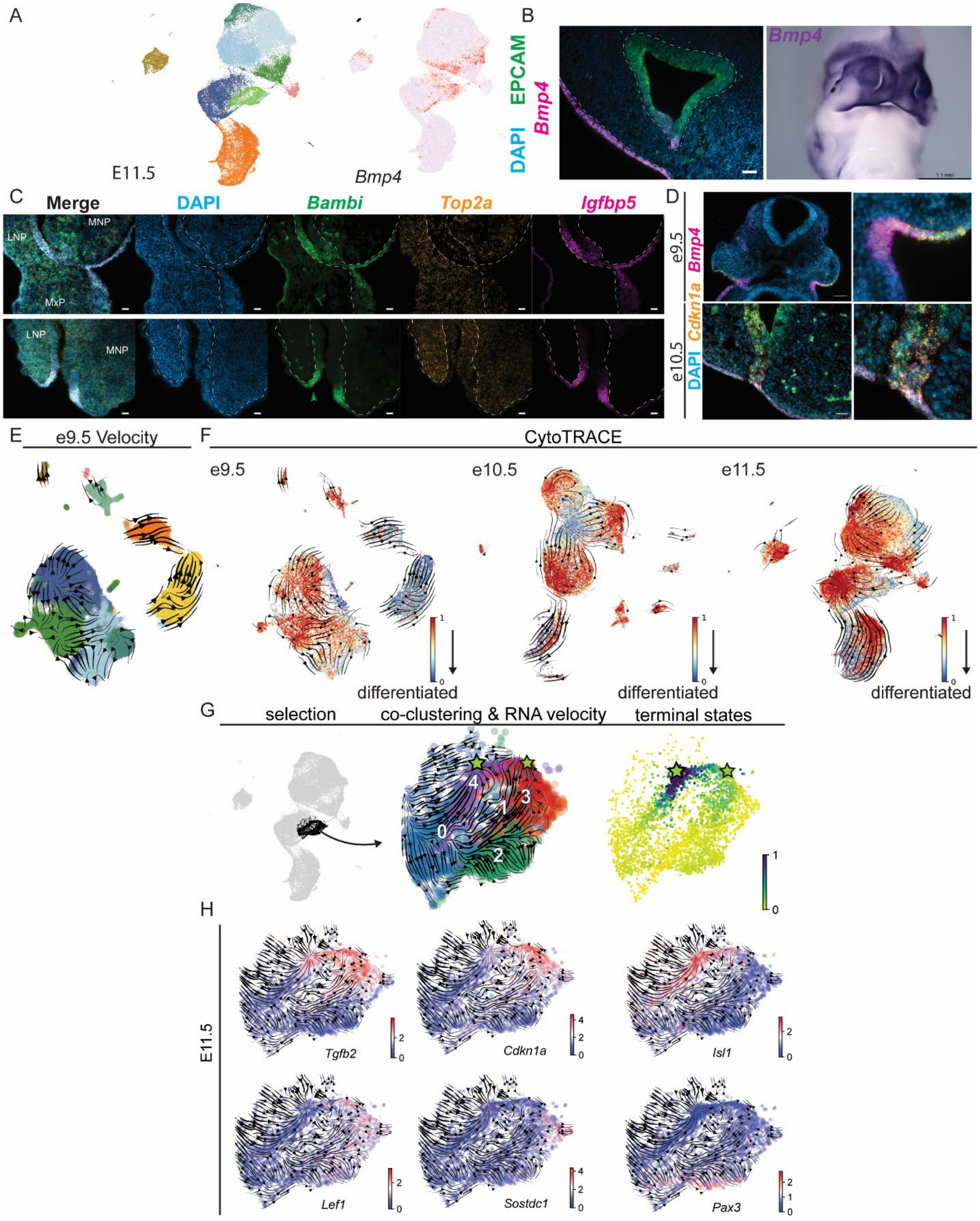


Figure 2.7.2: (Supplementary Figure 2)

A: UMAP of E11.5 of wild-type epithelium of the midface (left) and heatmap overlay of *Bmp4* expression at E11.5. *Bmp4* expression is (Figure caption continued on the next page.)

(Figure caption continued from the previous page.) expanded into ZL, PS1, PS2, and OR clusters. B: Transverse section of E11.5 midface. RNAscope probe *Bmp4* confirms expression in both prominence surface epithelium and at the fusion site ZL region (left). Scale bar 50µm. Wholemount *in-situ* hybridization chromogenic assay against *Bmp4* E11.75-E12.0. Expression of *Bmp4* beyond E11.5 is widespread throughout the midface, notably the medial and lateral nasal prominences, oral cavity, and developing vibrissae placodes (right). C: Together, other top DEGs for the ZL cluster are shown to localize at the λ fusion site in both coronal and transverse sections of the midface. RNAscope FISH against BMP antagonist *Bambi,* and IGF inhibitor *Igfbp5* confirm high levels of gene expression at the fusion site. Epithelium high in *Bambi* and *Igfbp5* transcript is largely depleted of *Top2a* gene expression, a marker of G2-M and mitotic division. Scale bars 50µm. C: *Bmp4* expression at E9.5 and E10.5 in transverse sections shows overlap with cell-cycle arrest gene *Cdkn1a* expression at the prior to and during ZL formation. Scale bars 50µm. D: RNA Velocity of E9.5 wild-type epithelium overlay E9.5 cluster UMAP. Endpoints of vectors are found primarily in neuroepithelial related ON and NSC clusters. E: CytoTRACE heatmap overlay of differentiated versus non-differentiated transcriptomes E9.5-E11.5. At E9.5 the most differentiated clusters, in line with RNA Velocity endpoints, are NSC and ON. By E10.5 the appearance of ZL cluster also coincides with a differentiated state in a non-neuronal cluster. At E11.5 CytoTRACE differentiated endpoints persist in ON, ZL, and portions of PS1 cluster. F: E11.5 ZL cluster selected for isolation (black), (left). Co-clustering and RNAvelocity on UMAP of ZL at E11.5 produces 5 subclusters with endpoints (green stars) found in subclusters 3 and 4 (center). Terminal states are mapped on UMAP in heatmap overlay (right). G: Heatmap overlay on UMAP of isolated ZL at E11.5 for select DEGs. Among the top genes found to be highly expressed at the predicted terminal states at E11.5 are *Tgfb2, Cdkn1a,* and *Isl1.* The same endpoints are largely devoid of *Lef1, Sostdc1,* and *Pax3* expression.

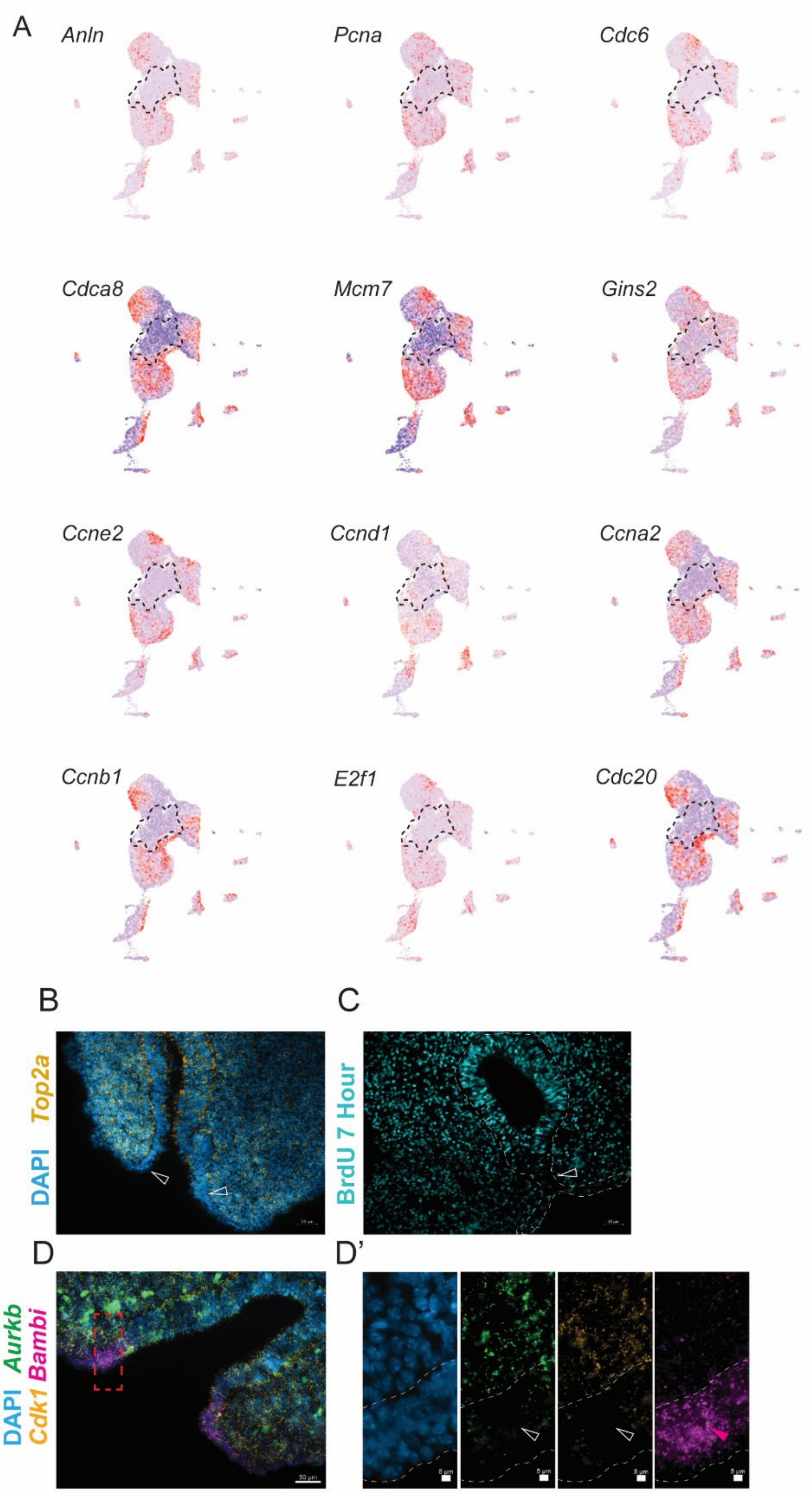


Figure 2.7.3: (Supplementary Figure 3)
A: Heatmap showing differential gene expression of additional cell cycle genes tied to specific phases of the cell cycle. All are shown to have expression in parts of clusters except for the ZL cluster (dashed line) where expression is absent. (Figure caption continued on the next page.)

(Figure caption continued from the previous page.) *Anln*- end of M phase, *Pcna*- S phase, *Cdc6*-S phase, *Cdca8*-G2 to M phase, *Mcm7*- G1 to S phase, *Gins2*-S phase, *Ccne2*-G1 to S phase, *Ccnd1*-G1 to S phase, *Ccna2*-S to G2 phase, *Ccnb1*-G2 to M phase, *E2f1*-G1-S transition, *Cdc20*-M phase. B: RNAscope FISH against *Top2a* in E10.5 transverse sections of MNP and LNP. C: 7-hour BrdU chase from E10.5 showing lack of BrdU integration at fusion site. D: RNAscope FISH against *Aurkb, Cdk1,* and *Bambi*. *Cdk1* and *Aurkb* are absent from region of high *Bambi* expression. Scale bars 50µm. Inset scale bars 5µm.

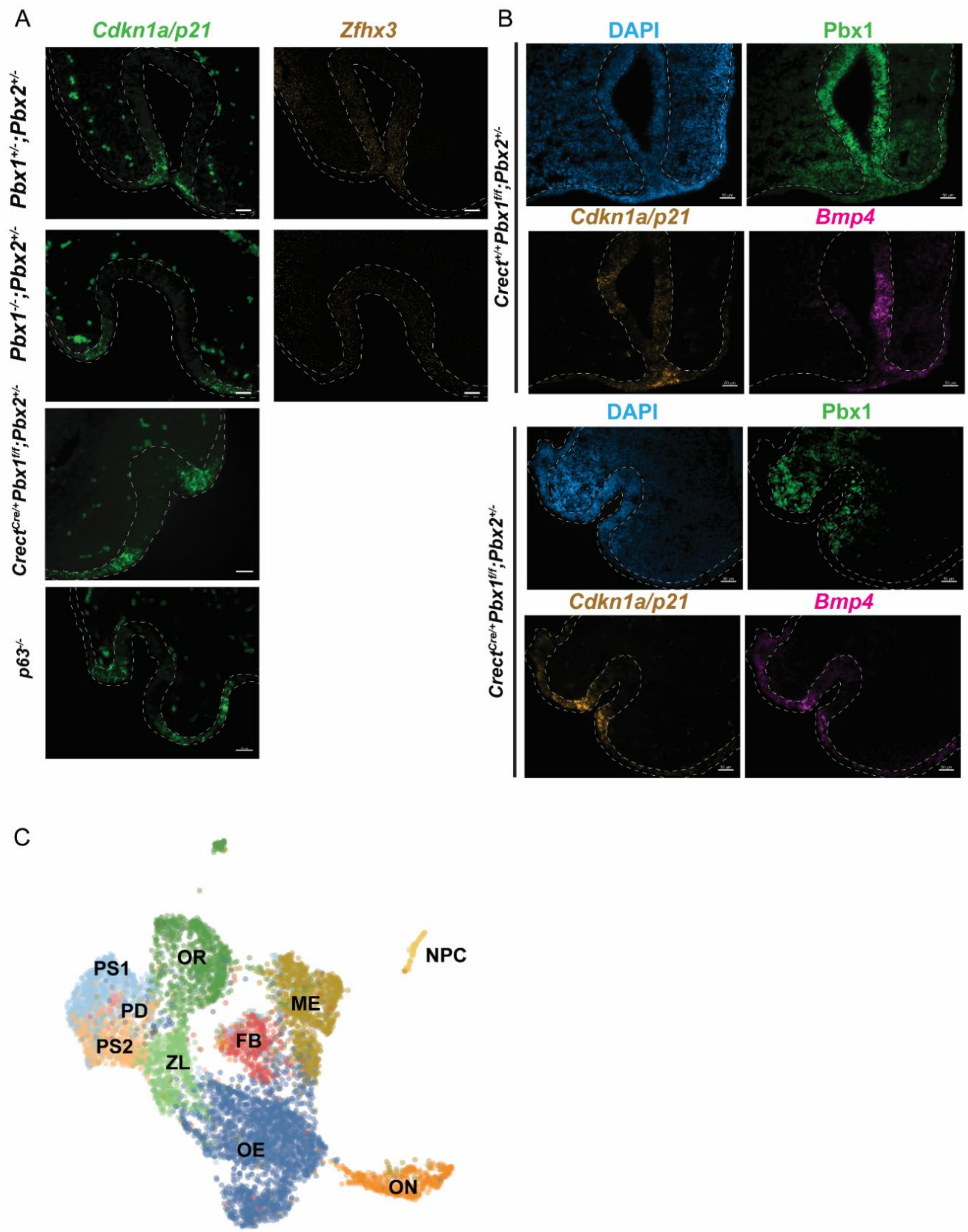


Figure 2.7.4: (Supplementary Figure 5)

A: RNAscope FISH against *Cdkn1a/p21* and *Zfhx3* in control (1st row), *Pbx1/2* global knockout (2nd row), *Pbx1* conditional knockout on *Pbx2* (Figure caption continued on the next page.)

(Figure caption continued from the previous page.) global deficient knockout (3rd row), and *p63* null (4th row) sections of LNP and MNP at lambdoidal junctions, showing high differential expression in all cases at the prominence tips or fusion site. *Zfhx3* expression (2nd column) is shown to be decreased in the global knockout *Pbx1/2* compared to controls. This expression is widespread but upregulated at the fusion site in controls (1st row). B: Validation of *Pbx1* conditional knockdown with *Crect-Cre*. Pbx1 immunofluorescence with *Cdkn1a* and *Bmp4* RNAscope FISH for identification of ZL cluster. (Top) *Crect-cre* negative control embryo section of LNP and MNP shows enrichment of Pbx1 in epithelium (dashed white border) compared to (Bottom) *Crect-Cre* positive *Pbx1/2* conditional mutant showing absent Pbx1 protein in epithelium. Note: Pbx1 is seen still expressed in olfactory neurons embedded in epithelium. C: Leiden clusters for E10.5 *Pbx1/2* conditional mutant versus control scRNAseq displayed in UMAP with cluster abbreviation labels. All scale bars = 50µm.

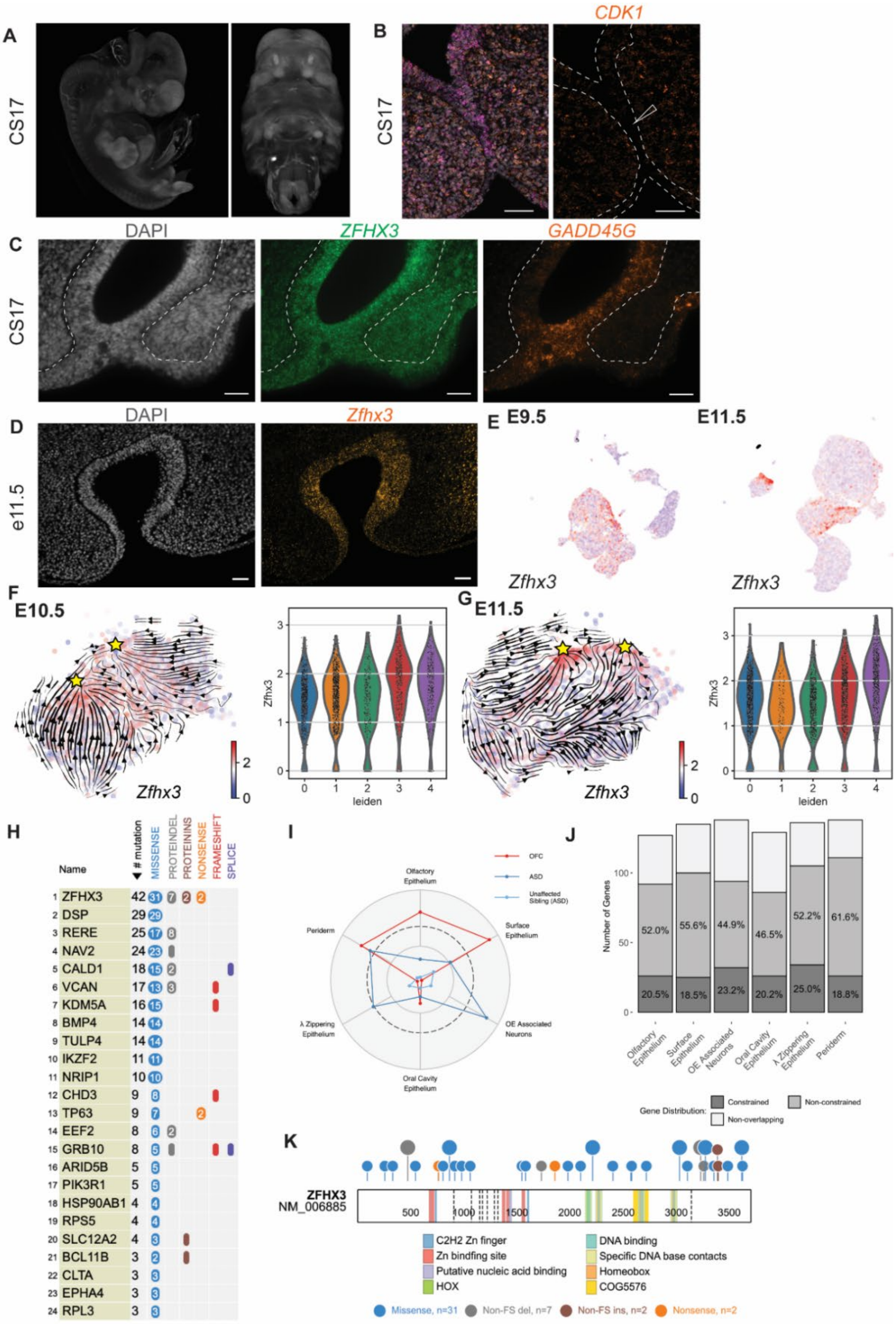

A
CS17
B
CS17
CDK1
C
CS17
DAPI
ZFHX3
GADD45G
D
e11.5
DAPI
Zfhx3
E
E9.5
E11.5
Zfhx3
Zfhx3
F
E10.5
Zfhx3
0
2
Zfhx3
leiden
0 1 2 3 4
G
E11.5
Zfhx3
0
2
Zfhx3
leiden
0 1 2 3 4
H
Name
mutation
MISSENSE
PROTEINDEL
PROTEININS
NONSENSE
FRAMESHIFT
SPLICE
1 ZFHX3 42 31 7 2 2
2 DSP 29 29
3 RERE 25 17 8
4 NAV2 24 23
5 CALD1 18 15 2
6 VCAN 17 13 3
7 KDM5A 16 15
8 BMP4 14 14
9 TULP4 14 14
10 IKZF2 11 11
11 NRIP1 10 10
12 CHD3 9 8
13 TP63 9 7 2
14 EEF2 8 6 2
15 GRB10 8 5
16 ARID5B 5 5
17 PIK3R1 5 5
18 HSP90AB1 4 4
19 RPS5 4 4
20 SLC12A2 4 3
21 BCL11B 3 2
22 CLTA 3 3
23 EPHA4 3 3
24 RPL3 3 3
I
OFC
ASD
Unaffected Sibling (ASD)
Olfactory Epithelium
Surface Epithelium
OE Associated Neurons
Oral Cavity Epithelium
λ Zippering Epithelium
Periderm
J
Number of Genes
0
50
100
52.0%
55.6%
44.9%
46.5%
52.2%
61.6%
20.5%
18.5%
23.2%
20.2%
25.0%
18.8%
Olfactory Epithelium
Surface Epithelium
OE Associated Neurons
Oral Cavity Epithelium
λ Zippering Epithelium
Periderm
Gene Distribution:
Constrained
Non-constrained
Non-overlapping
K
ZFHX3
NM_006885
500
1000
1500
2000
2500
3000
3500
C2H2 Zn finger
Zn bindfing site
Putative nucleic acid binding
HOX
DNA binding
Specific DNA base contacts
Homeobox
COG5576
Missense, n=31
Non-FS del, n=7
Non-FS ins, n=2
Nonsense, n=2

Figure 2.7.5: (Supplementary Figure 6)

A: High resolution episcopic microscopy image of CS17 human embryo lateral and frontal view. B: CS17 RNAscope FISH of *CDK1* and *CDKN1A/P21*. *CDK1* expression is absent from the fusing epithelium between LNP and MNP where *CDKN1A/P21* expression is enriched. C: CS17 RNAscope FISH against *ZFHX3* and *GADD45G* shows high expression in the epithelium near the fusion site of LNP and MNP at CS17. D: RNAScope FISH against *Zfhx3* in E11.5 mouse transverse section of MNP and LNP showing high expression in epithelium. E: E9.5 Heatmap of differential expression for *Zfhx3* overlaid on E9.5 wild-type UMAP. D: E11.5 Heatmap of differential expression for *Zfhx3* overlaid on E11.5 wild-type UMAP. F: *Zfhx3* expression heatmap overlay on isolated and reclustered E10.5 UMAP wildtype epithelium with RNA velocity vectors and endpoints of RNA velocity marked (yellow stars) (left). *Zfhx3* expression violin plot by individual subclusters for isolated E10.5 ZL cluster (right). G: *Zfhx3* expression heatmap overlay on isolated and reclustered E11.5 UMAP wildtype epithelium with RNA velocity vectors and endpoints of RNA velocity marked (yellow stars) (left). *Zfhx3* expression violin plot by subcluster for isolated E11.5 ZL cluster (right). H: ProteinPaint visualization of top 24 genes from intersection of scRNAseq of ZL cluster with rare variant analysis of human OFC cleft trios. Counts of all rare variant mutations per gene are shown with breakdown by class of variant. I: Radar plot of -log10 (p-values) for specific cluster enrichment of only *de novo* variations. The dotted line represents multiple test correction for 6 clusters (0.05/6, p=0.0083). J: Distribution of overlap in genes with rare variants in OFC probands and the top genes expressed in each scRNAseq cluster broken down by constrained, non-constrained, and non-overlapping categories. K. Lollipop plot illustrating all rare predicted damaging *ZFHX3* variants in any OFC probands. Full features for the *ZFHX3* locus are displayed.

CHAPTER 3: SUMMARY ON ANALYSIS OF MOUSE MUTANT EMBRYOS WITH *BMPR1A* ECTODERMAL DELETION

3.1 Introduction

Bmpr1a/Alk3 codes for a receptor in the TGFβ/BMP superfamily of ligands and receptors. It is a Ser/Thr kinase receptor in the type I category, along with related BMPR1B/ALK6. Once bound by a BMP protein, type II category BMPR2 recruits a type I receptor such as BMPR1A, phosphorylating the latter, and initiating a signal cascade through SMAD phosphorylation. BMPR1A has high affinity for BMP2, BMP4, but has interactions with the diverse array of BMP and GDF ligands as well in a complex combinatorial logic system involving receptor/ligand combinations and competitive ligand binding (Klumpe et al., 2022). Constitutive loss of *Bmpr1a* causes early embryonic lethality in the mouse, disrupting gastrulation and preventing formation of the mesoderm (Mishina et al., 1995). For this reason, conditional loss of *Bmpr1a* has been used to study context-specific genetic interactions and effects in mouse development.

Loss of *Bmpr1a* in the embryonic facial prominences was previously demonstrated to create a fully penetrant bilateral CL/P (Liu et al., 2005). Additionally, these mice had defects in tooth development as well as decreases in cell proliferation and increased apoptosis in the midfacial prominences. Experiments involving increased and decreased BMP signaling in chick facial primordia similarly caused changes to tissue proliferation (Ashique et al., 2002). Increases in BMP signaling favored cell death in the prominences, whereas exogenous Noggin, an inhibitor of BMP signaling, had a protective role. BMP4 is widely expressed in the midface and has strong expression at the lambdoidal junction as shown previously (see previous sections).

Conditional loss of *Bmpr1a* has been performed using *Nestin-Cre* recombination of the floxed allele (Liu et al., 2005). The *Nestin-Cre* expression is very broad in the midface, and not localized solely to the epithelium, except in some portions of the LNP and MNP. There are large areas of *Nestin-Cre* activity in the mesenchyme, especially of the MxP at E10.5, and strong activity in the forebrain. The *Nestin-Cre* line has been largely abandoned in the craniofacial biology field in favor of more specific mouse *Cre* lines. The findings of Liu et al 2005 cannot precisely untangle the mesenchymal *versus* epithelial roles of BMP signaling, despite showing that this signaling is critical for fusion of the face prominences. In contrast to *Nestin-Cre, Cre-ectoderm (Crect)* uses an ectoderm-specific enhancer of *Tfap2a* to drive expression in surface cephalic ectoderm as early as E8.25 (Reid et al., 2011). The expression at E10.5, during fusion of the midface prominences, is restricted to the epithelium only (Losa et al., 2018).

To decipher the role of BMPR1A signaling in the epithelium, I employed a novel cross between the conditional *Bmpr1a/Alk3* allele as described in Mishina et al 1995, and the *Crect* recombinase mouse line to conditionally remove *Bmpr1a* from the ectoderm of the developing embryo. Here I illustrate the results of this cross through gross anatomical dissection and observation of phenotypes at several stages of development. I find drastic structural defects in the face, limbs, gut, and eye resulting from conditional loss of *Bmpr1a* in the ectoderm. I also investigate gene expression and protein levels of cell cycle markers in the zippering lambda epithelial cluster at the face prominence tips during fusion.

3.2 Results:

Gross anatomical descriptions of Bmpr1a mutants

I collected several litters of embryos at E15.5 and P0 stages to further characterize the phenotype of *Bmpr1a* loss in the ectodermal tissues. At E15.5, craniofacial development should have progressed past both primary and secondary palatogenesis, allowing for assessment of cleft lip and palate phenotypes.

E15.5 litters were dissected, and mutants compared to littermate controls. The precise genotype of individuals was revealed following the gross anatomical observations. In all cases of ectoderm-specific homozygous loss of *Bmpr1a* ($Bmpr1a^{f/f}$; $Crect^{Cre/+}$), there was the appearance of a bilateral cleft lip with cleft palate phenotype, also known as fully penetrant CL/P **(Fig. 3.5.1B)**. In instances of heterozygous loss, the phenotype was not fully penetrant with both phenotypically normal, CL only, and CL/P phenotypes observed. In all instances of CL phenotype, there was accompanying melanocyte pigmentation concentrated above each nostril **(Fig3.5.1C)**. The pupils of mutant eyes were dysmorphic and accompanied by what appeared to be misshapen irises in all affected homozygous and heterozygous individuals.

The homozygous loss of *Bmpr1a* was also characterized by complete agenesis of both forelimb and hindlimbs at E15.5 **(Fig.3.5.2B)**. Heterozygous mutants possessed near complete agenesis of the forelimbs, exhibiting severely hypoplastic distal and proximal components. A stylopod, elbow joint, and zeugopod component was seen in these hypoplastic forelimbs **(Fig.3.5.2A)**. In the autopod there was severe oligodactyly, with up to two individual digits seen at most. Agenesis of the hindlimb was fully penetrant, with no components of the hindlimb detected at E15.5 in any homozygous and affected heterozygous mutant embryos. Homozygous mutants displayed edema of the outer ectodermal layers around the cranium which was not seen in heterozygous mutants with or without other defects **(Fig.3.5.2B)**.

At E15.5 the embryos exhibited ventral body wall closure defects indicated by the presence of omphaloceles **(Fig.3.5.2B)**. The intestines could be seen outside of the body cavity in all homozygous and affected heterozygous embryos. Extreme examples of omphaloceles were seen in homozygous mutants, where the ventral body wall had minimal closure, thus permitting the intestines and liver to grow outside of the body cavity **(Fig.3.5.2B)**.

Calcified remnants of embryos were found sporadically **(Fig.3.5.2C)**. The necrotic tissues were not easily phenotyped or genotyped. While I cannot account for early-stage reabsorptions, I find evidence that a homozygous calcified mutant remains possesses a bilateral cleft lip, limb agenesis, and eye dysplasia. Based on the maturity of the eye and face, it suggests that embryonic lethality occurred around E11.5-E12.5.

Bmpr1a ectoderm deletion mutants are born alive and possessing major defects.

Litters resulting from the cross of *Bmpr1a*$^{f/f}$ female mice to male *Bmpr1a*$^{f/+}$*; CrectCre* mice come to term and are delivered normally. At P0, pups with both mutant and wild-type phenotypes were found **(Fig. 3.5.3A, B)**. The observations largely mimicked those reported at E15.5, including a bilateral cleft lip and limb defects. One slight difference observed was the rare occurrence of rudimentary and hypoplastic hindlimbs in heterozygotes at P0, though these were even less developed than their forelimb counterparts **(Fig. 3.5.3A)**. In contrast to embryonic stage E15.5, P0 pups did not possess omphaloceles, and the ventral body wall closed even in homozygous mutants, suggesting a potential resolution of this defect embryonically post-E15.5. However, the bodies of mutant individuals were distended and bloated **(Fig. 3.5.3A)**. This was most noticeable in homozygous pups.

Behaviorally, mutant P0 pups appeared in distress compared to phenotypically normal littermates. While mutants were motile and alive, they were unable to perform any locomotion as a result of the severe limb defects. As a possible consequence of the distended and bloated body cavity, mutants had a noticeable gasping-like breathing pattern. For this reason, immediate euthanasia was performed on these pups in accordance with humane euthanasia guidelines. Milk spots were not seen in mutants, while seen in all healthy littermates **(Fig. 3.5.3A, B)**. The ability to suckle could have been compromised by CL/P, motility issues, or distress brought on by the bloating phenotype.

While an omphalocele was not present, the ventral body wall was thin, and intestines could be seen clearly through the skin in homozygous mutants **(Fig. 3.5.3A)**. Following euthanasia, the body cavities of the bloated individuals were investigated further. Upon removal of the ventral body skin, a cause of the distended and bloated appearance became apparent. Mutant intestines appeared to be filled with gas and were inflated in contrast to littermate controls **(Fig. 3.5.3C)**. Gas was released upon dissection but could be observed largely in the upper digestive tract including the small intestines and cecum **(Fig. 3.5.3C)**. A mechanistic reason for this gas accumulation was not readily apparent. An anus was present in these individuals, no visible blockages were found, and gas was primarily localized to the proximal small intestines. Deterioration of the intestines may have been occurring at the time; however, no cell and tissue necrosis was discovered. Together, these findings suggested that swallowed air may be accumulating in the upper tracts, however the stomach did not appear to be inflated. Similar findings were present in *Dlx1/2*$^{-/-}$ and *Dlx2*$^{-/-}$ mutant mice which had small intestine bowel air accumulation, as well as cleft palate (Wright et al., 2020). Heterozygous mutants without CP were also found to have gas accumulation, ruling out cleft palate as a possible source of air swallowing. Instead, it was described that *Dlx* mutants possess poor gut motility owing to enteric nervous system loss of downstream vasoactive intestinal peptide *(Vip)*. It should be noted that DLX1/2 transcription factors have been reported as under the control of BMP signaling in postnatal neural progenitors via *Olig1* (Sabo et al., 2017). Loss of *Bmpr1a* in this context could increase Olig1 and its repression of *Dlx1/2* leading to a lack of DLX1/2, giving rise to the observed phenotype. Lastly, it should be noted that *Cdkn1c/p57* null mutant mice die an early postnatal death, with shortened limbs, cleft palate, as well as an inflated gastrointestinal tract (Yan et al., 1997). This connection may implicate BMPR1A in expression of *Cdkn1c*, a cell cycle arrest gene, and thus also be linked to *Cdkn1c* expression in the epithelium of the midfacial prominences.

Gene expression of zippering lambda epithelium markers is mostly unchanged in Bmpr1a mutants.

I employ RNAscope fluorescent *in-situ* hybridization (FISH) probes against targets in midface sections of both *Bmpr1a* mutants and controls, in conjunction with immunofluorescence at E10.5 **(Fig. 3.5.4A)**. At this timepoint, a visible bilateral cleft lip is already present. A large gap between the MNP and LNP is seen. Notably, and in line with previous reports of proliferation differences when altering BMP signaling in the face, the prominences appear shorter in mutants versus controls.

The gene expression that characterizes the zippering lambda epithelium (ZL) is defined by cell cycle gene differential expression. In wild-type sections *Cdkn1a/p21* is highly expressed in this cell population, alongside lowered proliferation. In *Pbx1/2* and *p63* mutants with CL/P we did not find changes in *Cdkn1a/p21* expression in the developing midface epithelium and ZL. Likewise, here I demonstrate that *Cdkn1a/p21* expression is not abolished by the loss of *Bmpr1a* in the epithelium. Both transcript and protein are intact in *Bmpr1a* mutants without noticeable changes in expression *versus* controls **(Fig.3.5.4A)**. Lowered KI67 expression is a trait we uncovered in the murine ZL cluster, indicating lowered proliferation, which largely overlaps with areas characterized by *Cdkn1a/p21* high expression. In *Bmpr1a* mutants, the *Cdkn1a/p21* high area appears to be no different in KI67 expression than the surrounding epithelium, suggesting increased proliferation and that the previously described cell cycle arrest may be disrupted, although independently of *Cdkn1a/p21* expression.

Bambi, another marker of the ZL, is a gene encoding a pseudoreceptor for BMP signaling that is believed to inhibit BMP signaling. This inhibition of BMP signaling may play an important role in the midface prominence fusion event, preventing premature apoptosis of the epithelium caused by high BMP signaling at the prominence tips as shown in chick embryos (Ashique et al., 2002). *Bambi* loss does not cause an abnormal phenotype in mouse, including in the face (Chen et al., 2007). Here I examine *Bambi* expression in mutants with loss of *Bmpr1a* in the midface epithelium that exhibit CL/P. In control embryos, *Bambi* expression is higher in the epithelium, but can be detected also in the mesenchyme. Overall, using RNAscope, I find decreases in expression of *Bambi* in the midface epithelium and mesenchyme of embryos with loss of *Bmpr1a* in the midface epithelium **(Fig.3.5.4A, B)**.

3.3 Discussion

Bmpr1a/Alk3 loss has major effect on the development of the mouse embryo. This receptor is present throughout the body but has a major role in the ectodermal lineages. I generated a mouse line with conditional loss of *Bmpr1a* in the ectoderm alone. While the homozygous mutation does not cause embryonic lethality, it produces severe phenotypes resulting in gross anatomical defects and perinatal morbidity. *Bmpr1a* genetic ablation leads to fully penetrant agenesis of the forelimb and hindlimbs, fully penetrant bilateral CL/P, gastrointestinal motility loss, and eye defects. This is accompanied by other notable features including embryonic edema, embryonic ventral body cavity closure disruption, accumulation of melanocytes at the nostrils, and spontaneous embryonic lethality of some homozygous mutant embryos at the E11.5-E12.5 stage. The same abnormalities are present in most embryos with heterozygous ectodermal specific loss as well, with only the occasional phenotypically normal individual. In heterozygotes, notable differences include decreased gas/bloating accumulation, and severe hypoplasia, rather than agenesis, of the forelimbs only.

By studying the loss of *Bmpr1a* from the ectoderm alone I aimed to remedy the difficulties in interpreting previous studies that employed genetic deletion non-specifically in the epithelium, mesenchyme, and brain tissue. This is especially pertinent to BMP signaling, which is believed to utilize crosstalk between the mesenchyme and epithelium in normal development.

The findings of this analysis are not entirely surprising given the established roles of BMP and specifically BMP2/4 signaling in the development of the eye lens (Shu et al., 2021), enteric gut (Chalazonitis et al., 2008), limb (Bandyopadhyay et al 2006), and the face (Graf et al., 2016; Liu et al., 2005). The lack of a BA1 mandibular phenotype in embryos with ectodermal *Bmpr1a* deletion was striking, as *Bmp4* and *Bmpr1a* are both highly expressed in the mandible, and mandibular patterning and fusion is believed to involve BMP4 signaling in the epithelium (MacKenzie et al., 2009; Xu et al., 2019). Yet, no discernable phenotype involving the mandible was seen in this ectodermal *Bmpr1a* deletion.

I investigated the expression of genes enriched in the ZL cluster in E10.5 sections of the midface and found that in *Bmpr1a* ectodermal loss the marker genes I analyzed are largely intact. In addition, the decreased proliferation that characterizes the ZL cell cluster is not observed. Interestingly, this is the third CL/P model investigated, together with *Pbx1/2* and *p63* mutants, in which *Cdkn1a/p21*, a marker of cell cycle arrest and the ZL cluster, is not altered in its high expression at the prominence tips. How precisely *Cdkn1a/p21* expression comes to be and remains unchanged, despite the presence of striking cell cycle arrest in these cells, remains elusive and does not seem to be reliant on either the PBX-WNT-IRF6-P63 signaling cascade or BMP signaling in the epithelium.

The *Bmpr1a*$^{f/f}$;*Crect*$^{Cre/+}$ mutant line may be a candidate for further study, especially through transcriptomic approaches such the scRNAseq we employed in the epithelium of *Pbx1/2* CL/P mutants. Analysis of mesenchyme effects caused by *Bmpr1a* loss in the ectoderm may also be worthwhile in this model, as crosstalk between the mesenchyme and epithelium is vital, and the nature of BMP signaling relies on extracellular ligands which can spread beyond individual tissue types.

3.4 Materials and Methods

Mice

Bmpr1a/Alk3 conditional *(Bmpr1a*$^{f/f}$) mice were originally generated as described in Mishina et al., 1995, and were acquired in our lab as a gift from Dr. Elizabeth Lacy at Sloan-Kettering Institute for Cancer Research. *Crect*^{Cre} mice were generated by the lab of Dr. Trevor Williams. Our *Crect*^{Cre} mice are kept on an FVB background strain as opposed to its original C56BL6/J background. To generate litters, a male heterozygous stud *Bmpr1a*$^{f/+}$; *Crect*$^{Cre/+}$ is bred to homozygous *Bmpr1a*$^{f/f}$ dams. The heterozygous male stud was generated through careful selection of weanlings without the highly penetrant deleterious phenotypes described. All animals were housed according to UCSF LARC animal husbandry guidelines and in accordance with UCSF IACUC protocol regulations. Following detection of copulatory plugs, males were separated from females. Females that were permitted to carry litters to term were carefully monitored at day 21 of gestation for birth of the litter to detect and humanely dispatch mutants in distress in a timely manner. No discernable effect on litter size or fecundity was detected in these crosses. For P0 litter phenotyping, a timed litter drop was phenotypically assessed and imaged in an expedited manner before CO2 shock and euthanasia via decapitation.

Embryo collection

Timed pregnancies were calculated using the evidence of a copulatory plug as embryonic day 0.5. Pregnant dams were sacrificed via CO2 inhalation and cervical dislocation in accordance with UCSF IACUC regulations. Embryos were collected at respective E10.5 or E15.5 timepoints and were dissected into ice cold PBS. 4 litters of crosses were examined at E15.5. Embryos for histological examination were fixed overnight at 4°C in 4% PFA/PBS with gentle agitation. The following day, fixed embryos were washed briefly in PBS and placed in 30% w/v sucrose/PBS solution and allowed to sink. Heads were decapitated from the bodies and embedded in 100% OCT or NEG-50 solution before being frozen and stored at -80°C. E15.5 embryos were dissected following imaging of the whole bodies. To assess the presence of a cleft palate, mandibles were removed using surgical scalpels to reveal the secondary palate.

RNAscope and Immunofluorescence:

Frozen blocks were at 14-µm-thick cryosections that were air dried at −20 °C for 1 h and stored at −80 °C. Slides carrying hindlimb bud sections were thawed and washed with 1X PBS to remove excess freezing medium before use. Slides were assayed using an RNAscope™ Multiplex Fluorescent Reagent Kit V2 following modified manufacturer's instructions protocols. Antigen target retrieval steps were skipped, and protease treatment limited to treating slides with Protease Plus for no longer than 10 min to avoid fragile embryonic tissues damage. Probe mixes were hybridized for 2 h at 40 °C in a HybEZ™ II Oven (Advanced Cell Diagnostics, Newark, CA). The appropriate HRP channels were developed with Opal™ 520, TSA™ Cy3 Plus, and

Cy5 Plus (PerkinElmer) dyes. Following DAPI staining and mounting with ProLong™ Gold (Invitrogen).

For sections involving immunofluorescence in conjunction with RNAscope, immunofluorescence assays were carried out after completion of the last RNAscope step involving blocking of the final HRP channel. A blocking solution was applied to slides containing 5% Normal Donkey Serum, 0.1% Tween-20 in 1X PBS for 45 minutes. Following this, antibodies raised against Mki67, EPCAM, and P21 were diluted in an antibody buffer (1:5 dilution of blocking solution to PBS-0.1% Tween-20 solution) and applied to the slides O/N at 4°C. The following day, primary antibodies were washed away and the slides washed 3 times 15 minutes each using antibody buffer. Secondary antibodies were diluted in antibody buffer (1:500) along with DAPI (1:1000) and applied to the slides for 2 hours at room temperature. Following washes to remove secondary antibody, slides were flicked dry and mounted using Prolong Gold antifade and coverslips.

Microscopy

Whole embryos were imaged using a Leica M20F5A dissecting microscope with a 1x/0.03 lens and variable zoom settings. Images were processed and exported using LAS-X software package for all Leica microscope images. Sections were imaged using a Zeiss Axio Observer.Z1 or a Zeiss Axiovert 200M with a Plan-Apochromat 20x/0.8 objective. Several Z-slices of sections were acquired. Sections with visible focused expression of RNAscope and immunofluorescence signal, were individually chosen, as the two were often at differing focal ranges. Images were processed using ZEN 3.6. Adjustments were made to reduce the background autofluorescence of the tissues by adjusting the absolute black and absolute white levels using a Best Fit adjustment where necessary.

3.5 Figures

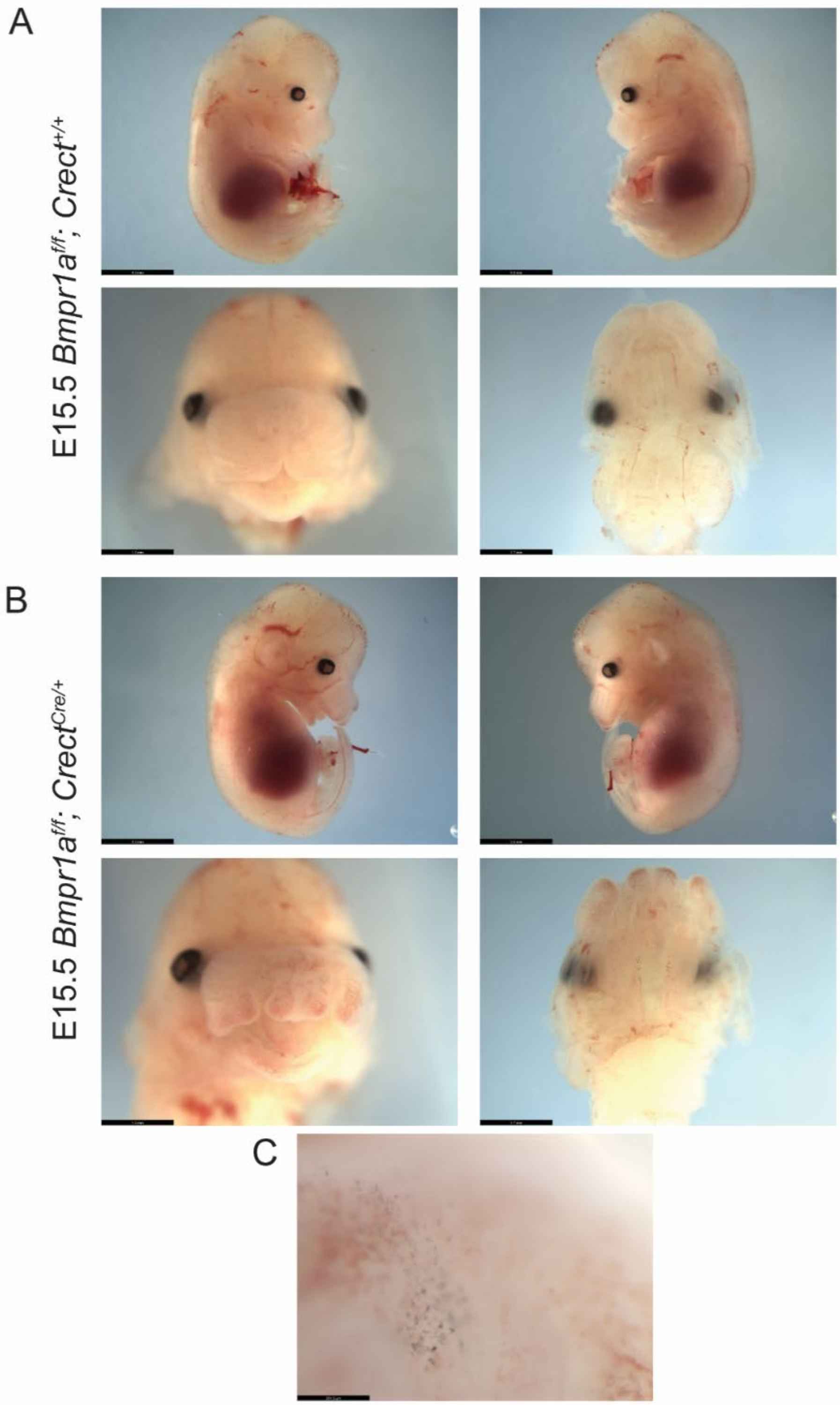
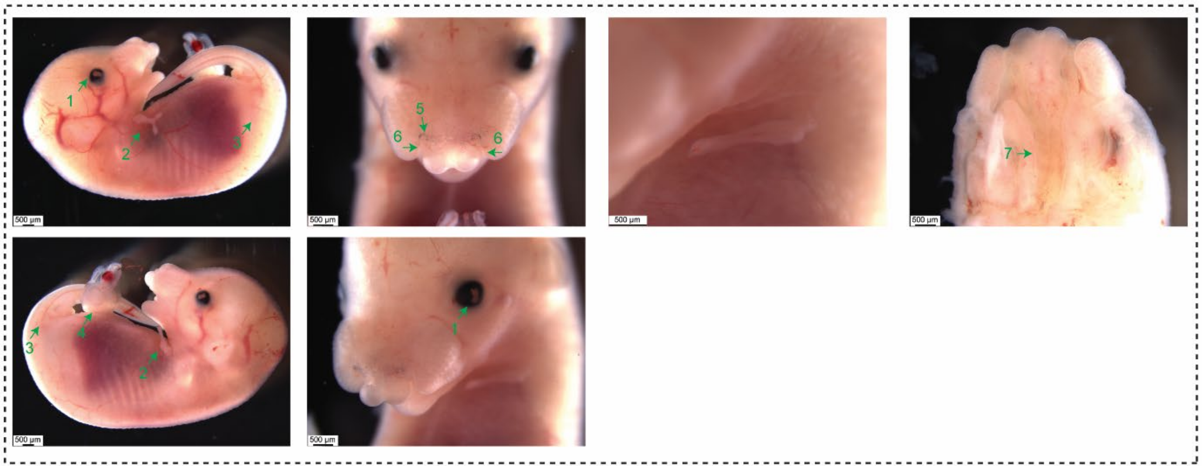


Figure 3.5.1: E15.5 Bmpr1a Homozygous Conditional Loss in the Epithelium Produces a Fully Penetrant Cleft Lip and Palate Phenotype
A comparison of conditional deletion of *Bmpr1a* using *Crect-Cre* recombinase to target ectodermal tissue demonstrates a craniofacial (Figure caption continued on the next page.)

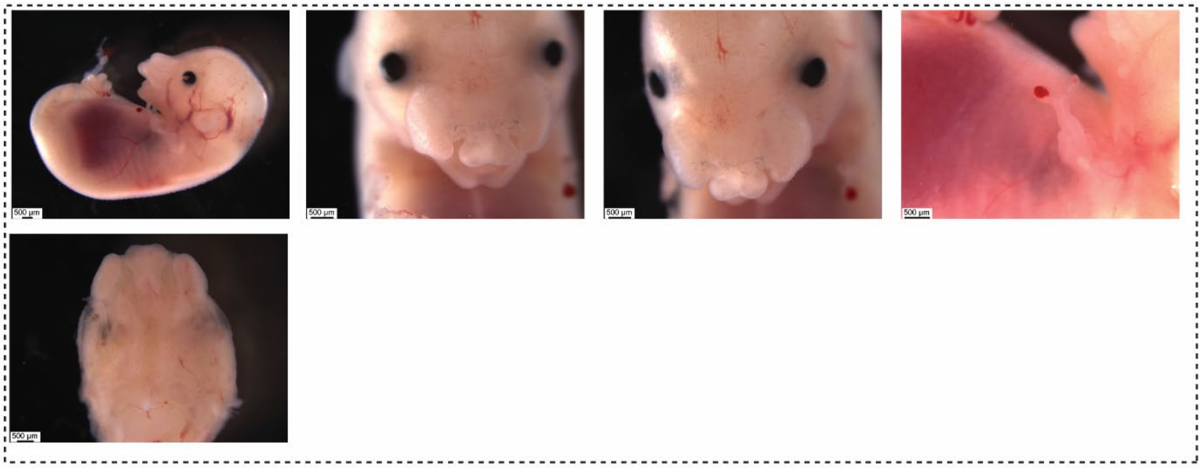
(Figure caption continued from the previous page.) phenotype similar to *Nestin-cre* driven deletion with fully penetrant bilateral cleft lip found in all cases of homozygous loss (n=7). A: *Crect-Cre* negative individuals are phenotypically normal and show no effect in the primary or secondary palate. No visible craniofacial defects are seen. B: Homozygous loss of *Bmpr1a* in the ectoderm produces the most severe craniofacial defects including a fully penetrant bilateral cleft lip and palate. C: A concentrated localization of pigmented melanocytes is found above the nostrils of all bilateral cleft lip mutants regardless of heterozygous or homozygous loss of *Bmpr1a* in the ectoderm.

A

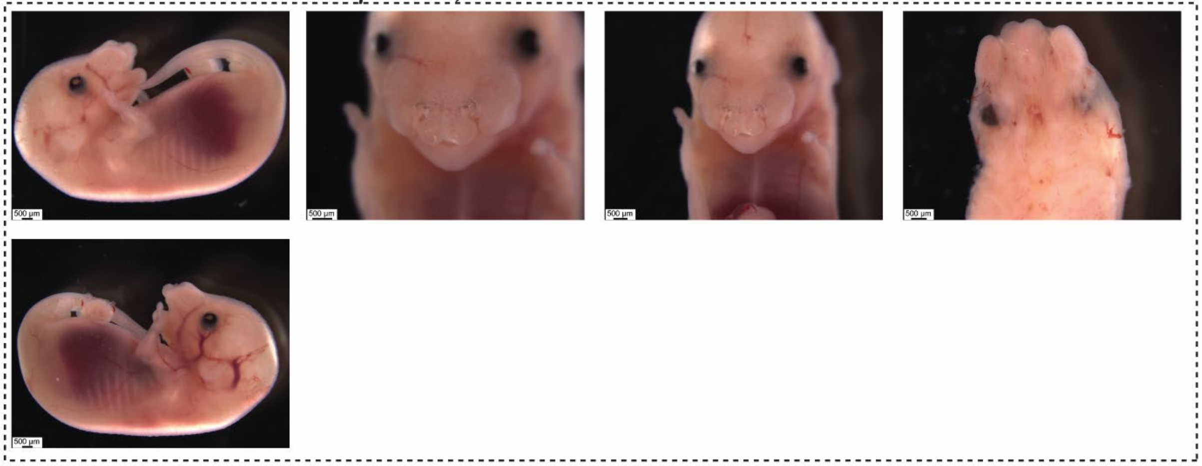
Individual 1: E15.5 *Bmpr1a*$^{f/+}$; *Crect*$^{Cre/+}$



Individual 2: E15.5 *Bmpr1a*$^{f/+}$; *Crect*$^{Cre/+}$



Individual 3: E15.5 *Bmpr1a*$^{f/+}$; *Crect*$^{Cre/+}$



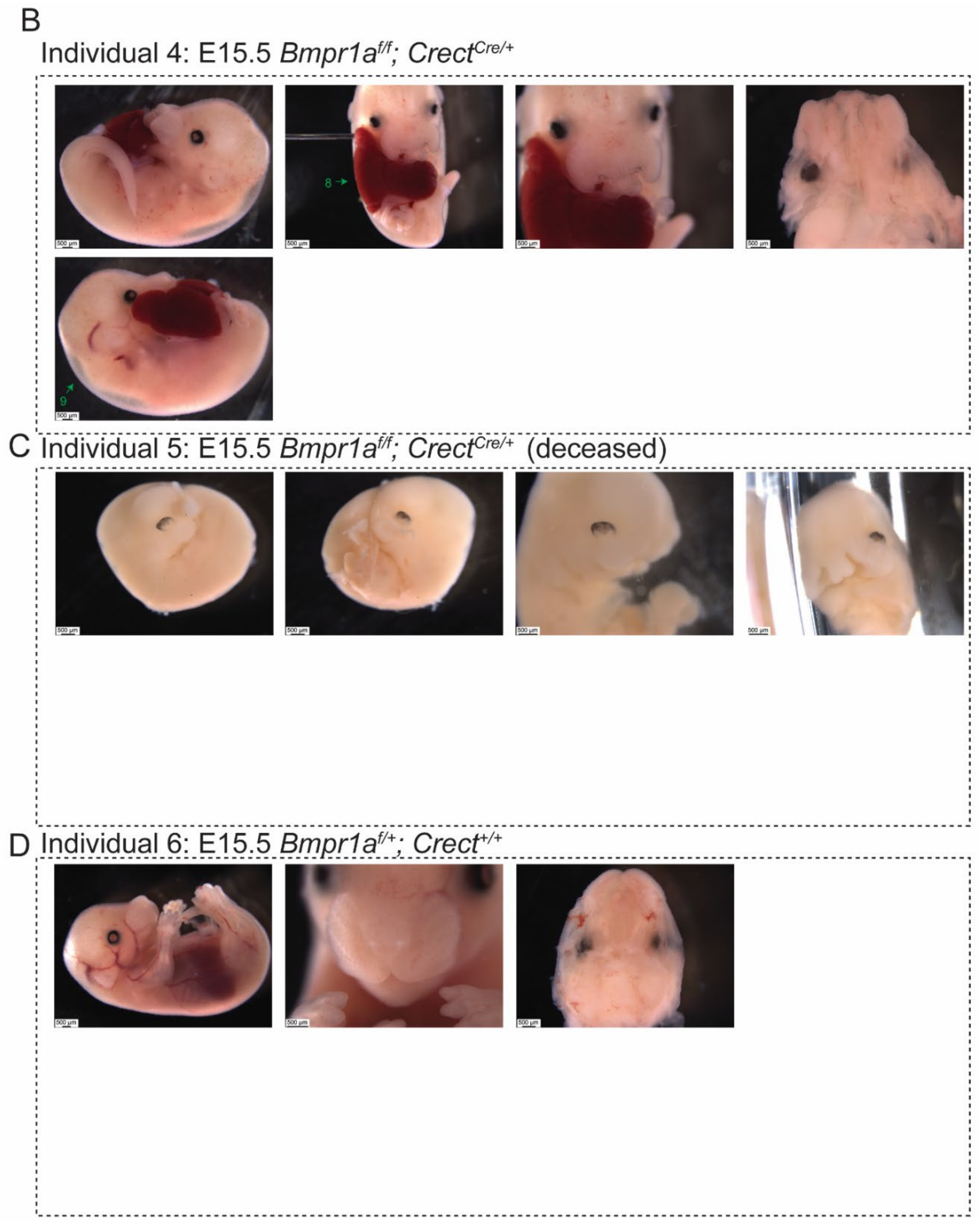


Figure 3.5.2: At E15.5 Ectodermal deletion of Bmpr1a produces an array of whole body phenotypes

Individuals of the same litter are shown at E15.5 including heterozygous and homozygous mutants. Across the embryo several key phenotypic traits are noted. A: A heterozygous individual is annotated with mutant traits as labeled: 1. Bilateral eye pupil asymmetry with the iris and/or lens appearing (Figure caption continued on the next page.)

(Figure caption continued from the previous page.) distorted and misshapen instead of round. 2. Forelimb buds are severely hypoplastic with oligodactyly of the autopod, and hypoplasia of sylopod and zeugopods 3. Hindlimbs or hindlimb buds are completely absent or undetectable at this scale. 4. Omphaloceles are present with intestines spilling out of a gap in the ventral body wall. 5. Concentrated clusters of melanocytes are seen above each nostril without clear association to any structure. 6. Bilateral cleft lip (primary palate) with no tissue connection seen between the lateral and nasal prominences. 7. Cleft palate (secondary palate) with palatal shelves separated by a large gap with no connecting tissue (mandible has been removed). Individual 2 and 3 are heterozygotes with variable hypoplasia of the forelimbs. B: Individual 4 is homozygous and possessed more extreme variations of the heterozygous phenotypes. 8: Omphalocele is larger and the entire liver is seen growing outside of the ventral body cavity. 9: Edema of the outer embryonic tissue layers is seen around the cranium. C: Individual 5 was a necrotic and calcified body with the appearance of an E11.5 or E12.5 embryo. This homozygous individual possessed bilateral cleft lip and no apparent limbs in line with the homozygous phenotype. D: Littermate control with no *Crect-Cre* allele and normal physiology. All scale bars 500μm.

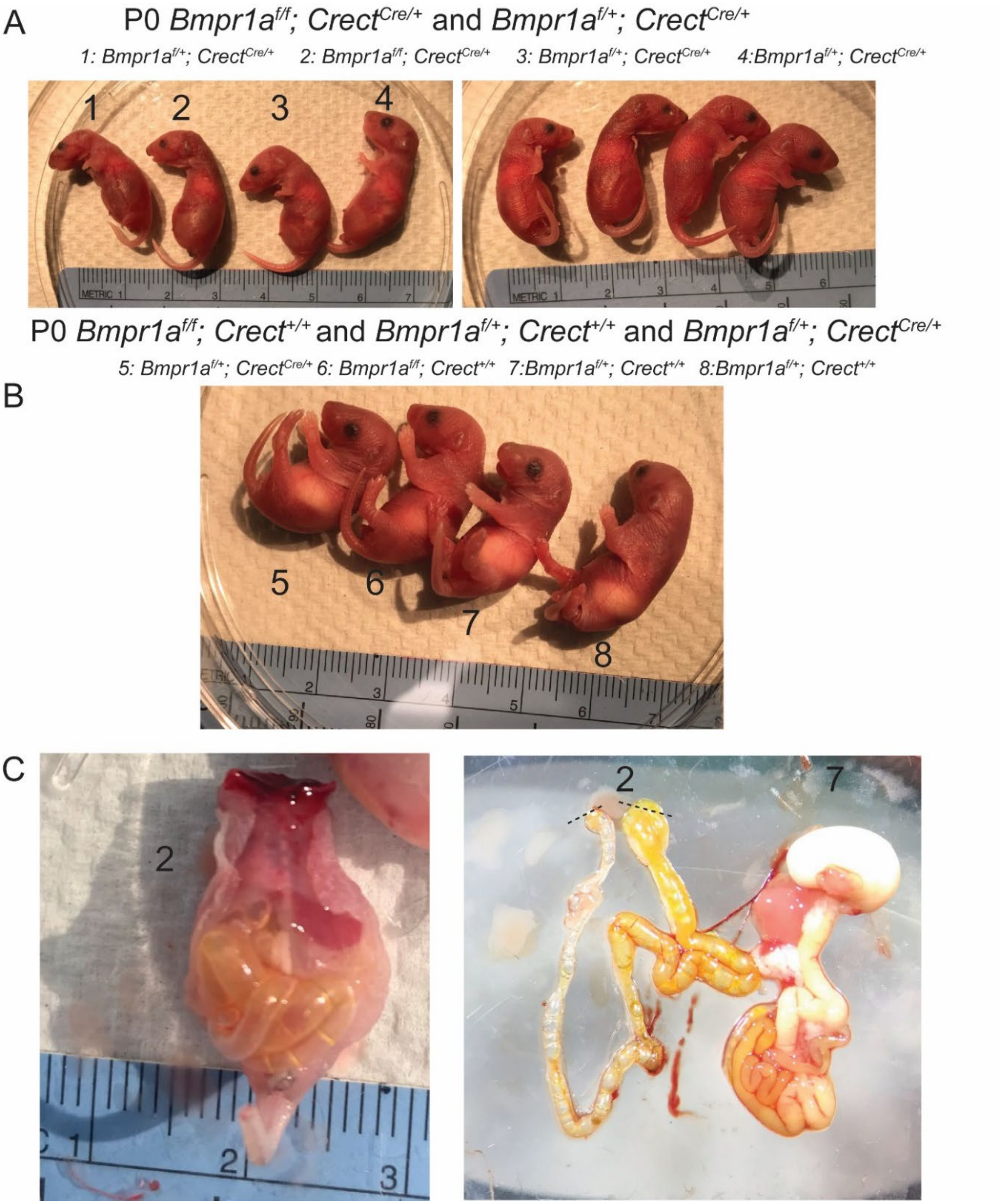


Figure 3.5.3: Mutant Bmpr1a litters are viable at P0 but with severe defects and perinatal morbidity

A single litter resulting from cross of a *Bmpr1a*$^{f/f}$ female to *Bmpr1a*$^{f/+}$*; Crect*$^{Cre/+}$ male. Litter size of 8 is shown at P0. Phenotypically normal individuals are arranged alongside littermates with visible gross anatomical defects. A: 4 pups at P0 with respective genotyping. All were born alive and with noticeable gasp-like breathing and bloating. Individuals 1, 3, and 4 are heterozygotes exhibiting variable hypoplastic forelimb (Figure caption continued on the next page.)

(Figure caption continued from the previous page.) phenotypes. Individual 1 possesses rudimentary hypoplastic hindlimbs. Individual 2 is homozygous and possessed the most extreme bloating. In lateral views, bulging intestines can be seen underneath the thin skin of the ventral body. 1-4 all possessed bilateral cleft lip and palate and did not possess obvious milk spots indicating poor suckling ability due to either mobility issues or facial clefts. B: Individuals 5-8 were phenotypically normal and had no signs of distress or labored breathing patterns. All possessed visible milk spots. Individual 5 is an example of non-full penetrance of the heterozygous conditional mutation phenotypes. Individual 5 had no anatomical abnormalities in the limbs, face, palate, or otherwise. C: Partially dissected individual 2 with ventral body wall removed showing gas filled intestines. No blockages were seen along the gastrointestinal tract. D: Individual 2 versus 7 gastrointestinal tract is displayed. During dissection, gas mostly was released from the upper small intestine sections of individual 2, yet distension of small intestine from gas pressure is seen above cecum (stomach has been removed). Individual 7 (right) gastrointestinal tract was phenotypically normal with no gas or bloating observed in the intestines. Milk is observed in the stomach unlike in individual 2 (not shown).

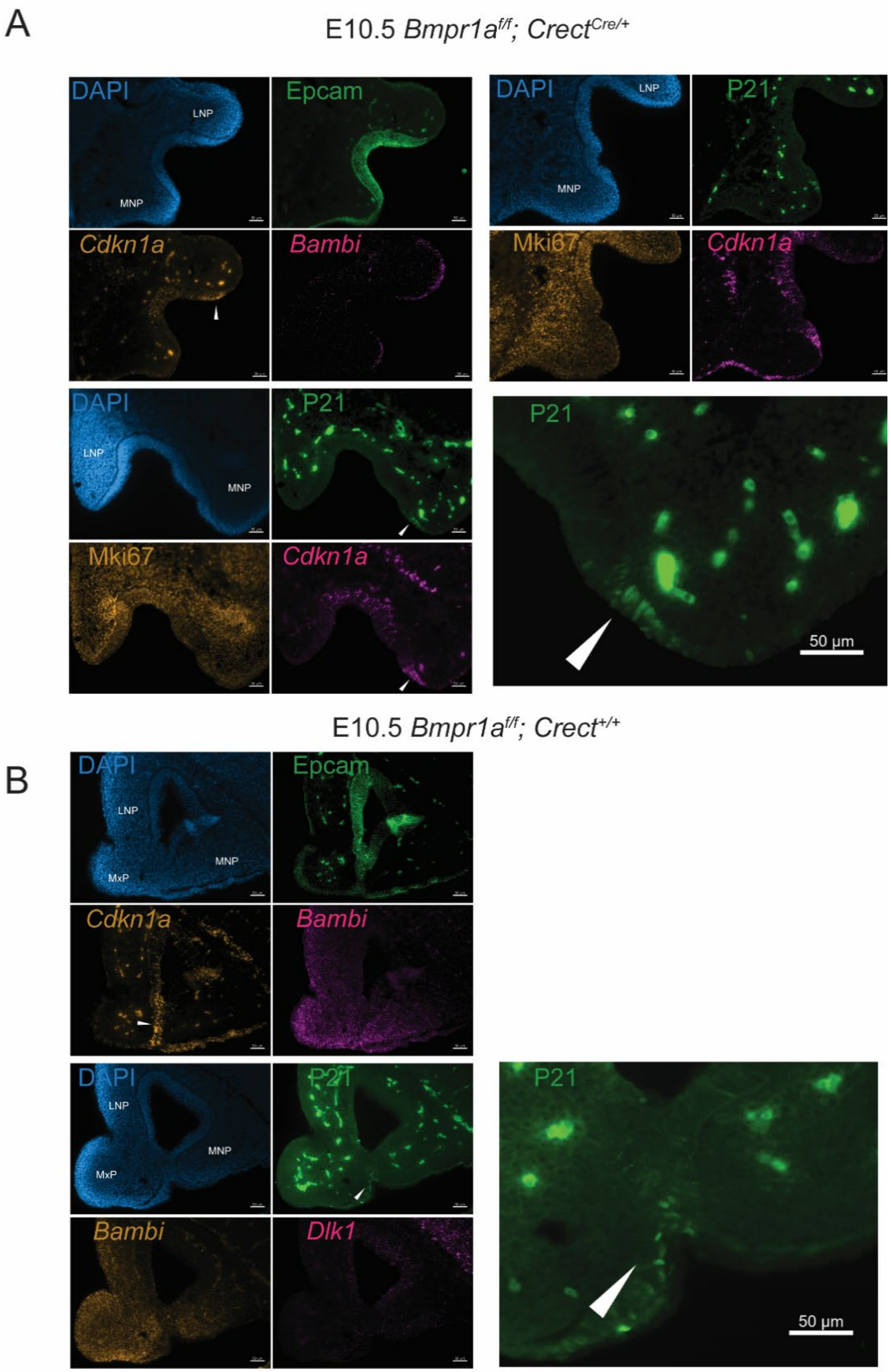

A
E10.5 Bmpr1a f/f; Crect Cre/+
DAPI
LNP
MNP
Epcam
DAPI
LNP
MNP
P21
Cdkn1a
Bambi
Mki67
Cdkn1a
DAPI
LNP
MNP
P21
Mki67
Cdkn1a
P21
50 µm
E10.5 Bmpr1a f/f; Crect +/+
B
DAPI
LNP
MNP
MxP
Epcam
Cdkn1a
Bambi
DAPI
LNP
MNP
MxP
P21
Bambi
Dlk1
P21
50 µm

Figure 3.5.4: Most cell cycle related gene expression during midface fusion is normal despite a cleft lip phenotype

A: transverse sections of midface prominences were created using homozygous bilateral cleft lip mutants and littermate controls with no discernable cleft lip phenotype at E10.5. A) The expression pattern of *Bambi,* and *Cdkn1a/p21,*both strongly DE in the ZL cluster of wildtype epithelium, was shown in homozygous *Bmpr1a* conditional mutant sections using RNAscope FISH. *Bambi* is localized to the epithelium of the prominence tips and can be seen in lower quantities in the underlying mesenchyme of lateral and medial prominences. *Cdkn1a/p21* expression is robust and highly localized at the distal epithelium of lateral and medial prominences in mutant sections. Immunofluorescence against EPCAM protein delineates the border of epithelium versus mesenchyme. No discernable disruption of the epithelium can be seen. KI67 expression is not shown to be lower in areas of high *Cdkn1a/p21* expression. P21 protein overlaps strongly with *Cdkn1a/p21* transcript but does not indicate an area of lowered KI67 signal. B: Littermate control sections of *Cre*-negative individual with fusing lambdoidal junction demonstrates normal expression patterns of *Bambi, Dlk1,* and *Cdkn1a/p21.* Normal Epcam denoted epithelial layers are detected using immunofluorescence. P21 expression overlapping a region of high *Cdkn1a/p21* transcript expression is seen at the fusion site between the medial and lateral nasal prominences. All scale bars 50µm.

CHAPTER 4: CONCLUDING REMARKS

To make a face as unique as oneself requires an elaborate coordination of developmental events performed by cellular behaviors that are dictated by genetic programming. At all levels, these events that shape the face can go astray leading to unexpected outcomes. In trying to decipher the role of the midface epithelium through the lens of transcriptomics during midface morphogenesis, I investigated a cross-section of cellular behavior, transcribed genetic instructions, and developmental outcomes. This analysis created a better understanding of both the foundational mechanisms of face shaping architecture and a starting point to compare against models of craniofacial birth defects. The research presented here demonstrated that through the use of mouse models of CL/P, subtle alterations in a discrete population of the cephalic epithelium can be tied to drastic outcomes in morphology.

This thesis highlights the first epithelium-only analysis of the lambdoidal junction from pre-formation to completion of face prominence fusion and in doing so creates an atlas of this tissue's inherent heterogeneity. The validation experiments performed highlight regions of distinct transcriptomic states that are precursors to mouth, nose, and eye formation. These regions are found to contribute to the formation of a transcriptomic state at the tips of the midface prominences, a region described as the zippering lambda epithelium. By demonstrating how multiple neighboring populations contribute to this cell cluster, it is thus believed that its transcriptome comes about through a mechanism at least partially dictated by physical location rather than a timed state preprogrammed to occur in all contributing predecessors. What leads to this state could be the circumstance of patterning gradients and signaling found throughout the head. Cells of the zippering lambda epithelium are located at the most distal region of the midfacial prominences, and by their nature are the furthest from the brain at the time of outgrowth. As the brain is a potent source of retinoic acid, several patterning systems rely on its gradient for both initial formation and maintenance. Other underlying signaling mechanisms, including paracrine signaling from underlying mesenchyme can be thought to also have a functional impact on the ZL cell cluster. While the ZL cluster was disrupted and shown to have decreased expression of several cell cycle arrest genes, it still was present in our scRNASeq data sets, and transcript levels for a major marker, *Cdkn1a/p21,* were unchanged in all CL/P mutant models. The formation of the ZL cluster therefore may be partially attributed to interplay of epithelial and mesenchymal signaling not completed abolished by epithelial specific deletions or may be heavily compensated by redundant mechanisms outside of the gene regulatory networks altered by our genetic ablations.

Developmental Senescence and a Senescence Associated Secretory Phenotype in the Zippering Lambda Epithelium: Analogs to Limb Development

A fascinating connection to the ZL cell cluster may lie in limb development. There are major similarities between the developing limbs and the facial prominences (B. T. Truong & Artinger, 2021). It is even thought that in some ways, the limbs may be rudimentary head-like appendages, or that the head is an advanced evolutionary cousin of the limb. Outgrowth of limbs begins with a trigger of HOX gene patterning allowing for initial thickening and outgrowth through FGF pathways. SHH allows for differentiation between anterior and posterior edges of the limb bud, and continued maintenance of the outgrowing bud relies on WNT and BMP signaling to sustain a complex patterning system involving the underlying mesenchyme and overlying AER epithelium. SHH, BMP, FGF, and WNT all contribute to the outgrowth and patterning of the midfacial prominences as well. Because of shared PBX-related phenotypes, the work of our lab has examined limb formation through the lens of gene regulatory networks involving PBX, HAND2, and transcriptional cofactors HOX and MEIS. These factors collaborate to drive the expression of critical downstream targets needed to properly create a limb. The disruption of midface prominence fusion via loss of PBX transcription factors can be seen as analogous to limb malformation from PBX deletion.

The AER (epithelium) of the limb possesses a close analog to the ZL cell cluster that cannot be overlooked and may be the key to understanding *why* the ZL cluster is formed. It was found that a senescence associated secretory phenotype (SASP) lies at the distal apical edge of the limb bud AER (Storer et al 2013). This region exhibits all the emerging hallmarks of developmental senescence including high expression of SAβ-Gal, low proliferation (measured through BrdU), high p53/p21 and p16/RB tumor suppressor network protein expression, and SASP signaling of these non-proliferative cells to their external environment through cytokines, growth factors, and extracellular matrix remodeling proteins. The senescent cells of the AER are fated to undergo apoptosis and macrophage-mediated clearance, not dissimilar to the apoptotic fate of many ZL cells in the midface. These AER senescent cells possess remarkable upregulation of *Cdkn1a, Cdkn2b, Bmp4, Igfbp5, Tgfb1, Wnt5a,* and *Rb1*, all of which are DEGs for the ZL cluster of the midface as well. P21 loss was believed to not present with developmental phenotypes (Brugarolas et al 1995, Deng et al 1995), yet it was shown that the AER of *Cdkn1a/p21*-deficient mice had dysregulation of *Fgf8, Msx2, Meis1,* and *Shh,* again, genes upregulated in and related to midface prominence development. P21 disruption limited the signaling abilities of the AER, and in *in-vitro* cells that were already senescent, knockdown of *Cdkn1a/p21* disrupted the pattern of patterning gene expression, suggesting that p21 is a key to maintenance of SASP.

The AER SASP phenotype was believed to be induced by pERK signaling in the underlying mesenchyme through paracrine signaling. Given the striking similarities of the midface ZL cell cluster expression patterns and the reported AER SASP, a logical next step in furthering our understanding of the ZL would be to investigate it through the lens of developmental senescence. An explanation for *Cdkn1a/p21* expression in the ZL may similarly lie in induction by ERK signaling in the underlying mesenchyme of the midface. Intriguingly, Storer et al 2013, despite staining for SAβ-Gal activity throughout the body of mouse and chick embryos at the stages of midfacial fusion, failed to remark on the face. As is the case with much published literature, observations of the embryo tend to leave the face out of descriptions of gene expression patterns and phenotyping, and are plagued by a tendency to not remove the head from the embryo to carefully observe the face. As the head is tucked into the body and often obscured by the tail and forelimbs that curl up during embryo phenotyping, most details of facial structures are lost. This may also be the case for *Cdkn1a/p21* loss of function in mouse embryos, in which initial phenotyping may have failed to perceive subtle changes to the shape of the midfacial prominences or primary palate.

Human OFC Risk Data and PI3K-AKT Signaling

Human CL/P risk alleles found through the intersection of the scRNAseq data sets with genomic analysis of OFC case trios created a wealth of information that can be used to further understand normal midface prominence fusion and its disruption in CL/P human patients. Several genes relating to PI3K-AKT signaling were found to be both upregulated in the ZL cell cluster and to carry variants associated with orofacial cleft risk, including *Grb10* and *Pik3r1*. Several PI3K-AKT signaling genes, which in turn are linked to insulin-growth factor (IGF) signaling, are also among the differentially expressed genes in *Pbx1/2* mutant epithelium, including *Pik3r1, Gsk3b, Foxo3,* and *Foxo1,* which are all expressed at significantly higher levels in mutant cephalic epithelium **(See Appendix 1, *Pik3r1* expression in the midface)**. These expression patterns may be linked simply to the known canonical role of PI3K-AKT signaling in regulation of cell proliferation, in which cell cycle arrest genes including *Cdkn1a/p21* and *Cdkn1b/p27* are downstream targets of FOXO transcription factors. Yet, increased expression of *Pik3r1, Gsk3b, Foxo3,* and *Foxo1* in *Pbx1/2* mutants would seemingly increase the cell cycle arrest state seen in the ZL cluster. Why these genes are upregulated in *Pbx1/2* mutants, in which the ZL cluster cell cycle arrest is disrupted, is mysterious and may indicate a compensatory mechanism at work through feedback loops. If there are human patients exhibiting OFC with disruptive mutations in *Pik3r1* or *Grb10,* it may be related to hyper-proliferation of cells and a failure to induce the ZL transcriptomic state. Further evidence is needed, including examination and validation of these genes in CL/P mutants as well as careful parsing of the mutant alleles found in OFC patients. To further explore any new OFC risk candidates, the next step may be the generation of mutant alleles in genetically engineered mice. This, in turn, may also yield an increase to the number of CL/P mouse models, which we are still sorely lacking.

Foxn3 Is a Potential Upstream Regulator of Cell Cycle Arrest that Warrants Further Analysis in Zippering Lambda Epithelium

An effort to incorporate the various gene expression patterns together into a unified pathway may be too simplistic in solving what most likely involves both parallel and crossed webs of signaling pathways in the ZL cluster alone. A proposed pathway involving genes I can reasonably implicate in cell cycle arrest, including *Zfhx3, Pi3kr1, Bcl11b, Foxo1, Foxo3, Cdkn1a, Rb1,* and *Gadd45g,* is just one of many that could prove true or not. I have added to this equation a gene that was not discussed yet, despite its major differential expression in the ZL cell cluster, *Foxn3*.

I have performed validation of *Foxn3* and shown through FISH that it is indeed enriched at the prominence fusion site **(Appendix 2)**. I have also demonstrated that it is expressed at the tips of face prominences in *Pbx1/2* conditional mutants with CL/P. *Foxn3* is not reported in relation to midface prominence fusion in the literature. *Foxn3* has remained understudied compared to other genes encoding FOX transcription factor family members but has ties to cell cycle arrest that are well documented and unique bispecific binding to both canonical forkhead and alternative forkhead-like binding motifs presumably through partners. *Foxn3* is expressed in and during the development of many organs including liver, pancreas, kidneys, lungs, and bone marrow (Kong et al., 2019). It has a role in suppressing the cell cycle through MYC (S. H. Lee et al., 2018). In *Drosophila,* it has been tied to S-phase cell cycle arrest and in yeast it promotes DNA-damage inducible cell cycle arrest at G2/M (Grassi et al., 2017; Robert et al., 2011). In cancer, it is believed to be an important tumor suppressor protein. It is strongly downregulated in multiple cancer types including melanoma, liver, lung, colon, and laryngeal cancer as well as lymphoma (Kong et al., 2019). It is also believed to be a suppressor of growth through inhibition of inhibitors of TGFβ (Kuki et al., 2017). Overexpression of *FOXN3* inhibits glioma cell growth via interactions with AKT/MDM2/P53 (C. Wang et al., 2021). Lastly, published ChIPSeq experiments using cancer cell lines demonstrate that *FOXN3* binds to the promoters of nearly all genes discussed in relation to the ZL cluster, including: *CDKN1A, CDKN1B, CDKN2A, CDKN2B, GADD45G, RB1, BCL11B, BAMBI, BMP4, ZFHX3, PBX3, TGFB1, TGFB2,* and *PIK3R1* (Rogers et al., 2019). Like many transcription factors, it may gain specificity through interactions with binding partners, but may also be upstream as a master regulator of cell cycle arrest that operates to protect cells from overgrowth in the face of DNA damage and oncogenesis.

We know that the FOX transcription factor family is large and varied. Some are implicated in cell cycle regulation, and some are found to interact with players such as PBX as well (Trasanidis et al., 2022; Zhu et al., 2017). However, no other FOX family member comes close to matching the differential expression of *Foxn3* in the ZL cell cluster. The highly localized expression of a tumor suppressor gene like *Foxn3* in the ZL has clear connections to cell cycle arrest. Further studies are needed to better understand the role of this transcription factor during the formation of the ZL cluster and cell cycle arrest. ChIPSeq experiments using FOXN3 as a target may reveal target genes and regulatory elements in midface epithelium. As the field continues to allow for lower material input for the interrogation of such binding through novel approaches such as cleavage under targets & release using nuclease (CUT&RUN) or & tagmentation (CUT&Tag), isolation of the small number of ZL cells in the lambdoidal junction epithelium may yield promising results. I made several attempts to isolate ZL cluster cells using FACS based approaches, targeting both ITGA6, BAMBI, as well as through the use of Hoechst/Pyronin Y-based sorting to enrich and isolate G0 arrested cells. While none of these approaches yielded enrichment for cells that matched the ZL cluster gene expression, optimization may improve the results. Simple, yet precise, microdissection, in the end, may yield the most promising results and allow for the assessment of genome-wide binding of FOXN3 and an array of other transcription factors in the ZL cell cluster.

The zippering lambda epithelium will necessitate careful and refined approaches to exploring and understanding the transient formation of this small cellular population. Lineage tracing experiments and labeling strategies to track this population in real-time, including the use of fluorescent reporter constructs that indicate cell cycle states such as FUCCI2A, are methods proposed that could be coupled with live-imaging and explant culture (Mort et al., 2014). Explant culture of the midfacial tissues and imaging of said tissues would be challenging but rewarding, especially if one could capture the fusion site and bear witness to the fusion of the lambdoidal junction, and thus formation of the ZL cluster.

Bibliography

Ahtiainen, L., Uski, I., Thesleff, I., & Mikkola, M. L. (2016). Early epithelial signaling center governs tooth budding morphogenesis. *J Cell Biol*, *214*(6), 753–767. https://doi.org/10.1083/jcb.201512074

Akiyama, R., Kawakami, H., Taketo, M. M., Evans, S. M., Wada, N., Petryk, A., & Kawakami, Y. (2014). Distinct populations within Isl1 lineages contribute to appendicular and facial skeletogenesis through the β-catenin pathway. *Developmental Biology*, *387*(1), 37–48. https://doi.org/10.1016/J.YDBIO.2014.01.001

Alkuraya, F. S., Saadi, I., Lund, J. J., Turbe-Doan, A., Morton, C. C., & Maas, R. L. (2006). SUMO1 haploinsufficiency leads to cleft lip and palate. *Science (New York, N.Y.)*, *313*(5794), 1751. https://doi.org/10.1126/SCIENCE.1128406

Ashique, A. M., Fu, K., & Richman, J. M. (2002). Endogenous bone morphogenetic proteins regulate outgrowth and epithelial survival during avian lip fusion. *Development (Cambridge, England)*, *129*(19), 4647–4660. https://dev.biologists.org/content/129/19/4647.long

Bergholz, J., & Xiao, Z. X. (2012). Role of p63 in Development, Tumorigenesis and Cancer Progression. *Cancer Microenvironment : Official Journal of the International Cancer Microenvironment Society*, *5*(3), 311–322. https://doi.org/10.1007/S12307-012-0116-9

Bhattacharya, S., Khanna, V., & Kohli, R. (2009). Cleft lip: The historical perspective. *Indian Journal of Plastic Surgery : Official Publication of the Association of Plastic Surgeons of India*, *42*(Suppl), S4. https://doi.org/10.4103/0970-0358.57180

Bishop, M. R., Diaz Perez, K. K., Sun, M., Ho, S., Chopra, P., Mukhopadhyay, N., Hetmanski, J. B., Taub, M. A., Moreno-Uribe, L. M., Valencia-Ramirez, L. C., Restrepo Muñeton, C. P., Wehby, G., Hecht, J. T., Deleyiannis, F., Weinberg, S. M., Wu-Chou, Y. H., Chen, P. K., Brand, H., Epstein, M. P., … Leslie, E. J. (2020). Genome-wide Enrichment of De Novo Coding Mutations in Orofacial Cleft Trios. *American Journal of Human Genetics*, *107*(1), 124–136. https://doi.org/10.1016/j.ajhg.2020.05.018

Bonaïti-Pellié, C., Briard-Guillemot, M. L., Feingold, J., & Frézal, J. (1975). Associated congenital malformations in retinoblastoma. *Clinical Genetics*, *7*(1), 37–39. https://doi.org/10.1111/J.1399-0004.1975.TB00360.X

Bornstein, S., Trasler, D. G., & Fraser, F. C. (1970). Effect of the uterine environment on the frequency of spontaneous cleft lip in CL/Fr mice. *Teratology*, *3*(4), 295–298. https://doi.org/10.1002/TERA.1420030403

Brantley, S. E., & Di Talia, S. (2021). Cell cycle control during early embryogenesis. *Development (Cambridge, England)*, *148*(13). https://doi.org/10.1242/DEV.193128

Brunskill, E. W., Potter, A. S., Distasio, A., Dexheimer, P., Plassard, A., Aronow, B. J., & Potter, S. S. (2014). A gene expression atlas of early craniofacial development. *Developmental Biology*, *391*(2), 133–146. https://doi.org/10.1016/J.YDBIO.2014.04.016

Bush, J. O., Lan, Y., & Jiang, R. (2004). The cleft lip and palate defects in Dancer mutant mice result from gain of function of the Tbx10 gene. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(18), 7022–7027. https://doi.org/10.1073/PNAS.0401025101

Cai, S., Si, N., Wang, Y., & Yin, N. (2023). Transcriptomic analysis of the upper lip and primary palate development in mice. *Frontiers in Genetics*, *13*. https://doi.org/10.3389/FGENE.2022.1039850

Campbell, G. J., Hands, E. L., & Van de Pette, M. (2020). The Role of CDKs and CDKIs in Murine Development. *International Journal of Molecular Sciences 2020, Vol. 21, Page 5343*, *21*(15), 5343. https://doi.org/10.3390/IJMS21155343

Capellini, T. D., Di Giacomo, G., Salsi, V., Brendolan, A., Ferretti, E., Srivastava, D., Zappavigna, V., & Selleri, L. (2006). Pbx1/Pbx2 requirement for distal limb patterning is mediated by the hierarchical control of Hox gene spatial distribution and Shh expression. *Development (Cambridge, England)*, *133*(11), 2263–2273. https://doi.org/10.1242/DEV.02395

Capellini, T. D., Zewdu, R., Di Giacomo, G., Asciutti, S., Kugler, J. E., Di Gregorio, A., & Selleri, L. (2008). Pbx1/Pbx2 govern axial skeletal development by controlling Polycomb and Hox in mesoderm and Pax1/Pax9 in sclerotome. *Developmental Biology*, *321*(2), 500–514. https://doi.org/10.1016/J.YDBIO.2008.04.005

Casimiro, M. C., Crosariol, M., Loro, E., Li, Z., & Pestell, R. G. (2012). Cyclins and Cell Cycle Control in Cancer and Disease. *Genes & Cancer*, *3*(11–12), 649. https://doi.org/10.1177/1947601913479022

Chalazonitis, A., Pham, T. D., Li, Z., Roman, D., Guha, U., Gomes, W., Kan, L., Kessler, J. A., & Gershon, M. D. (2008). Bone morphogenetic protein regulation of enteric neuronal phenotypic diversity: relationship to timing of cell cycle exit. *The Journal of Comparative Neurology*, *509*(5), 474–492. https://doi.org/10.1002/CNE.21770

Chen, J., Bush, J. O., Ovitt, C. E., Lan, Y., & Jiang, R. (2007). The TGF-beta pseudoreceptor gene Bambi is dispensable for mouse embryonic development and postnatal survival. *Genesis (New York, N.Y. : 2000)*, *45*(8), 482–486. https://doi.org/10.1002/DVG.20320

Choi, Y. H., & Kim, J. K. (2019). Dissecting Cellular Heterogeneity Using Single-Cell RNA Sequencing. *Molecules and Cells*, *42*(3), 189. https://doi.org/10.14348/MOLCELLS.2019.2446

Deng, C., Zhang, P., Wade Harper, J., Elledge, S. J., & Leder, P. (1995). Mice lacking p21CIP1/WAF1 undergo normal development, but are defective in G1 checkpoint control. *Cell*, *82*(4), 675–684. https://doi.org/10.1016/0092-8674(95)90039-X

Dohn, M., Zhang, S., & Chen, X. (2001). p63α and ΔNp63α can induce cell cycle arrest and apoptosis and differentially regulate p53 target genes. *Oncogene 2001 20:25*, *20*(25), 3193–3205. https://doi.org/10.1038/sj.onc.1204427

Dong, G., Ma, G., Wu, R., Liu, J., Liu, M., Gao, A., Li, X., Jun, A., Liu, X., Zhang, Z., Zhang, B., Fu, L., & Dong, J. T. (2020). ZFHX3 Promotes the Proliferation and Tumor Growth of ER-Positive Breast Cancer Cells Likely by Enhancing Stem-Like Features and MYC and TBX3 Transcription. *Cancers 2020, Vol. 12, Page 3415*, *12*(11), 3415. https://doi.org/10.3390/CANCERS12113415

Dong, X., Landford, W. N., Hart, J., Risolino, M., Kaymakcalan, O., Jin, J., Toyoda, Y., Ferretti, E., Selleri, L., & Spector, J. A. (2017). Toward Microsurgical Correction of Cleft Lip Ex Utero through Restoration of Craniofacial Developmental Programs. *Plastic and Reconstructive Surgery*, *140*(1), 75–85. https://doi.org/10.1097/PRS.0000000000003417

Dudas, M., Kim, J., Li, W. Y., Nagy, A., Larsson, J., Karlsson, S., Chai, Y., & Kaartinen, V. (2006). Epithelial and ectomesenchymal role of the type I TGF-beta receptor ALK5 during facial morphogenesis and palatal fusion. *Developmental Biology*, *296*(2), 298–314. https://doi.org/10.1016/J.YDBIO.2006.05.030

Fabik, J., Kovacova, K., Kozmik, Z., & Machon, O. (2020). Neural crest cells require Meis2 for patterning the mandibular arch via the Sonic hedgehog pathway. *Biology Open*, *9*(6). https://doi.org/10.1242/BIO.052043

Ferretti, E., Li, B., Zewdu, R., Wells, V., Hebert, J. M., Karner, C., Anderson, M. J., Williams, T., Dixon, J., Dixon, M. J., Depew, M. J., & Selleri, L. (2011). A Conserved Pbx-Wnt-p63-Irf6 Regulatory Module Controls Face Morphogenesis by Promoting Epithelial Apoptosis. *Developmental Cell*, *21*(4), 627–641. https://doi.org/10.1016/j.devcel.2011.08.005

Firnberg, N., & Neubüser, A. (2002). FGF signaling regulates expression of Tbx2, Erm, Pea3, and Pax3 in the early nasal region. *Developmental Biology*, *247*(2), 237–250. https://doi.org/10.1006/dbio.2002.0696

Francis-West, P. H., Tatla, T., & Brickell, P. M. (1994). Expression patterns of the bone morphogenetic protein genes Bmp-4 and Bmp-2 in the developing chick face suggest a role in outgrowth of the primordia. *Developmental Dynamics*, *201*(2), 168–178. https://doi.org/10.1002/AJA.1002010207

García-Gutiérrez, L., Delgado, M. D., & León, J. (2019). MYC Oncogene Contributions to Release of Cell Cycle Brakes. *Genes 2019, Vol. 10, Page 244*, *10*(3), 244. https://doi.org/10.3390/GENES10030244

Ghioni, P., Bolognese, F., Duijf, P. H. G., Bokhoven, H. van, Mantovani, R., & Guerrini, L. (2002). Complex Transcriptional Effects of p63 Isoforms: Identification of Novel Activation and Repression Domains. *Molecular and Cellular Biology*, *22*(24), 8659. https://doi.org/10.1128/MCB.22.24.8659-8668.2002

Gong, S. G., White, N. J., & Sakasegawa, A. Y. (2000). The Twirler mouse, a model for the study of cleft lip and palate. *Archives of Oral Biology*, *45*(1), 87–94. https://doi.org/10.1016/S0003-9969(99)00101-6

Gong, S.-G., & Eulenberg, R. L. (2001). Palatal development in Twirler mice. *The Cleft Palate-Craniofacial Journal : Official Publication of the American Cleft Palate-Craniofacial Association*, *38*(6), 622–628. https://doi.org/10.1597/1545-1569_2001_038_0622_PDITM_2.0.CO_2

Graf, D., Malik, Z., Hayano, S., & Mishina, Y. (2016). Common mechanisms in development and disease: BMP signaling in craniofacial development. *Cytokine & Growth Factor Reviews*, *27*, 129–139. https://doi.org/10.1016/J.CYTOGFR.2015.11.004

Grassi, D., Franz, H., Vezzali, R., Bovio, P., Heidrich, S., Dehghanian, F., Lagunas, N., Belzung, C., Krieglstein, K., & Vogel, T. (2017). Neuronal Activity, TGFβ-Signaling and Unpredictable Chronic Stress Modulate Transcription of Gadd45 Family Members and DNA Methylation in the Hippocampus. *Cerebral Cortex*, *27*(8), 4166–4181. https://doi.org/10.1093/CERCOR/BHX095

Green, R. M., Feng, W., Phang, T., Fish, J. L., Li, H., Spritz, R. A., Marcucio, R. S., Hooper, J., Jamniczky, H., Hallgrímsson, B., & Williams, T. (2015). Tfap2a-dependent changes in mouse facial morphology result in clefting that can be ameliorated by a reduction in Fgf8 gene dosage. *DMM Disease Models and Mechanisms*, *8*(1), 31–43. https://doi.org/10.1242/DMM.017616/-/DC1

Green, R. M., Leach, C. L., Diewert, V. M., Aponte, J. D., Schmidt, E. J., Cheverud, J. M., Roseman, C. C., Young, N. M., Marcucio, R. S., & Hallgrimsson, B. (2019). Nonlinear gene expression-phenotype relationships contribute to variation and clefting in the A/WySn mouse. *Developmental Dynamics : An Official Publication of the American Association of Anatomists*, *248*(12), 1232–1242. https://doi.org/10.1002/DVDY.110

Gritli-Linde, A. (2008). Chapter 2 The Etiopathogenesis of Cleft Lip and Cleft Palate: Usefulness and Caveats of Mouse Models. *Current Topics in Developmental Biology*, *84*, 37–138. https://doi.org/10.1016/S0070-2153(08)00602-9

Gulati, G. S., Sikandar, S. S., Wesche, D. J., Manjunath, A., Bharadwaj, A., Berger, M. J., Ilagan, F., Kuo, A. H., Hsieh, R. W., Cai, S., Zabala, M., Scheeren, F. A., Lobo, N. A., Qian, D., Yu, F. B., Dirbas, F. M., Clarke, M. F., & Newman, A. M. (2020). Single-cell transcriptional diversity is a hallmark of developmental potential. *Science*, *367*(6476), 405–411. https://doi.org/10.1126/SCIENCE.AAX0249/SUPPL_FILE/AAX0249_TABLE_S8.XLSB

Helton, E. S., Zhang, J., & Chen, X. (2007). The proline-rich domain in p63 is necessary for the transcriptional and apoptosis-inducing activities of TAp63. *Oncogene 2008 27:20*, *27*(20), 2843–2850. https://doi.org/10.1038/sj.onc.1210948

Helton, E. S., Zhu, J., & Chen, X. (2006). The unique NH2-terminally deleted (ΔN) residues, the PXXP motif, and the PPXY motif are required for the transcriptional activity of the ΔN variant of p63. *Journal of Biological Chemistry*, *281*(5), 2533–2542. https://doi.org/10.1074/jbc.M507964200

Hinrichsen, K. (1985). The early development of morphology and patterns of the face in the human embryo. *Advances in Anatomy, Embryology, and Cell Biology*, *98*. https://doi.org/10.1007/978-3-642-70754-4

Hoffmeyer, A., Piekorz, R., Moriggl, R., & Ihle, J. N. (2001). Gadd45gamma is dispensable for normal mouse development and T-cell proliferation. *Molecular and Cellular Biology*, *21*(9), 3137–3143. https://doi.org/10.1128/MCB.21.9.3137-3143.2001

Hopwood, N. (2000). Producing development: The anatomy of human embryos and the norms of Wilhelm His. *Bulletin of the History of Medicine*, *74*(1), 29–79. https://doi.org/10.1353/BHM.2000.0020

Hu, D., & Helms, J. A. (1999). The role of sonic hedgehog in normal and abnormal craniofacial morphogenesis. *Development (Cambridge, England)*, *126*(21), 4873–4884. https://doi.org/10.1242/DEV.126.21.4873

Hu, D., & Marcucio, R. S. (2009). Unique organization of the frontonasal ectodermal zone in birds and mammals. *Developmental Biology*, *325*(1), 200–210. https://doi.org/10.1016/J.YDBIO.2008.10.026

Hu, Q., Zhang, B., Chen, R., Fu, C., Jun, A., Fu, X., Li, J., Fu, L., Zhang, Z., & Dong, J. T. (2019). ZFHX3 is indispensable for ERβ to inhibit cell proliferation via MYC downregulation in prostate cancer cells. *Oncogenesis 2019 8:4*, *8*(4), 1–15. https://doi.org/10.1038/s41389-019-0138-y

Inubushi, T., Fujiwara, A., Hirose, T., Aoyama, G., Uchihashi, T., Yoshida, N., Shiraishi, Y., Usami, Y., Kurosaka, H., Toyosawa, S., Tanaka, S., Watabe, T., Kogo, M., & Yamashiro, T. (2022). Ras signaling and RREB1 are required for the dissociation of medial edge epithelial cells in murine palatogenesis. *DMM Disease Models and Mechanisms*, *15*(2). https://doi.org/10.1242/DMM.049093/273709/AM/RAS-SIGNALING-AND-ITS-EFFECTOR-RREB1-ARE-REQUIRED

Islam, S., Kjällquist, U., Moliner, A., Zajac, P., Fan, J. B., Lönnerberg, P., & Linnarsson, S. (2012). Highly multiplexed and strand-specific single-cell RNA 5' end sequencing. *Nature Protocols*, *7*(5), 813–828. https://doi.org/10.1038/NPROT.2012.022

Iwata, J. I., Hacia, J. G., Suzuki, A., Sanchez-Lara, P. A., Urata, M., & Chai, Y. (2012). Modulation of noncanonical TGF-β signaling prevents cleft palate in Tgfbr2 mutant mice. *The Journal of Clinical Investigation*, *122*(3), 873–885. https://doi.org/10.1172/JCI61498

Iwata, J. ichi, Suzuki, A., Pelikan, R. C., Ho, T. V., Sanchez-Lara, P. A., Urata, M., Dixon, M. J., & Chai, Y. (2013). Smad4-irf6 genetic interaction and TGFβ-mediated IRF6 signaling cascade are crucial for palatal fusion in mice. *Development (Cambridge)*, *140*(6), 1220–1230. https://doi.org/10.1242/DEV.089615/-/DC1

Jeong, J., Mao, J., Tenzen, T., Kottmann, A. H., & McMahon, A. P. (2004). Hedgehog signaling in the neural crest cells regulates the patterning and growth of facial primordia. *Genes & Development*, *18*(8), 937–951. https://doi.org/10.1101/GAD.1190304

Jernvall, J., Aberg, T., Kettunen, P., Keränen, S., & Thesleff, I. (1998). The life history of an embryonic signaling center: BMP-4 induces p21 and is associated with apoptosis in the mouse tooth enamel knot. *Development (Cambridge, England)*, *125*(2), 161–169. http://www.ncbi.nlm.nih.gov/pubmed/9486790

Jiang, R., Bush, J. O., & Lidral, A. C. (2006). Development of the upper lip: Morphogenetic and molecular mechanisms. In *Developmental Dynamics* (Vol. 235, Issue 5, pp. 1152–1166). John Wiley & Sons, Ltd. https://doi.org/10.1002/dvdy.20646

Jin, Y. R., Han, X. H., Taketo, M. M., & Yoon, J. K. (2012). Wnt9b-dependent FGF signaling is crucial for outgrowth of the nasal and maxillary processes during upper jaw and lip development. *Development*, *139*(10), 1821–1830. https://doi.org/10.1242/DEV.075796

Jung, C. G., Kim, H. J., Kawaguchi, M., Khanna, K. K., Hida, H., Asai, K., Nishino, H., & Miura, Y. (2005). Homeotic factor ATBF1 induces the cell cycle arrest associated with neuronal differentiation. *Development*, *132*(23), 5137–5145. https://doi.org/10.1242/DEV.02098

Jung, S. Y., Green, D. W., Jung, H. S., & Kim, E. J. (2018). Cell cycle of the enamel knot during tooth morphogenesis. *Histochemistry and Cell Biology*, *149*(6), 655–659. https://doi.org/10.1007/S00418-018-1666-9/FIGURES/2

Juriloff, D. M., Harris, M. J., Dewell, S. L., Brown, C. J., Mager, D. L., Gagnier, L., & Mah, D. G. (2005). Investigations of the genomic region that contains the clf1 mutation, a causal gene in multifactorial cleft lip and palate in mice. *Birth Defects Research. Part A, Clinical and Molecular Teratology*, *73*(2), 103–113. https://doi.org/10.1002/BDRA.20106

KIMMEL, C. B., MILLER, C. T., & KEYNES, R. J. (2001). Neural crest patterning and the evolution of the jaw. *Journal of Anatomy*, *199*(Pt 1-2), 105. https://doi.org/10.1046/J.1469-7580.2001.19910105.X

Klein, A. M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D. A., & Kirschner, M. W. (2015). Droplet Barcoding for Single-Cell Transcriptomics Applied to Embryonic Stem Cells. *Cell*, *161*(5), 1187–1201. https://doi.org/10.1016/J.CELL.2015.04.044

Klumpe, H. E., Langley, M. A., Linton, J. M., Su, C. J., Antebi, Y. E., & Elowitz, M. B. (2022). The context-dependent, combinatorial logic of BMP signaling. *Cell Systems*, *13*(5), 388-407.e10. https://doi.org/10.1016/J.CELS.2022.03.002

Kong, X., Zhai, J., Yan, C., Song, Y., Wang, J., Bai, X., Brown, J. A. L., & Fang, Y. (2019). Recent advances in understanding FOXN3 in breast cancer, and other malignancies. *Frontiers in Oncology*, *9*(MAY), 439326. https://doi.org/10.3389/FONC.2019.00234/BIBTEX

Kuki, K., Yamaguchi, N., Iwasawa, S., Takakura, Y., Aoyama, K., Yuki, R., Nakayama, Y., Kuga, T., Hashimoto, Y., Tomonaga, T., & Yamaguchi, N. (2017). Enhancement of TGF-β-induced Smad3 activity by c-Abl-mediated tyrosine phosphorylation of its coactivator SKI-interacting protein (SKIP). *Biochemical and Biophysical Research Communications*, *490*(3), 1045–1051. https://doi.org/10.1016/J.BBRC.2017.06.163

Kumari, R., & Jat, P. (2021). Mechanisms of Cellular Senescence: Cell Cycle Arrest and Senescence Associated Secretory Phenotype. *Frontiers in Cell and Developmental Biology*, *9*, 645593. https://doi.org/10.3389/FCELL.2021.645593/BIBTEX

Kurosaka, H., Iulianella, A., Williams, T., & Trainor, P. A. (2014). Disrupting hedgehog and WNT signaling interactions promotes cleft lip pathogenesis. *The Journal of Clinical Investigation*, *124*(4), 1660–1671. https://doi.org/10.1172/JCI72688

La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M. E., Lönnerberg, P., Furlan, A., Fan, J., Borm, L. E., Liu, Z., van Bruggen, D., Guo, J., He, X., Barker, R., Sundström, E., Castelo-Branco, G., … Kharchenko, P. V. (2018). RNA velocity of single cells. *Nature 2018 560:7719*, *560*(7719), 494–498. https://doi.org/10.1038/s41586-018-0414-6

Lee, S. H., Singh, I., Tisdale, S., Abdel-Wahab, O., Leslie, C. S., & Mayr, C. (2018). Widespread intronic polyadenylation inactivates tumour suppressor genes in leukaemia. *Nature 2018 561:7721*, *561*(7721), 127–131. https://doi.org/10.1038/s41586-018-0465-8

Lee, S. K., Sears, M. J., Zhang, Z., Li, H., Salhab, I., Krebs, P., Xing, Y., Nah, H. D., Williams, T., & Carstens, R. P. (2020). Cleft lip and cleft palate in Esrp1 knockout mice is associated with alterations in epithelial-mesenchymal crosstalk. *Development (Cambridge)*, *147*(21). https://doi.org/10.1242/DEV.187369/266886/AM/CLEFT-LIP-AND-CLEFT-PALATE-CL-P-IN-ESRP1-KO-MICE

Li, H., Jones, K. L., Hooper, J. E., & Williams, T. (2019). The molecular anatomy of mammalian upper lip and primary palate fusion at single cell resolution. *Development (Cambridge)*, *146*(12). https://doi.org/10.1242/DEV.174888/VIDEO-1

Li, M., Zhang, A., Zheng, Y., Li, J., & Zhao, J. (2020). ATBF1 Participates in Dual Functions of TGF-β via Regulation of Gene Expression and Protein Translocalization. *Biomolecules*, *10*(5). https://doi.org/10.3390/BIOM10050807

Li, W. Y., Dudas, M., & Kaartinen, V. (2008). Signaling through Tgf-beta type I receptor Alk5 is required for upper lip fusion. *Mechanisms of Development*, *125*(9–10), 874–882. https://doi.org/10.1016/J.MOD.2008.06.003

Li, Y. X., Yang, H. T., Zdanowicz, M., Sicklick, J. K., Qi, Y., Camp, T. J., & Diehl, A. M. (2007). Fetal alcohol exposure impairs hedgehog cholesterol modification and signaling. *Laboratory Investigation*, *87*(3), 231–240. https://doi.org/10.1038/LABINVEST.3700516

Li, Y., Zhao, H., Huang, X., Tang, J., Zhang, S., Li, Y., Liu, X., He, L., Ju, Z., Lui, K. O., & Zhou, B. (2018). Embryonic senescent cells re-enter cell cycle and contribute to tissues after birth. *Cell Research*, *28*(7), 775. https://doi.org/10.1038/S41422-018-0050-6

Liu, W., Sun, X., Braut, A., Mishina, Y., Behringer, R. R., Mina, M., & Martin, J. F. (2005). Distinct functions for Bmp signaling in lip and palate fusion in mice. *Development (Cambridge, England)*, *132*(6), 1453–1461. https://doi.org/10.1242/DEV.01676

Lorot-Marchand, A., Guerreschi, P., Pellerin, P., Martinot, V., Gbaguidi, C. C., Neiva, C., Devauchelle, B., Frochisse, C., Poli-Merol, M. L., & Francois-Fiquet, C. (2015). Frequency and socio-psychological impact of taunting in school-age patients with cleft lip-palate surgical repair. *International Journal of Pediatric Otorhinolaryngology*, *79*(7), 1041–1048. https://doi.org/10.1016/J.IJPORL.2015.04.024

Losa, M., Barozzi, I., Osterwalder, M., Hermosilla-Aguayo, V., Morabito, A., Chacón, B. H., Zarrineh, P., Girdziusaite, A., Benazet, J. D., Zhu, J., Mackem, S., Capellini, T. D., Dickel, D., Bobola, N., Zuniga, A., Visel, A., Zeller, R., & Selleri, L. (2023). A spatio-temporally constrained gene regulatory network directed by PBX1/2 acquires limb patterning specificity via HAND2. *Nature Communications 2023 14:1*, *14*(1), 1–20. https://doi.org/10.1038/s41467-023-39443-z

Losa, M., Risolino, M., Li, B., Hart, J., Quintana, L., Grishina, I., Yang, H., Choi, I. F., Lewicki, P., Khan, S., Aho, R., Feenstra, J., Vincent, C. T., Brown, A. M. C., Ferretti, E., Williams, T.,

& Selleri, L. (2018). Face morphogenesis is promoted by Pbx-dependent EMT via regulation of snail1 during frontonasal prominence fusion. *Development (Cambridge)*, *145*(5). https://doi.org/10.1242/dev.157628

Lu, Y., Liang, M., Zhang, Q., Liu, Z., Song, Y., Lai, L., & Li, Z. (2019). Mutations of GADD45G in rabbits cause cleft lip by the disorder of proliferation, apoptosis and epithelial-mesenchymal transition (EMT). *Biochimica et Biophysica Acta. Molecular Basis of Disease*, *1865*(9), 2356–2367. https://doi.org/10.1016/J.BBADIS.2019.05.015

MacKenzie, B. A., Wolff, R., Lowe, N., Billington, C. J., Peterson, A., Schmidt, B., Graf, D., Mina, M., Gopalakrishnan, R., & Petryk, A. (2009). Twisted gastrulation limits apoptosis in the distal region of the mandibular arch in mice. *Developmental Biology*, *328*(1), 13–23. https://doi.org/10.1016/j.ydbio.2008.12.041

Mai, C. T., Isenburg, J. L., Canfield, M. A., Meyer, R. E., Correa, A., Alverson, C. J., Lupo, P. J., Riehle-Colarusso, T., Cho, S. J., Aggarwal, D., & Kirby, R. S. (2019). National population-based estimates for major birth defects, 2010–2014. *Birth Defects Research*, *111*(18), 1420–1435. https://doi.org/10.1002/BDR2.1589

Maili, L., Letra, A., Silva, R., Buchanan, E. P., Mulliken, J. B., Greives, M. R., Teichgraeber, J. F., Blackwell, S. J., Ummer, R., Weber, R., Chiquet, B., Blanton, S. H., & Hecht, J. T. (2020). PBX-WNT-P63-IRF6 pathway in nonsyndromic cleft lip and palate. *Birth Defects Research*, *112*(3), 234–244. https://doi.org/10.1002/BDR2.1630

McInnes, L., Healy, J., & Melville, J. (2018). *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction*. https://arxiv.org/abs/1802.03426v3

Miller, A. J., & Cole, S. E. (2014). Multiple Dlk1 splice variants are expressed during early mouse embryogenesis. *The International Journal of Developmental Biology*, *58*(1), 65–70. https://doi.org/10.1387/IJDB.130316SC

Millicovsky, G., & Johnston, M. C. (1981). Hyperoxia and hypoxia in pregnancy: simple experimental manipulation alters the incidence of cleft lip and palate in CL/Fr mice. *Proceedings of the National Academy of Sciences of the United States of America*, *78*(9), 5722–5723. https://doi.org/10.1073/PNAS.78.9.5722

Mishina, Y., Suzuki, A., Ueno, N., & Behringer, R. R. (1995). Bmpr encodes a type I bone morphogenetic protein receptor that is essential for gastrulation during mouse embryogenesis. *Genes & Development*, *9*(24), 3027–3037. https://doi.org/10.1101/GAD.9.24.3027

Miura, Y., Kataoka, H., Joh, T., Tada, T., Asai, K., Nakanishi, M., Okada, N., & Okada, H. (2004). Susceptibility to Killer T Cells of Gastric Cancer Cells Enhanced by Mitomycin-C Involves Induction of ATBF1 and Activation of p21 (Waf1/Cip1) Promoter. *Microbiology and Immunology*, *48*(2), 137–145. https://doi.org/10.1111/J.1348-0421.2004.TB03491.X

Moens, C. B., & Selleri, L. (2006). Hox cofactors in vertebrate development. *Developmental Biology*, *291*(2), 193–206. https://doi.org/10.1016/J.YDBIO.2005.10.032

Mogollón, I., Moustakas-Verho, J. E., Niittykoski, M., & Ahtiainen, L. (2021). The initiation knot is a signaling center required for molar tooth development. *Development (Cambridge)*, *148*(9). https://doi.org/10.1242/DEV.194597/VIDEO-2

Mort, R. L., Ford, M. J., Sakaue-Sawano, A., Lindstrom, N. O., Casadio, A., Douglas, A. T., Keighren, M. A., Hohenstein, P., Miyawaki, A., & Jackson, I. J. (2014). Fucci2a: a bicistronic cell cycle reporter that allows Cre mediated tissue specific expression in mice. *Cell Cycle (Georgetown, Tex.)*, *13*(17), 2681–2696. https://doi.org/10.4161/15384101.2015.945381

Mossey, P. A., & Modell, B. (2012). Epidemiology of Oral Clefts 2012: An International Perspective. *Cleft Lip and Palate: Epidemiology, Aetiology and Treatment*, *16*, 1–18. https://doi.org/10.1159/000337464

Mukhopadhyay, N., Bishop, M., Mortillo, M., Chopra, P., Hetmanski, J. B., Taub, M. A., Moreno, L. M., Valencia-Ramirez, L. C., Restrepo, C., Wehby, G. L., Hecht, J. T., Deleyiannis, F., Butali, A., Weinberg, S. M., Beaty, T. H., Murray, J. C., Leslie, E. J., Feingold, E., & Marazita, M. L. (2020). Whole genome sequencing of orofacial cleft trios from the Gabriella Miller Kids First Pediatric Research Consortium identifies a new locus on chromosome 21. *Human Genetics*, *139*(2), 215–226. https://doi.org/10.1007/S00439-019-02099-1/FIGURES/3

Murthy, J., & Bhaskar, L. V. K. S. (2009). Current concepts in genetics of nonsyndromic clefts. *Indian Journal of Plastic Surgery : Official Publication of the Association of Plastic Surgeons of India*, *42*(1), 68–81. https://doi.org/10.4103/0970-0358.53004

Neubüser, A., Peters, H., Balling, R., & Martin, G. R. (1997). Antagonistic interactions between FGF and BMP signaling pathways: a mechanism for positioning the sites of tooth formation. *Cell*, *90*(2), 247–255. https://doi.org/10.1016/S0092-8674(00)80333-5

Nguyen, T. T., Mitchell, J. M., Kiel, M. D., Kenny, C. P., Li, H., Jones, K. L., Cornell, R. A., Williams, T. J., Nichols, J. T., & Van Otterloo, E. (2024). TFAP2 paralogs regulate midfacial development in part through a conserved ALX genetic pathway. *Development (Cambridge)*, *151*(1). https://doi.org/10.1242/DEV.202095/336058/AM/TFAP2-PARALOGS-REGULATE-MIDFACIAL-DEVELOPMENT-IN

Nolan, P. M., Banks, G., Bourbia, N., Wilcox, A. G., Bentley, L., Moir, L., Kent, L., Hillier, R., Wilson, D., Barrett, P., & Dumbell, R. (2023). A missense mutation in zinc finger homeobox-3 (ZFHX3) impedes growth and alters metabolism and hypothalamic gene expression in mice. *The FASEB Journal*, *37*(10), e23189. https://doi.org/10.1096/FJ.202201829R

Nottoli, T., Hagopian-Donaldson, S., Zhang, J., Perkins, A., & Williams, T. (1998). AP-2-null cells disrupt morphogenesis of the eye, face, and limbs in chimeric mice. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(23), 13714. https://doi.org/10.1073/PNAS.95.23.13714

O'Farrell, P. H. (2011). Quiescence: early evolutionary origins and universality do not imply uniformity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1584), 3498. https://doi.org/10.1098/RSTB.2011.0079

Osterburg, C., Ferniani, M., Antonini, D., Frombach, A. S., D'Auria, L., Osterburg, S., Lotz, R., Löhr, F., Kehrloesser, S., Zhou, H., Missero, C., & Dötsch, V. (2023). Disease-related p63 DBD mutations impair DNA binding by distinct mechanisms and varying degree. *Cell Death & Disease 2023 14:4*, *14*(4), 1–18. https://doi.org/10.1038/s41419-023-05796-y

Panaccione, A., Zhang, Y., Mi, Y., Yarbrough, W. G., Ivanov, S. V., Mitani, Y., El-Naggar, A. K., Yan, G., Prasad, M. L., McDonald, W. H., McDonald, W. H., Yarbrough, W. G., & Yarbrough, W. G. (2017). Chromosomal abnormalities and molecular landscape of metastasizing mucinous salivary adenocarcinoma. *Oral Oncology*, *66*, 38. https://doi.org/10.1016/J.ORALONCOLOGY.2016.12.011

Rajderkar, S. S., Paraiso, K., Amaral, M. L., Kosicki, M., Cook, L. E., Darbellay, F., Spurrell, C. H., Osterwalder, M., Zhu, Y., Wu, H., Afzal, S. Y., Blow, M. J., Kelman, G., Barozzi, I., Fukuda-Yuzawa, Y., Akiyama, J. A., Afzal, V., Tran, S., Plajzer-Frick, I., … Visel, A. (2023). Cell Type- and Tissue-specific Enhancers in Craniofacial Development. *BioRxiv*. https://doi.org/10.1101/2023.06.26.546603

Reed, S. C., & Snell, G. D. (1931). Harelip, a new mutation in the house mouse. *The Anatomical Record*, *51*(1), 43–50. https://doi.org/10.1002/AR.1090510104

Reid, B. S., Yang, H., Melvin, V. S., Taketo, M. M., & Williams, T. (2011). Ectodermal Wnt/β-catenin signaling shapes the mouse face. *Developmental Biology*, *349*(2), 261–269. https://doi.org/10.1016/J.YDBIO.2010.11.012

Richardson, M. K., & Keuck, G. (2022). The revolutionary developmental biology of Wilhelm His, Sr. *Biological Reviews*, *97*(3), 1131–1160. https://doi.org/10.1111/BRV.12834

Robert, T., Vanoli, F., Chiolo, I., Shubassi, G., Bernstein, K. A., Rothstein, R., Botrugno, O. A., Parazzoli, D., Oldani, A., Minucci, S., & Foiani, M. (2011). HDACs link the DNA damage response, processing of double-strand breaks and autophagy. *Nature 2011 471:7336*, *471*(7336), 74–79. https://doi.org/10.1038/nature09803

Rochlin, D. H., Ma, L. W., Sheckter, C. C., & Lorenz, H. P. (2022). Out-of-Pocket Costs and Provider Payments in Cleft Lip and Palate Repair. *Annals of Plastic Surgery*, *88*(4 Suppl 4), S343–S347. https://doi.org/10.1097/SAP.0000000000003081

Rogers, J. M., Waters, C. T., Seegar, T. C. M., Jarrett, S. M., Hallworth, A. N., Blacklow, S. C., & Bulyk, M. L. (2019). Bispecific Forkhead Transcription Factor FoxN3 Recognizes Two Distinct Motifs with Different DNA Shapes. *Molecular Cell*, *74*(2), 245-253.e6. https://doi.org/10.1016/J.MOLCEL.2019.01.019

Sabo, J. K., Heine, V., Silbereis, J. C., Schirmer, L., Levison, S. W., & Rowitch, D. H. (2017). Olig1 is required for noggin-induced neonatal myelin repair. *Annals of Neurology*, *81*(4), 560–571. https://doi.org/10.1002/ANA.24907

Sasaki, Y., Taya, Y., Saito, K., Fujita, K., Aoba, T., & Fujiwara, T. (2014a). Molecular contribution to cleft palate production in cleft lip mice. *Congenital Anomalies*, *54*(2), 94–99. https://doi.org/10.1111/CGA.12038

Sasaki, Y., Taya, Y., Saito, K., Fujita, K., Aoba, T., & Fujiwara, T. (2014b). Molecular contribution to cleft palate production in cleft lip mice. *Congenital Anomalies*, *54*(2), 94–99. https://doi.org/10.1111/CGA.12038

Satterstrom, F. K., Kosmicki, J. A., Wang, J., Breen, M. S., De Rubeis, S., An, J. Y., Peng, M., Collins, R., Grove, J., Klei, L., Stevens, C., Reichert, J., Mulhern, M. S., Artomov, M., Gerges, S., Sheppard, B., Xu, X., Bhaduri, A., Norman, U., … Buxbaum, J. D. (2020). Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell*, *180*(3), 568-584.e23. https://doi.org/10.1016/J.CELL.2019.12.036

Selleri, L., & Rijli, F. M. (2023). Shaping faces: genetic and epigenetic control of craniofacial morphogenesis. *Nature Reviews Genetics 2023 24:9*, *24*(9), 610–626. https://doi.org/10.1038/s41576-023-00594-w

Selleri, L., Zappavigna, V., & Ferretti, E. (2019). 'Building a perfect body': control of vertebrate organogenesis by PBX-dependent regulatory networks. *Genes & Development*, *33*(5–6), 258–275. https://doi.org/10.1101/GAD.318774.118

Shu, D. Y., Ng, K., Wishart, T. F. L., Chui, J., Lundmark, M., Flokis, M., & Lovicu, F. J. (2021). Contrasting roles for BMP-4 and ventromorphins (BMP agonists) in TGFβ-induced lens EMT. *Experimental Eye Research*, *206*. https://doi.org/10.1016/J.EXER.2021.108546

Siegenthaler, J. A., & Miller, M. W. (2005). Transforming Growth Factor β1 Promotes Cell Cycle Exit through the Cyclin-Dependent Kinase Inhibitor p21 in the Developing Cerebral Cortex. *The Journal of Neuroscience*, *25*(38), 8627. https://doi.org/10.1523/JNEUROSCI.1876-05.2005

Siewert, A., Reiz, B., Krug, C., Heggemann, J., Mangold, E., Dickten, H., & Ludwig, K. U. (2023). Analysis of candidate genes for cleft lip ± cleft palate using murine single-cell expression data. *Frontiers in Cell and Developmental Biology*, *11*, 1091666. https://doi.org/10.3389/FCELL.2023.1091666/FULL

Sock, E., Rettig, S. D., Enderich, J., Bösl, M. R., Tamm, E. R., & Wegner, M. (2004). Gene Targeting Reveals a Widespread Role for the High-Mobility-Group Transcription Factor Sox11 in Tissue Remodeling. *Molecular and Cellular Biology*, *24*(15), 6635–6644. https://doi.org/10.1128/MCB.24.15.6635-6644.2004

Soldatov, R., Kaucka, M., Kastriti, M. E., Petersen, J., Chontorotzea, T., Englmaier, L., Akkuratova, N., Yang, Y., Häring, M., Dyachuk, V., Bock, C., Farlik, M., Piacentino, M. L., Boismoreau, F., Hilscher, M. M., Yokota, C., Qian, X., Nilsson, M., Bronner, M. E., …

Adameyko, I. (2019). Spatiotemporal structure of cell fate decisions in murine neural crest. *Science (New York, N.Y.)*, *364*(6444). https://doi.org/10.1126/SCIENCE.AAS9536

Song, L., Li, Y., Wang, K., Wang, Y. Z., Molotkov, A., Gao, L., Zhao, T., Yamagami, T., Wang, Y., Gan, Q., Pleasure, D. E., & Zhou, C. J. (2009). Lrp6-mediated canonical Wnt signaling is required for lip formation and fusion. *Development*, *136*(18), 3161–3171. https://doi.org/10.1242/dev.037440

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B. B., Siddiqui, A., Lao, K., & Surani, M. A. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nature Methods*, *6*(5), 377–382. https://doi.org/10.1038/NMETH.1315

Teng, T., Teng, C. S., Kaartinen, V., & Bush, J. O. (2022). A unique form of collective epithelial migration is crucial for tissue fusion in the secondary palate and can overcome loss of epithelial apoptosis. *Development (Cambridge)*, *149*(10). https://doi.org/10.1242/DEV.200181/VIDEO-12

Thomason, H. A., Dixon, M. J., & Dixon, J. (2008). Facial clefting in Tp63 deficient mice results from altered Bmp4, Fgf8 and Shh signaling. *Developmental Biology*, *321*(1), 273–282. https://doi.org/10.1016/J.YDBIO.2008.06.030

Tirosh, I., Izar, B., Prakadan, S. M., Wadsworth, M. H., Treacy, D., Trombetta, J. J., Rotem, A., Rodman, C., Lian, C., Murphy, G., Fallahi-Sichani, M., Dutton-Regester, K., Lin, J. R., Cohen, O., Shah, P., Lu, D., Genshaft, A. S., Hughes, T. K., Ziegler, C. G. K., … Garraway, L. A. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science*, *352*(6282), 189–196. https://doi.org/10.1126/SCIENCE.AAD0501/SUPPL_FILE/TIROSH.SM.PDF

Trasanidis, N., Katsarou, A., Ponnusamy, K., Shen, Y. A., Kostopoulos, I. V., Bergonia, B., Keren, K., Reema, P., Xiao, X., Szydlo, R. M., Sabbattini, P. M. R., Roberts, I. A. G., Auner, H. W., Naresh, K. N., Chaidos, A., Wang, T. L., Magnani, L., Caputo, V. S., & Karadimitris, A. (2022). Systems medicine dissection of chr1q-amp reveals a novel PBX1-FOXM1 axis for targeted therapy in multiple myeloma. *Blood*, *139*(13), 1939–1953. https://doi.org/10.1182/BLOOD.2021014391

Truong, A. B., Kretz, M., Ridky, T. W., Kimmel, R., & Khavari, P. A. (2006). p63 regulates proliferation and differentiation of developmentally mature keratinocytes. *Genes & Development*, *20*(22), 3185–3197. https://doi.org/10.1101/GAD.1463206

Truong, B. T., & Artinger, K. B. (2021). The power of zebrafish models for understanding the co-occurrence of craniofacial and limb disorders. *Genesis*, *59*(1–2). https://doi.org/10.1002/DVG.23407

Villiard, É., Denis, J. F., Hashemi, F. S., Igelmann, S., Ferbeyre, G., & Roy, S. (2017). Senescence gives insights into the morphogenetic evolution of anamniotes. *Biology Open*, *6*(6), 891–896. https://doi.org/10.1242/BIO.025809/256768/AM/SENESCENCE-GIVES-INSIGHTS-INTO-THE-MORPHOGENETIC

Wang, C., Tu, H., Yang, L., Ma, C., Hu, J., Luo, J., & Wang, H. (2021). FOXN3 inhibits cell proliferation and invasion via modulating the AKT/MDM2/p53 axis in human glioma. *Aging*, *13*(17), 21587–21598. https://doi.org/10.18632/AGING.203499

Wang, K. Y., Chen, K. C., Chiang, C. P., & Kuo, M. Y. P. (1995). Distribution of p21ras during primary palate formation of non-cleft and cleft strains of mice. *Journal of Oral Pathology & Medicine : Official Publication of the International Association of Oral Pathologists and* *the American Academy of Oral Pathology*, *24*(3), 103–108. https://doi.org/10.1111/J.1600-0714.1995.TB01148.X

Watkins, S. E., Meyer, R. E., Strauss, R. P., & Aylsworth, A. S. (2014). Classification, epidemiology, and genetics of orofacial clefts. *Clinics in Plastic Surgery*, *41*(2), 149–163. https://doi.org/10.1016/J.CPS.2013.12.003

Welsh, I. C., Hagge-Greenberg, A., & O'Brien, T. P. (2007). A dosage-dependent role for Spry2 in growth and patterning during palate development. *Mechanisms of Development*, *124*(9–10), 746–761. https://doi.org/10.1016/J.MOD.2007.06.007

Welsh, I. C., Hart, J., Brown, J. M., Hansen, K., Rocha Marques, M., Aho, R. J., Grishina, I., Hurtado, R., Herzlinger, D., Ferretti, E., Garcia-Garcia, M. J., & Selleri, L. (2018). Pbx loss in cranial neural crest, unlike in epithelium, results in cleft palate only and a broader midface. *Journal of Anatomy*, *233*(2), 222–242. https://doi.org/10.1111/JOA.12821

Westfall, M. D., Mays, D. J., Sniezek, J. C., & Pietenpol, J. A. (2003). The ΔNp63α Phosphoprotein Binds the p21 and 14-3-3σ Promoters In Vivo and Has Transcriptional Repressor Activity That Is Reduced by Hay-Wells Syndrome-Derived Mutations. *Molecular and Cellular Biology*, *23*(7), 2264–2276. https://doi.org/10.1128/MCB.23.7.2264-2276.2003

Wong K. Chiminand Wu Lien-teh, W. L. (1932). *History of Chinese Medicine. Being a Chronicle of Medical Happenings in China from Ancient Times to the Present Period.* Tientsin Press, Ltd.

Wright, C. M., Garifallou, J. P., Schneider, S., Mentch, H. L., Kothakapa, D. R., Maguire, B. A., & Heuckeroth, R. O. (2020). Dlx1/2 mice have abnormal enteric nervous system function. *JCI Insight*, *5*(4). https://doi.org/10.1172/JCI.INSIGHT.131494

Wu, P., Jiang, T. X., Shen, J. Y., Widelitz, R. B., & Chuong, C. M. (2006). Morphoregulation of Avian Beaks: Comparative Mapping of Growth Zone Activities and Morphological Evolution. *Developmental Dynamics : An Official Publication of the American Association of Anatomists*, *235*(5), 1400. https://doi.org/10.1002/DVDY.20825

Wu, S., Wei, T., Fan, W., Wang, Y., Li, C., & Deng, J. (2021). Cell cycle during neuronal migration and neocortical lamination. *International Journal of Developmental Neuroscience*, *81*(2), 209–219. https://doi.org/10.1002/JDN.10091

Xu, J., Liu, H., Lan, Y., Adam, M., Clouthier, D. E., Potter, S., & Jiang, R. (2019). Hedgehog signaling patterns the oralaboral axis of the mandibular arch. *ELife*, *8*. https://doi.org/10.7554/ELIFE.40315

Yan, Y., Frisén, J., Lee, M. H., Massagué, J., & Barbacid, M. (1997). Ablation of the CDK inhibitor p57Kip2 results in increased apoptosis and delayed differentiation during mouse development. *Genes & Development*, *11*(8), 973–983. https://doi.org/10.1101/GAD.11.8.973

Yang, A., Schweitzer, R., Sun, D., Kaghad, M., Walker, N., Bronson, R. T., Tabin, C., Sharpe, A., Caput, D., Crum, C., & McKeon, F. (1999). p63 is essential for regenerative proliferation in limb, craniofacial and epithelial development. *Nature 1999 398:6729*, *398*(6729), 714–718. https://doi.org/10.1038/19539

Yu, W., Sun, Z., Sweat, Y., Sweat, M., Venugopalan, S. R., Eliason, S., Cao, H., Paine, M. L., & Amendt, B. A. (2020). Pitx2-Sox2-Lef1 interactions specify progenitor oral/dental epithelial cell signaling centers. *Development (Cambridge, England)*, *147*(11). https://doi.org/10.1242/DEV.186023

Zaw, K. T. T., Sato, N., Ikeda, S., Thu, K. S., Mieno, M. N., Arai, T., Mori, S., Furukawa, T., Sasano, T., Sawabe, M., Tanaka, M., & Muramatsu, M. (2017). Association of ZFHX3 gene variation with atrial fibrillation, cerebral infarction, and lung thromboembolism: An autopsy study. *Journal of Cardiology*, *70*(2), 180–184. https://doi.org/10.1016/J.JJCC.2016.11.005

Zhang, P., Liégeois, N. J., Wong, C., Finegold, M., Hou, H., Thompson, J. C., Silverman, A., Harper, J. W., DePinho, R. A., & Elledge, S. J. (1997). Altered cell differentiation and proliferation in mice lacking p57KIP2 indicates a role in Beckwith-Wiedemann syndrome. *Nature*, *387*(6629), 151–158. https://doi.org/10.1038/387151A0

Zhang, W., Hong, M., Bae, G. U., Kang, J. S., & Krauss, R. S. (2011). Boc modifies the holoprosencephaly spectrum of Cdo mutant mice. *DMM Disease Models and Mechanisms*, *4*(3), 368–380. https://doi.org/10.1242/DMM.005744/-/DC1

Zhang, Z., Wei, S., Du, H., Su, Z., Wen, Y., Shang, Z., Song, X., Xu, Z., You, Y., & Yang, Z. (2019). Zfhx3 is required for the differentiation of late born D1-type medium spiny neurons. *Experimental Neurology*, *322*. https://doi.org/10.1016/J.EXPNEUROL.2019.113055

Zhao, D., Ma, G., Zhang, X., He, Y., Li, M., Han, X., Fu, L., Dong, X. Y., Nagy, T., Zhao, Q., Fu, L., & Dong, J. T. (2016). Zinc Finger Homeodomain Factor Zfhx3 Is Essential for Mammary Lactogenic Differentiation by Maintaining Prolactin Signaling Activity. *The*

Journal of Biological Chemistry, *291*(24), 12809–12820. https://doi.org/10.1074/JBC.M116.719377

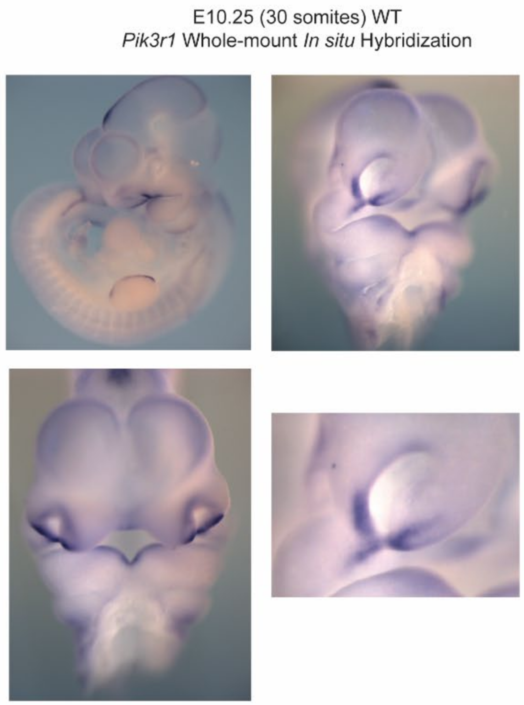
Zhao, H., Oka, K., Bringas, P., Kaartinen, V., & Chai, Y. (2008). TGF-beta type I receptor Alk5 regulates tooth initiation and mandible patterning in a type II receptor-independent manner. *Developmental Biology*, *320*(1), 19–29. https://doi.org/10.1016/J.YDBIO.2008.03.045

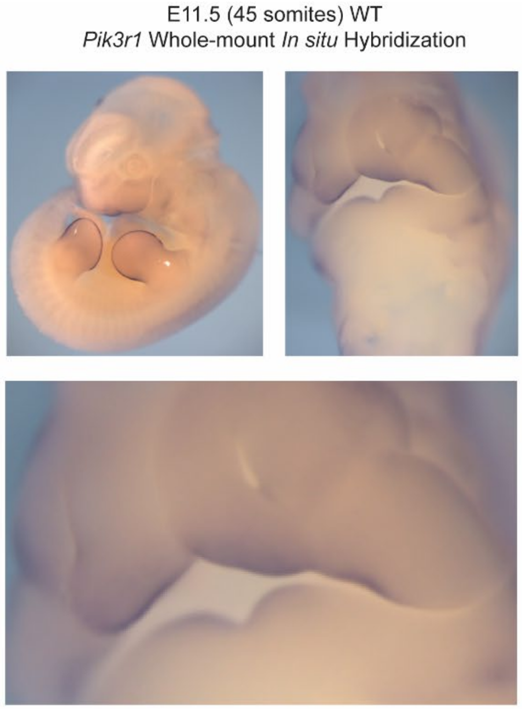
Zheng, G. X. Y., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., Ziraldo, S. B., Wheeler, T. D., McDermott, G. P., Zhu, J., Gregory, M. T., Shuga, J., Montesclaros, L., Underwood, J. G., Masquelier, D. A., Nishimura, S. Y., Schnall-Levin, M., Wyatt, P. W., Hindson, C. M., … Bielas, J. H. (2017). Massively parallel digital transcriptional profiling of single cells. *Nature Communications 2017 8:1*, *8*(1), 1–12. https://doi.org/10.1038/ncomms14049

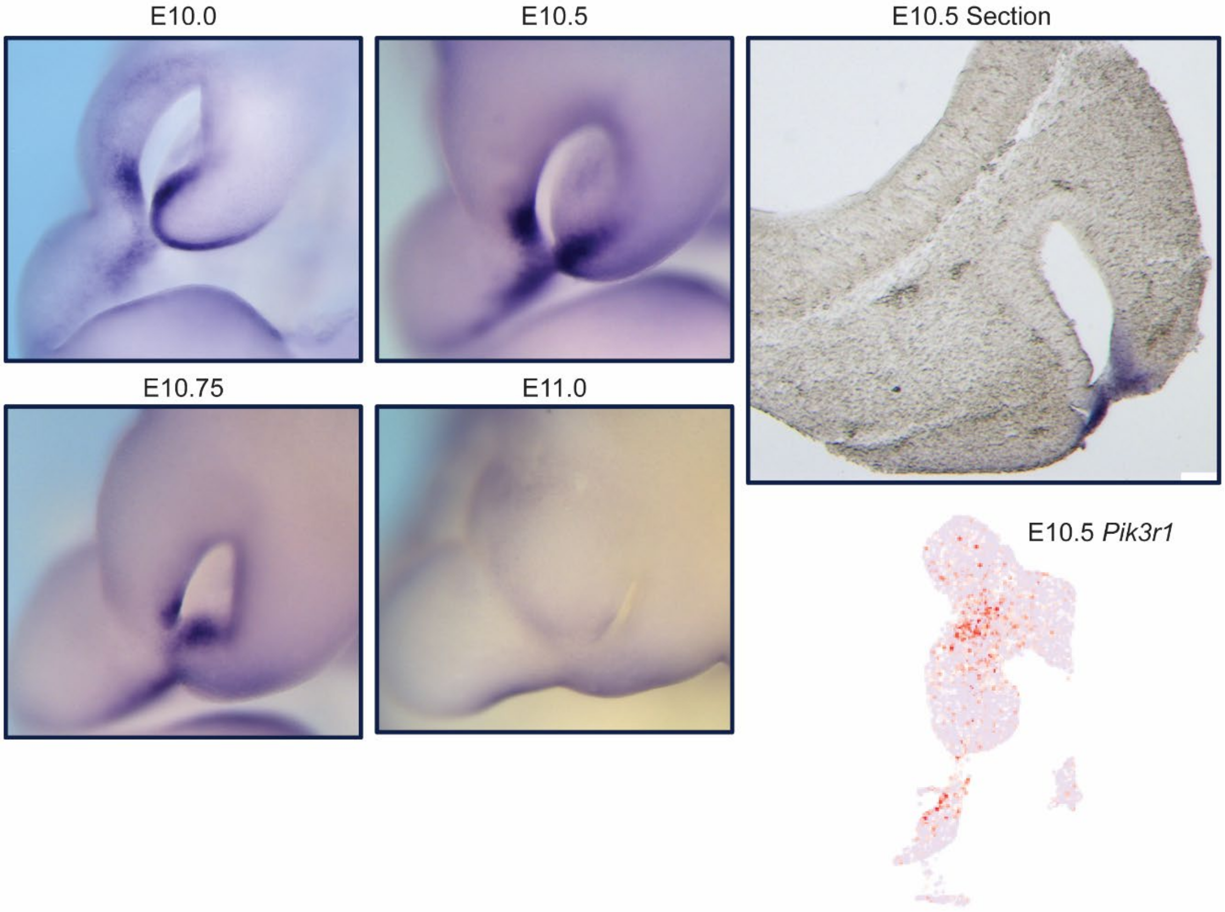
Zhou, H. M., & Conway, S. J. (2022). Pax3 Hypomorphs Reveal Hidden Pax7 Functional Genetic Compensation in Utero. *Journal of Developmental Biology*, *10*(2), 19. https://doi.org/10.3390/JDB10020019/S1

Zhu, X., Wei, L., Bai, Y., Wu, S., & Han, S. (2017). FoxC1 promotes epithelial-mesenchymal transition through PBX1 dependent transactivation of ZEB2 in esophageal cancer. *American Journal of Cancer Research*, *7*(8), 1642. /pmc/articles/PMC5574937/

APPENDIX 1


E10.25 (30 somites) WT
Pik3r1 Whole-mount *In situ* Hybridization


E11.5 (45 somites) WT
Pik3r1 Whole-mount *In situ* Hybridization


E10.0

E10.5

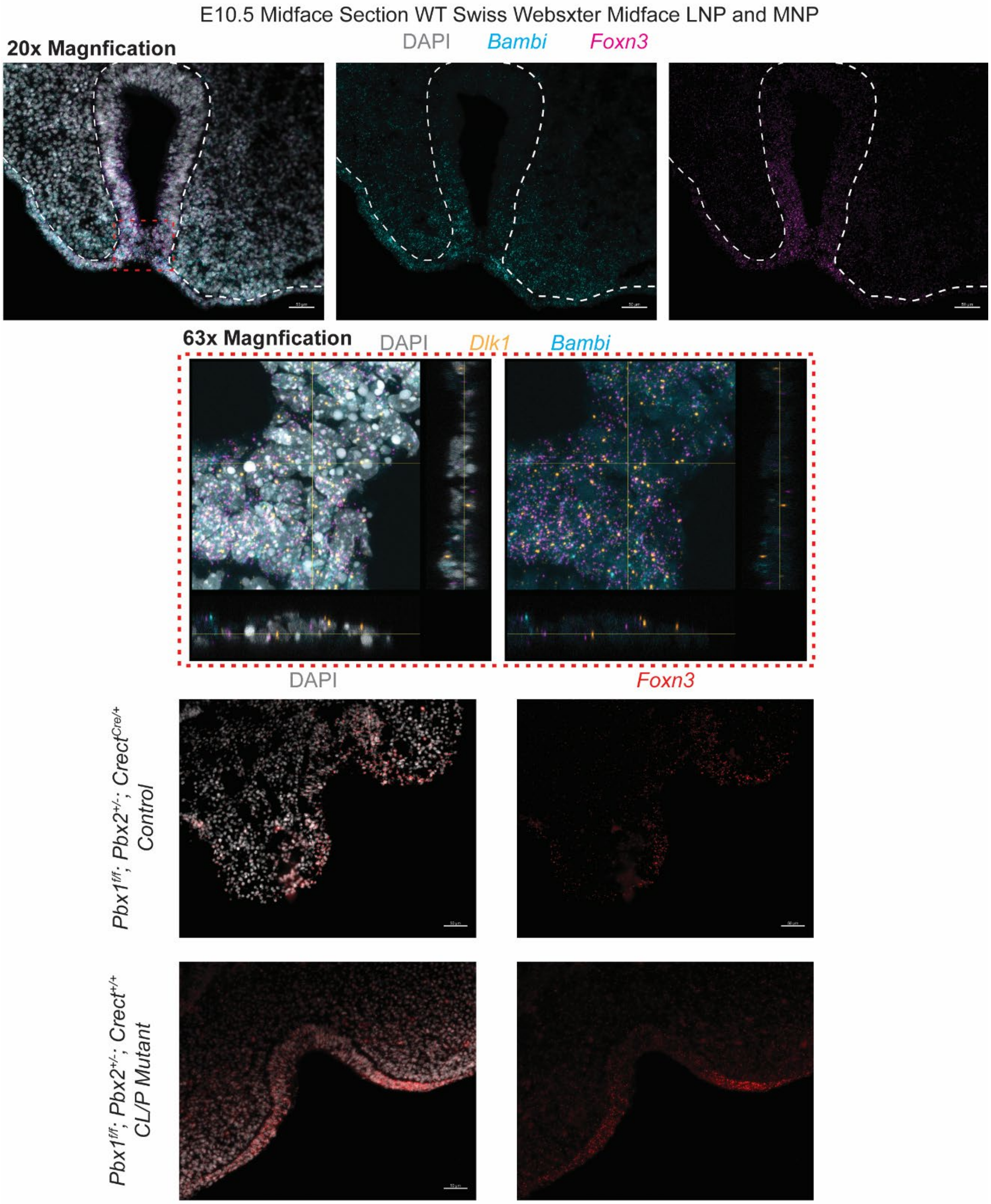
E10.5 Section

E10.75

E11.0

E10.5 *Pik3r1*

APPENDIX 2


E10.5 Midface Section WT Swiss Websxter Midface LNP and MNP
20x Magnfication
DAPI
Bambi
Foxn3
63x Magnfication
DAPI
Dlk1
Bambi
DAPI
Foxn3
Pbx1f/f; Pbx2+/-; CrectCre/+
Control
Pbx1f/f; Pbx2+/-; Crect+/+
CL/P Mutant

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

Author Signature

5/27/2024

Date