Université de Montréal

Induced Pluripotent Stem Cells as Modeling Tools to Understand Esophagus Development and Diseases

Par

Suleen Raad

Programmes de Biologie Moléculaire

Faculté de médecine

Thèse présentée en vue de l'obtention du grade de Philosophie Doctorat (PhD)

En Biologie Moléculaire, option générale

Juillet, 2022

Résumé

L'œsophage et la trachée proviennent du diverticule endodermique du tube de l'intestin antérieur au cours de l'embryogenèse. Des événements cellulaires et moléculaires bien régulés et organisés entraînent la séparation du tube de l'intestin antérieur en œsophage et trachée. Cette séparation est encore mal connue et la perturbation de ce processus se traduit par une anomalie congénitale sévère telle qu'une l'atrésie de l'œsophage avec ou sans fistule trachéo-œsophagienne (AO/FTO). L'AO/FTO est l'une des malformations congénitales gastro-intestinales les plus courantes affectant 1 naissance sur 3000. Cette malformation nécessite une intervention chirurgicale urgente à la naissance et est fréquemment associée à une morbidité à long terme. Les mécanismes sous-jacents au développement embryonnaire de l'AO/FTO sont mal compris. Les modèles animaux ont été largement utilisés pour comprendre les maladies humaines depuis des décennies et ont considérablement contribué à la compréhension du développement de l'œsophage. Cependant, des différences structurelles et morphologiques clés existent entre l'œsophage humain et animal, ce qui nécessite un modèle plus fiable pour comprendre le développement trachée-œsophagien.

Les cellules souches pluripotentes induites par l'homme ont été un outil précieux pour comprendre l'organogenèse en imitant le développement et en déchiffrant les mécanismes qui conduisent à des maladies congénitales et acquises. Cette thèse se concentre donc sur l'utilisation de cellules souches pluripotentes induites (IPS) par des patients pour déchiffrer les mécanismes de signalisation impliqués dans le développement de l'œsophage et les maladies congénitales telles que l'OA/FTO. Il étudie également l'une des maladies œsophagiennes acquises possibles, comme l'œsophage de Barrett. Nous avons orienté la différenciation des IPS saines et dérivées de patients vers différents stades de développement, tels que l'endoderme définitif, l'intestin antérieur, l'épithélium œsophagien et trachéal. De plus, l'épithélium œsophagien a été développé davantage dans un environnement tridimensionnel sans matrice pour générer des organoïdes œsophagiens matures. À chaque étape de la progression du développement, des analyses d'immunofluorescence, de qPCR et de séquençage d'ARN ont été effectuées. Nos résultats suggèrent que l'expression des marqueurs endodermiques *CXCR4, SOX17,* et *GATA4* était similaire dans les cellules différenciées des patients et des cellules saines. Cependant, au stade de l'intestin antérieur, nous avons observé une diminution significative de l'expression des gènes et des protéines du facteur transcriptionnel clé *SOX2* dans les cellules dérivées du patient. De plus, en utilisant le séquençage d'ARN à molécule unique, nous avons observé que les gènes critiques *GSTM1*, et *RAB37* impliqués dans la

morphogenèse cellulaire et associés à l'OA/FTO étaient dérégulés au stade de l'intestin antérieur dans les cellules dérivées du patient. Nous avons également observé une augmentation significative de l'expression du facteur de transcription *NKX2.1* habituellement exprimé uniquement dans les cellules trachéales, dans l'épithélium oesophagien dérivé du patient. *NKX2.1* est maintenue dans les organoïdes oesophagiens matures même après 2 mois.

Ensuite, nous voulions valider l'utilisation potentielle de nos organoïdes dérivés des IPS pour modéliser les maladies acquises de l'œsophage telles que l'œsophage de Barrett. Nous avons induit une métaplasie ou transformation épithéliale avec surexpression de BMP4 dans des organoïdes de l'œsophage sains et dérivés du patient sur une période d'un mois. Nos résultats préliminaires montrent que les organoïdes de l'œsophage dérivés des patients exprimaient des niveaux d'ARNm plus élevés de *MUC5AC*, un marqueur épithélial cylindrique par rapport au groupe sain. Cela suggère une plus grande sensibilité de l'organoïde de l'œsophage dérivé du patient aux changements epitheliales métaplasiques.

En conclusion, nous avons développé les premiers organoïdes œsophagiens tridimensionnels matures sans matrice différenciés des patients OA/FTO et identifié une signature moléculaire unique dans les cellules dérivées du patient au cours de la différenciation dirigée de l'œsophage. De plus, sur la base des résultats préliminaires, nous avons pu confirmer l'incidence plus élevée de l'œsophage de Barrett chez les patients OA/FTO par rapport au groupe sain.

Notre travail met donc en évidence l'importance de l'utilisation des IPS dérivées des patients pour modéliser les maladies œsophagiennes congénitales et acquises afin de fournir de nouvelles informations sur le développement des organes au cours de l'embryogenèse.

Mots clés : Atrésie œsophagienne/Fistule trachéo-œsophagienne, cellules souches pluripotentes induites, organoïdes œsophagiens, œsophage de Barrett.

Abstract

The esophagus and trachea originate from the endodermal diverticulum of the anterior foregut tube during embryogenesis. Well-regulated and organized cellular and molecular events result in the compartmentalization of the anterior foregut tube into the esophagus and trachea. This compartmentalization is still poorly understood and disruption in this process results in a severe congenital anomaly such as esophageal atresia with or without tracheoesophageal fistula (EA/TEF). EA/TEF is one of the most common gastrointestinal congenital defects affecting 1 in 3,000 births. This malformation requires urgent surgery at birth and is frequently associated with long-term morbidity. The mechanisms underlying the embryonic development of EA/TEF are poorly understood. Animal models have been widely used to understand human diseases for decades and have significantly contributed to the understanding of esophageal development. However, key structural and morphological differences exist between human and animal esophagus, thus necessitating a more reliable model to understand trachea-esophageal development. Human induced pluripotent stem cells (iPSC) have been a valuable tool to understand organogenesis by mimicking development and deciphering mechanisms that lead to congenital and acquired diseases. This thesis therefore focuses on the use of patient-derived induced pluripotent stem cells to decipher signaling mechanisms involved in esophageal development and congenital diseases such as EA/TEF. It also focuses on one of the possible acquired esophageal diseases, namely, Barrett's esophagus. We directed the differentiation of healthy and patient-derived iPSCs toward different developmental stages, such as definitive endoderm, anterior foregut, esophageal and tracheal epithelium. Furthermore, the esophageal epithelium was matured further in a matrix free 3-dimensional environment to generate mature esophageal organoids. At each stage of development progression, immunofluorescence, qPCR, and RNA sequencing analysis were performed. Our findings suggest that the expression of endodermal markers *CXCR4, SOX17,* and *GATA4*, were similar in both patient and healthy differentiated cells. However, at the anterior foregut stage, we observed a significant decrease in the gene and protein expression of key transcription factor SOX2 in patient-derived cells. Furthermore, using nanopore RNA sequencing, we observed that critical genes *GSTM1,* and *RAB7* involved in cellular morphogenesis and associated with EA/TEF to be dysregulated at the anterior foregut stage in patient-derived cells. We also observed a significant increase in the expression of transcription factor *NKX2.1*, usually

expressed only in tracheal cells, in the patient-derived esophageal epithelium. NKX2.1 expression was maintained in matured esophageal organoids even after 2 months.

Next, we wanted to validate the potential use of our PSC-derived organoids to model acquired esophagus diseases such as Barrett's esophagus (BE). We induced epithelial metaplasia with BMP4 overexpression in healthy and patient-derived esophagus organoids over a 1-month period. Our preliminary results show that patient-derived esophagus organoids expressed higher mRNA levels of *MUC5AC*, an epithelial columnar marker compared with the healthy group. This suggests a higher susceptibility of patient-derived esophagus organoid to metaplastic changes.

In conclusion, we developed the first matrix free mature 3-dimensional esophageal organoids differentiated from EA/TEF patient-derived and identified a unique molecular signature in patient derived cells during directed esophagus differentiation. Furthermore, based on the preliminary results, we could confirm the higher incidence of Barrett's esophagus in EA/TEF patients compared with the healthy group.

Our work therefore highlights the significance of using patient-derived iPSCs to model congenital and acquired esophageal diseases to yield new insights on organ development during embryogenesis. It lays the foundation for a personalized medical approach to other diseases and the ones affecting the whole gastrointestinal system in both children and adults.

Keywords: Esophageal atresia/Tracheoesophageal fistula, induced pluripotent stem cells, esophageal organoids, Barrett's esophagus.

Table of Contents

List of Figures and Tables

List of Abbreviations

AEG:	Anophthalmia-Esophageal-Genital
AF:	Anterior Foregut
AFE:	Anterior Foregut Endoderm
ASCIZ:	ATM Substrate Chk2-interacting Zn^{2+} finger protein
AXIN2:	Axin-related protein 2
BARX1:	Barx homeobox 1
BE:	Barrett's esophagus
BMP:	Bone Morphogenetic Protein
CDX1:	Caudal Homeobox 1
CDX2:	Caudal Homeobox 2
CHARGE:	Colomba, Heart defect, Atresia choanae, Restricted growth and development, Gential abnormality, Ear abnormality
CHD7:	Chromodomin Helicase DNA binding Protein 7
CMYC:	MYC proto-oncogene
CXCR4:	Chemokine receptor type 4
D/V:	Dorsal/Ventral
DE:	Definitive Endoderm
EA/TEF:	Esophageal Atresia/ Tracheoesophageal Fistula
EA1C5:	Esophageal Atresia Patient 1 Clone 5
EA2C5:	Esophageal Atresia Patient 2 Clone 5
EA3C6:	Esophageal Atresia Patient 3 Clone 6
EFTUD2:	Elongation Factor Tu GTP Binding Domain Containing 2
EPCAM:	Epithelial Cell Adhesion Molecule
EPCs:	Esophageal progenitor cells
ESCs:	Embryonic Stem Cells
FGF:	Fibroblast growth factors
FOXA2:	Forkhead Box protein A2
FOXF1:	Forkhead Box F1
GATA4:	GATA binding protein 4

GERD:	Gastroesophageal reflux disease
GLI 2:	GLI Family Zinc Finger 2
GLI3:	GLI Family Zin Finger 3
GSTM1:	Gluathione S-transferase Mu 1
HH:	Hedgehog
Ihh:	Indian Hedgehog signaling molecule
INV:	Involucrin
iPSCs:	induced Pluripotent stem cells
ISL1:	ISL LIM homeobox 1
ITGb:	Integrin Subunit Beta 1
KLF4:	Kruppel-like Factor 4
KRT 4:	Keratin 4
KRT 7:	Keratin 7
KRT 8:	Keratin 8
KRT 13:	Keratin 13
MEC:	Midline Epithelial Cells
MYCN:	MYCN Proto Oncogene
NANOG:	Nanog homeobox
NCC:	Neural Crest Cells
NKX2.1:	NK Homeobox 1
OCT3/4:	Octamer-binding transcription factor 3/4
OTX2:	Orthodenticle Homeobox 2
P63:	Tumor protein 63
PAX9:	Paired Box 9
PBMCs:	Peripheral blood mononuclear cells
RA:	Retinoic Acid
RAB37:	Ras related protein Rab-37
RAR:	Retinoic Acid Receptor
RNA:	Ribonucleic acid
Shh:	Sonic Hedgehog signaling molecule

SFRP1:	Secreted Frizzled Related Protein 1
SFRP2:	Secreted Frizzled Related Protein 2
SOX2:	SRY- Box Transcription factor 2
SOX9:	SRY- Box Transcription factor 9
SOX17:	SRY- Box Transcription factor 17
SSEA4:	Stage Specific Embryonic Antigen 4
TGFb:	Transforming growth factor beta
TRA1-60:	T cell Receptor Alpha locus
VACTERL:	Vertebral defects, Anal atresia, Cardiac defects, Tracheoesophageal fistula, Renal anomalies, Limb abnormalities
WNT:	Wingless-related Integration
YAP:	Yes-associated protein

This work is dedicated to my daughter Leona. You have made me stronger, better, and more fulfilled than I could have ever imagined. I hope you are proud of me.

Acknowledgment

I would like to express my sincere gratitude to my supervisor Prof. Christophe Faure. He continually conveyed a spirit of adventure and positivity in regard of research. I have been very fortunate to have an advisor who believed in me and was a source of constructive input, encouragement, support and understanding throughout my PhD.

My gratitude extends to Dr. Anu David for his treasured support. I am thankful to him for sharing his expertise and valuable guidance and encouragement during research, publications, and writing of this thesis.

For both of you Prof. Faure and Dr. David, thank you for not losing faith in me when the going got tough. It was not easy to become a mother during early stages of my PhD and ending the journey in a global pandemic while being pregnant with Eyas.

I would also like to thank Prof. Richard Bertrand for his presence, mentorship, and support. He went above and beyond in making himself available for guidance.

I am grateful to all those I have had the pleasure to work with during my PhD.

I would like to thank my parents, Hamed and Lina and my sister Farah, whose love is with me in whatever I pursue.

Most importantly, I wish to thank my husband Shady. He is my best friend, closest advisor, dedicated father, and the love of my life. He supported me every step of the way, as he always does, with great insight, humor, and love.

CHAPTER 1 – Introduction

Adapted and modified from: Genetic Mouse Models and Induced Pluripotent Stem Cells for Studying Tracheal-Esophageal Separation and Esophageal Development. Stem Cells Dev, Stem Cells Dev. 2020 Aug 1;29(15):953-966.

Raad S., David A., Faure C.

Authors' contribution: SR, AD and CF wrote the manuscript. SR illustrated the figures. AD and CF reviewed the manuscript.

1. Overview

The esophagus is a hollow muscular tube located between the pharynx and the cardia of the stomach and the only section of the gastrointestinal tract lacking any digestive or absorptive properties. The primary function of the esophagus is facilitating the passage of food from the mouth to the stomach and preventing acid refluxes. Mucus produced by the esophageal wall provides lubrication to ease the transport of ingested boli. Any residual material is propelled towards the stomach by active peristaltic contractions in the smooth muscles of the esophagus. During embryonic development, the esophagus originates from an endodermal diverticulum in the anterior region of the foregut tube. Highly regulated cellular behavior and molecular events are responsible for separating the common foregut tube into the esophagus and the trachea (Jacobs, Ku, and Que 2012), whereas epithelial biogenesis establishes proper epithelial morphogenesis of the esophagus and the trachea. Following foregut separation, the esophagus is lined with columnar cells, which then transform into a stratified squamous epithelium surrounded by two layers of smooth muscle with embedded enteric nerves and glial cells (Yu, Slack, and Tosh 2005). By contrast, the trachea is lined with pseudostratified epithelium and C-shaped cartilage rings (Perl et al. 2002). Disturbances in the crucial signaling mechanisms during embryo development may lead to severe esophageal congenital anomalies, such as esophageal atresia with or without tracheoesophageal fistula (EA/TEF) (van Lennep et al. 2019). The same signaling pathways have been associated later in life with the pathogenesis of esophageal diseases such as Barrett's esophagus that affect the epithelial cellular morphology (Peters and Avisar 2010).

2. Foregut separation and esophagus development

The esophagus and trachea originate from a common endodermal tube, the anterior foregut. At 4-5 weeks of human embryonic development, and mouse embryonic day 11.5 (E11.5), the anterior foregut is patterned down the sagittal axis, to separate dorsally into the esophagus (digestive system), and ventrally into the trachea (respiratory system) **(Figure 1)** (Jacobs, Ku, and Que 2012) and are surrounded by the splanchnic mesoderm. The bidirectional communication between the foregut endoderm and mesoderm plays a critical role in orchestrating esophageal and tracheal development (Han Lu et al. 2019). The nascent esophageal epithelium is made up of a single layer of cuboidal ciliated cells, which is gradually replaced with stratified squamous epithelial cells with a basal proliferative layer. Structural differences exist between the human and mouse esophageal epithelium such as the human esophageal epithelium is non-keratinized cell layers folded along papillae (Barbera et al. 2015), whereas, the murine esophageal epithelium is keratinized and contains fewer cell layers (Marques-Pereira and Leblond 1965). Submucosal glands are only found in the human esophageal epithelium serving as a source of progenitor cells in response to cell injury (Garman 2017). The trachea is lined with pseudostratified columnar epithelium in both humans and mice **(Figure 2).**

Very little is known about what initiates the anterior foregut compartmentalization into the esophagus and trachea and the exact molecular and cellular mechanisms involved. Several models of foregut compartmentalization have been proposed to explain this compartmentalization of the anterior foregut (Billmyre, Hutson, and Klingensmith 2015) and they all suggest a dynamic crosstalk between the endoderm and mesoderm during separation and subsequent epithelial morphogenesis: **1)** the outgrowth model (O'Rahilly and Muller 1984) , **2)** the watershed model (Sasaki, Kusafuka, and Okada 2001) and **3)** the septation model (Qi and Beasley 2000) in rat embryos. The septation model hypothesizes that separation initially occurs at the origination of lung buds and progresses in a rostral direction along the dorsal-ventral (D/V) midline to form a saddle separating the two tubes **(Figure 1B)** (Qi and Beasley 2000) and this epithelial saddle composed of cells from the lung buds and future esophagus is initiated at E9.5 foregut (Que 2015). Recent studies have shown that the fusion of lateral mesenchymal ridges mirrors the development of the esophagus and the trachea. Kim et al. (2019) observed a unique epithelial population termed midline epithelial cells (MECs) at the D/V boundary, which fuses to form a septum leading to the formation of the trachea and esophagus (Kim et al. 2019). In another study using mouse and

Xenopus models, Nasr et al. (2019) suggested that epithelial remodeling requires a small GTPase endosomal protein Rab11, and extracellular degradation allowing the mesenchyme to invade and resolve the septum using the Hedgehog/Gli signaling pathway (Nasr et al. 2019) Therefore, the septation model seems to be the most consistent with experimental evidence.

At week four of human embryogenesis (E9-9.5 in the mouse), neural crest cells (NCCs) scatter on the dorsal side of the foregut and encircle the gut. Muscle differentiation begins when NCCs enter the primitive gut (Obermayr et al. 2013) and by the end of week 6 (E14 in the mouse), NCCs have migrated within the mesenchyme surrounding the primitive esophagus and starts to form the myenteric plexus. In the *Mash1*$^{-/-}$ mouse knockout model which is characterized by the absence of enteric neurons in the anterior foregut, the development of the esophagus is not altered, suggesting that intrinsic neurons do not play a significant role in the separation of the foregut (Guillemot et al. 1993). By week 8 (E15 in the mouse), the NCCs are located external to the presumptive esophageal internal circular smooth muscle cells (Hao et al. 2016; Wallace and Burns 2005) and at birth, the lower segment of the esophagus consists of smooth muscle cells and the upper segment consists of striated muscles that originate from cranial mesoderm (Gopalakrishnan et al. 2015).

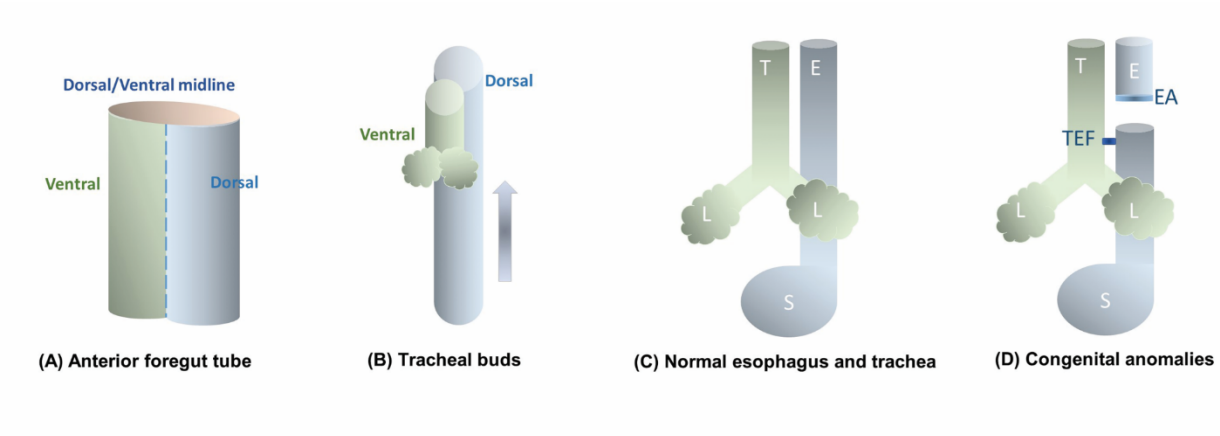


Figure 1. Development of esophagus and congenital anomalies: (A) At weeks 4-5 of human embryonic development, the anterior foregut tube is patterned down the sagittal axis to separate dorsally into the esophagus and ventrally into the trachea. (B) The ventral foregut endoderm evaginates and pushes the surrounding mesenchyme to form the two presumptive lung buds. The separation between ventral and dorsal endoderm starts and progresses in a rostral direction. (C) Intact normal separation of the esophagus in blue and trachea in green. (D) Patterning anomalies lead to the development of congenital anomalies such as esophageal atresia/Tracheoesophageal fistula. Adapted from (Raad et al. 2020)

T: Trachea, L: Lungs, E: Esophagus, S: Stomach, EA: Esophageal atresia, TEF: Tracheoesophageal fistula

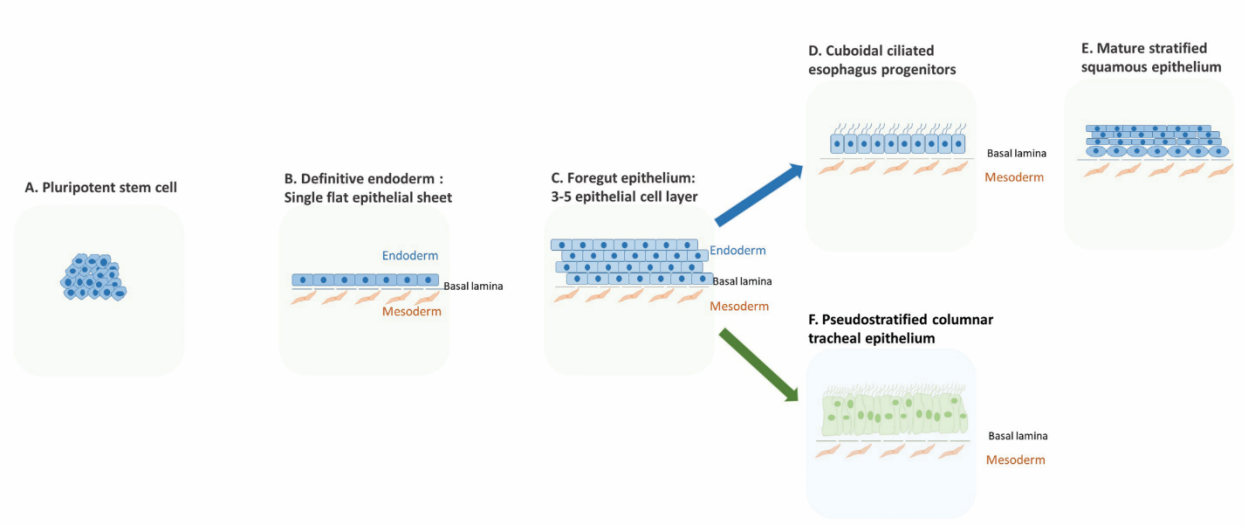


Figure 2. Epithelial morphogenesis of the esophagus and trachea: (A) Pluripotent stem cells with characteristics of high nucleus-cytoplasm ratio, prominent nucleoli, and formation of compact colonies. (B) The definitive endoderm is made up of a single flat epithelial sheet with a supporting mesoderm. (C) 3-5 epithelial cell layer makes up the foregut epithelium, which is then patterned along the D/V axis to give rise to (D) esophageal cuboidal ciliated progenitor epithelium. (E) Esophageal progenitors then further mature to become a stratified squamous epithelium with a basal proliferative layer and a suprabasal layer and (F) tracheal pseudostratified columnar tracheal epithelium. Adapted from Raad et. al (2020).

3. Signaling networks regulating foregut separation and subsequent epithelial morphogenesis

During embryonic development, the intrinsic potency and the surrounding environment determine the epithelial cell fate. Cells are organized by a process known as pattern formation, which depends on the regional specification caused by morphogen gradients of diffusing molecules. These molecules are involved in specific signaling pathways leading to the activation or inhibition of different transcription factors, thus determining cell fate and commitment. Mice esophageal development studies have shown that cells within the common foregut tube are patterned along the D/V axis to give rise to the esophagus and trachea on the dorsal and ventral side, respectively, depending on the transcription factors such as *SOX2* (dorsal side) and *NKX2.1* (ventral side). Other essential signaling molecules and pathways that are also involved in the bifurcation and subsequent epithelial morphogenesis of the esophagus and trachea include Bone

morphogenetic proteins (BMP), WNT/ß-catenin, Hedgehog, Fibroblast growth factors (FGF), and retinoic acid (RA) **(Figure 3)**.

3.1 Transcription factors SOX2 and NKX2.1: the key players in foregut specification

SOX2, a member of the SOX family, is a conserved transcription factor essential in maintaining stem cell pluripotency (Adachi et al. 2010) and is expressed in diverse and dynamic patterns during embryogenesis and also involved in early cell fate decisions (Sarkar and Hochedlinger 2013). In foregut derived organs, *SOX2* expression levels vary. It is expressed on the dorsal side of the endodermal foregut tube where the esophagus will develop, whereas *NKX2.1* is expressed on the ventral side, where the trachea arises **(Figure 3A).** The temporal and spatial regulation of expression of *SOX2* and *NKX2.1* in the dorsal and ventral regions of the foregut promotes proper foregut separation into the esophagus and trachea (Que et al. 2007). Defects in the expression levels of *SOX2* or *NKX2.1* results in abnormal foregut compartmentalization. The critical role of *SOX2* in esophageal development has been demonstrated in mice models with anterior foregut specific *SOX2* deficiency. Conditional deletion of *Sox2* in $Foxa2^{CreER}; Sox2^{loxp/loxp}$ mice, results in the loss of the esophagus (Teramoto et al. 2020a). Furthermore, when SOX2 is reduced to 30% of normal levels, more than 50% of mouse mutants displayed esophageal anomalies like EA/TEF (Que et al. 2007). In *NKX2.1* null mice, the foregut fails to separate, and the resulting tube is of esophageal identity with *SOX2* expressed throughout the foregut tube (Minoo et al. 1999; Que et al. 2007). Trisno et al. 2018 investigated the role of *SOX2* in esophagus specification and found that *SOX2* inhibits WNT signaling by promoting the expression of WNT inhibitors such as *SRRP1, SFRP2 and DKK1*. The WNT signaling as discussed below, is an essential pathway in promoting respiratory fate specification. Furthermore, they also observed that in *Sox2*-deletion mouse mutants, transcript levels of downstream target *Axin2* was upregulated in the dorsal foregut (Trisno et al. 2018). This supports the view on the critical role of *SOX2* in specifying esophagus fate during foregut development.

It is hypothesized that *SOX2* and *NKX2.1* repress each other; however, the mechanisms involved in this repression remain unclear (Jacobs, Ku, and Que 2012; Edwards et al. 2021). Recently it was demonstrated that *NKX2.1* binds silencer sequences near the *SOX2* gene and represses its expression. *NKX2.1* was also shown to directly repress other esophageal genes such

as *Klf5* (Kuwahara et al. 2020). However, more than 90% of respiratory-specific transcripts did not depend on *NKX2.1* expression levels, suggesting that other genes are yet to be explored (Kuwahara et al. 2020). Kim et al. (2019) observed that *Isl1* regulates the transcription of *NKX2.1*. The studies therefore suggest that proper D/V patterning and the expression of these transcription factors are essential to initiate normal separation of the foregut tube (Kim et al. 2019; Nasr et al. 2019; Trisno et al. 2018; Teramoto et al. 2020b).

Following foregut separation, the esophagus is lined with squamous epithelium and the trachea with columnar epithelium (Billmyre, Hutson, and Klingensmith 2015). The differentiation of the epithelium towards a columnar or a squamous fate is regulated by the expression of *SOX2*, tumor protein 63 (*p63*), and *NKX2.1* (Que et al. 2007; Que et al. 2006). *SOX2* and *p63* are responsible for squamous epithelial differentiation (Jiang et al. 2017), whereas *NKX2.1* is responsible for columnar epithelium (Jiang et al. 2017; Yuan et al. 2000). During esophageal development, high levels of *SOX2* stimulate the formation of stratified squamous epithelium and prevent pseudostratified columnar epithelium, lining the trachea. *SOX2* and *p63* are expressed in the basal proliferative layer of the stratified squamous epithelium, with lower expression levels in the suprabasal layers. Decreased *SOX2* expression has been linked to the downregulation of *p63* (Que et al. 2007) which is involved in the proliferation of the basal layer and the stratification of the epithelium. *P63* null mice have an esophagus lined with simple columnar epithelium lacking basal cells (Daniely et al. 2004) and Bailey et al. (2019) recently showed the critical role of the Yes-associated protein (*YAP*) in self-renewal and stratification of the epithelium to a multilayered stratified squamous epithelium (Bailey et al. 2019).

3.2. The BMP pathway promotes and spatially restricts respiratory fate in ventral anterior foregut epithelium (AFE)

Bone morphogenetic proteins (BMPs) belong to the Transforming growth factor-ß (TGF-ß) superfamily and are involved in stem cell proliferation and differentiation (Guasch et al. 2007; Oshimori and Fuchs 2012; Rahman et al. 2015). Signaling is normally initiated by ligand-induced oligomerization of serine/threonine receptor kinases. This results in phosphorylation of Smad2 and Smad3, cytoplasmic signaling molecules for the TGF-ß pathway or SMAD1/5/8 for the BMP pathway. Activated SMADs form a complex with SMAD 4 and then translocate to the nucleus

and regulate gene expression by partnering with transcription factors to modulate specific transcription. Dual inhibition of BMP/TGF-ß promotes the specification of anterior foregut endoderm to dorsal foregut endoderm, which further differentiates to esophagus (Zhang et al. 2018). BMP4 is mostly expressed in the ventral mesenchyme of the foregut tube right before separation, whereas *Noggin*, a BMP antagonist, is highly expressed in the dorsal endoderm (Que et al. 2006) **(Figure 3).** Conditional ablation of *BMP4* or BMP type 1 receptors (Domyan et al. 2011; Li et al. 2007) results in tracheal agenesis without alteration of *NKX2.1* expression (Li et al. 2007), but an expansion of *SOX2* expression in the ventral endoderm (Que 2015; Domyan et al. 2011). This also suggests the role of *BMP4* in inhibiting digestive fate and promoting respiratory fate in the ventral AFE (Domyan et al. 2011). Specifically, BMP signaling downstream effectors SMAD1/5/8 directly inhibit *SOX2* expression in the ventral foregut epithelium whereas the dorsal foregut secretes Noggin to restrict BMP activity and maintain *SOX2* expression (Domyan et al. 2011). Noggin misexpression results in failed foregut separation (Pinzon-Guzman et al. 2020) and ablation of *BMP7* (Li et al. 2007) or of *BMP4* function (Que et al. 2006) in *Noggin*-/- mice rescued unsuccessful foregut segregation. Moreover, in *Noggin*-/- mutants, the fistula connecting the trachea and stomach was made up of *NKX2.1* positive epithelial cells (Que et al. 2006). Noggin is also shown to be required for early notochord formation and delamination (Fausett, Brunet, and Klingensmith 2014). Dorsally, *Noggin* attenuates BMP signaling in the resolving notochord, a critical event for proper separation of the foregut tube (Fausett, Brunet, and Klingensmith 2014). Therefore, *BMP4* inhibits digestive fate and restricts the site of primary bud formation in the ventral AFE, whereas *Noggin* plays a vital role dorsally in proper notochord delamination during foregut separation (Domyan et al. 2011; Fausett, Brunet, and Klingensmith 2014).

After the esophagus is established from the foregut, inhibition of BMP by *Noggin* is crucial to allow the stratification of the squamous epithelium **(Figure 2E)**. Noggin levels in the squamous epithelium decrease once esophageal fate is determined. This results in the activation of the BMP pathway in the suprabasal differentiated layer, decreasing *SOX2* and *p63* expression levels (Zhang et al. 2017) with increasing levels of mature epithelial markers such as loricrin and Involucrin. Inhibition of BMP receptor1a results in a multilayered undifferentiated epithelium, lacking suprabasal markers' expression (Rodriguez et al. 2010). Guyot et al. also showed that the inhibition of the TGF-ß pathway along with the BMP pathway further promotes esophageal fate and basal esophageal cell expansion *in vitro* (Guyot and Maguer-Satta 2016). This suggests the dynamic role

of the BMP pathway during esophageal epithelium development and maturation which involved BMP inhibition for the stratification of the esophageal epithelium followed by its activation for the differentiation of the proliferative basal cells into a suprabasal differentiated layer.

3.3. The WNT pathway

WNT pathway is a conserved pathway that comprises a large family of 19 proteins in humans that include the canonical (WNT/β-catenin dependent) or the non-canonical (β-catenin independent) pathways. In the canonical WNT/β-catenin signaling pathway, when WNT ligands are absent, cytoplasmic β-catenin is phosphorylated by the destruction complex formed by three proteins: APC, AXIN, and GSK3β. Phosphorylated β-catenin is recognized by E3 ubiquitin ligase β-Trcp and is degraded. B-catenin is thus maintained at low levels in the cytoplasm. Signaling is initiated when WNT proteins interact with receptors of FZD family such as LRP5/6 or ROR2 to induce intracellular signaling pathway which passes the signal to the Dishevelled protein. In this context, β-catenin destruction complex is disrupted which prevents β-catenin proteasomal degradation. This will allow β-catenin to accumulate and localize to the nucleus to induce cellular response through gene transduction with the TCF/LEF transcription factors. Canonical WNT pathway is necessary for respiratory fate in the anterior foregut promoting lung development while inhibiting esophageal cell fate. The WNT ligands, WNT2, and WNT2b are expressed in the mesenchyme surrounding the ventral foregut **(Figure 3).** Suppressed Wnt signaling leads to lung and tracheal agenesis affecting *NKX2.1* expression in the ventral AFE (Goss et al. 2009; Harris-Johnson et al. 2009). *Wnt5a* has been shown to regulate fibroblast growth factor 10 (*Fgf10*) expression during lung formation and was shown to be upstream of *Sox2*, negatively regulating its expression (Volckaert et al. 2013). Continuous activation of the WNT pathway leads to the loss of *SOX2* expression needed for dorsal patterning and an increase in *NKX2.1* expression (Goss et al. 2009). However, the attenuation of the WNT pathway by SFRP1, SFRP2, and WNT inhibitors, ensures *SOX2* expression in dorsal foregut promoting esophageal fate (Trisno et al. 2018) (Woo et al. 2011) showed that WNT pathway is inhibited by the mesenchymal transcription factor, *Barx1*. However, Wnt/Fgf crosstalk is involved in the specification, regulation, and development of the basal cells and cartilage during tracheal development (Hou et al. 2019).

3.4. Hedgehog signaling pathway

Sonic Hedgehog (Shh) signaling pathway regulates important events during development and is essential for embryonic development, adult tissue maintenance, and regeneration. Shh proteins act in a concentration and time-dependent manner binding to canonical receptor Patched and its co-receptors, mediating downstream signal transduction, including dissociation of the transcription factors Gli2 and Gli3 proteins (Falkenstein and Vokes 2014). In the absence of a Shh signal, Gli2 is degraded, and Gli3 is cleaved into a transcriptional repressor. Shh ligands are expressed in the mouse ventral foregut endoderm at E9.5-E10.5 and mediate a paracrine communication to activate Gli2 and Gli3 in the surrounding mesoderm before shifting to the dorsal epithelium at E11.5 (Rodriguez et al. 2010). The role of Shh signaling in foregut development was first observed in Shh-mutant mice where the foregut fails to separate correctly, and the resulting esophagus is hypoplastic (Litingtung et al. 1998). Mahlapuu et al. (2001) suggested that SOX2/Shh signaling occurs via Foxf1 to support the development of an intact esophagus during foregut separation. Loss of Foxf1 has been linked to TEF and the development of small lungs. In mice with inactivation of Gli2 and low dosage of Gli3 (*Gli2* $^{-/-}$ *and Gli3* $^{+/-}$ mutants), a hypoplastic foregut was similarly observed, with a small esophagus and hypoplastic trachea and lungs (Motoyama et al. 1998). This occurs because Gli2 and Gli3 block the expression of Wnt2, Wnt2b and BMP4 in the surrounding mesenchyme resulting in failed respiratory specification and foregut separation (Rankin et al. 2016). Nasr et al. (2019) demonstrated that Shh/Gli pathway plays a critical role in the tracheoesophageal separation and initiates medial constriction by condensation of *Foxf1* mesodermal cells at the Sox2-Nkx2.1 boundary **(Figure 3B).**

3.5. Retinoic acid and Fibroblast growth factor 10

RA, an active derivative of Vitamin A, is known as a master regulator of patterning and differentiation of many organs. RA has been linked to posteriorizing the foregut and disrupting of RA results in abnormal development of posterior foregut organs (Wang et al. 2006; Rhinn and Dolle 2012). In *RARα1*$^{-/-}$ *β*$^{-/-}$ mutants, the midline septum in the foregut tube was absent, and the resulting lungs were hypoplastic (Mendelsohn et al. 1994; Luo et al. 1996). RA from foregut splanchnic mesoderm stimulates the expression of hedgehog ligands in the foregut endoderm epithelium. Hedgehog ligands Shh and Ihh, then signal back to the mesoderm to activate *Gli*

transcription factors that induce Wnt2 and Wnt2b, BMP4 in the ventral splanchnic mesoderm (Rankin et al. 2016). RA signaling has also been shown to induce the expression of *Fgf10*, an essential morphogen in tracheal differentiation (Chen et al. 2010) and is detected as early as E9.5 during primary lung bud formation in the anterior mesenchyme surrounding the future trachea and is highly expressed in budding sites during lung development in mice (Volckaert and De Langhe 2015; Chen et al. 2010) **(Figure 3A).** FGF10 has been shown to upregulate the expression of *NKX2.1* and downregulate *SOX2* (Que et al. 2007). *Fgf10* also influences the expression of other FGFs, BMPs, Wnts, and Shh during development while regulating polarity, adhesion, and directed cell migration indicating its critical role during foregut patterning (Yuan et al. 2018). In *Fgf10*-null mutant mice, normal separation of trachea and esophagus occurs despite the absence of primary bud lung formation, which could mean that separation of the foregut is not initiated by primary bud induction (Sala et al. 2011). RA has been shown to regulate the expression of Shh in the anterior side of the foregut tube before lung bud formation (Rankin et al. 2016) and cross-talks between factors such as RA, Shh, and *NKX2.1* are essential for foregut separation and subsequent trachea/lung development.

The mechanisms that regulate these signaling pathways, downstream transcription factors, and modulators responsible for foregut development and generation of a mature functional esophagus remain poorly understood.

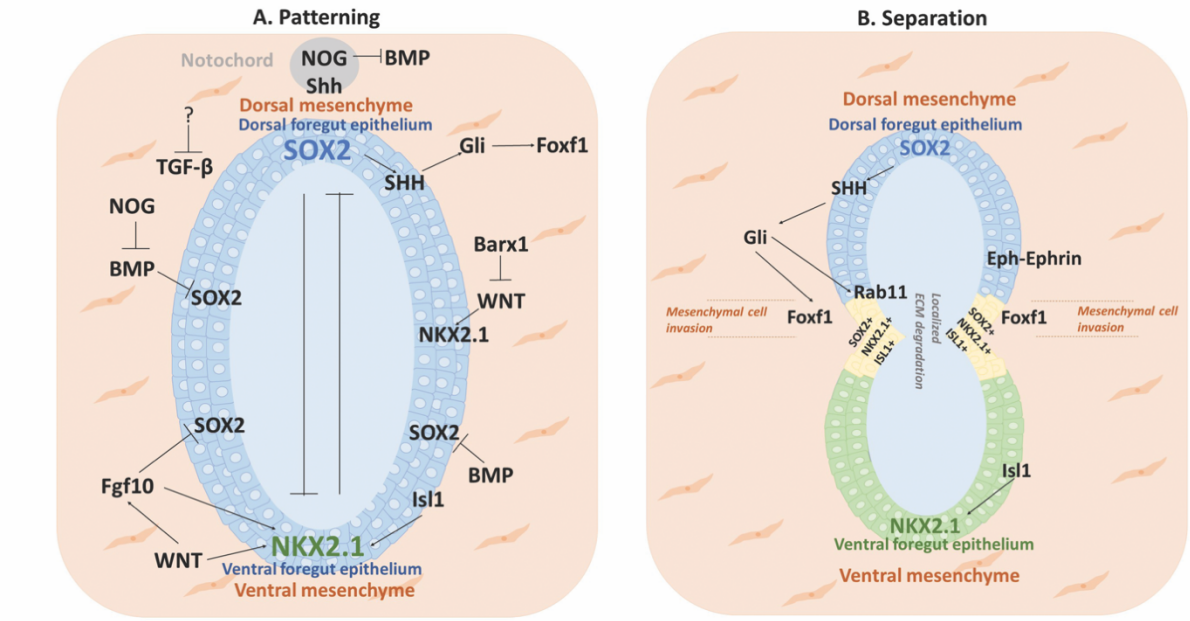


Figure 3. Signaling networks responsible for D/V patterning and subsequent separation of the anterior foregut into either esophageal or respiratory fate: (A) Patterning of the anterior foregut tube is regulated by a complex communication network between the mesenchyme (in orange) and the endoderm (in blue or green). Main signaling pathways such as BMP, WNT, FGF, SHH are vital in establishing patterning in the AFE. Any disruptions of these signaling pathways may lead to segregation defects. (B) Following patterning, segregation of the foregut tube is established by a specific epithelial population SOX2+NKX2.1+ ISL1+ orchestrated by mesenchymal cell invasion and localized extracellular matrix degradation allowing the intact separation of the foregut tube into the esophagus on the dorsal side and the trachea on the ventral side. Adapted from Raad et. al (2020).

4. Esophagus congenital anomaly: Esophageal atresia/tracheoesophageal fistula

EA/TEF is the most frequent congenital anomaly of the esophagus affecting 1 in 2,500 to 3,000 births (Nassar et al. 2012) where the upper segment is not connected to the lower part of the esophagus and the stomach. In most cases, EA is accompanied by an abnormal connection between the esophagus and the trachea, known as a fistula **(Figure 4).** This malformation requires urgent surgery with long-term morbidity, necessitating a lifelong follow-up (Cartabuke, Lopez, and Thota 2016; Krishnan et al. 2016). Several types of EA/TEF have been described based on the location of the malformation and the affected structures with the most common type being type C (80%), where the upper segment of the esophagus ends in a blind pouch, and a fistula connects the lower part to the trachea. Other less common subtypes include type A (8-10%) where no fistula exists, but the esophagus is disconnected, and type H (4%) where the esophagus connects typically to the stomach, but a TEF is connecting the esophagus to the trachea. The rarest malformations are types B and D. In type B, the upper segment of the esophagus is connected to the trachea, and the lower segment ends in a blind pouch whereas, in type D, TEF exists in the upper and lower part of the esophagus affecting only 1% **(Figure 4).** EA/TEF could be due to multiple factors, including genetics, environmental factors, fetal alcohol exposure, parental age, and maternal diabetes (Felix et al. 2008; Correa et al. 2008).

EA/TEF associated anomalies are reported in 30-50% of the cases (Galarreta, Vaida, and Bird 2020; van Lennep et al. 2019). Monogenic causes account for a minority of EA/TEF cases (5%) most often found in syndromic cases such as Anophthalmia-Esophageal-Genital (AEG) syndrome

(SOX2 mutations), Feingold syndrome (MYCN mutations), CHARGE syndrome (CHD7 mutations), Pallister-Hall syndrome (GLI3 mutations) and Mandibulofacial dysostosis (EFTUD2 mutations) (Stoll et al. 2009). Other associated anomalies include cardiac, anal, renal, limb, or vertebral malformations, which can be associated with the VACTER/VACTERL syndrome suggesting common development mechanisms between the different organs. Anomalies such as right aortic arch, aberrant subclavian artery, or vascular rings with double aortic arch, with or without cardiac malformations, are also frequently found in EA/TEF patients (Berthet et al. 2015). EA/TEF is frequently associated with congenital esophageal stenosis in the distal third of the esophagus (McCann et al. 2015). Tracheobronchial remnants are often found in such stenosis suggesting a localized abnormal dorsal/ventral patterning of the foregut. Finally, EA/TEF patients also have an increased prevalence of laryngeal clefts, which may be related to the abnormal development of the trachea-esophageal septum (Londahl et al. 2018).

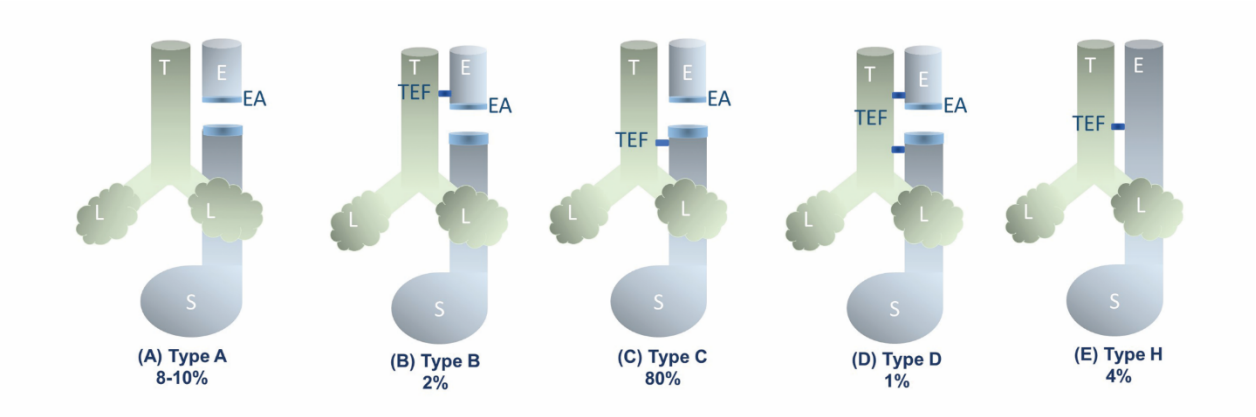


Figure 4. Types of Esophageal atresia and tracheoesophageal fistula: (A) Type A, a less common subtype of EA/TEF where the esophagus is discontinued without connecting the trachea. (B) Type B, one of the least common subtypes of EA/TEF where a fistula connects the upper part of the esophagus to the trachea, and the esophagus end in a blind pouch. (C) The most common type of EA/TEF occurring in 80% of the cases where the esophagus is discontinued and TEF occurs in the lower section of the esophagus. (D) Type D is the least common subtype of EA/TEF, where there is esophageal atresia with two fistulas between the trachea and the lower and upper esophageal pouches. (E) Type H is one of the less common subtypes of EA/TEF occurring in 4% of the cases where the esophagus connects typically to the stomach with a fistula connecting the esophagus to the stomach. Adapted from Raad et. al (2020).

T: trachea, L: lungs, E: esophagus, S: stomach, EA: esophageal atresia, TEF: tracheoesophageal fistula

4.1. Genes associated with developmental anomalies of the esophagus

Most cases of EA/TEF have not been associated with gene mutations. However, evidence suggests that EA/TEF may have a genetic background: **(a)** In twins, the concordance rate is higher in monozygotic twins with isolated EA than in dizygotic twins (Schulz et al. 2012); **(b)** although rare, hereditary EA/TEF has been reported (O'Rahilly and Muller 1984; Van Staey et al. 1984); **(c)** EA/TEF can be associated with single-gene disorders and karyotype anomalies (deletions or duplications of chromosomes involving genes implicated in trachea-esophagus development (de Jong et al. 2010; Felix et al. 2008; Brosens et al. 2014); **(d)** genetically modified mice that mimic EA/TEF have shown candidate genes in human esophageal development. Mouse knockout models have been used to mimic the genetics and morphogenetic regulations of foregut bifurcation and subsequent epithelial morphogenesis corresponding to human esophageal anomalies (Jacobs, Ku, and Que 2012). Mouse models of abnormal foregut development consist of two main categories depending on whether esophagus or trachea is the most affected. **Table 1** shows different foregut compartmentalization defects using rodent models with variant degrees of abnormal development of the esophagus and trachea. The first includes abnormal separation of the esophagus with variable hypoplasia in the esophagus leading to EA with or without TEF such as *Sox2*$^{GFP/COND}$, endodermal deletion of *SOX2, Noggin*$^{-/-}$, *Gli2,* and *Gli3 mutants*, *Shh* $^{-/-}$ *and Foxf1*$^{-/-}$. The other group encompasses impaired outgrowth of the respiratory tract with various levels of hypoplastic trachea and lungs, including *NKX2.1*$^{-/-}$, *Wnt2/2b* $^{-/-}$, *Ctnnb1* $^{flox/flox:\ Shh\text{-}Cre}$, *BMP4*^{condKO}, *BMPr1a*$^{-/-}$, *Gli 2* and *Gli3* mutants, and *Shh*$^{-/-}$ (**Table 1**).

In humans, except in the rare monogenic cases described above, the genetics in most of EA/TEF patients remain unexplained. Recent studies showed a multigenic architecture of rare variants in several genes, which discriminate EA/TEF cases from controls (Wang et al. 2021). However, the cause EA/TEF remains largely unknown and rare genetic variants are seldom reported in non-syndromic, isolated cases. Recently, a whole genome sequencing study observed a significant number of potential genes involved in EA/TEF. In this study of 185 EA/TEF patients, they identified novel risk genes in biological pathways linked to endocytosis, membrane dynamics, and intracellular transport such as *itsn1, ap1g2,* and *rab3gap2* (Zhong et al. 2021). Another study identified 19 de novo variants including a frameshift variant in EFTUD2 that has been previously reported to be mutated in 25 other EA/TEF patients (Wang et al. 2021; Gordon et al. 2012).

Table 1. Different foregut compartmentalization defects using mouse models with variant degrees of abnormal development of the esophagus and trachea. The specification refers to the specific expression of *SOX2* on the dorsal side and *NKX2.1* on the ventral side of the anterior foregut tube before separation. In contrast, septation refers to the physical separation of the anterior foregut into two distinct tubes, the esophagus, and the trachea. The esophagus and trachea indicate to their normal development. TEF indicates the presence of esophageal atresia or tracheoesophageal fistula described in the study. Lungs refer to the complete or incomplete (hypoplasia) development.

Abbreviations and symbols: NR Not reported, NA Not applicable because of the presence of a single tracheoesophageal tube without any septation, *Hypoplasia of the esophagus and TEF dose-dependent of *Sox2* expression, **Early Endodermal Deletion (E6.5) of *Sox2* Results in Esophageal Agenesis and Deletion of *Sox2* after initiation of esophageal development (E9.5 results in esophageal hypoplasia, ***Elevated *SOX2* expression in *NKX2.1* null embryo shown in Que 2007, **** Abnormal notochord delamination but septation preserved, *****not characterized, $ Columnar respiratory epithelium along the ventral surface, $$ Normal expression of *SOX2* in the dorsal foregut and abnormal expression of *NKX2.1* in the ventral foregut with ectopic expression of S*ox2* , $$$ The common tube is lined by columnar epithelium, # Tracheal agenesis related to decreased cell proliferation, ## *NKX2.1* expression present at E9.25 and lost afterwards with expansion of *SOX2* expression, ### Single tracheoesophageal tube that connects, to the stomach, #### Endodermal structures missing between the thymus and stomach levels.

	Specification	Septation	Esophagus	Trachea	TEF	Lungs	Reference
Normal	Yes	Yes	Yes	Yes	No	Yes	
$Sox2^{EGFP/COND}$	No	Abnormal	Hypoplastic*	Yes	Yes*	NR	Que et al. 2007
Endodermal Deletion of Sox2	No	Abnormal	Absent or Hypoplastic**	Yes	Yes*	Yes	Trisno et al. 2018 Teramoto et al. 2020
$NKX2.1^{-/-}$	No***	No	Yes	Hypoplastic	NA	Hypoplastic	Minoo et al. 1999 Que et al. 2007
Noggin $^{-/-}$	Yes	No****	Hypoplastic	Yes	Yes	Abnormal branching	Li et al. 2007 Que et al. 2006 Fausett et al. 2012
ephrin-B2 $^{lacZ/lacZ}$	NR	No	Yes*****	Yes*****	Yes	NR	Dravis et al. 2004
Barx 1 $^{-/-}$	Yes	No	Yes$	Yes	Yes	NR	Woo et al. 2011
$Wnt2/2b^{-/-}$ $Ctnnb1^{flox/flox:Shh-cre}$	No	No	Yes	No	NA	Agenesis	Goss et al. 2009; Harris-Johnson et al. 2009
Asciz $^{-/-}$ (Atmin KO)	Impaired$$	No	Yes	No	NA	Agenesis	Jurado et al. 2010
$RARα1^{-/-}β^{-/-}$	NR	No	No	Yes$$$	NA	Hypoplastic	Mendelsohn et al. 1994; Luo et al. 1996
$BMP4^{cond\ KO}$	Yes	No	Yes	Absent#	NA	Hypoplastic	Li et al. 2008
BMPr1a $^{-/-}$	Yes##	No	Yes	Absent	NA	Hypoplastic	Domyan et al. 2011
$Gli2^{-/-}$ $Gli2^{-/-}$ $Gli3^{+/-}$ $Gli2^{-/-}$ $Gli3^{-/-}$	NR NR NR	Yes No No####	Hypoplastic No### No####	Hypoplastic No### No####	No Yes NA	Hypoplastic Hypoplastic No####	Motoyama et al. 1998
$Shh^{-/-}$	NR	No	No###	No###	Yes	Hypoplastic	Litingtung et al. 1998
Foxf1 $^{-/-}$	NR	Abnormal	Hypoplastic	Yes	Yes	Hypoplastic	Mahlapuu et al. 2001

4.2. Possible mechanisms leading to EA/TEF

Defects in notochord resolution

Following gastrulation, a narrow band of cells delaminates from the endoderm to form the notochord, a major regulator of embryonic patterning (Jurand 1974). Little is known about the cellular dynamics of notochord separation from the endoderm, and it acts as a signaling center for proper patterning of the foregut, the neural tube, and surrounding tissues (Chamberlain et al. 2008). Notochord resolution from the dorsal foregut endoderm allows the endoderm to fold ventrally and form the gut tube. Once the notochord is resolved, lungs buds emerge on the ventral side of the anterior foregut tube commencing the bifurcation into the esophagus and trachea (Morrisey and Hogan 2010). Defects in notochord resolution result in EA/TEF, potentially linked to the excess of dorsal foregut endodermal tissue attached to the notochord, leaving insufficient dorsal foregut endoderm to form the esophagus (Li et al. 2007). BMP signaling attenuation is required in early notochord resolution, which seems to be a prerequisite for an intact foregut separation (Fausett, Brunet, and Klingensmith 2014). *Noggin* knockout mouse models have shown both notochord abnormalities and EA/TEF (Fausett, Brunet, and Klingensmith 2014).

Abnormal Dorsal/Ventral patterning

Dorsal/ventral (D/V) patterning defects also lead to the failure of foregut separation. The lack of fate commitment of either the esophagus or the trachea may result in an abnormal foregut separation. Multiple signaling pathways (described previously) are responsible for establishing proper D/V patterning prior to separation. The precise temporal and spatial distribution of signaling molecules within the foregut tube and surrounding mesoderm is critical for normal D/V patterning. It allows foregut separation and subsequent differentiation of epithelial progenitors into the esophageal or tracheal epithelium. For instance, *NKX2.1*$^{-/-}$, *SOX2*$^{GFP/Cond}$, *Asciz*$^{-/-}$, *Wnt2/2b*^{cond} are examples of genetic mouse models where the loss of D/V patterning results in EA/TEF along with tracheal and esophageal malformations (Minoo et al. 1999; Que et al. 2007; Jurado et al. 2010; Goss et al. 2009). **Table 1** shows different foregut compartmentalization defects using rodent models with variant degrees of abnormal development of the esophagus and trachea.

Abnormal septation

Foregut septation is a critical event in gut-respiratory development. Our knowledge of foregut compartmentalization comes mainly from mouse models. Mouse knockout models such as *Noggin*$^{-/-}$, *BMP4*^{condKO}, *and ephrin-B2*$^{lacZ/lacZ}$ fail to compartmentalize the anterior foregut leading to EA/TEF and have distinctive foregut phenotypes (Que et al. 2006; Li et al. 2008; Dravis and Henkemeyer 2011) (**Table 1**).

Of all the signaling pathways, the least studied is the Eph/ephrin pathway. Taylor et al. (2017) and O'Neill et al. 2016 revealed the role of Eph/ephrin in patterning, morphogenesis, self-organizing properties and is also implicated in multiple compartmentalization events during development (Dravis and Henkemeyer 2011). *Eph/ephrinB2*$^{LacZ/LacZ}$ embryos have foregut defects (Dravis and Henkemeyer 2011) resulting in the loss of anterior foregut compartmentalization. They also found that Eph/ephrin signaling is active specifically at the D/V midline and regulates cellular behavior during compartmentalization. D/V patterning in *ephrinB2* mutants is unknown.

Another example of mutants with failed foregut separation is *Barx1*-/- knockout mouse models. Barx1 is normally expressed in the stomach mesenchyme (Kim et al. 2005) and is required for the expression of Wnt antagonists such as secreted frizzled-related proteins 1 and 2, that suppress Wnt activity and direct towards the stomach fate (Kim et al. 2005). The deletion of *Barx1* gene results in EA/TEF with a significantly truncated esophagus (Kim et al. 2007). *Barx1* establishes esophageal identity by suppressing Wnt signaling in the surrounding mesenchyme (Woo et al. 2011). Interestingly, *Barx1* is also detected in specific cells in the dorsal foregut mesenchyme between the nascent esophagus and trachea, where the saddle is located (Woo et al. 2011).

Abnormal vascularization development

Abnormal vessel development could lead to vascular insufficiency and physical compression in the developing foregut, resulting in growth anomalies and atresia of the dorsal part of the anterior foregut. Anomalies such as right aortic arch, aberrant subclavian artery, or vascular rings with double aortic arch, with or without cardiac malformations, are frequently found in EA/TEF patients (Berthet et al. 2015). Aorta development occurs at gestation week three before the bifurcation of the foregut occurs at week 4 and is surrounded by vascular structures with a bilateral distribution.

Presumptive organs such as the trachea and esophagus are in contact with the developing heart and aortic arches that ultimately obliterate to form the aorta and great vessels (Larsen 1997). Isl1, a transcription factor involved in the multipotent and proliferative properties of the cardiac mesoderm, has been identified in the bifurcation area of the foregut tube into the esophagus and trachea (Kim et al. 2019). This suggests that common mechanisms may not only involve abnormal separation of the foregut but also vascular and cardiac malformations (Rao and Gershon 2018).

Abnormal muscle cell differentiation

Besides healthy vessel development, failure in normal esophageal development could be due to the developmental failure of cells/organs simultaneously developing with the esophageal epithelium. In type A EA/TEF, foregut separation is successful, but failure in the development of a continuous tube connecting the pharynx to the stomach can be observed. One hypothesis is that a localized abnormal smooth muscle malformation or failure of the enteric nerve cells to reach maturation in the esophagus could explain this phenomenon.

5. Acquired Esophageal disease: Barrett's esophagus

Different diseases affect the esophagus, and the most common ones include eosinophilic esophagitis (EE), gastroesophageal reflux disease (GERD), and Barrett's esophagus (BE). GERD is a normal physiological process and there are multiple mechanisms to protect from reflux such as anti-reflux barrier, esophageal clearance, mucosal resistance. However, when this reflux causes symptoms such as chest pain, a burning sensation, and complications such as difficulties in swallowing is considered a disease (GERD). A portion of GERD patients are at risk of developing BE, a precursor to esophageal adenocarcinoma. BE occurs within the distal esophagus when the stratified squamous esophagus epithelium changes into a columnar intestinal mucus secreting epithelium (Spechler and Souza 2014). This transition is believed to occur when the esophageal epithelium is continuously exposed to acids from the stomach causing inflammation. If left untreated, this metaplasia can become dysplasia and eventually cause esophageal adenocarcinoma (Behrens et al. 2011). It is widely accepted that the pathogenesis of BE begins with GERD, however, the origin of metaplastic columnar cells remains unclear. Different models for the origin

for BE have been suggested but none have been proven to be mutually exclusive and there may be more than one type of progenitor cells (Que et al. 2019). These include: **1.** transdifferentiation of esophageal epithelium into columnar epithelium, in which the mature esophageal cells fully differentiate into columnar cells. It is suspected that the differentiated cells reverse their differentiation status into a progenitor like state before changing into a columnar type (Wang 2017). **2.** Transcommitment differentiation in which progenitor cells proliferate and alter their differentiation pattern. The difference between transdifferentiation and transcommitment is that the latter starts with immature progenitor cells. It is not clear which progenitor cells gives rise to BE and these include basal cells of the esophageal epithelium (Shields et al. 1993), submucosal glands and their ducts (Epstein et al. 2017) or stomach cells such as progenitors of the gastric cardia that migrate to repair any acid-induced damage (Quante et al. 2012) or residual embryonic cells that may have persisted into adulthood and could provide progenitor cells for BE (Wang et al. 2011). Bone marrow cells that circulate in the blood stream to replace reflux-damaged esophagus epithelium are also suggested to be the possible progenitors for BE (Epperly et al. 2004) (Figure 5). However, to date, the exact cells of origin for BE remain to be known.

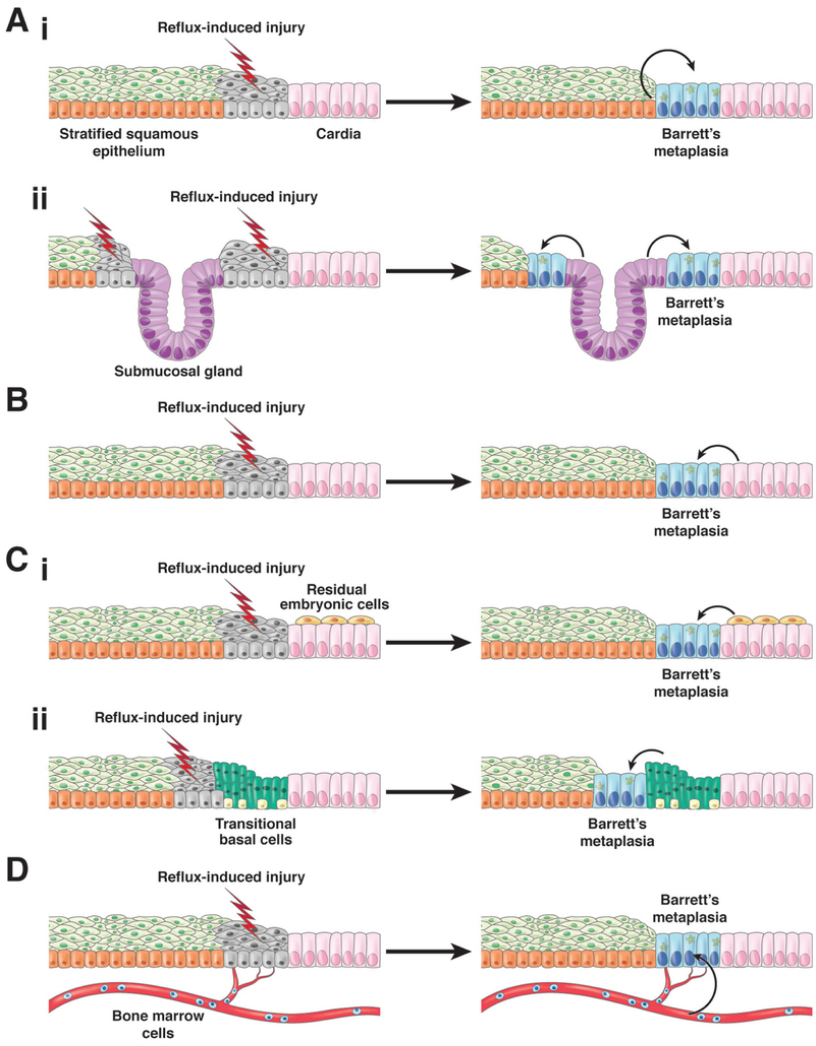


Figure 5: Proposed cells of origin for BE. A) Cells native to the esophagus including (Ai) squamous epithelial cells that undergo reflux-induced transdifferentiation or transcommitment to produce the columnar cells of Barrett's metaplasia and (Aii) Progenitor cells in esophageal submucosal glands and/or their ducts. B) Progenitor cells in the gastric cardia. C) Specialized populations of cells at the esophago-gastric junction migrate into the reflux-damaged esophagus including (Ci) residual embryonic cells (RECs) or (Cii) transitional basal cells (TBCs). D) Circulating bone marrow cells. For all of these proposed progenitor cells, reflux induced injury to the esophageal squamous mucosa is assumed to initiate the metaplastic process,by stimulating progenitor cell migration into the damaged esophagus via a wound-healing process. In addition, reflux is assumed to induce the transcommitment of the progenitor cells to produce the multiple columnar cell types of Barrett's metaplasia. Figure from Que et al. (2019). Authorization obtained from authors.

5.1. Genes and signaling pathways of Barrett's esophagus

The BMP pathway is not normally activated in the squamous esophageal epithelium. If inflammation occurs due to bile salts and acid, the BMP pathway is activated even before the

development of BE (van Baal et al. 2008; Quante et al. 2012). In fact, BMP4 and downstream SMAD proteins 1/5/8 were present in squamous epithelium with mild inflammatory esophagitis (Quante et al. 2012). *In vitro* studies with human esophageal cells treated with components mimicking gastroesophageal reflux resulted in an increased BMP4 expression (Que et al. 2006; Quante et al. 2012; Zhou et al. 2009). In BE stromal cell biopsies and BE animal models, BMP4 and activated SMAD proteins 1/5/8 were detected (Sasaki et al. 2004; van Baal et al. 2008). A shift in the cytokeratin expression was also observed when primary squamous esophageal epithelium was treated with BMP4 such as Krt13 and 14 expressing squamous epithelium became Krt7 and 8 expressing columnar epithelium (Quante et al. 2012; Zhou et al. 2009) **(Figure 6).**

Acid and bile salts can also activate the Hedgehog (HH) signaling pathway in esophageal epithelium. HH signaling is highly influential in the endoderm development (Krishnadath and Wang 2015). Rats lacking HH failed to develop a squamous epithelium and instead had an esophagus lined with columnar epithelium, suggesting a key role of squamous differentiation (Su et al. 2004). HH signaling contributes to the metaplastic development of the esophageal epithelium (Wang et al. 2010; Krishnadath and Wang 2015; Rimkus et al. 2016). Studies in mice have shown that the ectopic expression of SHH was correlated with the expression of BMP4 and SOX9 and induced columnar differentiation in the esophagus (Wang et al. 2010; Krishnadath and Wang 2015; Yang et al. 2012). HH signaling expression in the basal layer of the esophageal epithelium activates the BMP pathway in the underlying mesodermal layer and induces the expression of SOX9 during BE progression **(Figure 6).** SOX9 is a potential transcriptional driver of BE and is normally found in intestinal stem cells. It is also expressed in the basal layer of the esophageal cells during early development. However, this expression gradually decreases as the esophageal epithelium matures (Wang et al. 2010; Yang et al. 2012). There is a suggested mechanism by which BMP4 upregulates SOX9 in esophageal epithelia, driving a columnar differentiation during BE progression (Wang et al. 2010) **(Figure 6).**

WNT signaling progressively increases during the development of BE and aids in malignant transformation into adenocarcinoma. Reflux components activate WNT signaling, which subsequently activates SOX9 expression in the esophagus of mice (Ali et al. 2009; Goss et al. 2009; Huo et al. 2010). However, WNT is not involved in the development of BE but is important in

transitioning BE into esophageal adenocarcinoma via hyperproliferation (Anastas and Moon 2013; Niehrs 2012; Deng et al. 2013).

During BE progression, genes (*SOX2* and *p63*) involved in squamous cell differentiation are downregulated, and intestinal genes (*SOX9* and *CDX2*) are upregulated (Minacapelli et al. 2017) in the esophageal epithelium. Caudal type homeobox transcription factors *CDX1* and *CDX2* genes have also been implicated in the pathogenesis of BE (Stairs et al. 2008). The action of stomach acids and bile upregulates *CDX2* that explains the higher risk of BE recurrence in patients with GERD (Huo et al. 2010). *CDX2* plays a key role in intestinal epithelial differentiation and maintenance (Hryniuk et al. 2012) and its expression in esophageal epithelium shows that BE is complex at the molecular level.

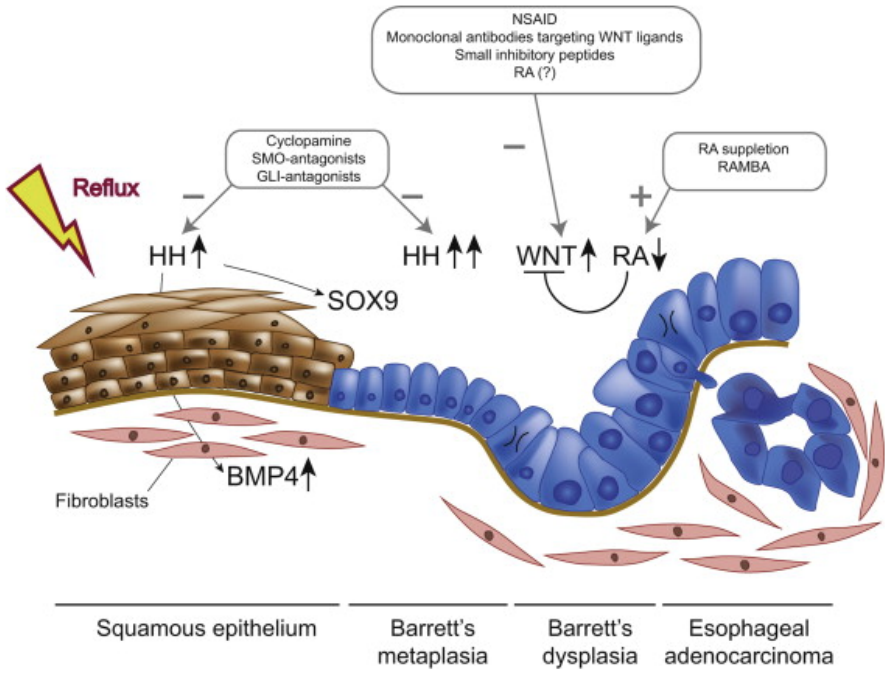


Figure 6: Dysregulation of signaling pathways BMP, HH, and WNT during BE development and malignant transformation. In case of bile and acid exposure, BMP pathway is activated with stromal BMP4 expression, initiating a columnar transdifferentiation of squamous esophagus epithelium. HH pathway is also involved in development of BE by stimulating BMP pathway and SOX9 expression. WNT signaling on the other hand is not involved in BE development but in its progression into esophageal adenocarcinoma. Figure from (Pavlov et al. 2014). Authorization obtained from authors.

5.2. Prevalence of Barrett's esophagus in esophageal atresia patients

EA/TEF patients are at an increased risk of GERD due to both intrinsic impaired motility and structural factors and includes an intrinsically weakened peristaltic movement (Courbette et al. 2020) as well an abnormal development of smooth muscle tissue (Al-Shraim et al. 2015). GERD is also common in EA/TEF patients following surgery especially, in long gap EA where they lose their anti-reflux barrier after anastomosis. Repeated exposure to refluxate in the esophageal mucosa induces metaplasia. Close to 50% of EA patients develop GERD, which can persist throughout their lives, but it is the most frequent during the first 5 years of their life (Koivusalo, Pakarinen, and Rintala 2007; Hassan and Mousa 2017) and chronic GERD leads to dysphagia, esophagitis, and BE. BE has been shown to be 10-fold higher in EA patients, even in younger children (the youngest reported is 2 years old) compared to the general population (Vergouwe, H, et al. 2018; Hsieh et al. 2017). Also, several reports describe esophageal adenocarcinoma in adult EA/TEF patients (Pultrum et al. 2005; Adzick et al. 1989; Vergouwe, Gottrand, et al. 2018) and there is a significant overlap of the signaling pathways involved in esophagus development and BE progression suggesting a shared etiology. Therefore, potential *de novo* genetic mutations involved in esophagus development might affect the esophageal tissue and BE progression to cancer in EA/TEF patients. However, it remains unclear if the incidence of BE in EA/TEF patients is higher due to the increased susceptibility of their esophageal epithelium to metaplastic changes.

6. Induced pluripotent stem cells as a model for understanding organogenesis and disease modeling

Takahashi et al. showed that somatic cells could be reprogrammed to generate induced pluripotent stem cells (iPSCs), through forced expression of several transgenes (Oct4, Sox2, Klf4 and c-myc) important for maintaining the defined properties of embryonic stem cells (ESCs) (Takahashi and Yamanaka 2006). Transcription factor-mediated reprogramming is initiated by changes in the transcriptome and chromatin structure of differentiated state into pluripotent state. Since then, several methods have been developed to generate iPSCs. The original method by Yamanaka (2006), was based on retroviral transduction of Oct4, Sox2, Klf4 and c-myc (Yamanaka factors) into mouse embryonic fibroblasts. Other methods include non-integrative viruses such as adenovirus, and sendai virus, or non-viral reprogramming methods such as mRNA transfection,

PiggyBac, and episomal plasmids. iPSCs have been extensively used in gaining insights on human embryonic development and disease modeling (Karagiannis et al. 2019). IPSCs, like ESCs, could be directed to differentiate into a specific cell fate, which involves the addition of growth factors, small molecules (activators or inhibitors) under specific culture conditions mimicking human development. Several types of cells have been derived from iPSCs such as intestinal organoids, hepatocytes, respiratory epithelium, and cardiomyocytes among others (Munera et al. 2017; Zhang et al. 2014; Huang et al. 2015; Burridge and Zambidis 2013). The first critical step in generating esophageal epithelial cells is to direct the differentiation of iPSCs to definitive endoderm (DE). Definitive endodermal cells are then guided into an anterior foregut (AF) fate from which the esophageal and tracheal progenitors arise when exposed to specific growth factors that regulate D/V patterning at this stage (Billmyre, Hutson, and Klingensmith 2015). This step involves the inhibition of BMP and TGF-ß pathways promotes an esophageal fate, whereas the activation of BMP, WNT, and RA pathways favors ventral tracheal fate **(Figure 7).**

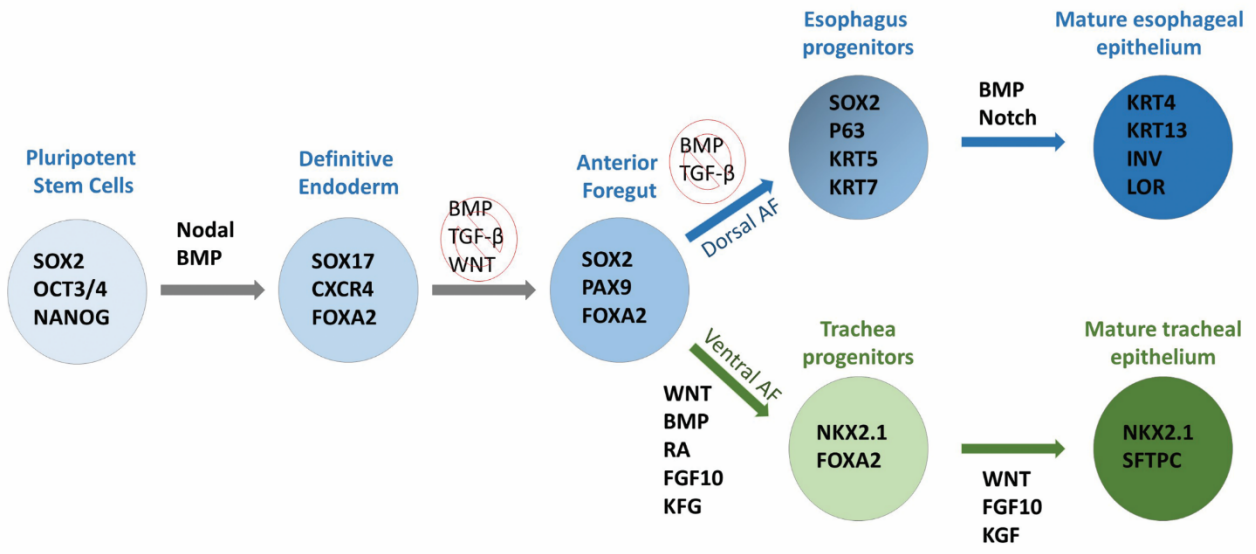


Figure 7. Scheme representing protocols for differentiation of iPSCs/hESCs into mature esophageal and tracheal epithelium: The first critical step is the formation of the definitive endoderm (DE) from pluripotent stem cells expressing SOX2, OCT3/4, and NANOG. DE is generated by activating the Nodal and BMP pathways and the expression of SOX17, CXCR4, and FOXA2. DE cell populations are then directed into an anterior foregut (AF) fate by the inhibition of BMP, TGF-ß, and WNT pathways. Trachea and esophagus derive from the anterior side of the foregut. Patterning of the anterior foregut tube along the sagittal axis results in the formation of the esophagus on the

dorsal side and the trachea on the ventral side. By mimicking the main signaling pathways involved in foregut patterning, AF can generate esophageal progenitors by the dual inhibition of BMP and TGFb pathways and the tracheal progenitors by activation the WNT, BMP, and RA signaling pathways. Esophageal progenitors can be further matured by activating the BMP and Notch signaling pathways to generate a stratified squamous epithelium expressing KRT4, KRT13, INV, and LOR. On the other hand, tracheal progenitors generate mature tracheal epithelium expressing NKX2.1 and SFTC by activating the Wnt pathway. Adapted from Raad et. al (2020).

6.1. Use of induced pluripotent stem cells to generate esophageal organoids

IPSC-derived cells follow their intrinsic developmental cues and self-organize into "organoids" that could mimic various histological and functional aspects of real tissues and organs they represent (Eiraku et al. 2011; Nakano et al. 2012; Haller et al. 2019). Organoids mimic the complexity of epithelial tissues composed of layers of interacting cells shaped into invaginations, compact folds, cysts, and tubes (Dedhia et al. 2016). Organoids of multiple organs such as the pancreas, intestine, liver, retina, the cerebral cortex have been successfully generated (Haller et al. 2019; Negoro et al. 2018; Takebe et al. 2013; DiStefano et al. 2018; Shi et al. 2012). The esophagus is the least studied segment of the gastrointestinal tract and critical modulators facilitating esophageal development and subsequent epithelial morphogenesis remains to be explored. Zhang et al. (2018) and Trisno et al. (2019) described stepwise protocols to differentiate iPSCs into esophageal organoids **(Figure 8).** The first critical step was to induce the iPSCs into an endodermal fate and the attempts to generate DE are heavily influenced by the specific cell type formation during embryogenesis. Nodal signaling is central to the generation of the endoderm precursors (Shen 2007). Trisno et al. (2018) subjected iPSCs to high concentrations of Activin A and BMP4 for one day, followed by Activin A only for two days with increasing serum concentrations. Zhang et al. (2018) used high levels of Activin A, FGF2, and BMP4 for two consecutive days. Both teams produced DE populations efficiently with decreased expression levels of pluripotency markers, NANOG and SOX2, and an increased expression of an endodermal marker, FOXA2. When the definitive endodermal fate was achieved, BMP, TGF-ß, and Wnt pathways were manipulated to promote anterior foregut endodermal (AFE) fate. Zhang et al. (2018) used inhibitors of BMP, TGF-ß, and Wnt pathways to promote the generation of SOX2+ dorsal AFE population. Increased levels of SOX2 expression indicate the specification of AFE toward esophageal fate. However, Trisno et al. (2018) inhibited the BMP pathway while activating the FGF pathway to generate AFE

spheroids. AFE was then directed into esophageal progenitor cells (EPCs) expressing markers for basal esophageal cells such as p63, PAX9, and FOXE1and the absence of NKX2.1expression. EPCs were obtained by adding FGF10 to activate FGF signaling (Trisno et al., 2018) or inhibiting BMP and TGF-ß signaling pathways (Zhang et al. 2018). Esophageal organoids were matured and maintained in culture for 2 months. *Zhang et al. (2018)* selected EPCs with the cell surface markers ITGb4+ EPCAM+ to isolate epithelial cells from the mixture with non-epithelial cells (Zhang et al. 2018). These isolated PSCs-derived EPCs then formed epithelial organoids expressing regional keratins with p63+ proliferative cells on the external periphery of the organoid and cells at the center highly expressing KRT13, a squamous suprabasal epithelium marker. Zhang et al. (2018) showed that sorted EPCs embedded in Matrigel could also form a tubular structure when implanted into the kidney capsule. This tubular structure was lined with p63+ stratified epithelium with differentiated KRT13 suprabasal layer similarly to the organoids *in vitro*.

Despite the advantages of human organoids, there are still many limitations before considering them an ideal system for disease modeling. For instance, the esophageal organoids generated in our lab resemble human esophageal epithelium but lack other cell components of the esophagus such as stromal cell, smooth muscle, immune cells, or enteric neuroglial cells. Esophageal organoids containing of all 3-germ layer derived cell types will allow a better modeling of the esophagus as well as of the congenital anomalies such as EA/TEF. Han et al. (2020) showed that hPSCs can be differentiated into splanchnic mesoderm that can give rise to esophageal mesenchyme *in vivo* (Han et al. 2020) and Eicher et al. (2021) derived germ layer components and built complex fundic and esophageal tissue (Eicher et al. 2021). Similar approaches of incorporating cells derived from different germ layers (endodermal and ectodermal) to generate intestinal epithelial organoids with enteric nerves have been studied (Eicher et al. 2021).

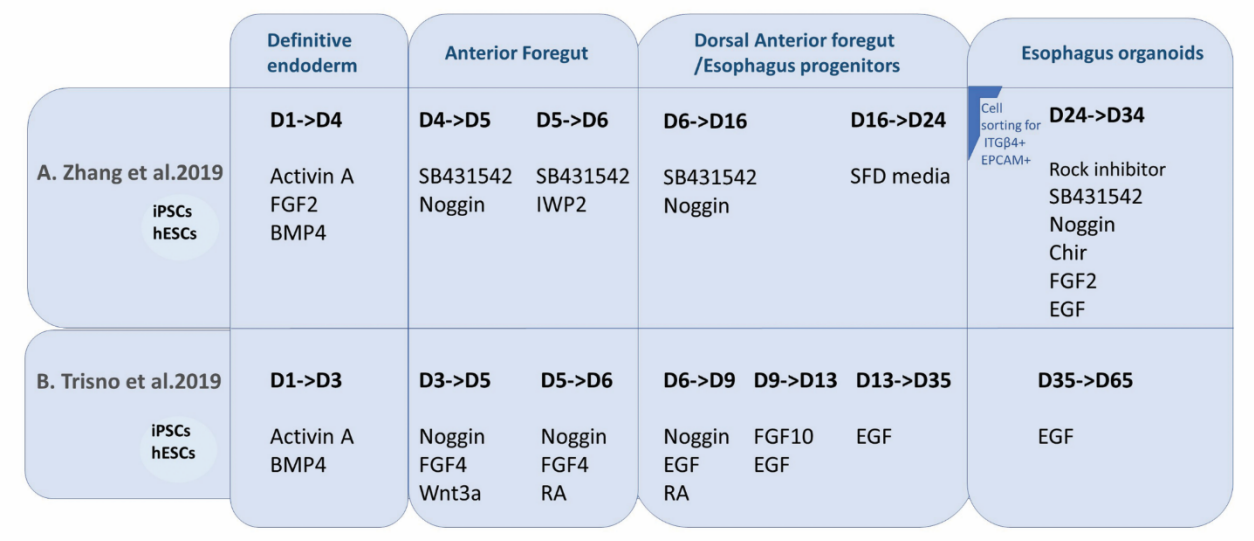


Figure 8. Schematic presentation of two published differentiation protocols of iPSCs/hESCs into esophageal organoids: (A) Zhang et al., 2019 differentiate iPSCs/hESCs into definitive endoderm using Activin A, FGF2, and BMP4 to form embryoid bodies. These embryoid bodies were then induced into anterior foregut by culturing in media supplemented with SB431542 and Noggin followed by IWP2 to inhibit the TGF-β, BMP, and WNT pathways respectively. After esophageal fate commitment, cells were sorted (ITGβ4+ EPCAM+) and resuspended to form esophageal organoids with an organoid culture medium with specific growth factors. (B) Trisno et al., 2019 propose a differentiation protocol where iPSCs/hESCs are subjected to Activin A and BMP4 for three days to generate definitive endoderm. Anterior foregut spheroids are then generated by inhibiting the BMP pathways and activating the WNT pathways for two consecutive days, followed by inhibition of BMP and activation of Retinoic acid (RA) for 24 hours. AF spheroids were then plated on Matrigel, and dorsal AF patterning achieved by the continued inhibition of BMP and activation of RA in addition to supplementing with EGF and FGF10. This step favored the efficiency of spheroid to organoid outgrowth. Esophageal organoids were maintained for two months.

Details: Activin A: activates Nodal Pathway, FGF2, FGF4, and FGF10: activates Fibroblast Growth pathway, EGF: activates Epidermal Growth Factor pathway, BMP4: activates Bone morphogenetic proteins pathway, CHIR: activates Wingless-related integration site pathway, RA: activates Retinoic acid pathway, SB431542: Inhibits Transformation Growth Factor-β pathway, Noggin: inhibits Bone Morphogenetic Proteins pathway, IWP2: inhibits Wingless-related integration site pathway, SFD: serum-free differentiation media, Rock inhibitor: Y-27632 selectively inhibits p160ROCK to increase survivability. Adapted from Raad et. al (2020).

6.2. Use of induced pluripotent stem cell-derived esophageal organoids to model EA/TEF

iPSC-derived organoids provide an efficient system to model and understand human organ development and malformations. Modeling the complex tissue morphogenesis of developing foregut remains challenging, however, using patient-specific iPSCs and derived organoids can help examine complex foregut diseases such as EA/TEF. Prior to our work and to the best of our knowledge, patient-derived esophagus organoids have not been used to study EA/TEF. In EA/TEF, the main signaling pathways and key transcription factors involved in normal foregut separation are dysregulated. In this context, using iPSCs generated from EA/TEF patients would allow deciphering dysregulated genes and signaling networks involved in the malformation. Recent studies using healthy iPSC-derived esophagus organoids has revealed some insights on EA/TEF (Trisno et al. 2018). For instance, downregulation of *SOX2* in esophagus organoids using CRISPR gene editing technique, revealed that *SOX2* regulates the dorsal expression of SFRP2, a WNT antagonist (Trisno et al. 2018). This suggests that inducing WNT signaling is one mechanism by which *SOX2* mutations in humans can cause EA/TEF. Using this approach, targeted mutation in healthy iPSC cell the variant in patient iPSC cell line can help validate whether patient variant mutation is causation of EA/TEF. However, in this thesis (chapters 2 and 3), we describe how we are the first to generate EA/TEF-derived iPSCs and derive patient-specific esophageal organoids. Our work revealed a dysregulation in gene expression of key signaling molecules in patient-derived cells. This is particularly valuable as it provides an excellent patient-specific complementary platform to study EA/TEF.

6.3 Use of induced pluripotent stem cell esophageal organoids to model Barrett's esophagus

Animal models have been extensively used model to study and understand BE (Kapoor et al. 2015) and the various structural and function differences between human and animal models need cautious consideration, thereby highlighting the need for an appropriate experimental model. The mouse esophagus consists of 4-6 layers of keratinized squamous epithelium and a single layer of basal cells, whereas the human esophagus consists of 30-40 layers of non-keratinized squamous epithelium containing elongations of lamina propria in the form of papillae and several layers of

basal cells (Rosekrans et al. 2015). Also, rodent esophagus epithelium lacks submucosal glands, one of the suggested precursor cells of BE. Identifying the molecular basis of BE development and cell of origin is crucial for disease management and development of new treatments. Therefore, using esophageal organoids generated from iPSCs could serve as a tool to model BE and study the role of various signaling pathways that have been previously linked to the disease such as BMP, WNT and Hedgehog signaling pathways. This can be achieved by incubating mature esophageal organoids in BMP4, b-Catenin or SHH proteins respectively and investigate the potential gene expression signatures reflecting metaplastic changes such as the loss of esophageal markers and the expression of gastric/intestinal markers. Furthermore, by using EA patient-derived esophageal organoids, it could be possible to not only understand the etiology of this disease but also develop a personalized treatment approach.

7. Rationale, hypothesis, and specific aims

The digestive and respiratory organ system derives from a common endodermal origin, the anterior foregut tube. The esophagus develops on the dorsal side and the trachea in the ventral region of the anterior foregut tube, in a process that remains poorly understood. Disruption of this tightly regulated process results in the most common congenital anomaly of the upper gastrointestinal tract: esophageal atresia/tracheoesophageal fistula (affecting 1 in 3,000 births). Our knowledge on the development of the esophagus remains limited and the exact molecular and cellular mechanisms that regulate the separation of the foregut into the esophagus and trachea remain poorly understood. Animal models have shown the involvement of signaling pathways and transcription factors involved in the development of the esophagus and trachea. However, there are structural and physiological differences between human and animals. Therefore, it is important to have a representative model of the human esophagus tissue. Recently, iPSC-derived cultures have gained importance as modeling tools to understand organ development and diseases. iPSCs can be differentiated into specific cell types by mimicking developmental pathways identified in animal models. These cultures can also undergo self-organization and form 3D organoids that resemble fetal tissue. Using patient-specific iPSCs and organoids can be used to study diseases and decipher potential mechanisms. Here, we are the first to describe the generation of iPSCs from EA/TEF type C patients. We propose that esophagus organoids from EA/TEF patients will serve as

complementary modeling tools to understand EA/TEF and other esophagus diseases such as Barretts esophagus.

In **Chapter 2,** we established three iPSC cell lines using peripheral blood mononuclear cells (PBMCs) from EA/TEF type C pediatric patients, using non-viral integrated Sendai virus vector. Pluripotency and trilineage differentiation capacity was confirmed by gene and protein expression profile and teratoma assay. The generated EA/TEF-specific cell lines serve as a useful tool to decipher mechanisms of EA/TEF and acquired associated diseases such as Barrett's esophagus (Published in *Stem Cell Research journal*).

In **Chapter 3**, we differentiate PSCs (ESCs and iPSCs) from healthy individuals and iPSCs from EA/TEF pediatric patients into mature esophagus organoids by manipulating key signaling pathways involved in esophagus development. We demonstrate that patient-derived cells exhibit a unique molecular signature, especially at the foregut stage. Additionally, SOX2, a critical esophagus fate regulator, is downregulated in patient-derived anterior foregut cells compared to the healthy group. Our study establishes a complementary system to understand morphogenesis and mechanisms involved in early esophagus development (*Under revision in Disease Models and Mechanisms journal*)

In **Chapter 4**, we sought to validate our esophagus organoids system to model Barrett's esophagus in EA/TEF patients. Our preliminary results show that patient-derived esophagus organoids are more susceptible to perturbation in gene expression when treated with BMP4 to induce BE (*Manuscript in preparation*).

In **Chapter 5,** we conclude that the current work generated a patient-specific esophagus organoid system to decipher potential mechanisms leading to EA/TEF and associated diseases, and potentially provide a platform for screening drugs to improve organ functionality.

CHAPTER 2

Generation of three induced pluripotent stem cell lines from patients with esophageal atresia/tracheoesophageal fistula type C

Raad Suleen, David Anu, Chung Wendy K, Faure Christophe

Published in *Stem Cell Research*, Volume 60, 2022, 102711, ISSN 1873-5061,
https://doi.org/10.1016/j.scr.2022.102711

Author affiliations

Suleen Raad, MSc

Esophageal Development and Engineering Laboratory, Sainte-Justine Research Centre, 3715 Côte Sainte Catherine, H3T1C5, Montreal, QC, Canada.

suleen.raad@umontreal.ca

Anu David, PhD

Esophageal Development and Engineering Laboratory, Sainte-Justine Research Centre, 3715 Côte Sainte Catherine, H3T1C5, Montreal, QC, Canada.

Anu.david@umontreal.ca

Co-corresponding author

Wendy Chung, MD PhD

Kennedy Family Professor of Pediatrics and Medicine

Chief, Clinical Genetics Columbia University, 1150 St. Nicholas Avenue, Room 620, New York, NY 10032

Wkc15@colombia.edu

Christophe Faure, MD

Esophageal Development and Engineering Laboratory, Sainte-Justine Research Centre,

Esophageal Atresia Clinic and Division of Pediatric Gastroenterology Hepatology and Nutrition, CHU Sainte Justine, 3715 Côte Sainte Catherine, Université de Montréal, H3T1C5, Montréal, QC, Canada.

Corresponding author: christophe.faure@umontreal.ca

Authors' contribution: AD, CF and SR designed the study. SR performed the experiments. SR and AD went through the data analysis. WC was responsible for genetic analysis. SR, AD and CF wrote the manuscript. AD and CF reviewed the manuscript.

Abstract

Esophageal atresia/tracheoesophageal fistula (EA/TEF) is the most common congenital anomaly of the upper gastrointestinal tract affecting 1 in 3,000 which could stem from a developmental anomaly of the foregut. The cause is not fully understood. We generated three iPSC cell lines using peripheral blood mononuclear cells (PBMCs) from EA/TEF type C patients. Pluripotency and trilineage differentiation capacity of these three iPSC cell lines were confirmed by gene and protein expression profiles and the differentiation ability into the three germ layers. The generated disease-specific cell lines could serve as a tool to investigate the mechanisms of EA/TEF and acquired associated diseases.

Resource table

Unique Stem Cell Line Identifier	CHUSJi001-A CHUSJi002-A

	CHUSJi003-A
Alternative name(s) of stem cell line	EA1C5 EA2C5 EA3C6
Institution	CHU-Sainte Justine, Montreal, QC, Canada
Contact information of distributor	Prof. Christophe Faure, christophe.faure@umontreal.ca
Type of cell lines	Human induced pluripotent stem cells (hiPSCs)
Origin	Human
Additional origin information	EA1C5 Age: 2 years Sex: Male EA2C5 Age: 6 years Sex: Male EA3C6 Age: 18 years Sex: Female
Cell source	Peripheral blood Mononuclear cells (PBMCs)
Clonality	Clonal
Method of reprogramming	Sendai virus expressing OCT3/4, SOX2, c-MYC and KLF4 genes
Genetic Modification	No
Type of genetic modification	N/A
Evidence of the reprogramming transgene loss	RT/qPCR loss of Sendai virus and exogenous transcription factors (TaqMan iPSC Sendai Detection kit)
Associated disease	Esophageal atresia Type C
Gene/locus	EA1C5: POLD3 c.1295_1296del p.Val432GlufsTer41 EA3C6: CEP350 c.4711_4712del p.Ser1571PhefsTer3
Date archived/stock date	None
Cell line repository/bank	https://hpscreg.eu/cell-line/CHUSJi001-A https://hpscreg.eu/cell-line/CHUSJi002-A https://hpscreg.eu/cell-line/CHUSJi003-A
Ethical approval	Ethics Committee of CHU Sainte-Justine Research Center (Protocol #2018-1670 (For iPSCs); 2019-2102 (For WES)

Resource Utility

Generation of esophageal organoids from healthy human iPSCs have been described (Zhang et al. 2018; Trisno et al. 2018; McCauley and Wells 2017).To the best of our knowledge no EA/TEF-patient derived iPSCs and its subsequent directed differentiation have been reported. EA/TEF patient-derived iPSC cell lines can be used to model esophageal congenital (EA/TEF) and acquired associated diseases (Barrett's esophagus, gastroesophageal reflux).

Resource details

EA/TEF is a congenital anomaly of the upper gastrointestinal tract affecting 1 in 3,000 births (van Lennep et al. 2019). EA/TEF affects the continuity of the esophagus, resulting in an incomplete open tube requiring surgery immediately following birth. EA can occur with or without an abnormal connection between the esophagus and the trachea. Type C EA/TEF is the most common type (>80% of cases), where the upper segment of the esophagus ends in a blind pouch, and a fistula connects the lower part to the trachea. EA/TEF patients can have chronic esophageal (such as Barrett esophagus) and respiratory symptoms requiring continuous physician visits throughout their lives. Some evidence suggests potential genetic causes of EA/TEF which includes syndromes resulting from mutation of a single gene, such as CHARGE syndrome, Feingold syndrome, Fanconi anemia; or chromosomal anomalies, such as Trisomy 21, Trisomy 18, Trisomy 13, and others. However, in most cases of EA/TEF, there has been no association with gene mutations. Our knowledge of the development of the esophagus remains limited and the exact molecular and cellular mechanisms that regulate the precise separation of the foregut tube into the esophagus and trachea still needs to be properly understood (Raad et al. 2020). Cells derived from EA/TEF patients could be an ideal model to mimic the development of the esophagus and trachea which would enable us to understand the role of key signaling molecules and pathways that regulate esophageal and tracheal development.

Here, we established three iPSC cell lines derived from peripheral blood mononuclear cells (PBMCs) of EA/TEF type C patients. We reprogrammed these PBMCs using non-viral integrating Sendai virus vector that includes 4 transcription/reprogramming factors *OCT3/4, SOX2, KLF4,* and *C-MYC* to generate iPSCs. iPSC colonies were manually picked into individual clones and passaged before characterization and differentiation. EA1C5, EA2C5, and EA3C6 display the morphology of embryonic stem cells with a prominent nucleolus, a high nucleus to cytoplasm ratio and pancake-shaped densely packed cells that grow in colonies **(Fig. 1A).** To characterize the 3

iPSC cell lines, we performed immunofluorescence staining for pluripotency markers SOX2, SSEA4, OCT3/4, and TRA1-60 (EA1C5 passage 17, EA2C5 passage 18 and EA3C6 passage 16) **(Fig.1B).** Reverse transcription-quantitative polymerase chain reaction (RT-qPCR) analysis confirmed the relative expression of pluripotency markers *OCT3/4, NANOG, SOX2* of the three iPSC cell lines, similarly to embryonic stem cell line (H9) (EA1C5 passage 15, EA2C5 passage 16 and EA3C6 passage 14) **(Fig. 1C).** We also performed RT-qPCR analysis using specific primers of Sendai virus to ensure the loss of residual episomal Sendai virus (Early passages: EA1C5 passage 4, EA2C5 passage 4, EA3C6 passage 2. Late passages: EA1C5 passage 22, EA2C5 passage 24 and EA3C6 passage 23) **(Fig. 1D).** All three iPSC cell lines have normal karyotype assessed by GTG-banding (EA1C5 passage 16, EA2C5 passage 18 and EA3C6 passage 14) **(Fig. 1E)**. Exome sequencing revealed no pathogenic genetic variants in established EA/TEF risk genes in any of the 3 EA/TEF patients. However, in patient 1, a heterozygous c.1295_1296del p.Val432GlufsTer41 variant of the gene *POLD3*, and in patient 3, a heterozygous c.4711_4712del p.Ser1571PhefsTer3 variant of *CEP350* were identified. Neither one of these two gene variants have been linked to EA/TEF. No genetic variants were detected in patient 2. Finally, EA1C5 passage 21, EA2C5 passage 20 and EA3 C6 passage 19 were able to differentiate into the 3 germ layers, endoderm, mesoderm, and ectoderm **(Fig 1F).**

Materials and Methods

I. Isolation of PBMCs and Reprogramming

PBMCs were isolated after collagenase dissociation and reprogrammed into iPSCs with the integration-free based Sendai virus system (Cytotune 2.0 kit catalog # A16517 from Life Technologies). iPSC colonies were manually picked and cultured in feeder-free conditions in Essential 8 flex medium on Geltrex-coated dishes (Life Technologies Catalog# A1569601). iPSC clones were maintained in Essential 8 flex medium (ThermoFisher, catalog#A1517001) in feeder-free conditions and passaged more than 20 times. iPSCs generation and characterization were performed in the iPSC cell reprogramming core facility of CHU-Sainte Justine.

II. Cell culture conditions

H9 (Human embryonic stem cells) and EA1C5, EA2 C5 and EA3 C6 were cultured in E8 flex medium and passaged using EDTA 0.5mL at a ratio of 1:6 every 4-5 days (ThermoFisher catalog#15575020). Culture conditions were at 37^{0}C, 5% CO_2 and 85% relative humidity. The media was changed every other day.

III. Immunofluorescence

iPSC colonies were fixed for 20min with paraformaldehyde (4%), then permeabilized by 0.5% Triton X-100 in PBS containing corresponding serum at RT. The cells were stained with the antibodies for anti-human SOX2, SSEA-4, OCT3/4, and TRA1-60 overnight at 4°C using the pluripotent Stem Cell 4-Marker Immunocytochemistry Kit (catalog # A24881 from Life Technologies), followed by several washes with PBS then incubated with an ALEXA secondary antibody for 30 minutes at room temperature (Table 2). Nuclei were counterstained with 4′,6-diamidino-2-phenylindole (DAPI).

IV. Exome sequencing

Exome sequence data were generated, and data were analyzed as previously reported. PMID: 32641753.

V. Teratoma formation

Cell clumps of EA1C5, EA2C5 and EA3C6 were first detached using the gentle cell dissociation reagent (StemCell Technologies Catalog# 100-0485) and a million cells were resuspended in cold phosphate-buffered saline (PBS) containing 50% Geltrex and kept on ice. Geltrex loaded cells were then injected under the renal capsule of SCID mice (n=1/patient). Eight weeks later, teratomas were dissected out and placed in OCT (Fisher Scientific, Catalog #23-730-571). Cryosections obtained were then stained with hematoxylin and eosin for morphological analysis to determine the three germ layers.

VI. RNA extraction and RT-qPCR

RNA was extracted using the Promega kit ReliaPrep™ RNA Cell Miniprep System. RNA was then reverse transcribed using Omniscipt™ RT kit (Qiagen catalog#205111) and the

complementary DNA (cDNA) obtained was used for real time quantitative PCR using LightCycler instrument (Roche Life Science). cDNA was quantified using TaqMan Gene expression assays and the TaqMan primers to target genes were purchased from Thermo Fisher Scientific listed in Table 2.

2. Karyotype analysis and short tandem repeat (STR) analysis
   The karyotype (G-banding) for the three iPSC lines at metaphase were counted and analyzed by the CHU Sainte-Justine cytogenetic department. The STR analysis of the three iPSCs were performed by Orchid PRO DNA, Laval, Quebec.

3. Mycoplasma detection
   The iPSC cultures (EA1C5 passage 7, EA2C5, passage 9, and EA3C6 passage 5) were tested for mycoplasma contamination using the lookout mycoplasma PCR detection kit (Sigma Aldrich).

Acknowledgements: We greatly thank Dr. Basma Benabdallah, iPSC Core at the CHU Sainte-Justine Research Centre and Drs. France Léveillé and Emmanuelle Lemyre, Cytogenetic Department at the CHU Sainte-Justine Hospital for their technical support. This work was funded by the Foundation CHU Sainte-Justine, and the Association Quebecoise de l'artresie de l'oesophage (AQAO) to Dr. Christophe Faure and 1P01HD093363 to Dr. Wendy K Chung.

Table 1. Characterization and validation

Classification	Test	Result	Data
Morphology	Microscopy	Normal	Fig. 1A
Phenotype	Qualitative analysis (Immunofluorescence staining) Quantitative analysis (RT-qPCR)	Expression of pluripotency markers SOX2, SSEA4, OCT3/4, TRA1-60 Expression of endogenous pluripotency markers: *SOX2, NANOG,* and *OCT3/4* Negative expression of exogenous markers: SeV, KOS, KLF4, MYC	Fig. 1B Fig. 1C Fig. 1D
Genotype	Karyotype (G-banding) and resolution	EA1C5: Normal, 46XY EA2C5: Normal, 46XY EA3C6: Normal, 46XX Resolution: 400	Fig. 1E
Identity	STR analysis	16 loci matched well	Submitted in archive with Journal
Mutation analysis	Exome sequencing	EA1C5: *POLD3* variant EA3C6: *CEP350* variant Both variants not linked to EA/TEF No variants detected in EA2C5	Available with authors
Microbiology and Virology	Mycoplasma testing by luminescence (Mycoplasma Detection Kit)	All three patient-derived iPSCs were negative	Available with authors
Differentiation potential	Teratoma formation	Histological staining for representative cell types for ectoderm, mesoderm and endoderm	Fig. 1F
Donor screening	HIV1+2 Hepatitis B, Hepatitis C	N/A	N/A
Genotype additional info	Blood group genotyping	N/A	N/A

	HLA tissue typing	N/A	N/A

Table 2. Reagent details

Antibodies used for Immunohistochemistry/Immunofluorescence				
	Antibody	**Dilution**	**Company Cat #**	**RRID**
Pluripotency marker	Rabbit anti-human SOX2	1 is to 200	Life Technologies Cat# 48-1400	AB_2533841
Pluripotency marker	Mouse anti-human OCT4	1 is to 100	Biolegend Cat #653702	AB_2561766
Pluripotency marker	Mouse anti-human SSEA4	1 is to 200	ThermoFisher Cat# 41-4000	AB_2533506
Pluripotency marker	Mouse anti-human TRA-1-60	1 is to 200	BioLegend Cat# 330602	AB_1186144
Secondary antibody	anti-rabbit Alexa 488	1 is to 1000	ThermoFisher Cat#A11070	AB_142134
Secondary antibody	Anti-mouse Alexa 488	1 is to 1000	ThermoFisher Cat# A28175	AB_2536161
Secondary antibody	Anti-mouse Alexa 594	1 is to 1000	ThermoFisher Cat# A11020	AB_141974
Nuclear counterstain	DAPI	2ug/mL	Thermo Fisher Cat# D1306	AB_2629482

	Primers			
	Target	Size of band	Thermo Fisher TaqMan assay ID	
Sendai virus and transgenes	Sendai-cMyc	89bp	Mr04269876_mr	
Sendai virus and transgenes	Sendai-OCT3/4	82bp	Mr04269878_mr	
Sendai virus and transgenes	Sendai-KLF4	74bp	Mr04269879_mr	
Sendai virus and transgenes	Sendai	59bp	Mr04269880_mr	
Sendai virus and transgenes	Sendai-SOX2	62bp	Mr04269881_mr	
Pluripotency marker	SOX2	86bp	Hs04234836_s1	
Pluripotency marker	NANOG	109bp	Hs02387400_g1	
Pluripotency marker	OTC4	77bp	Hs00999632_g1	
Housekeeping gene	GAPDH	157bp	Hs02786624_g1	

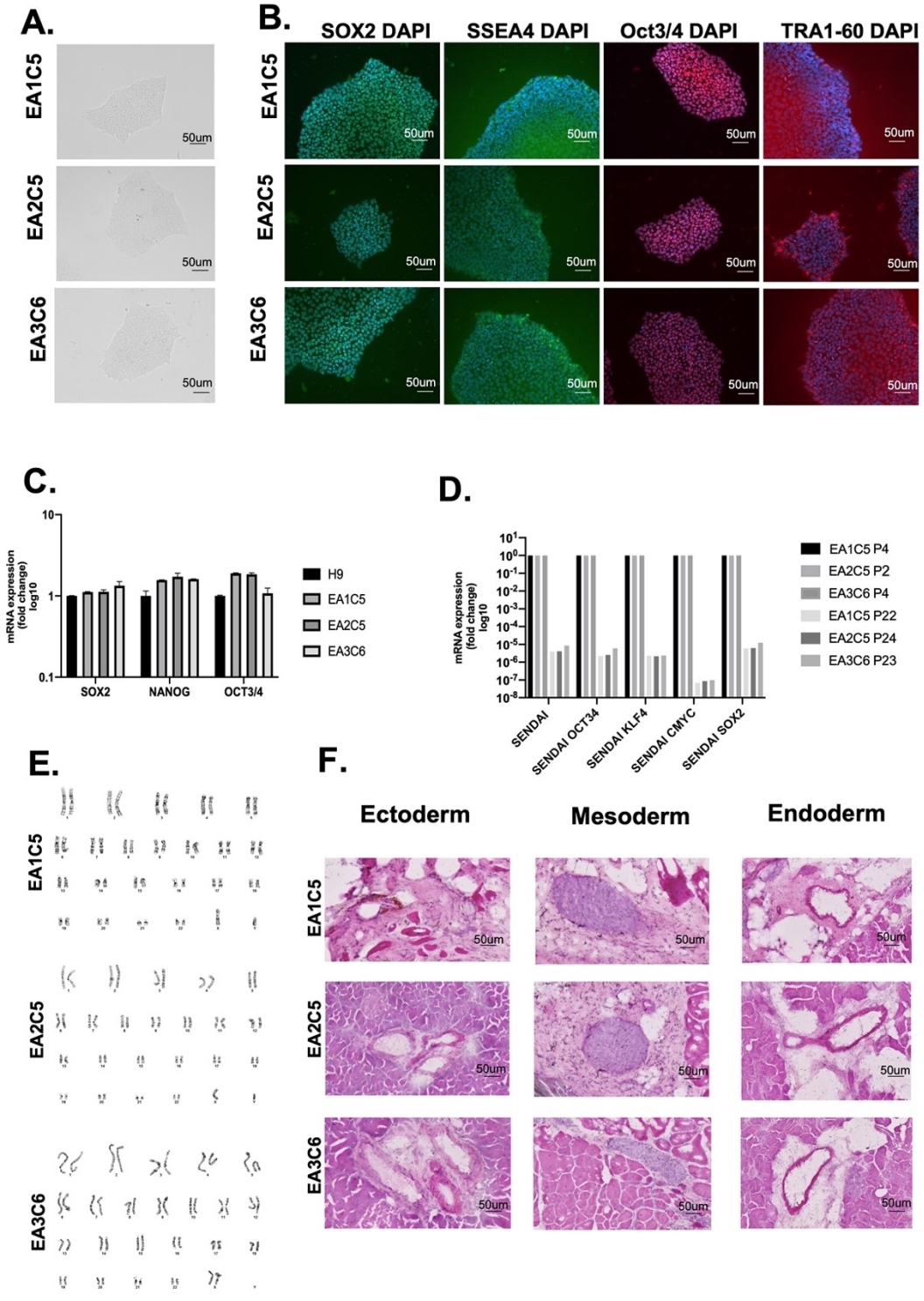


Figure 1: Characterization of EA1C5, EA2C5 and EA3C6.

Declaration of interests

☒The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

CHAPTER 3

Article under revision in Disease Models and Mechanisms Journal DMM/2022/049541

Directed differentiation of EA/TEF patient-derived induced pluripotent stem cells into esophageal epithelial organoids reveal SOX2 dysregulation at the anterior foregut stage

Raad S., David A., Sagniez M., Orfi Z., Dumont N., Smith M., Faure C.

Author affiliations

Suleen Raad, MSc

Esophageal Development and Engineering Laboratory

Sainte-Justine Research Centre

3175 Côte Sainte-Catherine, H3T 1C5, Canada

suleen.raad@umontreal.ca

Anu David, PhD

Esophageal Development and Engineering Laboratory

Centre Hospitalier Universitaire Sainte-Justine Research Centre

3175 Côte Sainte-Catherine, H3T 1C5, Canada

anu.david@umontreal.ca

Co-corresponding author

Melanie Sagniez

CHU Sainte-Justine Research Centre, Montréal, QC, Canada

Department of Biochemistry and Molecular Medicine, Faculty of Medicine, University of Montreal, QC, Canada

Zakaria Orfi, PhD

Centre Hospitalier Universitaire Sainte-Justine Research Centre

3175 Côte Sainte-Catherine, H3T 1C5, Canada

Nicolas A. Dumont, PhD

School of rehabilitation, Faculty of medicine, Université de Montréal

Centre Hospitalier Universitaire Sainte-Justine Research Centre

3175 Côte Sainte-Catherine, H3T 1C5, Canada

Martin Smith, PhD

CHU Sainte-Justine Research Centre, Montréal, QC, Canada

Department of Biochemistry and Molecular Medicine, Faculty of Medicine, University of Montreal, QC, Canada

Christophe Faure, MD

Esophageal Development and Engineering Laboratory, Sainte-Justine Research Centre,

Esophageal Atresia Clinic and Division of Pediatric Gastroenterology Hepatology and Nutrition, CHU Sainte Justine, 3715 Côte Sainte Catherine, Université de Montréal, H3T1C5, Montréal, QC, Canada.

christophe.faure@umontreal.ca

Corresponding author

Authors' contribution: AD, CF and SR designed the study. SR performed the experiments. SR and AD went through the data analysis. MS and MS performed the nanopore RNA sequencing, data acquisition and analysis. ND and ZO were responsible for myoblast differentiation. SR, AD and CF wrote the manuscript. AD and CF reviewed the manuscript.

Abstract

The esophagus and trachea originate from the endodermal diverticulum in the anterior foregut tube during embryogenesis. A series of well-regulated cellular and molecular events result in the compartmentalization of the anterior foregut tube into the esophagus and trachea. Disruption of the compartmentalization process leads to esophageal atresia/tracheoesophageal fistula (EA/TEF), a congenital anomaly of the esophagus affecting 1 in 3,000 births. The mechanisms that regulate trachea-esophageal separation are poorly understood. Induced pluripotent stem cells (iPSCs) thus provide an opportunity to understand human embryonic development and could be used to elucidate mechanisms associated with congenital diseases such as EA/TEF. Therefore, the objective of the current study is to differentiate pluripotent stem cells (PSCs), namely, embryonic stem cells [ESCs] and iPSCs from healthy individuals and iPSCs from three EA/TEF type C patients, into mature 3-dimensional esophageal organoids in matrix- and xenogeneic-free culture conditions. We directed the differentiation of PSCs toward different development stages, such as the definitive endoderm, anterior foregut, and mature esophageal epithelium and compared their protein and gene expression profiles by immunofluorescence, qPCR, and long-read RNA sequencing analysis. Expression of endodermal markers *CXCR4, SOX17,* and *GATA4*, was similar in both patient and healthy differentiated cells. At the anterior foregut stage, we observed a significant decrease in the expression of key transcription factor *SOX2* in patient-derived cells. Furthermore, by combining targeted gene expression and nanopore RNA sequencing, we observed that patient-derived cells exhibit unique molecular signatures, especially at the anterior foregut stage. This includes critical genes *GSTM1* and *RAB37* that are significantly lower in patient-derived cells. *GSTM1* regulates cellular morphogenesis and is associated with EA/TEF, whereas,

RAB37 is a key regulator of vesicle trafficking. Rab proteins have been linked to foregut malformations. Furthermore, we observed an abnormal expression of *NKX2.1*, a tracheal marker, in the patient-derived mature esophageal epithelial organoids. In conclusion, we developed the first 3-dimensional mature stratified squamous esophageal epithelial organoids from EA/TEF patient-derived iPSCs. We hypothesize that a transient dysregulation of key transcription factor *SOX2* and the abnormal expression of *NKX2.1* in patient-derived cells could be responsible for the abnormal foregut compartmentalization into the esophagus and trachea.

Introduction

The esophagus and trachea originate from the endodermal diverticulum in the anterior foregut tube. Well-regulated and organized cellular and molecular events result in the separation of the anterior foregut tube into the esophagus and trachea (Raad et al. 2020) (Billmyre, Hutson, and Klingensmith 2015). Disruption of the compartmentalization process results in severe esophageal congenital anomalies such as esophageal atresia with or without tracheoesophageal fistula (EA/TEF), affecting 1 in 3,000 newborns (van Lennep et al. 2019). Several types of EA/TEF have been described based on the location of the malformation and the affected structures, with the most common being type C (>80% of cases), where the upper segment of the esophagus ends in a blind pouch, and a fistula connects the lower part to the trachea. Other less common subtypes include type A (8-10% of cases) where no fistula exists, but the esophagus is disconnected (Clark 1999). EA/TEF-associated anomalies (cardiac, anal, renal, limb, or vertebral) are also reported in 30-50% of syndromic cases. Monogenetic causes account for a minority of EA/TEF cases (<5%) most often in syndromic cases such as Anophthalmia-Esophageal-Genital (AEG) syndrome (*SOX2* mutations), Feingold syndrome (*MYCN* mutations), CHARGE syndrome (*CHD7* mutations), Pallister-Hall syndrome (*GLI3* mutations) and Mandibulofacial dysostosis (*EFTUD2* mutations)(Stoll et al. 2009). Studies have also shown a multigenic architecture of rare variants in several genes, which discriminate EA/TEF cases from controls (Wang et al. 2021). However, the cause of EA/TEF remains largely unknown, and rare genetic variants are seldom reported in non-syndromic, isolated cases. EA/TEF is thus considered a multifactorial anomaly resulting from genetic and environmental factors (Brosens et al. 2014).

During embryogenesis, the esophagus and trachea arise after the separation of the anterior foregut endoderm common tube at week 4-5 in humans and embryonic day 9.5-11.5 in mice. In animal models (mouse and Xenopus), the dorsal/ventral patterning of the anterior foregut allows spatial specification of the two presumptive organs: the esophagus on the dorsal side of the anterior foregut tube (characterized by the expression of the transcription factor *SOX2*) and the trachea on the ventral side of the foregut tube (characterized by the expression of the *NKX2.1* gene) (Minoo et al. 1999; Que et al. 2007; Kim et al. 2019). Studies in mice have also demonstrated that the dorsal/ventral patterning is initiated by gradual expression of mesodermal Wnt2/2b, Bmp4, and Noggin along the dorsal-ventral axis. Wnt and Bmp signaling pathways inhibit *SOX2* expression

on the dorsal side of the anterior foregut (Domyan et al. 2011) and drive *Nkx2.1* expression toward the tracheal lineage. Functional genomic studies in mice and Xenopus have also been used to mimic the genetics and morphogenetic regulation of normal and abnormal foregut compartmentalization representing human esophageal anomalies such as the EA/TEF (Kim et al. 2019; Fausett and Klingensmith 2012; Raad et al. 2020). However, these studies utilize methods resulting in functional loss of specific genes that may not represent the genetic complexity observed in humans. The esophagus of the human differs structurally and morphologically from the esophagus of the mouse (Rosekrans et al. 2015). Therefore, there is a need to have a representative model of human esophagus development to not only decipher but also understand the possible mechanisms leading to EA/TEF. Induced pluripotent stem cells (iPSCs) offer an excellent tool in gaining insights not only on human embryonic and developmental ontologies but also to model diseases (Karagiannis et al. 2019; Rowe and Daley 2019) through the directed differentiation to specific organs originating from all three germ layers. To date, patient-derived iPSCs have not yet been used to study digestive malformations. Recently, patient-derived iPSCs have been used to study congenital heart diseases, where intrinsic defects were observed in differentiated cardiomyocytes derived from these iPSCs (Miao et al. 2020; Hrstka et al. 2017; Yang et al. 2017). Over the last few years, studies using healthy human iPSCs have been used to generate mature esophageal epithelium (Zhang et al. 2018; Trisno et al. 2018) and confirmed previous findings that *SOX2* is key to promoting esophageal specification and the critical role of BMP, TGF-ß, and WNT signaling pathways during esophageal development (Que et al. 2007; Teramoto et al. 2020a; Zhang et al. 2018; Trisno et al. 2018; Que et al. 2006; Li et al. 2007; Woo et al. 2011; Domyan et al. 2011).

Therefore, the objective of the current study is to differentiate embryonic stem cells (ESCs) and iPSCs from healthy individuals and iPSCs from three EA/TEF type C pediatric patients into mature esophageal organoids in matrix- and xenogeneic-free culture conditions. We adapted and modified a stepwise differentiation protocol (Matsuno et al. 2016; Zhang et al. 2018; Giroux et al. 2017; DeWard, Cramer, and Lagasse 2014) by manipulating key signaling pathways involved in esophagus development. We investigated the gene and protein expression profiles of key signaling molecules in patient cells and compared them to healthy cells at each developmental stage. Furthermore, by combining targeted gene expression and nanopore RNA sequencing, we demonstrated that patient-derived cells exhibit unique molecular signatures, especially at the

anterior foregut stage. Our study establishes a basic framework to understand the morphogenesis and mechanisms involved during early esophageal development by using patient-derived iPSCs.

Results:

Our modified protocol includes the stepwise differentiation of PSCs into mature esophageal organoids with checkpoints at four developmental stages: (i) definitive endoderm, (ii) anterior foregut, (iii) mature esophageal epithelium and (iv) 3-dimensional esophageal organoids **(Scheme 1).** We compared the differentiation potential to these stages between healthy PSCs (embryonic stem cell line [H9] (female)) and an iPSC derived from a non-familial healthy male and EA/TEF iPSCs (2 males and 1 female).

1. **Derivation of EA/TEF patient iPSCs**

iPSC lines from 3 pediatric EA/TEF patients were established by reprogramming peripheral blood mononuclear cells (PBMCs) in the Stem Cell core facility at CHU-Sainte Justine. Their pluripotency was confirmed by the mRNA expression of pluripotent genes *SOX2, NANOG*, and *OCT4* and immunofluorescence staining with SOX2, NANOG, OCT4, and SSEA4. All 3 iPSC cell lines had a normal karyotype, had no pathogenic genetic variants in established EA/TEF risk genes and showed the ability to differentiate into the three germ layers as evidenced through teratoma formation. **(Raad et al. 2021, submitted to Stem Cell Research).**

2. **Similar differentiation potential of healthy and EA/TEF patient PSCs into definitive endodermal (DE) cells**

The first critical step is the differentiation of PSCs into endodermal cells that give rise to the entire epithelial lining of the gastrointestinal tract, including the esophagus epithelium(Wells and Melton 1999). We evaluated the efficiency of endoderm differentiation by qPCR through gene expression of specific markers *CXCR4, GATA4,* and *SOX17* **(Fig. 1A-D).** There was no significant difference in gene expression levels between healthy and patient-derived definitive endoderm. At the protein level, CXCR4 and GATA4 were observed in the cytoplasm, whereas SOX17 was observed in the

cell nuclei, confirming DE commitment in both groups **(Fig. 1E).** Furthermore, we also verified that no ectodermal cells were present through the absence of *OTX2* **(Data not shown).**

3. Critical dorsal esophageal marker is downregulated in EA/TEF patient-derived anterior foregut cells

Developmentally, the anterior side of the foregut tube separates dorsally into the esophagus and ventrally into the trachea and therefore, to generate the dorsal side of the anterior foregut, we inhibited key signaling pathways shown to be critical for esophagus specification; the BMP, TGFb, and WNT pathways **(Fig. 2A).** Anterior foregut cells in both groups expressed PAX9, a foregut endodermal marker at the gene and protein levels **(Fig. 2B, D)**. The cells also expressed ISL1 **(Fig. 2C, E),** a recently identified critical marker that contributes to the specification of anterior foregut to both esophageal and tracheal epithelium (Kim et al. 2019). *ISL1* regulates the expression of *NKX2.1* and is required for esophagus-trachea separation (Zhang et al. 2018; Kim et al. 2019). The expression profiles of ISL1 and PAX9 were similar in both groups **(Fig. 2D, E).**

However, SOX2, a critical transcription factor necessary for foregut morphogenesis and expressed on the dorsal side of the anterior foregut, was downregulated in the patient-derived cells **(Fig. 2F, G)**. It is known that disruption of SOX2 expression leads to an abnormal separation of the anterior foregut into the esophagus and trachea (Teramoto et al. 2020a). Following quantification, we observed that percentage of SOX2 positive cells was significantly lower in all three patients (~10%) compared to the healthy foregut cells (45%) **(Fig. 2F, H).**

At the anterior foregut stage when the compartmentalization occurs, two critical transcription factors namely, *SOX2* and *NKX2.1,* have a reciprocal repressive function. *Nkx2.1* binds to silencer sequences near the *Sox2* gene and represses its transcription (Kuwahara et al. 2020; Kim et al. 2019; Trisno et al. 2018; Han et al. 2020). However, dysregulated *SOX2* did not affect the expression of *NKX2.1*, which remained undetected in both groups **(Fig. 2I).** Furthermore, at this stage, we also confirmed the absence of other lineage markers, specifically, a mid-hindgut marker, *CDX2* **(Data now shown)** and a posterior foregut marker, *HNF4a* **(Supplementary Fig. S1A)**

4. Novel transcript isoforms and distinct molecular signatures in patient-derived anterior foregut cells using nanopore sequencing

Low SOX2 expression levels in patient-derived anterior foregut cells led us to investigate what other genes could be involved in this stage of development. . We report 2 new isoforms of the SOX2 gene, one of which presents an intron in its 3'UTR. This filtered *de novo* assembly was used as a reference for sample-specific abundance estimation. The latter revealed a **potential batch-effect associated with 2 separate sample preparations that,** unfortunately, also coincided with the sex of the individual, representing 42% of the observed variation in the data **(PC1, Fig. 3A).** We performed batch correction using SVA (Leek et al. 2021) to mitigate this effect, which resulted in an effective separation of disease and healthy samples across the 2 remaining principal components **(Fig. 3B).** Sequencing validated our previous observations that SOX2 is less expressed at the anterior foregut stage in the patient group. The differential transcript expression with DeSeq2 identified 173 transcripts that presented an over 2-fold change in normalized expression with a p-value below 0.01 **(Supplementary Table 1).** We could identify gene expression signatures unique to both conditions **(Fig. 3C, D).** Specifically, both SOX2 transcripts that overlap the TaqMan probes used in qPCR (SOX2-201 and SOX2-201(o)-25276.2, **Fig. 3E**) presented an average TPM of 618 in healthy samples versus 253 in affected samples. GSTM1-201 was among the top differentially expressed isoforms in patient-derived cells (log2(fold change) = 5.61). It is significantly lower in all 3 patients compared to healthy samples **(Fig. 3D).** Previous work has shown that GSTM1 is associated with EA/TEF (Filonzi et al. 2010). Also, among the list is RAB37-204 (log2(fold change) = 1.08), an endosomal protein critical for vesicle trafficking regulation. Rab proteins have been previously linked to foregut malformations (Nasr et al. 2019; Nam et al. 2010; Edwards et al. 2021). Among the top differentially expressed transcripts are several non-coding RNAs, including Y-RNA and MEG3 **(Fig. 3D and Supplementary Table 1).**

5. NKX2.1, a tracheal marker, is expressed in EA/TEF patient-derived esophageal epithelium

We further differentiated the anterior foregut cells into esophagus epithelium by inhibiting the BMP and TGF-ß pathways (Que et al. 2006; Guyot and Maguer-Satta 2016) **(Fig. 4A**). Even with low SOX2 expression in patient-derived anterior foregut cells, we observed that these cells were committed to an esophageal fate. Specifically, we observed that esophageal epithelial cells derived from both groups expressed esophageal marker, *p63*, normally expressed in the basal proliferative layer of the developing esophagus **(Fig. 4B)**. Cells from both groups also expressed KRT4, an

esophageal squamous epithelial marker **(Fig. 4C, E)**. Interestingly, SOX2, a marker also expressed by the basal proliferative esophageal epithelium, was observed at similar levels in both groups **(Fig. 4D, E).**

The expression of SOX2 during esophagus differentiation of EA/TEF patient iPSCs differs greatly from healthy group. At the anterior foregut stage, we observed a temporal downregulation of *SOX2* expression in the 2 groups but was significantly more pronounced in patient cells **(Fig. 4F).** Interestingly, *SOX2* expression goes back to similar levels to that of the healthy group at the esophageal epithelial stage **(Fig. 4F).** However, at this stage though the *SOX2* expression returned to normal levels we observed a significantly higher expression of NKX2.1 in patient-derived esophageal epithelial cells at the gene and protein levels **(Fig. 4G, H).** 17% of the cells were positive for NKX2.1 in patient-derived esophageal epithelium **(Fig. 4I).** A recent study (Nasr et al., 2019) identified *ISL1* to be a regulator of *NKX2.1* during foregut separation**.** But we did not observe any significant difference in the expression of *ISL1* in both healthy and patient derived cells at both the anterior foregut and esophageal epithelial stage **(Supplementary Fig. S1B).**

6. Mature esophageal epithelial organoids express key markers: Involucrin (INV), Keratin-4 and -13 (KRT-4 and KRT-13), and p63

For further maturation and to allow for cellular organization of the esophageal epithelium into a stratified squamous epithelium, 3-dimensional organoids were generated and further matured. Cells were detached from their 2-dimensional culture conditions, matured in suspension **(Fig. 5A)** and within 48 hours the cells clustered together to form organoids **(Fig. 5B).** We observed no morphological and proliferative differences between the healthy and patient-derived organoids **(Fig. 5B, Supplementary Fig S2).** We observed high gene and protein expression of suprabasal markers such as KRT4, KRT13, and Involucrin in both healthy and patient-derived organoids closer to the fetal esophagus than to adult esophagus biopsy **(Fig. 5C, D, E, F)** in addition to p63, a basal proliferative marker **(Fig. 5G, H).** We expected the expression of these markers to be normal at this stage because the upper and lower end of the esophagus in EA/TEF patients is not affected morphologically.

7. Abnormal NKX2.1 expression is retained in EA/TEF patient-derived organoids

NKX2.1 is not normally expressed in human esophagus biopsies **(Fig. 6A).** However, our patient-derived esophageal organoids showed a positive expression of *NKX2.1* at levels close to the fetal trachea**.** As expected, no expression was observed in healthy esophageal organoids just like fetal esophagus and esophagus epithelial biopsies (**Fig. 6B).** NKX2.1 expression was interspersed in the KRT13 expressing suprabasal layers of the patient-derived esophageal organoids **(Fig. 6C, D).** A similar observation was made in TEF tissue from EA patients which showed abnormal expression of NKX2.1 (Brosens et al. 2020). NKX2.1 dysregulation was detected as early as day 16 in esophagus progenitor cells (**Fig. 6D**). This abnormal expression was retained in mature organoids after 60 days in culture (**Fig. 6D**).

8. Differentiation propensity of EA/TEF iPSCs into different organ lineages is like healthy iPSCs

To verify whether the dysregulation of SOX2 at the anterior foregut stage and the abnormal expression of NKX2.1 at the mature esophageal epithelial stage in patient-derived iPSCs is specific to the esophageal fate, we investigated whether EA/TEF patients-derived iPSCs could be differentiated into other organ lineages such as tracheal, liver, and muscle progenitor cells. By using published protocol (Huang et al. 2014), we differentiated PSCs into the ventral anterior foregut cells, thereby favoring a tracheal fate, which was confirmed by the expression of NKX2.1 in both the groups **(Supplementary Fig. S3A, B).**

We then differentiated patient-derived iPSCs into posterior foregut and then directed it towards the hepatic stage. All 3 EA/TEF patient-derived iPSCs generated hepatoblast cells. Patient-derived hepatoblasts expressed Alpha fetoprotein (AFP) **(Supplementary Fig. S4).**

Finally, we directed the differentiation of the patient-derived iPSCs toward the mesodermal cell stage for generating skeletal muscle progenitor cells using a previously published protocol (Shelton et al. 2016). Myogenic progenitor cells derived from both healthy and patient iPSCs expressed similar levels of PAX3 and PAX7, which are required for myogenic specification **(Supplementary Fig. S5).** These results therefore suggest that any abnormal expression of key factors is intrinsic to the esophagus and not any other organ as is observed in these type C EA/TEF patients.

Discussion

We report here the first *in vitro* generated matrix- and xenogeneic-free 3-dimensional mature stratified squamous esophageal epithelial organoids from EA/TEF patient-derived iPSCs. We observed a significant down regulation of SOX2 mRNA and protein expression in patient-derived anterior foregut cells. Studies have shown that SOX2 downregulation is linked to abnormal foregut separation resulting in EA/TEF (Teramoto et al. 2020a; Trisno et al. 2018; Ioannides et al. 2010; Que et al. 2007; Domyan et al. 2011). We also observed an abnormal expression of NKX2.1 in patient-derived cells at the esophageal epithelial stage until the end of the organoid cultures. We also observed a distinct transcript expression profile in all 3 patient-derived anterior foregut cells, the most critical developmental time-point where patterning and subsequent separation into the esophagus and trachea occur. This dysregulation in gene and protein expression was specific to the dorsal side of the anterior foregut and therefore of the esophageal fate. In fact, directed differentiation of EA/TEF iPSCs into posterior foregut derived cells (hepatoblasts) and mesodermal cells (myoblasts) revealed a similar gene and protein expression profile to the healthy group.

The downregulation of SOX2 specifically, however, was temporary and expression levels become similar in both groups when cells were further differentiated into the mature esophageal epithelium. The exact mechanisms regulating SOX2 expression in our patient-derived cells remain unclear. Though NKX2.1 and SOX2 are hypothesized to be co-repressive master regulators of foregut separation, NKX2.1 mRNA and protein levels remained unaffected at the anterior foregut stage. Interestingly, following nanopore sequencing we observed unannotated long non-coding RNA (*lncRNA-21751*) lying upstream of *SOX2* promoter that is significantly downregulated in all 3-patient derived anterior foregut cells. The exact role of *lncRNA-21751* in regulating *SOX2* expression at the anterior foregut stage remains unknown. We also speculated on the potential role of the long non-coding RNA *SOX2OT* at this critical stage. *SOX2OT* harbors the intronic region of SOX2 gene. It plays a positive role in regulating *SOX2* expression in a mechanism that remains largely unknown (Shahryari et al. 2014). So, we performed qPCR analysis of *SOX2OT* on the anterior foregut cells and observed its expression to be downregulated in all 3 EA/TEF anterior foregut cells (**Supplementary Fig.S6**). This down regulation of the *SOX2OT* could be one of the regulatory molecules involved in the expression of SOX2 in the anterior foregut cells.

NKX2.1 is normally absent in the esophagus epithelium. However, in the patient esophagus epithelium and organoids, NKX2.1 mRNA and protein levels were significantly high. *ISL1*, a recently identified transcription factor that regulates the expression of *NKX2.1*, was found at similar levels in both groups (**Supplementary Fig. S1B**). Although information on *SOX2* expression regulation is available, upstream regulation of SOX2 at the earliest stages of anterior foregut development is unknown. At the anterior foregut stage, *SOX2* and *NKX2.1* have a reciprocal repressive function. *Nkx2.1* binds to silencer sequences near the *Sox2* gene and represses its transcription. In our experimental system, dysregulated SOX2 did not affect the expression of *NKX2.1*, which remained undetected in both groups suggesting that *Nkx2.1* is not responsible of *SOX2* downregulation at that stage. One could hypothesize that *SOX2* expression is epigenetically regulated and/or lncRNAs *(lncRNA-21751, SOX2OT*) observed in the RNA sequencing play a role in SOX2 expression at the anterior foregut stage. It is also unknown whether a lower SOX2 expression in the anterior foregut leads to an abnormal esophageal development and abnormal maintenance of esophageal identity. A recent study suggests that misexpression of *Sox2* in gut precursors alters organ identity (Smith et al. 2022). The authors show that disruption of *SOX2* expression is fully sufficient to alter cell fate decisions by either leading to a loss of identity or completely changing cell fate. These authors showed that changes in key lineage-specific transcription factors binding are sufficient to alter chromatin accessibility patterns and drive subsequent changes in lineage fate decisions.

We identified around 173 RNA transcript isoforms that were significantly differentially expressed between the healthy and patient groups. *GSTM1* was one of the most differentially downregulated genes with its distinct functions in the detoxification of electrophilic compounds including carcinogens, therapeutic drugs, environmental toxins, and products of oxidative stress. *GSTM1* has a non-catalytic regulatory role in apoptotic ASK1-MAPK (mitogen-activated protein kinase) signaling cascade (Cho et al. 2001). Under non-stimulated conditions, *GSTM1* inhibits apoptotic cell death (Cho et al. 2001). There has been an increasing trend of linking xenobiotics to genes involved in detoxification in early embryonic development and specifically to EA/TEF. It is suspected that an altered detoxification process triggers an alteration of proliferation or apoptotic cellular behavior that may directly affect the separation process of the foregut into the esophagus and trachea (Filonzi et al., 2010). Another interesting gene which was differentially expressed was from the family of small GTPases, Rab, which are key regulators of intracellular membrane

trafficking. In a recent study, *Rab11* was shown to have a direct link to epithelial remodeling and extracellular matrix degradation during the foregut separation (Nasr et al. 2019). The work shown in xenopus and mouse demonstrates how the disruption of *Rab11*-mediated epithelial remodeling results in tracheoesophageal clefts (Nasr et al. 2019), providing a potential mechanistic framework for foregut separation in humans. In our patient-derived anterior foregut cells, however, we observe a significant downregulation of another Rab protein, *RAB37*. *Rab37* is a critical regulator of vesicle trafficking and play a potential role during human foregut compartmentalization, similar to what was observed in xenopus and mice with *Rab11*. In a suggested mechanism, *Rab37* mediates exocytosis of SFRP1 (secreted frizzled related protein 1) an antagonist of the Wnt pathway to suppress Wnt signaling in lung cancer cells *in vitro*.(Cho et al. 2018). Knowing the particular importance of the inhibition of Wnt signaling in the anterior foregut to favor an esophageal fate (Woo et al. 2011), raises the potential role of Rab37 at this developmental stage. Furthermore, the identification of numerous new transcript isoforms, including known and novel long non-coding RNAs, supports the observed regulatory complexity of esophagus and trachea development, as well as EA/TEF etiology, whilst suggesting that non-coding regulatory transcripts play a role in this process.

We cannot exclude a role of mesenchymal cells in the dysregulation of SOX2 and NKX2.1 in the present experimental setting. Although at minimal levels, we detected mRNA expression of brachyury, a transcription factor that regulates mesoderm formation (Herrmann et al. 1990) and vimentin, an intermediate filament expressed in mesenchymal cells, in our cultures during directed esophagus differentiation **(Supplementary Fig.S7A).** Additionally, after 2 months of culture, we observed vimentin protein expression in our mature esophageal organoids derived from both healthy and patient cells **(Supplementary Fig. S7B).** The influence of mesenchyme on foregut epithelium division has been previously demonstrated (Han et al. 2020). Dysregulation of SOX2 has been linked to mesenchymal development with respiratory characteristics (Teramoto et al. 2020a).

In conclusion, the experimental approach of using EA/TEF derived iPSCs, allowed us to mimic the initial developmental stages of the human esophagus to understand the origins of this malformation. We can conclude that the intrinsic defect observed in these cells are limited only to the esophagus. Our work is limited to isolated type C EA/TEF and thus cannot relate our results to

other types of EA or syndromic EA with associated malformations. More studies are required to understand the failure of key mechanisms and pathways involved during the critical interface of anterior foregut specification. This work therefore highlights the importance of using patient-derived iPSCs to model congenital diseases to yield new insights on organ development during embryogenesis.

Materials and methods

Blood collection for reprogramming

Blood was collected from the three pediatric patients after obtaining consent from their parents to reprogram the blood cells to PSCs for research purposes. This study was approved by the Institutional Review Board of CHU-Sainte Justine Research Center (Protocol #2018-1670 (For iPSCs); 2019-2102 (For WES).

Experimental Design

Using the Institutional iPSC Core facility, we reprogrammed PBMCs from three different EA/TEF type C patients. Control and patient cells were not matched for sex and ethnicity.

All iPSCs used for differentiation was between passages 20 and 35**.** Healthy and patient-derived iPSCs were differentiated simultaneously into mature esophageal organoids for every directed esophageal differentiation. The identity of the samples was not blinded to the investigator. Hepatic and myoblast differentiation were performed in Paganelli and Dumont labs, respectively. iPSCs derived from healthy subjects for esophageal differentiation, hepatic differentiation, and myoblast differentiation are different. The same clones of EA/TEF patient-derived iPSCs were used for all directed differentiations.

Human embryonic stem cell and induced pluripotent stem cells

Human embryonic stem cell (ESC) cell line, H9 was a kind gift from the Andelfinger lab at CHU Sainte Justine Research Center (Wunnemann et al. 2020). Healthy iPSC cell line (GHC4) and

EA/TEF patient-derived iPSCs (EA1, EA2, and EA3) were generated and obtained from the IPSC Core Facility at CHU-Sainte Justine research center. Patient 1 Age: 2 years Sex: Male, Patient 2 Age: 6 years Sex: Male, Patient 3 Age: 18 years Sex: Female.

Culture and expansion of ESCs and iPSCs

Both ESCs and iPSCs were cultured on feeder free and non-xenogeneic conditions. Cells are plated on human vitronectin VTN XF™ (STEMCELL Technologies, Canada. Catalog#100-0763) coated 100mm cell culture dishes. Cells were maintained at 37ºC with 5% CO_2 with daily replacement of Essential 8 media system (Thermofisher, Canada. Catalog#A1517001). Cells were passaged as aggregates every 3-4 days until they reach 60-70% confluency with 0.5mM EDTA diluted in PBS (Thermofisher, Canada. Catalog#15575020).

Differentiation protocol-Preparing cells for differentiation (Day -1 and Day 0)

Two days prior to differentiation (D-1), cells were dissociated into single cells using Accutase™ (Stem Cell technologies, Canada. Catalog# 07922) and transferred onto BIOLAMININ 521 LN (LN521, BioLamina, Sweden. Product# LN521-02) coated plates with E8 media supplemented with 10uM Rock inhibitor Y-27632 (Sigma-Aldrich, US. Product# SCM075). The following day (D0) (12-16 hours later), media was changed with E8 only. If survival rate of the cells were less than 50%, they were cultured for an additional 24 hours before starting the differentiation. ESC and iPSCs were maintained at 37ºC with 5% CO_2 throughout the differentiation process.

Endoderm differentiation (Day 1 to Day 3)

We modified and adapted previously published protocol for endoderm differentiation (Matsuno et al. 2016). Xeno-free media (XFM-) was prepared using 500 mL of RPMI 1640 medium without L-Glutamine (Thermo Fisher Scientific, Canada. Catalog# 11875101), 10mL B-27 Supplement minus insulin (Thermo Fisher Scientific, Canada. Catalog# A1895601), 5mL GlutaMax™ (Thermo Fisher Scientific, Canada. Catlog# 35050061), 5mL of KnockOut™ serum replacement (Thermo Fisher Scientific, Canada. Catalog# 10828010), 5mL Penicillin-Streptomycin (10,000

U/mL) (Thermo Fisher Scientific, Canada. Catalog# 15140148), 7.5mL HEPES (1M) Buffer (Thermo Fisher Scientific, Canada. Catalog# 15630130), 5mL of MEM Non- essential amino acids (100X) ((Thermo Fisher Scientific, Canada. Catalog# 11140050). Day 1 of differentiation cells are first washed with XFM- media to remove any residual E8 media, then cultured in XFM- with 100ng/mL Activin A (R&D systems, US. Catalog# 338-AC-010/CF) and 3uM CHIR99021 (Stem Cell technologies, Canada. Catlog# 72052). On Days 2 and 3 of culture, cells were first washed with XFM- media and the culture was continued with XFM- supplemented with 100ng/mL Activin A and 250nM of LDN193189 (Stemgent, US. Code# 04-0074). By day 3, a 70-80% confluent monolayer of endodermal cells should be observed under the microscope.

Anterior foregut differentiation (Day 4 and Day 5)

Endodermal cells were first washed with XFM- media and then cultured for 24 hours in XFM- supplemented with 1uM A8301 (Stemgent, US. Code# 04-0014) and 250nM of LDN193189. The following day cells were washed XFM- media and cultured in XFM- supplemented with 1uM A8301 and 1uM IWP2 (Stemgent, US. Code# 04-0034).

Esophagus differentiation (Day 6 to Day 24)

From day 6 to day 16 we switch to XFM+ containing the same components as XFM- and replacing b27 supplement minus insulin with b27 supplement with insulin (Thermofisher, Canada. Catalog# 17504044). To induce esophageal fate, we modified and adapted previously published protocol (Zhang et al. 2018; Trisno et al. 2018). Anterior foregut cells were cultured in XFM+ supplemented with 1uM A8301 and 250nM LDN193189 from day 6 until day 16, changing media daily. By day 16, we observed that cells had reached 100% confluency and the presence of dense cell clusters. On day 16 esophageal progenitor cells are cultured in XFM+ only until day 24.

Esophageal Organoid formation (2 months)

Organoids were generated in suspension using Nunclon™ Sphera™ low attachment 96-well plates (Thermofisher, Canada. Catalog# 174930) by modifying previously published esophageal studies

(Giroux et al. 2017; DeWard, Cramer, and Lagasse 2014). On day 24 of esophageal differentiation, cells were detached using TrypLE (Thermo Fisher Scientific, Canada. Catalog# 12604013) and gently resuspended in XFM+ media. Viable cells were counted using Trypan blue solution (Thermofisher, Canada. Catalog# 15250061) and 50,000 cells are then aliquoted in each well of the 96-well plate containing XFM+ supplemented with 1uM A8301, 250nM LDN, 3uM CHIR99021, 20ng/mL FGF2/bFGF (Peprotech, US. Catalog# AF100-18B) and 200ng/mL EGF (Thermofisher, Canada. Catalog# PHG0313).

Other organ-lineage differentiation

- **Trachea differentiation (Day 6 to Day 16)**

To induce tracheal fate, we again modified previously published protocols (Huang et al. 2014; Huang et al. 2015). Anterior foregut cells were cultured from day 6 to day 16 in XFM+ supplemented with 3uM CHIR99021, 10ng/mL human FGF10 (R&D systems, US. Catalog# 345-FG-025/CF), 10ng/mL human FGF7 (R&D systems, US. Catalog# 251-KG-010/CF), 10ng/mL BMP4 (Peprotech, US. Catalog# 120-05) and 50 nM RA (Tocris, UK. Catalog# 0695) from day 6 to day 16 changing media daily. Cells reached 100% confluency and formed two-layered cell clusters.

- **Hepatoblast differentiation (Day 0 to day 15)**

iPSCs were dissociated by TrypLE (Life Technologies, Catalog# 12604013) to single cells and seeded on human recombinant laminin 521 (Biolamina)-coated plates in Essential 8 Flex medium at a density of $7x10^{\wedge 5}$ cells/cm^2. Differentiation starts (day 0) when the cells reach around 70% confluency by changing the medium to RPMI B27 minus insulin (Life technologies) supplemented with 1% knockout serum replacement (KOSR, Life technologies, Catalog# 10828010). For the first 2 days, the cells were exposed to 100ng/mL Activin A (R&D systems, Catalog# 338-AC-010/CF) and 3uM CHIR99021 (Stem Cell Technologies, Catalog #72052) and then for the following 3 days to 100ng/mL Activin A alone. Subsequently, RPMI B27 minus insulin medium was supplemented with 20ng/mL BMP4 (Peprotech, Catalog# 120-05), 5ng/mL bFGF (PeproTech, Catalog# AF100-18B) and 4uM IWP2 (Tocris, Catalog# 3533) and 1uM A83-01(Tocris, Catalog# 2939/10) for 5 days, with daily medium change. At day 10, the medium was changed to RPMI B27 (Life

technologies, Catalog# 17504044), supplemented with 2% KOSR, 20ng/mL BMP4, 5ng/mL bFGF, 20ng/mL HGF (PeproTech, Catalog #100-39H) and 3uM CHIR99021 for 5 days, with daily medium change.

- **Myoblast differentiation (Day 0 to Day 20)**

The generation of iPSC-derived myoblast was adapted from a protocol published by M. Shelton *et al* with minor modifications (Shelton et al. 2016). Briefly, different growth factors and inhibitors are sequentially used to drive iPSCs toward the mesodermal lineage and promote their myogenic cell fate commitment. One day prior differentiation, human iPSCs (patients and control) were dissociated with TrypLE™ (Gibco, Catalog# 12604013) and 10^5 cell/well were plated as small colonies (10-20 cell/colony) on Vitronectin-coated 12-well plates using mTeSR1 media (STEMCELL Technologies, Catalog# 12604013) supplemented with 10 µM ROCK inhibitor (Y-27632, STEMCELL Technologies, Catalog# 12604013). Next day, the medium was changed to TeSR-E6 media (STEMCELL Technologies, Catalog#05946) supplemented with 7 µM CHIR 99021 (STEMCELL Technologies, Catalog #72052) for 3 days. After three days of CHIR99021 treatment, cells were gently washed with DPBS and cultured only in TeSR-E6 medium without any CHIR99021, and media was changed every day till day 7. At this time point, a broad expression of the somite markers PAX3 and MEOX1 can be detected. From days 10 to 20 of differentiation, 5ng/mL FGF2 (Wisent) is added to the TeSR-E6 medium to promote myogenic cell proliferation. At day 20, a significant proportion of cells express the muscle stem cell marker PAX7.

RT-qPCR

At each developmental stage (definitive endoderm, anterior foregut, esophageal progenitors, mature esophageal epithelium/organoids and other organ-lineages), cells were detached using Accutase™ and RNA was extracted using the Promega kit ReliaPrep ™ RNA Cell Miniprep System (Promega, US. Catalog# Z6011). RNA was reverse transcribed using Omniscipt ™ RT kit (Qiagen, US. Catalog# 205113) and the complementary DNA (cDNA) obtained was used for real time quantitative PCR using LightCycler instrument (Roche Life Science, Germany). cDNA was quantified using TaqMan Gene expression assays and the TaqMan primers to target genes were purchased from Thermo Fisher Scientific listed in **Table 1**. The transcript level of each gene was

normalized to GAPDH housekeeping gene using the 2^-DDCT method. Relative gene expression was calculated and reported as fold change compared to indicated samples using GAPDH normalized transcript level. The results include a mean of at least 3 technical replicates for each biological sample.

Immunofluorescence and microscopy imaging

For immunofluorescence staining, cells at each developmental stage (definitive endoderm, anterior foregut, esophageal progenitors, mature esophageal epithelium/organoids, and other organ-lineages) were fixed with 4% paraformaldehyde (PFA) (Thermo Scientific, Canada. Catalog# AAJ19943K2) for 20 minutes at room temperature then washed 3 times with DPBS. Cells were then permeabilized with 0.4% Triton X-100™ (Sigma- Aldrich, US. Catalog# 9002-93-1) in DPBS for 25 minutes at room temperature followed by washing with DPBS. Cells were then incubated with 3% blocking serum in DPBS for 1 hour at room temperature. Primary antibody was added to antibody dilution buffer (1% PBS-BSA, 0.3% Triton X-100, 0.3% serum) and were incubated overnight at 4°C. The following day, the cells were washed 3 times with DPBS. Secondary antibody was diluted in the same antibody dilution buffer, added to cells and incubated for 1 hour at room temperature (primary and secondary antibodies listed in **Table 3**). Following washing with DPBS, cells were stained with DAPI for 15 minutes at room temperature. Cover slips are mounted on top of a drop (7-8 uL) of ProLong™ Diamond Antifade Mountant (Thermo Fisher Scientific, Canada. Catalog# P36970).

RNA-sequencing assay

RNA was extracted as described in the RT-qPCR section which was followed by library preparation for Nanopore Sequencing was done using two different protocols. RNA (50 ng total RNA) was spiked reverse transcribed and amplified by PCR following the manufacturer's instructions for the cDNA-PCR Sequencing kit (SQK-PCS109, [Oxford Nanopore Technologies, UK) up until the PCR step (14 cycles, 500 seconds extension). The PCR reactions were then prepared for sequencing using the Genomic DNA by Ligation kit (SQK-LSK109 [Oxford Nanopore Technologies, UK]). Concentrations were quantified for RNA after elution and for

cDNA after the PCR step and before loading the flow cells using a Qubit Fluorometer with a Qubit RNA assay kit (High sensitivity) for RNA and a Qubit dsDNA assay kit (Broad range) for cDNA [Invitrogen, Waltham, Massachusetts, USA]. Final libraries, with the following cDNA concentrations: 26.2 ng/uL (Patient 1), 29 ng/uL (Patient 2), 23.8 ng/uL (Patient 3), 28ng/uL (iPSC), 13.4 ng/uL (H9), were loaded onto MinION flow cells (R9.4.1 - FLO-MIN006) and ran for 74 hours on GridION and MinION Mk1C sequencers. When required, the sequencing runs were refueled with 250 uL of FB buffer.

RNA-sequencing analysis

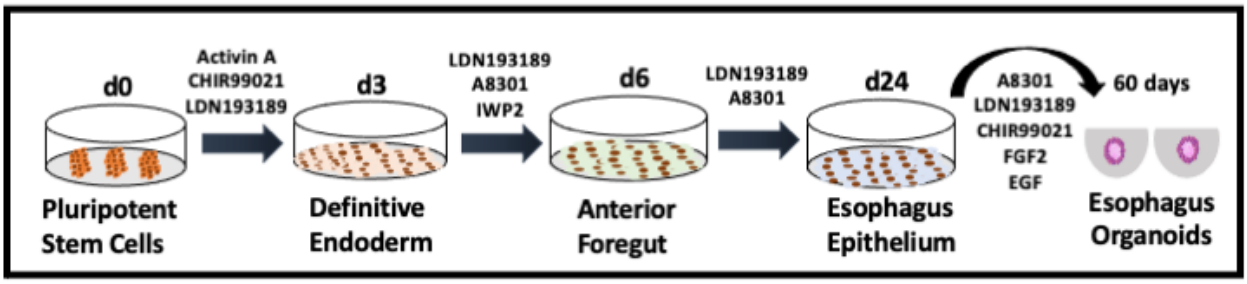
Raw fast5 files were basecalled during the sequencing run using Guppy v4.0.11 [https://nanoporetech.com/] in high accuracy mode. Fastq_pass and fastq_fail files for each sample were submitted to Pychopper v2.4.0 [https://github.com/nanoporetech/pychopper] to identify full-length reads, split ligation concatemers, rescue fused reads and reorient reads based on the stranded barcode adapters. Full length and rescued reads were then aligned to the human reference genome (GRCh38.p13)(Frankish et al. 2019), using Minimap2 v2.18(Li 2018) with -a -x splice --MD --secondary=no options. Alignments were converted to bam format and sorted with samtools v1.12 (Li et al. 2009). Resulting bam files were merged using "samtools merge" before being used as input for de novo assembly with Stringtie2 v2.1.4 (Kovaka et al. 2019). Gencode reference transcriptome v37 gtf format was used as input for Stringtie2's -G option and all transcripts were collapsed with the long reads -L parameter. GFFcompare v0.1.12.2(Pertea and Pertea 2020) was used to map the resulting GTF file to Gencode to evaluate the assembly and filter it. Transcript per million (TPM), classcode and exon number filters have been applied manually to select all isoforms that are TPM>0.2, "=" or "c" classcode (provided by GFFcompare) and all classcode if more than 1 exon. Fasta sequence corresponding to the filtered assembly annotation were retrieved using GFFread v0.12.7 (Pertea and Pertea 2020) and the samples fastq sequences obtained after pychopper were aligned again using minimap2 with -k 14 -a -N 100 options and the fasta sequence from the filtered assembly as a reference. Bam files were obtained using samtools and isoforms were quantified using Salmon v1.5.2 (Patro et al. 2017) in quant mode with -l SF –noErrorModel –noLengthCorrection options as recommended for Nanopore long reads. Read counts from all samples were merged into the same matrix using Salmon quantmerge option. Isoforms with more

than 3 samples with null read counts were filtered out manually to avoid artefacts. Normalization and differential expression analysis was performed using DESeq2 [Bioconductor version 3.13] (Love, Huber, and Anders 2014) on R v4.1.0 (R Core Team, 2021) to generate a normalized count matrix and statistics on differential expression. Batch effect correction was done using the sva package (Leek et al. 2021). Isoforms with a p-value < 0.01 and a |log2(FoldChange)|>1 after batch correction were considered statistically significant in the differential expression analysis. Isoforms with a p-value<0.01 and |log2(FoldChange)|>0.5 were used to perform a GO enrichment analysis with GOrilla from Gene Ontology (Eden et al. 2009).

Quantification and statistical analysis

All data quantification is presented as the mean ± SEM using GraphPad Software Prism 6. Statistical significance was determined by Student's t tests, Mann-Whitney test. When more than two groups are compared, multiple comparisons were performed using one-way ANOVA to compare the two groups. For each analysis, at least 3 technical replicates of each biological cell lines were included. Representative pictures shown are indicated in the legends. P values of 0.05 or less were considered statistically significant.

Scheme 1



Scheme 1: Stepwise differentiation protocol of human pluripotent stem cells into esophagus organoids. Illustration of different signaling pathways manipulated to differentiate pluripotent stem cells into each developmental stage starting from definitive endoderm, anterior foregut, esophagus epithelium and esophagus organoids.

Activin A: dimeric growth and differentiation factor activates Nodal/TGFb pathway

CHIR99021: aminopyrimidine derivative that is potent to GSK3 inhibitor, a key inhibitor of the WNT pathway. CHIR99021 activates the WNT pathway

LDN193189: dihydrochloride potent and selective ALK2 and ALK3 inhibitor. It inhibits BMP4 mediated Smad1/5/8 activation

A8301: potent inhibitor of the TGFb type I receptor ALK5 kinase, type I activin/nodal receptor ALK4 and type I nodal receptor ALK7

IWP2: inhibits WNT pathway at the level of the pathway activator Porcupine, leading to WNT secretion and signaling capability

FGF2: growth factor belonging to the FGF superfamily. It stimulates cell proliferation.

EGF: potent growth factor belonging to the EGF family. It induces cell proliferation, differentiation, and survival.

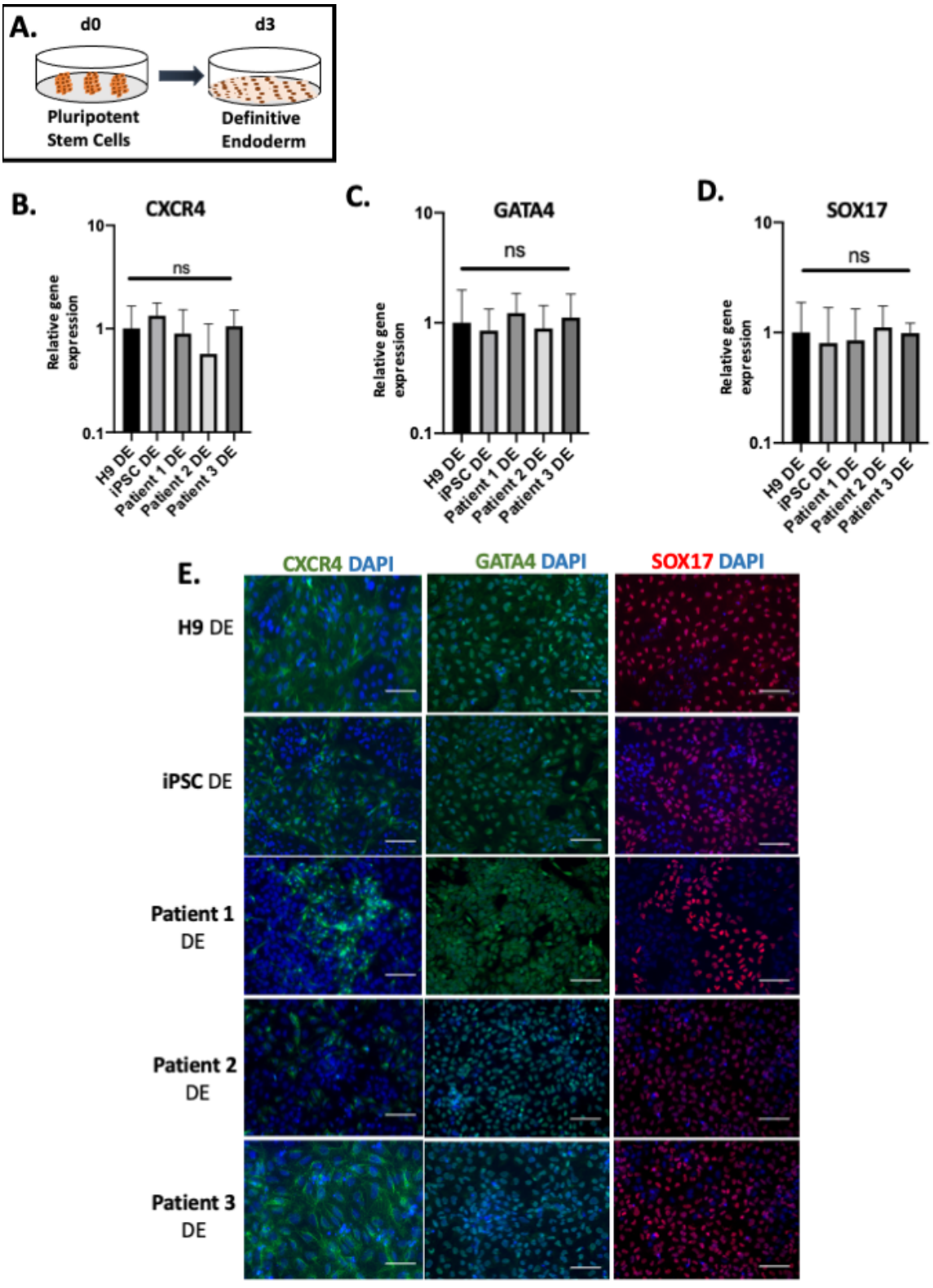


Figure 1: Differentiation of healthy and EA/TEF patient-derived pluripotent stem cells into definitive endoderm (DE). A) Illustration of first stage of differentiation into definitive endoderm.

B, C and D) Expression of key endodermal markers *CXCR4, GATA4* and *SOX17* was quantified by qPCR and reported as fold change. The fold change was generated by normalizing the transcript levels to those of healthy (H9) DE cells. Data representative mean ±SEM (3 technical replicates (3 different wells) from each of the 5 biological cell lines differentiated at the same time). **E)** DE cells from healthy and patients express similarly CXCR4, GATA4 and SOX17 at the protein level by immunofluorescence staining. Negative controls were included in each staining. Scale bar 50um.

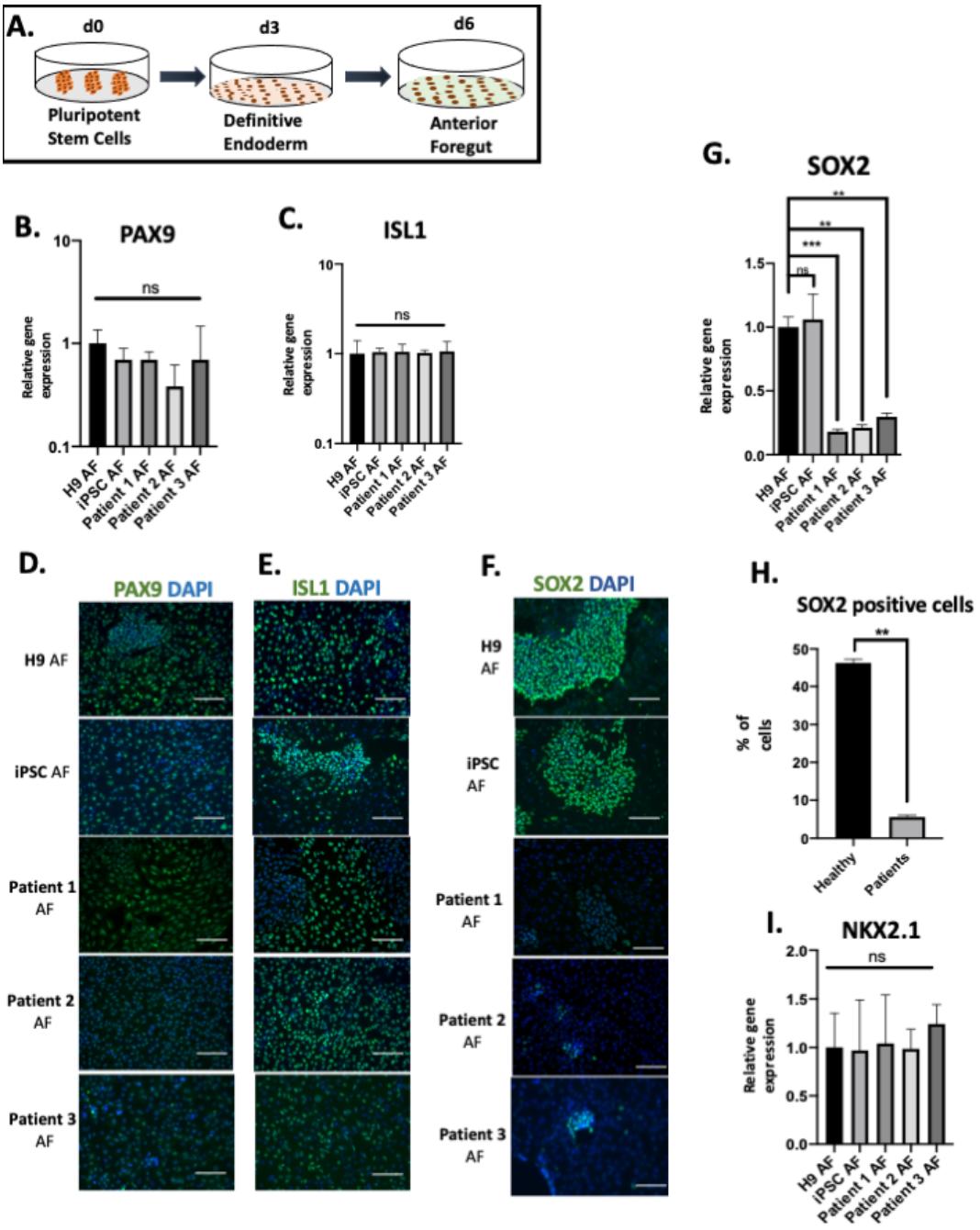


Figure 2: SOX2 expression significantly downregulated in patient-derived Anterior foregut (AF) cells. A) Schematic representation of differentiation from definitive endoderm into anterior foregut. **B and C)** *PAX9* and *ISL1* are expressed similarly by qPCR in both healthy and patient derived anterior foregut cells. The fold change was generated by normalizing the transcript levels to those of healthy (H9) DE cells. Data representative mean ±SEM (n>3 technical replicates for

each biological cell line). **D and E)** Healthy and patient-derived anterior foregut cells express PAX9 and ISL1 nuclear staining by immunofluorescence similarly. **F)** Low expression of SOX2 at the protein level by immunofluorescence in all 3 patients. To note, patients 2 and 3 have fewer cells with bright fluorescence and patient 1 has more cells with less fluorescence **G)** *SOX2* is downregulated in all three patient-derived anterior foregut cells. Transcript levels were compared to those of H9 AF. Data represents mean ± SEM (3 technical replicates (3 different wells) from each of the 5 biological cell lines differentiated at the same time) **$p<0.01$, *** $p<0.001$ by unpaired two tailed student's t test **H)** Cell counting of positive cells using Image J software to count positively stained cells revealed that less than 10% of the patient AF cells are positive for SOX2. Scale bar 50um. **I)** Lack of expression of NKX2.1 at the AF stage in both groups. The fold change was generated by normalizing the transcript levels to those of healthy (H9) DE cells. Data representative mean ±SEM (3 technical replicates (3 different wells) from each of the 5 biological cell lines differentiated at the same time).

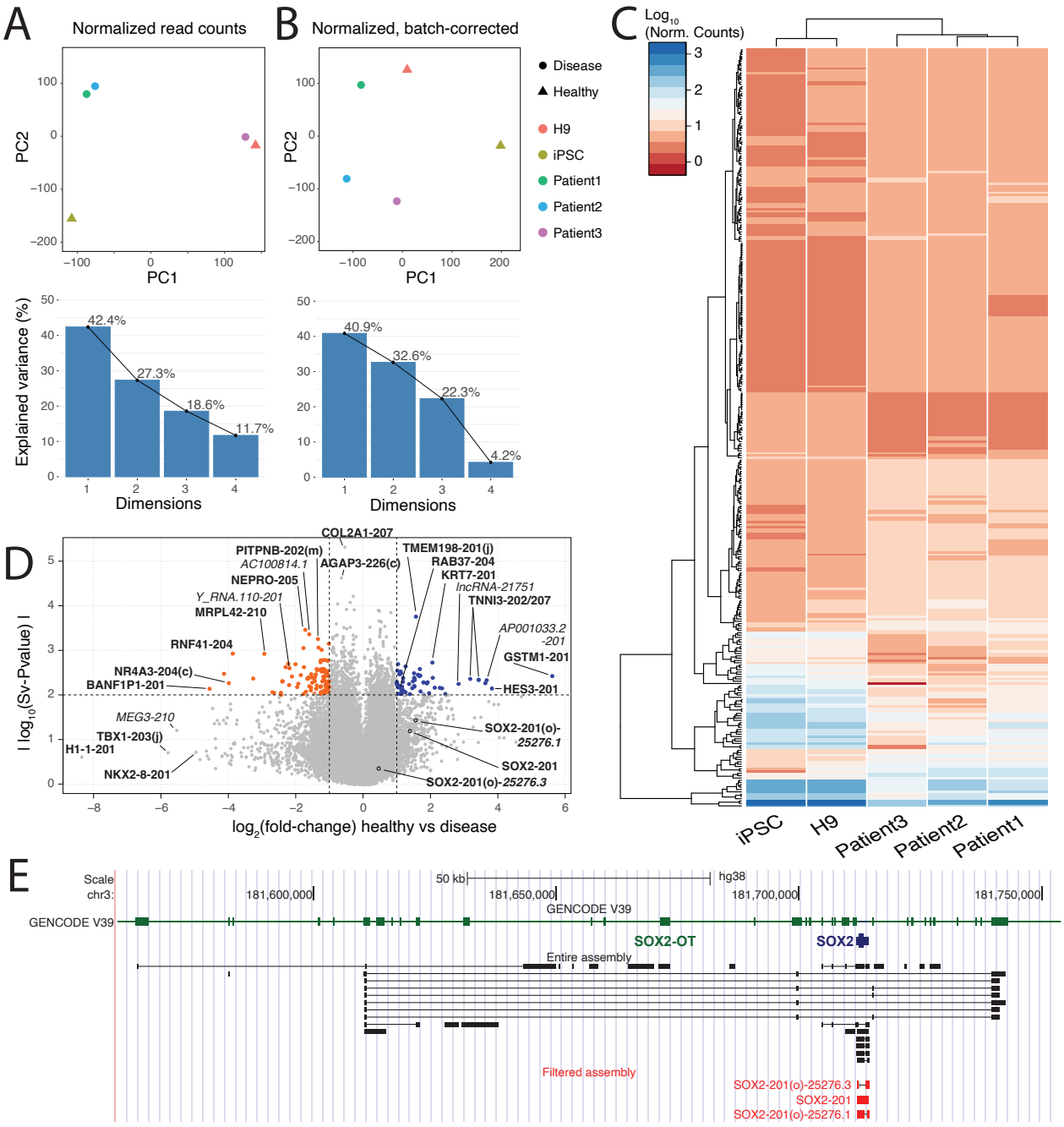


Figure 3: De novo assembly from long read RNA sequencing of patient versus healthy cells. **A)** Principal component analysis of sequencing data before and after **(B)** replicate- and sex-specific batch correction. **C)** Heatmap of **173 transcript isoforms with** batch-corrected P-value< 0.01 and **absolute fold-change > 2. D)** Volcano-plot of differentially expressed transcripts (horizontal threshold set at batch-corrected P-value = 0.01 and vertical thresholds set at log2 fold-change ±1).

E) UCSC Genome Browser view of the Sox2 locus displaying collapsed reference transcriptome (top panel), all *de novo* assembled transcript isoforms from this study (middle panel) and filtered isoforms (bottom panel)

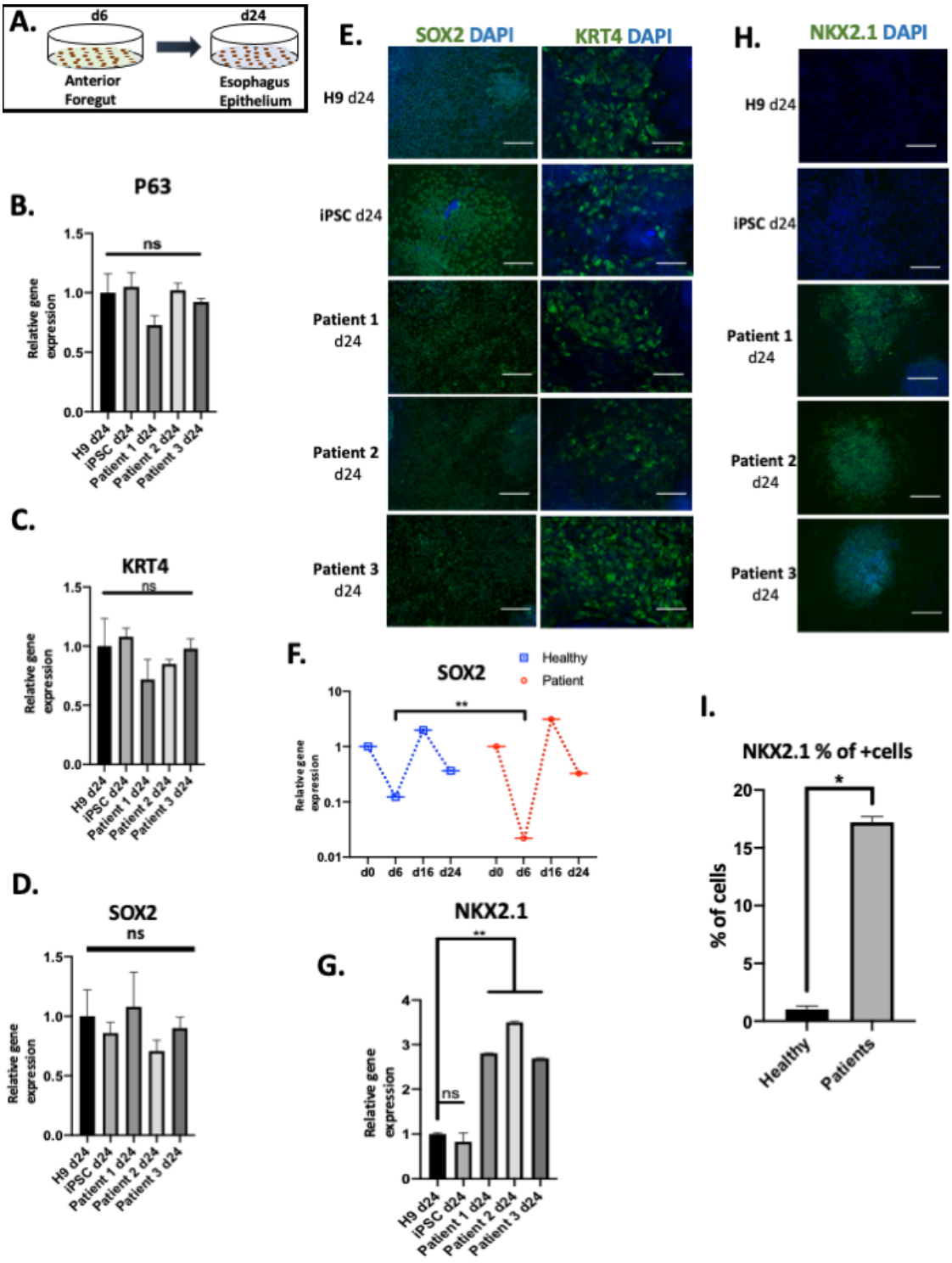


Figure 4.: Esophagus epithelial cells derived from EA/TEF patients express NKX2.1, a tracheal marker. A) Illustration representing the differentiation of anterior foregut cells into esophagus epithelium **B, C and D)** Transcript levels of *p63, KRT4* and *SOX2* shows similar expression in both healthy and patient groups. Relative expression was compared to that of H9 d24

esophagus epithelium. The fold change was generated by normalizing the transcript levels to those of healthy (H9) DE cells. Data representative mean ±SEM (3 technical replicates (3 different wells) from each of the 5 biological cell lines differentiated at the same time)**. E)** Immunofluorescence staining of d24 esophagus epithelium confirming the expression of SOX2 and KRT4 similarly in healthy and patient derived esophagus epithelium **F)** Relative expression of *SOX2* throughout differentiation. The fold change was generated by normalizing the transcript levels to those of healthy (H9) DE cells. Data representative mean ± (3 technical replicates (3 different wells) from each of the 5 biological cell lines differentiated at the same time) **p<0.01, *** p<0.001 by unpaired two tailed student's t test. **G)** Abnormal expression of *NKX2.1* in patient derived esophagus epithelium. Transcript levels were significantly higher in all 3 patient cells. Transcript levels were compared to that of H9 esophagus epithelium Data represents mean ±SEM (3 technical replicates (3 different wells) from each of the 5 biological cell lines differentiated at the same time) **p<0.01, *** p<0.001 by unpaired two tailed Student's t test. **H)** At the protein level, NKX2.1 was also expressed in patient-derived esophagus epithelial cells, confirming our findings at the RNA level. **I)** Cell counting of NKX2.1 positive cells by the software Image J counting cells positively staining for NKX2.1. Scale bar 50um

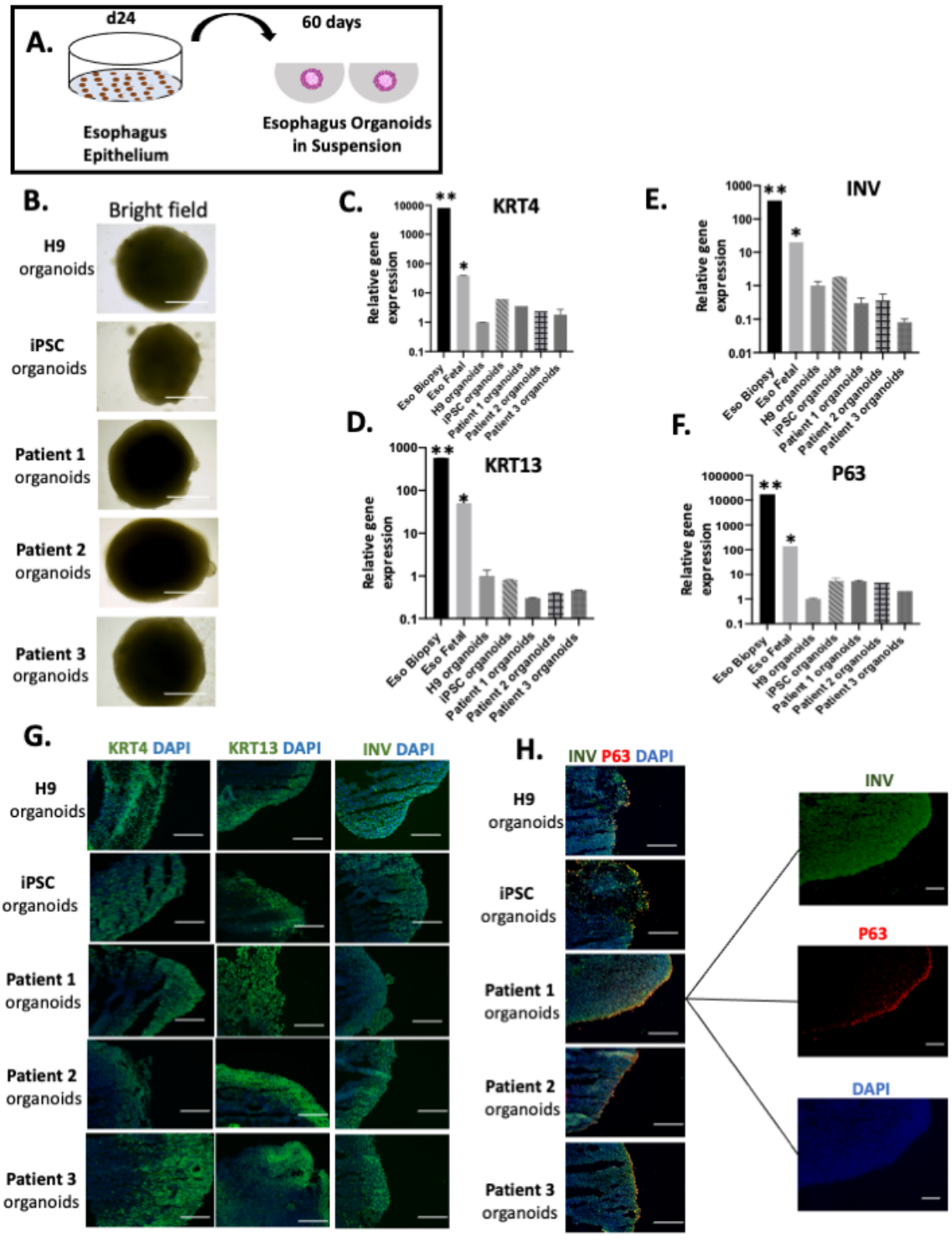


Figure 5: Healthy and EA/TEF patient PSCs can generate mature esophagus organoids. A) Illustration of differentiation from esophagus epithelium progenitors into esophagus organoids. **B)**

Bright field images of healthy and patient-derived esophagus organoids reveal similar morphology between the two groups. **C, D, E, F)** transcript levels of mature esophagus markers such as *KRT4, KRT13, INV* and *p63* reveal a similar expression between healthy and patient-derived esophagus organoids. Relative expression was compared to H9 derived esophagus organoids. Two references were included in the graphs, the esophagus biopsy and fetal esophageal tissue. Data represents mean ±SEM (3 technical replicates (3 different wells) from each of the 5 biological cell lines differentiated at the same time) **$p<0.01$, *** $p<0.001$ by unpaired two tailed student's t test. **G)** Immunofluorescence staining for KRT4, KRT13 and INV showing a positive expression in all generated esophagus organoids **H)** Dual immunofluorescence staining for INV and p63 showing a basal proliferative layer positive for p63 and a suprabasal layer positive for INV, a specific marker for stratified esophagus epithelium. Scale bar 50um.

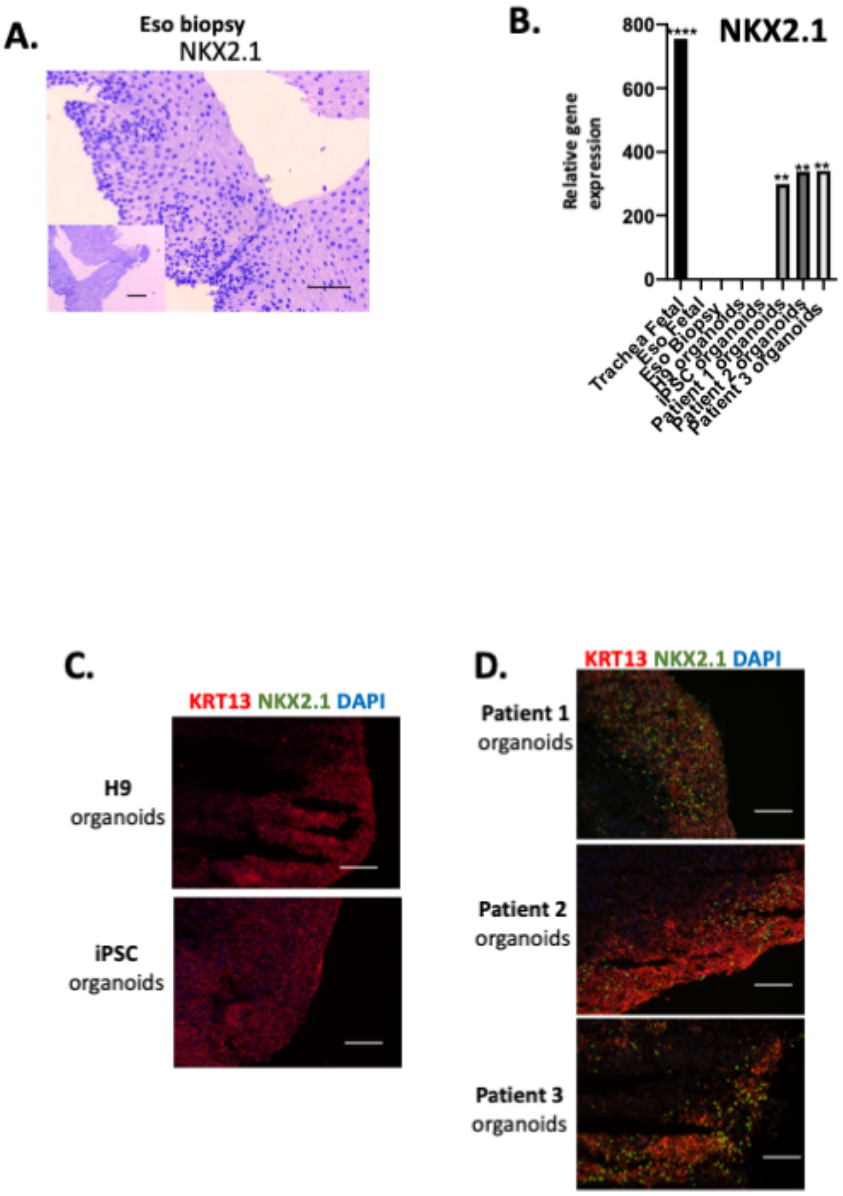


Figure 6.: Retained abnormal expression of NKX2.1 in patient-derived esophagus organoids **A)** Immunohistochemistry of esophagus biopsy against NKX2.1 **B)** Transcript levels of *NKX2.1* is significantly higher in all 3 patient-derived esophagus organoids. Expression of NKX2.1 was absent in healthy derived organoids and esophagus biopsies. Transcript levels were compared to those of H9 derived organoids. Fetal trachea was included as a reference in our analysis. Data represents mean ±SEM (3 technical replicates (3 different wells) from each of the 5 biological cell

lines differentiated at the same time) **p<0.01, *** p<0.001 by unpaired two tailed Student's t test, **C and D)** Dual staining by Immunofluorescence confirms positive KRT13 staining in both groups, however, NKX2.1 is expressed in all 3 patient-derived esophagus organoids and is absent in healthy organoids. Scale bar 50um.

Table 1: List of Taqman gene expression assays for qPCR analysis

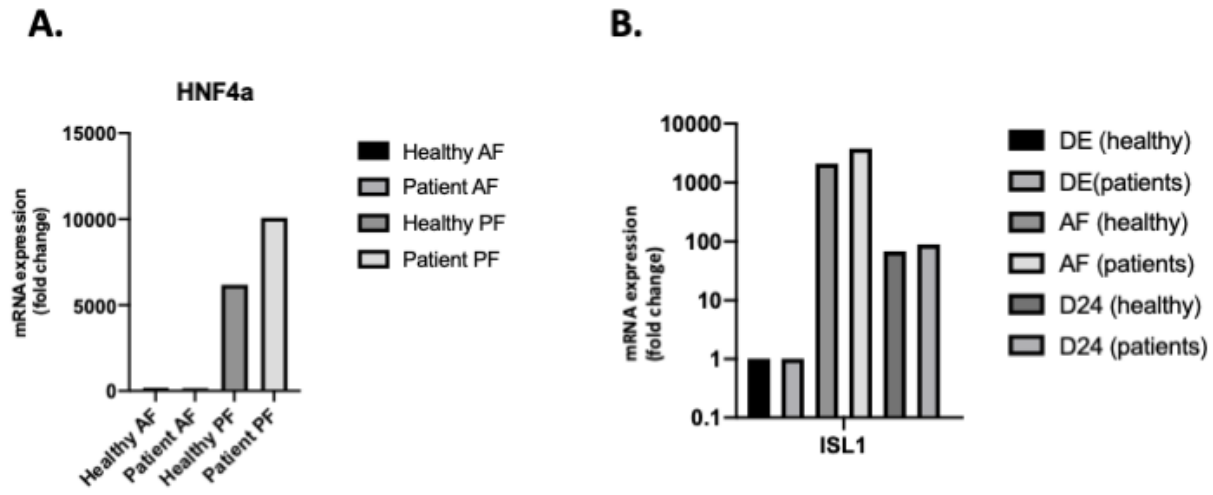
Gene name	**Source**	**Identifier**
CXCR4	ThermoFrisher	Hs00607978_s1
GATA4	ThermoFrisher	Hs00171403_m1
SOX17	ThermoFrisher	Hs00751752_s1
SOX2	ThermoFrisher	Hs04234836_s1
PAX9	ThermoFrisher	Hs00196354_m1
ISL1	ThermoFrisher	Hs00158126_m1
NKX2.1	ThermoFrisher	Hs00968940_m1
KRT4	ThermoFrisher	Hs00361611_m1

KRT13	ThermoFrisher	Hs02558881_s1
P63	ThermoFrisher	Hs00978340_m1
INV	ThermoFrisher	Hs00846307_s1

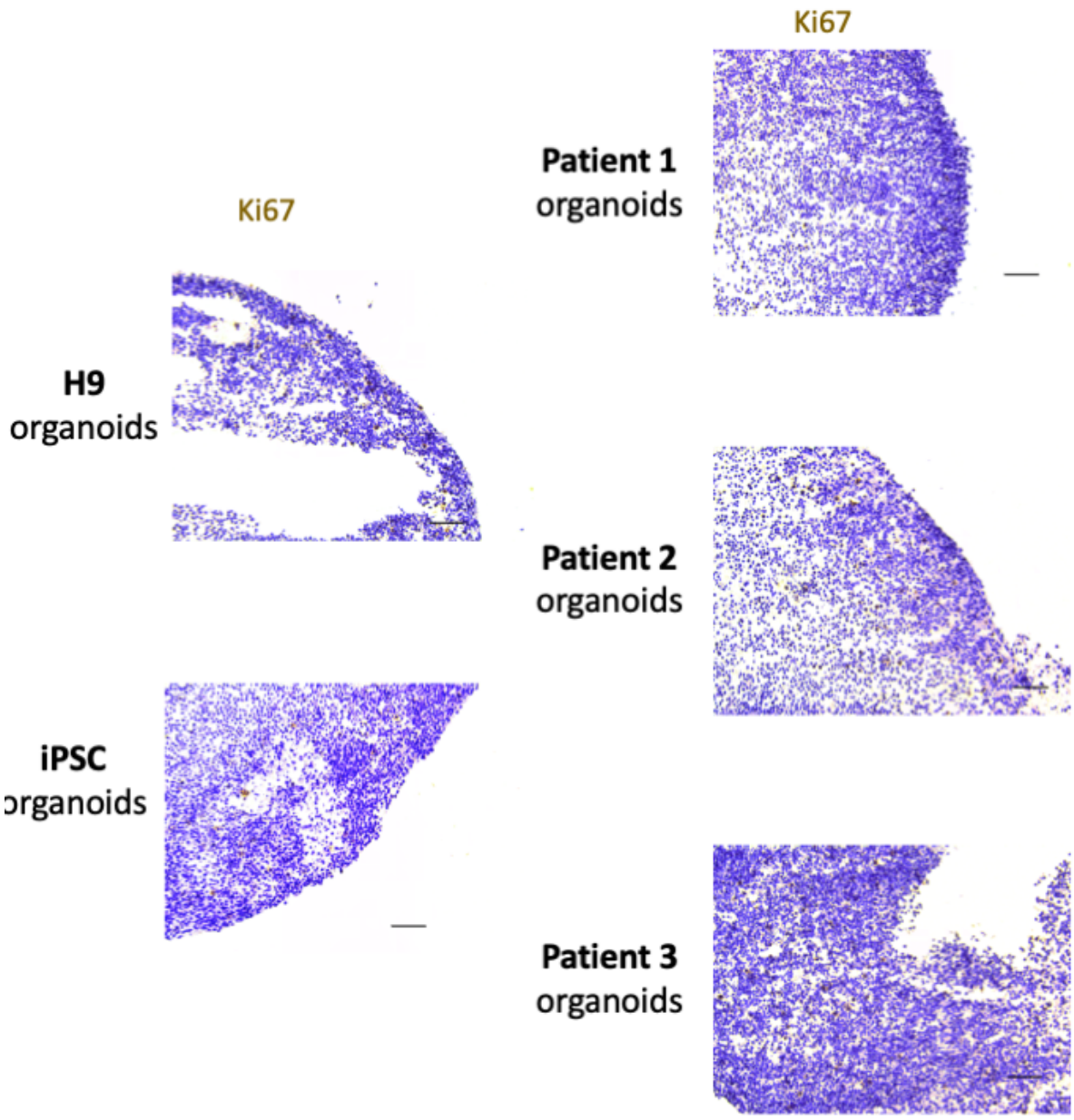
Table 2: List of primary and secondary antibodies

Antibodies	**Source**	**Identifier**
Mouse monoclonal antibody to TFF1	Abcam	ab72876
Rabbit monoclonal antibody to CXCR4	Abcam	ab181020
Mouse monoclonal antibody to SOX2	Abcam	ab79351
Rat monoclonal antibody to PAX9	Abcam	ab28538
Rabbit monoclonal antibody to cytokeratin 4	Abcam	ab51599
Mouse monoclonal antibody to SOX17	Abcam	ab84990
Rabbit monoclonal antibody to ISL1	Abcam	ab109517
Rabbit polyclonal antibody to GATA4	ThermoFisher	PA1-102
Mouse monoclonal antibody to INV	Abcam	ab68
Rabbit polyclonal antibody to P63	Abcam	ab53039
Rabbit monoclonal antibody to KRT13	Abcam	ab92551
Rabbit polyclonal antibody to AFP (ready to use)	DAKO omnis	GA50061-2
Secondary antibodies		
Goat anti-Mouse IgG H&L (Alexa Fluor 488)	Abcam	ab150113
Goat anti-Mouse IgG H&L (Alexa Fluor 594)	Abcam	ab150120
Donkey anti-Rabbit IgG H&L (Alexa Fluor 488)	ThermoFisher	A21206
Goat anti-Rabbit IgG H&L (Alexa Fluor 488)	Abcam	ab150077

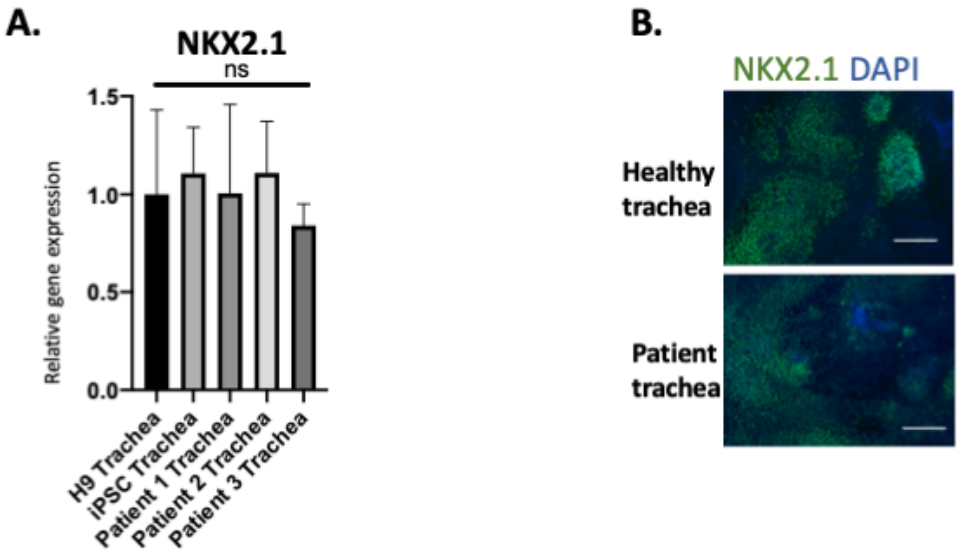
Supplementary Figures:



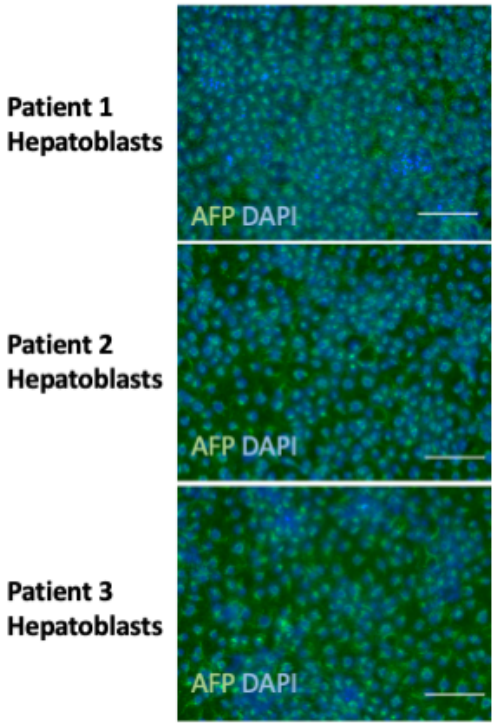
Supplementary figure S1: mRNA expression levels of *HNF4a, ISL1* in healthy and patient derived cells. A) Expression levels of *HNF4a*, a posterior foregut marker, by qPCR is absent in healthy and patient derived anterior foregut cells compared to posterior foregut cells. mRNA levels were represented by the fold change compared to healthy posterior foregut cells**. B)** Expression of *ISL1* mRNA levels by qPCR at definitive endoderm, anterior foregut and esophagus epithelium at day 24 are similar between healthy and patient cells. Fold change relative expression was compared to healthy definitive endodermal cells **n=3 biological replicates.**



Supplementary figure S2: Healthy and Patient-derived organoids can proliferate after 2 months of culture. Immunohistochemical staining for ki67, a proliferative marker, shows a similar proliferative capacity in both healthy and patient derived organoids after 2 months of culture. Scale bar 50um.

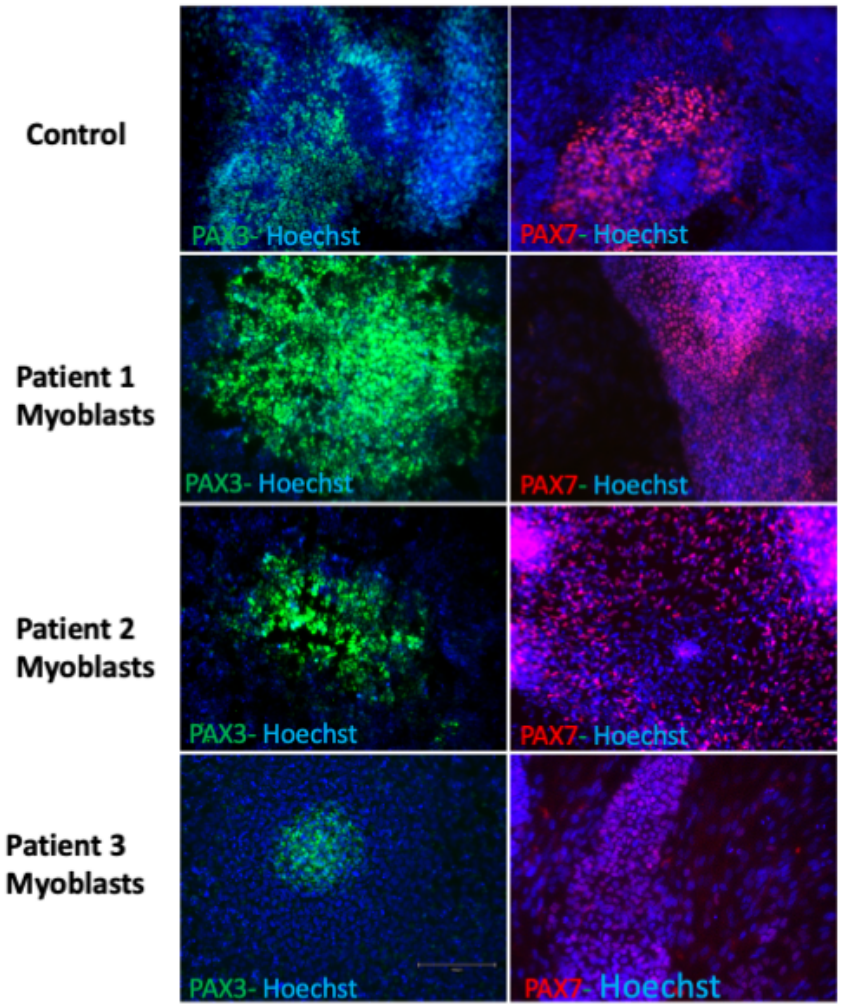


Supplementary figure S3: Healthy and Patient-derived iPSCs can generate tracheal epithelium. A) Expression of *NKX2.1* transcription factor in tracheal epithelial cells derived from healthy and patient iPSCs reveal similar expression levels. mRNA relative expression was compared to the healthy group. Data represents mean +- SEM (3 technical replicates (3 different wells) from each of the 5 biological cell lines differentiated at the same time)**. B)** Immunofluorescence staining of tracheal epithelium from healthy and patient cells reveal expression of NKX2.1 in both groups. Negative controls were included for each staining. Scale bar 50um.



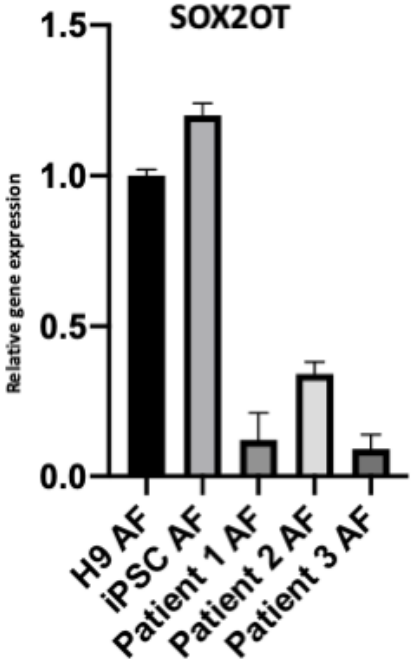
Supplementary figure S4: Hepatic differentiation of healthy and patient derived iPSCs.

Immunofluorescence staining of all 3-patient derived hepatoblasts reveal expression of AFP a specific hepatic marker. Scale bar 50um



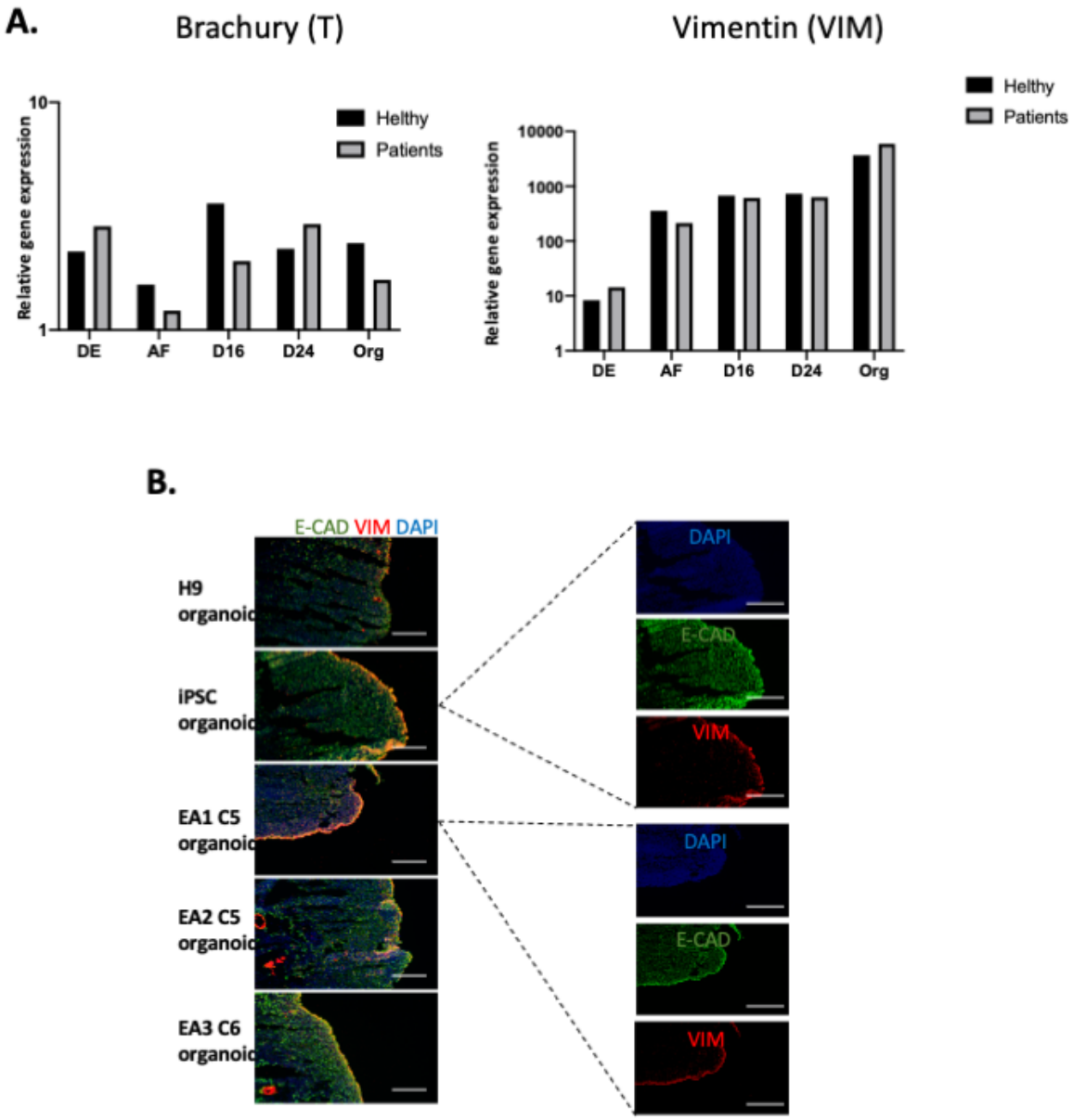
Supplementary figure S5: Differentiation of iPSCs into myoblast cells.

Immunofluorescence staining of 3 patient-iPSC derived myoblasts reveal similar protein expression of PAX3 and PAX7 similarly to the control group. Scale bar 100um



Supplementary figure S6: Long non-coding RNA *SOX2OT* a potential regulator of SOX2 at the anterior foregut stage.

mRNA levels by qPCR reveal significant downregulation of *SOX2OT* in all 3 patient-derived anterior foregut cells. Relative expression was compared to H9 AF cells. N=3 biological replicates.



Supplementary figure S7: 2 months mature esophagus organoids express mesenchymal markers brachyury and vimentin. A) Transcript levels of Brachyury (*T*) and Vimentin (*Vim*) throughout esophagus differentiation is similar in both healthy and patient derived cells. Transcript levels were compared to those of day 0 iPSCs n=3 biological replicates. B) Healthy and patient derived esophagus organoids express by immunofluorescence vimentin VIM in the outermost layer of the esophagus whereas the rest of the esophagus is positive for E-cadherin, an epithelial marker.

Esophagus organoids are organized with an inner core of epithelial cells and outer layer of mesenchymal cells. Scale bar 50um.

Supplementary Table 1a: 40 top differentially expressed transcripts in patient-derived anterior foregut cells.

Long read single molecule cDNA sequencing analysis revealed a list of 173 transcripts differentially expressed in patient derived anterior foregut cells. The 40 most differentially expressed are described here (40 transcripts with the highest fold change value and batch-corrected P-value < 0.01).

Transcript Name	log2 FC	p-value	Gene Name	Linked disease	Function
GSTM1-201 (=) - STRG.5167.1	5.61	0.004	Glutathione S-Transferase Mu 1	Asbestosis and Oral Leukoplakia	Glutathione conjugates formation
BANF1P1-201 (=) - STRG.80167.2	-4.56	0.007	Barrier To Autointegration Factor 1 Pseudogene 1	Leiomyoma	Pseudogene
LRRC37B-211 (o) - STRG.92784.13	-4.13	0.003	Leucine Rich Repeat Containing 37B	Chromosome 17Q11.2 Deletion Syndrome and Chromosome 15Q26-Qter Deletion Syndrome	Protein Coding gene
NR4A3-204 (c) - STRG.57246.1	-3.99	0.005	Nuclear Receptor Subfamily 4 Group A Member 3	Chondrosarcoma, Extraskeletal Myxoid and Chondrosarcoma	May act as a transcriptional activator

RNF41-204 (=) - STRG.72097.13	-3.87	0.001	Ring Finger Protein 41	Conversion Disorder and Prader-Willi Syndrome	Encodes for E3 ubiquitin ligase that plays a role in type 1 cytokine receptor signaling
FGF16-201 (=) - STRG.112642.1	3.83	0.007	Fibroblast Growth Factor 16	Metacarpal 4-5 Fusion and Syndactyly, Type Iii	Proper heart development
TNNI3-207 (=) - STRG.103424.2	3.67	0.005	Troponin I3, Cardiac Type	Cardiomyopathy, Dilated, 2A and Cardiomyopathy, Familial Hypertrophic, 7	Inhibitory subunit blocking actin-myosin interactions and thereby mediating striated muscle relaxation
TNNI3-202 (=) - STRG.103424.1	3.63	0.005	Troponin I3, Cardiac Type	Cardiomyopathy, Dilated, 2A and Cardiomyopathy, Familial Hypertrophic, 7	Inhibitory subunit blocking actin-myosin interactions and thereby mediating striated muscle relaxation
MCUR1P1-201 (=) -	3.43	0.005	MCUR1 Pseudogene 1	N/A	Pseudogene

STRG.91073.1					
AP001033.2-201 (=) - STRG.96560.1	3.43	0.005	Long non-coding RNA	N/A	Long non-coding RNA
CHD1-212 (i) - STRG.34005.1	-3.26	0.004	Chromodomain Helicase DNA Binding Protein 1	Pilarowski-Bjornsson Syndrome and Schizophrenia 8	Alters gene expression possibly by modification of chromatin structure
HES3-201 (=) - STRG.307.1	3.18	0.004	Hes Family BHLH Transcription Factor 3	Chromosome 1P36 Deletion Syndrome	Transcriptional repressor of genes that require a bHLH protein for their transcription
MRPL42-210 (=) - STRG.73269.8	-2.93	0.001	Mitochondrial Ribosomal Protein L42	Somatization Disorder	Encodes a protein identified as belonging to both the 28S and the 39S subunits of the mitochondrial ribosome
STRG.217511.25 (u)	2.83	0.006	New intergenic isoform	N/A	N/A
PITPNM2-201 (c) -	-2.69	0.009	Phosphatidylinositol Transfer	Retinal Degeneration	Catalyzes the transfer of phosphatidylinositol and phosphatidylcholine

STRG.74884.1			Protein Membrane Associated 2		between membranes (in vitro). Binds calcium ions.
OSMR-201 (n) - STRG.32126.5	-2.64	0.009	Oncostatin M Receptor	Amyloidosis, Primary Localized Cutaneous, 1 and Primary Cutaneous Amyloidosis	Associates with IL31RA to form the IL31 receptor. Binds IL31 to activate STAT3 and possibly STAT1 and STAT5. Capable of transducing OSM-specific signaling events
PRICKLE2-DT-201 (x) - STRG.21751.7	2.44	0.010	Prickle Planar Cell Polarity Protein 2	Sensory Ataxic Neuropathy, Dysarthria, And Ophthalmoparesis and Ataxia Neuropathy Spectrum	Encode for a homolog of Drosophila prickle
AC016542.2-201 (=) - STRG.61654.1	-2.44	0.009	To be Experimentally Confirmed	N/A	N/A
MYO1B-213 (=) - STRG.16666.6	-2.43	0.010	Myosin IB	Colorectal Cancer and Deafness, Autosomal	Motor protein that may participate in process critical to neuronal development and function

				Dominant 48	such as cell migration, neurite outgrowth and vesicular transport
AC004491.1-201 (=) - STRG.46144.21	-2.39	0.005	To be Experimentally Confirmed	N/A	N/A
KLHL3-209 (j) - STRG.35076.10	-2.39	0.005	Kelch Like Family Member 3	Pseudohypoaldosteronism, Type Iid and Pseudohypoaldosteronism, Type Iie	Regulator of ion transport in the distal nephron
AC069366.1-201 (=) - STRG.92410.1	-2.39	0.005	Pseudogene	N/A	N/A
NFATC4-201 (=) - STRG.78293.3	2.34	0.007	Nuclear Factor Of Activated T Cells 4	Leukostasis and Trichothiodystrophy 6, Nonphotosensitive	Encoded for a protein that is part of a DNA-binding transcription complex
FGD1-201 (n) - STRG.111935.8	-2.30	0.002	FYVE, RhoGEF And PH Domain	Aarskog-Scott Syndrome and Scott Syndrome	Regulates the actin cytoskeleton and cell shape

			Containing 1		
ZFR2-204 (=) - STRG.99186.1	2.28	0.007	Zinc Finger RNA Binding Protein 2	Malignant Essential Hypertension	Zinc Finger RNA Binding Protein
ASAP2-201 (n) - STRG.10409.8	-2.22	0.002	ArfGAP With SH3 Domain, Ankyrin Repeat And PH Domain 2	Bulbar Polio	Activates the small GTPases ARF1, ARF5 and ARF6. Regulates the formation of post-Golgi vesicles and modulates constitutive secretion. Modulates phagocytosis mediated by Fc gamma receptor and ARF6. Modulates PXN recruitment to focal contacts and cell migration
SAR1B-206 (=) - STRG.34986.7	-2.19	0.006	Secretion Associated Ras Related GTPase 1B	Chylomicron Retention Disease and Hypobetaliproteinemia, Familial, 1	Involved in transport from the endoplasmic reticulum to the Golgi apparatus.
Y_RNA.110-201 (=) - STRG.65888.19	-2.18	0.003	miscellaneous RNA	N/A	Miscellaneous small RNA

ZFR2-201 (j) - STRG.99186.2	2.13	0.007	Zinc Finger RNA Binding Protein 2	Malignant Essential Hypertension	Zinc Finger RNA Binding Protein
FNBP4-205 (c) - STRG.65888.2	-2.11	0.004	Formin Binding Protein 4	Microphthalmia With Limb Anomalies and Cerebral Amyloid Angiopathy, Itm2b-Related, 2	Regulation of cytoskeletal dynamics during cell division and migration and maintenance of membrane curvature at sites of nascent vesicle formation
PPM1N-206 (=) - STRG.102456.3	2.07	0.004	Protein Phosphatase, Mg2+/Mn2+ Dependent 1N (Putative)	N/A	Protein Phosphatase
KRT7-201 (=) - STRG.71856.1	2.06	0.002	Keratin 7	Pseudomyxoma Peritonei and Signet Ring Cell Adenocarcinoma	Blocks interferon-dependent interphase and stimulates DNA synthesis in cells. Involved in the translational regulation of the human papillomavirus type 16 E7 mRNA (HPV16 E7)

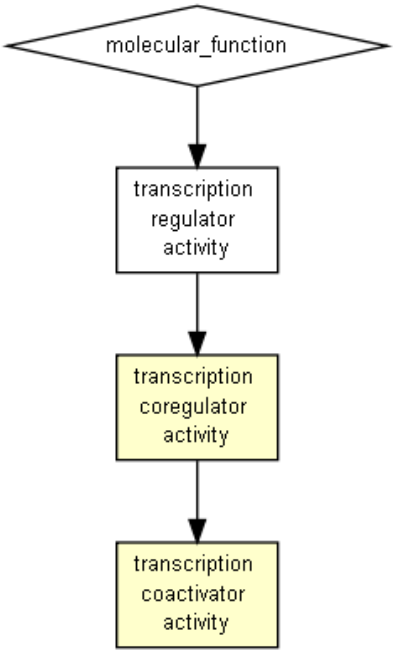
CSNK1G1-209 (o) - STRG.84144.5	-2.01	0.002	Casein Kinase 1 Gamma 1	Aortic Valve Prolapse and Gm1-Gangliosidosis, Type I	Cell cycle checkpoint arrest in response to stalled replication forks by phosphorylating Claspin
AL391005.1-201 (=) - STRG.64011.1	1.96	0.009	To be Experimentally Confirmed	N/A	N/A
MTURN-201 (j) - STRG.44254.3	-1.96	0.007	Maturin, Neural Progenitor Differentiation Regulator Homolog	Polycystic Kidney Disease 2 With Or Without Polycystic Liver Disease	Promotes megakaryocyte differentiation and represses NF-kappa-B transcriptional activity
LRRC37B-204 (j) - STRG.92784.12	-1.92	0.006	Leucine Rich Repeat Containing 37B	Chromosome 17Q11.2 Deletion Syndrome and Chromosome 15Q26-Qter Deletion Syndrome	Protein Coding gene
POLD1-212 (i) -	1.92	0.005	DNA Polymeras	Mandibular Hypoplasia	Role in DNA replication and repair

STRG.103029.2			e Delta 1, Catalytic Subunit	, Deafness, Progeroid Features, And Lipodystrophy Syndrome and Colorectal Cancer 10	
ARSI-201 (=) - STRG.35725.1	1.89	0.006	Arylsulfatase Family Member I	Autosomal Recessive Spastic Paraplegia Type 66 and Gastric Dilatation	Protein encoded by this gene is thought to be secreted, and to function in extracellular space
LINC02334-202 (j) - STRG.76096.1	-1.85	0.004	Long Intergenic Non-Protein Coding RNA 2334	N/A	Long non-coding RNA
IFITM1-202 (=) - STRG.64176.1	1.80	0.005	Interferon Induced Transmembrane Protein 1	Influenza and West Nile Virus	Restricts cellular entry by diverse viral pathogens, such as influenza A virus, Ebola virus and Sars-CoV-2

Supplementary Table 1b: GO enrichment analysis results.

Most differentially expressed transcripts uncovered by long read single molecule cDNA sequencing (Oxford Nanopore) revealed an enrichment of the transcription coactivator activity

function with 8 genes involved in this function found differentially expressed (genes coding for transcripts with batch-corrected P-value < 0.01 and |log2(FoldChange)|>0.5 analyzed with GOrilla from Gene Ontology (Eden et al., 2009).



GO term	Description	P-value	FDR q-value	Enrichment (N, B, n, b)	Genes
GO:0003713	transcription coactivator activity	6.93E-5	6.62E-2	3.87 (348,8,90,8)	HCFC1 SMARCC1 BCL9L SMARCA4 MED23 DDX17 NCOA1 MTA3
GO:0003712	transcription coregulator activity	8.55E-4	4.08E-1	2.97 (348,13,90,10)	HCFC1 SMARCC1 BCL9L

					SMARCA4 MED23 DDX17 NSD1 BTG1 NCOA1 MTA3

Acknowledgments:

We greatly thank Dr. Basma Benabdallah, iPSC Core at CHU Sainte-Justine Research Centre and Dr. Silvia Selleri (Paganelli Lab) from CHU Sainte-Justine Research Center.

This work was finded by the Foundation CHU Sainte-Justine and the Association Quebecoise de l'artresie de L'oesophage (AQAO) to Dr. Christophe Faure.

Z.O. is supported by an award from MITACS. N.A.D. is supported by FRQS (Fonds de recherche du Québec – Santé) Junior-2 award and by research grants from the Canadian Institutes of Health Research (PJT-156408, PJT-174993), Natural Sciences and Engineering Research Council of Canada (RGPIN-2018-05979), Canada Foundation for Innovation (37622), and the Quebec Cell, Tissue and Gene Therapy Network –ThéCell (a thematic network supported by the FRQS).

MAS is supported by a FRQS Junior 1 fellowship and establishment award (295760).

CHAPTER 4: Barrett's esophagus in Esophageal Atresia Patients

Manuscript in preparation (preliminary results)

Raad S., David A., Faure C.

Author affiliations

Suleen Raad, MSc

Esophageal Development and Engineering Laboratory

Sainte-Justine Research Centre

3175 Côte Sainte-Catherine, H3T 1C5, Canada

suleen.raad@umontreal.ca

Anu David, PhD

Esophageal Development and Engineering Laboratory

Centre Hospitalier Universitaire Sainte-Justine Research Centre

3175 Côte Sainte-Catherine, H3T 1C5, Canada

anu.david@umontreal.ca

Christophe Faure, MD

Esophageal Development and Engineering Laboratory, Sainte-Justine Research Centre,

Esophageal Atresia Clinic and Division of Pediatric Gastroenterology Hepatology and Nutrition, CHU Sainte Justine, 3715 Côte Sainte Catherine, Université de Montréal, H3T1C5, Montréal, QC, Canada.

christophe.faure@umontreal.ca

Corresponding author

Authors' contribution: AD, CF and SR designed the study. SR performed the experiments. SR and AD went through the data analysis. SR, AD and CF wrote the manuscript. AD and CF reviewed the manuscript.

Abstract

The derivation of induced pluripotent stem cells (iPSCs) has ushered a new era to study human organ development, disease modelling, and drug discovery. The esophagus has been the least studied section of the digestive tract. A common acquired esophagus condition is Barrett's esophagus (BE), a precursor to esophageal adenocarcinoma. In patients with BE, esophageal squamous mucosa is damaged by gastroesophageal reflux disease (GERD) and is replaced by metaplastic columnar mucosa. Clinically, esophageal atresia patients present a higher risk for GERD, and developing BE compared to the general population. We have previously succeeded in generating esophagus organoids from iPSCs derived from EA/TEF patients. Here, we suggest, for the first time, that EA/TEF esophagus organoids treated with 100ng/ml BMP4 expressed high levels of BE columnar epithelial markers *FOXA2* and *MUC5AC* compared to healthy organoids. This model provides a platform to have extensive cellular, molecular, and clinical data to understand the susceptibility of EA/TEF patients to developing BE and potentially becoming a pre-clinical tool for drug testing.

Introduction

Barrett's esophagus (BE) is a potential serious complication of gastroesophageal reflux disease (GERD) causing chronic injury to the esophagus epithelium (Spechler and Souza 2014). The prevalence of BE ranges from 1 to 5%, which progresses into esophageal adenocarcinoma in 0.1 to 0.3% cases every year (Bhat et al. 2011; Hvid-Jensen et al. 2011). In BE, the squamous epithelium in the distal part of the esophagus is replaced with columnar epithelium with gastric and

intestinal features (Spechler and Souza 2014). The underlying cellular origin and molecular mechanisms that cause the replacement of squamous epithelium with metaplastic columnar mucosa remain poorly understood. It has been previously demonstrated that the BMP signaling is essential for the morphogenesis of the esophagus. Specifically, BMP4 is generated in the surrounding mesenchyme during early foregut development. Conditional knockout of Bmp4 causes failure of foregut separation with a single tube connecting the pharynx to the stomach (Li et al. 2008). At later stages of esophagus development, Noggin, a BMP antagonist, regulates the transition from simple columnar to stratified squamous epithelium in the developing esophagus (Rodriguez et al. 2010). However, abnormal BMP signaling has been directly associated with Barrett's esophagus progression with a direct role in the metaplastic transformation of normal squamous cells into columnar epithelial cells (Que et al. 2019)

Esophageal atresia tracheoesophageal fistula (EA/TEF) is a relatively common congenital anomaly of the upper gastrointestinal tract affecting 1 in 3,000 births (van Lennep et al. 2019). Long-term complications and outcomes of EA/TEF include respiratory and digestive problems such as tracheomalacia, eosinophilic esophagitis, and GERD among others (Kovesi and Rubin 2004). This raises concerns about the increased risk of developing BE and esophageal adenocarcinoma. The prevalence of BE in EA/TEF patients is relatively high (Tullie et al. 2021) with an incidence in EA/TEF patients as young as 2 years old (Hsieh et al. 2017). The exact mechanisms that make EA/TEF patients are more prone to BE complications and risks remain poorly understood.

In this chapter, we used iPSC derived-esophageal organoids **(discussed in chapter 3)** to study EA/TEF patients' susceptibility to BE. We hypothesize that esophageal organoids derived from EA/TEF patients are more susceptible to developing BE. We treated mature esophageal organoids with human recombinant BMP4 for one-month and compared their gene expression profiles with healthy esophageal organoids. Our preliminary results show that EA/TEF derived-esophageal organoids express columnar marker, MUC5AC, an indication of metaplastic changes occurring in stratified squamous esophageal epithelium. MUC5AC expression was observed in healthy organoids.

Materials and Methods

Human embryonic stem cell and induced pluripotent stem cells

Human embryonic stem cell (ESC) line, H9 was a kind gift from the Andelfinger lab at CHU Sainte-Justine research center (Wunnemann et al. 2020). Healthy iPSC cell line (GHC4) and EA/TEF patient-derived iPSCs (EA1, EA2, and EA3) were generated and obtained from the IPSC Core Facility at CHU-Sainte Justine research center.

Culture and expansion of ESCs and iPSCs

Both ESCs and iPSCs were cultured on feeder free and non-xenogeneic conditions. Cells were plated on human vitronectin VTN XF™ (STEMCELL Technologies, Canada. Catalog#100-0763) coated 100mm cell culture dishes. Cells were maintained at 37°C with 5% CO_2 with daily replacement of Essential 8 media system (Thermofisher, Canada. Catalog#A1517001). Cells were passaged as aggregates every 3-4 days until they reach 60-70% confluency with 0.5mM EDTA diluted in PBS (Thermofisher, Canada. Catalog#15575020).

Differentiation protocol: Preparing cells for differentiation (Day -1 and Day 0)

Two days prior to differentiation (D-1), cells were dissociated into single cells using Accutase™ (Stem Cell technologies, Canada. Catalog# 07922) and transferred onto BIOLAMININ 521 LN (LN521, BioLamina, Sweden. Product# LN521-02) coated plates with E8 media supplemented with 10uM Rock inhibitor Y-27632 (Sigma-Aldrich, US. Product# SCM075). The following day (D0) (12-16 hours later), media was changed with E8 only. If survival rate of the cells were less than 50%, they were cultured for an additional 24 hours before starting the differentiation. ESC and iPSCs were maintained at 37°C with 5% CO_2 throughout the differentiation process.

Endoderm differentiation (Day 1 to Day 3)

We modified and adapted previously published protocol for endoderm differentiation (Matsuno et al. 2016). Xeno-free media (XFM-) was prepared using 500 mL of RPMI 1640 medium without

L-Glutamine (Thermo Fisher Scientific, Canada. Catalog# 11875101), 10mL B-27 Supplement minus insulin (Thermo Fisher Scientific, Canada. Catalog# A1895601), 5mL GlutaMax™ (Thermo Fisher Scientific, Canada. Catlog# 35050061), 5mL of KnockOut™ serum replacement (Thermo Fisher Scientific, Canada. Catalog# 10828010), 5mL Penicillin-Streptomycin (10,000 U/mL) (Thermo Fisher Scientific, Canada. Catalog# 15140148), 7.5mL HEPES (1M) Buffer (Thermo Fisher Scientific, Canada. Catalog# 15630130), 5mL of MEM Non-essential amino acids (100X) ((Thermo Fisher Scientific, Canada. Catalog# 11140050). Day 1 of differentiation cells are first washed with XFM- media to remove any residual E8 media, then cultured in XFM- with 100ng/mL Activin A (R&D systems, US. Catalog# 338-AC-010/CF) and 3uM CHIR99021 (Stem Cell technologies, Canada. Catlog# 72052). On Days 2 and 3 of culture, cells were first washed with XFM- media and the culture was continued with XFM- supplemented with 100ng/mL Activin A and 250nM of LDN193189 (Stemgent, US. Code# 04-0074). By day 3, a 70-80% confluent monolayer of endodermal cells should be observed under the microscope.

Anterior foregut differentiation (Day 4 and Day 5)

Endodermal cells were first washed with XFM- media and then cultured for 24 hours in XFM- supplemented with 1uM A8301 (Stemgent, US. Code# 04-0014) and 250nM of LDN193189. The following day cells were washed XFM- media and cultured in XFM- supplemented with 1uM A8301 and 1uM IWP2 (Stemgent, US. Code# 04-0034).

Esophagus differentiation (Day 6 to Day 24)

From day 6 to day 16 we switch to XFM+ containing the same components as XFM- and replacing b27 supplement minus insulin with b27 supplement with insulin (Thermofisher, Canada. Catalog# 17504044). To induce esophageal fate, we modified and adapted published protocol (Zhang et al. 2018; Trisno et al. 2018). Anterior foregut cells were cultured in XFM+ supplemented with 1uM A8301 and 250nM LDN193189 from day 6 until day 16, changing media daily. By day 16, we observed that cells had reached 100% confluency and the presence of dense cell clusters. On day 16 esophageal progenitor cells are cultured in XFM+ only until day 24.

Esophageal organoid formation (2 months)

Organoids were generated in suspension using Nunclon™ Sphera™ low attachment 96-well plates (Thermofisher, Canada. Catalog# 174930) by modifying previously published esophageal studies (Giroux et al. 2017; DeWard, Cramer, and Lagasse 2014). On day 24 of esophageal differentiation, cells were detached using TrypLE (Thermo Fisher Scientific, Canada. Catalog# 12604013) and gently resuspended in XFM+ media. Viable cells were counted using Trypan blue solution (Thermofisher, Canada. Catalog# 15250061) and 50,000 cells are then aliquoted in each well of the 96-well plate containing XFM+ supplemented with 1uM A8301, 250nM LDN, 3uM CHIR99021, 20ng/mL FGF2/bFGF (Peprotech, US. Catalog# AF100-18B) and 200ng/mL EGF (Thermofisher, Canada. Catalog# PHG0313).

BMP4 induction in Esophageal Organoids (1 month)

To evaluate the effect of BMP4 signaling, 2 months old mature esophageal organoids from both healthy and EA/TEF groups were incubated in XFM+ media supplemented with either 25ng/mL or 100ng/mL BMP4 recombinant protein (Peprotech US. Catalog #120-05ET) for a month.

RT-qPCR

At each developmental stage (definitive endoderm, anterior foregut, esophageal progenitors, mature esophageal epithelium/organoids, and other organ-lineages), cells were detached using Accutase™ and RNA was extracted using the Promega kit ReliaPrep™ RNA Cell Miniprep System (Promega, US. Catalog# Z6011). RNA was reverse transcribed using Omniscipt™ RT kit (Qiagen, US. Catalog# 205113) and the complementary DNA (cDNA) obtained was used for real time quantitative PCR using LightCycler instrument (Roche Life Science, Germany). cDNA was quantified using TaqMan Gene expression assays and the TaqMan primers to target genes were purchased from Thermo Fisher Scientific listed in Table 2. The transcript level of each gene was normalized to GAPDH housekeeping gene using the 2^-DDCT method. The relative gene expression was calculated and reported as fold change compared to indicated samples using GAPDH normalized transcript level. The results include a mean of at least 3 technical replicates of

each biological sample. All data quantification is presented as the mean ± SEM using GraphPad Software Prism 6.

Results and discussion

Induction of metaplasia to create BE from healthy and EA patient esophageal organoids

In this study, we examined the acquired esophagus disease BE in our generated esophagus organoid model. We used the same esophageal differentiation protocol discussed in **chapter 3,** to direct the differentiation of PSCs (embryonic stem cells (H9) and iPSCs derived from a healthy subject) and 3 EA/TEF iPSCs into esophageal organoids.

To investigate the effect of BMP signaling on our system, mature esophagus organoids (after 2 months of culture without passaging) from both groups were treated with BMP4 recombinant protein for 30 days at 2 different concentrations: 25ng/mL and 100ng/mL. No morphological changes were observed in BMP4 treated esophagus organoids in both groups **(Fig. 1).**

The genomic signature of BE shares similarities with that of normal intestinal tissue. This includes transcription factors such as *FOXA2*. *FOXA2* is normally expressed in columnar epithelium of embryonic development and the adult intestine and is considered a BE progression marker (Wang et al. 2014). mRNA expression levels of *FOXA2* in all 3 patient derived esophagus organoids were significantly higher in healthy organoids after 1 month exposure to 100ng/mL of BMP4 **(Fig. 2A).** Untreated esophagus organoids from both healthy and patient PSCs did not express FOXA2 (**supplementary figure S1**).

We then looked into the expression of *MUC5AC*, columnar epithelial marker, known to be strongly expressed in BE epithelium (Arul et al. 2000). Healthy derived organoids expressed *MUC5AC* at very low levels after BMP4 treatment at 25ng/mL or at 100ng/mL whereas EA/TEF-derived organoids revealed high expression of *MUC5AC* after one month exposure to BMP4 but only at 100ng/mL **(Fig 2B, 2C).** Untreated esophagus organoids derived from both groups did not express *FOXA2* or *MUC5AC*. We did not detect the expression of *MUC2*, intestinal epithelial mucin marker **(data not shown),** but suspect that longer exposure at higher concentrations of BMP4 (at 200ng/mL) might induce its expression in esophagus organoids.

Conclusion

Over the last decade, BE has dramatically risen in incidence (Runge, Abrams, and Shaheen 2015) and the mechanism by which BE develops remains unknown. Moreover, concerns exist with EA/TEF patients who have a high prevalence of developing BE compared to the normal population. Our preliminary data reveal that EA/TEF derived esophagus organoids are more susceptible to developing BE upon overexpression of BMP4 compared to healthy organoids. After a 1-month treatment of healthy and EA/TEF derived organoids with 100ng/mL of BMP4, EA/TEF derived esophagus organoids expressed FOXA2 and MUC5AC significantly higher than healthy derived esophagus organoids. Our *in vitro* esophageal organoid system is a good tool to investigate the crosstalk and signaling pathways such as BMP, Hedgehog and WNT that have been suggested to play a key role in BE progression. This work mainly focused on the epithelial transformation as our organoids lack submucosal glands, gastroesophageal junction epithelium, and residual embryonic cells. However, our system can be used as a patient-specific model to understand the cellular and molecular mechanisms affected in EA/TEF patients which could make these patients more susceptible to BE progression.

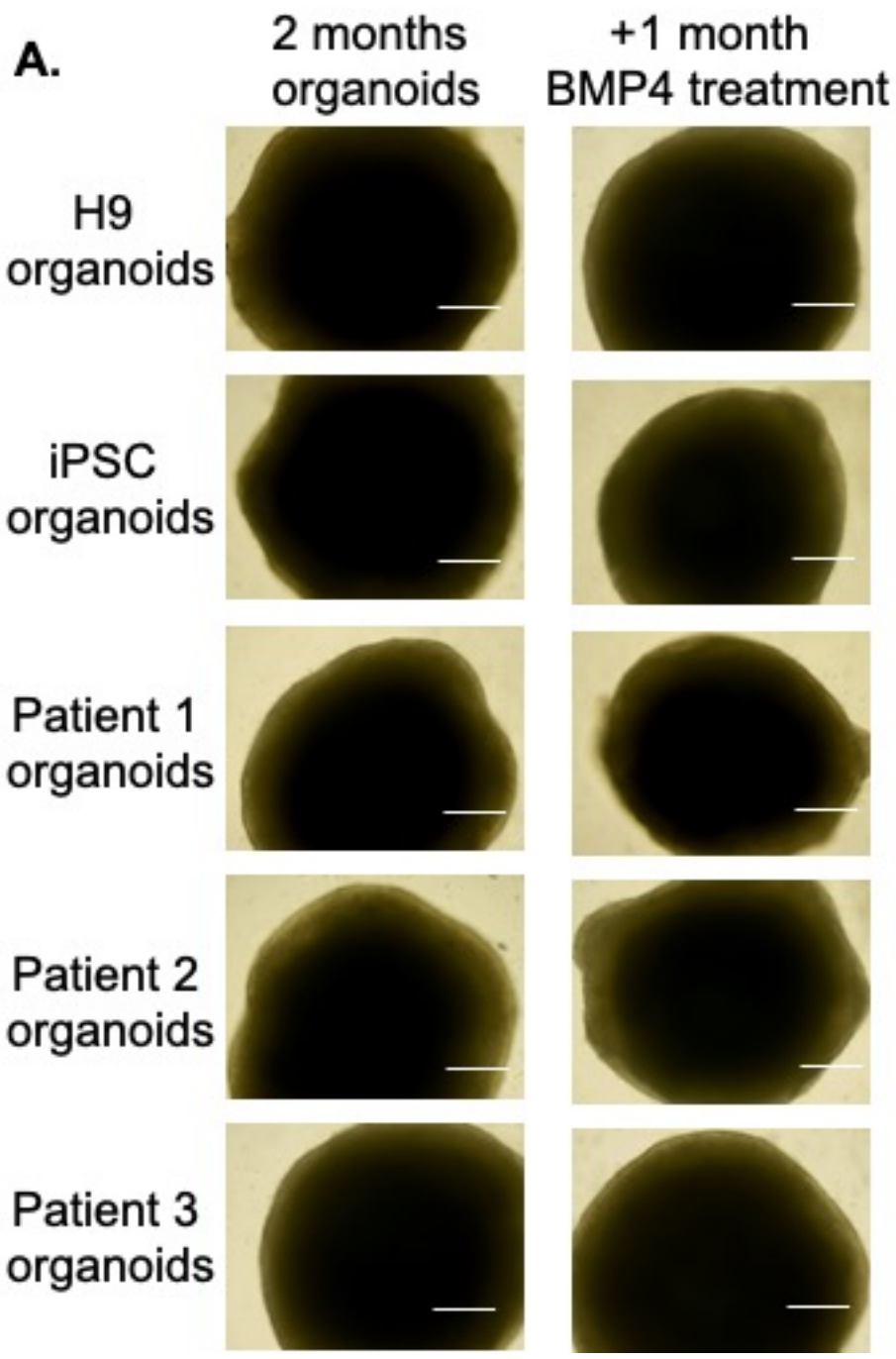


Figure 1: PSC derived esophagus organoids after 1-month BMP4 treatment. Representative bright field images of esophagus organoids from healthy and patients PSCs. No morphological differences were detected after 1-month BMP4 treatment. Scale bar 50um

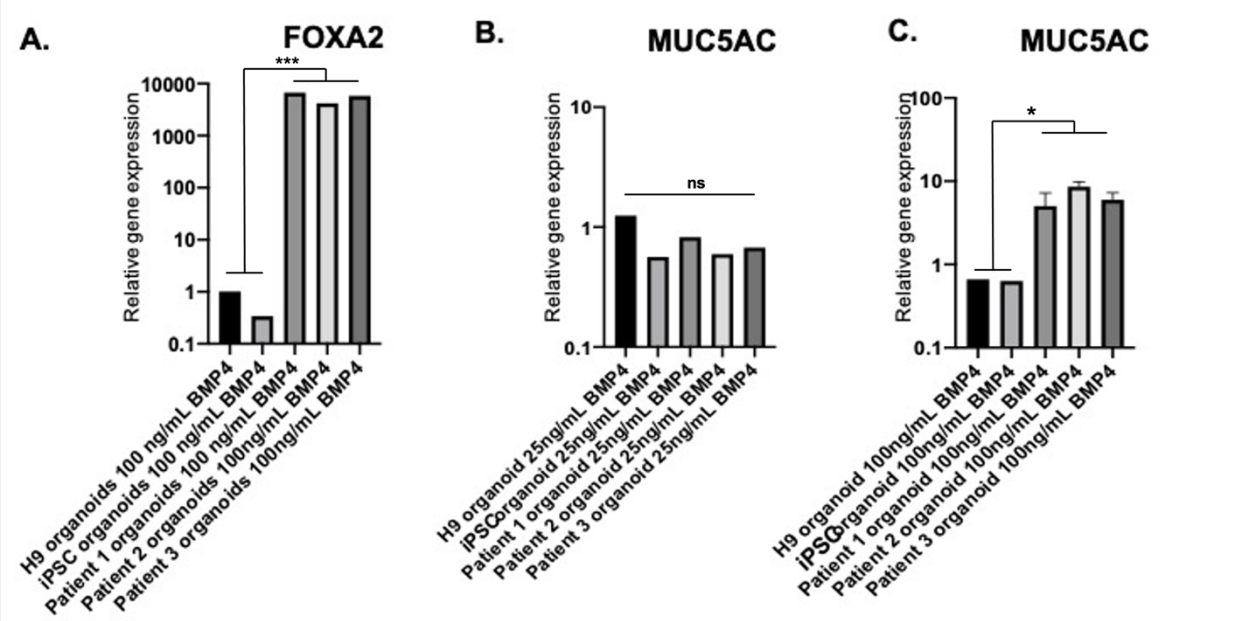
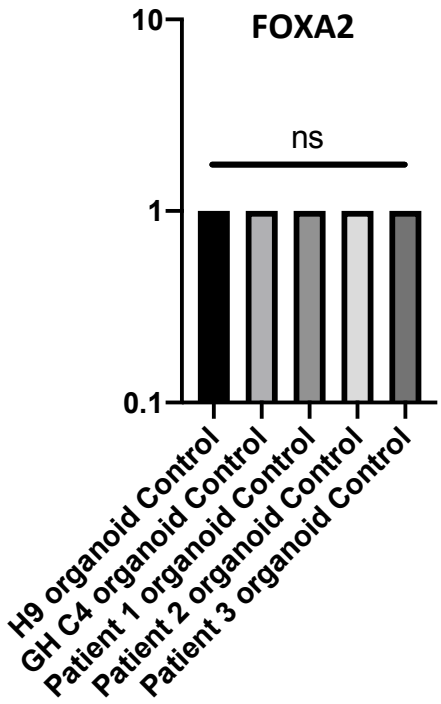


Figure 2: Gene expression of key BE markers in healthy and EA/TEF derived esophagus organoids after 1 month treatment with BMP4. A, B and C) mRNA expression levels of FOXA2 transcription factor and MUC5AC by qPCR. The fold change (y axis) was generated by normalizing the transcript levels to those of the healthy esophagus organoids (H9). Data representative mean ± SEM (n=3 technical replicates for each biological cell line) **p<0.01, *** p<0.001 by unpaired two tailed Student's t test.



Supplementary Figure S1: Gene expression of FOXA2 in untreated control healthy and EA/TEF derived esophagus organoids. Absent expression of FOXA2 transcription factor by qPCR. The fold change (y axis) was generated by normalizing the transcript levels to those of the healthy esophagus organoids (H9). Data representative mean ± SEM (n=3 technical replicates for each biological cell line).

CHAPTER 5: Conclusions and Future Directions

Discussion

During embryonic development, a single primordial anterior foregut tube separates into two structurally and functionally distinct tubular organs: the esophagus, that allows the passage of food to the stomach, and the trachea, which branches into the lungs to allow air passage. Efforts in understanding this rapid and complex division of the foregut tube into the esophagus and the trachea remain to this day. Esophageal atresia and tracheoesophageal fistula (EA/TEF) are common abnormalities arising from a disruption of the proper separation of the foregut (1 in 3500 births). Several key genes and signaling pathways involved in this segregation have been described, mostly from animal models. These include *SOX2* and *NKX2.1* transcription factors and BMP, TGF-ß, WNT and HH signaling pathways which are critically involved in the specification, regulation, and development of the esophagus. However, many questions as to how cellular behavior accompanied by these signaling queues orchestrate intact foregut separation, remain unanswered. In addition, mouse models exhibit a limitation as they are structurally different from the human esophagus. This led us to think of induced pluripotent stem cells. Since their discovery, iPSCs have been used to model the development of a variety of human tissues and organs. Also, iPSCs through reprogramming of patient cells are being used to model human diseases. A fundamental understanding of the development of the esophagus helps recognize how and why congenital malformation and esophageal diseases may occur. Here, we successfully generated the first reported iPSCs from the PBMCs of 3 type C EA/TEF pediatric patients to model not only an esophageal congenital disease such as EA/TEF but also an acquired disease such as Barrett's esophagus. We developed a protocol in which we differentiate iPSCs into esophagus organoids by tightly regulating developmental pathways such as Nodal, Wnt, BMP, and TGF-ß, to drive the differentiation first into definitive endoderm, then the anterior foregut to finally induce esophageal commitment and maturation into esophagus organoids. Using EA/TEF patient iPSCs we generated matrix- and xenogeneic-free 3-dimensional mature stratified squamous esophageal epithelial organoids and compared them to esophageal organoids derived from healthy individuals. The generated 2 months old esophageal stratified squamous epithelial organoids expressed key mature markers such as KRT4, KRT13, INV and p63 in both the groups. However, at the crucial stage of anterior foregut differentiation, we observed a significant downregulation of transcription factor,

SOX2, a dorsal marker for esophageal fate, in EA/TEF patient-derived cells. Though the patient-derived cells achieved maturity similar to healthy cells, we observed an abnormal expression of transcription factor *NKX2.1*, a ventral marker for trachea in esophageal organoids. Both factors, *SOX2* and *NKX2.1* are known to play key roles during anterior foregut patterning and subsequent separation. Using nanopore RNA sequencing of anterior foregut stage in healthy and patient cells, we were able to identify 173 differentially expressed transcripts. These include *GSTM1*, an enzyme that plays a role in detoxification, and was linked to EA/TEF. Another one is *RAB37* from the Rab family of proteins and have also been linked to anterior foregut malformations. This unique molecular signature was specific for patient-derived cells directed toward the esophageal fate only. When we directed the differentiation towards other cellular fates such as the trachea, the hepatoblast and myoblasts, we did not observe any significant differences in their differentiation potential in both health and EA/TEF patient-derived cells. The strategic approach of using EA/TEF derived iPSCs, allowed us to mimic the initial developmental stages of the human esophagus to understand the origins of this malformation. We can conclude that the intrinsic defect observed in these cells are limited only to the esophagus. Our work is limited to isolated type C EA/TEF and thus cannot relate our results to other types of EA or syndromic EA with associated malformations. More studies are required to understand the failure of key mechanisms and pathways involved during the critical interface of anterior foregut specification.

We also used our iPSC-derived esophageal organoids to model Barrett's esophagus, an acquired esophageal disease. It is known that EA/TEF patients are more prone to BE. In fact, the prevalence of BE in adults with EA/TEF is four times higher than the general population and is present at a younger median age (43 vs 60 years). Several pathways, such as the BMP pathway, involved in esophagus development are also linked to BE development. Previous work demonstrated that EA patients have an intrinsic cellular and molecular susceptibility to BE (Ten Kate et al. 2022). Therefore, using our EA/TEF patient-derived organoids it would be possible to identify genetic risk factors in EA/TEF patients as well as decipher the underlying mechanisms associated with BE. Our preliminary results showed that EA/TEF patient-derived organoids express early BE columnar epithelium markers FOXA2 and MUC5AC, compared to the healthy when exposed to high levels of BMP4. However, more studies are required to understand the mechanisms regulating this higher susceptibility of EA/TEF patients to develop BE.

Our work therefore highlights the advantage of using patient-derived iPSCs to model congenital diseases which allows to gain new insights on organ development during embryogenesis.

Despite the exciting findings in our work, we still lack the exact understanding and mechanistic causing the dysregulation in gene and protein expression in patient-derived cells. The next steps would be to help define the regulatory network and potential downstream effectors causing this unique molecular signature in patient-derived cells.

Future directions

This study just touches the tip of an iceberg and therefore more in-depth studies are required to unravel the mechanisms involved in the dysregulation of key factors during embryonic esophageal development in EA/TEF patients.

To understand SOX2 downregulation, we sought the role of a long non- coding RNA, SOX2OT and we observed that SOX2OT was downregulated in EA/TEF patient-derived cells. SOX2OT is a long non-coding RNA, that harbors SOX2 gene in its intronic region and is expressed in embryonic stem cells which is downregulated during cellular differentiation. SOX2OT is involved in the regulation of transcription factor, SOX2 (Shahryari et al.2014) and we hypothesize that SOX2OT may play a crucial role during foregut/esophagus differentiation. This could be determined by performing knockout studies of SOXOT expression in healthy cells and investigate dysregulation of SOX2 expression at the anterior foregut stage. One can also use editing tools like CRISPR/Cas9 and consider editing SOX2OT promoter *in vitro* in healthy iPSCs and relate its effect on the SOX2 expression during directed esophagus differentiation.

It would also be important to use CRIPR/Cas9 gene editing to correct identified dysregulated gene transcripts such as *GSTM1* and *RAB37* in the patient iPSCs, to determine if it is sufficient to cause a defective phenotype of generated esophageal organoids. This will help further improve our understanding on the underlying mechanisms and signaling networks regulating foregut separation. These transcripts can also be studied in animal models to validate the *in vitro* findings and thus better assess their involvement alone or with other genes causing disruption in cellular behavior during esophageal development.

Another limitation in our model was the lack of the different cell types that make up the esophagus. Cells originating from all 3 germ layers interact throughout organogenesis during embryonic

development of individual organs. Organoids of multicellular origins can recapitulate accurately the aspects of human physiology by controlling *in vitro* environments during cell differentiation. This approach has been successfully used to create complex liver organoids by incorporating mesenchymal cells, enteric nervous cells, and vessels (Takebe et al. 2013). Eicher et al. (2021) generated complex gastric organoids by combining gastric epithelial mesenchymal cells with neural crest cells to model gastrointestinal tract (Eicher et al. 2021). This approach could be used provide insights in the role mesenchymal and neural cells during esophageal development. Despite these limitations, our patient-derived organoid system can be used as an ideal model for high throughput screening of drugs or therapeutic options to have a personalized and precise treatment approach. This can further improve esophageal functionality such as dysmotility, gastroesophageal reflux disease, and recurring stenosis.

Finally, multicellular esophageal organoids could be used to engineer an artificial tube to repair discontinued esophagus in EA/TEF patients. Using corrected patient-derived esophageal epithelial and mesenchymal cells in a biomimetic material can bridge the gap between the two ends of the esophagus, especially, in long gap EA/TEF patients. This bioengineering approach could provide a long-term alternative for gastric replacement and the associated postoperative problems.

REFERENCES

Adachi, K., H. Suemori, S. Y. Yasuda, N. Nakatsuji, and E. Kawase. 2010. 'Role of SOX2 in maintaining pluripotency of human embryonic stem cells', *Genes Cells*, 15: 455-70.

Adzick, N. S., J. H. Fisher, H. S. Winter, R. H. Sandler, and W. H. Hendren. 1989. 'Esophageal adenocarcinoma 20 years after esophageal atresia repair', *J Pediatr Surg*, 24: 741-4.

Al-Shraim, M. M., R. A. Eid, A. O. Musalam, K. Radad, A. H. Ibrahim, and T. A. Malki. 2015. 'Ultrastructural Changes of the Smooth Muscle in Esophageal Atresia', *Ultrastruct Pathol*, 39: 413-8.

Ali, I., P. Rafiee, Y. Zheng, C. Johnson, B. Banerjee, G. Haasler, H. Jacob, and R. Shaker. 2009. 'Intramucosal distribution of WNT signaling components in human esophagus', *J Clin Gastroenterol*, 43: 327-37.

Anastas, J. N., and R. T. Moon. 2013. 'WNT signalling pathways as therapeutic targets in cancer', *Nat Rev Cancer*, 13: 11-26.

Arul, G. S., M. Moorghen, N. Myerscough, D. A. Alderson, R. D. Spicer, and A. P. Corfield. 2000. 'Mucin gene expression in Barrett's oesophagus: an in situ hybridisation and immunohistochemical study', *Gut*, 47: 753-61.

Bailey, D. D., Y. Zhang, B. J. van Soldt, M. Jiang, S. Suresh, H. Nakagawa, A. K. Rustgi, S. S. Aceves, W. V. Cardoso, and J. Que. 2019. 'Use of hPSC-derived 3D organoids and mouse genetics to define the roles of YAP in the development of the esophagus', *Development*, 23:146 dev178855.

Barbera, M., M. di Pietro, E. Walker, C. Brierley, S. MacRae, B. D. Simons, P. H. Jones, J. Stingl, and R. C. Fitzgerald. 2015. 'The human squamous oesophagus has widespread capacity for clonal expansion from cells at diverse stages of differentiation', *Gut*, 64: 11-9.

Behrens, A., O. Pech, F. Graupe, A. May, D. Lorenz, and C. Ell. 2011. 'Barrett's adenocarcinoma of the esophagus: better outcomes through new methods of diagnosis and treatment', *Dtsch Arztebl Int*, 108: 313-9.

Berthet, S., E. Tenisch, M. C. Miron, N. Alami, J. Timmons, A. Aspirot, and C. Faure. 2015. 'Vascular Anomalies Associated with Esophageal Atresia and Tracheoesophageal Fistula', *J Pediatr*, 166: 1140-44.

Bhat, S., H. G. Coleman, F. Yousef, B. T. Johnston, D. T. McManus, A. T. Gavin, and L. J. Murray. 2011. 'Risk of malignant progression in Barrett's esophagus patients: results from a large population-based study', *J Natl Cancer Inst*, 103: 1049-57.

Billmyre, K. K., M. Hutson, and J. Klingensmith. 2015. 'One shall become two: Separation of the esophagus and trachea from the common foregut tube', *Dev Dyn*, 244: 277-88.

Brosens, E., M. Ploeg, Y. van Bever, A. E. Koopmans, I. Jsselstijn H, R. J. Rottier, R. Wijnen, D. Tibboel, and A. de Klein. 2014. 'Clinical and etiological heterogeneity in patients with tracheo-esophageal malformations and associated anomalies', *Eur J Med Genet*, 57: 440-52.

Burridge, P. W., and E. T. Zambidis. 2013. 'Highly efficient directed differentiation of human induced pluripotent stem cells into cardiomyocytes', *Methods Mol Biol*, 997: 149-61.

Cartabuke, R. H., R. Lopez, and P. N. Thota. 2016. 'Long-term esophageal and respiratory outcomes in children with esophageal atresia and tracheoesophageal fistula', *Gastroenterol Rep (Oxf)*, 4: 310-14.

Chamberlain, C. E., J. Jeong, C. Guo, B. L. Allen, and A. P. McMahon. 2008. 'Notochord-derived Shh concentrates in close association with the apically positioned basal body in

neural target cells and forms a dynamic gradient during neural patterning', *Development*, 135: 1097-106.
Chen, F., Y. Cao, J. Qian, F. Shao, K. Niederreither, and W. V. Cardoso. 2010. 'A retinoic acid-dependent network in the foregut controls formation of the mouse lung primordium', *J Clin Invest*, 120: 2040-8.
Cho, S. G., Y. H. Lee, H. S. Park, K. Ryoo, K. W. Kang, J. Park, S. J. Eom, M. J. Kim, T. S. Chang, S. Y. Choi, J. Shim, Y. Kim, M. S. Dong, M. J. Lee, S. G. Kim, H. Ichijo, and E. J. Choi. 2001. 'Glutathione S-transferase mu modulates the stress-activated signals by suppressing apoptosis signal-regulating kinase 1', *J Biol Chem*, 276: 12749-55.
Cho, S. H., I. Y. Kuo, P. F. Lu, H. T. Tzeng, W. W. Lai, W. C. Su, and Y. C. Wang. 2018. 'Rab37 mediates exocytosis of secreted frizzled-related protein 1 to inhibit Wnt signaling and thus suppress lung cancer stemness', *Cell Death Dis*, 9 (9): 868-81.
Clark, D. C. 1999. 'Esophageal atresia and tracheoesophageal fistula', *Am Fam Physician*, 59: 910-6, 19-20.
Correa, A., S. M. Gilboa, L. M. Besser, L. D. Botto, C. A. Moore, C. A. Hobbs, M. A. Cleves, T. J. Riehle-Colarusso, D. K. Waller, and E. A. Reece. 2008. 'Diabetes mellitus and birth defects', *Am J Obstet Gynecol*, 199(3): 237-2379.
Courbette, O., T. Omari, A. Aspirot, and C. Faure. 2020. 'Characterization of Esophageal Motility in Children With Operated Esophageal Atresia Using High-resolution Impedance Manometry and Pressure Flow Analysis', *J Pediatr Gastroenterol Nutr*, 71: 304-09.
Daniely, Y., G. Liao, D. Dixon, R. I. Linnoila, A. Lori, S. H. Randell, M. Oren, and A. M. Jetten. 2004. 'Critical role of p63 in the development of a normal esophageal and tracheobronchial epithelium', *Am J Physiol Cell Physiol*, 287: 171-81.
de Jong, E. M., J. F. Felix, A. de Klein, and D. Tibboel. 2010. 'Etiology of esophageal atresia and tracheoesophageal fistula: "mind the gap"', *Curr Gastroenterol Rep*, 12: 215-22.
Dedhia, P. H., N. Bertaux-Skeirik, Y. Zavros, and J. R. Spence. 2016. 'Organoid Models of Human Gastrointestinal Development and Disease', *Gastroenterology*, 150: 1098-112.
Deng, Y., Q. Su, J. Mo, X. Fu, Y. Zhang, and E. H. Lin. 2013. 'Celecoxib downregulates CD133 expression through inhibition of the Wnt signaling pathway in colon cancer cells', *Cancer Invest*, 31: 97-102.
DeWard, A. D., J. Cramer, and E. Lagasse. 2014. 'Cellular heterogeneity in the mouse esophagus implicates the presence of a nonquiescent epithelial stem cell population', *Cell Rep*, 9: 701-11.
DiStefano, T., H. Y. Chen, C. Panebianco, K. D. Kaya, M. J. Brooks, L. Gieser, N. Y. Morgan, T. Pohida, and A. Swaroop. 2018. 'Accelerated and Improved Differentiation of Retinal Organoids from Pluripotent Stem Cells in Rotating-Wall Vessel Bioreactors', *Stem Cell Reports*, 10: 300-13.
Domyan, E. T., E. Ferretti, K. Throckmorton, Y. Mishina, S. K. Nicolis, and X. Sun. 2011. 'Signaling through BMP receptors promotes respiratory identity in the foregut via repression of Sox2', *Development*, 138: 971-81.
Dravis, C., and M. Henkemeyer. 2011. 'Ephrin-B reverse signaling controls septation events at the embryonic midline through separate tyrosine phosphorylation-independent signaling avenues', *Dev Biol*, 355: 138-51.
Eden, E., R. Navon, I. Steinfeld, D. Lipson, and Z. Yakhini. 2009. 'GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists', *BMC Bioinformatics*, 10-48.

Edwards, N. A., V. Shacham-Silverberg, L. Weitz, P. S. Kingma, Y. Shen, J. M. Wells, W. K. Chung, and A. M. Zorn. 2021. 'Developmental basis of trachea-esophageal birth defects', *Dev Biol*, 477: 85-97.

Eicher, A. K., D. O. Kechele, N. Sundaram, H. M. Berns, H. M. Poling, L. E. Haines, J. G. Sanchez, K. Kishimoto, M. Krishnamurthy, L. Han, A. M. Zorn, M. A. Helmrath, and J. M. Wells. 2021. 'Functional human gastrointestinal organoids can be engineered from three primary germ layers derived separately from pluripotent stem cells', *Cell Stem Cell*. 29(1),36-51.

Eiraku, M., N. Takata, H. Ishibashi, M. Kawada, E. Sakakura, S. Okuda, K. Sekiguchi, T. Adachi, and Y. Sasai. 2011. 'Self-organizing optic-cup morphogenesis in three-dimensional culture', *Nature*, 472: 51-6.

Epperly, M. W., H. Guo, H. Shen, Y. Niu, X. Zhang, M. Jefferson, C. A. Sikora, and J. S. Greenberger. 2004. 'Bone marrow origin of cells with capacity for homing and differentiation to esophageal squamous epithelium', *Radiat Res*, 162: 233-40.

Epstein, J. A., H. Cosby, G. W. Falk, M. A. Khashab, R. Kiesslich, E. A. Montgomery, J. S. Wang, and M. I. Canto. 2017. 'Columnar islands in Barrett's esophagus: Do they impact Prague C&M criteria and dysplasia grade?', *J Gastroenterol Hepatol*, 32: 1598-603.

Falkenstein, K. N., and S. A. Vokes. 2014. 'Transcriptional regulation of graded Hedgehog signaling', *Semin Cell Dev Biol*, 33: 73-80.

Fausett, S. R., L. J. Brunet, and J. Klingensmith. 2014. 'BMP antagonism by Noggin is required in presumptive notochord cells for mammalian foregut morphogenesis', *Dev Biol*, 391: 111-24.

Fausett, S. R., and J. Klingensmith. 2012. 'Compartmentalization of the foregut tube: developmental origins of the trachea and esophagus', *Wiley Interdiscip Rev Dev Biol*, 1: 184-202.

Felix, J. F., M. F. van Dooren, M. Klaassens, W. C. Hop, C. P. Torfs, and D. Tibboel. 2008. 'Environmental factors in the etiology of esophageal atresia and congenital diaphragmatic hernia: results of a case-control study', *Birth Defects Res A Clin Mol Teratol*, 82: 98-105.

Filonzi, L., C. Magnani, G. L. de' Angelis, S. Dallaglio, and F. Nonnis Marzano. 2010. 'Evidence that polymorphic deletion of the glutathione S-transferase gene, GSTM1, is associated with esophageal atresia', *Birth Defects Res A Clin Mol Teratol*, 88: 743-7.

Frankish, A., M. Diekhans, A. M. Ferreira, R. Johnson, I. Jungreis, J. Loveland, J. M. Mudge, C. Sisu, J. Wright, J. Armstrong, I. Barnes, A. Berry, A. Bignell, S. Carbonell Sala, J. Chrast, F. Cunningham, T. Di Domenico, S. Donaldson, I. T. Fiddes, C. Garcia Giron, J. M. Gonzalez, T. Grego, M. Hardy, T. Hourlier, T. Hunt, O. G. Izuogu, J. Lagarde, F. J. Martin, L. Martinez, S. Mohanan, P. Muir, F. C. P. Navarro, A. Parker, B. Pei, F. Pozo, M. Ruffier, B. M. Schmitt, E. Stapleton, M. M. Suner, I. Sycheva, B. Uszczynska-Ratajczak, J. Xu, A. Yates, D. Zerbino, Y. Zhang, B. Aken, J. S. Choudhary, M. Gerstein, R. Guigo, T. J. P. Hubbard, M. Kellis, B. Paten, A. Reymond, M. L. Tress, and P. Flicek. 2019. 'GENCODE reference annotation for the human and mouse genomes', *Nucleic Acids Res*, 47: D766-D73.

Galarreta, C. I., F. Vaida, and L. M. Bird. 2020. 'Patterns of malformation associated with esophageal atresia/tracheoesophageal fistula: A retrospective single center study', *Am J Med Genet A*, 182: 1351-63.

Garman, K. S. 2017. 'Origin of Barrett's Epithelium: Esophageal Submucosal Glands', *Cell Mol Gastroenterol Hepatol*, 4: 153-56.

Giroux, V., A. A. Lento, M. Islam, J. R. Pitarresi, A. Kharbanda, K. E. Hamilton, K. A. Whelan, A. Long, B. Rhoades, Q. Tang, H. Nakagawa, C. J. Lengner, A. J. Bass, E. P. Wileyto, A. J. Klein-Szanto, T. C. Wang, and A. K. Rustgi. 2017. 'Long-lived keratin 15+ esophageal progenitor cells contribute to homeostasis and regeneration', *J Clin Invest*, 127: 2378-91.

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

Gordon, C. T., F. Petit, M. Oufadem, C. Decaestecker, A. S. Jourdain, J. Andrieux, V. Malan, J. L. Alessandri, G. Baujat, C. Baumann, O. Boute-Benejean, R. Caumes, B. Delobel, K. Dieterich, D. Gaillard, M. Gonzales, D. Lacombe, F. Escande, S. Manouvrier-Hanu, S. Marlin, M. Mathieu-Dramard, S. G. Mehta, I. Simonic, A. Munnich, M. Vekemans, N. Porchet, L. de Pontual, S. Sarnacki, T. Attie-Bitach, S. Lyonnet, M. Holder-Espinasse, and J. Amiel. 2012. 'EFTUD2 haploinsufficiency leads to syndromic oesophageal atresia', *J Med Genet*, 49: 737-46.

Goss, A. M., Y. Tian, T. Tsukiyama, E. D. Cohen, D. Zhou, M. M. Lu, T. P. Yamaguchi, and E. E. Morrisey. 2009. 'Wnt2/2b and beta-catenin signaling are necessary and sufficient to specify lung progenitors in the foregut', *Dev Cell*, 17: 290-8.

Guasch, G., M. Schober, H. A. Pasolli, E. B. Conn, L. Polak, and E. Fuchs. 2007. 'Loss of TGFbeta signaling destabilizes homeostasis and promotes squamous cell carcinomas in stratified epithelia', *Cancer Cell*, 12: 313-27.

Guillemot, F., L. C. Lo, J. E. Johnson, A. Auerbach, D. J. Anderson, and A. L. Joyner. 1993. 'Mammalian achaete-scute homolog 1 is required for the early development of olfactory and autonomic neurons', *Cell*, 75: 463-76.

Guyot, B., and V. Maguer-Satta. 2016. 'Blocking TGF-beta and BMP SMAD-dependent cell differentiation is a master key to expand all kinds of epithelial stem cells', *Stem Cell Investig*, 3:88-91.

Haller, C., J. Piccand, F. De Franceschi, Y. Ohi, A. Bhoumik, C. Boss, U. De Marchi, G. Jacot, S. Metairon, P. Descombes, A. Wiederkehr, A. Palini, N. Bouche, P. Steiner, O. G. Kelly, and R. C. Kraus M. 2019. 'Macroencapsulated Human iPSC-Derived Pancreatic Progenitors Protect against STZ-Induced Hyperglycemia in Mice', *Stem Cell Reports*, 12: 787-800.

Han Lu , H. Koike, P. Chaturvedi, K. Kishimoto, K. Iwasawa, K. Giesbrecht, PC. Witcher, A. Eicher, T. Nasr, L. Haines, JM. Shannon, M. Morimoto, JM. Wells, T. Takebe, and AM. Zorn. 2019. 'Single cell transcriptomics reveals a signaling network coordinating endoderm and mesoderm lineage diversification during foregut organogenesis' *Nat Commun*,11(1):4158.

Hao, M. M., J. P. Foong, J. C. Bornstein, Z. L. Li, P. Vanden Berghe, and W. Boesmans. 2016. 'Enteric nervous system assembly: Functional integration within the developing gut', *Dev Biol*, 417: 168-81.

Harris-Johnson, K. S., E. T. Domyan, C. M. Vezina, and X. Sun. 2009. 'beta-Catenin promotes respiratory progenitor identity in mouse foregut', *Proc Natl Acad Sci U S A*, 106: 16287-92.

Hassan, M., and H. Mousa. 2017. 'Impedance Testing in Esophageal Atresia Patients', *Front Pediatr*, 5: 85.

Hou, Z., Q. Wu, X. Sun, H. Chen, Y. Li, Y. Zhang, M. Mori, Y. Yang, J. Que, and M. Jiang. 2019. 'Wnt/Fgf crosstalk is required for the specification of basal cells in the mouse trachea', *Development*, (3)171496-146.

Hrstka, S. C., X. Li, T. J. Nelson, and Group Wanek Program Genetics Pipeline. 2017. 'NOTCH1-Dependent Nitric Oxide Signaling Deficiency in Hypoplastic Left Heart Syndrome Revealed Through Patient-Specific Phenotypes Detected in Bioengineered Cardiogenesis', *Stem Cells*, 35: 1106-19.

Hryniuk, A., S. Grainger, J. G. Savory, and D. Lohnes. 2012. 'Cdx function is required for maintenance of intestinal identity in the adult', *Dev Biol*, 363: 426-37.

Hsieh, H., A. Frenette, L. Michaud, U. Krishnan, D. B. Dal-Soglio, F. Gottrand, and C. Faure. 2017. 'Intestinal Metaplasia of the Esophagus in Children With Esophageal Atresia', *J Pediatr Gastroenterol Nutr*, 65: 1-4.

Huang, S. X., M. D. Green, A. T. de Carvalho, M. Mumau, Y. W. Chen, S. L. D'Souza, and H. W. Snoeck. 2015. 'The in vitro generation of lung and airway progenitor cells from human pluripotent stem cells', *Nat Protoc*, 10: 413-25.

Huang, S. X., M. N. Islam, J. O'Neill, Z. Hu, Y. G. Yang, Y. W. Chen, M. Mumau, M. D. Green, G. Vunjak-Novakovic, J. Bhattacharya, and H. W. Snoeck. 2014. 'Efficient generation of lung and airway epithelial cells from human pluripotent stem cells', *Nat Biotechnol*, 32: 84-91.

Huo, X., H. Y. Zhang, X. I. Zhang, J. P. Lynch, E. D. Strauch, J. Y. Wang, S. D. Melton, R. M. Genta, D. H. Wang, S. J. Spechler, and R. F. Souza. 2010. 'Acid and bile salt-induced CDX2 expression differs in esophageal squamous cells from patients with and without Barrett's esophagus', *Gastroenterology*, 139: 194-203.

Hvid-Jensen, F., L. Pedersen, A. M. Drewes, H. T. Sorensen, and P. Funch-Jensen. 2011. 'Incidence of adenocarcinoma among patients with Barrett's esophagus', *N Engl J Med*, 365: 1375-83.

Ioannides, A. S., V. Massa, E. Ferraro, F. Cecconi, L. Spitz, D. J. Henderson, and A. J. Copp. 2010. 'Foregut separation and tracheo-oesophageal malformations: the role of tracheal outgrowth, dorso-ventral patterning and programmed cell death', *Dev Biol*, 337: 351-62.

Jacobs, I. J., W. Y. Ku, and J. Que. 2012. 'Genetic and cellular mechanisms regulating anterior foregut and esophageal development', *Dev Biol*, 369: 54-64.

Jiang, M., H. Li, Y. Zhang, Y. Yang, R. Lu, K. Liu, S. Lin, X. Lan, H. Wang, H. Wu, J. Zhu, Z. Zhou, J. Xu, D. K. Lee, L. Zhang, Y. C. Lee, J. Yuan, J. A. Abrams, T. C. Wang, A. R. Sepulveda, Q. Wu, H. Chen, X. Sun, J. She, X. Chen, and J. Que. 2017. 'Transitional basal cells at the squamous-columnar junction generate Barrett's oesophagus', *Nature*, 550: 529-33.

Jurado, S., I. Smyth, B. van Denderen, N. Tenis, A. Hammet, K. Hewitt, J. L. Ng, C. J. McNees, S. V. Kozlov, H. Oka, M. Kobayashi, L. A. Conlan, T. J. Cole, K. Yamamoto, Y. Taniguchi, S. Takeda, M. F. Lavin, and J. Heierhorst. 2010. 'Dual functions of ASCIZ in the DNA base damage response and pulmonary organogenesis', *PLoS Genet*, 6: e1001170.

Jurand, A. 1974. 'Some aspects of the development of the notochord in mouse embryos', *J Embryol Exp Morphol*, 32: 1-33.

Kapoor, H., K. R. Lohani, T. H. Lee, D. K. Agrawal, and S. K. Mittal. 2015. 'Animal Models of Barrett's Esophagus and Esophageal Adenocarcinoma-Past, Present, and Future', *Clin Transl Sci*, 8: 841-7.

Karagiannis, P., K. Takahashi, M. Saito, Y. Yoshida, K. Okita, A. Watanabe, H. Inoue, J. K. Yamashita, M. Todani, M. Nakagawa, M. Osawa, Y. Yashiro, S. Yamanaka, and K. Osafune. 2019. 'Induced Pluripotent Stem Cells and Their Use in Human Models of Disease and Development', *Physiol Rev*, 99: 79-114.

Kim, B. M., G. Buchner, I. Miletich, P. T. Sharpe, and R. A. Shivdasani. 2005. 'The stomach mesenchymal transcription factor Barx1 specifies gastric epithelial identity through inhibition of transient Wnt signaling', *Dev Cell*, 8: 611-22.
Kim, B. M., I. Miletich, J. Mao, A. P. McMahon, P. A. Sharpe, and R. A. Shivdasani. 2007. 'Independent functions and mechanisms for homeobox gene Barx1 in patterning mouse stomach and spleen', *Development*, 134: 3603-13.
Kim, E., M. Jiang, H. Huang, Y. Zhang, N. Tjota, X. Gao, J. Robert, N. Gilmore, L. Gan, and J. Que. 2019. 'Isl1 Regulation of Nkx2.1 in the Early Foregut Epithelium Is Required for Trachea-Esophageal Separation and Lung Lobation', *Dev Cell*, 51: 675-83.
Koivusalo, A., M. P. Pakarinen, and R. J. Rintala. 2007. 'The cumulative incidence of significant gastrooesophageal reflux in patients with oesophageal atresia with a distal fistula--a systematic clinical, pH-metric, and endoscopic follow-up study', *J Pediatr Surg*, 42: 370-4.
Kovaka, S., A. V. Zimin, G. M. Pertea, R. Razaghi, S. L. Salzberg, and M. Pertea. 2019. 'Transcriptome assembly from long-read RNA-seq alignments with StringTie2', *Genome Biol*, 20:(1) 278.
Kovesi, T., and S. Rubin. 2004. 'Long-term complications of congenital esophageal atresia and/or tracheoesophageal fistula', *Chest*, 126: 915-25.
Krishnadath, K. K., and K. K. Wang. 2015. 'Molecular pathogenesis of Barrett esophagus: current evidence', *Gastroenterol Clin North Am*, 44: 233-47.
Krishnan, U., H. Mousa, L. Dall'Oglio, N. Homaira, R. Rosen, C. Faure, and F. Gottrand. 2016. 'ESPGHAN-NASPGHAN Guidelines for the Evaluation and Treatment of Gastrointestinal and Nutritional Complications in Children With Esophageal Atresia-Tracheoesophageal Fistula', *J Pediatr Gastroenterol Nutr*, 63: 550-70.
Kuwahara, A., A. E. Lewis, C. Coombes, F. S. Leung, M. Percharde, and J. O. Bush. 2020. 'Delineating the early transcriptional specification of the mammalian trachea and esophagus', *Elife*, 9:e55526.
Larsen, William J. 1997. 'Human Embryology 2nd edition.' in, *Vasculature* (Churchill Livingstone: New York).
Li, H. 2018. 'Minimap2: pairwise alignment for nucleotide sequences', *Bioinformatics*, 34: 3094-100.
Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and Subgroup Genome Project Data Processing. 2009. 'The Sequence Alignment/Map format and SAMtools', *Bioinformatics*, 25: 2078-9.
Li, Y., J. Gordon, N. R. Manley, Y. Litingtung, and C. Chiang. 2008. 'Bmp4 is required for tracheal formation: a novel mouse model for tracheal agenesis', *Dev Biol*, 322: 145-55.
Li, Y., Y. Litingtung, P. Ten Dijke, and C. Chiang. 2007. 'Aberrant Bmp signaling and notochord delamination in the pathogenesis of esophageal atresia', *Dev Dyn*, 236: 746-54.
Litingtung, Y., L. Lei, H. Westphal, and C. Chiang. 1998. 'Sonic hedgehog is essential to foregut development', *Nat Genet*, 20: 58-61.
Londahl, M., A. L. Irace, K. Kawai, N. D. Dombrowski, R. Jennings, and R. Rahbar. 2018. 'Prevalence of Laryngeal Cleft in Pediatric Patients With Esophageal Atresia', *JAMA Otolaryngol Head Neck Surg*, 144: 164-68.
Love, M. I., W. Huber, and S. Anders. 2014. 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biol*, 15: 550-8.

Luo, J., H. M. Sucov, J. A. Bader, R. M. Evans, and V. Giguere. 1996. 'Compound mutants for retinoic acid receptor (RAR) beta and RAR alpha 1 reveal developmental functions for multiple RAR beta isoforms', *Mech Dev*, 55: 33-44.

Marques-Pereira, J. P., and C. P. Leblond. 1965. 'Mitosis and Differentiation in the Stratified Squamous Epithelium of the Rat Esophagus', *Am J Anat*, 117: 73-87.

Matsuno, K., S. I. Mae, C. Okada, M. Nakamura, A. Watanabe, T. Toyoda, E. Uchida, and K. Osafune. 2016. 'Redefining definitive endoderm subtypes by robust induction of human induced pluripotent stem cells', *Differentiation*, 92: 281-90.

McCann, F., L. Michaud, A. Aspirot, D. Levesque, F. Gottrand, and C. Faure. 2015. 'Congenital esophageal stenosis associated with esophageal atresia', *Dis Esophagus*, 28: 211-5.

McCauley, H. A., and J. M. Wells. 2017. 'Pluripotent stem cell-derived organoids: using principles of developmental biology to grow human tissues in a dish', *Development*, 144: 958-62.

Mendelsohn, C., D. Lohnes, D. Decimo, T. Lufkin, M. LeMeur, P. Chambon, and M. Mark. 1994. 'Function of the retinoic acid receptors (RARs) during development (II). Multiple abnormalities at various stages of organogenesis in RAR double mutants', *Development*, 120: 2749-71.

Miao, Y., L. Tian, M. Martin, S. L. Paige, F. X. Galdos, J. Li, A. Klein, H. Zhang, N. Ma, Y. Wei, M. Stewart, S. Lee, J. R. Moonen, B. Zhang, P. Grossfeld, S. Mital, D. Chitayat, J. C. Wu, M. Rabinovitch, T. J. Nelson, S. Nie, S. M. Wu, and M. Gu. 2020. 'Intrinsic Endocardial Defects Contribute to Hypoplastic Left Heart Syndrome', *Cell Stem Cell*, 27: 574-89.

Minacapelli, C. D., M. Bajpai, X. Geng, C. L. Cheng, A. A. Chouthai, R. Souza, S. J. Spechler, and K. M. Das. 2017. 'Barrett's metaplasia develops from cellular reprograming of esophageal squamous epithelium due to gastroesophageal reflux', *Am J Physiol Gastrointest Liver Physiol*, 312: 615-22,.

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

Morrisey, E. E., and B. L. Hogan. 2010. 'Preparing for the first breath: genetic and cellular mechanisms in lung development', *Dev Cell*, 18: 8-23.

Motoyama, J., J. Liu, R. Mo, Q. Ding, M. Post, and C. C. Hui. 1998. 'Essential function of Gli2 and Gli3 in the formation of lung, trachea and oesophagus', *Nat Genet*, 20: 54-7.

Munera, J. O., N. Sundaram, S. A. Rankin, D. Hill, C. Watson, M. Mahe, J. E. Vallance, N. F. Shroyer, K. L. Sinagoga, A. Zarzoso-Lacoste, J. R. Hudson, J. C. Howell, P. Chatuvedi, J. R. Spence, J. M. Shannon, A. M. Zorn, M. A. Helmrath, and J. M. Wells. 2017. 'Differentiation of Human Pluripotent Stem Cells into Colonic Organoids via Transient Activation of BMP Signaling', *Cell Stem Cell*, 21: 51-64.

Nakano, T., S. Ando, N. Takata, M. Kawada, K. Muguruma, K. Sekiguchi, K. Saito, S. Yonemura, M. Eiraku, and Y. Sasai. 2012. 'Self-formation of optic cups and storable stratified neural retina from human ESCs', *Cell Stem Cell*, 10: 771-85.

Nam, K. T., H. J. Lee, J. J. Smith, L. A. Lapierre, V. P. Kamath, X. Chen, B. J. Aronow, T. J. Yeatman, S. G. Bhartur, B. C. Calhoun, B. Condie, N. R. Manley, R. D. Beauchamp, R. J. Coffey, and J. R. Goldenring. 2010. 'Loss of Rab25 promotes the development of intestinal neoplasia in mice and is associated with human colorectal adenocarcinomas', *J Clin Invest*, 120: 840-9.

Nasr, T., P. Mancini, S. A. Rankin, N. A. Edwards, Z. N. Agricola, A. P. Kenny, J. L. Kinney, K. Daniels, J. Vardanyan, L. Han, S. L. Trisno, S. W. Cha, J. M. Wells, M. J. Kofron, and A.

M. Zorn. 2019. 'Endosome-Mediated Epithelial Remodeling Downstream of Hedgehog-Gli Is Required for Tracheoesophageal Separation', *Dev Cell*, 51: 665-74.
Nassar, N., E. Leoncini, E. Amar, J. Arteaga-Vazquez, M. K. Bakker, C. Bower, M. A. Canfield, E. E. Castilla, G. Cocchi, A. Correa, M. Csaky-Szunyogh, M. L. Feldkamp, B. Khoshnood, D. Landau, N. Lelong, J. S. Lopez-Camelo, R. B. Lowry, R. McDonnell, P. Merlob, J. Metneki, M. Morgan, O. M. Mutchinick, M. N. Palmer, A. Rissmann, C. Siffel, A. Sipek, E. Szabova, D. Tucker, and P. Mastroiacovo. 2012. 'Prevalence of esophageal atresia among 18 international birth defects surveillance programs', *Birth Defects Res A Clin Mol Teratol*, 94: 893-9.
Negoro, R., K. Takayama, K. Kawai, K. Harada, F. Sakurai, K. Hirata, and H. Mizuguchi. 2018. 'Efficient Generation of Small Intestinal Epithelial-like Cells from Human iPSCs for Drug Absorption and Metabolism Studies', *Stem Cell Reports*, 11: 1539-50.
Niehrs, C. 2012. 'The complex world of WNT receptor signalling', *Nat Rev Mol Cell Biol*, 13: 767-79.
O'Rahilly, R., and F. Muller. 1984. 'Chevalier Jackson lecture. Respiratory and alimentary relations in staged human embryos. New embryological data and congenital anomalies', *Ann Otol Rhinol Laryngol*, 93: 421-9.
Obermayr, F., R. Hotta, H. Enomoto, and H. M. Young. 2013. 'Development and developmental disorders of the enteric nervous system', *Nat Rev Gastroenterol Hepatol*, 10: 43-57.
Oshimori, N., and E. Fuchs. 2012. 'Paracrine TGF-beta signaling counterbalances BMP-mediated repression in hair follicle stem cell activation', *Cell Stem Cell*, 10: 63-75.
Patro, R., G. Duggal, M. I. Love, R. A. Irizarry, and C. Kingsford. 2017. 'Salmon provides fast and bias-aware quantification of transcript expression', *Nat Methods*, 14: 417-19.
Pavlov, K., C. Meijer, A. van den Berg, F. T. Peters, F. A. Kruyt, and J. H. Kleibeuker. 2014. 'Embryological signaling pathways in Barrett's metaplasia development and malignant transformation; mechanisms and therapeutic opportunities', *Crit Rev Oncol Hematol*, 92: 25-37.
Perl, A. K., S. E. Wert, A. Nagy, C. G. Lobe, and J. A. Whitsett. 2002. 'Early restriction of peripheral and proximal cell lineages during formation of the lung', *Proc Natl Acad Sci U S A*, 99: 10482-7.
Pertea, G., and M. Pertea. 2020. 'GFF Utilities: GffRead and GffCompare', *F1000Res*, 9.
Peters, J. H., and N. Avisar. 2010. 'The molecular pathogenesis of Barrett's esophagus: common signaling pathways in embryogenesis metaplasia and neoplasia', *J Gastrointest Surg*, 14 Suppl 1: S81-7.
Pinzon-Guzman, C., S. Sangadala, K. M. Riera, E. Y. Popova, E. Manning, W. J. Huh, M. S. Alexander, J. S. Shelton, S. D. Boden, and J. R. Goldenring. 2020. 'Noggin regulates foregut progenitor cell programming and mis-expression leads to esophageal atresia', *J Clin Invest*, 130(8), 4396-4410.
Pultrum, B. B., C. M. Bijleveld, Z. J. de Langen, and J. T. Plukker. 2005. 'Development of an adenocarcinoma of the esophagus 22 years after primary repair of a congenital atresia', *J Pediatr Surg*, 40: 1-4.
Qi, B. Q., and S. W. Beasley. 2000. 'Stages of normal tracheo-bronchial development in rat embryos: resolution of a controversy', *Dev Growth Differ*, 42: 145-53.
Quante, M., G. Bhagat, J. A. Abrams, F. Marache, P. Good, M. D. Lee, Y. Lee, R. Friedman, S. Asfaha, Z. Dubeykovskaya, U. Mahmood, J. L. Figueiredo, J. Kitajewski, C. Shawber, C. J. Lightdale, A. K. Rustgi, and T. C. Wang. 2012. 'Bile acid and inflammation activate

gastric cardia stem cells in a mouse model of Barrett-like metaplasia', *Cancer Cell*, 21: 36-51.
Que, J. 2015. 'The initial establishment and epithelial morphogenesis of the esophagus: a new model of tracheal-esophageal separation and transition of simple columnar into stratified squamous epithelium in the developing esophagus', *Wiley Interdiscip Rev Dev Biol*, 4: 419-30.
Que, J., M. Choi, J. W. Ziel, J. Klingensmith, and B. L. Hogan. 2006. 'Morphogenesis of the trachea and esophagus: current players and new roles for noggin and Bmps', *Differentiation*, 74: 422-37.
Que, J., K. S. Garman, R. F. Souza, and S. J. Spechler. 2019. 'Pathogenesis and Cells of Origin of Barrett's Esophagus', *Gastroenterology*, 157: 349-64.
Que, J., T. Okubo, J. R. Goldenring, K. T. Nam, R. Kurotani, E. E. Morrisey, O. Taranova, L. H. Pevny, and B. L. Hogan. 2007. 'Multiple dose-dependent roles for Sox2 in the patterning and differentiation of anterior foregut endoderm', *Development*, 134: 2521-31.
Raad, S., A. David, J. Que, and C. Faure. 2020. 'Genetic Mouse Models and Induced Pluripotent Stem Cells for Studying Tracheal-Esophageal Separation and Esophageal Development', *Stem Cells Dev*, 29: 953-66.
Rahman, M. S., N. Akhtar, H. M. Jamil, R. S. Banik, and S. M. Asaduzzaman. 2015. 'TGF-beta/BMP signaling and other molecular events: regulation of osteoblastogenesis and bone formation', *Bone Res*, 3: 15005.
Rankin, S. A., L. Han, K. W. McCracken, A. P. Kenny, C. T. Anglin, E. A. Grigg, C. M. Crawford, J. M. Wells, J. M. Shannon, and A. M. Zorn. 2016. 'A Retinoic Acid-Hedgehog Cascade Coordinates Mesoderm-Inducing Signals and Endoderm Competence during Lung Specification', *Cell Rep*, 16: 66-78.
Rao, M., and M. D. Gershon. 2018. 'Enteric nervous system development: what could possibly go wrong?', *Nat Rev Neurosci*, 19: 552-65.
Rhinn, M., and P. Dolle. 2012. 'Retinoic acid signalling during development', *Development*, 139: 843-58.
Rimkus, T. K., R. L. Carpenter, S. Qasem, M. Chan, and H. W. Lo. 2016. 'Targeting the Sonic Hedgehog Signaling Pathway: Review of Smoothened and GLI Inhibitors', *Cancers (Basel)*, 8(2),22.
Rodriguez, P., S. Da Silva, L. Oxburgh, F. Wang, B. L. Hogan, and J. Que. 2010. 'BMP signaling in the development of the mouse esophagus and forestomach', *Development*, 137: 4171-6.
Rosekrans, S. L., B. Baan, V. Muncan, and G. R. van den Brink. 2015. 'Esophageal development and epithelial homeostasis', *Am J Physiol Gastrointest Liver Physiol*, 309: G216-28.
Rowe, R. G., and G. Q. Daley. 2019. 'Induced pluripotent stem cells in disease modelling and drug discovery', *Nat Rev Genet*, 20: 377-88.
Runge, T. M., J. A. Abrams, and N. J. Shaheen. 2015. 'Epidemiology of Barrett's Esophagus and Esophageal Adenocarcinoma', *Gastroenterol Clin North Am*, 44: 203-31.
Sala, F. G., P. M. Del Moral, C. Tiozzo, D. A. Alam, D. Warburton, T. Grikscheit, J. M. Veltmaat, and S. Bellusci. 2011. 'FGF10 controls the patterning of the tracheal cartilage rings via Shh', *Development*, 138: 273-82.
Sarkar, A., and K. Hochedlinger. 2013. 'The sox family of transcription factors: versatile regulators of stem and progenitor cell fate', *Cell Stem Cell*, 12: 15-30.
Sasaki, T., T. Kusafuka, and A. Okada. 2001. 'Analysis of the development of normal foregut and tracheoesophageal fistula in an adriamycin rat model using three-dimensional image reconstruction', *Surg Today*, 31: 133-9.

Sasaki, Y., N. Iwai, T. Tsuda, and O. Kimura. 2004. 'Sonic hedgehog and bone morphogenetic protein 4 expressions in the hindgut region of murine embryos with anorectal malformations', *J Pediatr Surg*, 39: 170-3; discussion 70-3.

Schulz, A. C., E. Bartels, R. Stressig, J. Ritgen, E. Schmiedeke, M. Mattheisen, M. Draaken, M. Ludwig, S. Bagci, A. Muller, U. Gembruch, A. Geipel, C. Berg, A. Heydweiller, H. Bachour, J. Schumacher, P. Bartmann, M. M. Nothen, and H. Reutter. 2012. 'Nine new twin pairs with esophageal atresia: a review of the literature and performance of a twin study of the disorder', *Birth Defects Res A Clin Mol Teratol*, 94: 182-6.

Shahryari, A., M. R. Rafiee, Y. Fouani, N. A. Oliae, N. M. Samaei, M. Shafiee, S. Semnani, M. Vasei, and S. J. Mowla. 2014. 'Two novel splice variants of SOX2OT, SOX2OT-S1, and SOX2OT-S2 are coupregulated with SOX2 and OCT4 in esophageal squamous cell carcinoma', *Stem Cells*, 32: 126-34.

Shelton, M., A. Kocharyan, J. Liu, I. S. Skerjanc, and W. L. Stanford. 2016. 'Robust generation and expansion of skeletal muscle progenitors and myocytes from human pluripotent stem cells', *Methods*, 101: 73-84.

Shen, M. M. 2007. 'Nodal signaling: developmental roles and regulation', *Development*, 134: 1023-34.

Shi, Y., P. Kirwan, J. Smith, H. P. Robinson, and F. J. Livesey. 2012. 'Human cerebral cortex development from pluripotent stem cells to functional excitatory synapses', *Nat Neurosci*, 15: 477-86.

Shields, H. M., F. Zwas, D. A. Antonioli, W. G. Doos, S. Kim, and S. J. Spechler. 1993. 'Detection by scanning electron microscopy of a distinctive esophageal surface cell at the junction of squamous and Barrett's epithelium', *Dig Dis Sci*, 38: 97-108.

Spechler, S. J., and R. F. Souza. 2014. 'Barrett's esophagus', *N Engl J Med*, 371: 836-45.

Stairs, D. B., H. Nakagawa, A. Klein-Szanto, S. D. Mitchell, D. G. Silberg, J. W. Tobias, J. P. Lynch, and A. K. Rustgi. 2008. 'Cdx1 and c-Myc foster the initiation of transdifferentiation of the normal esophageal squamous epithelium toward Barrett's esophagus', *PLoS One*, 3: e3534.

Stoll, C., Y. Alembik, B. Dott, and M. P. Roth. 2009. 'Associated malformations in patients with esophageal atresia', *Eur J Med Genet*, 52: 287-90.

Su, Y., X. Chen, M. Klein, M. Fang, S. Wang, C. S. Yang, and R. K. Goyal. 2004. 'Phenotype of columnar-lined esophagus in rats with esophagogastroduodenal anastomosis: similarity to human Barrett's esophagus', *Lab Invest*, 84: 753-65.

Takahashi, K., and S. Yamanaka. 2006. 'Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors', *Cell*, 126: 663-76.

Takebe, T., K. Sekine, M. Enomura, H. Koike, M. Kimura, T. Ogaeri, R. R. Zhang, Y. Ueno, Y. W. Zheng, N. Koike, S. Aoyama, Y. Adachi, and H. Taniguchi. 2013. 'Vascularized and functional human liver from an iPSC-derived organ bud transplant', *Nature*, 499: 481-4.

Ten Kate, C. A., A. de Klein, B. M. de Graaf, M. Doukas, A. Koivusalo, M. P. Pakarinen, R. van der Helm, T. Brands, I. Jsselstijn H, Y. van Bever, R. M. H. Wijnen, M. C. W. Spaander, and E. Brosens. 2022. 'Intrinsic Cellular Susceptibility to Barrett's Esophagus in Adults Born with Esophageal Atresia', *Cancers (Basel)*, 14(3)-513.

Teramoto, M., R. Sugawara, K. Minegishi, M. Uchikawa, T. Takemoto, A. Kuroiwa, Y. Ishii, and H. Kondoh. 2020a. 'The absence of SOX2 in the anterior foregut alters the esophagus into trachea and bronchi in both epithelial and mesenchymal components', *Biol Open*, 9: 048728.

———. 2020b. 'The absence of SOX2 in the anterior foregut alters the esophagus into trachea and bronchi in both epithelial and mesenchymal components', *Biol Open*, 9(2).
Trisno, S. L., K. E. D. Philo, K. W. McCracken, E. M. Cata, S. Ruiz-Torres, S. A. Rankin, L. Han, T. Nasr, P. Chaturvedi, M. E. Rothenberg, M. A. Mandegar, S. I. Wells, A. M. Zorn, and J. M. Wells. 2018. 'Esophageal Organoids from Human Pluripotent Stem Cells Delineate Sox2 Functions during Esophageal Specification', *Cell Stem Cell*, 23: 501-15.
Tullie, L., A. Kelay, G. S. Bethell, C. Major, and N. J. Hall. 2021. 'Barrett's oesophagus and oesophageal cancer following oesophageal atresia repair: a systematic review', *BJS Open*, 5(4).
van Baal, J. W., F. Milana, A. M. Rygiel, C. M. Sondermeijer, C. A. Spek, J. J. Bergman, M. P. Peppelenbosch, and K. K. Krishnadath. 2008. 'A comparative analysis by SAGE of gene expression profiles of esophageal adenocarcinoma and esophageal squamous cell carcinoma', *Cell Oncol*, 30: 63-75.
van Lennep, M., M. M. J. Singendonk, L. Dall'Oglio, F. Gottrand, U. Krishnan, S. W. J. Terheggen-Lagro, T. I. Omari, M. A. Benninga, and M. P. van Wijk. 2019. 'Oesophageal atresia', *Nat Rev Dis Primers*, 5: 26.
Van Staey, M., S. De Bie, M. T. Matton, and J. De Roose. 1984. 'Familial congenital esophageal atresia. Personal case report and review of the literature', *Hum Genet*, 66: 260-6.
Vergouwe, F. W., M. Gottrand, B. P. Wijnhoven, I. Jsselstijn H, G. Piessen, M. J. Bruno, R. M. Wijnen, and M. C. Spaander. 2018. 'Four cancer cases after esophageal atresia repair: Time to start screening the upper gastrointestinal tract', *World J Gastroenterol*, 24: 1056-62.
Vergouwe, F. W. T., I. Jsselstijn H, K. Biermann, N. S. Erler, R. M. H. Wijnen, M. J. Bruno, and M. C. W. Spaander. 2018. 'High Prevalence of Barrett's Esophagus and Esophageal Squamous Cell Carcinoma After Repair of Esophageal Atresia', *Clin Gastroenterol Hepatol*, 16: 513-21.
Volckaert, T., A. Campbell, E. Dill, C. Li, P. Minoo, and S. De Langhe. 2013. 'Localized Fgf10 expression is not required for lung branching morphogenesis but prevents differentiation of epithelial progenitors', *Development*, 140: 3731-42.
Volckaert, T., and S. P. De Langhe. 2015. 'Wnt and FGF mediated epithelial-mesenchymal crosstalk during lung development', *Dev Dyn*, 244: 342-66.
Wallace, A. S., and A. J. Burns. 2005. 'Development of the enteric nervous system, smooth muscle and interstitial cells of Cajal in the human gastrointestinal tract', *Cell Tissue Res*, 319: 367-82.
Wang, D. H. 2017. 'The Esophageal Squamous Epithelial Cell-Still a Reasonable Candidate for the Barrett's Esophagus Cell of Origin?', *Cell Mol Gastroenterol Hepatol*, 4: 157-60.
Wang, D. H., N. J. Clemons, T. Miyashita, A. J. Dupuy, W. Zhang, A. Szczepny, I. M. Corcoran-Schwartz, D. L. Wilburn, E. A. Montgomery, J. S. Wang, N. A. Jenkins, N. A. Copeland, J. W. Harmon, W. A. Phillips, and D. N. Watkins. 2010. 'Aberrant epithelial-mesenchymal Hedgehog signaling characterizes Barrett's metaplasia', *Gastroenterology*, 138: 1810-22.
Wang, D. H., A. Tiwari, M. E. Kim, N. J. Clemons, N. L. Regmi, W. A. Hodges, D. M. Berman, E. A. Montgomery, D. N. Watkins, X. Zhang, Q. Zhang, C. Jie, S. J. Spechler, and R. F. Souza. 2014. 'Hedgehog signaling regulates FOXA2 in esophageal embryogenesis and Barrett's metaplasia', *J Clin Invest*, 124: 3767-80.
Wang, J., P. R. Ahimaz, S. Hashemifar, J. Khlevner, J. A. Picoraro, W. Middlesworth, M. M. Elfiky, J. Que, Y. Shen, and W. K. Chung. 2021. 'Novel candidate genes in esophageal

atresia/tracheoesophageal fistula identified by exome sequencing', *Eur J Hum Genet*, 29: 122-30.
Wang, X., H. Ouyang, Y. Yamamoto, P. A. Kumar, T. S. Wei, R. Dagher, M. Vincent, X. Lu, A. M. Bellizzi, K. Y. Ho, C. P. Crum, W. Xian, and F. McKeon. 2011. 'Residual embryonic cells as precursors of a Barrett's-like metaplasia', *Cell*, 145: 1023-35.
Wang, Z., P. Dolle, W. V. Cardoso, and K. Niederreither. 2006. 'Retinoic acid regulates morphogenesis and patterning of posterior foregut derivatives', *Dev Biol*, 297: 433-45.
Wells, J. M., and D. A. Melton. 1999. 'Vertebrate endoderm development', *Annu Rev Cell Dev Biol*, 15: 393-410.
Woo, J., I. Miletich, B. M. Kim, P. T. Sharpe, and R. A. Shivdasani. 2011. 'Barx1-mediated inhibition of Wnt signaling in the mouse thoracic foregut controls tracheo-esophageal septation and epithelial differentiation', *PLoS One*, 6: e22493.
Wunnemann, F., A. Ta-Shma, C. Preuss, S. Leclerc, P. P. van Vliet, A. Oneglia, M. Thibeault, E. Nordquist, J. Lincoln, F. Scharfenberg, C. Becker-Pauly, P. Hofmann, K. Hoff, E. Audain, H. H. Kramer, W. Makalowski, A. Nir, S. S. Gerety, M. Hurles, J. Comes, A. Fournier, H. Osinska, J. Robins, M. Puceat, Mibava Leducq Consortium principal investigators, O. Elpeleg, M. P. Hitz, and G. Andelfinger. 2020. 'Loss of ADAMTS19 causes progressive non-syndromic heart valve disease', *Nat Genet*, 52: 40-47.
Yang, C., Y. Xu, M. Yu, D. Lee, S. Alharti, N. Hellen, N. Ahmad Shaik, B. Banaganapalli, H. Sheikh Ali Mohamoud, R. Elango, S. Przyborski, G. Tenin, S. Williams, J. O'Sullivan, O. O. Al-Radi, J. Atta, S. E. Harding, B. Keavney, M. Lako, and L. Armstrong. 2017. 'Induced pluripotent stem cell modelling of HLHS underlines the contribution of dysfunctional NOTCH signalling to impaired cardiogenesis', *Hum Mol Genet*, 26: 3031-45.
Yang, L., L. S. Wang, X. L. Chen, Z. Gatalica, S. Qiu, Z. Liu, G. Stoner, H. Zhang, H. Weiss, and J. Xie. 2012. 'Hedgehog signaling activation in the development of squamous cell carcinoma and adenocarcinoma of esophagus', *Int J Biochem Mol Biol*, 3: 46-57.
Yu, W. Y., J. M. Slack, and D. Tosh. 2005. 'Conversion of columnar to stratified squamous epithelium in the developing mouse oesophagus', *Dev Biol*, 284: 157-70.
Yuan, B., C. Li, S. Kimura, R. T. Engelhardt, B. R. Smith, and P. Minoo. 2000. 'Inhibition of distal lung morphogenesis in Nkx2.1(-/-) embryos', *Dev Dyn*, 217: 180-90.
Yuan, T., T. Volckaert, D. Chanda, V. J. Thannickal, and S. P. De Langhe. 2018. 'Fgf10 Signaling in Lung Development, Homeostasis, Disease, and Repair After Injury', *Front Genet*, 9: 418.
Zhang, R. R., T. Takebe, L. Miyazaki, M. Takayama, H. Koike, M. Kimura, M. Enomura, Y. W. Zheng, K. Sekine, and H. Taniguchi. 2014. 'Efficient hepatic differentiation of human induced pluripotent stem cells in a three-dimensional microscale culture', *Methods Mol Biol*, 1210: 131-41.
Zhang, Y., M. Jiang, E. Kim, S. Lin, K. Liu, X. Lan, and J. Que. 2017. 'Development and stem cells of the esophagus', *Semin Cell Dev Biol*, 66: 25-35.
Zhang, Y., Y. Yang, M. Jiang, S. X. Huang, W. Zhang, D. Al Alam, S. Danopoulos, M. Mori, Y. W. Chen, R. Balasubramanian, S. M. Chuva de Sousa Lopes, C. Serra, M. Bialecka, E. Kim, S. Lin, A. L. R. Toste de Carvalho, P. N. Riccio, W. V. Cardoso, X. Zhang, H. W. Snoeck, and J. Que. 2018. '3D Modeling of Esophageal Development using Human PSC-Derived Basal Progenitors Reveals a Critical Role for Notch Signaling', *Cell Stem Cell*, 23: 516-29.

Zhou, G., Y. G. Sun, H. B. Wang, W. Q. Wang, X. W. Wang, and D. C. Fang. 2009. 'Acid and bile salt up-regulate BMP4 expression in human esophageal epithelium cells', *Scand J Gastroenterol*, 44: 926-32.