Title
Transcriptional Regulation of Tracheoesophageal Fate Specification in the Mammalian Foregut

Author
Kuwahara, Akela

Publication Date
2020

Peer reviewed|Thesis/dissertation

Transcriptional Regulation of Tracheoesophageal Fate Specification in the Mammalian Foregut

by
Akela Alderdice Kuwahara

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:
Nadav Ahituv
77BA96E7DAE34F4...

Nadav Ahituv

Chair

DocuSigned by:
Jeffrey Bush
[illegible]...

Jeffrey Bush

DocuSigned by:
Licia Selleri
[illegible]...

Licia Selleri

DocuSigned by:
Ophir Klein
F985C63D8CE3437...

Ophir Klein

Committee Members

Dedicated to all of my family,
who have shown me the many ways to live a full life.

Acknowledgements

I am honored to share this accomplishment with the many amazing people who have influenced me scientifically and personally throughout my life, all of whom have contributed to the realization of this thesis in one way or another. I will be forever grateful.

To Dr. Jeff Bush, my Ph.D. advisor, mentor, and friend, thank you for everything. Your scientific guidance has forever shaped the way I approach new questions, ideas, and challenges. I have learned so much from you experimentally and conceptually. Thank you for investing in me, and voicing your beliefs in my capabilities as a scientist. They are the highest compliment and have buoyed me countless times. Writing with you has been one of the most enjoyable parts of my PhD, thank you for including me in grant applications from the start and for giving me the foundation to write the manuscript for this project. People say you lead by example, and I could not agree more. I will continue to look to the example you have set of commitment, perseverance, and engagement in all that I do. Thank you.

To all of my other advisors and mentors at UCSF, thank you for your consistent support. My thesis committee chair, Dr. Nadav Ahituv, and my thesis committee members Dr. Licia Selleri and Dr. Ophir Klein – you have all provided me with invaluable scientific and professional advice and encouragement. Thank you for sharing your time and knowledge with me. To Dr. Dan Hart, thank you for giving me the freedom to explore new scientific interests and for nurturing my curiosity as a young graduate student. To Dr. Todd Nystul, Dr. Robert Blelloch, and Demian Sainz, thank you for making the DSCB program a fantastic place to learn.

I am lucky to have had many incredible mentors throughout my life, all of whom have shaped my goals and character, thank you to all them. I am especially thankful to Dr. Jacob Varkey of Humboldt State University who recognized in me potential that I didn't know I had and is one of the key reasons I began this PhD program. Thank you to Dr. Amy Sprowles, also of Humboldt State University, who patiently taught me the fundamentals of bench science that I have carried with me through many of the experiments described in this thesis and who will continue to be a role model of scientific, teaching, and personal excellence to me.

To my colleagues at UCSF, you have been amazing role models, friends, and supporters throughout this process, thank you. In particular, I owe immense gratitude to Ace Lewis, my scientific partner-in-crime, co-captain of Team Foregut and Akela and the Jets. Thank you for welcoming me into this project, and for teaching, supporting, and laughing with me along the way, and allowing me to do the same for you. To my fellow Bush Lab members, it has been an honor to work alongside such wonderful minds, I cannot wait to hear about the wonderful places your life takes you. To all of my DSCB peers, particularly my classmates Daniel, Buddy, Bonnie, and Ashley, I am so grateful to have gone through this process with you and have amazing memories of launching into grad school together. To fellow members of the Program in Craniofacial Biology, thank you for all of your scientific and personal support and for being my teammates, in and out of lab. To my mentor, dear friend, and collaborator Dr. Michelle Percharde, thank you for investing in my growth as a scientist, and for being there for me from the very start of my PhD. I am so grateful for your continued support and friendship.

To Mali'o, thank you. Your sisterhood has carried me across oceans. I think you have read every scholarship, internship, and graduate school application I've written since I was 18 years old. You've weathered every crisis and celebrated every victory. I am so honored to share this with you, my anemone.

To Travis, I cannot thank you enough for joining me in this endeavor. Your strength has helped to push me up and over this hill, and I am so grateful. Thank you for knowing when to give perspective, motivation, and empathy, and for believing in me always. You have taken on all of my joys, stress, burdens, and ambitions with an open heart and for that and so many other things, I will be forever grateful.

To my family, the Alderdice's and Kuwahara's, thank you for being examples of excellence, each in your own way. Because of you, I understand the many ways to love, learn, grow, and contribute to this world. To Nana and Tata, you will always inspire me, and I can only hope to live a life as full as yours.

Above all, thank you to my parents Ken and Lauren. You raised me in a world full of love and adventure and have been my biggest supporters in all that I do. Thank you for every opportunity you have provided for me or have encouraged me to pursue. Thank you for being engaged in my journey and invested in my future. All of this has been possible because you have selflessly given me the best parts of yourselves. I am so proud to be your daughter. Thank you for everything, I love you.

Contributions:

All of the work presented in this dissertation was performed under supervision of Dr. Jeffrey Bush in the Program in Craniofacial Biology and Department of Cell and Tissue Biology at the University of California, San Francisco. I performed all experimental procedures presented in here with direct assistance from Dr. Jeffrey Bush, Ace Lewis, Coohleen Coombes, Fang-Shiuan Leung and Dr. Michelle Percharde. Dr. Jeffrey Bush and Ace Lewis assisted with E10.5 and E11.5 embryonic dissections for single cell RNA-sequencing and ChIP-sequencing experiments. Coohleen Coombes performed E13.5 embryonic dissections, tissue preparations, RNAscope/immunofluorescent staining, and E18.5 skeletal preparations. Fang-Shiuan Leung assisted with cryosectioning and imaging. Dr. Michelle Percharde assisted with FACS purification of foregut cells. All sequencing data analysis presented here was performed by me and Dr. Michelle Percharde. ChIP-sequencing protocol and advice was provided by Dr. Marta Losa-Llabata and Dr. Licia Selleri's laboratory. Dr. Nadav Ahituv, Dr. Ophir Klein, and Dr. Licia Selleri provided valuable advice and direction throughout this project. This dissertation includes modifications of data and writing from a submitted manuscript titled "Delineating the early transcriptional specification of the mammalian trachea and esophagus."

Transcriptional Regulation of Tracheoesophageal Fate Specification in the Mammalian Foregut

Akela Alderdice Kuwahara

Abstract

Specification of the trachea and esophagus from the embryonic foregut is critical for the function of the respiratory and digestive systems. Congenital birth defects associated with improper tracheoesophageal development in humans are common and have severe consequences to respiration and feeding that require immediate surgical intervention. During development, tracheoesophageal specification is dependent on proper dorsoventral patterning of the foregut endoderm tube. This involves the establishment of ventral NKX2.1 and dorsal SOX2 expression domains that are thought to promote tracheal and esophageal fate, respectively. Loss of *Nkx2.1* results in failed foregut separation, and adoption of SOX2 expression throughout the common foregut tube. Similarly, loss of *Sox2* results in a common foregut tube expressing NKX2.1, leading to the previous conclusion that both of these transcription factors are master regulators of tracheal and esophageal fate, respectively. However, our understanding of tracheoesophageal fate specification is limited by a lack of information on dorsoventral foregut patterning at the transcriptome level. In this study, we use genome-wide methods to understand how tracheal and esophageal lineages are specified during mouse embryonic development. We use single cell RNA-sequencing to define transcriptomic profiles of early developing mouse trachea, lung, and esophagus, and discover a multitude of previously unknown markers of these tissues. Transcriptomic analysis of

Nkx2.1$^{-/-}$ mutant foreguts reveals that NKX2.1 loss does not result in lineage conversion to esophagus as previously hypothesized and exposes an NKX2.1-independent tracheoesophageal program. Using ChIP-seq against NKX2.1, we identify direct NKX2.1 regulatory targets and interrogate their combinatorial regulation by NKX2.1 and SOX2 in compound mouse mutant analysis. Amongst the novel targets we identify are *Shh* and *Wnt7b*, which we demonstrate are regulated by NKX2.1 to control tracheal and esophageal mesenchyme specification to cartilage and smooth muscle. Together, these data dramatically revise our understanding of how tracheal and esophageal cell types are specified during development and uncover a limited yet critical role for *Nkx2.1* in this process.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Overview and significance of foregut development

Proper development of the mammalian trachea and esophagus is critical for the function of the respiratory and digestive systems. The trachea serves as the tube that provides air flow to the lungs and is comprised of secretory, basal, and ciliated cells that moisten the airway and provide a barrier for pathogens and debris. The trachea is surrounded ventrally by cartilaginous rings that provide structure of the airway, and dorsally by the trachealis, a band of smooth muscle that connects the c-shaped cartilage rings and allows for flexibility and airway expansion/contraction (Reviewed in Brand-Saberi and Schäfer, 2014). Conversely, the esophagus serves as the tube that connects the oral cavity to the stomach and is comprised of squamous epithelial cells surrounded by striated and smooth muscle that support peristalsis (Reviewed in Kuo and Urma, 2006).

During development, the trachea and esophagus are specified from the ventral and dorsal domains of the foregut endoderm tube **(Fig. 1.1)**. This process is dependent on signals from the surrounding splanchnic mesoderm-derived mesenchyme that result in lung bud induction and dorsoventral specification, followed by physical separation of the trachea and esophagus (**Fig. 1.1a**, Billmyre et al., 2015; Cardoso and Lü, 2006; Morrisey et al., 2013; Rankin et al., 2016, 2017; Shannon et al., 1998). In mice, this specification and physical separation occurs between embryonic days 9.0-11.5 (E9.0-E11.5), corresponding to the third and fourth week of human development. At E9.5, primitive lung buds begin to evaginate from the common foregut tube. Over the course of the next 48 hours, the common foregut tube separates into trachea and esophagus, beginning at the point of lung bud evagination. By E18.5, the trachea consists of ciliated

and secretory cells within a pseudostratified columnar epithelium surrounded by cartilaginous rings, whereas the esophagus is a stratified squamous epithelium surrounded by smooth muscle.

Improper tracheoesophageal specification or separation can result in an unseparated common foregut tube and is associated with several congenital birth defects in humans including tracheoesophageal fistula (TEF), esophageal atresia (EA), and tracheal agenesis (TA). TEF is characterized by an improper connection between the trachea and esophagus and occurs in approximately 1 in every 3000 live births in humans (Brunner and Bokhoven, 2005; Que et al., 2006). TEF often occurs in conjunction with EA, which is characterized by a blind-ended esophagus that extends from the pharynx but does not connect to the stomach (Brunner and Bokhoven, 2005). Tracheal agenesis is characterized by partial or complete loss of the trachea, with the oral cavity connecting through a common tube to the lungs and stomach (Sher and Liu, 2016; van Veenendaal et al., 2000). These foregut malformations result in an inability to eat or breathe properly and, when detected early, can be repaired surgically. Currently, genetic or environmental causes of this defect are poorly understood, in part due to our limited knowledge of genes involved in normal tracheoesophageal development. Therefore, an improved understanding of mammalian tracheoesophageal specification and separation is essential for addressing the biology of foregut malformations in humans.

The developing foregut provides an exciting system in which to study general mechanisms of fate specification and morphogenesis during embryonic development. The relatively simple system of a single common tube becoming dorsoventrally specified to give rise the tracheal and esophageal tubes allows us to dissect how fate decisions are

made within a tissue, and the consequences that these decisions have on tissue morphogenesis and function. Additionally, fate specification in the foregut is tightly temporally coupled to morphogenetic process of foregut separation, and the majority of genes that are known to drive specification also influence foregut separation. In studying foregut development, we can address fundamental questions about how cell fate boundaries are formed at the transcriptional level, the cell biological level, and the tissue level during embryonic development.

Establishment of dorsoventral identity in the common foregut

While our understanding of tracheoesophageal epithelial identity remains limited, NKX2.1 and SOX2 are two known transcription factors that delineate the dorsoventral axis in the developing foregut. NKX2.1 is the earliest known marker of the respiratory lineage and is detectable in the ventral foregut endoderm and evaginating lung bud as early as E9.0 (Guazzi et al., 1990; Mizuno et al., 1991). NKX2.1 expression remains restricted to the ventral and future lung and tracheal domains throughout foregut separation. In contrast, the SOX2 transcription factor is expressed uniformly throughout the dorsoventral extent of the foregut endoderm epithelial tube until the onset of NKX2.1 expression. At this time, SOX2 expression is reduced in the ventral relative to dorsal endoderm, a difference that is still observed following the separation of the trachea and esophagus at E11.5 (Que et al., 2007a). The opposing expression patterns of NKX2.1 and SOX2 are believed to be maintained through a mutually repressive relationship between these two transcription factors, although a direct regulatory relationship in the foregut has not been established.

The dorsoventral expression patterns of SOX2 and NKX2.1 are established by multiple signaling pathways during embryonic development. The developing foregut is surrounded by splanchnic mesoderm-derived mesenchyme and located between the notochord and the developing cardiac mesenchyme, all of which provide critical signals that instruct foregut specification. Establishment of ventral foregut identity requires BMP and WNT signaling from the surrounding mesenchyme. These signals are induced by a retinoic acid, sonic hedgehog signaling cascade in early foregut development that induces *Bmp4* and *Wnt2/2b* expression in the ventral mesenchyme (Rankin et al., 2016, 2017). Both BMP and WNT signaling pathways are required for the expression of the respiratory-specific transcription factor NKX2.1. *Bmp4* is expressed in the ventral mesenchyme and signals to the endoderm via its receptors BMPR1A and BMPR1B (Domyan et al., 2011; Li et al., 2008; Rankin et al., 2017; Stevens et al., 2017). Loss of *Bmp4* in the ventral mesenchyme leads to a failure of foregut separation, maintenance of respiratory identity as indicated by NKX2.1 expression (Li et al., 2008). Additionally, loss of *Bmpr1a; Bmpr1b* leads to a lack of NKX2.1 expression in the ventral foregut with an expansion of SOX2, P63, and mesenchymal smooth muscle (Domyan et al., 2011). *Nkx2.1* expression is rescued by ventral deletion of *Sox2,* indicating that BMP signaling acts to repress ventral expression of *Sox2*, and create a permissive environment for activation of *Nkx2.1* expression (Domyan et al., 2011; Rankin et al., 2017; Stevens et al., 2017).

Notochord expression of *Noggin (Nog)*, a BMP-antagonist, is essential for initial foregut delamination from the notochord, and subsequent foregut patterning (Fausett et al., 2014; Li et al., 2007; Que et al., 2006). Mice lacking *Nog* exhibit defects in notochord and foregut morphogenesis, with evidence of improper separation of the two tissues. This

phenotype is rescued by attenuating the expression of *Bmp4*, indicating that proper levels of BMP activity are essential for foregut development. This also suggests that the relative position of the foregut along the dorsal-ventral axis of the embryo, and therefore relative to mesodermal signaling cues, is important for proper signal reception and fate specification.

Establishment of *Nkx2.1* expression in the ventral foregut epithelium additionally requires *Wnt2/2b* expression in the ventral mesenchyme and *Ctnnb1*-dependent induction of *Nkx2.1* (Goss et al., 2009; Harris-Johnson et al., 2009; Ostrin et al., 2018; Stevens et al., 2017). Loss of *Wnt2;2b,* results in an unseparated foregut tube that lacks expression of *Nkx2.1* as well as *Wnt7b*, another marker of the respiratory epithelium (Goss et al., 2009). Inactivation of *Ctnnb1* (*beta-Catenin*) in the ventral epithelium also results in a lack of respiratory specification as indicated by *Nkx2.1* expression and failure of foregut separation (Harris-Johnson et al., 2009). Together, these data have led to a model wherein respiratory fate is induced through combined repression of *Sox2* by BMP signaling and activation of *Nkx2.1* by WNT signaling. These pathways are likely to be conserved in humans because generation of *Nkx2.1*-expressing respiratory progenitor cells from human embryonic stem cells (hESCs) requires WNT and BMP signaling (Dye et al., 2015; Ostrin et al., 2018; Trisno et al., 2018). At the regulatory level, the *Nkx2.1* promoter can be activated by FOXA1, SP1, SP3, GATA6, and HOXB3 in various cell lines (Boggaram, 2009), however whether these transcription factors help to initiate NKX2.1 expression in the foregut is unknown. Nonetheless, induction of NKX2.1 expression and respiratory fate within the SOX2-expressing foregut endoderm is dependent on signaling from the surrounding mesenchyme. The resulting opposing NKX2.1 and SOX2

expression patterns are then thought to be maintained through mutual repression between these two transcription factors.

NKX2.1 and SOX2 as regulators of tracheoesophageal fate

In addition to being a robust marker of the respiratory lineage, *Nkx2.1* has important functional roles in the development of the trachea and esophagus. In mice lacking *Nkx2.1*, the foregut tube fails to separate to trachea and esophagus and, by E11.5, rudimentary lung buds that exhibit branching and specification defects protrude from the common foregut tube (Kimura et al., 1996; Minoo et al., 1999a; Yuan et al., 2000). High expression of SOX2 as well as another marker of esophageal cells at this stage, P63, is observed throughout the *Nkx2.1* mutant common foregut tube (Que et al., 2007a). Specification of the foregut mesenchyme to tracheal cartilage, trachealis smooth muscle, and esophageal smooth muscle is critical for maintaining the structure and function of the trachea and esophagus and is also perturbed in *Nkx2.1*$^{-/-}$ mutants. Whereas the foregut mesenchyme of wild-type embryos develops into cartilaginous rings around the ventral side of the trachea and smooth muscle around the esophagus, the common foregut tube of *Nkx2.1*$^{-/-}$ mutants lacks organized tracheal cartilage and is mostly surrounded by smooth muscle (Minoo et al., 1999a; Que et al., 2007a). These phenotypic data most closely resemble the human congenital disease tracheal agenesis, where a single foregut tube leads to the lungs and stomach. In mice, the studies described above have led to the hypothesis that NKX2.1 is a master regulator of tracheoesophageal fate, and that its loss results in a trachea-to-esophageal fate conversion.

This hypothesis is consistent with studies in lung adenocarcinoma and in the distal lung (Little et al., 2019; Snyder et al., 2013). In lung adenocarcinoma, low expression of NKX2.1 is associated with an increase in gastric markers, indicating a transition from a more respiratory-like cell to a gastric-like cell (Snyder et al., 2013). In lung alveolar type 1 (AT1) cells, loss of *Nkx2.1* results in an increase in expression of gastrointestinal genes and adoption of gastric morphology (Little et al., 2019). These studies suggest that NKX2.1 is important in directing and maintaining cells in a respiratory fate instead of a gastric fate. However, there has been little focus on the programs downstream of NKX2.1 during initial tracheal specification and we lack an understanding of how NKX2.1 regulates respiratory fate at the genome-wide scale.

Several transcriptional targets of NKX2.1 have been identified, along with several NKX2.1 regulatory partners in the lung. However, few have been explored in the context of respiratory development or within the trachea. NKX2.1 has been shown to interact with FOXA1 in lung epithelial or adenocarcinoma cells to regulate respiratory genes *SpC* and *Ccsp, Lmo3* (Minoo et al., 2007; Watanabe et al., 2013). NKX2.1 can also interact *in vitro* with GATA4/6 to induce *SpC* expression at the *SpC* promoter (Liu et al., 2002). NKX2.1 binding and transcriptional regulation *in vitro* has also been demonstrated in the respiratory genes *Cldn18*, *Abca3, Ugrp1*, and *Wnt7b,* yet whether these genes are regulated by NKX2.1 during respiratory development *in vitro* is unknown (Besnard et al., 2007; Niimi et al., 2001a, 2001b; Weidenfeld et al., 2002). While these studies have informed lung cell type specification, lung homeostasis, and lung adenocarcinoma biology, very little effort has been directed at understanding NKX2.1 regulation of early respiratory identity. This is a critical step in respiratory development, and a genome-wide

understanding of the role of NKX2.1 in tracheal and lung specification will reveal new insights on respiratory development and disease.

The transcription factor SOX2 is expressed in an opposing, yet less restricted pattern than NKX2.1 during foregut development and is also critical for tracheoesophageal development. Hypomorphic reduction of *Sox2* expression in mice results in an unseparated foregut tube with an increase in dorsal NKX2.1 expression and several differentiated respiratory cell markers (eg. SCGB1A1, alcian blue staining of mucus-secreting cells) and reduction of dorsal P63 expression, consistent with the conversion of the stratified esophageal epithelium to a simple columnar epithelium that more resembles the trachea as documented by histology (Que et al., 2007a). Additionally, loss of *Sox2* throughout the foregut endoderm results in reduced smooth muscle in the surrounding mesenchyme and an expansion of the tracheal mesenchymal markers *Sox9* and *Tbx4* (Teramoto et al., 2019). Furthermore, knockdown of SOX2 in human induced pluripotent stem cell (hiPSC)-derived dorsal foregut cells results in an upregulation of NKX2.1, and forced expression of SOX2 in hiPSC-derived ventral foregut cells represses NKX2.1 (Trisno et al., 2018). These data support a model in which NKX2.1 and SOX2 form a co-repressive master regulatory switch to define tracheal and esophageal cell types.

The extent to which NKX2.1 directly induces respiratory genes and represses esophageal genes, or acts indirectly through repression of *Sox2* is unknown, and is difficult to assay at the whole-genome scale with our current limited understanding of initial tracheoesophageal patterning. Recent studies have applied single cell-transcriptomic approaches to understand early definitive endoderm and splanchnic

mesenchyme differentiation (Han et al., 2019; Nowotschin et al., 2019), but have not focused on elucidating trachea and esophagus endoderm-specific transcriptomic profiles. Further, though significant work has focused on identifying direct targets of NKX2.1 regulation in other respiratory contexts including lung adenocarcinomas (Sawaya et al., 1993; Snyder et al., 2013; Tagne et al., 2012a; Watanabe et al., 2013; Yamaguchi et al., 2012), lung epithelial cell culture (Bohinski et al., 1994; Bruno et al., 1995; Kelly et al., 1996; Liu et al., 2002; Oguchi and Kimura, 1998; Zhu et al., 2004), and embryonic distal lungs (Little et al., 2019; Tagne et al., 2012a), direct targets of NKX2.1 and SOX2 during early tracheoesophageal specification are currently unknown.

Specification of foregut mesenchymal cell fates

Mesenchymal specification to tracheal cartilage, trachealis smooth muscle, and esophageal smooth muscle during foregut development is also critical for proper function of the trachea and esophagus. C-shaped tracheal cartilaginous rings surround the ventral trachea and are critical for maintaining tracheal shape and rigidity (**Fig. 1.1a,c**, Cardoso and Whitsett, 2008; Hines et al., 2013; Morrisey and Hogan, 2010). These cartilage rings are derived from SOX9+ cartilage progenitor cells that begin to be specified in the ventral foregut mesenchyme at E10.5 and exhibit periodicity of the cartilage rings along the rostral caudal axis of the trachea by E13.5 (Bi et al., 1999; Snowball et al., 2015a). The resulting cartilage rings are connected at the dorsal trachea by trachealis smooth muscle, which provides flexibility for the expansion and contraction of the trachea. In contrast, the esophagus is surrounded by striated and smooth muscle derived from the cardiopharyngeal mesoderm (Gopalakrishnan et al., 2015). Both esophageal and

trachealis smooth muscle specification can be observed by E13.5 in mice with the expression of smooth muscle actin (SMA), one of the earliest markers of smooth muscle.

Transcriptional regulation of tracheal and esophageal endodermal fates is tightly coupled to the induction of the corresponding mesenchymal cell type (Hines et al., 2013). Consistent with this, loss of *Nkx2.1* from the endoderm resulted in reduction and malformation of tracheal cartilage, with ventral expansion of smooth muscle (Minoo et al., 1999b; Que et al., 2007a). Conversely, loss of *Sox2* resulted in a reduction of smooth muscle and an expansion of tracheal cartilage markers (Teramoto et al., 2019). This suggests that mesenchymal fate specification in the foregut is partially dependent on epithelial-to-mesenchymal crosstalk downstream of the NKX2.1-SOX2 regulatory program.

Some of the epithelial signals that regulate mesenchymal cell fate have been identified. Loss of epithelial WNT secretion from the ventral foregut endoderm in $Wls^{lox/lox}$; Shh^{Cre} mutant embryos resulted in a lack of SOX9 expression in the tracheal mesenchyme and subsequent lack of tracheal cartilage rings (Snowball et al., 2015a). Similarly, loss of mesenchymal WNT reception in $Ctnnb1^{lox/lox}$; $Dermo1^{Cre}$ embryos resulted in a nearly complete loss of tracheal cartilage markers and a corresponding expansion of smooth muscle differentiation in the ventral aspects of tracheal mesenchyme (Hou et al., 2019; Kishimoto et al., 2019; Snowball et al., 2015a). Thus, epithelial-mesenchymal WNT signaling is critical for tracheal mesenchyme specification to cartilage, yet transcriptional regulation of WNT ligand expression in the foregut is unknown.

Esophageal mesenchyme specification to striated and smooth muscle occurs

concomitant with tracheal cartilage and trachealis specification. The esophageal striated muscle arises from the cranial mesoderm, and is induced via a transcriptional cascade including TBX1, ISL1, MYF5, and MYOD (Comai et al., 2019; Gopalakrishnan et al., 2015). Given that *Sox2* mutants exhibit a loss of esophageal smooth muscle, it is likely that signals from the developing esophageal epithelium influence muscle differentiation. SHH is a known regulator of smooth muscle formation across contexts (Huycke et al., 2019a; Mao et al., 2010), and is also expressed in the esophageal epithelium from E11.5 (Litingtung et al., 1998). Mice lacking *Shh* exhibit severe morphological defects including a lack of smooth muscle formation in the airway (Kim et al., 2015; Litingtung et al., 1998; Pepicelli et al., 1998) and improper patterning of tracheal cartilage (Miller et al., 2004; Sala et al., 2011). Thus, SHH is a promising candidate for the coupling of esophageal epithelial and mesenchymal fates, yet this connection has not yet been demonstrated. In summary, while WNT and SHH signaling are likely mediators of mesenchymal specification in the developing trachea and esophagus, how these signals are regulated to induce coordinated mesenchymal differentiation is currently unknown.

Research aims

This study uses genome-wide approaches to elucidate how early tracheoesophageal patterning occurs. Specifically, this study aims to determine 1) what the transcriptional profiles are that define the dorsal-ventral domains of the foregut that give rise to the trachea and esophagus, and 2) how this dorsal-ventral identity is regulated by NKX2.1 and SOX2. We perform single cell RNA-sequencing (scRNA-seq) during and immediately following tracheoesophageal separation to identify transcriptomic profiles

that distinguish trachea from lung and esophagus. To understand the extent to which NKX2.1 regulates tracheoesophageal identity, we performed bulk RNA-sequencing on *Nkx2.1*$^{-/-}$ mutant tracheas, and NKX2.1 chromatin immunoprecipitation and sequencing (ChIP-seq) to identify NKX2.1 dependent and NKX2.1-independent tracheal programs. To decouple the role of NKX2.1 in directly regulating tracheoesophageal fate from its role in repressing *Sox2* and therefore indirectly regulating tracheoesophageal fate, we utilized our sequencing data and genetic mouse models to examine the consequences of loss of both *Nkx2.1* and *Sox2* on tracheoesophageal patterning. Finally, we show that NKX2.1 regulates epithelial-to-mesenchymal signaling via novel targets *Wnt7b* and *Shh* to specify mesenchymal cartilage and smooth muscle development. Together these data provide a framework for understanding transcriptional regulation of tracheoesophageal fate specification of the epithelial and mesenchymal lineages.

Figures

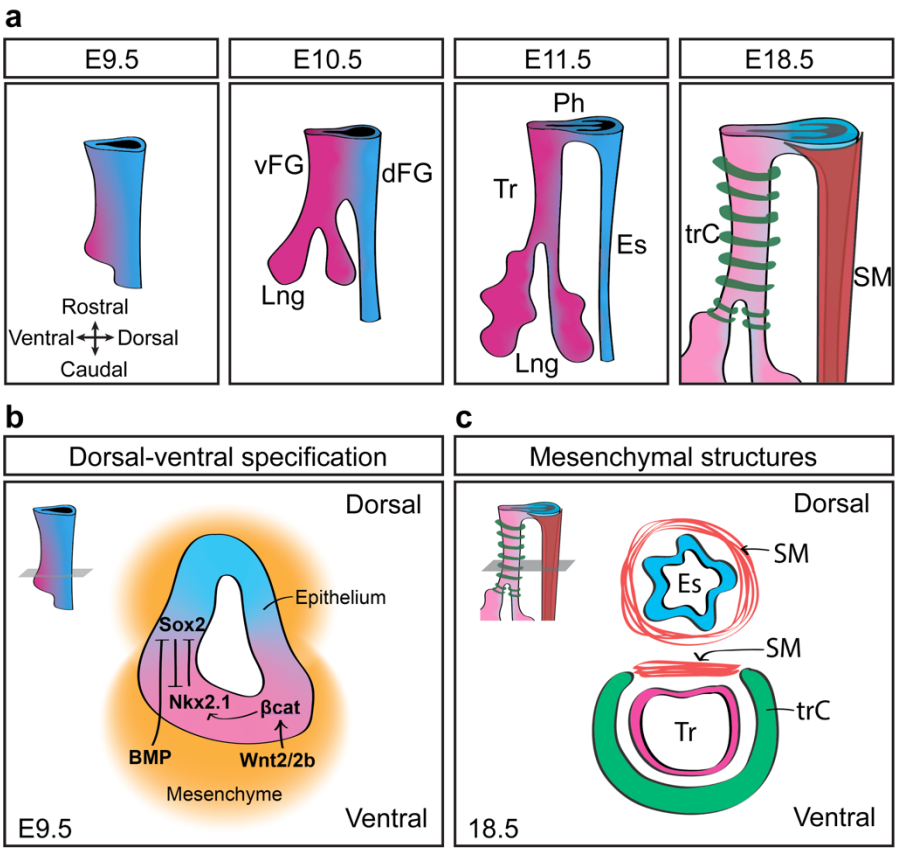


Figure 1.1 Overview of foregut development in the mouse embryo. a. Schematic of dorsoventral specification and physical separation of the foregut endoderm to lungs, trachea, and esophagus between embryonic day (E) 9.0-18.5, and mesenchymal specification of tracheal cartilage (trC) and esophageal smooth muscle (SM) at E18.5. vFG: ventral foregut, dFG: dorsal foregut, Lng: lung, Tr: trachea, Es: esophagus. **b.** Mesenchymal signals (BMP, WNT2/2b) that establish Sox2 and Nkx2.1 expression along the dorsoventral axis of the common foregut tube. **c.** Depiction of mesenchymal tissue types surrounding the trachea and esophagus at E18.5.

Chapter 2

Single cell transcriptomics of the developing trachea and esophagus

Rationale

Currently, the tracheal and esophageal transcriptional programs are unknown at the transcriptome-wide scale, and are instead largely defined by the expression of *Nkx2.1* in the ventral foregut/trachea and lung, and *Sox2* enrichment in the dorsal foregut/esophagus (Ishii et al., 1998; Minoo et al., 1999a; Que et al., 2007a; Yuan et al., 2000). The progress of single cell sequencing technologies in the past several years has facilitated the exploration of cell populations that could not otherwise be elucidated using bulk RNA-sequencing (RNA-seq) approaches. For the study of foregut specification, we posited that single cell sequencing experiments would be particularly useful in identifying dorsal and ventral cell populations within the common foregut tube, which could not be physically dissected or distinguished using bulk RNA-seq experiments. To understand the extent to which cells in the developing foregut differ on a transcriptome-wide scale and identify new cell-type specific markers for the study of fate specification, we performed single-cell RNA sequencing (scRNA-seq) on E10.5 mid-separation and E11.5 post-separation dissected mouse foreguts.

Identification of cell populations in the developing foregut epithelium

To generate single-cell transcriptomes of the developing foregut, we isolated live foregut epithelial cells via fluorescence activated cell sorting (FACS) and sequenced using droplet-based single-cell RNA-sequencing **(Fig. 2.1)**. We then identified transcriptionally distinct cell clusters using single-nearest neighbor clustering through Seurat (V3), a publicly available analysis platform based in R (Stuart et al., 2018). The analysis of single cell sequencing data involves the identification of cell “clusters”, groups

of cells that share similar transcriptomic profiles and can thus be classified as cells of similar fate. The identification of these cell clusters is driven by genes that show high variability between cells across the entire sample, and thus can be used to distinguish different cell types. In developmental systems where most cells are dividing to contribute to the growth of the embryo, transcriptional signatures of cell cycle often dominate this group of variable genes. Indeed, this was the case for our samples, where the cell clusters demarcated cells in different stages of the cell cycle, rather than the biological signatures of different cell fates. Thus, we applied a linear regression model to regress out cell cycle state, exposing more biologically relevant variable genes and allowing us to identify clusters that more clearly resembled the cell types present in the developing foregut.

Through our scRNA-seq analysis, we generated 6,407 single-cell transcriptomes at E10.5, comprising 5 cell clusters, and 10,493 single-cell transcriptomes at E11.5, comprising 7 cell clusters, and visualized these clusters using Uniform Manifold Approximation and Projection (UMAP) dimensional reduction (**Fig. 2.2a,b,** Becht et al., 2018; Stuart et al., 2018). We delineated the dorsal-ventral axis in our scRNA-seq data at E10.5 and E11.5 by projecting the expression levels of *Nkx2.1* and *Sox2* on the UMAP **(Fig. 2.2c,d)**. Similarly, *Sox9*, a marker of developing lung (Herriges et al., 2012; Perl et al., 2005; Rockich et al., 2013), marked a subset of *Nkx2.1*-positive respiratory cells, enabling us to distinguish distal lung from trachea **(Fig. 2.2c,d)**. However, the scRNA-seq analysis identified clusters at a higher resolution than these three markers could provide and required further validation. We therefore computationally performed differential expression analysis for each cell cluster using the FindAllMarkers function in Seurat (Stuart et al., 2018) to identify top marker genes for each cluster **(Table 2.1, 2.2)**. To

determine the spatial expression pattern of these markers, we next performed RNAscope fluorescent in-situ hybridization for these markers in the E10.5 and E11.5 embryonic foregut. Through these approaches, we identified the cell population corresponding to the pharynx, marked by *Foxe1* **(Fig. 2.3a)**. *Foxe1* was expressed in the pharyngeal epithelium and absent from the epithelium of the common foregut tube at E10.5 and separated trachea and esophagus at E11.5. We also identified a cell cluster corresponding to the ultimobranchial body, a derivative of the pharyngeal endoderm that gives rise to the follicular cells of the thyroid (Nilsson and Fagman, 2017), marked by *Crabp1* **(Fig. 2.3b)**. Additionally, within the E11.5 lung, we identified a cluster with unique markers including the distal lung marker *Bmp4* (**Table 2.1**, Weaver et al., 1999), demonstrating that the proximodistal lung axis can be identified by markers in our dataset. Signatures of proliferation, including the expression of *Hist1h1b*, *Hist1h2ap*, and *Mki67,* sub-divided the trachea and lung clusters at E11.5 **(Fig. 2.2a,b, Table 2.1)**. Other genes with previously characterized differential expression patterns within the foregut endoderm, including *Wnt7b, Bmp4, Slc2a3, Dlk1* further confirmed this clustering approach **(Table 2.1)**. Thus, our scRNA-seq data captures expression patterns of the few known dorsal-ventral markers of the foregut and identifies transcriptome-wide signatures of cell clusters of the trachea, esophagus, and lung.

Discovery of novel markers of the developing trachea and esophagus

We next leveraged our scRNA-seq data to expand our limited knowledge of dorsoventral patterning of the developing foregut at the whole transcriptome scale. Using differential expression analysis of cell clusters, we identified markers of ventral foregut,

lung, and dorsal foregut cells at E10.5 **(Fig. 2.4a)** as well as markers of trachea, lung and esophagus at E11.5 **(Fig. 2.4b)**. We used RNAScope to visualize the spatial expression pattern of some of these genes in the undivided E10.5 foregut and lung, and the E11.5 trachea, esophagus, and lung. We identified several genes, including *Wnt7b*, *Irx2*, *Crlf1*, and *Etv5* that showed similar expression patterns to *Nkx2.1* in that they spanned all respiratory cells at E10.5 and E11.5 **(Fig. 2.4, 2.5)**. Notably, we also found genes that exhibited enrichment in the ventral foregut at E10.5 and specifically marked the trachea and proximal airway at E11.5 but were not markers of the lung at either stage, such as *Tppp3*, *Pcdh10*, *Ly6h*, and *Cldn18* **(Fig. 2.4, 2.6)**. This indicates that the trachea and proximal airway become transcriptionally distinct from the lung during early respiratory development and suggests that trachea and lung may each be actively specified early, rather than a lung-specific program being built upon a general respiratory program. Additionally, we identified genes such as *Klf5*, *Dcn*, *Krt19*, and *Pitx1* that were enriched in the dorsal foregut/esophageal cells **(Fig. 2.4, 2.7)**. Our identification of *Krt19* as a marker of the dorsal foregut at E10.5 and esophageal epithelium at E11.5 suggests that, in addition to cell identity, specification of epithelial structure begins early in development, and may be investigated using this scRNA-seq analysis. As a whole this analysis provides a thorough characterization of multiple dorsoventral markers along the rostral-caudal axis of the developing foregut. The resolution of this data allows for the selection of new markers of specific populations within the foregut, as well as candidate genes for further investigation into their role in foregut development.

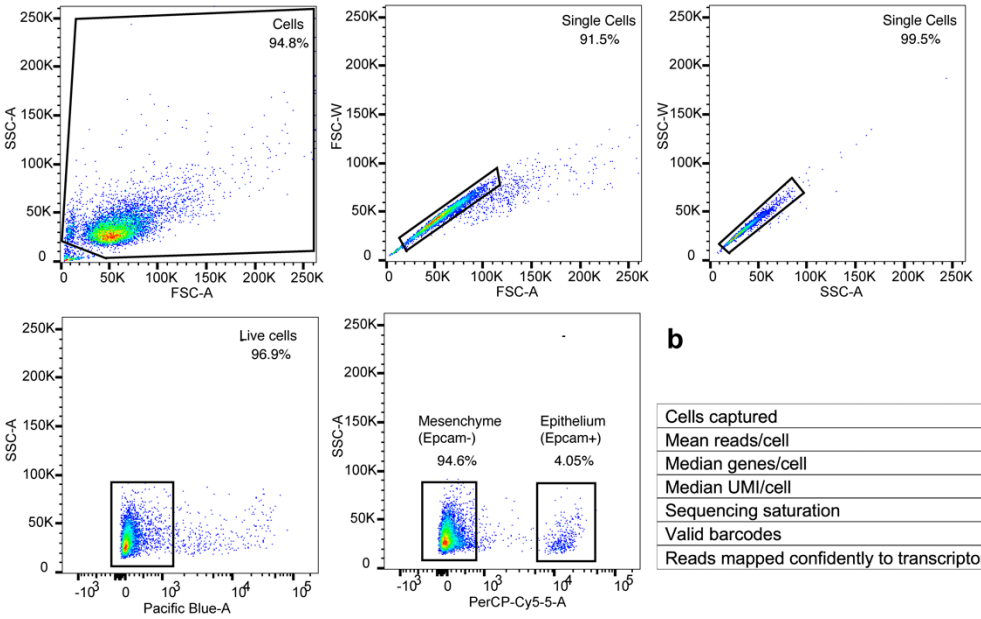
Conclusions and discussion

This scRNA-seq analysis profiles foregut specification prior to and immediately after tracheoesophageal specification. We identified new and robust markers of the dorsal and ventral domains of the unseparated foregut, and the newly separated trachea and esophagus. We also distinguish trachea/proximal airway from distal lung cells during early establishment of respiratory identity. While genes specific to differentiated cell types of the trachea later in embryonic development have been identified previously (Morrisey and Hogan, 2010), we identified a unique tracheal transcriptional profile during early stages of respiratory development and propose that early tracheal specification is active, and distinct from lung differentiation.

Of note, all of the tracheal and esophageal markers we identified and validated were dorsoventrally restricted in the common foregut tube of E10.5 embryos prior to physical separation of the trachea and esophagus, consistent with fate specification beginning before tracheoesophageal separation. Additionally, with the exception of *Tppp3* and *Ly6h*, the vast majority of genes we validated obeyed the NKX2.1 expression boundary, whether they were co-expressed with NKX2.1 in the ventral foregut or directly opposing NKX2.1 in the dorsal foregut. None of the genes we examined followed the same expression pattern as SOX2, with high expression in the dorsal foregut and lower expression extending into the NKX2.1+ ventral domain. This suggests that the SOX2 expression pattern is unique and requires closer examination of upstream regulators of *Sox2* expression during foregut development. Additionally, this shows that NKX2.1 is a more robust delineator of the future tracheal and esophageal domains than SOX2, and identifies other genes that can be used to distinguish these two populations.

Interestingly, there were several populations of cells we expected to resolve using our single cell sequencing approach that we were unable to capture. First, because the dorsal and ventral domains of the E10.5 foregut are actively separating during this time, we hypothesized that we may identify a transcriptionally distinct population of actively separating foregut cells, as proposed by others (Han et al., 2019; Kim et al., 2019). In addition to an actively separating population at E10.5, we also expected to identify multiple cell clusters within the E11.5 trachea that resemble heterogeneity or regional specification. Indeed, through immunofluorescent staining, we detect SOX2, NKX2.1 double positive cells at the boundary where separation occurs in E10.5 foreguts, as well as differences in SOX2 expression along the dorsal-ventral axis of the E11.5 trachea. However, even with increased resolution within the clustering parameters, we did not resolve a separating boundary population at E10.5 or multiple populations within the trachea that would indicate dorsal-ventral heterogeneity at E11.5. This is likely because, while these populations are unique in that they express both *Nkx2.1* and *Sox2*, there are either not other detectable genes, or too few genes, that are specific to these cells that would identify them as a unique population. This also suggests that while NKX2.1 and SOX2 likely regulate a morphogenetic program to influence separation, the separation process itself may rely primarily on post-transcriptional activity that is not detectable through RNA-seq approaches. With an increased number of tracheal cells or deeper sequencing, future experiments may be able to detect an actively separating population and/or heterogeneity within the trachea. Nonetheless, our data reliably identified tracheal, esophageal, and lung cell clusters of the developing foregut, and together these data uncover a multitude of previously unknown genes that define these cell populations.

a



b

	E10.5 foregut	E11.5 foregut	E11.5 lung
Cells captured	6,603	4,386	6,705
Mean reads/cell	101,904	162,554	91,862
Median genes/cell	4,164	4,781	4,586
Median UMI/cell	16,591	21,923	20,925
Sequencing saturation	64.7%	67.9%	50.6%
Valid barcodes	97.2%	97.4%	97.6%
Reads mapped confidently to transcriptome	64.7%	62.8%	65.3%

Figure 2.1. Quality control metrics and characterization of single-cell RNA sequencing experiments. a. Representative gating strategy for FACS of mouse embryonic foreguts. Top row shows gating for single cells, and bottom row shows gating for sytox-negative live cells (left) and epCAM-positive epithelial cells (right). **b.** Quality control metrics of scRNA-seq experiments identified with CellRanger.

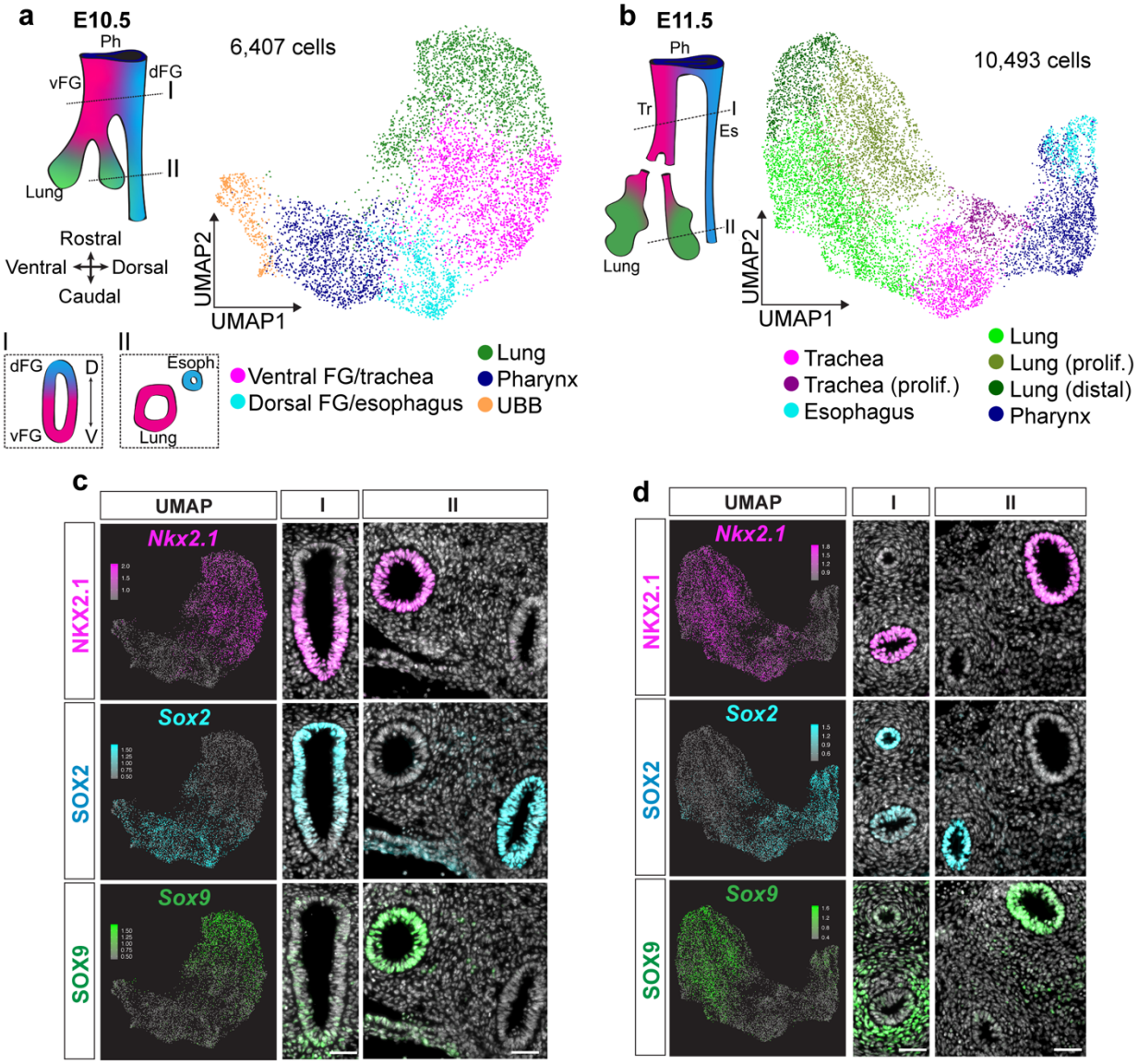


Figure 2.2. scRNA-seq captures cell types of the developing foregut. Dissected, FACS-purified foregut epithelial cells were subjected to droplet-based single-cell RNA sequencing at E10.5 and E11.5. **a.** UMAP representation of 6,407 cells identified at E10.5 and **b.** 10,493 cells identified at E11.5. Colors represent cell populations identified using shared nearest neighbor clustering. vFG: ventral foregut, dFG: dorsal foregut, Ph: pharynx, UBB: ultimobranchial body, Tr: trachea, Es: esophagus. **c-d.** Immunofluorescent staining of NKX2.1 (magenta), SOX2 (cyan), and SOX9 (green) in **c.** E10.5 and **d.** E11.5 foregut and lung. First column shows UMAP of gene expression level of *Nkx2.1*, *Sox2*, and *Sox9* as detected by scRNA-seq. Second and third columns represent regions of the foregut along the rostral-caudal axis as depicted in the schematic of **a** and **b.** Scale bar = 50um.

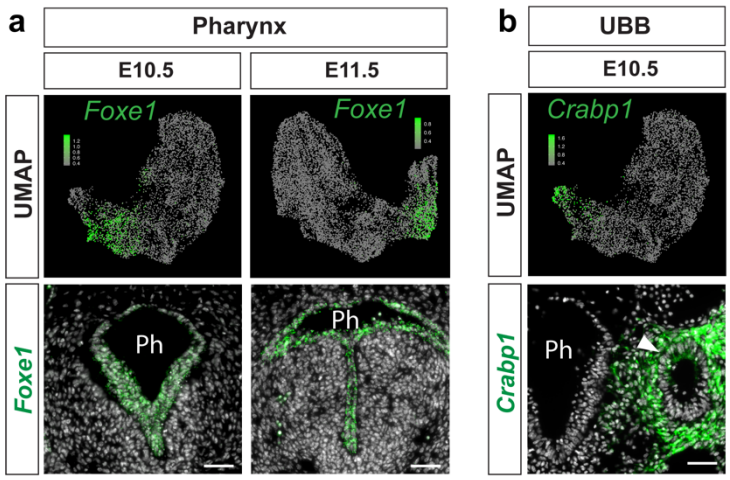


Figure 2.3. Validation of additional cell clusters identified with scRNA-seq of the foregut. a. Identification of pharynx cell cluster. Top row: projection of *Foxe1* RNA expression as determined by E10.5 and E11.5 scRNA-seq on UMAP (top row). Bottom row: RNA localization of *Foxe1* in the E10.5 and E11.5 pharynx (Ph). **b.** Identification of ultimobranchial body (UBB) cell cluster. Top row: projection of *Crabp1* RNA expression as determined by E10.5 scRNA-seq on UMAP. Bottom row: RNA localization of *Crabp1* in the E10.5 ultimobranchial body (UBB) (arrowhead). Scale bar = 50um.

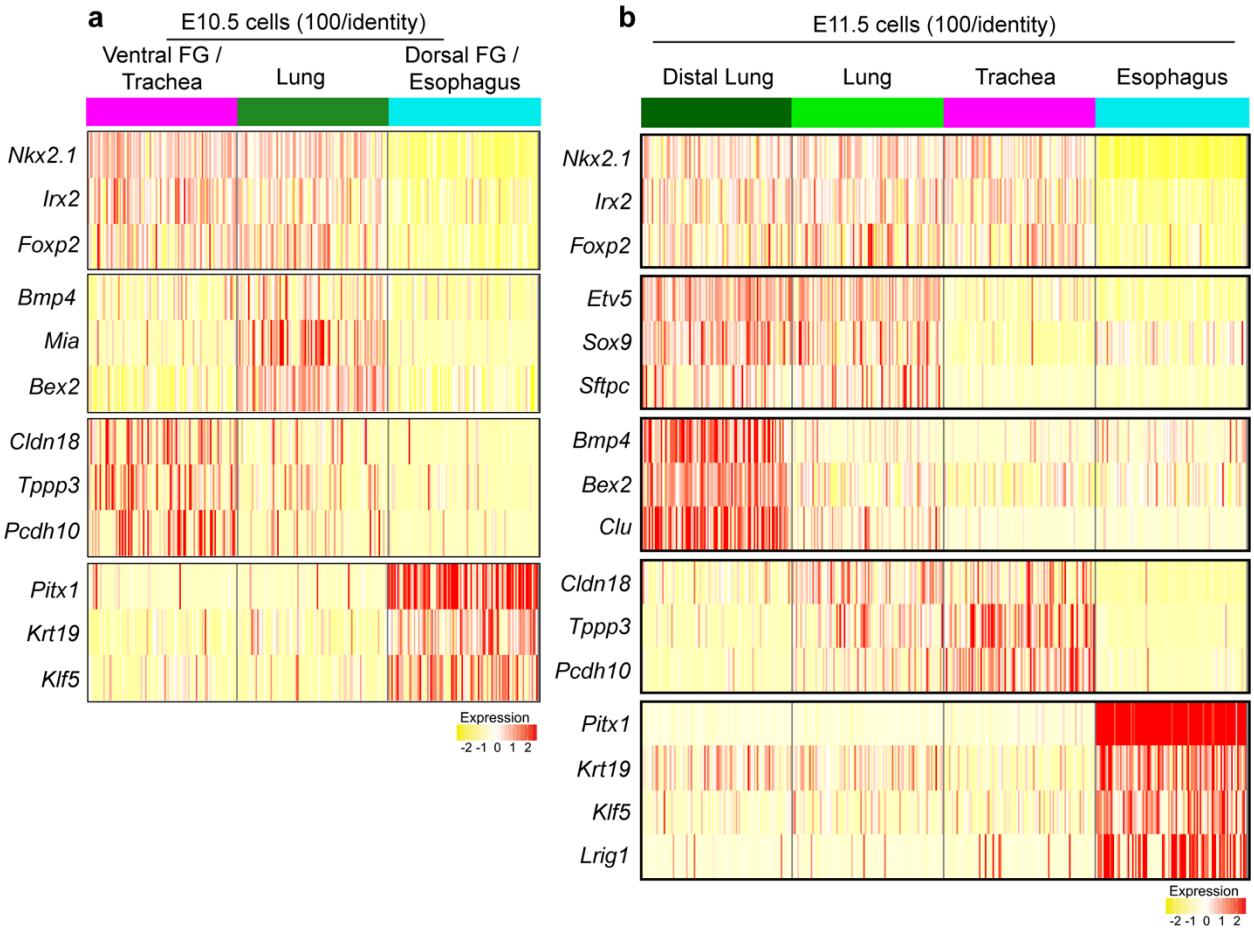


Figure 2.4. Differentially expressed genes between trachea, esophagus, and lung cells. a. Heatmap of selected marker gene expression across 100 E10.5 cells each of ventral foregut/trachea, lung, and esophagus/dorsal foregut as identified by scRNA-seq. Selected genes are markers of respiratory, lung, ventral FG/trachea, and dorsal FG/esophagus cells (top to bottom). b. Heatmap of selected marker gene expression across 100 E11.5 cells each of distal lung, lung, trachea, and esophagus cells as identified by scRNA-seq. Selected genes are markers of respiratory, lung, distal lung, trachea, and esophagus cells (top to bottom).

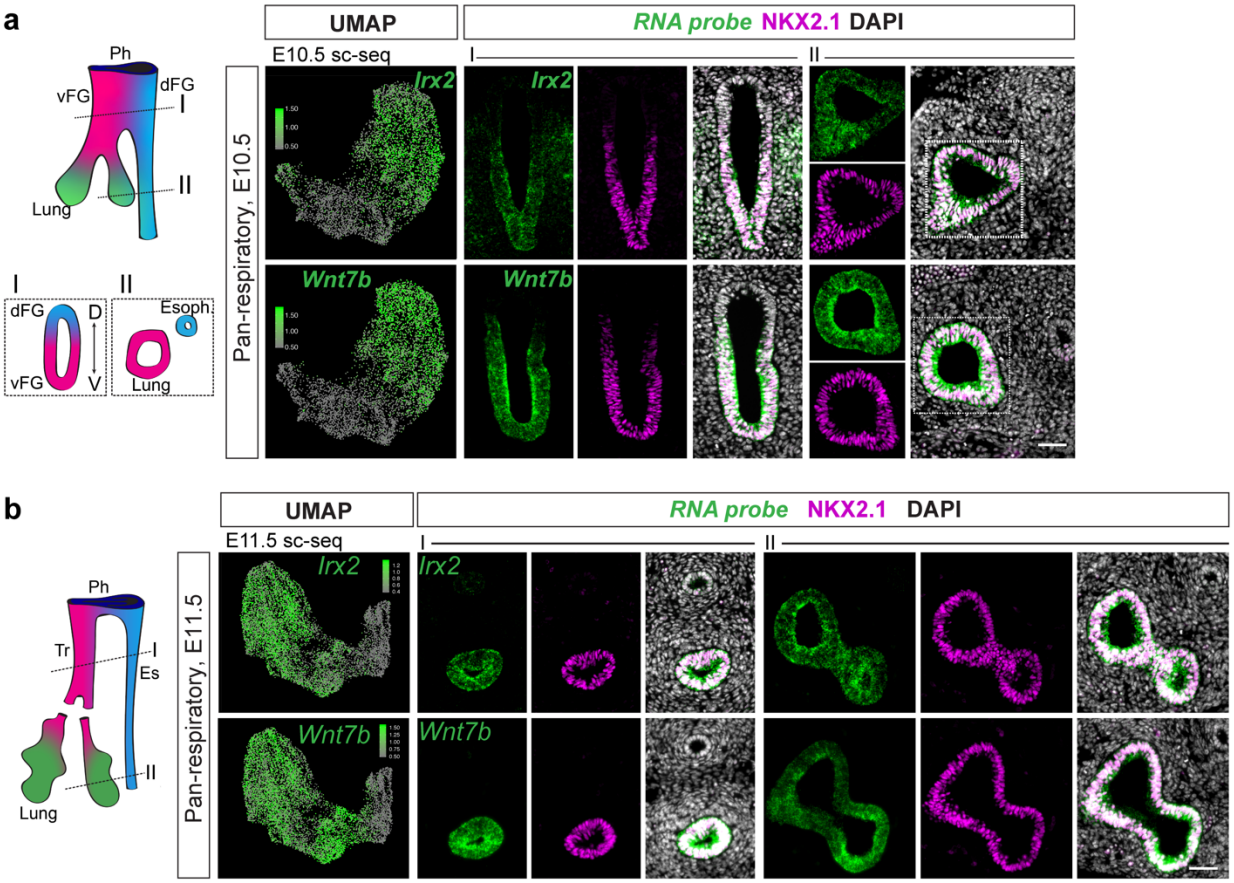


Figure 2.5. Pan-respiratory markers identified by scRNA-seq. RNA localization of pan-respiratory genes *Irx2* and *Wnt7b* identified from differential expression analysis of cell populations in **a.** E10.5 and **b.** E11.5 scRNA-seq data. First column shows projection of RNA expression level as determined by scRNA-seq on UMAP. Staining panels show RNA expression of marker gene (green) and NKX2.1 expression (magenta) at the positions pictured in the schematic on the right. Scale bar = 50um.

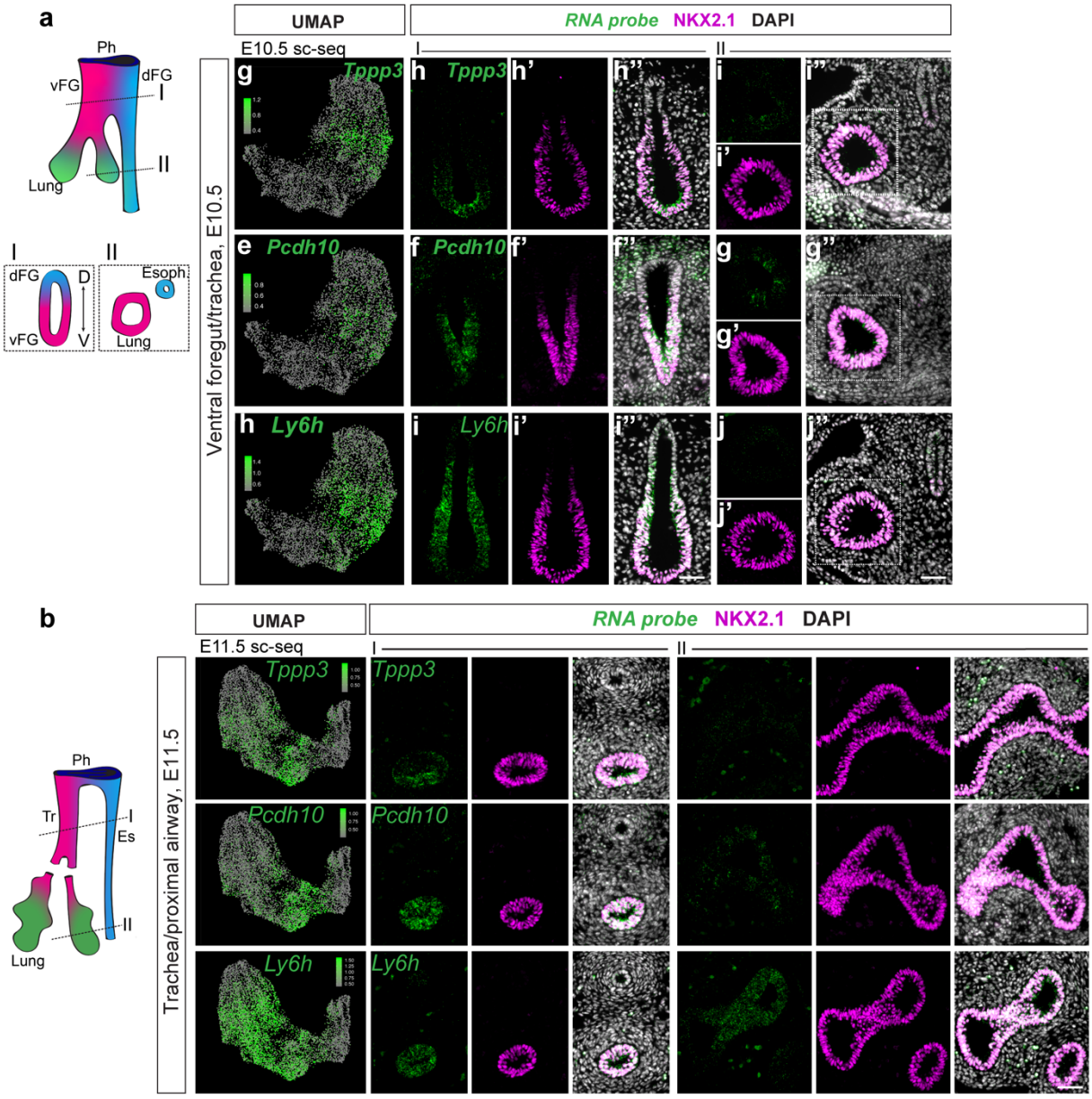


Figure 2.6. Ventral foregut/trachea markers identified by scRNA-seq. RNA localization of ventral foregut/trachea enriched genes *Tppp3, Pcdh10* and *Ly6h* identified from differential expression analysis of cell populations in **a.** E10.5 and **b.** E11.5 scRNA-seq data. First column shows projection of RNA expression level as determined by scRNA-seq on UMAP. Staining panels show RNA expression of marker gene (green) and NKX2.1 expression (magenta) at the positions pictured in the schematic on the right. Scale bar = 50um.

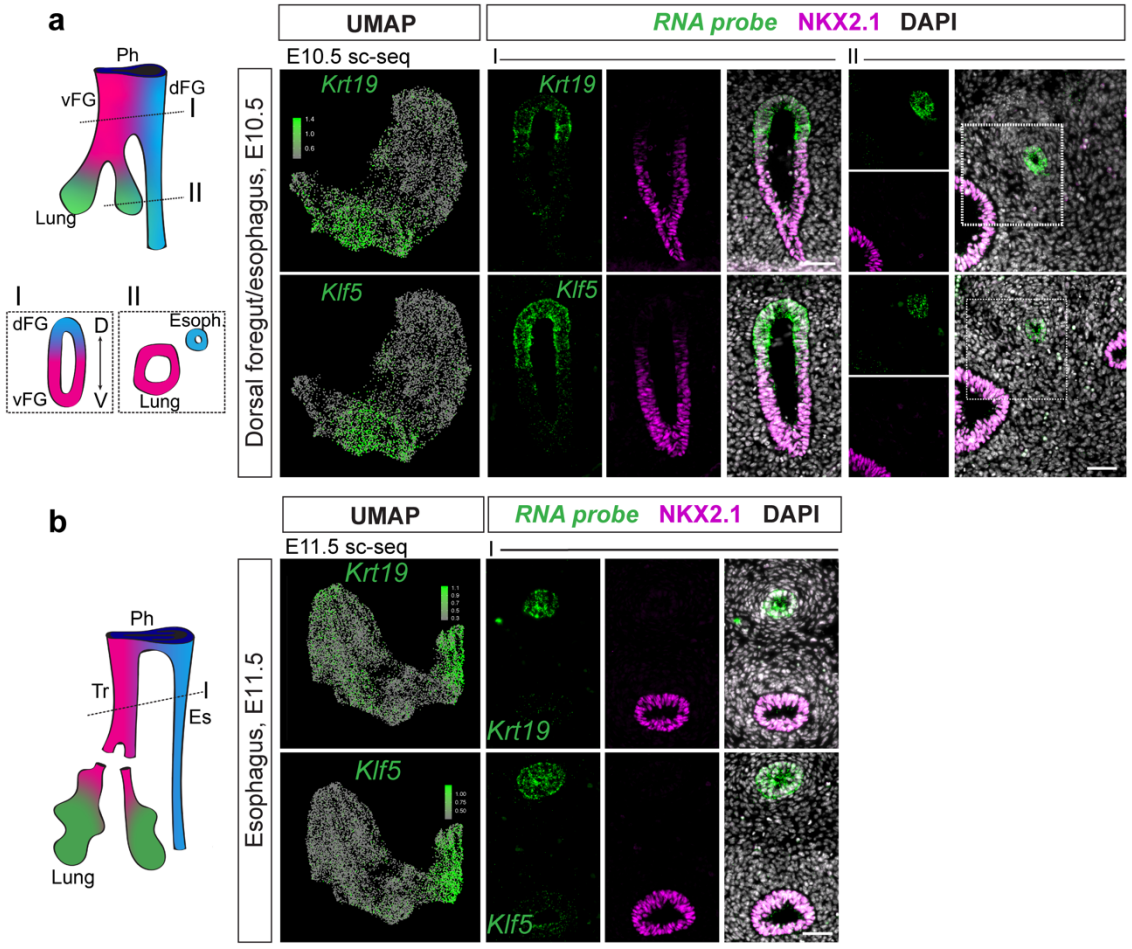


Figure 2.7. Dorsal foregut/esophagus markers identified by scRNA-seq. RNA localization of ventral foregut/trachea enriched genes *Krt19* and *Klf5* identified from differential expression analysis of cell populations in **a.** E10.5 and **b.** E11.5 scRNA-seq data. First column shows projection of RNA expression level as determined by scRNA-seq on UMAP. Staining panels show RNA expression of marker gene (green) and NKX2.1 expression (magenta) at the positions pictured in the schematic on the right. Scale bar = 50um.

Table 2.1. Markers of foregut populations in scRNA-seq data at E10.5. Cell cluster markers identified by differential expression analysis using FindMarkers in Seurat in E10.5 scRNA-seq data. Cluster=cell cluster identity, Gene=official gene symbol of marker gene, avg_logFC=average log fold change of gene enrichment compared to all other clusters, p_val_adj=adjusted p value (Wilcoxon Rank Sum test).

E10.5 scRNA-seq cluster markers

Cluster	Gene	avg_logFC	p_val_adj
Lung	Dlk1	1.48	0
	Slc2a3	1.26	0
	Bex2	0.92	0
	Bex4	0.92	0
	Crlf1	0.88	0
	Sox9	0.82	0
	Id2	0.71	0
	Epha4	0.66	0
	Tgfb2	0.65	0
	Wnt7b	0.64	0
	Sftpb	0.63	0
	Atox1	0.54	0
	Actr3	0.51	0
	Calm1	0.49	0
	Etv5	0.53	4.55E-299
	Oat	0.54	6.32E-299
	Meg3	0.67	4.74E-270
	Ccnd1	0.52	1.29E-264
	Clu	0.50	1.41E-228
	Apoe	0.77	4.98E-118
Trachea/ ventral foregut	Ly6h	0.73	0
	Cdkn1c	0.69	0
	Nkx2-1	0.66	0
	Ces1d	0.50	8.53E-290
	Irx3	0.54	1.14E-282
	Add3	0.52	2.56E-278
	H2-Q1	0.48	1.64E-227
	H2-Q4	0.54	6.48E-227
	Cldn18	0.52	1.65E-213
	Capn6	0.44	1.46E-211
	Emid1	0.42	6.18E-207
	Tppp3	0.46	7.56E-207
	Sparc	0.55	4.14E-205
	Irx1	0.46	1.36E-202
	Irx2	0.43	1.10E-199
	Rasl11b	0.49	2.10E-185
	Mecom	0.40	2.30E-164
	Irx5	0.36	6.10E-163
	Igfbp2	0.39	6.41E-122
	Map1b	0.37	2.68E-81
Esophagus/ dorsal foregut	Pitx1	1.08	0
	Serpina1a	0.72	1.13E-286
	Vtn	0.71	2.86E-218
	Ntn1	0.35	7.53E-201
	Sox2	0.67	1.21E-194
	Igfbp2	0.73	7.07E-175
	Ifitm1	0.62	4.87E-169
	Rbp1	0.56	7.33E-140
	Krt7	0.45	3.86E-113
	Gsto1	0.41	2.10E-107
	Klf5	0.43	1.50E-104
	Bcl11a	0.37	4.36E-99
	Hoxc4	0.39	5.81E-99
	Wfdc2	0.44	6.45E-80
	Krt19	0.35	2.11E-78
	Krt8	0.35	1.09E-71
	Krt18	0.35	3.39E-61
	Mt1	0.50	1.24E-56
	Arl4a	0.34	3.24E-52
	Mt2	0.45	7.22E-33

Cluster	Gene	avg_logFC	p_val_adj
Pharynx	Dcn	1.70	0
	Meis2	0.93	0
	Hadh	0.83	0
	Sfrp1	0.73	0
	Foxe1	0.61	0
	Trp63	0.54	0
	Perp	0.66	4.57E-297
	Sfn	0.62	1.06E-295
	Emb	0.70	2.32E-265
	Siva1	0.55	2.76E-246
	Shisa2	0.50	4.13E-239
	Sox2	0.55	3.20E-227
	Ifitm3	0.61	3.91E-222
	4631405K08	0.54	6.44E-222
	Fxyd3	0.53	3.70E-221
	Hspb1	0.73	2.90E-209
	Csrp2	0.56	5.95E-193
	Isl1	0.55	1.25E-188
	Krt19	0.57	1.42E-174
	Bcam	0.50	9.60E-165
Ultimobranchial body	Aldh1a1	1.99	0
	Crabp1	1.09	0
	Pax9	1.04	0
	Mfap4	0.99	0
	Col1a2	0.81	0
	Hoxb1	1.23	3.10E-300
	Edn1	0.93	1.34E-279
	Mest	1.61	3.88E-236
	Nkx2-5	0.80	1.58E-186
	Meis2	1.06	9.37E-179
	Mpped2	0.85	1.28E-168
	Crabp2	1.00	1.43E-168
	Gpx3	0.98	2.34E-164
	Tmsb4x	1.02	1.92E-161
	Sh3glb1	0.87	1.90E-147
	Shisa2	0.80	2.60E-139
	Igfbp5	1.05	2.55E-134
	Ripply3	0.89	2.69E-127
	Hoxb4	0.89	8.24E-112
	Arg1	0.77	4.83E-70

Table 2.2. Markers of foregut populations in scRNA-seq data at E11.5. Cell cluster markers identified by differential expression analysis using FindMarkers in Seurat in E11.5 scRNA-seq data. Cluster=cell cluster identity, Gene=official gene symbol of marker gene, avg_logFC=average log fold change of gene enrichment compared to all other clusters, p_val_adj=adjusted p value (Wilcoxon Rank Sum test).

E11.5 scRNA-seq cluster markers

Cluster	Gene	avg_logFC	p_val_adj
Lung	Cdkn1c	0.51316735	0
	Ccnd1	0.45595656	0
	Npc2	0.31783079	0
	Tuba1a	0.2969975	1.72E-258
	Tgfb2	0.40112806	9.97E-246
	Nnat	0.28024743	1.67E-207
	Galk1	0.27244832	1.29E-204
	Sparc	0.28323914	3.48E-195
	Cldn18	0.30929867	1.01E-187
	Sesn3	0.29856412	3.98E-185
	Irx1	0.28709998	8.17E-185
	Btg2	0.3164168	3.03E-180
	Ly6h	0.28974267	2.06E-174
	Krt18	0.26531135	2.45E-167
	Tmsb4x	0.28561535	1.39E-166
	Sftpc	0.47630693	3.29E-142
	Thbs1	0.25636556	1.84E-140
	Smoc2	0.28538315	7.75E-135
	Ckb	0.26866931	2.31E-120
	Etv5	0.25733023	3.43E-111
Lung (prolif.)	Hist1h1b	0.71036195	0
	Sftpc	0.70433264	0
	Hist1h1e	0.64966184	0
	Hist1h2ap	0.64104285	0
	2810417H13	0.60868437	0
	Top2a	0.59574189	0
	Cdk1	0.44083774	0
	Tyms	0.39217147	0
	Tubb5	0.32894869	0
	Foxp2	0.42481269	1.44E-281
	Fbxo5	0.33790608	9.14E-280
	Etv5	0.41726634	4.96E-264
	Mia	0.32629996	1.60E-235
	Spc24	0.32851715	1.28E-225
	Nkx2-1	0.33384303	4.24E-221
	Hmgb2	0.3419074	1.50E-197
	Ube2c	0.39832736	8.89E-194
	Sox9	0.33162432	4.01E-169
	Rrm2	0.35412861	1.54E-157
	Hbb-y	0.46771701	8.73E-05
Lung (distal)	Dlk1	1.7280603	0
	Bex4	1.12988596	0
	Apoe	1.10048662	0
	Bmp4	1.01948068	0
	Clu	0.84780196	0
	Bex1	0.6148226	0
	Slc2a3	0.58799674	0
	Atox1	0.51199168	0
	Bex2	0.80923235	8.17E-298
	Actr3	0.57063229	1.06E-296
	Lin7a	0.52535779	5.78E-271
	Vim	0.60558366	7.09E-264
	Selm	0.55847815	3.32E-258
	Id2	0.92200924	1.14E-241
	Hspa5	0.51761895	6.73E-227
	Crlf1	0.50963883	9.55E-227
	Etv5	0.51314341	6.89E-183
	Meg3	0.71235874	5.15E-179
	Peg3	0.58389342	2.42E-175
	Sox9	0.52018129	1.02E-165

Cluster	Gene	avg_logFC	p_val_adj
Trachea	Rps28	0.42999605	0
	Nr2f2	0.54270012	4.39E-292
	Hadh	0.5378054	1.38E-271
	Nrp2	0.37694762	1.58E-265
	Ccnd2	0.55617024	1.71E-255
	6330403K07	0.46314586	5.43E-244
	Lrrn3	0.37437738	2.06E-239
	Ptn	0.55521563	2.77E-238
	Pcdh10	0.4710236	4.12E-229
	Map1b	0.59607025	2.56E-228
	Mecom	0.46725265	3.13E-227
	Golim4	0.51681328	1.16E-225
	Ddit4	0.49054311	2.13E-216
	Anxa2	0.43099979	2.79E-201
	Igfbp5	0.55022363	1.17E-199
	Rps27rt	0.37248395	4.63E-198
	Id3	0.48149991	9.03E-197
	Fstl1	0.40149157	1.80E-178
	Tppp3	0.45937602	4.36E-143
	Id1	0.37312865	7.39E-111
Trachea (prolif.)	Hist1h2ap	1.25189966	1.03E-211
	Pcdh10	0.54797743	3.34E-141
	Nr2f2	0.5873678	3.08E-133
	Top2a	0.75563423	5.63E-133
	Hist1h2ae	0.70650668	5.71E-129
	Ptn	0.61794801	5.23E-125
	Malat1	0.4447495	3.23E-109
	Hist1h1b	0.54001552	7.13E-107
	Hadh	0.45583401	4.46E-97
	Mki67	0.52850552	1.37E-84
	Id1	0.53919562	3.42E-84
	Mecom	0.38921699	1.21E-80
	Hist1h1e	0.43021137	4.67E-76
	Incenp	0.38350599	2.53E-73
	Smc4	0.4015453	1.49E-69
	Sox2	0.38050536	2.33E-69
	Nusap1	0.42644973	1.28E-67
	Golim4	0.43427162	3.82E-64
	Cenpf	0.40028535	2.01E-54
	Xist	0.50356922	7.70E-09
Esophagus	Pitx1	1.79075559	0
	Dcn	1.5354889	0
	Fxyd3	0.96738619	0
	Anxa1	0.86846217	0
	Fam43b	0.72754829	0
	Klf4	0.71786781	5.97E-292
	Igfbp2	1.15982733	7.26E-284
	Ifitm1	0.85360745	1.25E-283
	Igfbp4	1.20521195	3.74E-272
	4631405K08	0.9456269	2.09E-269
	Sfn	0.7925042	3.44E-255
	Ifitm3	1.0629267	8.32E-251
	Ctnnb1	0.78449445	3.41E-245
	Krt15	0.73517445	3.66E-215
	Uqcr10	0.61749497	1.02E-212
	Fam210b	0.6807958	1.98E-211
	Perp	0.60063767	2.78E-206
	Isyna1	0.73395614	3.04E-188
	Sox2	0.73967935	3.27E-174
	Igfbp5	0.67688193	1.61E-131

Cluster	Gene	avg_logFC	p_val_adj
Pharynx	Hspb1	1.58216506	0
	Dcn	1.23153553	0
	Sfn	1.16219829	0
	Perp	0.93103326	0
	Fxyd3	0.92313058	0
	Ifitm3	0.91732006	0
	Anxa8	0.86335589	0
	Anxa2	0.83283698	0
	Krt15	0.82740568	0
	Siva1	0.80170828	0
	Meis2	0.79241114	0
	4631405K08	0.78488638	0
	Klf5	0.74051424	0
	Id1	0.72803025	0
	Bcam	0.71181802	0
	Pdgfa	0.68042724	0
	Upk3bl	0.64287081	0
	S100a10	0.63476409	0
	Sox2	0.6139526	0
	Wfdc2	0.60914954	9.35E-304

Chapter 3

Regulation of tracheoesophageal fate specification by Nkx2.1

Rationale

With our new transcriptome-wide understanding of the genes that define the developing trachea and esophagus, we sought to examine how tracheoesophageal identity is dysregulated upon *Nkx2.1* loss. *Nkx2.1*$^{-/-}$ embryos display a failure of tracheoesophageal separation and an adoption of some esophageal characteristics including high expression of SOX2 and P63 throughout the common tube, dysmorphic or absent tracheal cartilages, and ventral expansion of smooth muscle (Minoo et al., 1999a; Que et al., 2007b; Yuan et al., 2000). However, while *Nkx2.1* is the earliest known marker of respiratory fate and thought to be a master regulator of respiratory specification, we know very little about the extent and composition of the NKX2.1 regulatory program. The mesenchymal defects in *Nkx2.1*$^{-/-}$ mutant foreguts indicates that the NKX2.1 regulatory program must include mediators of epithelial to mesenchymal signaling to influence specification of the foregut mesenchyme. Additionally, the subsequent gain of SOX2 expression in *Nkx2.1*$^{-/-}$ mutant foreguts and our lack of understanding of NKX2.1 binding in the developing trachea makes it difficult to determine whether NKX2.1-regulated genes are under direct regulation by NKX2.1, or indirect regulation mediated by SOX2. The experiments described in this chapter define the NKX2.1-regulatory program, identify examples of NKX2.1 and SOX2 co-regulation as well as SOX2-independent NKX2.1 regulation, and suggest a potential mechanism for NKX2.1 regulation of mesenchymal fates.

Transcriptomic analysis of *Nkx2.1*$^{-/-}$ mutant foreguts

To identify the NKX2.1 transcriptional program, we performed RNA-seq on dissected, FACS-purified foregut epithelium from E11.5 *Nkx2.1*$^{-/-}$ and WT embryos **(Fig. 3.1a)**. We removed lung tissue during the dissection to focus on tracheoesophageal-specific changes. Differential expression analysis between the *Nkx2.1*$^{-/-}$ and WT foregut epithelium identified 257 NKX2.1-regulated genes, with 109 genes upregulated and 148 genes downregulated in *Nkx2.1*$^{-/-}$ foreguts (**Fig. 3.1b, Table 3.1, 3.2,** DESeq2, log2FC>0.7, padj<0.05). Our discovery of both an upregulated and downregulated set of genes in *Nkx2.1*$^{-/-}$ mutants agrees with studies in the developing mouse brain and alveolar cell *in vitro* (Little et al., 2019; Sandberg et al., 2016, 2018) that suggest NKX2.1 can act to activate and repress target genes.

To determine whether the NKX2.1 transcriptional program exhibited tracheoesophageal specificity at a transcriptome-wide level, we examined their expression in our scRNA-seq data. We calculated the expression level of NKX2.1-regulated genes as a percentage of all reads expressed in each cell and visualized this expression across each cell cluster at E11.5 **(Fig. 3.2a-e)**. As a whole, genes that were upregulated in *Nkx2.1*$^{-/-}$ mutants were enriched in cells of the esophagus and pharynx **(Fig. 3.2a,d)**, and genes that were downregulated in *Nkx2.1*$^{-/-}$ were enriched in tracheal and lung cells **(Fig. 3.2b-e)**. This supports the hypothesis that NKX2.1 positively regulates tracheal genes and negatively regulates esophageal genes at E11.5. We performed a similar analysis examining the expression of the NKX2.1 transcriptional program in E10.5 foreguts, to understand how these genes were expressed at an earlier stage of foregut development **(Fig. 3.2f-j)**. We found that genes positively regulated by NKX2.1 at E11.5

were also expressed in the esophagus and pharynx at E10.5 **(Fig. 3.2f,i)**, and genes negatively regulated by NKX2.1 at E11.5 were expressed in the E10.5 trachea and lung **(Fig. 3.2g,j)**. This agrees with our E11.5 analysis, and also suggest that NKX2.1 likely regulates these genes at earlier stages of foregut development as well. Interestingly, genes that increase in *Nkx2.1*$^{-/-}$ mutants also appear to be expressed in cells of the distal lung at E11.5 and lung at E10.5 in our scRNA-seq data **(Fig. 3.2a,d,f,i)**. While we are limited by the fact that our *Nkx2.1-/-* analysis excluded lung tissue, it is possible and likely that the regulatory role of NKX2.1 varies within respiratory cell types, and that NKX2.1 may have a different regulatory program in the trachea, proximal lung, and distal lung.

Identification of an NKX2.1-independent transcriptional program.

Our transcriptome-wide analysis of *Nkx2.1*$^{-/-}$ foreguts allows us to address the current prevailing model that the common foregut tube in *Nkx2.1*$^{-/-}$ mutants undergoes an esophageal fate conversion (Minoo et al., 1999a; Que et al., 2007b; Yuan et al., 2000). Surprisingly, our scRNA-seq dataset presented in Chapter 2 identified many genes that mark tracheal and esophageal cells that did not appear to change in expression in our *Nkx2.1*$^{-/-}$ mutant RNA-seq analysis **(Table1.1, 3.1, 3.2)**. We examined the spatial expression of several of these NKX2.1-independent genes using an *Nkx2.5-cre* strain which mediates recombination in the ventral foregut (**Fig. 3.3a**, Stanley et al., 2004) to generate *Nkx2.1*$^{lox/lox}$*; Nkx2.5*$^{cre/+}$ (*Nkx2.1*$^{tr/tr}$) embryos lacking NKX2.1 in the trachea. Using RNAscope, we found *Irx2*, *Ly6h*, and *Nrp2* to be tracheal-specific and maintained in the ventral epithelium of the unseparated foregut tube in E11.5 *Nkx2.1*$^{tr/tr}$ embryos **(Fig. 3.3b)**. Likewise, we found *Dcn*, *Ackr3*, and *Meis2* to be esophageal-specific and

maintained in the dorsal region of the common foregut tube in *Nkx2.1*$^{tr/tr}$ foreguts **(Fig. 3.3c)**. Immunofluorescent staining also showed that the esophageal genes LRIG1 and PITX1 were maintained in the dorsal region of *Nkx2.1*$^{-/-}$ foreguts **(Fig. 3.3d)**. These data are the first evidence of tracheal and esophageal genes that maintain their patterning in the absence of *Nkx2.1* and suggest that *Nkx2.1*$^{-/-}$ mutant foreguts do not undergo a complete tracheal-to-esophageal fate conversion.

We next examined the extent to which tracheoesophageal patterning is independent of NKX2.1 by generating bulk transcriptional profiles of tracheal and esophageal epithelium from dissected, FACS-purified WT trachea and esophagus at E11.5 **(Fig. 3.4a)**. Differential expression analysis on these datasets identified 1126 genes enriched in the trachea and 809 genes enriched in the esophagus of WT embryos **(Fig. 3.4b)**, many of which were also captured in our single cell analysis. Interestingly, only 11% of tracheal-enriched genes and 6% of esophageal-enriched genes in WT foreguts were affected by loss of *Nkx2.1*, and the majority of tracheal- or esophageal-enriched genes retained their dorsal-ventral patterning in *Nkx2.1*$^{-/-}$ foreguts **(Fig. 3.4b)**. We additionally performed principal component analysis on the *Nkx2.1*$^{-/-}$ mutant foreguts, WT foregut (trachea+esophagus epithelium), and the separated trachea and esophagus epithelium. We observed that, for the first principal component, the *Nkx2.1*$^{-/-}$ mutant epithelium fell between that of the trachea and esophagus, rather than clustering with the esophageal epithelium as a fate conversion hypothesis would suggest. While it is important to note that the *Nkx2.1*$^{-/-}$ mutant RNA-seq experiment **(Fig. 3.1a)** was performed separately from the trachea and esophagus RNA-seq experiment **(Fig. 3.4a)**, the experimental variation is likely captured in the second principal component, as it

separates these two experiments (y-axis of **Fig. 3.4c**). As a whole, these data identify an NKX2.1-independent dorsoventral gene expression program in the foregut that is contrary to a fate conversion hypothesis and indicates that NKX2.1 is not the sole master regulator of tracheal/esophageal fates.

Direct binding of NKX2.1 near NKX2.1-regulated genes.

To identify genomic regions directly bound by NKX2.1 and gain insight into the direct regulation of the NKX2.1 transcriptional program, we performed NKX2.1 chromatin immunoprecipitation followed by sequencing (ChIP-seq) on 175 pooled E11.5 WT trachea per replicate. As NKX2.1 is an epithelial-restricted transcription factor, we did not need to separate the epithelial and mesenchymal tissues prior to the ChIP. Instead, we dissected the entire E11.5 respiratory system, removed the lungs, and performed the ChIP on pooled E11.5 trachea. After sequencing, we performed peak calling using MACS to identify genomic regions where the NKX2.1 ChIP-seq library was significantly enriched over the input DNA. We identified 15,861 genomic regions (peaks) shared between two biological replicates that showed NKX2.1 binding (**Fig. 3.5a**, FDR<0.00001). We performed motif analysis on these shared peaks and confirmed that these peaks were centrally enriched for the known NKX2.1 motif (**Fig. 3.5b**, p=3.4e-37). Thus, we concluded that our ChIP-seq data captured NKX2.1 binding in the trachea and could be utilized for further exploration of tracheoesophageal development.

In parallel to our NKX2.1 ChIP-seq experiments, we performed ChIP-seq for the histone modifications H3K27ac that marks active enhancers and H3K27me3 that marks regions of heterochromatin. These experiments were performed using 175 manually separated E11.5 trachea and esophagus, with no replicates. Since the mesenchyme

comprises ~95% of dissected foregut tissue (eg. **Fig. 2.1a**), these datasets primarily capture the profile of these histone marks in the E11.5 tracheal and esophageal mesenchyme. Additionally, due to the quantity of material required for successful ChIP-sequencing experiments, we were unable to successfully sequence the esophageal H3K27me3 ChIP library. While genomic profiling of mesenchymal tissues was not the focus of this study, these datasets provide a valuable resource for future work identifying candidate enhancers for genes involved in mesenchymal specification.

To identify genes associated with NKX2.1 peaks and understand larger trends of NKX2.1 binding in the trachea, we generated peak-gene associations with GREAT (McLean et al., 2010) using the basal-plus-extension rule. We then performed gene ontology analysis on the top 1000 peaks by fold enrichment over input. Interestingly, the top biological process of genes associated with these NKX2.1 peaks was actin filament-based movement **(Fig. 3.6a)**. Other biological processes associated with these peaks epithelial tube branching in lung morphogenesis, lung morphogenesis, mesenchymal cell proliferation in lung development, and hair follicle development. The target gene associations with these morphological terms suggests that NKX2.1 is an important regulator of foregut morphogenesis, in addition to fate specification.

In a parallel project in the laboratory, we have explored regulation of foregut morphogenesis by the Eph-ephrin signaling ligand, EPHRIN-B2. In this project, we have determined that *Efnb2* is expressed in the dorsal foregut and esophagus, and is critical for foregut separation, but appears to not affect dorsal-ventral fate specification. We have also determined that *Efnb2* is negatively regulated by NKX2.1 during foregut development. Thus, we utilized our ChIP-seq data to determine whether this regulation

may be direct via binding by NKX2.1 near the *Efnb2* locus. Indeed, we saw 5 NKX2.1 ChIP-seq peaks within 70kb of the *Efnb2* locus, and confirmed NKX2.1 binding to these regions using ChIP-PCR for NKX2.1 in the E11.5 trachea, with a positive control region of conserved NKX2.1 binding observed in the developing brain (**Fig. 3.6b,c,** Sandberg et al., 2016). Ongoing work in the laboratory will focus on the functional role of these binding sites and the effect of loss of these regions on foregut morphogenesis. This provides evidence for one potential mechanism by which NKX2.1 regulates tracheoesophageal morphogenesis.

We further examined binding of NKX2.1 near select genes identified to be regulated by NKX2.1 in our RNA-seq analysis and marking specific cell types based on our scRNA-seq analysis. We observed multiple peaks at the promoter and within 10kb of the *Nkx2.1* gene, consistent with previous data suggesting that it may autoregulate its own expression (**Fig. 3.7a**, row1, Nakazato et al., 1997; Oguchi and Kimura, 1998; Tagne et al., 2012b). We also observed multiple NKX2.1 peaks at the promoter and within the locus of the *Sox2* gene, suggesting direct repression of *Sox2* by NKX2.1 (**Fig. 3.7a**, row2). In addition, NKX2.1 binding was observed at the promoter of genes such as *Pcdh10, Tppp3,* and *Klf5* that we identified as specific markers of the tracheal and esophageal lineages by scRNA-seq (**Fig. 3.7a**, row3-5), suggesting that these genes may also be direct targets of NKX2.1 regulation. We did not observe cluster-specific gene expression of genes associated with the top 1000 NKX2.1 ChIP-seq peaks (by log2FC) in our scRNA-seq data, suggesting that peak-gene associations alone do not infer regulatory specificity **(Fig. 3.7b)**.

We next asked whether NKX2.1 binding was more frequent amongst NKX2.1 regulated genes at the genome-wide scale. We compared our NKX2.1 ChIP-seq dataset with our *Nkx2.1*$^{-/-}$ mutant RNA-seq dataset and revealed that NKX2.1-regulated genes are associated with NKX2.1 ChIP-seq peaks at a higher frequency than observed at random (**Fig. 3.7c**, Fisher's exact test, $p<0.0001$). Furthermore, when we divided the NKX2.1 transcriptional program into genes that were upregulated or downregulated in *Nkx2.1*$^{-/-}$ mutants, we found that genes that are downregulated in *Nkx2.1*$^{-/-}$ mutants are more frequently associated with NKX2.1 ChIP-seq peaks (**Fig. 3.7c**, bottom row, Fisher's exact test, $p<0.0001$). These data suggest that whereas NKX2.1 is a direct positive regulator of tracheal-specific genes, repression of esophageal-specific genes may more often be indirect.

Analysis of co-regulation of the tracheoesophageal program by Nkx2.1 and Sox2.

Based on previous studies, NKX2.1 indirect regulation may be mediated through repression of *Sox2*. Indeed, given the genetically co-repressive relationship of SOX2 and NKX2.1 in the foregut, it is difficult to determine whether the transcriptional changes we observed in *Nkx2.1*$^{-/-}$ mutants are solely due to the loss of NKX2.1 or also due to the subsequent gain of SOX2. Thus, we devised a genetic strategy to uncouple NKX2.1 and SOX2 regulation by generating compound mutant embryos and determining whether loss of SOX2 leads to loss of regulation by NKX2.1. To achieve this, we utilized *Nkx2.5-cre* to generate *Nkx2.*$^{lox/lox}$*; Nkx2.5*$^{cre/+}$ (*Nkx2.1*$^{tr/tr}$) embryos lacking Nkx2.1, and *Nkx2.1*$^{lox/lox}$*; Sox2*$^{lox/lox}$ *; Nkx2.5*$^{cre/+}$ (*Nkx2.1*$^{tr/tr}$*; Sox2*$^{tr/tr}$) embryos lacking both *Nkx2.1* and *Sox2* in the ventral foregut/trachea cells **(Fig. 3.8a-c)**. We then examined NKX2.1-regulated genes

that we determined in our scRNA-seq analysis to be enriched in either the E11.5 trachea or the esophagus. The expression of the novel tracheal-specific, NKX2.1-dependent genes *Pcdh10*, and *Tppp3* was lost in the epithelium of *Nkx2.1*$^{tr/tr}$ foreguts, as expected from our RNA-seq analysis **(Fig. 3.8d,e,g,h)**. This loss of expression persisted in the epithelium of *Nkx2.1*$^{tr/tr}$*; Sox2*$^{tr/tr}$ foreguts, indicating that these genes are dependent on NKX2.1 and independent of SOX2 regulation **(Fig. 3.f,i)**. Conversely, expression of the esophageal-specific, NKX2.1-regulated genes *Klf5* and *Has2* was adopted throughout the epithelium in *Nkx2.1*$^{tr/tr}$ foreguts, as predicted from our RNA-seq analysis **(Fig. 3.8j,k,m,n)**. Interestingly, whereas *Klf5* expression was also increased in the ventral foregut of *Nkx2.1*$^{tr/tr}$*; Sox2*$^{tr/tr}$ mutants **(Fig. 3.8l)**, *Has2* expression was not, with the exception of a few cells that retain SOX2 expression ventrally (arrowhead in **Fig. 3.8o,o'**). Thus, while *Klf5* and *Has2* were both repressed by NKX2.1, upregulation of *Has2* in the ventral foregut of *Nkx2.1*$^{tr/tr}$ embryos appears to also depend on the upregulation of SOX2 in this region. Together, these findings revealed that NKX2.1 regulation results from both direct activation of tracheal genes and repression of esophageal genes, as well as the indirect suppression of target genes through the repression of *Sox2*, illustrating the complex relationship between these two transcription factors with opposing expression patterns.

Nkx2.1 regulation of foregut morphogenesis

The mutant embryos generated using *Nkx2.5-cre* for the previous analyses provided us with a unique opportunity to visualize morphogenetic changes under NKX2.1 and SOX2 regulation. The recombination pattern of *Nkx2.5-cre* encompasses the ventral

foregut and overlaps with the NKX2.1 expression pattern, thus providing a ventral deletion of NKX2.1. However, the *Nkx2.5-cre* recombination domain does not fully encompass the NKX2.1 expression domain. NKX2.1 expression extends slightly further dorsal than *Nkx2.5-cre* recombination, resulting in a small domain which retains NKX2.1 expression in these *Nkx2.1*$^{tr/tr}$ mutants **(Fig. 3.3a)**. Interestingly, these *Nkx2.1* $^{tr/tr}$ mosaic mutants exhibit ectopic evagination of the regions that retain NKX2.1 expression **(Fig. 3.9a,b)**. We observed that some of these regions develop into fully lumenized tubes that protrude from the mutant foregut (see Example 1, **Fig. 3.9b**). We have observed a similar phenotype in a separate project tin the laboratory that has used a *Foxa2-creER* to create mosaic deletions of *Nkx2.1*. These mutant foreguts exhibit a similar, yet more random, phenotype of ectopic NKX2.1+ tubes (data not shown). This phenotype corroborates the morphogenetic gene ontology associated with NKX2.1 ChIP-seq peaks, NKX2.1 regulation of *Efnb2* during foregut morphogenesis, as well as the NKX2.1-regulated morphogenetic genes identified in our Nkx2.1 RNA-seq experiments. Together, these data support the hypothesis that NKX2.1 regulates a morphogenetic program to instruct tracheoesophageal separation.

Interestingly, the ectopic evaginations observed in *Nkx2.1*$^{tr/tr}$ mosaic mutants appears to be abrogated in *Nkx2.1*$^{tr/tr}$*; Sox2*$^{tr/tr}$ mutants **(Fig. 3.9c)**. While *Nkx2.1*$^{tr/tr}$*; Sox2*$^{tr/tr}$ mutants exhibit patches of NKX2.1-positive cells and SOX2-positive cells, the morphology of these regions is continuous with the rest of the NKX2.1; SOX2-negative epithelium **(Fig. 3.9c)**. While these results indicate that SOX2 expression is required for NKX2.1-dependent morphogenesis, we are unable to determine whether this requirement is within or adjacent to the NKX2.1-positive cells. Further studies manipulating SOX2 and

NKX2.1 expression patterns within the foregut epithelium are necessary to determine the cellular-specific requirement of SOX2 in foregut morphogenesis.

Nkx2.1 regulation of epithelial-mesenchymal signaling.

The idea of a tracheoesophageal fate conversion in *Nkx2.1*$^{-/-}$ mutants is supported by the transformation of ventral mesenchymal cell fates from tracheal cartilage to smooth muscle (Minoo et al., 1999a; Que et al., 2007b; Yuan et al., 2000). However, previous reports describing the mesenchymal phenotypes in *Nkx2.1*$^{-/-}$ mutants have been limited and conflicting. One report shows that tracheal cartilage is present but reduced and disorganized in *Nkx2.1*$^{-/-}$ mutant foreguts (Minoo et al., 1999b). Another study suggests that tracheal cartilage is completely lost in *Nkx2.1*$^{-/-}$ mutants and is replaced by esophageal smooth muscle surrounding the common foregut tube (Que et al., 2007b). To understand the mechanism behind the mesenchymal phenotypes in *Nkx2.1*$^{-/-}$ mutants, we first needed a more thorough characterization of this phenotype. We therefore performed skeletal preparations on multiple E18.5 WT and *Nkx2.1*$^{-/-}$ mutant embryos to observe the cartilage defects along the rostral-caudal axis of the foregut **(Fig. 3.10a)**, as well as immunofluorescent staining at E13.5 to visualize cartilage and smooth muscle differentiation **(Fig. 3.10b)**. We observed a dramatic reduction and disorganization of tracheal cartilage with expansion of smooth muscle in *Nkx2.1*$^{-/-}$ mutants, with greater loss of cartilage at the caudal end of the trachea, and malformation of the thyroid and cricoid cartilages at the rostral end of the foregut **(Fig. 3.10)**.

These mesenchymal defects in *Nkx2.1*$^{-/-}$ mutants indicate that NKX2.1 regulates epithelial-mesenchymal signaling to coordinate fate decisions between these two tissues, since NKX2.1 expression is restricted to the epithelium. However, how this occurs is

currently unknown. We explored our *Nkx2.1*$^{-/-}$ mutant RNA-seq dataset and our single cell datasets to identify potential candidate mediators of epithelial-mesenchymal signaling under NKX2.1 regulation. Notably, the signaling genes *Wnt7b* and *Shh* were among the NKX2.1-regulated genes identified in our RNA-seq data; *Wnt7b* decreased and *Shh* increased in *Nkx2.1*$^{-/-}$ foregut epithelium compared to WT **(Fig. 3.1b, 3.11a)**. In WT E11.5 and E13.5 embryos, *Wnt7b* is expressed in the tracheal epithelium and *Shh* is expressed more strongly in the esophageal epithelium (**Fig. 3.11a**, Gerhardt et al., 2018; Litingtung et al., 1998; Rajagopal et al., 2008; Snowball et al., 2015b). Both signaling molecules are important in foregut mesenchymal differentiation; *Wnt7b* mutants exhibit disrupted cartilage formation in the ventral tracheal mesenchyme (Gerhardt et al., 2018; Rajagopal et al., 2008; Snowball et al., 2015b), and *Shh* mutants exhibit a loss of smooth muscle, and disorganization of tracheal cartilages(Litingtung et al., 1998; Miller et al., 2004; Pepicelli et al., 1998; Sala et al., 2011).

Given our finding that *Wnt7b* and *Shh* are regulated by NKX2.1, we examined our ChIP-seq data to determine whether NKX2.1 might directly regulate *Wnt7b* and *Shh*. Our ChIP-seq data identified multiple NKX2.1 ChIP-seq peaks near the *Wnt7b* gene, though these did not appear to include the *Wnt7b* promoter **(Fig. 3.11b)**. Binding of NKX2.1 to the promoter and gene body of the *Shh* gene was observed as well, consistent with the possibility of direct suppression by NKX2.1 **(Fig. 3.11b)**. Further functional analysis will help to determine which of these NKX2.1 binding sites are biologically relevant for regulation of *Wnt7b* and *Shh* expression in the foregut.

To determine whether NKX2.1 regulation of *Wnt7b* and *Shh* is impacted by changes in ventral SOX2 expression, we examined *Nkx2.1*; *Sox2* compound mutants.

Wnt7b expression was lost in both *Nkx2.1*$^{tr/tr}$ and *Nkx2.1*$^{tr/tr}$*; Sox2*$^{tr/tr}$ foreguts, suggesting that NKX2.1 regulates *Wnt7b* independently of SOX2 regulation **(Fig. 3.12a-d)**. However, the gain of *Shh* in the ventral foregut of *Nkx2.1*$^{tr/tr}$ embryos was lost in *Nkx2.1*$^{tr/tr}$*; Sox2*$^{tr/tr}$ foreguts **(Fig. 3.12e-h)**, suggesting that SOX2 is required for activation of *Shh*. By utilizing the mosaic nature of these mutants, we can also see that while *Shh* expression requires SOX2, NKX2.1 still represses *Shh*, even in the presence of SOX2 (arrowheads in **Fig. 3.12g**). Thus, this is more likely an example of dual regulation of *Shh* by NKX2.1 and SOX2, as opposed to SOX2 serving as a mediator of NKX2.1 indirect regulation.

To determine whether NKX2.1 regulation of *Wnt7b* has effects on WNT signaling from the epithelium to the mesenchyme, we examined the expression of *Wnt7b* and downstream canonical WNT signaling transcriptional target *Axin2* in E11.5 and E13.5 *Nkx2.1*$^{-/-}$ and WT foreguts **(Fig. 3.13)**. In WT foreguts, *Wnt7b* expression was restricted to the tracheal epithelium, and *Axin2* expression was highest in the ventral mesenchyme where SOX9+ cartilage progenitor cells are found **(3.13a,c,d)**. We also observed *Axin2* expression in the esophageal epithelium of E11.5 WT foreguts (left panel of **Fig. 3.13a**). In *Nkx2.1*$^{-/-}$ embryos, we confirmed a decrease in *Wnt7b* throughout the ventral epithelium of the common foregut tube compared to WT trachea at E11.5 and E13.5 (right panels of **Fig. 3.13a,c**). Furthermore, *Axin2* expression was decreased in *Nkx2.1*$^{-/-}$ mutants and correlated with disorganized Sox9 expression which did not extend as far dorsally as in WT trachea **(Fig. 3.13a,c,d)**. The decrease in *Axin2* expression in the mesenchyme of E13.5 *Nkx2.1*$^{-/-}$ foreguts was more dramatic in the caudal foregut, consistent with the greater reduction of cartilage specification in this region **(Fig. 3.10a)** and as visualized by SOX9 localization *Nkx2.1*$^{-/-}$ foreguts **(Fig. 3.13d)**. Thus, the loss of *Wnt7b* expression

observed in *Nkx2.1*$^{-/-}$ foreguts correlates with a decrease in WNT signaling in the ventral foregut mesenchyme and defects in cartilage specification in these mutants.

We next examined the expression of *Shh* and downstream transcriptional target *Ptch1* at E11.5 and E13.5 to determine whether Nkx2.1 regulation of *Shh* had effects on hedgehog signaling during mesenchymal specification. In WT foreguts, we observed *Shh* expression in the esophageal epithelium, and *Ptch1* expression in the surrounding esophageal mesenchyme, colocalizing with developing smooth muscle. We visualized smooth muscle specification with the expression of SMA **(Fig. 3.13f)**. In *Nkx2.1*$^{-/-}$ foreguts, increased expression of *Shh* in the ventral foregut epithelium mirrored an increase in *Ptch1* expression in the surrounding mesenchyme **(Fig. 3.13a,e)**. This increase in *Ptch1* expression also corresponded with the increase in smooth muscle surrounding *Nkx2.1*$^{-/-}$ foreguts compared to WT **(Fig. 3.13e,f)**. Together, these findings show that, in addition to regulating a subset of genes that characterize tracheoesophageal epithelial identity, NKX2.1 and SOX2 regulate epithelial-mesenchymal crosstalk via WNT and SHH signaling to instruct mesenchymal differentiation in the developing foregut.

Conclusions and discussion

Aside from gross phenotypic characterization and analysis of few known tracheal and esophageal markers (Minoo et al., 1999; Que et al., 2007), the regulatory role of NKX2.1 in the developing foregut has remained largely unexplored. In this chapter, we define the NKX2.1 regulatory program at the genome-wide scale with RNA-sequencing of *Nkx2.1*$^{-/-}$ foreguts and NKX2.1 ChIP-seq. Additionally, we uncover a substantial NKX2.1-independent program of tracheoesophageal specification. Within the NKX2.1

transcriptional program, we identify examples of SOX2-dependent and -independent regulation, exposing the complex relationship between these two transcription factors. Lastly, we show that NKX2.1 influences mesenchymal specification through direct binding to and transcriptional regulation of *Wnt7b* and *Shh*, which has downstream effects on mesenchymal WNT and SHH signaling, leading to defects in foregut cartilage and smooth muscle formation. Together, these findings dramatically broaden our understanding of the earliest stages of tracheoesophageal development, while revising and advancing our perspective on the roles of NKX2.1 in this process **(Fig. 3.14)**.

This NKX2.1 transcriptional program included transcription factors, signaling molecules, and morphogenetic factors. Further dissection of this transcriptional program may uncouple key processes in foregut development such as separation morphogenesis, growth and elongation, signaling and innervation, establishment of tissue identity, and establishment of cell-type specific identity. For example, *Msn, Mical2*, and *Tppp3* are all positively regulated by NKX2.1, have known roles in regulating the cytoskeleton, epithelial integrity, and signal transduction, and are exciting candidates for the study of foregut morphogenesis, epithelial maintenance, and signal transduction (Giridharan et al., 2012; Neisch and Fehon, 2011; Speck et al., 2003; Torisawa et al., 2019; Zhou et al., 2010). Further, the intestinal stem cell marker *Lgr5* (Barker et al., 2007) was negatively regulated by NKX2.1, and the tracheal ciliated cell gene *Ccno* (Funk et al., 2015) was positively regulated by NKX2.1 suggesting a role for NKX2.1 in regulating cell type specific identity. Continued exploration of these NKX2.1-regulated genes with known or potentially exciting unknown functions in respiratory and gut development will provide further clarity into the distinct roles of NKX2.1 in foregut development.

While the mutually repressive relationship between *Sox2* and *Nkx2.1* has been described previously, the direct or indirect nature of this relationship had not been explored. Our finding of NKX2.1 binding to the *Sox2* locus in the trachea is the first evidence supporting direct regulation of *Sox2* expression by NKX2.1. It is important to note that in WT trachea, SOX2 and NKX2.1 are co-expressed, and SOX2 has important roles in tracheal growth and development (Que et al., 2007b, 2009). Thus, it is likely that the regulatory relationship between NKX2.1 and SOX2 cannot be explained by simple mutual repression. Further exploration of the regulation of *Sox2* in the trachea will likely expose factors that influence *Sox2* expression beyond repression by NKX2.1.

Using a candidate approach, we examine the regulatory relationship between NKX2.1 and SOX2 during foregut development and identify examples of both SOX2-dependent and -independent NKX2.1 regulation. Interestingly, of the NKX2.1-regulated genes we examined, we identified SOX2-dependence only amongst esophageal genes that were gained in *Nkx2.1*$^{-/-}$ mutants, and not tracheal genes that were lost in *Nkx2.1*$^{-/-}$ mutants. This strengthens the hypothesis that NKX2.1 directly activates tracheal genes, and more often represses esophageal genes indirectly. Additionally, our identification of other transcription factors regulated by NKX2.1 such as KLF5 provides candidates for further exploration of NKX2.1 indirect repression of esophageal genes. The candidate genes we examined further indicate that NKX2.1 and SOX2 regulation in foregut development is more nuanced than the binary, co-repressive switch that has been previously suggested (Billmyre et al., 2015; Que et al., 2007b). Future transcriptomic analyses of the transcriptional program downstream of SOX2 in the developing trachea

and esophagus will provide exciting avenues to further understand the regulatory roles of both SOX2 and NKX2.1.

The apparent fate conversion of the *Nkx2.1*$^{-/-}$ mutant foregut to esophagus is supported by the transformation of ventral mesenchymal cell fates from tracheal cartilage to smooth muscle (Minoo et al., 1999a; Que et al., 2007a; Yuan et al., 2000). While our identification of an NKX2.1-independent program challenges the fate conversion model, we suggest that the mesenchymal transformation observed in *Nkx2.1*$^{-/-}$ foreguts occurs via a loss of NKX2.1 regulation of epithelial-mesenchymal WNT and SHH signaling. This hypothesis is supported by findings that WNT signaling from the tracheal epithelium to the mesenchyme is required for the formation of tracheal cartilage (Gerhardt et al., 2018; Hou et al., 2019; Kishimoto et al., 2019; Snowball et al., 2015a), and hedgehog signaling is a key regulator of smooth muscle development across contexts (Huycke et al., 2019a; Mao et al., 2010). A recent report suggested that epithelial-mesenchymal WNT signaling during cartilage specification occurs through NKX2.1-independent mechanisms (Kishimoto et al., 2019). Our results differ from this finding but do support the possibility that an NKX2.1 independent mechanism also contributes to tracheal cartilage specification, as some disorganized cartilage still forms in *Nkx2.1*$^{-/-}$ mutants. It will be important to leverage our RNA-seq datasets to identify additional regulators of epithelial-mesenchymal signaling during foregut mesenchyme specification.

At the transcriptome level, loss of *Nkx2.1* does not result in a tracheal-to-esophageal fate conversion as previously thought, but rather results in changes in only a subset of dorsoventrally restricted genes. Our discovery of an NKX2.1-independent tracheal transcriptional program indicates that NKX2.1 is not the sole master regulator of

tracheal fate and indicates the presence of other transcriptional regulators of tracheal fate specification. The function of the larger NKX2.1-independent program, and how it is established, remains to be explored. As loss of canonical WNT signaling from the early splanchnic mesoderm results in a loss of NKX2.1 expression and a complete loss of lung bud outgrowth (Goss et al., 2009; Harris-Johnson et al., 2009), it will be exciting to learn whether Wnt/β-catenin signaling directly induces a wider tracheal transcriptional program. Interestingly, the ISL1 transcription factor has been recently identified to exhibit ventrally-restricted expression in the trachea and is required for normal NKX2.1 expression (Kim et al., 2019). Further, our data reveal that *Isl1* is independent of NKX2.1, indicating that it resides hierarchically upstream of NKX2.1. It remains to be determined; however, the extent to which ISL1 regulates tracheal transcriptional identity. Several IRX transcription factors are also highly expressed in tracheal cells, independently of NKX2.1. These transcription factors regulate a variety of other developmental processes including tissue specification and morphogenesis (Kim et al., 2012; van Tuyl et al., 2006) and could be interesting candidates for future study of foregut specification. The esophageal-specific transcription factor PITX1 is also NKX2.1-independent and is an exciting candidate for future exploration of esophageal specification.

It is important to note that our findings are specific to the trachea and esophagus, and do not capture NKX2.1 regulation of lung development. In fact, others have shown that Nkx2.1 regulates key genes in the distal lung (Little et al., 2019) and in lung adenocarcinoma (Snyder et al., 2013) that loss of NKX2.1 results in a more thorough adoption of gastrointestinal features in these cells, consistent with distinct roles during early trachea and lung specification.

Figures

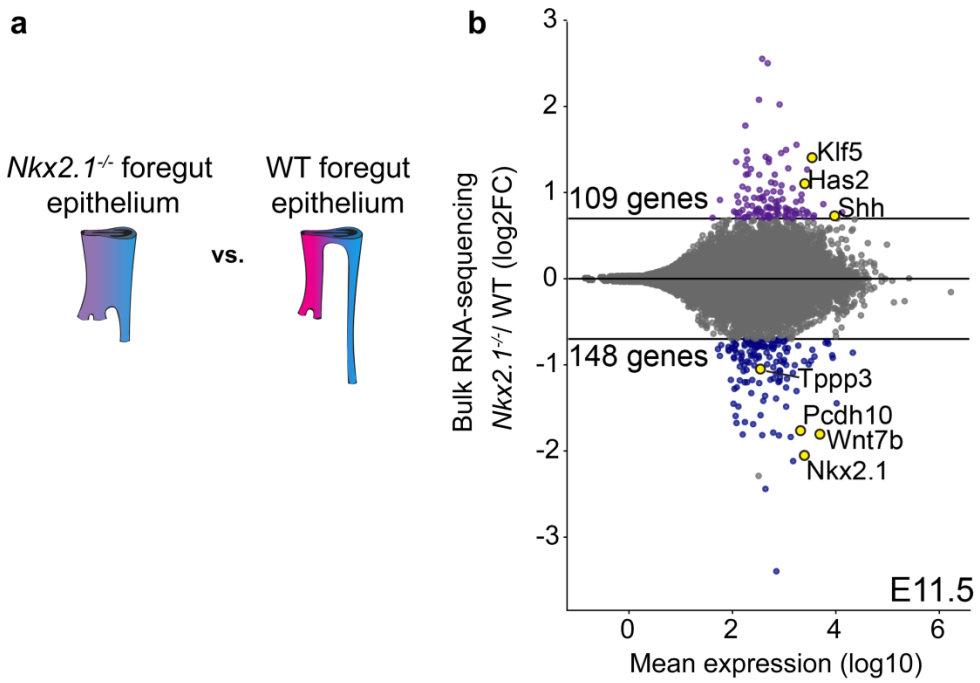


Figure 3.1. Identification of the NKX2.1 transcriptional program. a. Dissected, FACS-purified epithelium from E11.5 *Nkx2.1*$^{-/-}$ and WT foreguts was sequenced and analyzed for differential gene expression. **b.** 109 genes were increased and 148 genes were decreased in *Nkx2.1*$^{-/-}$ foreguts compared to WT (DESeq2, log2FC>0.7, padj<0.05). Labeled genes show examples of tracheal genes that decrease in *Nkx2.1*$^{-/-}$ mutants (*Tppp3, Pcdh10, Wnt7b, Nkx2.1*), and esophageal genes that increase in *Nkx2.1*$^{-/-}$ mutants (*Klf5, Has2, Shh*).

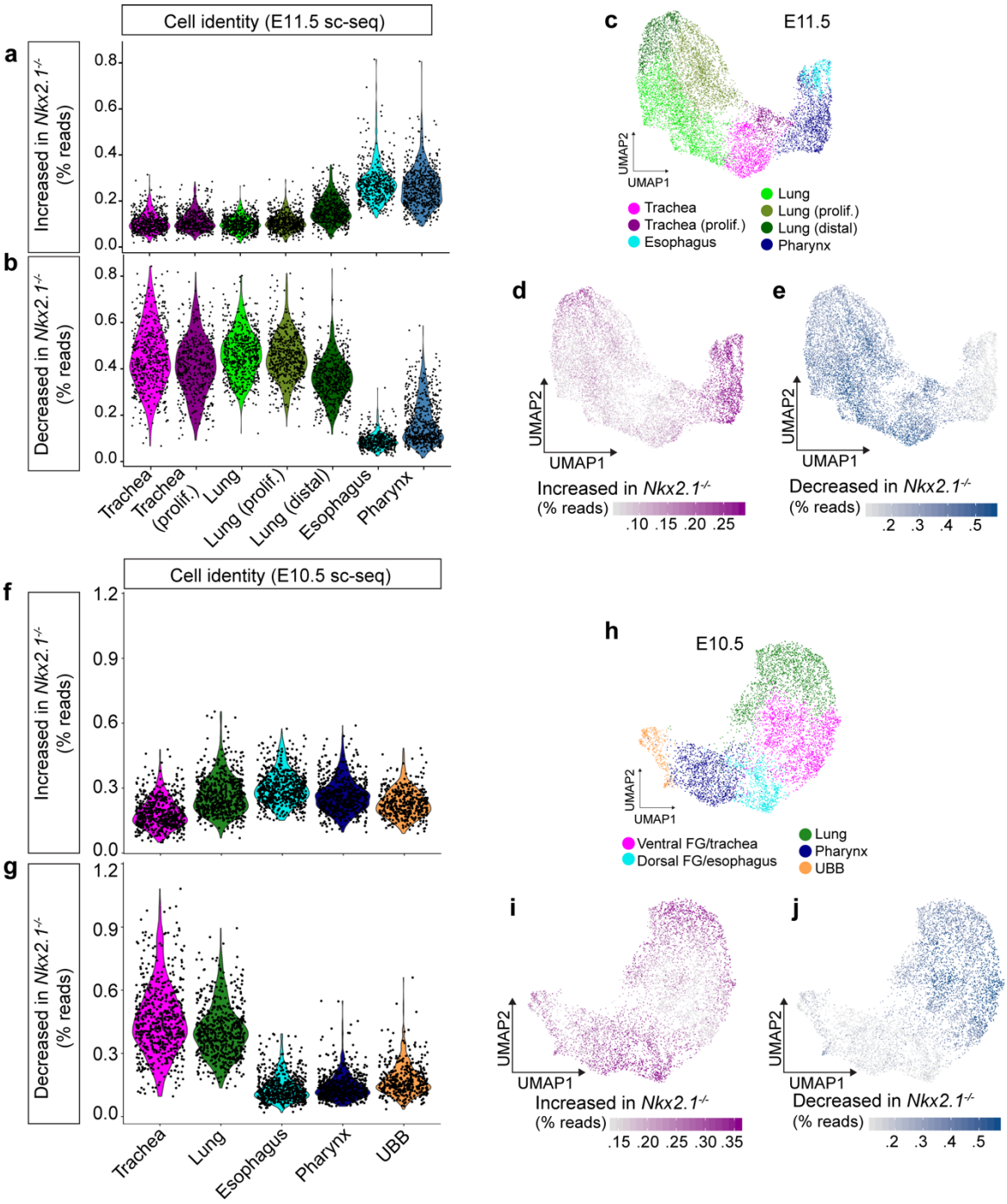


Figure 3.2. NKX2.1 positively regulates tracheal genes and negatively regulates esophageal genes. a-b. Violin plot depicting combined expression of all genes that increase (**a**) or decrease (**b**) in *Nkx2.1*$^{-/-}$ foreguts as a percentage of all reads in E11.5 scRNA-seq data. Cells are grouped by their assigned cluster and clusters were subsampled to 600 cells/cluster for visualization. **c-e.** Combined expression of all genes that increase (**d**) or decrease (**e**) in *Nkx2.1*$^{-/-}$ foreguts as a percentage of all reads in E11.5 scRNA-seq data, projected onto the E11.5 UMAP. Cluster identities depicted in **c. f-j.** Same as above, for E10.5 scRNA-seq data.

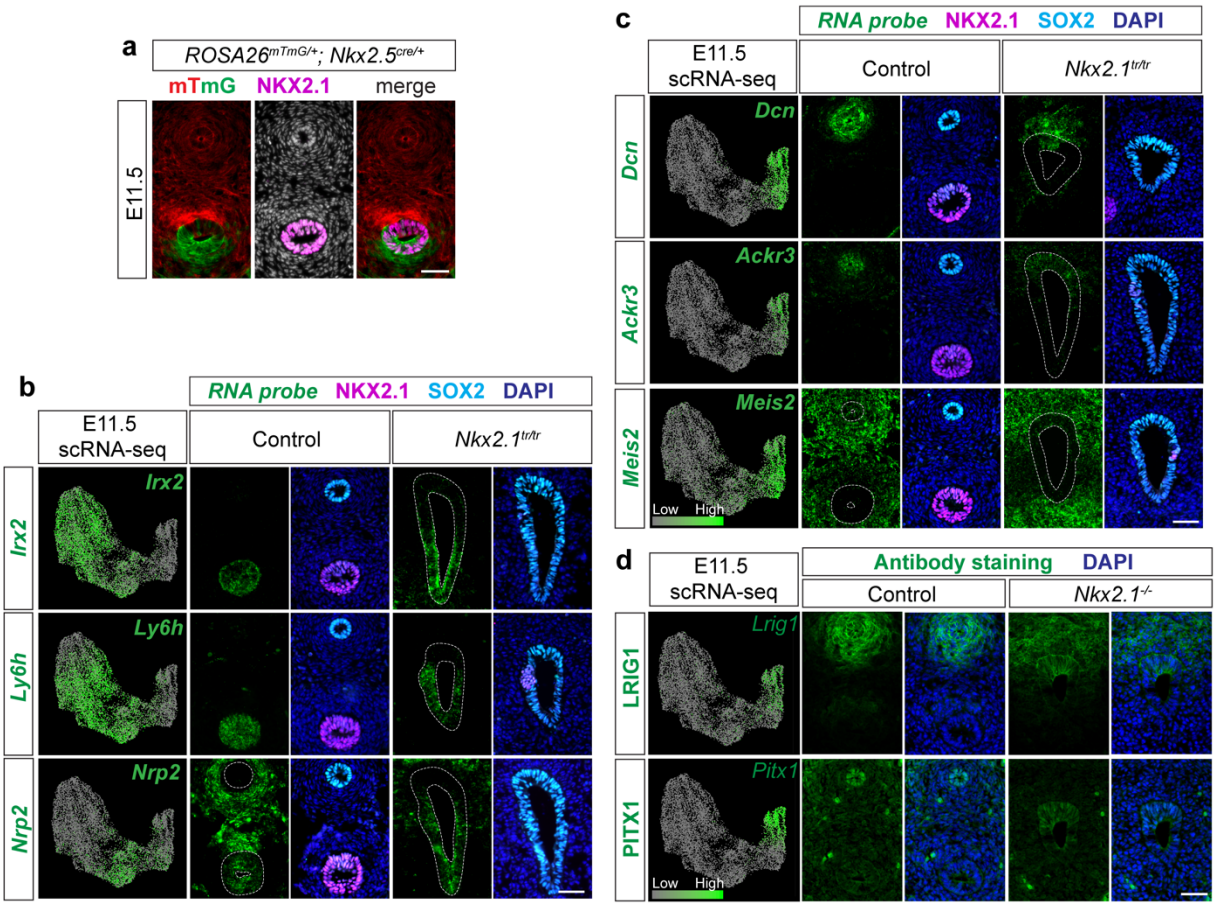


Figure 3.3. Identification of NKX2.1-independent tracheal and esophageal genes. a. Lineage trace of *Nkx2.5-cre* recombination pattern in E11.5 foreguts using the *ROSA26*^{mTmG} reporter (red/green) and NKX2.1 immunofluorescent staining (magenta). **b.** RNA localization of NKX2.1-independent tracheal genes *Irx2, Ly6h*, and *Nrp2* identified by scRNA-seq and *Nkx2.1*$^{-/-}$ mutant RNA-seq data at E11.5. First column shows projection of RNA expression level as determined by scRNA-seq on UMAP. Second-fifth columns show RNA localization (green) in E11.5 control and *Nkx2.1*$^{tr/tr}$ embryos with immunofluorescent staining of NKX2.1 (magenta) and SOX2 (cyan). **c.** RNA-localization of NKX2.1-independent esophageal genes *Dcn, Ackr3* and *Meis2.* **d.** Protein expression of NKX2.1-independent genes LRIG1 and PITX1 in E11.5 control and *Nkx2.1*$^{-/-}$ embryos. Scale bar = 50um.

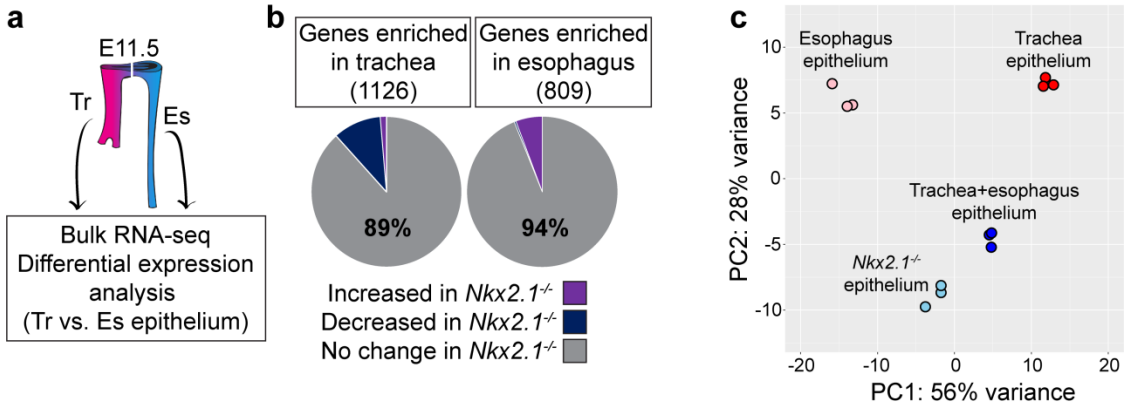


Figure 3.4. Genomic analysis of the NKX2.1-independent transcriptional program. **a.** Schematic of experimental procedures for RNA-seq and differential expression analysis of WT trachea and esophagus for comparison with *Nkx2.1*$^{-/-}$ mutant RNA-seq data. **b**. NKX2.1-dependent genes as a portion of genes enriched in WT trachea (left) and esophagus (right). NKX2.1-independent genes make up 89% of tracheal-enriched genes and 94% of esophageal enriched genes. **c.** Principal component analysis of all bulk RNA-seq samples (esophagus epithelium, trachea epithelium, *Nkx2.1*$^{-/-}$ epithelium, and WT trachea+esophagus epithelium). Principal component 1 (PC1) depicted on the x-axis captures 56% of variance between samples, and principal component 2 (PC2) captures 28% of variance between samples.

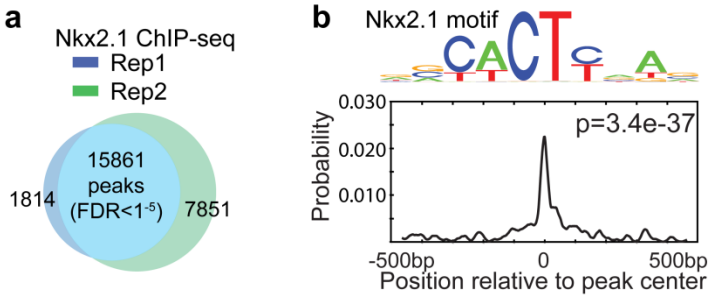


Figure 3.5. Identification of NKX2.1 DNA-binding regions in the trachea. a. ChIP-seq for NKX2.1 in E11.5 foreguts identified 15,861 NKX2.1-bound genomic regions (peaks) shared between two biological replicates (FDR<0.00001). **b.** Motif analysis of NKX2.1 ChIP-seq data shows peaks are enriched for the NKX2.1 motif (p=3.4e-37).

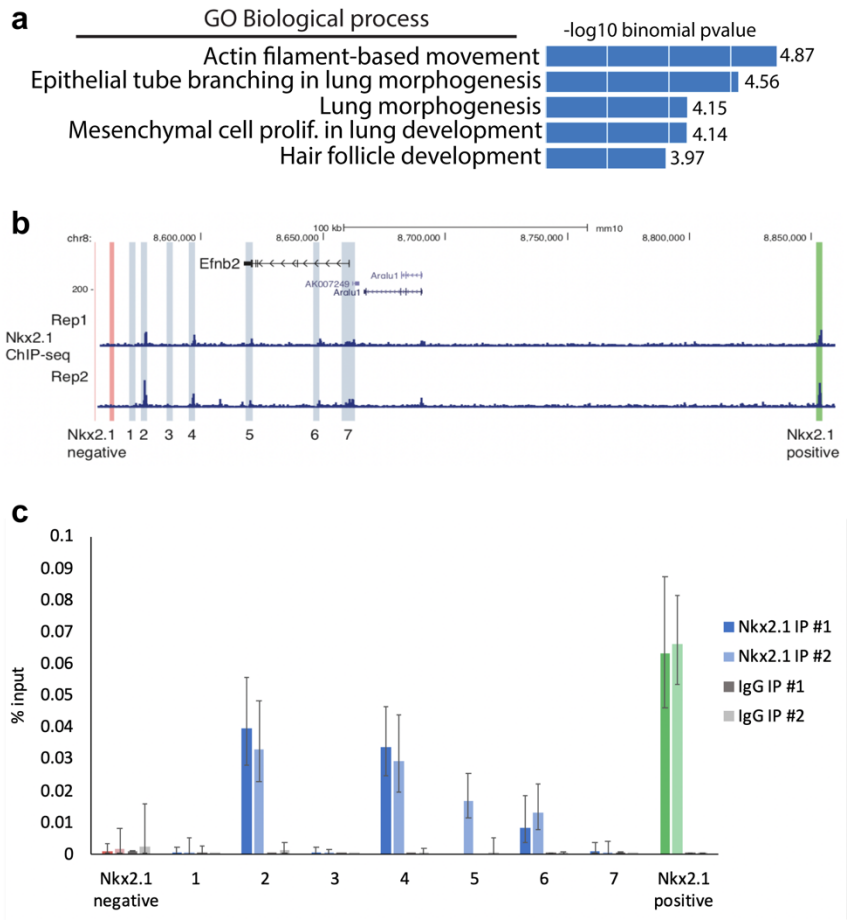


Figure 3.6. NKX2.1 binds near genes associated with morphogenesis, including *Efnb2.* **a.** Gene ontology analysis of top 1000 NKX2.1 ChIP-seq peaks by logFC over input shared between both replicates using GREAT (basal plus extension rule). Plot shows all biological processes of genes associated with NKX2.1 peaks in order of decreasing -log10 binomial p-value. **b.** NXK2.1 ChIP-seq tracks at *Efnb2* gene. Highlighted regions show location of NKX2.1 peaks (blue) and negative (red) and positive (green) genomic regions used for ChIP-PCR validation. **c.** ChIP-qPCR validation of NKX2.1 binding at selected NKX2.1 ChIP-peaks (blue) and positive (green) and negative (red) control regions. IgG control shown in gray.

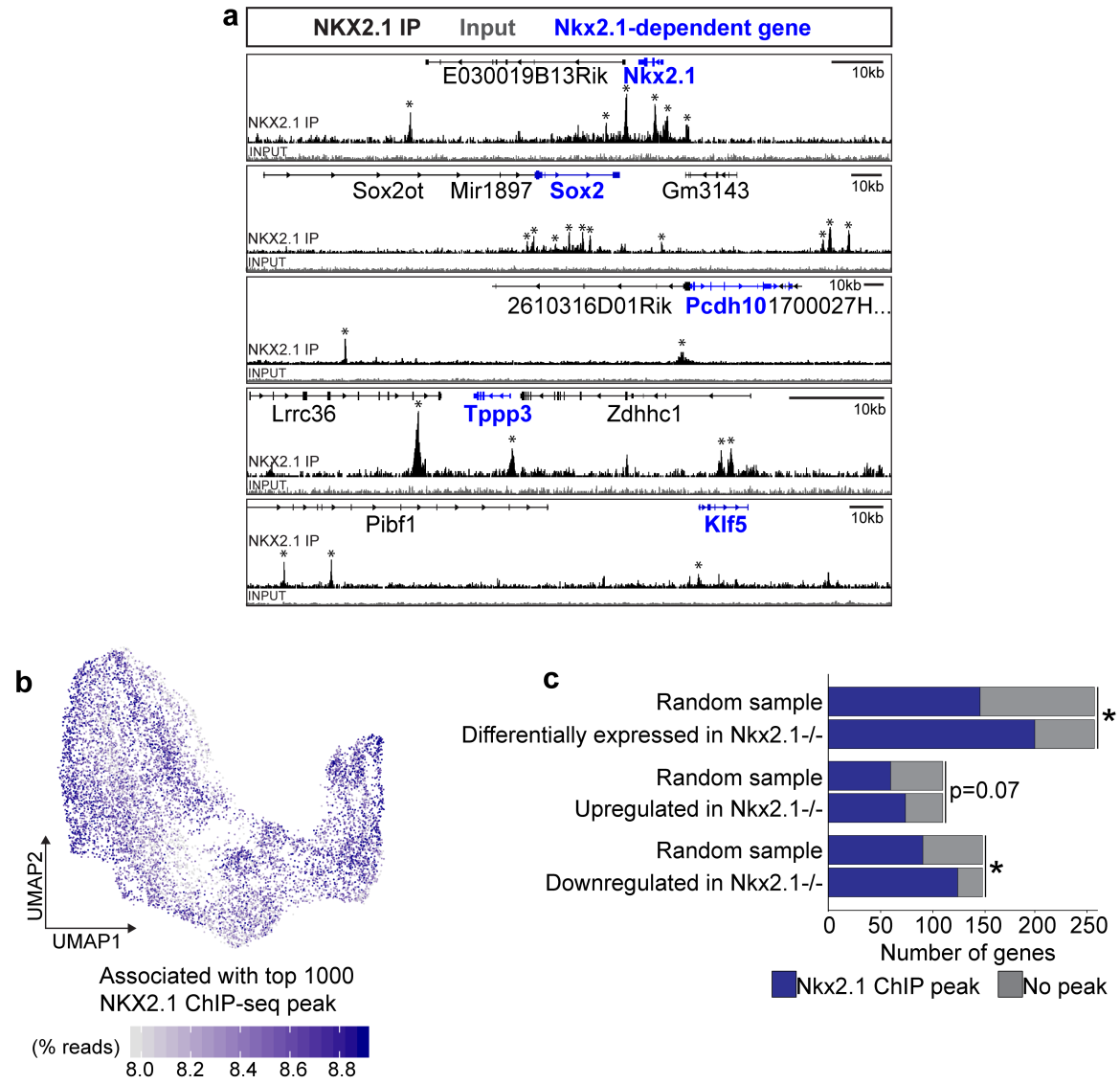


Figure 3.7. NKX2.1 directly binds near NKX2.1-regulated genes. a. NKX2.1 ChIP-seq (black) and input (grey) tracks near loci of select NKX2.1-regulated genes (blue) *Nkx2.1, Sox2, Pcdh10, Tppp3,* and *Klf5*. Scale bar = 10kb. **b.** Combined expression of all genes associated with the top 1000 NKX2.1 ChIP-seq peaks (by log2FC over input) as a percentage of all reads in E11.5 scRNA-seq data, projected onto the E11.5 UMAP. Peak-gene associations generated with GREAT basal-plus-extension rule. **c.** Genome-wide comparison of NKX2.1 ChIP-seq with *Nkx2.1*$^{-/-}$ mutant RNA-seq. Overlap of NKX2.1 ChIP-seq associated genes with all NKX2.1-regulated genes (top), genes increased in *Nkx2.1*$^{-/-}$ mutants (middle) and genes decreased in *Nkx2.1*$^{-/-}$ mutants (lower). *asterisk=p<0.0001, Fisher's exact test.

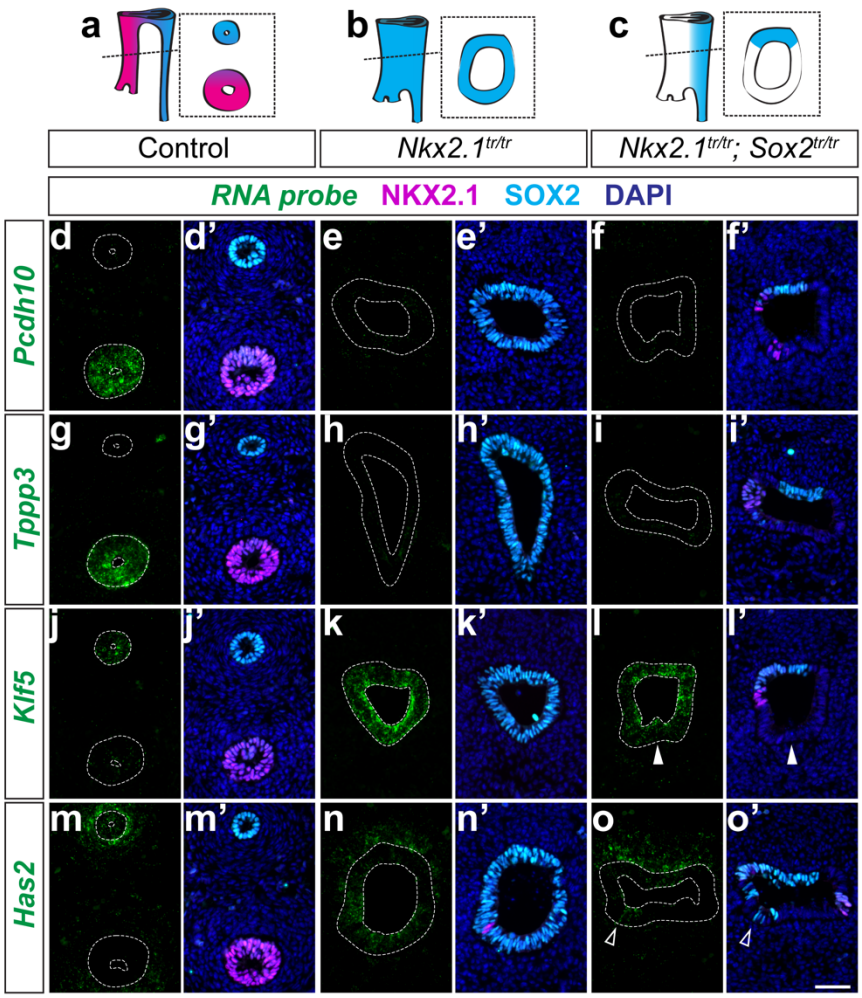


Figure 3.8. NKX2.1 regulates target genes in a SOX2-dependent and independent manner. a-c. Schematic of *Nkx2.1; Sox2* compound mutant analysis phenotypes and resulting NKX2.1 and SOX2 expression patterns in E11.5 control embryos (**a**), *Nkx2.1*$^{tr/tr}$ mutants (**b**), *Nkx2.1*$^{tr/tr}$*, Sox2*$^{tr/tr}$ mutants (**c**). **d-f.** RNA localization of NKX2.1-dependent, SOX2-independent gene *Pcdh10* in control (**d**), *Nkx2.1*$^{tr/tr}$ (**e**), and *Nkx2.1*$^{tr/tr}$*, Sox2*$^{tr/tr}$ (**f**) embryos with immunofluorescent staining of NKX2.1 (magenta) and SOX2 (cyan). **g-i.** Tracheal, NKX2.1-dependent, SOX2-independent gene *Tppp3*. **j-l.** Esophageal, NKX2.1-regulated, SOX2-independent gene *Klf5*. Solid arrowheads indicate SOX2-negative, *Klf5*-positive ventral cell. **m-o.** Esophageal, NKX2.1-regulated, SOX2-dependent gene *Has2*. Arrowheads indicate ventral SOX2-positive, *Has2*-positive cells. Scale = 50um.

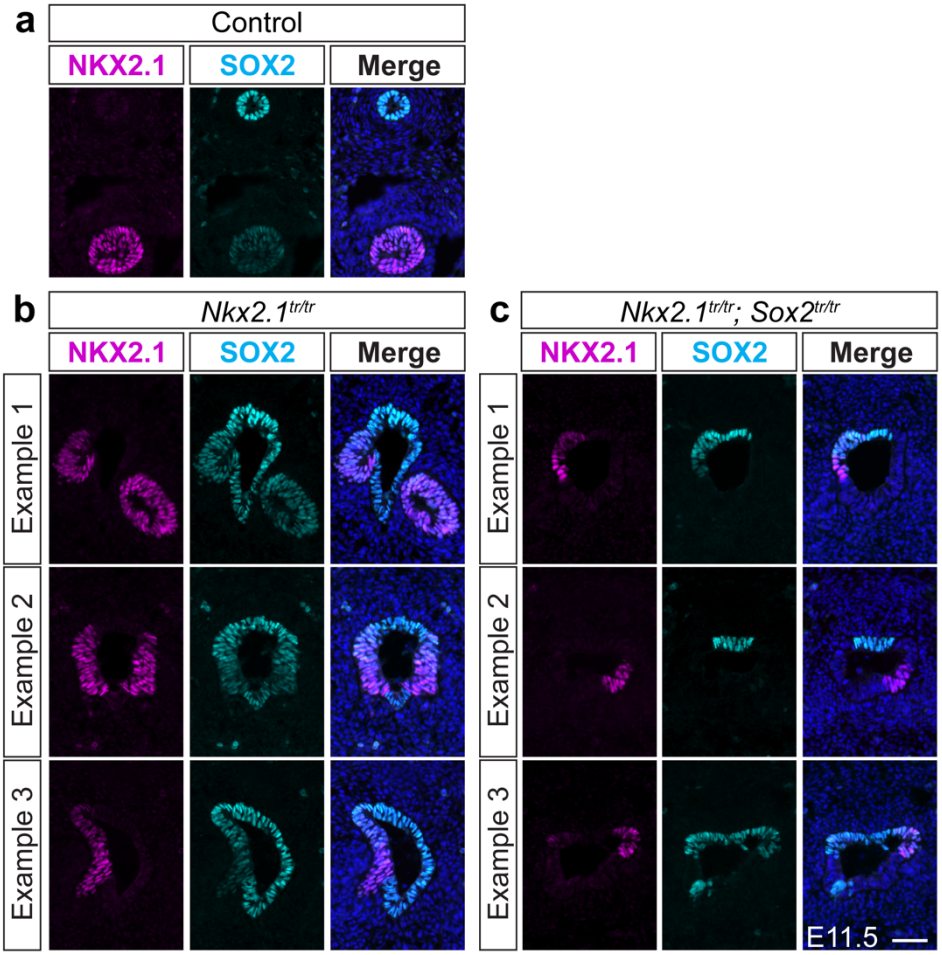


Figure 3.9. NKX2.1 mosaicism results in ectopic foregut evaginations. a. Immunofluorescent staining for NKX2.1 (magenta) and SOX2 (cyan) with DAPI (blue) in E11.5 control trachea and esophagus. **b.** Immunofluorescent staining for NKX2.1 (magenta) and SOX2 (cyan) with DAPI (blue) in E11.5 *Nkx2.1*$^{tr/tr}$ foreguts. Three examples shown. **c.** Immunofluorescent staining for NKX2.1 (magenta) and SOX2 (cyan) with DAPI (blue) in E11.5 *Nkx2.1*$^{tr/tr}$*; Sox2*$^{tr/tr}$ foreguts. Three examples shown. Scale bar=50um.

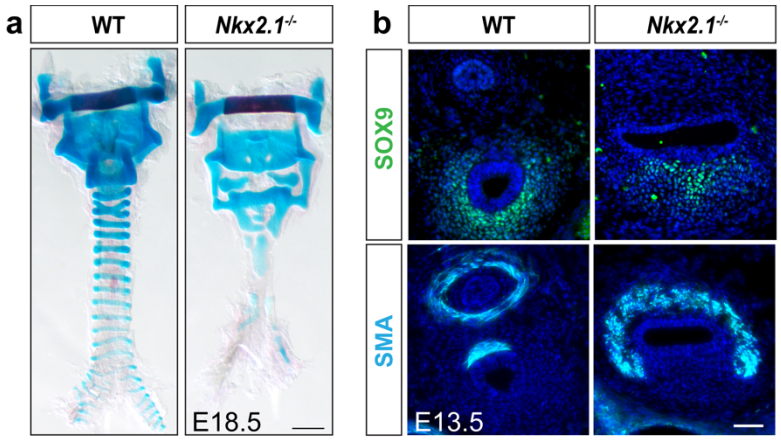


Figure 3.10. Mesenchymal phenotype in *Nkx2.1*$^{-/-}$ mutant foreguts. a. Alcian blue staining of tracheal cartilage in E18.5 control and *Nkx2.1*$^{-/-}$ embryos. **b.** SOX9 staining of cartilage progenitors and SMA staining of smooth muscle in E13.5 control and *Nkx2.1*$^{-/-}$ embryos.

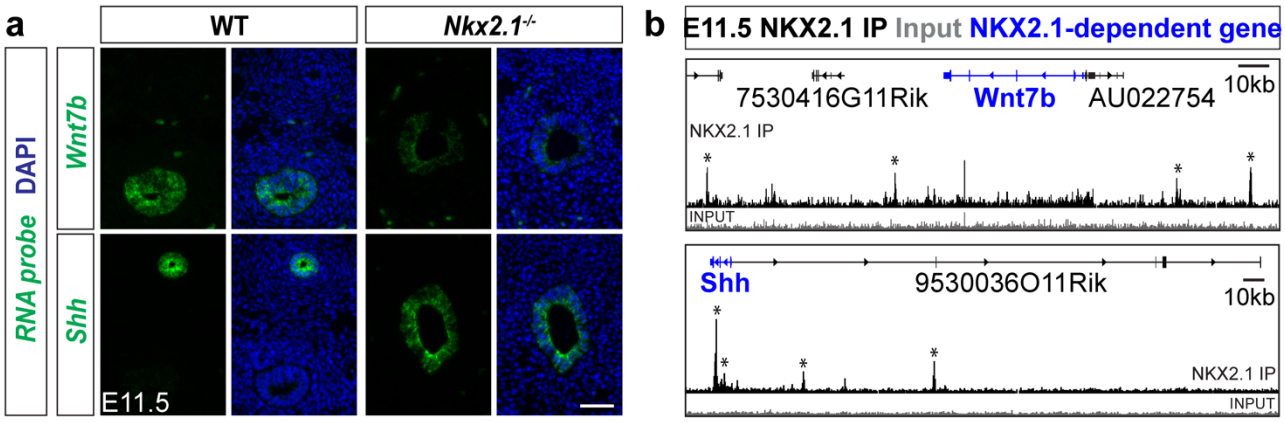


Figure 3.11 Wnt7b and Shh are direct targets of NKX2.1 in the foregut. a. RNA-localization of *Wnt7b* (top) and *Shh* (bottom) in WT and *Nkx2.1*$^{-/-}$ foreguts. Scale=50um **b.** ChIP-seq of direct NKX2.1 binding near *Wnt7b* and *Shh* loci. Scale = 10kb.

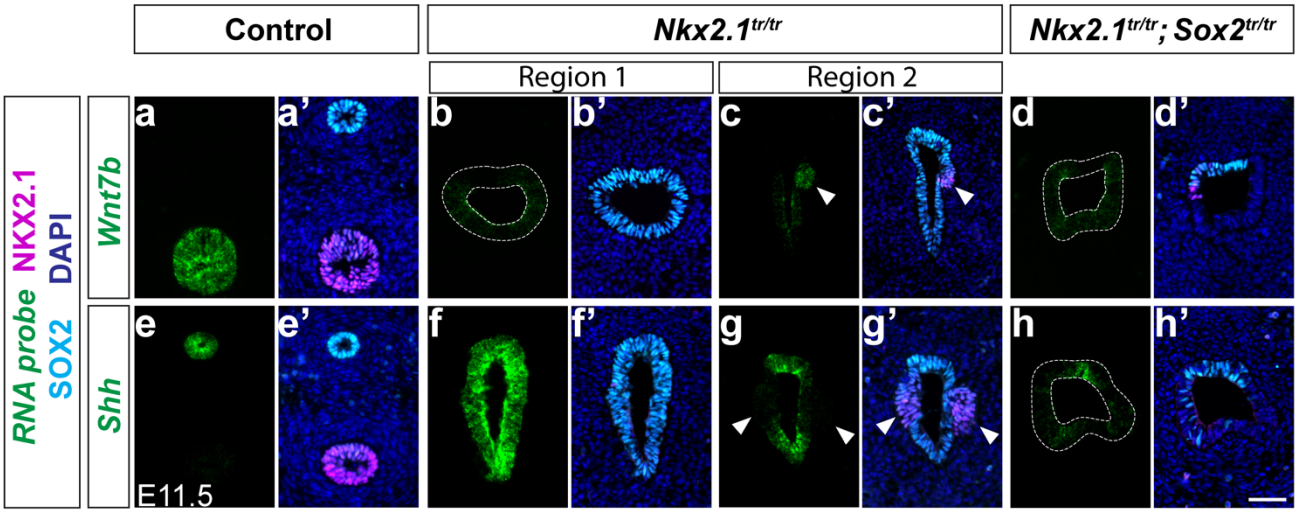


Figure 3.12. *Wnt7b* and *Shh* expression in *Nkx2.1;Sox2* compound mutants. a-d. RNA localization (green) of *Wnt7b* in control (a,a'), *Nkx2.1*$^{tr/tr}$ (b-c'), and *Nkx2.1*$^{tr/tr}$*, Sox2*$^{tr/tr}$ (d,d') E11.5 embryos with immunofluorescent staining of NKX2.1 (magenta) and SOX2 (cyan). Two regions of the common foregut tube in *Nkx2.1*$^{tr/tr}$ embryos are shown to illustrate mosaicism of NKX2.1 expression (arrowheads) and regulatory consequences. **e-h**. RNA localization (green) of *Shh* in control (e,e'), *Nkx2.1*$^{tr/tr}$ (f-g'), and *Nkx2.1*$^{tr/tr}$*, Sox2*$^{tr/tr}$ (h,h') E11.5 embryos with immunofluorescent staining of NKX2.1 (magenta) and SOX2 (cyan). Two regions of the common foregut tube in *Nkx2.1*$^{tr/tr}$ embryos are shown to illustrate mosaicism of NKX2.1 expression (arrowheads) and regulatory consequences.

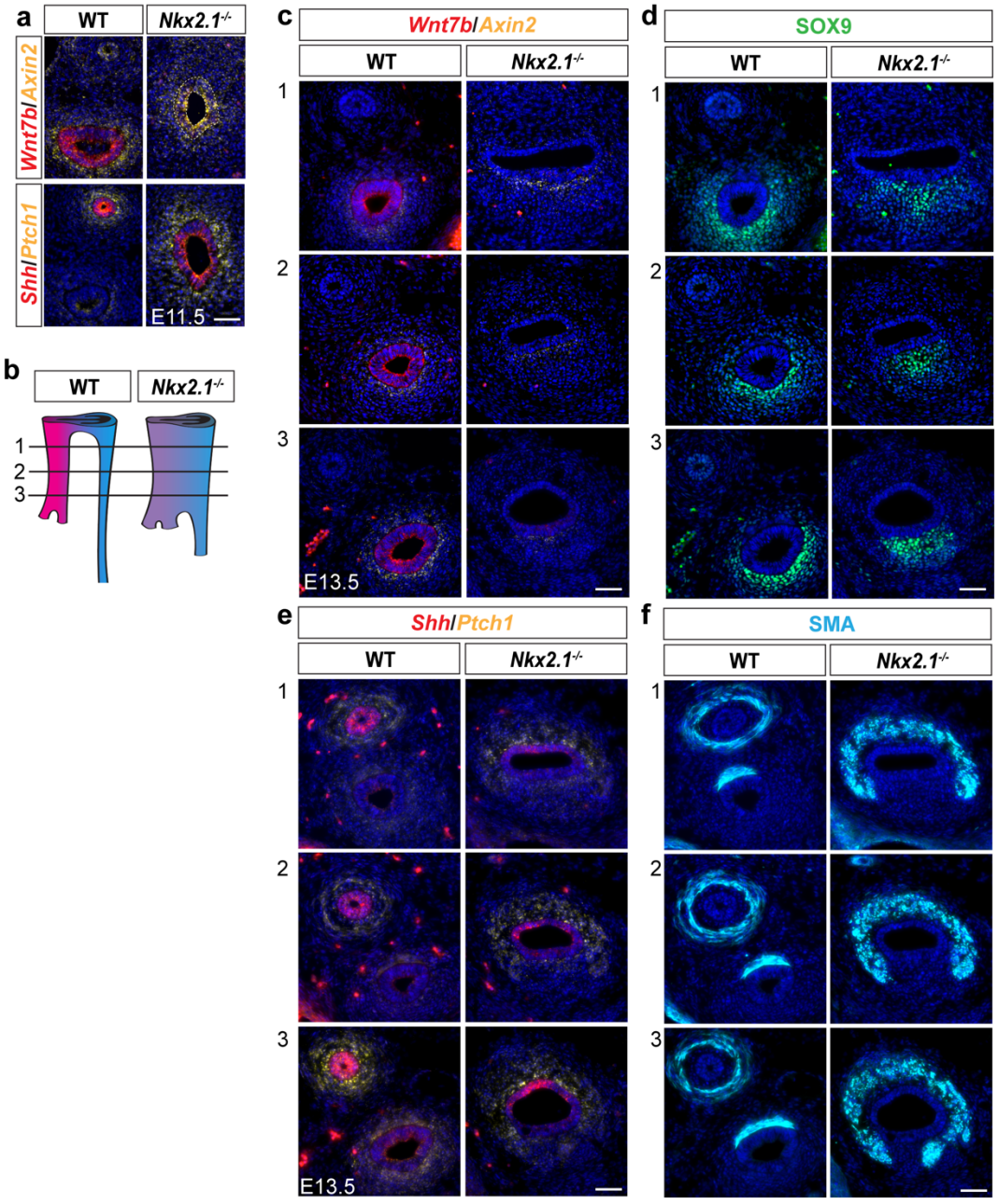


Figure 3.13. NKX2.1 regulates epithelial-mesenchymal WNT and SHH signaling. a. RNA localization of *Wnt7b* (red) and *Axin2* (yellow), top row, and *Shh* (red) and *Ptch1* (yellow), bottom row in WT and *Nkx2.1*$^{-/-}$ E11.5 foreguts. **b.** Schematic of rostral-caudal position of images in **c-f**. **c-f.** RNA localization of *Wnt7b* and *Axin2* (**c**), *Shh* and *Ptch1* (**e**), and protein localization of SOX9 (**d**), and smooth muscle actin (SMA) (**f**) along the rostral-caudal axis of E13.5 WT and *Nkx2.1*$^{-/-}$ foreguts.

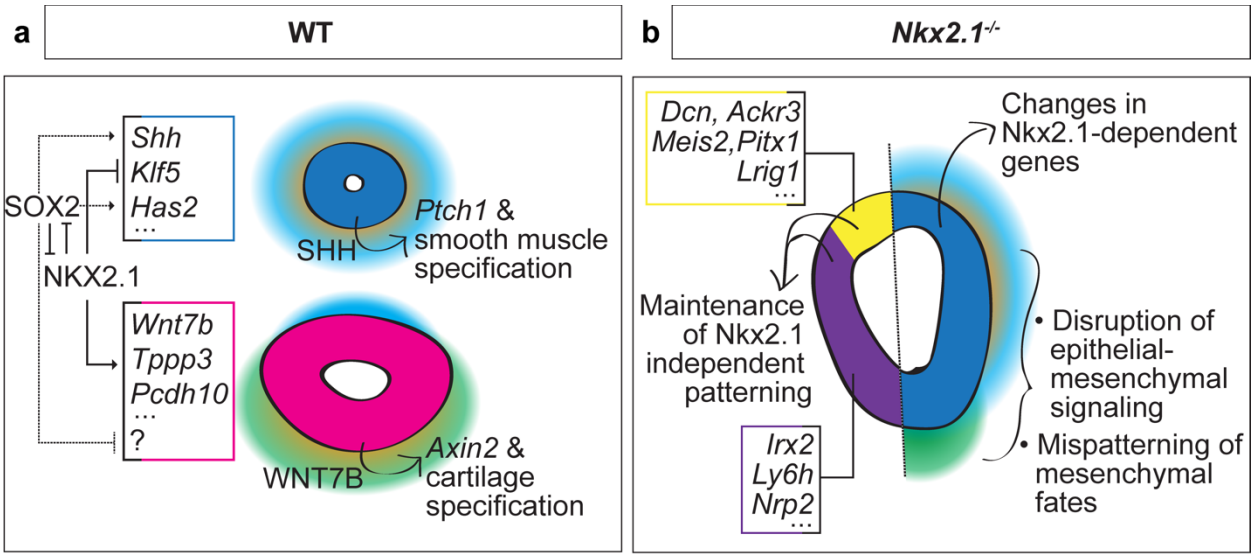


Figure 3.14 Proposed model of NKX2.1-dependent and -independent tracheoesophageal specification. a. Schematic of NKX2.1 and SOX2 regulation of gene expression in WT trachea and esophagus, and epithelial-mesenchymal signaling downstream of NKX2.1. NKX2.1 negatively regulates esophageal genes *Shh, Klf5* and *Has2,* and positively regulates tracheal genes *Wnt7b, Tppp3,* and *Pcdh10.* SOX2 is required for expression of *Shh* and *Has2*. NKX2.1 regulation of *Shh* and *Wnt7b* influences mesenchymal SHH and WNT response required for smooth muscle and tracheal cartilage development. b. Schematic of *Nkx2.1*$^{-/-}$ mutant phenotype. Maintenance of the NKX2.1-independent transcriptional program includes dorsal expression of *Dcn, Ackr3, Meis2, Pitx1,* and *Lrig1* and ventral expression of *Irx2, Ly6h,* and *Nrp2*. NKX2.1-regulated transcriptional changes include increased expression of esophageal genes and decreased expression of tracheal genes in the ventral foregut, accompanied by changes in epithelial-mesenchymal signaling and mesenchymal differentiation.

Table 3.1 Genes increased in E11.5 *Nkx2.1*$^{-/-}$ vs WT foreguts. Differential expression analysis of bulk RNA-seq of E11.5 *Nkx2.1-/-* and WT foregut epithelium with DESeq2. Table shows genes with log2 fold change (Log2FC) > 0.7, adjusted p-value (padj, Wald test, Benjamini-Hochberg adjusted) < 0.05. lfcSE: log fold change standard error.

Genes increased in Nkx2.1-/- vs WT							
Gene	log2FC	lfcSE	padj	Gene	log2FC	lfcSE	padj
Wnt6	2.56	0.19	4.37E-38	Sh3bgrl2	0.83	0.16	8.98E-06
Sp8	2.51	0.17	5.16E-45	Pi4k2b	0.83	0.18	0.0002
Lmo2	2.08	0.20	5.69E-22	Hyal1	0.82	0.13	5.59E-08
Cdkn1a	2.03	0.15	7.33E-38	Sgpp2	0.82	0.17	9.38E-05
Rbp4	1.78	0.20	5.83E-16	Nfe2l3	0.82	0.16	1.25E-05
Lgr5	1.56	0.13	4.75E-32	Map3k13	0.82	0.19	0.0006
Slc2a3	1.50	0.18	8.61E-15	Zim1	0.81	0.16	2.13E-05
Ccdc85a	1.48	0.20	6.11E-11	D630039A03	0.81	0.20	0.0025
Ednra	1.46	0.14	2.60E-24	Fry	0.81	0.14	2.07E-06
Ryr3	1.42	0.20	5.35E-10	Clmn	0.81	0.14	8.10E-07
Klf5	1.41	0.10	2.19E-42	Abhd2	0.80	0.13	1.65E-07
Slc40a1	1.39	0.18	7.05E-13	1110032F04I	0.80	0.20	0.0026
Dnahc5	1.39	0.20	1.60E-09	Fbxo32	0.80	0.16	1.75E-05
Kcnd2	1.36	0.20	2.87E-09	Myof	0.79	0.15	7.58E-06
Arap2	1.36	0.19	1.37E-10	Grb14	0.79	0.20	0.0025
Adamtsl5	1.32	0.20	9.11E-09	Unc13d	0.78	0.19	0.0019
Sp5	1.28	0.15	1.52E-15	Filip1l	0.78	0.13	3.53E-07
Mtap6	1.28	0.20	2.44E-08	Fzd6	0.78	0.11	5.39E-10
Abi3bp	1.27	0.19	9.26E-09	Nell1	0.78	0.20	0.0041
Pld6	1.22	0.20	1.54E-07	Tnfrsf19	0.78	0.10	1.29E-13
Cyp2s1	1.21	0.16	9.53E-12	Pvrl4	0.78	0.19	0.0012
9130014G24Rik	1.21	0.20	1.64E-07	Sh3bp5	0.77	0.19	0.0012
Plxnc1	1.21	0.19	1.25E-08	Kcns1	0.77	0.20	0.0038
Fosl2	1.12	0.18	1.35E-07	Zfhx3	0.77	0.07	1.73E-23
Ap1s3	1.11	0.15	2.02E-11	Lypd6b	0.77	0.16	4.79E-05
Has2	1.11	0.11	1.40E-19	Pof1b	0.77	0.20	0.0039
Tspan12	1.09	0.17	2.73E-08	Rin2	0.76	0.13	1.30E-07
Gm16197	1.07	0.21	1.25E-05	Rbms1	0.76	0.09	3.07E-14
Spock2	1.05	0.15	1.59E-10	Ghr	0.75	0.11	2.87E-09
Slc39a4	1.04	0.20	1.46E-05	Syne1	0.75	0.13	8.67E-07
Vtn	1.03	0.20	2.51E-05	Sptlc3	0.75	0.20	0.0057
Dkk3	1.02	0.16	1.12E-08	Nceh1	0.75	0.20	0.0057
Tshz2	1.01	0.16	3.35E-08	Anxa1	0.75	0.18	0.0010
Ptrf	1.00	0.19	9.93E-06	Epb4.1l4b	0.74	0.17	0.0009
Prickle2	0.98	0.12	3.63E-14	Bex2	0.74	0.13	2.97E-06
S100a6	0.96	0.20	0.0001	Parm1	0.74	0.16	0.0003
Gata3	0.96	0.20	9.51E-05	Psap	0.74	0.08	5.54E-16
Parvb	0.94	0.20	0.0002	Btbd11	0.73	0.18	0.0013
Gm8909	0.94	0.19	2.70E-05	Shh	0.73	0.06	1.07E-28
Ifitm3	0.94	0.19	7.61E-05	Macc1	0.73	0.15	6.37E-05
Mreg	0.93	0.20	0.0003	Palm3	0.73	0.17	0.0006
Atp7b	0.93	0.16	3.03E-07	Pex5l	0.72	0.20	0.0099
Inf2	0.93	0.17	4.07E-06	Fam181b	0.72	0.20	0.0080
Mgat4c	0.92	0.19	0.0001	Gramd1b	0.72	0.18	0.0019
Alpl	0.91	0.13	2.36E-10	8430427H17	0.72	0.10	7.06E-10
Dclk1	0.91	0.17	1.01E-05	Tgm2	0.72	0.16	0.0005
Ahr	0.90	0.16	1.02E-06	Plcd1	0.72	0.18	0.0025
Pdlim7	0.89	0.12	1.12E-11	Serping1	0.72	0.20	0.0097
Camk1d	0.89	0.15	2.09E-07	Msx2	0.71	0.20	0.0117
Mitf	0.89	0.20	0.0006	Sema3e	0.71	0.10	1.70E-10
Efnb2	0.87	0.08	4.24E-22	Dach2	0.71	0.17	0.0018
Rgs6	0.86	0.20	0.0010	Wdr72	0.71	0.18	0.0019
Vps45	0.85	0.20	0.0007	Pcyt1b	0.71	0.18	0.0038
Phlda1	0.84	0.18	0.0002103	Anxa8	0.70	0.14	4.72E-05
Gsto1	0.84	0.17	4.72E-05				

Table 3.2 Genes decreased in E11.5 *Nkx2.1*$^{-/-}$ vs WT foreguts. Differential expression analysis of bulk RNA-seq of E11.5 *Nkx2.1-/-* and WT foregut epithelium with DESeq2. Table shows genes with log2 fold change (Log2FC) < -0.7, adjusted p-value (padj, Wald test, Benjamini-Hochberg adjusted) < 0.05. lfcSE: log fold change standard error.

Genes decreased in Nkx2.1-/- vs WT												
Gene	log2FC	lfcSE	padj	Gene	log2FC	lfcSE	padj	Gene	log2FC	lfcSE	padj
Crlf1	-3.40	0.18	8.79E-73	Muc1	-0.99	0.20	2.50E-05	Mtap1a	-0.78	0.19	0.0017
Hhip	-2.44	0.19	5.78E-35	Ptprm	-0.97	0.14	1.19E-09	Cntn2	-0.77	0.20	0.0035
Msn	-2.12	0.14	6.83E-47	Dgkk	-0.97	0.20	0.0001	Col5a2	-0.77	0.20	0.0028
Nkx2-1	-2.05	0.13	2.50E-52	Thsd4	-0.97	0.15	1.71E-08	Micalcl	-0.77	0.19	0.0024
Mical2	-1.84	0.14	5.43E-38	Creb5	-0.96	0.20	0.0001	Snora81	-0.77	0.20	0.0045
Col4a6	-1.82	0.15	9.75E-32	Maf	-0.96	0.20	0.0001	Dnm1	-0.77	0.20	0.0038
Ror1	-1.82	0.20	5.22E-17	Col4a5	-0.96	0.11	7.90E-17	Sh3kbp1	-0.77	0.20	0.0033
Rspo1	-1.81	0.20	3.26E-16	Dnajb4	-0.96	0.14	1.91E-09	Slc6a4	-0.77	0.20	0.0050
Wnt7b	-1.81	0.10	6.59E-68	Irx1	-0.96	0.11	1.29E-14	Sall4	-0.76	0.17	0.0002
Lrrn3	-1.79	0.20	2.58E-17	Iqgap2	-0.94	0.16	8.41E-07	Snord22	-0.76	0.18	0.0010
Pcdh10	-1.77	0.20	3.16E-16	Pkhd1	-0.94	0.20	0.0001	Serpinb9	-0.76	0.20	0.0054
Etv5	-1.75	0.21	4.99E-15	Ankrd33b	-0.93	0.20	0.0003	Mamld1	-0.76	0.20	0.0043
Ptprt	-1.69	0.20	3.88E-14	Gpr85	-0.93	0.20	0.0002	Cgnl1	-0.76	0.08	8.86E-18
Wnt11	-1.68	0.20	6.77E-14	Rai14	-0.92	0.10	2.66E-18	Mmp16	-0.75	0.19	0.0027
Ptgis	-1.67	0.20	8.68E-14	Islr	-0.92	0.19	9.12E-05	Map3k14	-0.75	0.18	0.0010
Tmem132d	-1.60	0.20	7.87E-13	Pgf	-0.91	0.20	0.0003	Golim4	-0.75	0.10	6.24E-11
A930038C07	-1.59	0.19	8.68E-14	Tub	-0.91	0.16	2.67E-06	Slc16a2	-0.74	0.20	0.0053
Atp8a1	-1.49	0.16	2.25E-17	Sparc	-0.91	0.11	6.67E-13	Hist1h2bp	-0.74	0.20	0.0080
Cdh8	-1.49	0.21	8.47E-11	Sema4f	-0.90	0.20	0.0005	Fhl1	-0.74	0.20	0.0058
Slc24a3	-1.45	0.20	1.98E-10	Gldn	-0.90	0.18	4.77E-05	Daam2	-0.74	0.20	0.0077
Fbn2	-1.45	0.16	4.97E-17	Samd9l	-0.90	0.19	9.51E-05	BC018473	-0.74	0.17	0.0008
Rem2	-1.42	0.20	9.12E-11	Lepr	-0.89	0.20	0.0006	Cacna1g	-0.73	0.20	0.0096
Itga9	-1.39	0.17	4.53E-14	Add3	-0.89	0.09	1.55E-19	Tmem38b	-0.73	0.18	0.0016
Gpr126	-1.36	0.19	2.09E-10	Fndc3c1	-0.89	0.17	1.25E-05	Lass2	-0.72	0.14	1.21E-05
Leprel1	-1.33	0.20	1.12E-08	Snora23	-0.89	0.19	0.0003	Adamts10	-0.72	0.14	3.90E-05
Npc1	-1.31	0.13	1.75E-20	Lpl	-0.88	0.20	0.0004	Ppp1r18	-0.71	0.18	0.0020
Abcc6	-1.28	0.20	2.95E-08	Gprin3	-0.86	0.20	0.0005	Zfp128	-0.71	0.20	0.0088
Mtus2	-1.26	0.14	2.77E-17	Kcnb2	-0.86	0.20	0.0010	Npr3	-0.71	0.20	0.0105
Tecrl	-1.26	0.21	8.31E-08	Smad9	-0.86	0.18	0.0001	Dpyd	-0.71	0.18	0.0023
Adcy7	-1.21	0.20	1.54E-07	Rmrp	-0.86	0.18	6.03E-05	Armcx3	-0.71	0.17	0.0009
Adamts3	-1.20	0.20	4.30E-07	Pcdh18	-0.85	0.14	1.01E-07	D630003M21	-0.71	0.20	0.0122
Dll4	-1.20	0.20	4.03E-07	Fbln5	-0.85	0.20	0.0012	Trpm3	-0.71	0.20	0.0123
Eps8	-1.20	0.17	5.39E-10	Ntf3	-0.84	0.20	0.0013	Cckar	-0.70	0.15	0.0001
Epha4	-1.19	0.17	3.22E-10	Elfn1	-0.83	0.20	0.0011	Kcp	-0.70	0.20	0.0114
Sh3rf3	-1.18	0.20	3.53E-07	Fbln7	-0.83	0.19	0.0009	Tril	-0.70	0.19	0.0062
Ets1	-1.18	0.21	8.95E-07	Foxp2	-0.83	0.09	5.22E-17	Fhod3	-0.70	0.20	0.0110
Ccno	-1.18	0.19	7.98E-08	Socs2	-0.82	0.20	0.0019	Tcea3	-0.70	0.19	0.0058
Dclk2	-1.17	0.17	8.27E-10	Bik	-0.82	0.19	0.0005	Mrc2	-0.70	0.13	2.55E-06
Fmod	-1.16	0.13	2.14E-16	Sept10	-0.82	0.17	9.51E-05				
Scn5a	-1.16	0.18	9.53E-09	Efna1	-0.81	0.15	1.06E-05				
Tle2	-1.15	0.15	5.99E-12	Bmp5	-0.81	0.19	0.0006				
Thbs3	-1.10	0.16	1.81E-09	Hivep2	-0.81	0.14	9.90E-07				
Epha7	-1.09	0.13	7.53E-15	Fads3	-0.80	0.19	0.0012				
Syt7	-1.08	0.15	1.60E-10	Gprc5a	-0.80	0.19	0.0013				
Fam176a	-1.08	0.21	1.15E-05	Entpd3	-0.80	0.18	0.0004				
Rtkn2	-1.07	0.16	1.03E-08	Dpp4	-0.80	0.20	0.0021				
Tmem108	-1.06	0.20	1.04E-05	Snora21	-0.80	0.19	0.0012				
Tppp3	-1.05	0.20	1.33E-05	B3gnt7	-0.79	0.19	0.0011				
Edn3	-1.03	0.15	1.21E-09	Col4a4	-0.79	0.19	0.0016				
Osr2	-1.02	0.21	4.01E-05	Heg1	-0.79	0.18	0.0008				
Col22a1	-1.01	0.19	8.61E-06	Palm2	-0.79	0.18	0.0005				
St6gal1	-1.01	0.12	2.71E-15	Ces1d	-0.78	0.18	0.0005				
Col25a1	-1.00	0.20	5.52E-05	Zeb2	-0.78	0.18	0.0007				
Tns1	-1.00	0.16	2.06E-08	Slitrk5	-0.78	0.21	0.0043				
Tgfb2	-1.00	0.20	3.13E-05	Nhsl2	-0.78	0.13	5.66E-08				

Chapter 4

Future studies and broader impacts

Our ability to understand the extent of NKX2.1 regulation in the foregut has been limited by our lack of understanding of normal tracheoesophageal specification. In Chapter 2 we address this gap in knowledge using scRNA-seq to identify tracheal and esophageal transcriptional programs. In Chapter 3, we utilize this expanded knowledge in combination with *Nkx2.1* genomic analyses to define the NKX2.1 transcriptional program and profile NKX2.1 binding in the foregut genome. Further, we identify a substantial NKX2.1-independent program in the developing foregut. The data presented here greatly expand our understanding of tracheoesophageal specification the mammalian embryo and open many avenues for future exploration.

Several additional experiments will expand on our data describing the relationship between NKX2.1 regulation and *Wnt7b/Shh*-mediated regulation of cartilage and smooth muscle. Functional analysis of the NKX2.1 binding sites near *Wnt7b* and *Shh* identified by our NKX2.1 ChIP-seq data will provide evidence that NKX2.1 directly regulates these genes at these sites. These analyses likely be best accomplished using a reporter knock-in to a safe harbor locus in mice to determine whether NKX2.1 peaks near *Wnt7b* activate reporter expression in the trachea and NKX2.1 peaks near *Shh* repress reporter expression in the trachea. Further, genetic interaction and rescue experiments manipulating *Wnt7b* and *Shh* expression in *Nkx2.1*$^{-/-}$ embryos will determine whether the mesenchymal phenotypes observed *Nkx2.1*$^{-/-}$ mutants are explained solely by changes in *Wnt7b* and *Shh* expression. In fact, several other genes involved in WNT signaling pathways were regulated by NKX2.1 in our *Nkx2.1*$^{-/-}$ vs WT transcriptional analysis including *Wnt6* and *Dkk3* and are promising candidates for exploration of the WNT signaling pathway during foregut mesenchymal specification. Lastly, as disorganized

cartilage still forms in *Nkx2.1*$^{-/-}$ mutant embryos, it is likely that there are multiple pathways involved in foregut mesenchymal specification. Our sequencing datasets provide a resource for identification of such pathways including other epithelial-mesenchymal signaling molecules as well as their upstream regulators.

Our finding that *Nkx2.1*$^{-/-}$ mutant foreguts exhibit both positive and negative transcriptional changes compared to WT indicates that NKX2.1 may have multiple binding partners to confer these differences in activity. This is further supported by previous reports that NKX2.1 can act as an activator and repressor of target genes in the developing mouse brain and alveolar cells *in vitro* (Little et al., 2019; Sandberg et al., 2016). It will be important for future work to dissect these two modes of regulation. A preliminary motif analysis of NKX2.1 peaks near upregulated or downregulated genes identified by our NKX2.1 RNA-seq and ChIP-seq analyses did not reveal significantly differentially enriched motifs that may distinguish positive and negative regulation by NKX2.1. However, our scRNA-seq analysis identified many new transcription factors present in the trachea and respiratory cells that may act with NKX2.1 to regulate tracheal fates. Using these newly-identified transcription factors as candidates and a closer examination of NKX2.1 ChIP-seq peaks, we may further uncover modulators of positive and negative regulation by NKX2.1 in the trachea.

Our discovery of an NKX2.1-independent program during foregut development opens many new avenues for future exploration. We do not yet understand the role that the NKX2.1-independent program plays in tracheoesophageal specification, separation, or differentiation. A phenotypic and transcriptional examination of *Nkx2.1*$^{-/-}$ mutant foreguts in later embryonic development will show us whether this independent program

persists throughout foregut development and will expose the phenotypic consequences of such a program. For example, are dorsal-ventral differences in epithelial structure or cell-type composition observed in *Nkx2.1*$^{-/-}$ mutant foreguts at E18.5? In this way, we may be able to identify specific processes regulated by NKX2.1, such as foregut morphogenesis, and understand which aspects of tracheoesophageal development are independent of NKX2.1.

The role of SOX2 in tracheoesophageal development has not yet been explored using genome-wide approaches. SOX2 is undoubtedly important in tracheoesophageal development, as mice lacking *Sox2* in the foregut endoderm exhibit esophageal agenesis, a failure of foregut separation, and apparent mesenchymal transformation to a tracheal-like phenotype (Que et al., 2007b, 2009; Teramoto et al., 2019). Future RNA-sequencing experiments that identify the SOX2 transcriptional program are required to fully understand the role of SOX2 in foregut development. It will be interesting to compare these SOX2 analyses with our profiling of NKX2.1 in the foregut, to understand whether these two transcription factors have reciprocal or cooperative roles, or a combination of both. Additionally, as SOX2 is expressed in the ventral foregut and trachea in addition to the dorsal foregut and esophagus, it will be interesting to explore the different roles of SOX2 between these two tissues.

The majority of existing mouse mutants that exhibit tracheoesophageal phenotypes target mesenchymal signaling pathways that establish dorsal and ventral identity. These models have been extremely useful in uncovering the signaling cascades that instruct initial foregut specification. However, many developmental processes that are critical for tracheoesophageal development occur within the epithelium, after

dorsoventral specification. These include tracheoesophageal separation, growth and elongation, specification of epithelial structure, and differentiation of specialized cell types. Our profiling of pre- and post-separation trachea and esophagus epithelial transcriptomes therefore provides an exciting dataset in which to explore genes that may have more specific effects on these processes in tracheoesophageal development. Further, with the increased ease of generating genomic perturbations in mice (eg. Takabayashi et al., 2018), we can more readily screen potentially important genes and generate mouse models that allow us to probe specific processes within foregut development.

Many aspects of foregut development are conserved across species. For example, the dorsoventral expression pattern of NKX2.1 and SOX2 is also observed in *Xenopus* embryos, and the phenotypic consequences of *Nkx2.1* loss in the frog mimic those observed in the mouse (Kim et al., 2019; Nasr et al., 2019; Small et al., 2000). Indeed, many of the signaling pathways that establish dorsoventral patterning in the foregut have been extensively studied *Xenopus* foregut development as a model system (eg. Rankin et al., 2017). Additionally, key players in gut tube formation are conserved in other vertebrates such as zebrafish and invertebrates such as *Drosophila*, *C. elegans*, sea urchins, and ascidians (reviewed in Stainier, 2002). While these organisms exhibit conservation within the endoderm, their respiratory organs, if present, differ greatly. These model organisms, along with the increased ease of genomic analyses, provide an interesting ground in which to study respiratory development from an evolutionary context.

Conservation of key regulators of foregut development is also illustrated in human cell culture models. In human iPSC and ESC models, retinoic acid, WNT, and BMP signaling are required to produce anterior foregut monolayers, and organoid models of respiratory, esophagus, and stomach tissue (Dye et al., 2015; McCauley et al., 2017; McCracken et al., 2011; Múnera and Wells, 2017; Trisno et al., 2018). In iPSC/ESC-derived models of ventral foregut and lung development, the presence of NKX2.1-positive cells is often used to indicate achievement of respiratory fate. Our data suggests that, rather than relying on NKX2.1 as the marker of respiratory fate, these *in vitro* differentiation protocols could be better informed with the examination of a panel of NKX2.1-regulated and NKX2.1-independent respiratory markers. Further, human iPSC/ESC models of respiratory development are largely focused on generating lung tissue, or specific differentiated lung cell types (Miller and Spence, 2017). Fewer efforts, however, have focused on generating tracheal tissue, which also plays a significant role in respiratory function and disease. Our transcriptomic profiling of tracheal cells provides a foundation on which to begin generating methods for tracheal differentiation from iPSC/ESCs. It is possible that tracheal-like cells may be generated during initial stages of lung differentiation in iPSC/ESCs, and that identification and purification of these cells using our new markers may facilitate the generation of *in vitro* tracheal cell models.

Foregut malformations in humans are common and include defects in both foregut epithelial and mesenchymal formation. For example, tracheoesophageal fistula (TEF), tracheal agenesis (TA), esophageal atresia (EA), and esophageal agenesis are characterized by improper connections between the trachea and esophagus, or failure of full tracheal or esophageal formation (Brunner and Bokhoven, 2005; Lee, 2018).

Tracheomalacia and tracheal stenosis are characterized by a reduction in tracheal cartilage, leading to tracheal collapse (Brunner and Bokhoven, 2005; Sher and Liu, 2016). Epithelial and mesenchymal foregut malformations can occur on their own, with other foregut malformations, or as a part of a larger syndrome. For example, EA and TEF are often observed in patients with VACTERL (Vertebral anomalies, Anal atresia, Cardiac malformations, Tracheo-Esophageal fistula, Renal and Limb malformations) which encompasses multiple other congenital defects. The genetic causes of foregut malformations in humans are not well understood but do suggest that mouse models will help us understand these defects. For example, *Sox2* has been demonstrated to be critical in mouse foregut development, and mutations in *SOX2* in humans are associated with EA/TEF (Fantes et al., 2003; Hagstrom et al., 2005; Williamson et al., 2006). Additionally, mutations in *NKX2.1* in humans have been associated with brain-lung-thyroid syndrome, a developmental disease that includes choreoathetosis, hypothyroidism, and respiratory distress (Carré et al., 2009; Krude et al., 2002). Future studies in human patients focused on genetic causes of foregut malformations will benefit greatly from advances made in mouse and human stem cell models. Studies such as the one presented in this dissertation that profile the development of normal tracheoesophageal specification will provide a valuable resource in which to inform human genome-wide association studies. They will also contribute to our ultimate ability to treat these conditions with hESC-based tissue engineering approaches.

As a whole, this study uses genomic approaches to understand fate specification and tissue morphogenesis during embryonic development. We utilize genetic mouse models to manipulate fate specification and examine the consequences of these

manipulations at the transcriptomic, morphogenetic, and individual gene level. Not only do the data presented here expose the role of a key regulator and identify new genes involved in foregut development, they also illustrate the power of bringing high throughput experiments back into the embryo to begin to make sense of the complex and remarkable processes that occur during embryonic development.

Chapter 5

Materials and Methods

Mice

All animal procedures were performed at the University of California San Francisco (UCSF) under approval from the UCSF Institutional Animal Care and Use Committee (AN182040). Mouse embryos were collected from pregnant females via cesarean section at the described timepoint following observation of a vaginal plug. Noon the day of the plug was considered embryonic day 0.5. For single cell sequencing experiments, timed-pregnant CD1 female mice were obtained from Harlan/Envigo (Cat: 030) and embryos were staged using somite counts. For mutant analysis, the following alleles were used: $Nkx2.1^{lox/lox}$ (MGI: 3653645), $Sox2^{lox/lox}$ (MGI: 4366453), Nkx2.5-Cre (MGI: 2448972), Actin-Cre (MGI: 2176050).

Immunofluorescence

Mouse embryos were dissected at E10.5 or E11.5 in cold PBS and fixed overnight in 4% paraformaldehyde at 4C. For cryopreservation, embryos were subjected to a sucrose gradient of 12.5% sucrose in PBS for 8 hours, followed by 25% sucrose in PBS overnight at 4C, and embedded in OCT. Tissue sections of 12um were cut using a cryostat and used for immunofluorescence, or in-situ hybridization followed by immunofluorescence with standard protocols. Primary antibodies used for immunofluorescence were: NKX2.1 (Millipore 07601, 1:200), SOX2 (Neuromics GT15098, 1:250), SOX9 (Santa Cruz, sc-20095, 1:250), ECAD (Invitrogen 13-1900, 1:300).

In-situ hybridization

For in-situ hybridization, 12um cryosections were generated as described in "Immunofluorescence". Sections were stained using the RNAScope Multiplex Fluorescent Reagent Kit v2 (Advanced Cell Diagnostics, cat# 323100) with the following adjustments to the manufacturer's protocol: antigen retrieval step was bypassed, protease step used ProteasePlus for 10 minutes. Following in-situ hybridization, slides were washed 2x in PBS and subjected to immunofluorescent staining as described in "Immunofluorescence" above. Probes used for in-situ hybridization against mouse RNA were obtained from Advanced Cell Diagnostics as follows: *Irx2* (519901), *Wnt7b* (401131), *Pcdh10* (477781-C3), *Tppp3* (586631), *Ly6h* (587811), *Krt19* (402941), *Klf5* (444081), *Foxe1* (509641), *Crabp1* (474711-C3), *Bmp4* (401301-C2), *Nrp2* (500661), *Dcn* (413281-C3), *Has2* (465171-C2), *Meis2* (436371-C3), *Ackr3* (482561-C2), *Axin2* (400331-C3), *Lef1* (441861), *Shh* (314361), *Ptch1* (402811-C2), *Gli1* (311001-C2). Additional probes not included in figures: *Ccno* (546521), *Dkk3* (400931-C2), *Gli1* (311001-C2), *Isl1* (451931-C2), *Klf4* (451931-C2), *Lef1* (441861), *Mia* (498011), *Sfrp1* (404981), *Wnt6* (401111-C2).

Skeletal preparation

Skeletal preps were performed as previously described (Martin et al., 1995).

Dissociation and FACS of embryonic tissue

Foregut tissue was dissected in cold PBS and dissociated to single cells using TrypLE Express (phenol-red free, Thermo cat# 12604013) at 37C for 5-minutes, followed by

trituration for 1-3 minutes at 37C. Cells were washed twice with FACS buffer (2mM EDTA and 5% fetal bovine serum in phenol-red free HBSS). To identify epithelial cells, cells were stained with PerCP/Cy5.5 anti-mouse CD326/EpCAM (BioLegend, cat# 118219, used at 1:100) at 4C for 30 minutes followed by two washes with FACS buffer. Cells were resuspended in FACS buffer with Sytox Blue nucleic acid stain (Thermo, S11348, used at 1uM) to stain dead cells, and passed through a 35uM cell strainer. Cells were sorted using a BD FACS Aria II. Single live epithelial cells were collected after size selection and gating for Sytox-negative, EpCAM-positive cells. For scRNA-seq, cells were sorted into EDTA-free FACS buffer and processed as described below. For bulk-RNA-seq, cells were sorted into RNA lysis buffer (Qiagen RNeasy Micro kit, cat# 74004) with 1% beta-mercaptoethanol and processed as described below.

Single-cell RNA sequencing

Foregut tissue was dissected from 20 embryos at E10.5 (6-9ts) and 28 embryos at E11.5 (16-20ts). To ensure for representation of tracheal and esophageal cells at E11.5, lung tissue was separated from E11.5 foreguts and processed in parallel. Tissue from each timepoint was pooled and single-cell suspension and epithelial purification was performed as described above. 25,000 live epithelial cells from each sample were loaded into individual wells for single-cell capture using the Chromium Single Cell 3' Reagent Kit V2 (10X Genomics). Library preparation for each sample was also performed using the Chromium Single Cell 3' Reagent Kit V2, and each sample was given a unique i7 index. Libraries were pooled and subjected to sequencing in a single lane of an Illumina

NovaSeq6000. Sequencing data was processed, and downstream analysis performed as described below.

RNA sequencing

For bulk RNA-sequencing experiments, whole foregut tissue was dissected from E11.5 Nkx2.1-/- or WT embryos, and lungs were removed at the time of dissection. For RNA-sequencing of WT trachea and esophagus, trachea and esophagus were manually separated at the time of dissection. Foreguts of individual embryos were dissociated, stained, and sorted as described above. Each biological replicate consisted of RNA pooled from two Nkx2.1-/- or WT embryos, with a total of 3 biological replicates. RNA was purified using the RNeasy Micro kit (Qiagen, cat# 74004) and quantification was performed using the RNA 600 Pico kit (Agilent, cat# 5067-1513) on an Agilent 2100 Bioanalyzer. RNA-sequencing libraries were prepared from 4ng of input RNA using the SMARTer Stranded Total RNAseq kit V2 (Takara, cat# 634411) with 13 amplification cycles with a BioRad C1000 ThermalCycler. Library size and quality was checked using an Agilent 2100 Bioanalyzer with the High Sensitivity DNA kit (Agilent, cat# 5067-4626), and library concentration was determined with the QuBit dsDNA HS Assay kit (Invitrogen, cat# Q32854). Libraries were normalized to 7nM, pooled, and sequenced across two lanes of an Illumina HiSeq 4000 to generate single-end reads. Data processing and downstream analysis was performed as described below.

Chromatin immunoprecipitation and sequencing

For ChIP-seq experiments, trachea were dissected from E11.5 mouse embryos and cross-linked with 1% cold PFA for 9min. Cross-linking was stopped with glycine for a final concentration of 0.125M. Crosslinked tissue was washed 2x in PBS and stored at -80C. For each replicate 175 trachea were pooled and tissue was dissociated in cold PBS by passing through a 25G needle until fully dissociated. Cells were lysed in 500ul lysis buffer (50mM Tris-HCl pH8 + 2mM EDTA pH8 + 0.1% NP-40 + 10% glycerol in DNase/RNase-free water) with protease inhibitors (Aprotinin + Pepstatin A + Leupeptin + 1mM PMSF) for 5 minutes on ice. Nuclei were pelleted by spinning cells at 845xg for 5 minutes at 8C. Nuclei were lysed with 500ul SDS lysis buffer (50mM Tris-HCl + 10mM EDTA + 1% SDS in sterile water) for 5 minutes on ice. Chromatin from lysed nuclei in SDS lysis buffer was sheared to obtain 200-500bp DNA fragments using a Diagenode Bioruptor UCD-200 with 35 cycles (30sec on/off) submerged in cold water. Fragment size was determined by running a 20ul aliquot of decrosslinked chromatin on a 1.5% agarose gel. Sheared chromatin was diluted 1:10 with ChIP dilution buffer (50mM Tris-HCl + 2mM EDTA + 0.5M NaCl + 0.1% SDS + 1% SDS in sterile water) then pre-cleared with washed Dynabeads Protein G for 1h at 4C. Dynabeads were magnetically isolated from chromatin and 1% of chromatin was separated and decrosslinked to be used as input. 75% of the remaining sample was incubated with Nkx2.1 antibody (Millipore 07601), and 25% was incubated with anti-rabbit IgG (Cell Signaling 2729) overnight at 4C. To isolate antibody-bound DNA, washed Dynabeads Protein G were added to each sample (50ul beads/sample) and incubated for 30 minutes at 4C. Dynabeads with antibody-bound chromatin were isolated magnetically and subjected to 3x washes each with wash buffer (10mM Tris-HCl + 2mM

EDTA + 0.5M NaCl + 0.1% SDS + 1% NP-40 in sterile water), LiCl buffer (10mM Tris-HCl + 2mM EDTA + 0.5M LiCl + 0.1% SDS + 1% NP-40), and TE buffer (1mM Tris-HCl + 1mM EDTA) for 5 minutes on ice. Chromatin was eluted in 100ul of 2% SDS in TE on a 65C heatblock with vigorous shaking (1400rpm) for 15 minutes. Input DNA and immunoprecipitated DNA were decrosslinked by adding 5ul 5M NaCl to 100ul eluate and incubating overnight at 65C. Decrosslinked DNA was purified using MicroChIP Diapure columns (Diagenode, cat# C03040001) and eluted in 10ul of elution buffer. The entire eluate of Nkx2.1 immunoprecipitated DNA and 0.5ng of input DNA were used to prepare ChIP libraries. Libraries were prepared using the Microplex Library Preparation Kit v2 (Diagenode, cat# C05010012) according to manufacturers instructions. Library quality and size were calculated using an Agilent 2100 Bioanalyzer with the High Sensitivity DNA kit (Agilent, cat# 5067-4626), and library concentration was quantified with the QuBit dsDNA HS Assay kit (Invitrogen, cat# Q32854). Libraries were pooled to 5nM and sequenced on an Illumina HiSeq 4000. Sequencing data was processed, and downstream analysis was performed as described below.

Analysis of scRNA-seq data

We used the Cell Ranger v2.1.1 pipeline from 10X Genomics 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. Filtered gene-barcode matrices were then analyzed using the Seurat v3.0 R package(Stuart et al., 2018). Seurat objects were generated with CreateSuratObject (min.cells = 10, min.features = 200) for E10.5 foregut

and lung cells, E11.5 foregut cells, and E11.5 lung cells. E11.5 foregut and lung cells were merged to create a single gene-barcode matrix. Cells were further filtered based on the distribution of number of genes (nFeature) and percent mitochondrial genes (percent.mito) per cell across the dataset as follows. nFeature_RNA (E10.5): >2000, <7000, nFeature_RNA (E11.5): >2500, <8500, percent.mito: >0.5, <7.5. Data was normalized for sequencing depth, log-transformed, and multiplied by a scale factor of 10000 using the default parameters of NormalizeData. Linear regression was performed to eliminate variability across cell cycle stage (CellCycleScoring) and mitochondrial content using ScaleData. For E11.5 merged foregut and lung, nCount_RNA was also regressed out as these datasets retained slight variability in sequence depth that was not eliminated with ScaleData. The top 2000 variable genes within each dataset were selected based on a variance stabilizing transformation (FindVariableGenes, selection.method = “vst”) and used in downstream principal component analysis (PCA). The principal components (PCs) were identified with RunPCA and PCs to include in downstream analysis were empirically determined with visualization of PCs in an ElbowPlot. Cell clusters were identified by construction of a shared nearest neighbor graph (FindNeighbors) and a modularity optimization-based clustering algorithm (FindClusters) using the PCs determined by PCA (E10.5 dims = 1:20, E11.5 dims = 1:12). Clustering was performed at multiple resolutions between 0.2 and 2, and optimal resolution was determined empirically based on the expression of known population markers and the FindMarkers function (E10.5 resolution = 0.55, E11.5 resolution = 0.45). Several outlying clusters of mesenchymal contamination were removed, and cells were re-clustered for visualization purposes. Cells and clustering were visualized using Uniform

Manifold Approximation and Projection (UMAP) dimensional reduction (RunUMAP). Markers for each cluster were identified with FindAllMarkers using default parameters, and cluster identity was determined based on the presence of known markers, as well as experimental evidence of RNA localization in specific cell types.

Analysis of bulk RNA-seq data

Analysis of RNA-seq reads was performed as described previously (Percharde et al., 2018). Differential expression analysis (Nkx2.1-/- vs WT, WT trachea vs WT esophagus) was performed using DESeq2(Love et al., 2014) (test=c("Wald"), betaPrior=T) and genes with a log2 fold change >0.7 or <-0.7 and an adjusted p-value>0.05 were determined to be differentially expressed. Differential expression results were visualized using the ggplot2 package.

Analysis of ChIP-seq data

FASTQ files of raw sequencing reads for Nkx2.1 ChIP and input libraries were processed using a custom script (github.com/mpercharde/ChIPseq). Trimming and generation of fastqc files to examine sequence quality were performed using Trim Galore. Trimmed reads were aligned to the mouse genome (mm9) using bowtie2 and sorted deduplicated bam files were generated using samtools. Peak calling was performed with MACS2 using a false discovery rate less than 1e-5 (macs2 callpeak -t chip.sorted.bam -c input.sorted.bam -f BAM -q 0.00001 -g mm -n nkx_peaks --outdir macs/). Peaks shared between both replicates were identified by finding the intersection of both replicates using Galaxy. Motif analysis to test for the enrichment of the Nkx2.1 motif was performed with

MEME ChIP using a 500bp region flanking the peak summit for all peaks shared between both replicates. Nkx2.1 binding at specific loci was visualized in the Integrative Genomics Viewer.

Acknowledgement of UCSF cores and other resources used

All cell sorting was performed at the UCSF Parnassus Flow Cytometry Core supported in part by Grant NIH P30 DK063720. Sequencing of RNA-seq libraries and ChIP-seq libraries was performed at the UCSF Center for Advanced Technology. Single cell sequencing including 10X capture, library preparation, sequencing, and CellRanger processing was performed by the UCSF Institute for Human Genetics. Data analysis was performed using the XSEDE PSC Bridges Computing Resource.

References

Barker, N., Es, J.H. van, Kuipers, J., Kujala, P., Born, M. van den, Cozijnsen, M., Haegebarth, A., Korving, J., Begthel, H., Peters, P.J., et al. (2007). Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature *449*, 1003–1007.

Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. Nature Biotechnology *37*, 38.

Besnard, V., Xu, Y., and Whitsett, J.A. (2007). Sterol response element binding protein and thyroid transcription factor-1 (Nkx2.1) regulate Abca3 gene expression. American Journal of Physiology-Lung Cellular and Molecular Physiology *293*, L1395–L1405.

Bi, W., Deng, J.M., Zhang, Z., Behringer, R.R., and de Crombrugghe, B. (1999). Sox9 is required for cartilage formation. Nat. Genet. *22*, 85–89.

Billmyre, K.K., Hutson, M., and Klingensmith, J. (2015). One shall become two: Separation of the esophagus and trachea from the common foregut tube. Dev. Dyn. *244*, 277–288.

Boggaram, V. (2009). Thyroid transcription factor-1 (TTF-1/Nkx2.1/TITF1) gene regulation in the lung. Clin Sci (Lond) *116*, 27–35.

Bohinski, R.J., Di Lauro, R., and Whitsett, J.A. (1994). The lung-specific surfactant protein B gene promoter is a target for thyroid transcription factor 1 and

hepatocyte nuclear factor 3, indicating common factors for organ-specific gene expression along the foregut axis. Mol. Cell. Biol. *14*, 5671–5681.

Brand-Saberi, B.E.M., and Schäfer, T. (2014). Trachea: Anatomy and Physiology. Thoracic Surgery Clinics *24*, 1–5.

Brunner, H.G., and Bokhoven, H. van (2005). Genetic players in esophageal atresia and tracheoesophageal fistula. Current Opinion in Genetics & Development *15*, 341–347.

Bruno, M.D., Bohinski, R.J., Huelsman, K.M., Whitsett, J.A., and Korfhagen, T.R. (1995). Lung cell-specific expression of the murine surfactant protein A (SP-A) gene is mediated by interactions between the SP-A promoter and thyroid transcription factor-1. J. Biol. Chem. *270*, 6531–6536.

Cardoso, W.V., and Lü, J. (2006). Regulation of early lung morphogenesis: questions, facts and controversies. Development *133*, 1611–1624.

Cardoso, W.V., and Whitsett, J.A. (2008). Resident cellular components of the lung: developmental aspects. Proc Am Thorac Soc *5*, 767–771.

Carré, A., Szinnai, G., Castanet, M., Sura-Trueba, S., Tron, E., Broutin-L'Hermite, I., Barat, P., Goizet, C., Lacombe, D., Moutard, M.-L., et al. (2009). Five new TTF1/NKX2.1 mutations in brain–lung–thyroid syndrome: rescue by PAX8 synergism in one case. Hum Mol Genet *18*, 2266–2276.

Comai, G., Heude, E., Mella, S., Paisant, S., Pala, F., Gallardo, M., Langa, F., Kardon, G., Gopalakrishnan, S., and Tajbakhsh, S. (2019). A distinct cardiopharyngeal mesoderm genetic hierarchy establishes antero-posterior patterning of esophagus striated muscle. ELife *8*, e47460.

Domyan, E.T., Ferretti, E., Throckmorton, K., Mishina, Y., Nicolis, S.K., and Sun, X. (2011). Signaling through BMP receptors promotes respiratory identity in the foregut via repression of Sox2. Development *138*, 971–981.

Dye, B.R., Hill, D.R., Ferguson, M.A., Tsai, Y.-H., Nagy, M.S., Dyal, R., Wells, J.M., Mayhew, C.N., Nattiv, R., Klein, O.D., et al. (2015). In vitro generation of human pluripotent stem cell derived lung organoids. ELife *4*, e05098.

Fantes, J., Ragge, N.K., Lynch, S.-A., McGill, N.I., Collin, J.R.O., Howard-Peebles, P.N., Hayward, C., Vivian, A.J., Williamson, K., van Heyningen, V., et al. (2003). Mutations in SOX2 cause anophthalmia. Nat. Genet. *33*, 461–463.

Fausett, S.R., Brunet, L.J., and Klingensmith, J. (2014). BMP antagonism by Noggin is required in presumptive notochord cells for mammalian foregut morphogenesis. Developmental Biology *391*, 111–124.

Funk, M.C., Bera, A.N., Menchen, T., Kuales, G., Thriene, K., Lienkamp, S.S., Dengjel, J., Omran, H., Frank, M., and Arnold, S.J. (2015). Cyclin O (Ccno) functions during deuterosome-mediated centriole amplification of multiciliated cells. The EMBO Journal *34*, 1078–1089.

Gerhardt, B., Leesman, L., Burra, K., Snowball, J., Rosenzweig, R., Guzman, N., Ambalavanan, M., and Sinner, D. (2018). Notum attenuates Wnt/β–catenin signaling to promote tracheal cartilage patterning. Developmental Biology *436*, 14–27.

Giridharan, S.S.P., Rohn, J.L., Naslavsky, N., and Caplan, S. (2012). Differential regulation of actin microfilaments by human MICAL proteins. J Cell Sci *125*, 614–624.

Gopalakrishnan, S., Comai, G., Sambasivan, R., Francou, A., Kelly, R.G., and Tajbakhsh, S. (2015). A Cranial Mesoderm Origin for Esophagus Striated Muscles. Dev. Cell *34*, 694–704.

Goss, A.M., Tian, Y., Tsukiyama, T., Cohen, E.D., Zhou, D., Lu, M.M., Yamaguchi, T.P., and Morrisey, E.E. (2009). Wnt2/2b and beta-catenin signaling are necessary and sufficient to specify lung progenitors in the foregut. Dev. Cell *17*, 290–298.

Guazzi, S., Price, M., De Felice, M., Damante, G., Mattei, M.G., and Di Lauro, R. (1990). Thyroid nuclear factor 1 (TTF-1) contains a homeodomain and displays a novel DNA binding specificity. EMBO J. *9*, 3631–3639.

Hagstrom, S.A., Pauer, G.J.T., Reid, J., Simpson, E., Crowe, S., Maumenee, I.H., and Traboulsi, E.I. (2005). SOX2 mutation causes anophthalmia, hearing loss, and brain anomalies. Am. J. Med. Genet. A *138A*, 95–98.

Han, L., Koike, H., Chaturvedi, P., Kishimoto, K., Iwasawa, K., Giesbrecht, K., Witcher, P.C., Eicher, A., Nasr, T., Haines, L., et al. (2019). Single cell transcriptomics

reveals a signaling roadmap coordinating endoderm and mesoderm lineage diversification during foregut organogenesis. BioRxiv 756825.

Harris-Johnson, K.S., Domyan, E.T., Vezina, C.M., and Sun, X. (2009). beta-Catenin promotes respiratory progenitor identity in mouse foregut. Proc. Natl. Acad. Sci. U.S.A. *106*, 16287–16292.

Herriges, J.C., Yi, L., Hines, E.A., Harvey, J.F., Xu, G., Gray, P.A., Ma, Q., and Sun, X. (2012). Genome-scale study of transcription factor expression in the branching mouse lung. Developmental Dynamics *241*, 1432–1453.

Hines, E.A., Jones, M.-K.N., Verheyden, J.M., Harvey, J.F., and Sun, X. (2013). Establishment of smooth muscle and cartilage juxtaposition in the developing mouse upper airways. Proc. Natl. Acad. Sci. U.S.A. *110*, 19444–19449.

Hou, Z., Wu, Q., Sun, X., Chen, H., Li, Y., Zhang, Y., Mori, M., Yang, Y., Que, J., and Jiang, M. (2019). Wnt/Fgf crosstalk is required for the specification of basal cells in the mouse trachea. Development *146*, dev171496.

Huycke, T.R., Miller, B.M., Gill, H.K., Nerurkar, N.L., Sprinzak, D., Mahadevan, L., and Tabin, C.J. (2019a). Genetic and Mechanical Regulation of Intestinal Smooth Muscle Development. Cell *179*, 90-105.e21.

Huycke, T.R., Miller, B.M., Gill, H.K., Nerurkar, N.L., Sprinzak, D., Mahadevan, L., and Tabin, C.J. (2019b). Genetic and Mechanical Regulation of Intestinal Smooth Muscle Development. Cell *179*, 90-105.e21.

Ishii, Y., Rex, M., Scotting, P.J., and Yasugi, S. (1998). Region-specific expression of chicken Sox2 in the developing gut and lung epithelium: Regulation by epithelial-mesenchymal interactions. Developmental Dynamics *213*, 464–475.

Kelly, S.E., Bachurski, C.J., Burhans, M.S., and Glasser, S.W. (1996). Transcription of the lung-specific surfactant protein C gene is mediated by thyroid transcription factor 1. J. Biol. Chem. *271*, 6881–6888.

Kim, E., Jiang, M., Huang, H., Zhang, Y., Robert, J., Gilmore, N., Gan, L., and Que, J. (2019). Isl1 Regulation of Nkx2.1 in the Early Foregut Epithelium Is Required for Trachea-Esophageal Separation and Lung Lobation (Rochester, NY: Social Science Research Network).

Kim, H.Y., Pang, M.-F., Varner, V.D., Kojima, L., Miller, E., Radisky, D.C., and Nelson, C.M. (2015). Localized Smooth Muscle Differentiation Is Essential for Epithelial Bifurcation during Branching Morphogenesis of the Mammalian Lung. Dev. Cell *34*, 719–726.

Kim, K.-H., Rosen, A., Bruneau, B.G., Hui, C., and Backx, P.H. (2012). Iroquois homeodomain transcription factors in heart development and function. Circ. Res. *110*, 1513–1524.

Kimura, S., Hara, Y., Pineau, T., Fernandez-Salguero, P., Fox, C.H., Ward, J.M., and Gonzalez, F.J. (1996). The T/ebp null mouse: thyroid-specific enhancer-binding protein is essential for the organogenesis of the thyroid, lung, ventral forebrain, and pituitary. Genes Dev. *10*, 60–69.

Kishimoto, K., Furukawa, K.T., LuzMadrigal, A., Yamaoka, A., Matsuoka, C., Habu, M., Alev, C., Zorn, A.M., and Morimoto, M. (2019). Induction of tracheal mesoderm and chondrocyte from pluripotent stem cells in mouse and human. BioRxiv 758235.

Krude, H., Schütz, B., Biebermann, H., Moers, A. von, Schnabel, D., Neitzel, H., Tönnies, H., Weise, D., Lafferty, A., Schwarz, S., et al. (2002). Choreoathetosis, hypothyroidism, and pulmonary alterations due to human NKX2-1 haploinsufficiency. J Clin Invest *109*, 475–480.

Kuo, B., and Urma, D. (2006). Esophagus - anatomy and development. GI Motility Online.

Lee, S. (2018). Basic Knowledge of Tracheoesophageal Fistula and Esophageal Atresia. Adv Neonatal Care *18*, 14–21.

Li, Y., Litingtung, Y., Dijke, P.T., and Chiang, C. (2007). Aberrant Bmp signaling and notochord delamination in the pathogenesis of esophageal atresia. Developmental Dynamics *236*, 746–754.

Li, Y., Gordon, J., Manley, N.R., Litingtung, Y., and Chiang, C. (2008). Bmp4 is required for tracheal formation: A novel mouse model for tracheal agenesis. Developmental Biology *322*, 145–155.

Litingtung, Y., Lei, L., Westphal, H., and Chiang, C. (1998). Sonic hedgehog is essential to foregut development. Nat Genet *20*, 58–61.

Little, D.R., Gerner-Mauro, K.N., Flodby, P., Crandall, E.D., Borok, Z., Akiyama, H., Kimura, S., Ostrin, E.J., and Chen, J. (2019). Transcriptional control of lung alveolar type 1 cell development and maintenance by NK homeobox 2-1. Proc. Natl. Acad. Sci. U.S.A.

Liu, C., Glasser, S.W., Wan, H., and Whitsett, J.A. (2002). GATA-6 and thyroid transcription factor-1 directly interact and regulate surfactant protein-C gene expression. J. Biol. Chem. *277*, 4519–4525.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Mao, J., Kim, B.-M., Rajurkar, M., Shivdasani, R.A., and McMahon, A.P. (2010). Hedgehog signaling controls mesenchymal growth in the developing mammalian digestive tract. Development *137*, 1721–1729.

Martin, J.F., Bradley, A., and Olson, E.N. (1995). The paired-like homeo box gene MHox is required for early events of skeletogenesis in multiple lineages. Genes Dev. *9*, 1237–1249.

McCauley, K.B., Hawkins, F., Serra, M., Thomas, D.C., Jacob, A., and Kotton, D.N. (2017). Efficient Derivation of Functional Human Airway Epithelium from Pluripotent Stem Cells via Temporal Regulation of Wnt Signaling. Cell Stem Cell *20*, 844-857.e6.

McCracken, K.W., Howell, J.C., Wells, J.M., and Spence, J.R. (2011). Generating human intestinal tissue from pluripotent stem cells in vitro. Nat Protoc *6*, 1920–1928.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis - regulatory regions. Nat Biotechnol *28*, 495–501.

Miller, A.J., and Spence, J.R. (2017). In Vitro Models to Study Human Lung Development, Disease and Homeostasis. Physiology (Bethesda) *32*, 246–260.

Miller, L.-A.D., Wert, S.E., Clark, J.C., Xu, Y., Perl, A.-K.T., and Whitsett, J.A. (2004). Role of Sonic hedgehog in patterning of tracheal-bronchial cartilage and the peripheral lung. Dev. Dyn. *231*, 57–71.

Minoo, P., Su, G., Drum, H., Bringas, P., and Kimura, S. (1999a). Defects in Tracheoesophageal and Lung Morphogenesis inNkx2.1(−/−) Mouse Embryos. Developmental Biology *209*, 60–71.

Minoo, P., Su, G., Drum, H., Bringas, P., and Kimura, S. (1999b). Defects in tracheoesophageal and lung morphogenesis in Nkx2.1(-/-) mouse embryos. Dev. Biol. *209*, 60–71.

Minoo, P., Hu, L., Xing, Y., Zhu, N.L., Chen, H., Li, M., Borok, Z., and Li, C. (2007). Physical and Functional Interactions between Homeodomain NKX2.1 and Winged Helix/Forkhead FOXA1 in Lung Epithelial Cells. Molecular and Cellular Biology *27*, 2155–2165.

Mizuno, K., Gonzalez, F.J., and Kimura, S. (1991). Thyroid-specific enhancer-binding protein (T/EBP): cDNA cloning, functional characterization, and structural identity with thyroid transcription factor TTF-1. Mol. Cell. Biol. *11*, 4927–4933.

Morrisey, E.E., and Hogan, B.L.M. (2010). Preparing for the First Breath: Genetic and Cellular Mechanisms in Lung Development. Developmental Cell *18*, 8–23.

Morrisey, E.E., Cardoso, W.V., Lane, R.H., Rabinovitch, M., Abman, S.H., Ai, X., Albertine, K.H., Bland, R.D., Chapman, H.A., Checkley, W., et al. (2013). Molecular Determinants of Lung Development. Annals ATS *10*, S12–S16.

Múnera, J.O., and Wells, J.M. (2017). Generation of Gastrointestinal Organoids from Human Pluripotent Stem Cells. Methods Mol. Biol. *1597*, 167–177.

Nakazato, M., Endo, T., Saito, T., Harii, N., and Onaya, T. (1997). Transcription of the thyroid transcription factor-1 (TTF-1) gene from a newly defined start site: positive regulation by TTF-1 in the thyroid. Biochem. Biophys. Res. Commun. *238*, 748–752.

Nasr, T., Mancini, P., Rankin, S.A., Edwards, N.A., Agricola, Z.N., Kenny, A.P., Kinney, J.L., Daniels, K., Vardanyan, J., Han, L., et al. (2019). Endosome-Mediated Epithelial Remodeling Downstream of Hedgehog-Gli Is Required for Tracheoesophageal Separation. Developmental Cell *51*, 665-674.e6.

Neisch, A.L., and Fehon, R.G. (2011). Ezrin, Radixin and Moesin: key regulators of membrane–cortex interactions and signaling. Current Opinion in Cell Biology *23*, 377–382.

Niimi, T., Nagashima, K., Ward, J.M., Minoo, P., Zimonjic, D.B., Popescu, N.C., and Kimura, S. (2001a). claudin-18, a Novel Downstream Target Gene for the T/EBP/NKX2.1 Homeodomain Transcription Factor, Encodes Lung- and Stomach-Specific Isoforms through Alternative Splicing. Molecular and Cellular Biology *21*, 7380–7390.

Niimi, T., Keck-Waggoner, C.L., Popescu, N.C., Zhou, Y., Levitt, R.C., and Kimura, S. (2001b). UGRP1, a Uteroglobin/Clara Cell Secretory Protein-Related Protein, Is a Novel Lung-Enriched Downstream Target Gene for the T/EBP/NKX2.1 Homeodomain Transcription Factor. Mol Endocrinol *15*, 2021–2036.

Nilsson, M., and Fagman, H. (2017). Development of the thyroid gland. Development *144*, 2123.

Nowotschin, S., Setty, M., Kuo, Y.-Y., Liu, V., Garg, V., Sharma, R., Simon, C.S., Saiz, N., Gardner, R., Boutet, S.C., et al. (2019). The emergent landscape of the mouse gut endoderm at single-cell resolution. Nature *569*, 361–367.

Oguchi, H., and Kimura, S. (1998). Multiple transcripts encoded by the thyroid-specific enhancer-binding protein (T/EBP)/thyroid-specific transcription factor-1 (TTF-1) gene: evidence of autoregulation. Endocrinology *139*, 1999–2006.

Ostrin, E.J., Little, D.R., Gerner-Mauro, K.N., Sumner, E.A., Ríos-Corzo, R., Ambrosio, E., Holt, S.E., Forcioli-Conti, N., Akiyama, H., Hanash, S.M., et al. (2018). β-Catenin maintains lung epithelial progenitors after lung specification. Development *145*.

Pepicelli, C.V., Lewis, P.M., and McMahon, A.P. (1998). Sonic hedgehog regulates branching morphogenesis in the mammalian lung. Curr. Biol. *8*, 1083–1086.

Percharde, M., Lin, C.-J., Yin, Y., Guan, J., Peixoto, G.A., Bulut-Karslioglu, A., Biechele, S., Huang, B., Shen, X., and Ramalho-Santos, M. (2018). A LINE1-Nucleolin Partnership Regulates Early Development and ESC Identity. Cell *174*, 391-405.e19.

Perl, A.-K.T., Kist, R., Shan, Z., Scherer, G., and Whitsett, J.A. (2005). Normal lung development and function after Sox9 inactivation in the respiratory epithelium. Genesis *41*, 23–32.

Que, J., Choi, M., Ziel, J.W., Klingensmith, J., and Hogan, B.L.M. (2006). Morphogenesis of the trachea and esophagus: current players and new roles for noggin and Bmps. Differentiation *74*, 422–437.

Que, J., Okubo, T., Goldenring, J.R., Nam, K.-T., Kurotani, R., Morrisey, E.E., Taranova, O., Pevny, L.H., and Hogan, B.L.M. (2007a). Multiple dose-dependent roles for Sox2 in the patterning and differentiation of anterior foregut endoderm. Development *134*, 2521–2531.

Que, J., Okubo, T., Goldenring, J.R., Nam, K.-T., Kurotani, R., Morrisey, E.E., Taranova, O., Pevny, L.H., and Hogan, B.L.M. (2007b). Multiple dose-dependent roles for Sox2 in the patterning and differentiation of anterior foregut endoderm. Development *134*, 2521.

Que, J., Luo, X., Schwartz, R.J., and Hogan, B.L.M. (2009). Multiple roles for Sox2 in the developing and adult mouse trachea. Development *136*, 1899.

Rajagopal, J., Carroll, T.J., Guseh, J.S., Bores, S.A., Blank, L.J., Anderson, W.J., Yu, J., Zhou, Q., McMahon, A.P., and Melton, D.A. (2008). Wnt7b stimulates embryonic lung growth by coordinately increasing the replication of epithelium and mesenchyme. Development *135*, 1625–1634.

Rankin, S.A., Han, L., McCracken, K.W., Kenny, A.P., Anglin, C.T., Grigg, E.A., Crawford, C.M., Wells, J.M., Shannon, J.M., and Zorn, A.M. (2016). A Retinoic Acid-Hedgehog Cascade Coordinates Mesoderm-Inducing Signals and Endoderm Competence during Lung Specification. Cell Reports *16*, 66–78.

Rankin, S.A., McCracken, K.W., Luedeke, D.M., Han, L., Wells, J.M., Shannon, J.M., and Zorn, A.M. (2017). Timing is everything: Reiterative Wnt, BMP and RA signaling regulate developmental competence during endoderm organogenesis. Dev. Biol.

Rockich, B.E., Hrycaj, S.M., Shih, H.P., Nagy, M.S., Ferguson, M.A.H., Kopp, J.L., Sander, M., Wellik, D.M., and Spence, J.R. (2013). Sox9 plays multiple roles in the lung epithelium during branching morphogenesis. Proc Natl Acad Sci USA *110*, E4456.

Sala, F.G., Del Moral, P.-M., Tiozzo, C., Alam, D.A., Warburton, D., Grikscheit, T., Veltmaat, J.M., and Bellusci, S. (2011). FGF10 controls the patterning of the tracheal cartilage rings via Shh. Development *138*, 273–282.

Sandberg, M., Flandin, P., Silberberg, S., Su-Feher, L., Price, J.D., Hu, J.S., Kim, C., Visel, A., Nord, A.S., and Rubenstein, J.L.R. (2016). Transcriptional Networks Controlled by NKX2-1 in the Development of Forebrain GABAergic Neurons. Neuron *91*, 1260–1275.

Sandberg, M., Taher, L., Hu, J., Black, B.L., Nord, A.S., and Rubenstein, J.L.R. (2018). Genomic analysis of transcriptional networks directing progression of cell states during MGE development. Neural Dev *13*, 21.

Sawaya, P.L., Stripp, B.R., Whitsett, J.A., and Luse, D.S. (1993). The lung-specific CC10 gene is regulated by transcription factors from the AP-1, octamer, and hepatocyte nuclear factor 3 families. Mol. Cell. Biol. *13*, 3860–3871.

Shannon, J.M., Nielsen, L.D., Gebb, S.A., and Randell, S.H. (1998). Mesenchyme specifies epithelial differentiation in reciprocal recombinants of embryonic lung and trachea. Dev. Dyn. *212*, 482–494.

Sher, Z.A., and Liu, K.J. (2016). Congenital tracheal defects: embryonic development and animal models. Genetics 2016, Vol. 3, Pages 60-73.

Small, E.M., Vokes, S.A., Garriock, R.J., Li, D., and Krieg, P.A. (2000). Developmental expression of the Xenopus Nkx2-1 and Nkx2-4 genes. Mechanisms of Development *96*, 259–262.

Snowball, J., Ambalavanan, M., Whitsett, J., and Sinner, D. (2015a). “Endodermal Wnt signaling is required for tracheal cartilage formation.” Developmental Biology *405*, 56–70.

Snowball, J., Ambalavanan, M., Whitsett, J., and Sinner, D. (2015b). “Endodermal Wnt signaling is required for tracheal cartilage formation.” Developmental Biology *405*, 56–70.

Snyder, E.L., Watanabe, H., Magendantz, M., Hoersch, S., Chen, T.A., Wang, D.G., Crowley, D., Whittaker, C.A., Meyerson, M., Kimura, S., et al. (2013). Nkx2-1 represses a latent gastric differentiation program in lung adenocarcinoma. Mol. Cell *50*, 185–199.

Speck, O., Hughes, S.C., Noren, N.K., Kulikauskas, R.M., and Fehon, R.G. (2003). Moesin functions antagonistically to the Rho pathway to maintain epithelial integrity. Nature *421*, 83–87.

Stainier, D.Y.R. (2002). A glimpse into the molecular entrails of endoderm formation. Genes Dev. *16*, 893–907.

Stanley, E.G., Biben, C., Elefanty, A., Barnett, L., Koentgen, F., Robb, L., and Harvey, R.P. (2004). Efficient Cre-mediated deletion in cardiac progenitor cells conferred by a 3’UTR-ires-Cre allele of the homeobox gene Nkx2-5. Int. J. Dev. Biol. *46*, 431–439.

Stevens, M.L., Chaturvedi, P., Rankin, S.A., Macdonald, M., Jagannathan, S., Yukawa, M., Barski, A., and Zorn, A.M. (2017). Genomic integration of Wnt/β-catenin and BMP/Smad1 signaling coordinates foregut and hindgut transcriptional programs. Development *144*, 1283–1295.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Stoeckius, M., Smibert, P., and Satija, R. (2018). Comprehensive integration of single cell data. BioRxiv 460147.

Tagne, J.-B., Gupta, S., Gower, A.C., Shen, S.S., Varma, S., Lakshminarayanan, M., Cao, Y., Spira, A., Volkert, T.L., and Ramirez, M.I. (2012a). Genome-wide analyses of Nkx2-1 binding to transcriptional target genes uncover novel regulatory patterns conserved in lung development and tumors. PLoS ONE *7*, e29907.

Tagne, J.-B., Gupta, S., Gower, A.C., Shen, S.S., Varma, S., Lakshminarayanan, M., Cao, Y., Spira, A., Volkert, T.L., and Ramirez, M.I. (2012b). Genome-wide analyses of Nkx2-1 binding to transcriptional target genes uncover novel regulatory patterns conserved in lung development and tumors. PLoS ONE *7*, e29907.

Takabayashi, S., Aoshima, T., Kabashima, K., Aoto, K., Ohtsuka, M., and Sato, M. (2018). i- GONAD (improved genome-editing via oviductal nucleic acids delivery), a convenient in vivo tool to produce genome-edited rats. Sci Rep *8*, 1–12.

Teramoto, M., Sugawara, R., Minegishi, K., Uchikawa, M., Takemoto, T., Kuroiwa, A., Ishii, Y., and Kondoh, H. (2019). The absence of SOX2 in the anterior foregut alters the esophagus into trachea and bronchi in both epithelial and mesenchymal components. BioRxiv 739714.

Torisawa, T., Ishihara, S., and Oiwa, K. (2019). Tubulin Polymerization-Promoting Protein Family Member 3 (Tppp3) Facilitates Microtubule Bundling and Network Formation via its Weak Interaction with Microtubules. Biophysical Journal *116*, 257a.

Trisno, S.L., Philo, K.E.D., McCracken, K.W., Catá, E.M., Ruiz-Torres, S., Rankin, S.A., Han, L., Nasr, T., Chaturvedi, P., Rothenberg, M.E., et al. (2018). Esophageal Organoids from Human Pluripotent Stem Cells Delineate Sox2 Functions during Esophageal Specification. Cell Stem Cell *23*, 501-515.e7.

van Tuyl, M., Liu, J., Groenman, F., Ridsdale, R., Han, R.N.N., Venkatesh, V., Tibboel, D., and Post, M. (2006). Iroquois genes influence proximo-distal morphogenesis during rat lung development. American Journal of Physiology-Lung Cellular and Molecular Physiology *290*, L777–L789.

van Veenendaal, M.B., Liem, K.D., and Marres, H.A. (2000). Congenital absence of the trachea. Eur. J. Pediatr. *159*, 8–13.

Watanabe, H., Francis, J.M., Woo, M.S., Etemad, B., Lin, W., Fries, D.F., Peng, S., Snyder, E.L., Tata, P.R., Izzo, F., et al. (2013). Integrated cistromic and expression analysis of amplified NKX2-1 in lung adenocarcinoma identifies LMO3 as a functional transcriptional target. Genes Dev. *27*, 197–210.

Weaver, M., Yingling, J.M., Dunn, N.R., Bellusci, S., and Hogan, B.L. (1999). Bmp signaling regulates proximal-distal differentiation of endoderm in mouse lung development. Development *126*, 4005.

Weidenfeld, J., Shu, W., Zhang, L., Millar, S.E., and Morrisey, E.E. (2002). The WNT7b Promoter Is Regulated by TTF-1, GATA6, and Foxa2 in Lung Epithelium. J. Biol. Chem. *277*, 21061–21070.

Williamson, K.A., Hever, A.M., Rainger, J., Rogers, R.C., Magee, A., Fiedler, Z., Keng, W.T., Sharkey, F.H., McGill, N., Hill, C.J., et al. (2006). Mutations in SOX2 cause anophthalmia-esophageal-genital (AEG) syndrome. Hum. Mol. Genet. *15*, 1413–1422.

Yamaguchi, T., Yanagisawa, K., Sugiyama, R., Hosono, Y., Shimada, Y., Arima, C., Kato, S., Tomida, S., Suzuki, M., Osada, H., et al. (2012). NKX2-1/TITF1/TTF-1-Induced ROR1 is required to sustain EGFR survival signaling in lung adenocarcinoma. Cancer Cell *21*, 348–361.

Yuan, B., Li, C., Kimura, S., Engelhardt, R.T., Smith, B.R., and Minoo, P. (2000). Inhibition of distal lung morphogenesis in Nkx2.1(−/−) embryos. Developmental Dynamics *217*, 180–190.

Zhou, W., Li, J., Wang, X., and Hu, R. (2010). Stable knockdown of TPPP3 by RNA interference in Lewis lung carcinoma cell inhibits tumor growth and metastasis. Mol Cell Biochem *343*, 231–238.

Zhu, N.L., Li, C., Xiao, J., and Minoo, P. (2004). NKX2.1 regulates transcription of the gene for human bone morphogenetic protein-4 in lung epithelial cells. Gene *327*, 25–36.

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.

DocuSigned by:
Akela Kuwahara
C43665CA96C3461...
Author Signature

2/21/2020
Date