TECHNISCHE UNIVERSITÄT MÜNCHEN

FAKULTÄT FÜR MEDIZIN

Identification of Transcriptional Networks Governing Embryonic Lung Development via Directed Differentiation of Human Pluripotent Stem Cells

Chaido Ori

Vollständiger Abdruck der von der
Fakultät für Medizin der Technischen Universität München
zur Erlangung des akademischen Grades einer
Doktorin der Naturwissenschaften (Dr. rer. nat.)
genehmigten Dissertation.

Vorsitz: Prof. Dr. Alessandra Moretti

Prüfer der Dissertation:

1. Prof. Dr. Heiko Lickert
2. Prof. Dr. Wolfgang Wurst (schriftliche Beurteilung)
   apl. Prof. Dr. Thomas Floss (mündliche Prüfung)

Die Dissertation wurde am 03.03.2022 bei der Technischen Universität München eingereicht und durch die Fakultät für Medizin am 09.08.2022 angenommen.

Parts of this thesis have been published as a pre-print on the preprint server bioRxiv (Ori et al., 2021).

ABSTRACT

During embryogenesis, cells commit to a singular lineage among multiple alternate fate choices by passing through dynamic transcriptional states. The elucidation of the pathways underlying human lung organogenesis is essential for modeling lung development and advancing therapies for respiratory disorders. Human pluripotent stem cells (hPSCs) constitute a unique tool for studying the cellular and molecular mechanisms governing embryonic lineage specification, especially in very early stages, which are inaccessible to study *in vivo*. The transcription factor Nk2 homeobox 1 (NKX2-1) distinguishes cells fated to become lung and thyroid in the foregut endoderm, and hence serves as a hallmark to identify PSC-derived progenitors that can generate functional respiratory cells. To date, the molecular mechanisms prompting human lung endoderm specification and patterning remain elusive.

To investigate how individual stem cells exit pluripotency and transit towards lung lineages, I generated NKX2-1+ lung progenitors from human induced pluripotent stem cells (hiPSCs) via a stepwise differentiation protocol, which recapitulates key mechanisms of lung development, and used single cell RNA sequencing (scRNA-seq) with high temporal resolution. I profiled the transcriptional hierarchy of lung specification from hPSCs and delineated the underlying transcriptional trajectories. To this end, I validated major developmental stages, and substantiated the multipotent capacity of NKX2-1+ progenitors by generating respiratory organoids that share transcriptomic characteristics with primary fetal lung. My work revealed that activation of Sonic hedgehog, TGF-β, and Notch signaling pathways is required in a FOXA2-ISL1-NKX2-1 trajectory that generates respiratory progenitors during the foregut endoderm stage. Additionally, I demonstrated that the induction of the transcription factor HHEX marks an alternative trajectory in the early definitive endoderm that precedes the lung lineage and gives rise to a major hepatoblast population. The high temporal resolution of single cell trajectory analysis allowed me to construct a hierarchical model of gene expression changes towards two lineages from hPSC; lung and liver, respectively.

In summary, my thesis outlines a novel comprehensive transcriptional roadmap underlying the human embryonic lung specification and provides fundamental knowledge for improving lung differentiation strategies with applications in regenerative medicine and modeling diseases with developmental origins.

ZUSAMMENFASSUNG

Während der Embryogenese durchlaufen Zellen dynamische Transkriptionszustände, die zur endgültigen Festlegung des Zellschicksals führen. Die Aufklärung der Wege, die der Organogenese der menschlichen Lunge zugrunde liegen, ist für die Entwicklung von Modellen zur Lungenentstehung und neuer Ansätze zur Behandlung von Atemwegserkrankungen von entscheidender Bedeutung. Humane pluripotente Stammzellen (hPSCs) sind ein einzigartiges Instrument zur Untersuchung der die Differenzierung steuernden zellulären und molekularen Mechanismen, insbesondere in sehr frühen Stadien, die der Untersuchung *in vivo* nicht zugänglich sind. Zellen, die sich im Entoderm des Vorderdarms zu Lungen- und Schilddrüsenzellen differenzieren, ist die Expression des Transkriptionsfaktors Nk2 homeobox 1 (NKX2-1) gemein; dieser dient daher als Merkmal zur Identifizierung von Vorläuferzellen aus hPSC, die funktionelle Atemwegszellen bilden können. Bis heute sind die molekularen Mechanismen, die zur Spezifizierung und Musterbildung innerhalb des menschlichen Lungenentoderms führen, nicht vollständig bekannt.

Zur Klärung der Frage wie einzelne Stammzellen ihre Pluripotenz zugunsten einer Spezifizierung Richtung Lungenzellinien verlieren, habe ich NKX2-1+ Lungenvorläufer aus humanen induzierten pluripotenten Stammzellen (hiPSCs) durch ein schrittweises Differenzierungsprotokoll erzeugt, das Schlüsselmechanismen der Lungenentwicklung rekapituliert und Einzelzell-RNA-Sequenzierung (scRNA-seq) mit hoher zeitlicher Auflösung verwendet. Hierbei habe ich die transkriptionelle Hierarchie der Lungenspezifizierung von hPSCs aufgeklärt und die zugrunde liegenden transkriptionellen Trajektoren beschrieben. Zu diesem Zweck habe ich unser Zellkulturmodell zu den wichtigsten Entwicklungsstadien validiert. Die multipotente Fähigkeit der NKX2-1+ Vorläuferzellen konnte durch die Erzeugung von Lungenorganoiden nachgewiesen werden, welche transkriptomische Merkmale mit der primären fötalen Lunge teilen. Meine Arbeit hat gezeigt, dass die Aktivierung von Sonic Hedgehog, TGF-β und Notch-Signalwegen in einer FOXA2-ISL1-NKX2-1-Abfolge erforderlich ist, um respiratorische Progenitoren während des Entodermstadiums des Vorderdarms hervorzubringen. Auch markiert die Induktion des Transkriptionsfaktors HHEX ein alternatives Zellschicksal im frühen definitiven Entoderm, die der Lungenspezifikation vorausgeht und eine große Hepatoblastenpopulation hervorbringt. Die hohe zeitliche Auflösung der Einzelzell-Trajektoranalyse ermöglichte es, ein hierarchisches Modell der

Genexpressionsveränderungen während der Differenzierung in Richtung zweier Zellinien aus hPSC zu konstruieren: Lunge und Leber.

Insgesamt skizziert meine Arbeit zum ersten Mal den umfassenden Transkriptionsfahrplan, der der Spezifizierung der menschlichen embryonalen Lunge zugrunde liegt und liefert grundlegende Erkenntnisse für die Verbesserung von Strategien zur Lungendifferenzierung, die ihre Anwendung in der regenerativen Medizin und der Modellierung von Krankheiten mit entwicklungsbedingtem Ursprung finden sollten.

ACKNOWLEDGEMENTS

Throughout my Ph.D. study I have received great support and assistance. It is a great pleasure to thank those who made the successful completion of this degree possible.

I would first like to sincerely thank my supervisor Dr. Micha Drukker for giving me the chance to carry out this project in ISF-P group and for the guidance and the freedom he offered me in order to explore my scientific curiosity and pursue my project.

I am profoundly thankful to my thesis committee members, Prof. Dr. Heiko Lickert, Prof. Dr. Wolfgang Wurst, and Dr. Darcy Wagner for their immense support, advice, and valuable feedback throughout my project. I would also like to thank the members from CPC Research School for welcoming me in Munich and making the Ph.D. program so constructive and enjoyable.

Special thanks to my colleagues Dr. Maria Forstner, Dr. Friederike Matheus, Dr. Markus Grosch, Dr. Anna Pertek, Dr. Dmitry Shaposhnikov, Sebastian Ittermann, Polyxeni Nteli for the stimulating discussions, the great working atmosphere and for all the fun we have had in the last years. I would also like to acknowledge Karin Landgraf for supporting my project as Master student. Moreover, I wish to express my deepest gratitude to Dr. Fatma Uzbas for the unfailing support, encouragement and invaluable contribution throughout this dissertation, to Ejona Rusha for the technical assistance and the insightful suggestions; and to Christina Koupourtidou for the scientific and moral support whenever needed and above all our friendship. My sincere thanks go to Karen Biniossek for taking care of all organizational matters and for the administrative support.

I would also like to gratefully acknowledge my research partners Dr. Ilias Angelidis and Meshal Ansari for our wonderful collaboration and for making possible the scRNA-sequencing experiments and analysis, respectively. Their expertise was invaluable in addressing my research questions and methodology. Meshal Ansari performed the bioinformatic processing of the scRNA-sequencing data, which played a pivotal role in the completion of the project.

Getting though my Ph.D. required more than academic guidance. Many thanks to my friends inside and outside Munich, for the fun we had together, the emotional support and friendship.

My research would not have been possible without the persistent encouragement, unparalleled love and unabated support from my family. I am grateful to my brother Jiannis for his eternal friendship and spiritual support in all steps of my life. My deep and heartfelt gratitude goes to my parents, Christos and Vasiliki for offering me the opportunities and experiences that have made me who I am. I am extremely grateful to my patient and supportive fiancé Christos for the understanding and encouragement along the journey.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

3D : three-dimensional

AATD : α1-antitrypsin deficiency

ALI : Air Liquid Interface

A-P : Anterior-Posterior

ATI : alveolar type I

ATII : alveolar type II

BADJs : bronchioalveolar duct junctions

BASCs : bronchioalveolar stem cells

BM : Basal Media

BR : BMP4, RA

BSA : Bovine Serum Albumin

CDX2 : caudal-type homeobox-2

CF : Cystic Fibrosis

CGRP : calcitonin gene-related peptide

CKF : CHIR99021, KGF, FGF-10

COPD : chronic obstructive pulmonary disease

CPM : Carboxypeptidase M

DCI : dexamethasone, 8-bromo-cAMP, isobutylmethylxanthine

DCICS : DCI, CHIR99021, SB431542

DE : Definitive Endoderm

dox : doxycycline

DSM : Dorsomorphin

EB : embryoid body

ECM : extracellular matrix

EMT : epithelial-to-mesenchymal transition

ESC(s) : Embryonic Stem Cell(s)

FBS : Fetal Bovine Serum

FE : Foregut Endoderm

FGF : Fibroblast Growth Factor

FOXJ1 : Forkhead box J1

GFR : growth factor reduced

GO : gene ontology

h : hour(s)

HB : hepatoblasts

HCA : Human Cell Atlas

hESCs : human embryonic stem cells

hiPSCs : human induced pluripotent stem cells

HPS : Hermansky–Pudlak syndrome

IBMX : isobutylmethylxanthine

ICM : inner cell mass

iPSCs : induced pluripotent stem cells

KRT14 : cytokeratin 14

KRT5 : cytokeratin 5

LB : Luria Bertani

LNEP : lineage-negative epithelial progenitor cells

LO : Lung Organoids

LP(s) : Lung Progenitor(s)

mESCs : mouse embryonic stem cells

min : minutes

MTG : α-Monothioglycerol

NE : Neuroendocrine cell

NEBs : Neuroendocrine Bodies

NGFR : Nerve Growth Factor

NGS : Normal Goat Serum

O.C.T. : Tissue-Tek® O.C.T.™ Compound

PAGA : Partition-based Graph Abstraction

PB : PiggyBac

pcw : post conception week

P-D : Proximal-Distal

PDPN : Podoplanin

PFA : paraformaldehyde

PI : Propidium Iodide

PSC(s) : Pluripotent Stem Cell(s)

RA : retinoic acid

RAGE : advanced glycosylation end product–specific receptor

RQ : Relative Quantification

SD : Standard Deviation

SB : sytox blue

SCGB1A1 : secretoglobin family 1A member 1

sc-RNA seq : single cell-RNA sequencing

SFTPB : Surfactant protein B

SFTPC : Surfactant protein C

SHH : Sonic hedgehog

SSEA-3/4 : Stage-specific embryonic antigen-3/4

T : Brachyury

TE : Trophoectoderm

TF(s) : Transcription factor(s)

TRA-1-60/1-81 : Tumor rejections antigen-1-60/1-81

UMAP : Uniform Manifold Approximation and Projection

WHO : World Health Organization

1. INTRODUCTION

1.1. THE RESPIRATORY SYSTEM

The respiratory system consists of specific structures and organs and is responsible for the exchange of oxygen and carbon dioxide between the external environment and the circulating blood. In human, it is structurally divided into the upper respiratory tract, which encompasses the nasal cavity, pharynx and larynx, and the lower respiratory tract consisting of the trachea, the bronchi, the bronchioles, and the alveoli, which make up the lungs. The tracheobronchial or respiratory tree is formed beyond the trachea, as a result of extensive branching morphogenesis. The trachea bifurcates into the right and left main bronchi, which in turn subdivide into the secondary (lobar) bronchi. Further rounds of branching form the tertiary bronchi and then the bronchioles, which progressively become smaller and more numerous, leading to terminal bronchioles. Terminal bronchioles branch further forming the respiratory bronchioles that lead to alveolar ducts and the alveoli (**Figure 1**). The airways from nose to terminal bronchioles constitute a conducting zone, which transmits gases into and out of the lungs, as well as filters, warms, and moistens the incoming air. The bronchioles terminate into the respiratory zone (alveoli), which includes the structures that are involved directly in the gas exchange process (Amador et al., 2022; Craft et al., 2019) (**Figure 1**).

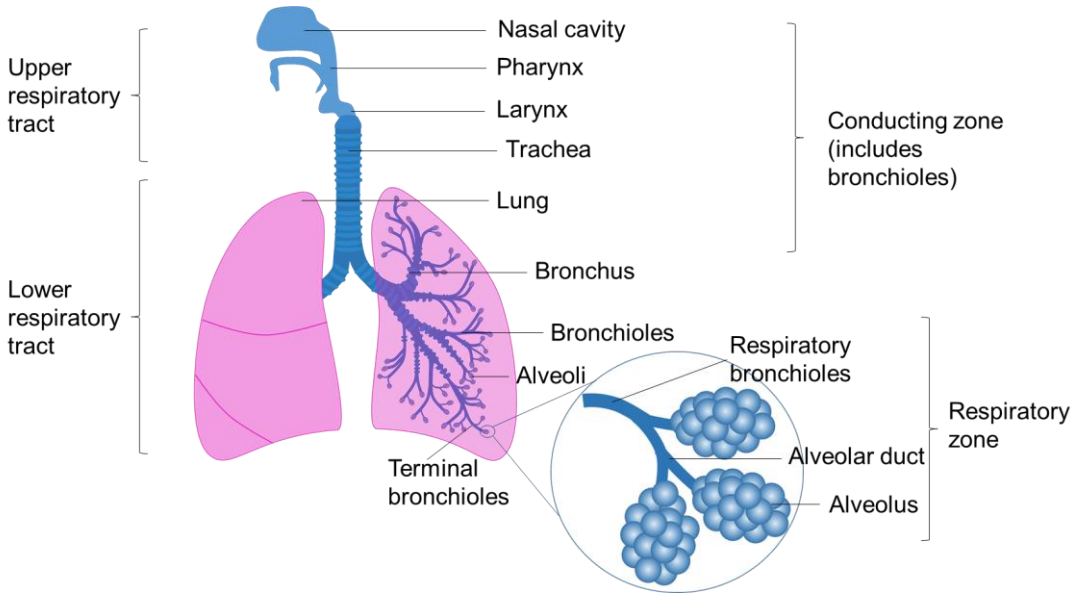


Figure 1. Description of the respiratory system. Structural and functional organization of the respiratory system. The upper and lower respiratory tract comprise the conducting and respiratory zone, respectively.

Preservation of the epithelial cell types lining each of the anatomical airway compartments is necessary to maintain their functions throughout life and is carried out by diverse pools of regional stem and progenitor cells that exist in specific niches across the pulmonary epithelial tree. To achieve such a complexity, the development of the lung is a precise and progressive progress, initiating at the early embryonic development in the endoderm epithelium coupled with several events of proximal-distal patterning that are strongly dependent on cell-to-cell interactions and events of branching morphogenesis. Defects in any of these processes can harm the function of the respiratory system and eventually lead to respiratory diseases (Hogan et al., 2014).

1.1.1. Respiratory epithelium composition

The mature lung is lined by epithelial cells, the composition and functional specialization of which varies along the proximal-distal (P-D) axis to fulfill the regional functionalities and maintain organ integrity. Thus, region-specific epithelial progenitor cell populations that reside in distinct niches exist to maintain homeostasis and can also display lineage plasticity in response to injury (Hogan et al., 2014).

In human, airways are lined by a pseudostratified epithelium composed of basal and luminal cells, including ciliated cells, secretory goblet and club (Clara) cells, and neuroendocrine cells (**Figure 2**). Basal cells extend throughout the airways and are characterized by the expression of tumor protein 63 (TP63), cytokeratin 5 and 14 (KRT5, KRT14), nerve growth factor receptor (NGFR) and podoplanin (PDPN, also known as T1a) (Daniely et al., 2004; Rock et al., 2009; Schoch et al., 2004; Tata et al., 2013). Studies in mice and human lungs showed that basal cells can self-renew and function as multipotent progenitors that generate secretory and ciliated cells during steady state and upon injury (Hong et al., 2004; Rock et al., 2009; Teixeira et al., 2013). Secretory cells define a more heterogeneous cell population. Goblet cells produce mucus glycoprotein mucin 5AC (MUC5AC) (Hovenberg et al., 1996); and club cells are marked by the expression of the secretoglobin 1A1 (SCGB1A1, also known as CCSP or CC10), and the surfactant protein-B (SFTPB), subtypes of which exist along the axis of the airway (Phelps and Floros, 1991; Singh et al., 1988). In particular, a subset of secretory club cells that exhibits naphthalene-resistance in murine models and are positive for SCGB1A1 and negative for cytochrome p450, termed variant club cells, have been identified as progenitor cells that can re-populate damaged epithelium by giving rise to new club and ciliated cells (Hong et al., 2001). Another subpopulation of club cells, named

bronchioalveolar stem cells (BASCs), exist at the border of the conducting and respiratory zones, are positive for SCGB1A1 and surfactant protein-C (SFTPC), exhibit stem cell characteristics and differentiate into club and alveolar epithelial cells (Kim et al., 2005). The locations of variant club cells around clusters of neuroendocrine cells, referred as neuroendocrine bodies (NEBs) (Reynolds et al., 2000), and of both variant club cells and BASCs at the bronchioalveolar duct junctions (BADJs) (Giangreco et al., 2002) indicate the existence of micro-environments that regulate the niches of progenitor cells in the airways.

Ciliated cells are characterized by the presence of multi-cilia on their apical surface that contribute to the mucociliary clearance of the airways, and by the expression of the forkhead box J1 (FOXJ1) transcription factor (Tichelaar et al., 1999). Neuroendocrine (NE) cells are marked by the expression of calcitonin gene-related peptide (CGRP) (Song et al., 2012). Lineage tracing studies have shown that also NE cells can self-renew under steady conditions and can differentiate into club and ciliated cells only upon injury (Song et al., 2012).

Alveolar epithelium is lined by cuboidal surfactant-producing alveolar type II cells (ATII) and squamous gas exchanging alveolar type I cells (ATI) (**Figure 2**). ATI cells cover around 95% of the alveoli surface area and are marked by the expression of aquaporin 5 (AQP5), PDPN, as well as advanced glycation end product receptor (RAGE) (Kreda et al., 2001; Shirasawa et al., 2004). ATI cells create a thin interface with the underlying vascular endothelium, allowing the gas exchange. ATII cells, on the other hand, produce the pulmonary surfactant to reduce surface tension and prevent collapse of the alveoli upon expiration. SFTPC is the only surfactant protein, which is exclusively expressed in ATII cells of the adult human lung; thereby it is a marker for the identification of ATII cells. Notably, ATII cells serve as progenitors that can proliferate and give rise to ATI cells, maintaining the homeostasis of the alveolar epithelium during the steady state and upon injury (Barkauskas et al., 2013; Driscoll et al., 1995).

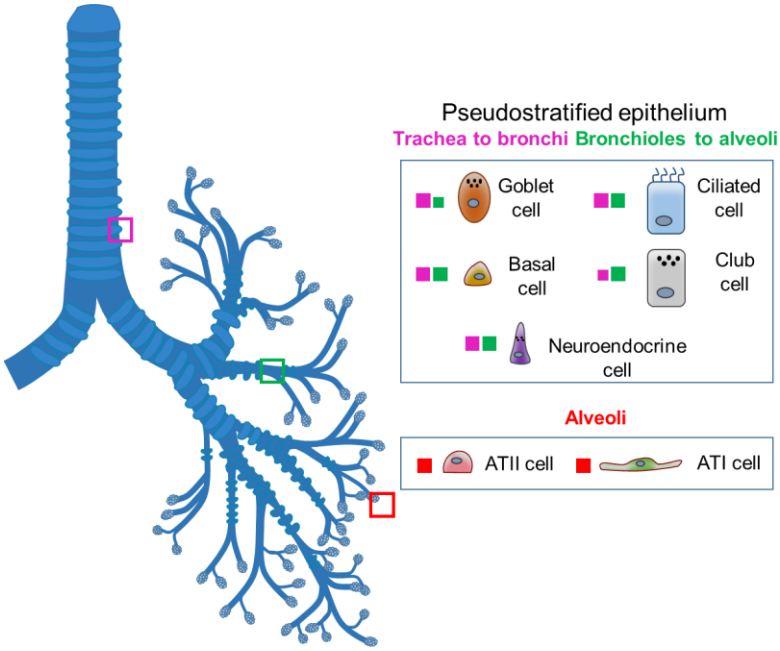


Figure 2. Epithelium composition of human lung. Schematic illustration of the major epithelial cell types in different regions of the human lung.

1.1.2. Lung development - endoderm formation and pattering

Lung development initiates early in human embryonic development with the establishment of respiratory cell fate in the primitive gut tube and continues after birth through a process defined by multiple developmental stages. These stages rely on reciprocal interactions between the endoderm and the surrounding mesoderm (Kimura et al., 1996; Lazzaro et al., 1991; Minoo et al., 1999; Morrisey and Hogan, 2010).

1.1.2.1. Formation of primitive gut tube

Human embryonic development begins with the fusion of sperm and oocyte during a process known as fertilization (conception), resulting in the formation of zygote. After series of mitotic cell divisions, the morula - a solid sphere of cells, is being formed and developed further to the blastocyst. The first distinct types of cells begin to be recognizable, as the first overt differentiation event has occurred, the trophectoderm (TE), a surrounding cell layer which gives rise to the placenta, and the inner cell mass (ICM) which will give rise to the embryo proper (epiblast) and extra-embryonic tissues (hypoblast) (Gilbert, 2000).

Following implantation of the blastocyst into the uterine wall, gastrulation takes place in the 3rd week of development, a process by which epiblast cells are allocated to the three germ layers – endoderm, mesoderm and ectoderm. Gastrulation initiates with the formation of the primitive streak (PS) on the posterior side of the embryo and involves sets of morphogenetic movements coupled with cell proliferation and differentiation, revealing a morphologically obvious anterior-posterior (A-P) axis. Pluripotent epiblast cells ingress through the primitive streak and give rise to mesoderm and definitive endoderm (DE) (Hashimoto and Nakatsuji, 1989), whereas the remaining epiblast cells become ectoderm (Bénazéraf and Pourquié, 2013). Increasing evidence suggests that similar to mouse (Lickert et al., 2002; Tada, 2005), human DE and mesoderm might arise from a common transient and bipotential precursor in the primitive streak named mesendoderm (D'Amour et al., 2005; Tada, 2005). Once the three germ layers are established, cells interact and relocate to produce tissues and organs in a process named organogenesis (Gilbert, 2000) (**Figure 3**).

Following gastrulation, complex morphogenetic movements lead to the transformation of definitive endoderm into the primitive gut tube that is patterned along the A-P axis into the foregut, midgut and hindgut domains (**Figure 3**), driven by reciprocal epithelial-mesenchymal signaling interactions (Zorn and Wells, 2009). The initial patterning of the gut endoderm takes place during late gastrulation, as the sheet of newly induced endoderm starts to shape a gut tube (Carlson, 2014). Then, restricted expression patterns of transcription factors (TFs) in the foregut, midgut, and hindgut define organ-specific domains. During organogenesis, primordial organ buds arise as outgrowths from these domains, which ultimately differentiate into functional organs **(Figure 3)**. The foregut develops into thyroid, liver, biliary system, lungs, esophagus, trachea, stomach, pancreas, whereas the midgut and hindgut form the small and large intestines, respectively (Zorn and Wells, 2009) (**Figure 3**).

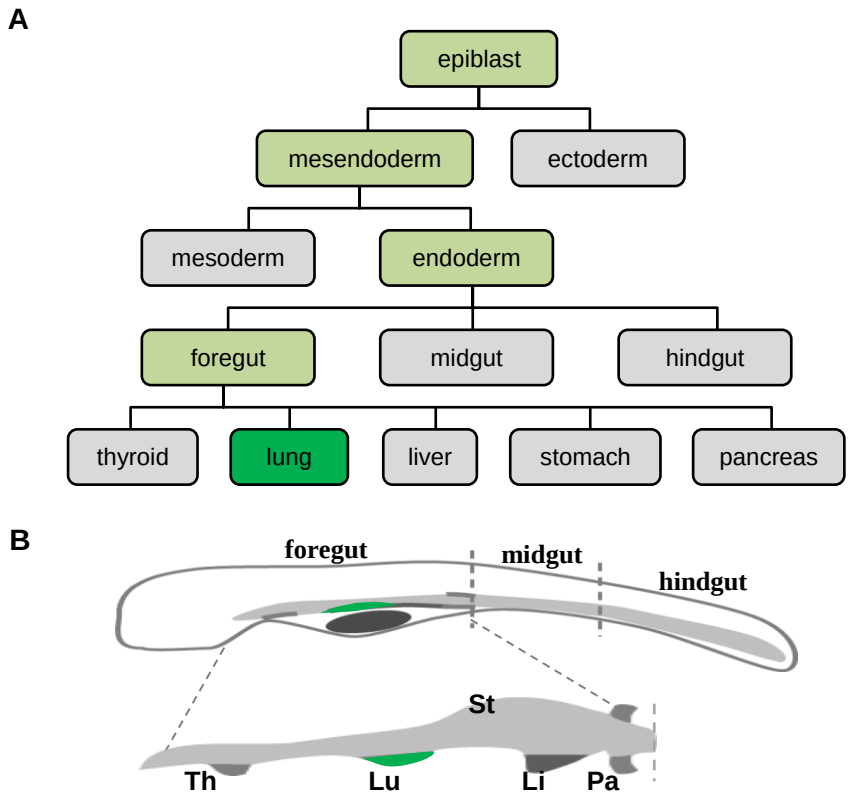


Figure 3. Overview of lung specification process. (A) The major events in endoderm organ specification. **(B)** Schematic illustration of the primitive gut tube. **Th**: thyroid, **Lu**: lung, **St**: stomach, **Li**: liver, **Pa**: pancreas

1.1.2.2. Lung organogenesis

Human lung development is divided into five stages defined mainly by morphological criteria: embryonic, pseudoglandular, canalicular, saccular and alveolar stage (Burri, 1984; Schittny, 2017). All phases of lung development overlap and do not have sharp borders.

During the **embryonic stage** (4-7 pcw) the primary left and right lung buds arise from the ventral side of the anterior foregut endoderm. They branch off from the trachea, invade the adjacent mesoderm, and elongate (Brauer, 2003). Repetitive events of stereotypic and asymmetrical branching initiate to generate the arborised respiratory tree (Warburton et al., 2000).

The **pseudoglandular stage** (5-17 pcw) is defined by continuous rounds of branching morphogenesis that leads to the formation of the conducting airways up to the terminal bronchioles. The airway tree is laid down and smooth muscle and vasculature develop. As the lung epithelium

grows out, differentiation of the respiratory epithelium initiates with the first ciliated, goblet and basal cells appearing in the proximal airways (Schittny, 2017).

The **canalicular stage** (16-26 pcw) is characterized by the morphological distinction between conducting and respiratory airways, as the basic gas-exchange portion is formed along with its vascularization. In particular, the most distal epithelial tubes widen into the airspaces leading to a canalization of the surrounding mesenchyme. An air-blood (alveolar-capillary) barrier develops as the mesenchymal capillary networks come into close proximity to the distal epithelium and first events of alveolar cell differentiation take place (Schittny, 2017). Yet, it is not clear whether ATI and ATII lineages arise from a bipotent progenitor or if alveolar cells are specified in a sequential manner with the transdifferentiation of ATII to ATI cells. Studies on alveolar cell specification at single-cell level identified a population of bipotent alveolar progenitors that co-express ATI and ATII markers (Desai et al., 2014; Treutlein et al., 2014). However, more recently, lineage tracing studies in mice demonstrated that those bipotent progenitors exist in mice but are rather rare and possibly constitute residual undifferentiated cells that do not contribute essentially to the mature alveolar epithelium (Frank et al., 2019).

At **saccular stage** (24-38 pcw) branching morphogenesis ceases and terminal airways are widening and growing in length. At the ends of the airways, increased numbers of thin-walled saccules (primitive alveoli) are clustered, enlarge and become covered by a capillary bilayer. Alveolar epithelial differentiation continues while maturation of the surfactant system occurs and the surfactant-producing lamellar bodies become detectable in differentiating ATII cells (Khoor et al., 1994; Schittny, 2017).

During postnatal or **alveolar stage** (36 pcw – young adulthood) the septation of the distal saccules leads to the formation of mature alveoli and to a dramatic increase of the gas exchange surface area (Burri, 1984). Although alveolar development was considered to happen until 3 years of age, advances in imaging technologies estimate that alveolar formation likely continues into later childhood and even adolescence (Narayanan et al., 2012). At the end of the branching morphogenesis, the human respiratory tree comprises 23 generations of vascularized branching conducting airways that serve to conduct air to and from the gas-exchanging epithelium of the alveolar bed (Weibel, 1963).

1.1.3. Molecular basis of lung development

Lung development has been extensively studied in recent years. Most of our current understanding has been derived from animal models, and little has been validated with respect to the molecular regulation of human fetal lung development, especially in the very early and late developmental stages, during which the lung tissue is inaccessible to studies. Nevertheless, recent advances in molecular techniques enabled more detailed characterizations of the human tissue. Notably, combination of new technologies with collaborative efforts of tissue banks and repositories allowed an international exertion to map all human cells, named Human Cell Atlas (HCA) (Regev et al., 2018), part of which is the Human Developmental Cell Atlas (Behjati et al., 2018), and the LungMap consortium that focuses specifically on the developing lung (Ardini-Poleske et al., 2017).

Series of morphological changes and cell specification events take place upon fertilization. In the early blastocyst, the exclusive expression pattern of the transcription factors POU domain, class 5 transcription factor 1 (*Pou5f1*, also known as *Oct4*) and caudal-type homeobox-2 (*Cdx2*), mediate the cell fate decision to ICM and TE, respectively (Niwa et al., 2005). The involvement of Hippo pathway has been shown to regulate this early segregation process (Nishioka et al., 2009). The second cell fate decision results in the segregation of the epiblast from the primitive endoderm (hypoblast) in the ICM, a process that was shown to rely on the activity of FGF signaling (Arman et al., 1998; Chazaud et al., 2006; Feldman et al., 1995). Lineage-tracing experiments have shown that ICM of early blastocyst contains already distinct populations of cells that express the transcription factors *Nanog* and *Gata6* in a "salt and pepper" pattern and are committed to develop into epiblast or primitive endoderm, respectively (Chazaud et al., 2006). The resulted pluripotent epiblast cells express *Nanog*, *Oct4*, and *Sox2* (Avilion et al., 2003; Hart et al., 2004; Yeom et al., 1996), while the primitive endoderm is characterized by the gene expression of *Gata4*, *Gata6*, *Sox17* and *Sox7* (Artus et al., 2011). Although pluripotent cells represent a transient state in the early embryo, cultures of human pluripotent embryonic stem cell lines can be established from the ICM (Thomson, 1998), and propagated *in vitro* under specific pathway modulators that allow them to indefinitely self-renew and maintain their potential to differentiate (Smith, 2001). The maintenance of pluripotency requires the activity of Nodal, a TGF-β related factor (Vallier et al., 2009).

As gastrulation proceeds, the primitive streak is formed in the posterior epiblast, marked by the expression of the transcription factor *Brachyury* (*T*) (Herrmann, 1991; Wilkinson et al., 1990). Genetic studies have demonstrated that Nodal and Wnt3 are essential for the formation of primitive

streak (Conlon et al., 1994; Liu et al., 1999). In the anterior part of primitive streak resides a mesendodermal population and is characterized by the expression of *T*, *Gsc*, *Foxa2*, *Eomes* and *Mixl1* (Kinder et al., 2001; Robb and Tam, 2004). Nodal was shown to have a dose-dependent effect on the mesendoderm specification, i.e. higher levels are required for definitive endoderm commitment whereas low levels promote mesoderm induction (Zorn and Wells, 2009). Furthermore, Activin, a member of TGF-β superfamily, acts similarly to Nodal, as they share a common downstream signaling cascade, binding to the same receptors (Gamer and Wright, 1995; Gray et al., 2003; Henry et al., 1996; Smith et al., 1990; Thomsen et al., 1990); thus, activation of Activin/Nodal signaling is broadly used in the differentiation protocols of human and mouse embryonic stem cells (hESCs and mESCs, respectively) into endoderm (D'Amour et al., 2005; Kubo, 2004).

Morphogenetic movements of the endoderm lead to the formation of the primitive gut tube surrounded by mesoderm. A complex interplay of signaling pathways mainly from the surrounding mesoderm pattern the endoderm along the A-P axis into domains ultimately, specifying cell fate. The regionalization of the developing gut tube is indicated by the reemergence of *Sox2* expression in the foregut domain together with *Hhex* and *Foxa2*, and the expression of *Cdx2* posteriorly at a broad hindgut, followed by the expression of *Pdx1* at the posterior foregut/midgut region. Subsequently, distinct expression of TFs along the A-P and D-V axis subdivides the foregut into organ specific regions (Zorn and Wells, 2009).

The earliest known marker of the lung primordia is the homeodomain protein NKX2-1 (Kimura et al., 1996; Minoo et al., 1999). The transcription factor Nkx2-1 is expressed in a subset of progenitors on the ventral side of anterior foregut endoderm, the presumptive lung endoderm, where *Sox2* expression is downregulated, introducing a D-V patterning along the gut tube (Que et al., 2007). Mice with Nkx2-1 mutation seem to partially preserve features of proximal differentiation. In particular, they have abnormal lungs with two main bronchi that develop to cystic structures lined by columnar cells with scattered cilia, and the epithelium does not express any of the surfactant-protein genes that are normally found in the distal lung (Minoo et al., 1999).

Genetic studies in animal models have shown that among the signaling pathways involved in the establishment of the A-P identity of the gut tube retinoic acid (RA), Wnt/β-catenin, FGF, BMP and sonic hedgehog (SHH) are of critical importance. Notably, these signaling pathways often demonstrate spatio- temporal- and gradient- dependent effects during patterning of the gut endoderm,

where the molecules may perform opposite functions in different time points or concentrations, thus specifying each region consecutively.

In particular, RA gradient was shown to pattern the gut tube, whereby the anterior part is exposed to less RA than the posterior, inducing different transcription factors in different regions along the tube (Bayha et al., 2009). Also, disruption of RA signaling affects the foregut compartmentalization and leads to failure of oesophagotracheal septum formation and lung agenesis (Chen et al., 2007; Desai et al., 2006; Mollard et al., 2000; Wang et al., 2006). RA has been shown to regulate *Fgf10* expression in the mesoderm surrounding the lung field, which in turn modulates the lung bud outgrowths (Desai et al., 2004).

Wnt/β-catenin signaling is also involved in foregut specification. Initially, it promotes posterior endoderm (Sherwood et al., 2011), and later it is crucial for the anterior endoderm specification and for the D-V patterning of anterior foregut. Wnt/β-catenin promotes expression of *Nkx2-1* in the ventral anterior foregut endoderm, and restriction of the *Sox2* expression. In particular, this has been shown by disruption of Wnt during that critical time window, which resulted in the absence of *Nkx2-1* expression and expansion of *Sox2* expression, leading to abnormal tracheal formation, and lung agenesis. Conversely, forced activation of Wnt/β-catenin signaling resulted in ectopic generation of NKX2-1+ progenitors in the posterior gut endoderm (Goss et al., 2009; Harris-Johnson et al., 2009).

BMP signaling has a spatiotemporal role in the foregut patterning as well. More specifically, *Bmp4* is produced by the ventral mesenchyme (Weaver et al., 1999) and acts through the BMP receptors *Bmpr1a*, *b* in the ventral endoderm. In absence of Bmp4 or its receptors, respiratory specification does not take place due to the expanded expression of *Sox2* in the ventral endoderm at the expense of *Nkx2-1* (Domyan et al., 2011). In addition, the BMP antagonist, Noggin is being produced by the dorsal foregut mesenchyme, suppressing BMP signaling and allowing the expression of *Sox2* (Domyan et al., 2011; Que et al., 2006).

Fibroblast growth factor signaling pathway (FGF) plays an important role in defining specific regions along the A-P axis of the developing endoderm. Different thresholds of FGFs pattern the foregut endoderm into different derivatives. Concretely, FGF4 initially promotes posterior endoderm cell fate, in a concentration dependent manner. FGF4 inhibits the expression of *Nkx2-1* and disrupts foregut morphogenesis (Dessimoz et al., 2006). Then, gradients of FGF signaling from the cardiac mesoderm induce further patterning within the foregut domain. For instance, studies

using ventral foregut endoderm explants have demonstrated that higher levels of FGF2 promote *Nkx2-1* expression while moderate levels induce the expression of albumin (*Alb*) which is an hepatic marker (Serls, 2004; Wang et al., 2011). Another example is FGF10 directing the branching morphogenesis of the lung through the fibroblast growth factor receptor 2b (Fgfr2b) (Bellusci et al., 1997b; Ohuchi et al., 2000; Sekine et al., 1999). Genetic ablation of either *Fgf10* or *Fgfr2b* in mice leads to lung agenesis and abrogation of branching, although lung specification and the separation of trachea from the esophagus take place (De Moerlooze et al., 2000; Sekine et al., 1999).

Sonic hedgehog was also shown to be critical for lung development. Mice lacking Shh or related signaling components, such as *Gli2/3* double-null mice, demonstrate defects in the establishment of the tracheoesophageal septum, as well as in the development of the lung and the epithelial branching (Litingtung et al., 1998; Motoyama et al., 1998). It has been suggested that SHH establishes molecular boundaries within the fore-midgut domain by inhibiting pancreas formation (Deutsch et al., 2001; Hebrok et al., 1998; Kim and Melton, 1998). However, SHH signaling in later lung development acts in a contradictory way, as during the lung expansion its role is to downregulate *Fgf10* expression (Bellusci et al., 1997a).

Upon respiratory specification, the primary lung buds extend into the mesenchyme, while lung epithelial cells receive signals from the surrounding mesenchyme which pattern the lung in a P-D manner and regulate branching morphogenesis, leading to the formation of the respiratory tree. The localized expression of *Fgf10* in the mesenchyme, adjacent to the distal tip endoderm is the main driver of that process, which triggers secondary and subsequent budding, acting through Fgfr2b in the endoderm (De Moerlooze et al., 2000; Ohuchi et al., 2000; Sekine et al., 1999). Crosstalk between FGF10 and other signaling pathways including Wnts, SHH and BMP4 is essential for the regulation of the progenitors and the determination of the ultimate size and shape of the lung. In a feedback response, the distal epithelial tips express Shh which acts by downregulating *Fgf10* expression as the lung develops, limiting the bud outgrowth (Bellusci et al., 1997b). However, SHH also induces hedgehog-interacting protein (Hhip), which in turn releases the repression of *Fgf10* by inhibiting SHH signaling (Chuang et al., 2003). During branching, sprouty (SPRY) proteins are induced in the distal epithelium and SPRY2 acts also by inhibiting FGF-mediated budding (Mailleux et al., 2001; Tefft et al., 1999). The effect of BMP signaling seems to depend on whether BMP signaling is activated via an autocrine or a paracrine mechanism, where BMP4 signaling by the mesoderm (paracrine) stimulates lung bud branching, whereas autocrine activation limits the *Fgf10*-

mediated budding (Cardoso, 2006). RA is active during the primary lung bud formation, but further airway branching requires downregulation of RA pathways (Malpel et al., 2000). Although the role of Wnt signaling in the process of branching morphogenesis is still controversial, Wnt seems to be important for the subsequent P-D patterning by promoting distal lung endoderm (McCauley et al., 2017; Mucenski et al., 2003).

As the lung develops, the P-D patterning of the endoderm becomes apparent. The leading distal tips of the primary lung buds co-express *Sox9* and *Id2*, and initially act as a pool of multipotent progenitors that generates the bronchial and alveolar epithelium. At later stage, they generate only alveolar epithelial lineages, including alveolar epithelial type I and alveolar epithelial type II cells (Rawlins et al., 2009). In mice, it has been shown that during the pseudoglandular stage, cells that exit the distal tips downregulate *Sox9* expression and upregulate *Sox2* in a process that generates the proximal lineages. In human embryos of the respective stage this process seems to be different as *SOX2* and *SOX9* are co-expressed in the tips of the developing lung throughout the pseudoglandular stage (Miller et al., 2018; Nikolić et al., 2017). Additionally to Sox9/Id2 other transcription factors mark the distal progenitors, including Foxp1/2, N-myc, Etv4/5, and members of the Iroquois protein family (Irx1, Irx2, Irx3) (Rawlins et al., 2009; Swarr and Morrisey, 2015). Notch signaling is critical for balancing the differentiation of secretory cells and ciliated cells within the developing airways (Guseh et al., 2009). Loss of Notch signaling causes a defect in secretory club cell differentiation, while the ciliated lineage expands to populate the proximal airways (Tsao et al., 2009). Notch has also be shown to be involved in alveolar morphogenesis (Tsao et al., 2016). Furthermore, Yap an effector of Hippo signaling pathway seems to regulate P-D patterning, control *Sox2* expression and is required to promote airway epithelial cell fate at the proximal lung (Mahoney et al., 2014).

1.2. STEM CELL APPICATIONS IN LUNG REPAIR

During homeostasis, the cellular turnover of the lung is very low. However, upon injury a specialized group of endogenous lung progenitor cells can become activated and replenish the damaged respiratory epithelium maintaining lung homeostasis and regeneration. Disruption of this process leads to fibrosis which can be the cause of several lung diseases (Wansleeben et al., 2013). Thus, comprehending the mechanisms that govern the regenerative capacity of the lung paves the way for the development of novel therapeutic targets.

1.2.1. Developmental pathways in lung regeneration

Numerous developmental pathways have been shown to display critical regulatory function in lung regeneration. Wnt/β-catenin signaling, for instance, one of the important regulatory pathways in lung endoderm specification, is re-activated in the adult lung upon injury and during repair and regeneration (Al Alam et al., 2011; Flozak et al., 2010). The role of Notch signaling pathway in stimulating secretory cell differentiation during development is also recapitulated after injury, in the response of basal cells (Xing et al., 2012). In addition, Hippo pathway seems to also play an important role in triggering lung regeneration. Specifically, Yap, a downstream effector of Hippo pathway has been shown to control stem cell proliferation and differentiation during homeostasis and repair (Lange et al., 2015; Zhao et al., 2014). Thus, the investigation of lung development has pivotal implications for understanding and trying to manipulate the remodeling process and regeneration of the adult organ.

1.2.2. Lung disorders

The lung is the most vulnerable internal organ to airborne infections and injury due to its exposure to particles, chemicals, and infectious organisms from the external environment. In fact, common lung diseases such as chronic obstructive pulmonary disease (COPD), lower respiratory infections, and lung cancers are among the top ten leading causes of death globally (World Health Organization, 2020) (**Figure 4**). Furthermore, genetic conditions, such as α1-antitrypsin deficiency (AATD) and cystic fibrosis (CF), are common hereditary disorders affecting the lung. As of today, the therapeutic options for several lung diseases provide no cure and are limited to symptom reduction and delay of the disease progression. At end-stage lung disease, transplantation is the only option (Arcasoy and Kotloff, 1999). However, the donor organ shortage (Van Raemdonck et al., 2009), as well as the

poor survival rates due to both acute and chronic rejection upon transplantation (Yusen et al., 2015) make necessary the development of innovative therapies. Advances in stem cell treatments are promising with the potential to reverse or cure some lung diseases.

Leading causes of death globally in 2019	
1	Ischaemic heart disease
2	Stroke
3	Chronic obstructive pulmonary disease
4	Lower respiratory infections
5	Neonatal conditions
6	Trachea, bronchus, lung cancers
7	Alzheimer's disease and other dementias
8	Diarrhoeal diseases
9	Diabetes mellitus
10	Kidney diseases

Figure 4. Top 10 causes of death globally (WHO, 2020). Data was taken from WHO, https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death.

1.2.3. Pluripotent Stem cells

Pluripotent stem cells (PSCs) have the ability to self-renew and give rise to progenitors of all germ layers, and eventually to all cells of the adult body. Thereby, they provide an extremely powerful tool for understanding human organogenesis, as well as for developing applications in regenerative medicine, disease modeling and drug discovery. There are two types of PSCs; embryonic stem cells (ESCs) which are derived from the inner cell mass of preimplantation embryos (Evans and Kaufman, 1981; Martin, 1981; Thomson, 1998), and induced pluripotent stem cells (iPSCs) which refer to adult somatic cells that have been converted to an embryonic stem cell-like stage via a process called reprogramming (Takahashi and Yamanaka, 2006). Even though PSCs provide an additional source of cells for regenerative medicine through application of cell-replacement therapies, the use of ESCs faces obstacles due to ethical and legal barriers that result from the destruction of human embryos, as well as due to immunological rejection upon allogeneic cell transplantation. Thus, the breakthrough of Takahashi and Yamanaka, who identified a combination of four transcription factors Oct4, Sox2, Klf4, and c-Myc that were sufficient to reprogram first mouse and then human somatic cells into PSCs paved the way for the generation of

patient-specific iPSCs (Takahashi et al., 2007; Takahashi and Yamanaka, 2006) (**Figure 5**). Human PSCs are characterized by the expression of certain markers, including the surface markers stage-specific embryonic antigen-3 (SSEA-3) and -4 (SSEA-4) and tumor rejections antigen-1-60 (TRA-1-60) and -1-81 (TRA-1-81), as well as the transcription factors NANOG, OCT4 and SOX2 (Takahashi et al., 2007; Thomson, 1998).

An ultimate aim of lung differentiation protocols is to generate pure populations of progenitors with the ability to expand *in vitro* in yields that allow lung tissue engineering. However, further studies on the properties of human PSCs need to be performed in order to allow the comprehension of the basic biology of pluripotency and cellular differentiation, as well as overcome all the different issues linked to therapeutic applications. Furthermore, amelioration of the existing approaches has to be performed to achieve clinical-grade human PSCs for safe cell therapies (**Figure 5**). Application of adult stem and progenitor cells for the treatment of lung disorders represents a tempting outspread of the principles of developmental biology towards novel therapeutic approaches (**Figure 5**).

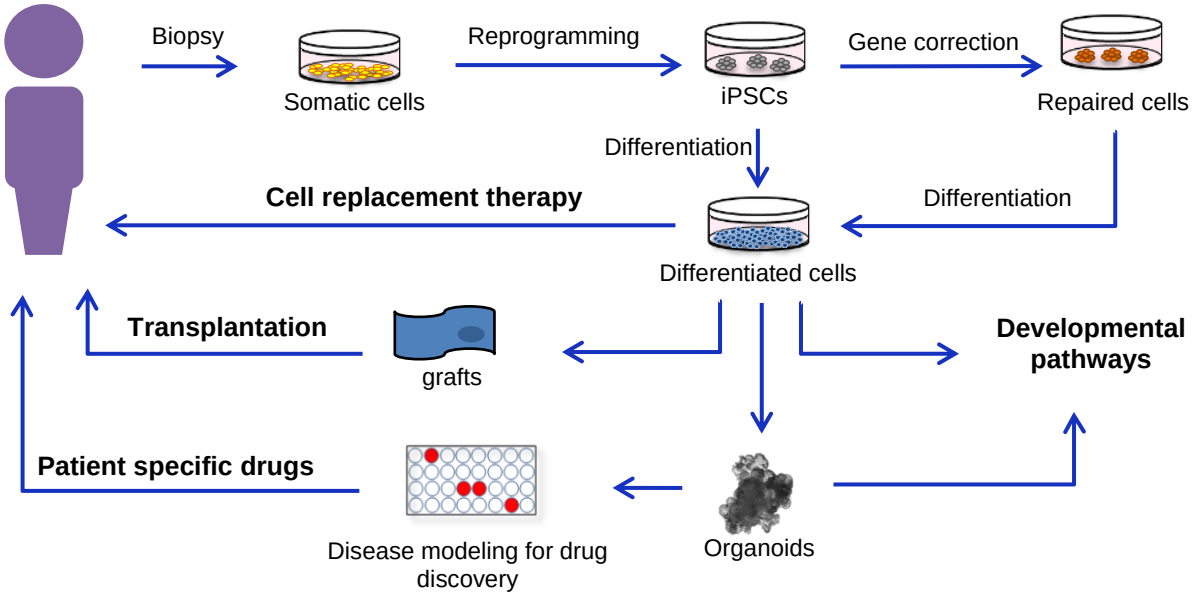


Figure 5. Applications of human Pluripotent Stem Cells. Patient specific hiPSCs, generated by reprogramming of somatic cells, can be differentiated into specialized cell types e.g. lung cells. hiPSC-derived specialized lung cells can be used for developmental studies, disease modeling, drug screening and discovery, as well as for cell/tissue replacement therapies upon (if required) genetic engineering, thereby improving further the existing therapies and/or enabling the development of novel therapy options.

1.2.3.1. Differentiation of hPSCs into lung epithelial progenitor cells

Directed differentiation of hPSCs aims to recapitulate processes of human embryonic development by consecutively applying timed signals to prompt specification and maturation of hPSCs into specific lineages. Establishment of such protocols has so far been almost exclusively developed based on the principles of developmental biology from animal models, due to the lack of available human data in such early stages. In fact, establishment of directed PSC differentiation protocols to lung epithelium relies on transitioning through progenitor states as they have been described in mouse.

The first major developmental stage in lung organogenesis is the formation of definitive endoderm. In accordance with the patterning of the primitive streak, directed differentiation regimens mimic definitive endoderm commitment. This was shown the first time by Kubo and colleagues in 2004, using mouse ESCs (mESCs), by inducing the Nodal signaling pathway using Activin-A (Kubo, 2004). A year later, D'Amour et al., 2005 successfully differentiated hESCs into endodermal cells that express *FOXA2*, *SOX17* and the DE-specific cell surface marker CXCR4.

With respect to embryonic development, the next step includes the specification of definitive endoderm to an anterior or posterior fate, similarly to the A-P patterning of the primitive gut tube. Green and colleagues in 2011 demonstrated that dual inhibition of TGF-β and BMP signaling pathways highly enriches anterior foregut endoderm (Green et al., 2011). This population was characterized by co-expression of the transcription factors *FOXA2, HHEX* and *SOX2* and absence of *CDX2*. Since then, the dual SMAD inhibition approach has been widely adopted (Firth et al., 2014; Gotoh et al., 2014; Mou et al., 2012). An alternative strategy was described by Wong et al. (2015) who exposed DE to SHH and FGF2, and later from Huang et al. (2014) and Dye et al. (2015) who applied additionally to the BMP/TGF-β inhibition, either simultaneous inhibition of WNT or activation of the WNT, SHH and FGF4 pathways, respectively.

The next major developmental milestone in lung differentiation protocols was the induction of NKX2-1+ lung progenitor cells that are able to differentiate further towards proximal or distal lung lineages. NKX2-1+ cells were generated upon application of factor combinations, mainly activating Wnt, BMP, FGF and RA signaling pathways (Gotoh et al., 2014; Green et al., 2011; Hawkins et al., 2017; Huang et al., 2014; Mou et al., 2012). Further progress has been achieved in the field with the specialization of iPSC-derived lung progenitor cells into functional bronchiolar or alveolar lineages (Dye et al., 2015; Gotoh et al., 2014; Hawkins et al., 2017), using approaches such as air liquid interface (ALI)-cultures (Firth et al., 2014; Konishi et al., 2016), by injecting cells into the flank or kidney capsule of immune-deficient mice (Y.-W. Chen et al., 2017; Green et al., 2011; Huang et al., 2014), seeding cells in decellularized matrices (Ott et al., 2010; Petersen et al., 2010) or generating lung organoids which mimic the 3D organization of lung tissue (Dye et al., 2015; Gotoh et al., 2014; Hawkins et al., 2017). More specifically, the treatment of lung progenitors with a mixture of small molecules dexamethasone, 8-bromo-cAMP, and isobutylmethylxanthine (IBMX) known as DCI, that is known to induce maturation of fetal primary ATII and club cells (Ballard et al., 1991; Berg et al., 2002; Gonzales et al., 2002; Oshika et al., 1998), is well established for the promotion of ATII cell differentiation *in vitro* (Katsirntaki et al., 2015). In addition, groups have accomplished to differentiate hPSCs to CFTR-expressing airway epithelium, paving the way for generation and use of cystic fibrosis (CF)-disease-specific iPS cells for potential drug screening and disease modeling applications (Crane et al., 2015; McCauley et al., 2017). Last, inhibition of Notch signaling pathway has been shown to benefit airway multi-ciliated epithelium generation from

hPSCs in ALI cell cultures (Firth et al., 2014; Konishi et al., 2016), strategy based on principles of mouse lung development as outlined above (Rock et al., 2011).

Despite the progress that has been made on lung differentiation from PSCs, extensive variability still exists in the differentiation efficiency between different labs, as well as between human PSC lines. Several studies focused on increasing the yields of PSC-derived lung progenitors. Hawkins and colleagues demonstrated that during lung differentiation protocol NKX2-1- cells were more prone to generate non-lung lineage cells, such as liver and intestinal like cells. The authors sought to overcome the heterogeneity to achieve a more efficient differentiation downstream to mature lung cell lineages. Thus, fluorescent reporter cell lines or surface markers have been used to isolate the lung epithelial cells. Surface markers including Carboxypeptidase M (CPM) or the combination of $CD47^{hi}/CD26^{lo}$ were identified to isolate NKX2-1+ lung progenitor cells and to give rise to lung epithelial organoids (Gotoh et al., 2014; Hawkins et al., 2017). However, despite purifying the NKX2-1+ progenitors, the obtained cells seem to be transcriptionally heterogeneous and tend to differentiate further to both lung and non-lung cell types (Hurley et al., 2020; McCauley et al., 2018).

1.2.3.2. Transdifferentiation

Large scale generation of mature cells characterized by high purity is necessary for clinical applications. As mentioned above, traditional strategies for hPSC differentiation focus on sequential administration of signaling molecules to mimic *in vivo* development. However, these strategies often face major limitations such as low purity, inefficiency, time-consuming culture conditions and optimizations. Transdifferentiation is the direct conversion of one cell type to another one without entering a stem-like stage. Ectopic expression of tissue specific transcription factors can be used to induce the transdifferentiation process into specific cell types (Morita et al., 2015), following the example of the reprograming of somatic cells into iPSCs (Takahashi et al., 2007; Takahashi and Yamanaka, 2006). In addition, transdifferentiation can be achieved via targeted silencing or upregulation of endogenous genes using methods that focus on the direct manipulation of DNA or the epigenetic environment, such as CRISPR/Cas9 (Chakraborty et al., 2014; Z. Chen et al., 2017; Rubio et al., 2016). Last, transdifferentiation can occur by chemical treatments without the forced expression of TFs, thus representing an alternative, transgene-free approach that may be more cost-effective and easier to control (Thoma et al., 2014). Direct lineage conversion into lung-like cells

would provide an alternative cell source for the production of high number of functional lung cells suitable for cell/tissue replacement approaches in regenerative medicine (Wong et al., 2019).

1.2.3.3. Lung Organoids

Tissue function in adult organs ensues from the synergistic interaction of several cell types which are spatially arranged within a three-dimensional (3D) structure, supported by extracellular matrix (ECM). In order to mimic this structural organization *in vitro* and generate more physiologically relevant phenotypes for translational research, as well as to recapitulate some organ function, strategies have been developed for the production of small multicellular 3D-organ-like structures termed "organoids". Organoids can be derived from either primary progenitor cells or PSCs that grow in suitable 3D culture microenvironments comprised of mixtures of ECM proteins allowing the self-organization, as well as spatially restricted lineage commitment. This structural organization recapitulates biologically relevant cell-cell and cell-matrix interactions allowing the study of complex intercellular communication and organization networks, which are restricted in monolayer systems (Lancaster and Knoblich, 2014).

Generation of organoids from adult lung tissue was achieved from at least four regions of the lung, including trachea/bronchus (Hild and Jaffe, 2016; Rock et al., 2009), bronchioles (Peng et al., 2015; Vaughan et al., 2006), BADJ (Kim et al., 2005) and alveolous (Barkauskas et al., 2013) enabling the study of human lung progenitor cells, their plasticity and potential to affect repair, as well as the examination of factors that are required for self-renewal and differentiation. For instance, it was found that treatment of basal-derived bronchospheres with antibodies for Notch2 promoted induction of ciliated cells at the expense of secretory goblet cells, and thereby identifying Notch as a potential therapeutic target for goblet cell metaplasia disease (Danahay et al., 2015) or potentially other relevant lung diseases in which the proportion of ciliated versus secretory cells is disrupted. In addition, Rawlins lab identified the optimal conditions for maintaining self-renewing lung organoids derived from human fetal lung (Nikolić et al., 2017).

Organoids that recapitulate airway and alveolar lineages have been successfully generated also from PSCs (Y.-W. Chen et al., 2017; Gotoh et al., 2014; Hurley et al., 2020; Konishi et al., 2016; Korogi et al., 2019; McCauley et al., 2017; Porotto et al., 2019; Strikoudis et al., 2019). This resource of cells provides a promising tool for improving our understanding and treatment of lung diseases, as it enables the development of patient-specific treatments and allows the modeling of the

progression of genetic diseases for which only end-stage samples are typically available. Proof-of-concept studies include the generation of cystic fibrosis patient-derived iPSCs that were shown to recapitulate known *in vitro* disease phenotypes, followed by precisely correction of the mutation and restoring of CFTR function in those cells after lung differentiation (Crane et al., 2015; McCauley et al., 2017) or the modeling of diseases such as the Hermansky–Pudlak syndrome type (HPS) using lung organoids (Korogi et al., 2019). In addition, human iPSCs (hiPSCs) offer a source of progenitors that can efficiently be expanded upon CRISPR/Cas9 gene editing and reliably differentiate into functional organoids enabling the investigation of the function of human specific genes towards a desired lineage.

Importantly, stepwise differentiation protocols also allow the analysis of the kinetics and pathways underlying lung development in stages that are inaccessible for investigation *in vivo*. For instance, McCauley and colleagues (McCauley et al., 2017) revealed that Wnt signaling acts in human lung organogenesis by inhibiting proximal and promoting distal cell fate upon the respiratory specification, a mechanism that had been also observed previously in the mouse lung where Wnt seems to regulate P-D patterning and alveolar proliferation (Li et al., 2002; Shu et al., 2005).

1.2.3.4. Bioengineered human lung scaffolds for transplantation

Lung transplantation is the only curative treatment option for patients suffering from end-stage respiratory disease. However, in addition to the shortage of available lung donors, transplant efficacy remains a clinical issue due to the high rejection rates and the complications resulting from the immunosuppressive drugs (Yusen et al., 2015). In order to address these problems new, alternative approaches are needed. Recent advances in regenerative medicine and *ex vivo* lung tissue engineering attempt to increase the amount of donor lungs suitable for transplantation (Gilpin et al., 2016). Novel approaches have been developed and are being continuously improved making use of biologically derived or synthetic lung scaffolds which are seeded with cells and cultured *ex vivo* to generate functional organs for transplantation. Biologic acellular lung scaffolds are obtained by a procedure called decellularization that involves removal of all cellular material from native lung tissue without affecting the mechanical integrity, ECM scaffold structure, and composition (**Figure 6**) (Uhl et al., 2017; Wagner et al., 2013; Wu et al., 2017). Repopulation of the acellular lung scaffold with autologous stem or progenitor cells obtained from the intended transplant recipient via recellularization would construct then a personalized neo-organ or graft, eliminating the long-term

requirements for immunosuppressive drugs and risk of rejection (Tsuchiya et al., 2014). In addition, a decellularized organ has the benefit of preserving the original architecture and natural ECM which is important for the reconstruction of the complex branching structures and provides a chance to investigate cell-ECM interactions in a more *in vivo*-like setting, and in particular when acellular scaffolds are produced from diseased lungs (Parker et al., 2014; Sokocevic et al., 2013; Sun et al., 2016). Alternatively, artificial lung scaffolds produced by synthetic and natural polymers also demonstrate high potential in the field of lung bioengineering (Andrade et al., 2007; Ling et al., 2014; Singh et al., 2013). Despite the great potential of differentiated PSCs as deposit for material of recellularization (Ghaedi et al., 2013), and specifically in patients suffering from lung diseases caused by known gene alterations for which patient-derived iPSC could be gene-corrected prior to subsequent recellularisation (Firth et al., 2015), efficient differentiation of PSCs to proximal and distal epithelial cells remains an issue, and they have also been shown to form teratomas (Takahashi and Yamanaka, 2006). Therefore further optimizations are required prior to clinical use.

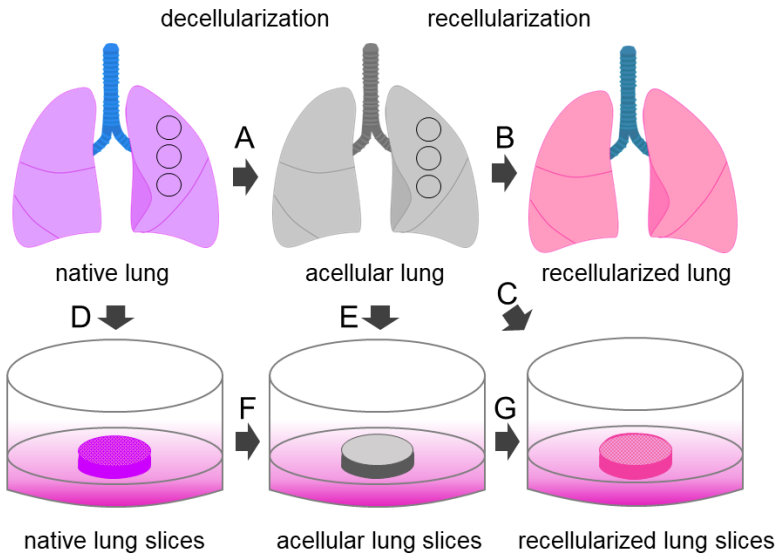


Figure 6. Schematic illustration of lung de- and re-cellularization approaches. (A, B) De- and re-cellularization of the whole lung or **(C-G)** using cut lung slices (shown as rings in **A**, **B**). The figure was adapted and modified from (Bölükbas et al., 2019).

1.3. MODELING HUMAN LUNG DEVELOPMENT USING PSCS

Studying human lung development has been challenging due to the limited access to embryonic and fetal tissue. Our current understanding of human lung development has been mainly shaped by molecular and genetic studies in mice. Studies in human lung developmental staging coupled with transcriptomic analysis of human fetal lung tissues followed (Kho et al., 2010), and more recently our knowledge was significantly boosted by analyzing the human fetal lung tissues and *in vitro* human lung developmental models at single-cell resolution (Nikolić et al., 2017). Despite the extremely high similarity between human and mouse lungs in terms of the cell type composition, maturation and function, studies have revealed significant cellular and molecular differences as well, including genes that have been used as cell-type specific markers, such as the co-expression of SOX2/SOX9 in human embryonic lung tips during pseudoglandular stage. This implies differences in the signaling requirements for culture maintenance (Danopoulos et al., 2018; Miller et al., 2018; Nikolić et al., 2017).

In vitro experimental systems offer important insights into human lung development. Notably, Nikolic and colleagues demonstrated that human embryonic epithelial lung tips are multipotent progenitors, and identified suitable conditions to grow them as long-term self-renewing organoids that can be directed to differentiate *in vitro* towards alveolar and bronchial cells, providing a genetic system for studying human development (Nikolić et al., 2017). However, this method is low-throughput and can be rather used for studying lung development only after respiratory cell fate specification. On the other hand, directed differentiation of hPSCs provides an unlimited supply enabling the generation of large numbers of lineage-specific differentiated cells for high-throughput experiments (**Figure 7**). This approach also allows studying of the kinetics of human embryonic development via the multistage, stepwise differentiation protocols and the research of cell fate decisions, in stages that remain elusive to scientific study.

1.3.1. Single-cell RNA-sequencing for studying early human development using iPSCs

During embryonic development cells acquire distinct fates by passing through dynamic transcriptional states. In order to investigate human embryogenesis and examine lineage commitments it is necessary to understand the sequence of events that take place and the signaling pathways involved in the process, as well as how histologically identical cells make diverse lineage choices to end up as different cell types. Human PSCs can give rise to specialized cell types through

stepwise changes in the transcriptional networks that coordinate fate choices from pluripotency into differentiated states, providing a reliable tool for understanding lineage-specific mechanisms that underlie development.

Pioneering computational and genome-wide profiling approaches are nowadays broadly used to uncover the molecular programs that underlie developmental processes. Although gene expression analysis of bulk RNA samples provide a virtual average of the diverse constituent cell types and does not allow distinction of the signals that lead a progenitor cell to a particular differentiation pathway, single cell profiling reveals cellular and transcriptional heterogeneity, as well as developmental dynamics of complex tissues such as a tumor or developing organ (Potter, 2018). In addition to the detailed transcriptomic characterization, which allows us to draw a comprehensive roadmap for the human embryonic lung development, single-cell RNA-sequencing (scRNA-seq) can illustrate unsuccessful differentiation events or off-target cell fates that may occur alongside the target cell type in directed differentiation approaches (Treutlein et al., 2016), providing crucial insights for protocol optimization, which usually requires costly and time-consuming trial-and-error experiments (Loh et al., 2014). Furthermore, scRNA-seq allows the computational reconstruction of differentiation trajectories in order to the define cell fates (Haghverdi et al., 2016; Trapnell et al., 2014).

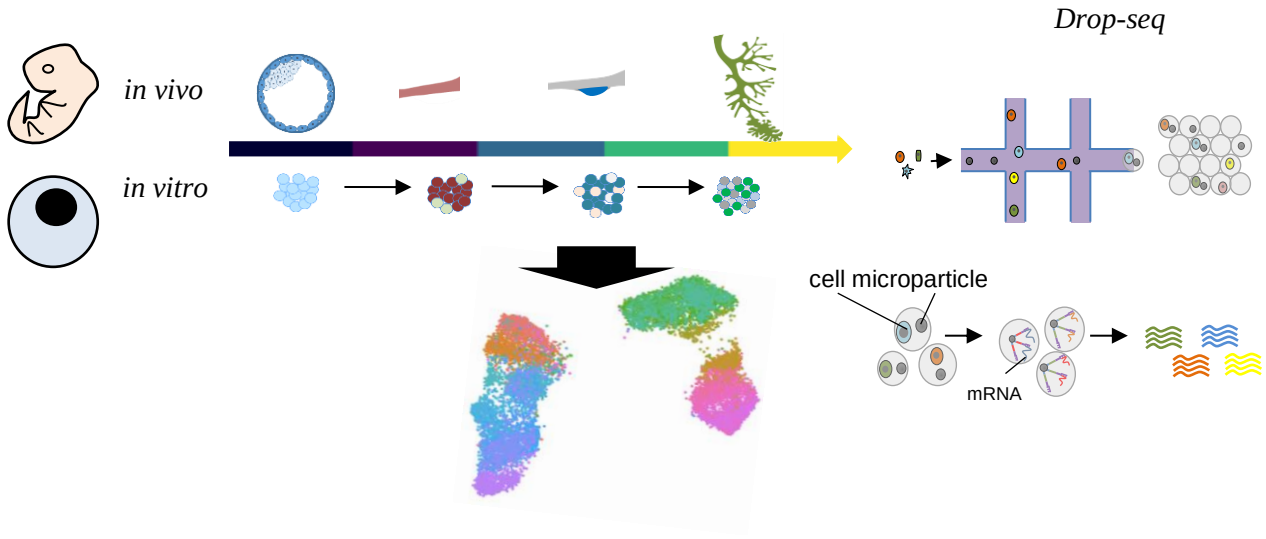


Figure 7. Schematic illustration of single cell analysis workflow during development or *in vitro* differentiation. Drop-seq scheme was adapted and modified from (Macosko et al., 2015).

2. AIM AND IMPACT OF THE STUDY

The mechanisms underlying the differentiation of human lung progenitors have not been examined systematically to date. The moderate efficiency and reproducibility of the differentiation protocols, as well as the narrow window that the early lung progenitors exist *in vivo* limit the study of human lung organogenesis and consequently the development of innovative therapeutic approaches. The overall objective of this thesis is to elucidate the transcriptional networks governing human lung specification using PSCs, as well as to resolve the cellular heterogeneity that defines the lung differentiation approaches.

I focused on the spatio-temporal regulation of transcriptional networks from pluripotency to NKX2-1 positive progenitors and revealed cellular identities and mechanisms with unprecedented resolution and timing, as well as other endodermal lineages that appear alongside lung progenitors. Prior to this undertaking, I confirmed the expression of stage specific signatures for the mouse and human developing lung in my hPSC differentiation protocol, and established that it is a reliable tool for modeling lung development. Furthermore, I reconstructed the molecular pathways leading to early lung progenitors *in vitro*. My analysis revealed the simultaneous emergence of hepatocytes and the underlying mechanisms. This finding offers a high resolution roadmap for the generation of hepatoblasts as well, contributing to the fundamental knowledge required for the improvement of hepatocyte generation from hPSCs.

My work represents the first detailed hierarchical model of gene expression changes along lung specification including lineage commitments and signaling pathways involved in the process. Likewise, dissecting the mechanisms of lung specification and understanding the fundamental programs driving the differentiation process will advance studies on tissue engineering and regeneration, and will further inform the optimization of robust and effective differentiation protocols, which typically require costly and time-consuming trial-and-error experiments.

3. MATERIALS AND METHODS

3.1. MATERIALS

All materials used in the study are further specified in respective sections of the experimental procedures, where they have been applied.

3.1.1. Cell lines

Human ESC line, H9 (WA09, WiCELL Research Institute), and the hiPSC lines, HMGU #1 and MHHi006-A-2 (or hHSC_F1285_T-iPS2-NKX2.1eGFP, hereafter referred to as *NKX2-1*$^{eGFP/+}$) (Olmer et al., 2019) were used in this study. The *NKX2-1*$^{eGFP/+}$ reporter cell line was generated using TALENs gene editing approach, targeting exon 3 of the human *NKX2-1* locus, and in this study it is used for monitoring the expression of NKX2-1 during directed differentiation. BJ cells (ATCC, Cat. No. CRL-2522™) are the human fibroblasts used for the validation of the fibroblast differentiation experiments as described in section **3.2.2.2.**

3.1.2. Cell culture media, supplements, chemicals and consumables

All reagents and chemicals routinely used are listed in **Table 1.**

Table 1. Media and supplements used in cell culture experiments.

Reagent	Manufacturer	Cat. No.
3-Isobutyl-1-methylxanthine (IBMX)	Sigma-Aldrich	I5879
8-Bromoadenosine 3′,5′-cyclic monophosphate sodium salt (cAMP)	Sigma-Aldrich	B7880
Accutase® solution	Sigma-Aldrich	A6964
all trans-Retinoic Acid (RA)	Sigma-Aldrich	R2625
B-27 Supplement	ThermoFisher Scientific	17504044
Bovine Albumin Fraction V (7.5% solution, BSA)	ThermoFisher Scientific	15260037
CHIR99021 trihydrochloride	R&D Systems	4953/50
Corning® Matrigel® Growth Factor Reduced (GFR) Basement Membrane Matrix	Corning	354230
CryoStor® cell cryopreservation media	Sigma-Aldrich	C2874
DAPT	TOCRIS	2634
Dexamethasone (Dex)	Sigma-Aldrich	D4902
DMEM/F12	Gibco™	11320074
Donkey serum	Sigma-Aldrich	D9663-10ML
Dorsomorphin dihydrochloride (DSM)	TORCIS	3093
Doxycycline hyclate	Sigma-Aldrich	D9891
Dulbecco's Phosphate Buffered Saline (DPBS)	ThermoFisher Scientific	14190-094

Fetal Bovine Serum (FBS)	HyClone™	SH30071.03
Formaldehyde 16% (w/v), Methanol-free	ThermoFisher Scientific	10321714
Gelatin powdered, pure Ph. Eur., NF	AppliChem GmbH	A1693
Geltrex Basement Membrane Matrix	ThermoFisher Scientific	A1413302
Gentle Cell Dissociation Reagent	STEMCELL technologies	07174
GluataMAX, 100X	Gibco™	35050061
Goat serum	Sigma-Aldrich	G9023-10ML
Ham's F-12 Nutrient Mix	Gibco™	21765029
Hygromycin B (50 mg/mL)	Gibco™	10687010
IMDM	Gibco™	12440053
KnockOut™ Serum Replacement (KSR)	Gibco™	10828028
L-Ascorbic Acid 2-Phosphate Magnesium	Sigma-Aldrich	A8960
MEM Non-Essential Amino Acids Solution (NEAA, 100X)	Gibco™	11140050
N2 Supplement	Gibco™	17502048
Penicillin-Streptomycin	Gibco™	15070063
PneumaCult™-ALI Medium with 6.5 mm Transwell® +Inserts	STEMCELL Technologies	5022
Propidium iodide (PI)	Sigma-Aldrich	P4170
Recombinant Human BMP-4 Protein	R&D Systems	314-BP
Recombinant Human FGF-10	R&D Systems	345-FG
Recombinant Human FGF-basic (154 a.a.; bFGF)	Peprotech	100-18B
Recombinant Human KGF/FGF-7 Protein	R&D Systems	251-KG
Recombinant Human Sonic Hedgehog (SHH)	Peprotech	100-45
Recombinant Human/Mouse/Rat Activin A Protein	R&D Systems	338-AC
RPMI 1640 Medium, no glutamine	Gibco™	31870025
SB431542 in Solution	Miltenyi Biotec	130-106-543
Sodium Butyrate	Sigma-Aldrich	B5887
StemMACS iPS-Brew XF	Miltenyi Biotech	130-104-368
Sucrose	Sigma-Aldrich	S0389
SYTOX™ Blue Dead Cell Stain (SB)	ThermoFisher Scientific	S34857
Tissue-Tek® O.C.T.™ Compound (O.C.T.)	Sakura Finetek	4583
Triton™ X-100	Sigma-Aldrich	X100-500ML
Trypsin-EDTA (0.25%), phenol red	ThermoFisher Scientific	25200056
Valproic Acid (VPA)	Cayman Chemical	13033
Y-27632 (RI)	R&D Systems	1254/10
α-Monothioglycerol (MTG)	Sigma-Aldrich	M6145

Table 2. Consumables

Consumable	Manufacturer	Cat. No.
10/20 µl XL TipOne® filter tips	Starlab	S1120-3810-C
1000 µl TipOne® filter tips	Starlab	S1126-7810-C
200 µl TipOne® filter tips	Starlab	S1120-8810
384-well PCR plates	Kisker	G034-ABI
5ml Polystyrene Round-Bottom Tube with Cell Strainer Cap	Corning	352235
Cryotubes	ThermoFisher Scientific	10577391
Falcon® 5 mL Tubes with Cell Strainer Cap	Corning	352235
Neubauer Hacmacytometer	Heinz Herenz	1080340
PCR Plate, 96-well, non-skirted	ThermoFisher Scientific	AB0600
Peel-A-Way embedding molds	Sigma-Aldrich	E6032-1CS
Sealing Tape Aluminium Foil	Starlab	E2796-9792
Superfrost® Plus Microscope Slides	ThermoFisher Scientific	J1800AMNZ
µ-Slide 8 well, ibidi-treat	Ibidi	80826
Conical Tubes (50 mL)	Invitrogen	AM12502
Falcon cell strainer tubes 35 µM	Corning	FALC352235
Safe-Lock Tubes 1.5 ml PCR clean	Eppendorf	0030123328
Safe-Lock Tubes 2.0 ml	Eppendorf	0030123344

3.1.3. Antibodies

All antibodies used in this study are listed below in **Table 3**.

Table 3. Antibodies

Antigen	Species	Dilution rate	Manufacturer	Cat. No.
Primary antibodies				
OCT-4	Rabbit	1:100	CST	2840
AFP	Mouse	1:100	Sigma Aldrich	WH0000174M1
NANOG	Rabbit	1:100	CST	4903
SOX2	Rabbit	1:100	CST	2748s
FOXA2	Mouse	1:100	Abcam	ab60721
SFTPC	Rabbit	1:100	Millipore	AB3786
SFTPB	Rabbit	1:100	Abcam	ab40876
SCGB1A1	Mouse	1:100	Santa Cruz	sc-365992
MUC5AC	Mouse	1:100	Abcam	ab79082
NKX2-1	Mouse	1:100	Santa Cruz	sc-53136
Vimentin (V9)	Mouse	1:100	Santa Cruz	sc-6260

Conjugated antibodies

EPCAM, APC-conjugated	Mouse	1:100	ThermoFisher Scientific	BDB347200
cKIT, APC-conjugated	Mouse	1:100	ThermoFisher Scientific	CD11705
CXCR4, PE-conjugated	Mouse	1:200	ThermoFisher Scientific	MHCXCR404

Secondary antibodies and Isotope controls

Alexa Fluor 488	goat anti-mouse	1:1000	ThermoFisher Scientific	A-11001
Alexa Fluor 546	goat anti-rabbit	1:1000	ThermoFisher Scientific	A-11035
Alexa Fluor 647	goat anti-rabbit	1:1000	ThermoFisher Scientific	A-21244
Alexa Fluor 594	donkey anti-mouse	1:1000	ThermoFisher Scientific	A-21203
PE	mouse IgG2a kappa isotype control	1:1000	eBioscience	12-4724-41
APC	mouse IgG2a kappa isotype control	1:1000	eBioscience	17-4724-81

3.1.4. Oligonucleotides

Oligonucleotides used for the experimental procedures (RT-qPCR and cloning) were purchased from Sigma-Aldrich. Primers were designed using the open source program Primer3web v. 4.1.0 (Koressaar and Remm, 2007). Oligonucleotides were reconstituted with H_2O to a stock concentration of 100 µM and are listed in **Table 4**.

Table 4. Primers used for RT-qPCR

Gene target	Sequence
FOXA2	F: GGGAGCGGTGAAGATGGA R: TCATGTTGCTCACGGAGGAGTA
SOX17	F: GGCGCAGCAGAATCCAGA R: CCACGACTTGCCCAGCAT
AFP	F: GCTTACACAAAGAAAGCCC R: TAATAATGTCAGCCGCTCC
CXCR4	F: CACCGCATCTGGAGAACCA R: GCCCATTTCCTCGGTGTAGTT
OCT4	F: CAATTTGCCAAGCTCCTGAAG R: AAAGCGGCAGATGGTCGTT
SOX2	F: CCTCCGGGACATGATCAGCATGTA R: GCAGTGTGCCGTTAATGGCCGTG
NANOG	F: CCTTCCTCCATGGATCTGCTT R: CTTGACCGGGACCTTGTCTTC
NKX2-1	F: CTTCCCCGCCATCTCCCGCTTC R: GCCGACAGGTACTTCTGTTGCTTG
PAX8	F: ACTACAAACGCCAGAACCCTACCA R: CCGGATGATTCTATTAATGGAG
PAX6	F: GCGGAGTTATGATACCTACACC R: GAAATGAGTCCTGTTGAAGTGG
MUC5AC	F: GCCTACGAGGATTTTAACAT R: CAGGACCGGGTGGCCGTTGA
SFTPC	F: GTTCTGGAGATGAGCATTGGG R: GCGATCAGCAGCTGCTGGTAGT

SFTPB	F: CTGGCCCAAGGTCGCCGA R: TGGAGCATTGCCTGTGGTATGG
SCGB1A1	F: ATGGACACACCCTCCAGTTATG R: TGGGCTATTTTTTCCATGAGC
GAPDH	F: GCTCATTTCCTGGTATGACAACG R: GAGATTCAGTGTGGTGGGGG
VIMENTIN	F: TCCACGAAGAGGAAATCCA R: CAGGCTTGGAAACATCCAC
FSP1	F: AGTACGTGTTGATCCTGACTG R: ACTTGTGGAAGGTGGACAC
ACTA2 (SMA)	F: CCTTCATCGGGATGGAGTC R: CCTTCCTGATGTCAATATCACAC
GATA6	F: GACTTGCTCTGGTAATAGCA R: CTGTAGGTTGTGTTGTGGG
EpCAM	F: CGAGTGAGAACCTACTGGA R: TGATCTCCTTCTGAAGTGCA
E-Cad	F: AGGAATTCTTGCTTTGCTAATTC R: CGAAGAAACAGCAAGAGCAGC

Table 5. Primers used to construct the overexpression vectors.

	Sequence
eGFP	
Fragment 1	F: ACACTTACCGCATTGACAAGCACGCCTCACGGGAGCTCCA
	R: GGTGAACAGCTCCTCGCCCTTGCTCACCATGGATCCGAGCTCGGTACCAAGCTTA
Fragment 2	F: TCCACCTTGCTATACGGTCGGACCTGGTGACTGCAGCGCGGGGATCTC
	R: TGGAGCTCCCGTGAGGCGTGCTTGTCAATGCGGTAAGTGT
Fragment 3	F: AAACTTAAGCTTGGTACCGAGCTCGGATCCATGGTGAGCAAGGGCGAGGAGCTGT
	R: GTGCTTTGGACTCATCGACTTGTCGTCATCGTCTTTGTAGTC
Fragment 4	F: TCCCATGGACTACAAAGACGATGACGACAAGTCGATGAGTCCAAAGCAC
	R: GAACTCCAGCATGAGATCCCCGCGCTGCAGTCACCAGGTCCGACCGTA
mCherry	
Fragment 1	F: TCCACCTTGCTATACGGTCGGACCTGGTGACTGCAGCGCGGGGATCTC
	R: TGGAGCTCCCGTGAGGCGTGCTTGTCAATGCGGTAAGTGT
Fragment 2	F: ACACTTACCGCATTGACAAGCACGCCTCACGGGAGCTCCA
	R: CTCCTCGCCCTTGCTCACCATGGATCCGAGCTCGGTAC
Fragment 3	F: GTACCGAGCTCGGATCCATGGTGAGCAAGGGCGAGGAGGAT
	R: GCTCCGGATCCCTTGTACAGCTCGTCCATGCCGCC
Fragment 4	F: ATGGACGAGCTGTACAAGGGATCCGGAGCCACG
	R: GAACTCCAGCATGAGATCCCCGCGCTGCAGTCACCAGGTCCGACCGTA
Primers used for the sequencing of vector constructs	
eGFP	TGAGCAAAGACCCCAACGAG
mCherry	CGCCTACAACGTCAACATCA

3.1.5. Commercial Kits

Kits that were routinely used in the study are listed in **Table 6**.

Table 6. Commercial Kits routinely used in this study.

Kit name	Manufacturer	Cat. No.
5-alpha Competent *E. coli*	NEB	C2987I
GeneJET Plasmid-Miniprep-Kit	ThermoFisher Scientific	K0502
Gibson Assembly® Master Mix	NEB	E2611
Agilent High Sensitivity DNA Kit	Agilent Technologies	5067-4626
MinElute Reaction Cleanup Kit	Qiagen	28206
P3 Primary Cell 4D- Nucleofector® X Kit	Lonza	V4XP-3024
PureLink™ HiPure Plasmid Filter Maxiprep Kit	ThermoFisher Scientific	K210017
Q5® High-Fidelity 2X Master Mix	NEB	M0492
QIAquick PCR Purification Kit	Qiagen	28104
QuantSeq 3′ mRNA-Seq Library Prep Kit for Illumina (REV)	Lexogen	015.24
RNeasy MinElute cleanup kit	Qiagen	74204
RNeasy Mini Kit	Qiagen	74106
Verso cDNA synthesis kit	ThermoFisher Scientific	AB1453A

3.1.6. Enzymes

Table 7. Enzymes and enzyme master-mixes used in the study.

Enzyme	Manufacturer	Cat. No.
Collagenase, type IV	Gibco™	17104019
Dispase® II	Sigma-Aldrich	D4693
EcoRI-HF	NEB	R3101
Gibson Assembly Master-Mix	NEB	E2611 L
Power SYBR® Green PCR Master Mix	ThermoFisher Scientific	4367659
Proteinase K Solution 20 mg/ml	ThermoFisher Scientific	AM2546
Q5® High-Fidelity 2X Master Mix	NEB	M0492
RNase A (20 mg/ml)	ThermoFisher Scientific	12091021

3.2. EXPERIMENTAL PROCEDURES

3.2.1. Maintenance of PSCs

3.2.1.1. Stem cell cultivation and passaging

All human iPSC and ESC lines (*NKX2-1*$^{eGFP/+}$, HMGU #1, H9) were maintained in feeder-free conditions, on Geltrex Basement Membrane Matrix coated plates in StemMACS iPS-Brew XF. Passaging was performed in routine basis by washing the cells with DPBS and harvesting them after incubation at 37°C for 10 min with Accutase or filtered 2 mg/ml collagenase IV in DMEM/F12. The *NKX2-1*$^{eGFP/+}$ reporter cell line was kindly provided from Hannover Medical School (Olmer et al., 2019). Cells were maintained in a HERAcell 240i incubator at 37°C and 5% CO_2. All centrifugation steps were performed on a Megafuge 40R centrifuge (ThermoFisher Schientific) at room temperature (RT) for 3 minutes (min) and 200 x *g*.

3.2.1.2. Freezing and thawing

hPSCs were cryopreserved in CryoStor® cell cryopreservation media. When cells reached 80% confluency in a 6-well plate, they were washed with DPBS followed by 10 min incubation with Accutase, which was then quenched with fresh medium. Upon centrifugation, cells pellets were resuspended in 1 ml CryoStor®, transferred in cryotubes which were cooled to -80°C in freezing containers and 24 hours later placed into liquid nitrogen tanks for long-term storage. For thawing, cells were placed in water bath at 37°C, centrifuged and resuspended in fresh culture medium. Cells were then transferred for maintenance on tissue culture plates with fresh medium enriched with 10μM Y-27632 for the first 24 hours and cultured as described before.

3.2.2. Directed differentiation of PSCs

3.2.2.1. Differentiation to lung epithelial cells

For the directed differentiation of PSCs into lung progenitor cells and organoids, a stepwise approach was used, to recapitulate the respiratory lineage patterning, via definitive endoderm, foregut endoderm and lung progenitor stage as shown in **Figure 8**.

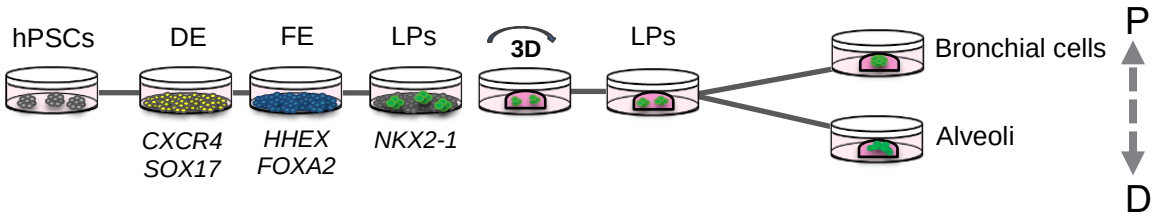


Figure 8. Schematic illustration of PSC differentiation to lung epithelial cells. Respiratory lineage patterning via a stage specific approach. Representative stage specific markers are shown below the scheme. DE: definitive endoderm, FE: foregut endoderm, LP: lung progenitor stage and P-D: proximal-distal axis. Adapted from (Ori et al., 2021).

3.2.2.1.1. Differentiation of PSCs to definitive endoderm

For the formation of DE, hPSCs that reached 80% confluency (day 0) were rinsed with DPBS and incubated with Accutase for 10 minutes, at 37°C. Single-cell suspensions were then seeded onto 24-well plates, which were pre-coated with 1:40 Growth Factor Reduced (GFR) Matrigel, in a density of 2 x 10^5 cells/cm^2. Cells were treated with 100 ng/ml Activin-A, 1 μM CHIR99021, and 10 μM Y-27632, in Definitive Endoderm Basal Media (DE-BM or iii in section **4.1.1.**) consisting of RPMI1640 medium, 1 x B27 supplement, 1 x NEAA and GlutaMAX. Growth media and supplements are listed on **Table 1**. On days 1-6 the DE-BM was supplemented with 100 ng/ml Activin-A, 1 μM CHIR99021 and 0.25 mM (day 1) and 0.125 mM (days 2-6) sodium butyrate (Gotoh et al., 2014; Ori et al., 2021).

Additional conditions had been tested (see section **4.1.1.**) for the induction of DE. The approach from Huang et al. (2014) (i) includes the use of ultra-low attachment plates for EB formation containing Serum-free differentiation (SFD) media supplemented with 10 μM Y-27632, 10 ng/ml Wnt3a and 3 ng/ml BMP4 for 24h, followed by SFD with 10 μM Y-27632, 100 ng/ml Activin-A, 0.5 ng/ml BMP4 and 2.5 ng/ml FGF2 for 4 additional days. SFD medium contained 75% IMDM, 25% of F-12, 0.05% BSA, 1% GlutaMAX, 0.5% N2, 1% B27, 1% penicillin-streptomycin, 50 μg/ml L-ascorbic acid and 0.04 μl/ml MTG. Furthermore, the approach from Firth et al. (2014) (ii) refers to RPMI basal media supplemented with 25 ng/ml Wnt3a and 100 ng/ml Activin-A for 24h, followed by RPMI supplemented with 1% FBS and 100 ng/ml Activin-A for 3 additional days.

3.2.2.1.2. DE patterning into foregut endoderm

For the induction of FE stage two Basal Mediums (FE-BM1 and FE-BM2) were tested and prepared as shown in **Table 8**. On day 6 of differentiation DE cells were collected with Accutase and re-plated in a density of 1:3 onto GFR Matrigel-coated plates and in FE-BM. Then, cells were treated for 2 days (day 6, 7) with 50 ng/ml SHH, 2 µM DSM and 10 µM SB431542, supplemented with 10 µM Y-27632 for the first 24 hours. On day 8 the medium was changed to FE-BM1 or 2 respectively, with 2 µM DSM and 10 µM SB431542 (Ori et al., 2021).

3.2.2.1.3. Generation of LPs

To generate lung progenitor cells, the medium was switched to FE-BM1 or 2 containing 20 ng/ml recombinant human BMP4 (rhBMP4), 50 nM retinoic acid (RA), 3 µM CHIR99021 and 20 ng/ml rhFGF10. For the inhibition of TGF-β and Notch pathways, 10 µM SB431542 and 100 µM DAPT were additionally used at the LP stage, respectively (Ori et al., 2021).

3.2.2.1.4. Generation of LP-derived lung organoids

On day 15 of differentiation, the cells were rinsed twice with DPBS and eGFP+ colonies were precisely dissected and transferred to 80% Matrigel (diluted in FE-BM2). Drops of 40 µl were pipetted into the centre of each well of a 24-well tissue culture plate and incubated for 30 min at 37°C (**Figure 9**) (Ori et al., 2021).

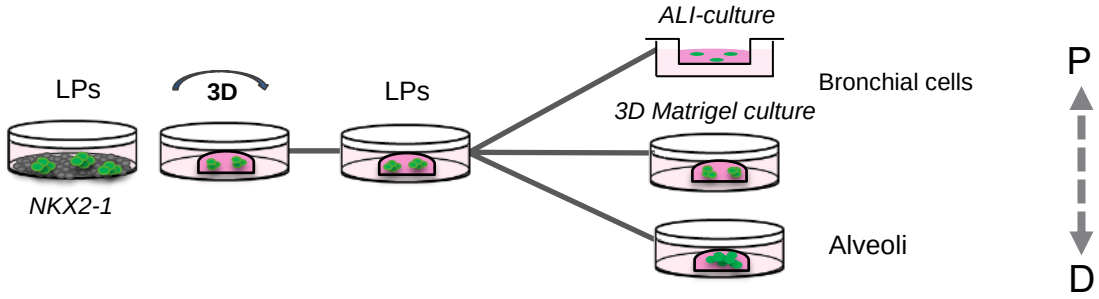


Figure 9. Schematic illustration of lung organoid generation from lung progenitor cells.

Then, FE-BM2 media supplemented with 10 ng/ml FGF10, 10 ng/ml KGF and 3 µM CHIR9902 was added to each well and changed daily for 7 days. On day 22 the medium was switched to FE-BM2 media supplemented with 50nM dexamethasone, 0.1 mM 8-Bromoadenosine 30,50-cyclic monophosphate sodium salt (cAMP), 0.1 mM 3-Isobutyl-1-methylxanthine (IBMX) and ± 10 µM SB431542, 3 µM CHIR99021 (DCI±SC). The PneumaCult™ Air Liquid Interface (ALI)-medium

was used to induce more proximal lineages. On day 15, eGFP+ colonies were transferred in 1:2 GFR Matrigel/FE-BM2 and placed in 6.5mm inserts, in a total amount of 100 µl. After incubation for 20 min at 37°C, FE-BM2 supplemented with 10 ng/ml FGF10, 10 ng/ml KGF and 3 µM CHIR9902 was added in the lower chamber until day 27. On day 28 of the differentiation protocol, the medium was switched to PneumaCult™-ALI Maintenance Medium which was prepared according to the manufacturer's instructions supplemented with 10 µM DAPT (Konishi et al. 2016; Firth et al. 2014). All growth media and supplements are listed in **Table 1**. For all the 3D Matrigel culture conditions the medium was changed every third day (Ori et al., 2021). For splitting, organoids were first recovered from Matrigel and then dissociated. After passaging, organoids were grown in Matrigel and FE-BM2 media supplemented with DCICS and 10 ng/ml KGF.

3.2.2.2. Generation of secondary fibroblasts from PSCs

For the generation of PSC-derived fibroblasts, on day 0 PSCs were treated with 2 mg/ml of collagenase IV solution and were plated as colonies onto low attachment plates containing fibroblast medium (FB-BM, **Table 8**), to allow embryoid body (EB) formation. On day 7 of culture, EBs were transferred to 0.1% gelatin / 2% fetal bovine serum (FBS) coated 6-well plates, and kept in 37°C for 7 more days. On day 14, cells were dissociated with 0.25% Trypsin and transferred onto plates pre-coated with 1% Matrigel for maintenance (Togo et al., 2011). BJ cells were used to validate the differentiation efficiency.

3.2.2.3. Ectopic expression of NKX2-1 in PSCs and secondary fibroblasts

Cells were subjected to fibroblast differentiation include (*iNKX2-1-eGFP-)H9* and (*iNKX2-1-mCherry-)NKX2-1*$^{eGFP/+}$ lines.

3.2.2.3.1. Acceleration of lung differentiation by ectopic expression of NKX2-1 in PSCs

On day 0, hPSCs cells that had reached 80% confluency were cultured in FE-BM2 supplemented with 3 µM CHIR99021, 50 nM RA, 20 ng/ml BMP4, and 20 ng/ml FGF10 for 24 hours. The day of pre-treatment is referred to as day 14. On day 1, cells were incubated for 1.5 hours in collagenase IV and then were gently detached, collected and centrifuged at 100 x *g* for 1 min. Cells were re-suspended as clumps in differentiation medium and distributed in tubes containing 15-30 cell clumps each. Upon removal of the excess medium, cell clumps were resuspended in 40 µl of undiluted Matrigel, pipetted as drops onto the wells of a 24-well plate and incubated for 20 min at 37°C. Fresh

medium (with composition depending on the experimental setup) was added to the wells. For all 3D Matrigel cultures the medium was changed every 3 days. Induction intervals of ectopic *NKX2-1* were tested as demonstrated in section **4.4.**

3.2.2.3.2. Transdifferentiation of secondary fibroblasts into lung epithelial cells via ectopic expression of NKX2-1

Media composed of basal medium FE-BM2 containing 3 μM CHIR99021, 10 ng/ml KGF, 10 ng/ml FGF-10 (CKF) ± 20 ng/ml BMP4, 50 nM RA (CKF+BR) or 2 μM DAPT (CKF+D) were tested. Induction intervals for the over-expression of *NKX2-1* and treatment with 1mM VPA were tested as shown in section **4.4.**

Table 8. Cell culture media mixtures used in differentiation experiments.

Medium name	Base	Supplements
DE-BM (iii)	RPMI	1 x NEAA, 1 x Glutamax, 1 x B27
FE-BM1 (Gotoh et al., 2014)	DMEM/F12	1 x N2, 1 x B27, 1 x GlutaMAX, 0.05 mg/ml of L-Ascorbic Acid, 50 U/ml Pen-Strep, 0.4 mM of MTG
FE-BM2 (Hawkins at el., 2017)	75% IMDM	25% Ham's F-12, 0.05% BSA, 0.5 x N2, 0.5 x B27, 1 x GlutaMAX, 0.05 mg/ml of L-Ascorbic Acid, 50 U/ml Pen-Strep, 0.4 mM of MTG
ALI	PneumaCult ALI	Acc. to manufacturer's instructions
FB-BM	DMEM	10% FBS, 10% KSR, 1 x NEAA, 1 x GlutaMAX, 0.05 mg/ml of L-Ascorbic Acid, 0.45 mM MTG, 4.0 ng/ml bFGF

3.2.3. Flow cytometry analysis

Flow cytometry analysis was performed using a BD FACSAria™ III cell sorter (BD Biosciences) and final data analysis was performed on FlowJo software. All antibodies used for this study are listed on **Table 3**.

3.2.3.1. Cell preparation for Flow Cytometric Analysis

Cells were rinsed with DPBS and incubated in Gentle Cell Dissociation Medium/0.05% Trypsin (2vol:1vol) for 10 min, at 37°C. Then, the detached cells were diluted in FACS buffer, consisted of 2% FBS/DPBS and centrifuged at 200 x *g* for 3 min. Then, the cell pellets were washed with FACS buffer, centrifuged again and resuspended in FACS buffer.

3.2.3.2. Immunofluorescence staining for Flow Cytometric Analysis

The samples were incubated (if necessary) with conjugated antibodies diluted in FACS buffer

(section **3.2.3.1.**) for 25 min on ice, washed twice with FACS buffer and analyzed. Live cells were distinguished by staining with propidium iodide (PI) or Sytox blue (SB). Isotype controls were used for gating stained cells, whereas undifferentiated cells that do not express eGFP were used as negative control for gating eGFP+ cells (Ori et al., 2021).

3.2.3.3. Fluorescence-Activated Cell Sorting

For sorting, single cell suspensions were obtained as described above and resuspended in FACS buffer stained with PI. The dissociated cells were sorted into 2% FBS/DMEM.

3.2.4. Immunofluorescence staining

Immunofluorescence staining was performed protected from light. All antibodies used for the study are listed on **Table 3**.

3.2.4.1. For monolayer culture

For imaging 2D cultures, cells were cultivated on removable 8-well chamber glass slides (Ibidi), rinsed twice with DPBS and fixed with 4% paraformaldehyde (PFA) for 20 min at RT. Cells were washed three times with DPBS, permeabilized with 0.2% Triton X-100 for 5 min and blocked with 3% bovine serum albumin (BSA) for 30 min at RT. Next, cells were incubated with primary antibodies diluted in solution containing 0.1% Triton X-100/3% BSA overnight at 4°C. After three washes with DPBS, cells were incubated with secondary antibody in 0.1% Triton X-100/3% BSA solution for 1 hour at RT and washed again three times with DPBS. The removable wells were then removed and ProLong™ Glass Antifade Mountant with NucBlue Stain was used for coverslip mounting (Ori et al., 2021).

3.2.4.2. For organoids

To image lung organoids, the samples were first incubated with 2 mg/ml Dispase II, at 37°C for 1 hour in order to be extracted from the Matrigel. Organoids were washed twice with DPBS, fixed with 4% PFA for 2 hours at 4°C and incubated in 30% sucrose/DPBS overnight at 4°C. The samples were then embedded in O.C.T. compound frozen and stored in -20°C. Cryosections with 6 μm thickness were permeabilized and blocked with 0.5% Triton X-100/5% Normal Goat or Donkey Serum (NGS or NDS) for 30 min at RT. Afterwards, the slices were incubated with primary antibodies, diluted in 0.5% Triton X-100/5% NGS or NDS, overnight at 4°C, followed by washing three times with DPBS

(10 min each), and stained with secondary antibodies for 3 hours at RT. Upon two more washing steps, ProLong™ Glass Antifade Mountant with NucBlue Stain was used for coverslip mounting (Ori et al., 2021).

3.2.4.3. Imaging

Imaging was performed with an AXIO Observer.Z1 microscope inverted epifluorescence microscope (Zeiss). Immunofluorescence images were processed and edited using FIJI (ImageJ) software.

3.2.5. Transcriptomic analysis

3.2.5.1. RNA isolation

RNA extraction was performed using the RNeasy Mini Kit according to the manufacturer's instructions. For lysis of the samples, RLT buffer was either added a) directly onto adherent cells grown in 24-well cell culture plates, b) to cell pellets that had been harvested with Accutase, or c) to lung organoids after recovering from Matrigel via Dispase II, as described above. Isolated RNA was stored at -80°C.

3.2.5.2. Reverse transcription - cDNA synthesis

RNA concentration was measured using the Nanodrop ND-1000 system. Equal RNA concentrations (100-500 ng) were reverse transcribed into cDNA using the Verso cDNA Synthesis Kit for each experiment, according to manufacturer's instructions.

3.2.5.3. qPCR

qPCR was performed in 384-well plates using Power SYBR® Green PCR Master Mix in a total reaction volume of 10 µl on a QuantStudio 12k Flex (Life Technologies). Following cycling conditions were applied: 2 min at 50 °C, 10 min at 95 °C, 40 cycles of 15 sec at 95 °C and 1 min at 60°C. All Ct values were normalized to the respective values of housekeeping gene *GAPDH* before calculating the relative quantification. Primers are listed on **Table 4**.

3.2.6. Generation of *iNKX2-1-eGFP-H9* and *iNKX2-1-mCherry-NKX2-1$^{eGFP/+}$* overexpression cell lines

3.2.6.1. Construction of vectors

For the generation of overexpression vectors, fragments with overlapping sequences were amplified via PCR and isolated. PiggyBac backbone PB-GFP-FLAG-SRF was obtained from Dr. Shaposhnikov (Helmholtz Center Munich). A PB-iNKX2-1-eGFP construct was generated using as a template the PB-GFP-FLAG-SRF backbone which harbors an eGFP coding sequence followed by a P2A signal upstream of the inserted sequence, enabling the cleavage of the translated eGFP and NKX2-1 peptides. It also includes a tetracycline-inducible promoter system, a Hygromycin resistance gene and a transposon, enabling genomic integration upon co-transfection with a transposase plasmid. The primers and templates used for amplification are listed in **Table 5**. Amplification of DNA fragments was performed using the Q5 High-Fidelity 2X Master Mix according to the manufacturer's instructions. Following cycling conditions were used: (1) 2 min at 98°C; (2) 10 cycles of 10 sec at 98°C, 40 sec at 70 °C, 1 min at 72°C; (3) 25 cycles of 10 sec at 98°C, 40 sec at 60°C, 1 min at 72°C; (4) 3 min at 72°C. For the amplification of the NKX2-1, cDNA product (on day 15 of differentiation) was isolated and purified with QIAquick PCR Purification Kit according to manufacturer's instructions. For DNA purification the MinElute Cleanup Kit was used according to manufacturer's specifications. Purified fragments were ligated using the Gibson Assembly Master Mix according to manufacturer's instructions to final constructs. To generate the PB-iNKX2-1-mCherry, the eGFP sequence of the generated PB-iNKX2-1-eGFP was exchanged with a gene sequence for the red fluorescent protein mCherry. mCherry sequence was amplified from FUW-tetO-mCherry (also obtained from Dr. Shaposhnikov, Helmholtz Center Munich) with overhangs complementary to PB-iNKX2-1-eGFP for the subsequent assembly.

3.2.6.2. Transformation and plasmid isolation

NEB® 5-alpha competent *E. coli* bacteria were inoculated with ligated plasmid and plated on Luria-Bertani (LB) agar plates containing 50 μm/ml ampicillin and incubated at 37°C overnight. Bacteria colonies were picked and propagated; and plasmid DNA was isolated using the GeneJET Plasmid Miniprep Kit according to manufacturer's instructions. The correct integrations were verified by Sanger sequencing (at GATC Biotech, Konstanz, Germany) using the primers listed in **Table 5**. Subsequently, upon validation, a clone with the correct construct was expanded in 100 ml LB

medium containing 50 μm/ml ampicillin, incubated at 37°C overnight and plasmids were then isolated using the PureLink HiPure Plasmid FP Maxiprep Kit, according to manufacturer's specifications. DNA concentration was measured using the Nanodrop ND-1000 system and plasmids were stored at -20°C.

3.2.6.3. Nucleofection

PSCs were dissociated and harvested using Accutase and counted with a Neubauer Chamber. 1.0 x 10^6 cells were nucleofected with 6 μg of DNA in total (including the 3 μg of a plasmid expressing the PiggyBac transposase, obtained from Dr. Shaposhnikov) on a 4D-Nucleofector™ System (Lonza) using the P3 Primary Cell 4D-Nucleofector Kit according to manufacturer's instructions. Cells were plated on 1:100 Matrigel or Geltrex/DMEM/F-12-coated plates containing StemMACS iPS-Brew XF supplemented with 10 μM Y-27632 for 24 hours. The medium was changed every day with fresh StemMACS iPS-Brew XF and cells underwent selection with 50 ng/ml Hygromycin B for at least two weeks before subjected to an experiment. To validate the stably integrated clones, overexpression was induced in cells via treatment with 1 μg/ml dox for 24 hours, and eGFP or mCherry expression was assessed via fluorescence microscopy. Expression of *NKX2-1* was tested by RT-qPCR and flow cytometry.

3.2.7. Transcriptomic analysis – sequencing experiments

3.2.7.1. Bulk mRNA-seq library construction, sequencing and analysis

Transcriptome analysis of undifferentiated hiPSCs (day 0), sorted eGFP+ and eGFP- (day 15) and DCI (±CS) lung organoids (day 35/50) was performed using the QuantSeq 3′ mRNA-Seq Library Prep Kit for Illumina with Custom Sequencing Primer (Lexogen) following the manufacturer's instructions. The 3′ mRNA sequencing libraries were prepared from 110 ng of total input RNA per sample which was isolated with the RNeasy Mini Kit, as described in section **3.2.5.1.** (Ori et al., 2021). The quality of the libraries was evaluated on an Agilent 2100 Bioanalyzer using the High Sensitivity DNA Kit. Then, samples were sequenced using HiSeq2500 machine. Analyses were carried out using SeqMonk (Babraham Bioinformatics (https://www.bioinformatics.babraham.ac.uk/projects/seqmonk/) and differential expression analysis was performed using DESeq2 (Love et al., 2014). GO terms analysis was performed using Panther (Thomas et al., 2003) and String platforms (Szklarczyk et al., 2019; Ori et al., 2021).

3.2.7.2. Single cell RNA-seq (Droplet-sequencing)

3.2.7.2.1. Generation of single-cell suspensions

Accutase was used for the daily sampling of cells during differentiation. The detached cells were centrifuged for 5 min at 300 × *g* (4 °C), counted with a Neubauer chamber and assessed for single-cell separation and viability. Then, 250,000 cells were aliquoted in 2.5 ml PBS supplemented with 0.04% BSA and loaded for DropSeq at a final concentration of 100 cells/µl (Ori et al., 2021).

3.2.7.2.2. Single-cell RNA sequencing

Dropseq experiments were carried out based on the original Dropseq protocol (Macosko et al., 2015; Ziegenhain et al., 2017). Single cell (100/µl) suspensions were co-encapsulated in droplets with barcoded beads (120b/µl, purchased from ChemGenes Corporation, Wilmington, MA) at rates of 4000 µl/h, using a microfluidic polydimethylsiloxane device (Nanoshift). Droplet emulsions were collected for 15 min/each prior to droplet breakage by perfluorooctanol (Sigma-Aldrich) and upon breakage, beads were harvested and the hybridized mRNA transcripts were reverse transcribed (Maxima RT, Thermo Fisher). Exonuclease I (New England Biolabs) was added for the abolition of the unused primers, after which beads were washed, counted, and aliquoted for pre-amplification (2000 beads/reaction, equals ~100 cells/reaction) using a total of 10 PCR cycles. PCR details: (Smart PCR primer: AAGCAGTGGTATCAACGCAGAGT (100 µM), 2× KAPA HiFi Hotstart Readymix (KAPA Biosystems), cycle conditions: 3 min 95 °C, 4 cycles of 20s 98 °C, 45s 65 °C, 3 min 72 °C, followed by 8 cycles of 20s 98 °C, 20s 67 °C, 3 min 72 °C, then 5 min at 72 °C) (Macosko et al. 2015; Ori et al., 2021).

PCR products of every sample were combined and purified twice by 0.6 × clean-up beads (CleanNA), according to the manufacturer's instructions. In order to obtain transcript integrity, purity, and amount, complementary DNA (cDNA) samples were loaded on a DNA High Sensitivity Chip on the 2100 Bioanalyzer (Agilent) before the tagmentation. 1 ng of pre-amplified cDNA from approximately 600 cells was tagmented by Nextera XT (Illumina) with a custom P5 primer (Integrated DNA Technologies) for every sample. The single-cell libraries were sequenced in a 100 bp paired-end run on the Illumina HiSeq4000 using 0.2 nM denatured sample and 5% PhiX spike-in. For priming of read 1, 0.5 µM Read1CustSeqB was used (primer sequence: GCCTGTCCGCGGAAGCAGTGGTATCAACGCAGAGTAC) (Ori et al., 2021).

3.2.7.2.3. Bioinformatics Processing of the data set

For the creation of the count matrices, the Drop-seq computational pipeline (version 2.0) was used, as previously elaborated (Macosko et al., 2015). Mapping was performed by STAR (version 2.5.2a) (Dobin et al., 2013) and the reads were aligned to the hg19 reference genome (provided by the Drop-seq group, GSE63269). Seurat Version 2.3 (Satija et al., 2015) was used for downstream analyses, upon aggregating the ensuing count matrices into one combined object. For barcode filtering, the barcodes with less than 400 identified genes were omitted. Given that 1000 cells were anticipated per sample, the first 1200 cells, sorted by number of transcripts per cell, were used further. A high proportion of transcript counts obtained from mitochondria-encoded genes might point out dying or stressed cells, so cells with a percentage of mitochondrial genes of above 20% were deducted from downstream analysis. Additional filtering was conducted driven by histograms of quality metrics. Cells with many UMI counts may depict doublets, so only cells with less than 5000 UMIs were used in downstream analysis (Ori et al., 2021).

The common pre-processing procedure followed upon the initial filtering of cells. For the normalization and scaling of the expression matrices, Seurat's NormalizeData() and ScaleData() functions were used. With the aim to reduce the effects of unwanted sources of cell-to-cell variation, cell-cycle effects, percentage of mitochondrial reads and number of counts were regressed out CellCycleScoring() and ScaleData(). The 600 top variable genes across the data set were selected via FindVariableGenes() using the default parameters, by which known cell cycle genes were omitted. Those variable genes constituted the outset of the principal component analysis. The first 50 components were the input for Seurat's function FindClusters() at a resolution of 2, producing 13 clusters. In order to visualize the clustering outcome of the high dimensional single-cell data, the UMAP was produced using again 50 components as input for the Seurat function RunUMAP() with a number of neighbors set to 20. To additionally clean up the data set, few cells disagreeing in the Louvain cluster and UMAP embedding were omitted (Ori et al., 2021).

3.2.7.2.4. Selection of genes with significant association over time

To account for the severe disruption in gene expression caused by the medium change, the subsequent analyses were divided into three stages corresponding to the DE (days 0 - 6), FE (days 7 - 10) and LP stage (days 11 - 15). The stage wise temporal analyses were performed using the Python package Scanpy (Wolf et al., 2018). Scanpy's pp.pca() and pp.neighbors() were used for the

recalculation of the neighborhood graph based on the first 10 components and the principal components for every individual stage. After generating the diffusion maps via the tl.diffmap() function setting the root cells after manual inspection, the diffusion pseudotime was calculated with tl.dpt(). The diffusion pseudotime for the whole data set was defined by adding the stage-wise pseudotimes and scaled to a value between 0 and 1. In order to determine global connectivity, the Louvain clusters were used as input for Partition-based graph abstraction (PAGA) tl.paga(). The weighted edges indicate a statistical measure of connectivity between the groups. Edges with a weight < 0.05 are omitted (Ori et al., 2021).

To identify genes that demonstrate significantly changing expression patterns over time, the following strategy was used by the means of the R packages limma (Ritchie et al., 2015) and splines. Only positive cells for either *eGFP* or *NKX2-1* from days 11 to 15 were chosen for the calculation and demonstration. Moreover, ribosomal derived genes were omitted from this analysis. A regression model was applied, in which splines are used to model non-linear effects of continuous variables to fit the time-course data. For every gene a natural cubic spline with 4 knots was fit while using the time point of extraction as explanatory variable. UMI counts of each cell were included as a covariate in the model to account for differences in library size. Genes were ordered depending on their adjusted p-values. Given that it is non-trivial to interpret p-values along time, we employ the adjusted p-value as a ranking for the genes and displayed the top 200 genes in **Figure 29C** (Ori et al., 2021).

3.2.7.2.5. Stage-wise hierarchical clustering of genes

The clustering analysis was performed for DE and the combined DE and FE. Genes for both versions with appealing expression kinetics were chosen using the regression model based on spline fits. As an input for hierarchical clustering, using the hdist() from the R package stats, the scaled expression levels of genes with an adjusted p-value of less than 0.005 were used. The dendrogram tree was divided into 10 clusters that were then re-annotated into 6 groups for every developmental stage of the differentiation protocol based on their average expression per sampled time point. Given that the cluster sizes can vary, we sought to feature the average expression of the 100 genes with the lowest adjusted p-value per cluster in **Figure 32B** and **35B**. For the second clustering analyses, we display the average expression and kinetic profiles for day 6 to day 10 (Ori et al., 2021).

3.2.7.2.6. Potential branching into lung and non-lung progenitor cells

In order to pinpoint genes correlated to the potential differentiation branches, the signatures from the differential gene expression analysis between eGFP+ cells versus eGFP- cells from the bulk experiment were used. As variable genes for further analysis those genes were chosen, which exhibited a log fold change of above 1 and below -1. The effect of media change after day 10 was limited by excluding genes that were differentially expressed between the FE and LP stage. To control the dimensionality limitation these 1294 genes were used as input for the PCA. UMAPs and diffusion maps were calculated using Scanpy's workflow with 50 principal components and the number of neighbors set to 10. Then filtering for the two branches of interest was performed, which started in the FE stage and had their endpoint in the LP stage. Cells were scored by their similarity to the bulk signature via tl.score(), to estimate whether the two potential cell fate trajectories correspond to the branches that were obtained (Ori et al., 2021).

The Python package diffxpy (https://github.com/theislab/diffxpy) was used to determine genes showing temporally altered expression patterns along the two potential branches. All cells from FE stage were selected as a starting population for the two trajectories (lung and hepatocyte progenitors). Cells from the DE were divided according their Louvain clusters upon visual inspection of the diffusion map and known lineage specific marker gene expression, while the diffusion pseudotime was calculated on these two trajectories individually. Diffxpy's test.continuous_1d() function is based on a test with a spline basis in order to permit smooth trends. For that the combination of the two trajectories was used, with the pseudotime serving as a constant covariate and a categorical annotation "trajectory" (indicating in which branch each cell belong) as factor to control. The trajectory-wise pseudotimes were binned into "source" and 4 groups each in order to obtain a better illustration outcome. The heatmap in **Figure 37A** is the result of plotting the average expression of cells per bin for the top 100 genes ordered by adjusted p-value from the diffxpy (Ori et al., 2021).

3.2.7.2.7. Pairwise gene expression profiles along pseudotime

Cells ordered by diffusion pseudotime were divided into 9 bins and 10 bins for the liver and the lung branch, respectively, as described above. Pairwise gene correlations along the trajectories were determined with the cor() function (default parameters, R version 3.6.1) using the average expression in each bin, and illustrated using corrplot package in R (version 0.84) (**Figure 36**) (Ori et al., 2021).

3.2.8. Statistical analysis and data visualization

Sample sizes for the flow cytometry, gene expression analysis experiments and organoid size measuring are indicated in respective figure legends, and plots have been generated with GraphPad Prism 7.0 software. Statistical analysis was conducted using GraphPad Prism 7.0 software. P-values below 0.05 ($p<0.05$) were considered significant. P-value: *<0.05, **<0.01, ***<0.001. Star war plots and heatmaps displaying data of bulk mRNA sequencing analyses were generated in SeqMonk (Babraham Bioinformatics; (https://www.bioinformatics.babraham.ac.uk/projects/seqmonk/). For the circular bar chart (shown in **Figure 30**) statistical and bioinformatics operations, such as normalization, pattern recognition and pathway enrichment analysis were performed with Perseus software package (Tyanova et al., 2016; Ori et al., 2021).

3.2.9. Data availability

RNA sequencing data has been deposited in the Gene Expression Omnibus database under accession number GSE167011.

4. RESULTS

4.1. STEPWISE DIFFERENTIATION OF PSCS INTO EARLY LUNG PROGENITORS

The stepwise strategy used for the differentiation of human PSCs towards lung progenitor cells aimed at imitating the molecular pathways that underlie human embryonic lung development *in vitro*. In particular, lung lineage specification is induced by sequential application of developmental signals, starting with the formation of definitive endoderm (DE), followed by patterning to foregut endoderm (FE), and subsequent generation of early lung progenitors (LPs) marked by the expression of the transcription factor NKX2-1. Despite the progress in stem cell differentiation approaches towards lung epithelial cells, cell heterogeneity remains a barrier. In order to achieve high differentiation efficiency, it is crucial to begin with the generation of definitive endoderm with high purity as a basis for the subsequent protocols.

4.1.1. Generation of high purity definitive endoderm

To establish an approach for the efficient generation of definitive endoderm, I used the published differentiation approach from Huang et al. (2014) as preliminary basis, which utilizes EB outgrowths for definitive endoderm derivation, as well as from Firth et al. (2014) and Gotoh et al. (2014) that carry out monolayer cultures (**Figure 10A**). Quantification of the DE-associated cell surface markers by flow cytometry demonstrated that the approach adopted from Gotoh et al. (2014) generated high purity DE using both human ESCs and iPSCs, with 98.6% of CXCR4+, 99.1% EpCAM+ and 84% cKIT+ for H9 cells and 96% of CXCR4+, 97% EpCAM+ and 96% cKIT+ for HMGU #1 iPSCs, respectively, on day 6 of differentiation (**Figure 10B**). In order to track the expression of NKX2-1 during the subsequent steps of differentiation into epithelial lung progenitors, I used a hiPSC line integrated with the *eGFP* gene downstream to the endogenous promoter of *NKX2-1* (*NKX2-1*$^{eGFP/+}$) (Olmer et al., 2019). Given the high variation in differentiation efficiency of different human PSC lines towards desired lineages including endoderm (Huang et al., 2014), it was important to confirm the successful derivation of definitive endoderm using the *NKX2-1*$^{eGFP/+}$ hiPSC line. The protocol resulted in DE cell population with 92% of co-expressing CXCR4+/cKIT+ cells and 97.6% EpCAM+ cells, on day 6 of differentiation (**Figure 10C**). The induction efficiency was also investigated by gene expression analysis for the DE-associated transcription factors *FOXA2* and *SOX17* (**Figure 10D**) (Ori et al., 2021), which showed that all DE-markers were highly induced on that stage of differentiation. Importantly, when the differentiation conditions were applied on hiPSCs

(HMGU #1 and *NKX2-1*$^{eGFP/+}$) the addition of NEAA and GlutaMAX to the DE-basal medium (DE-BM) was required to ensure the survival of the cells (data not shown).

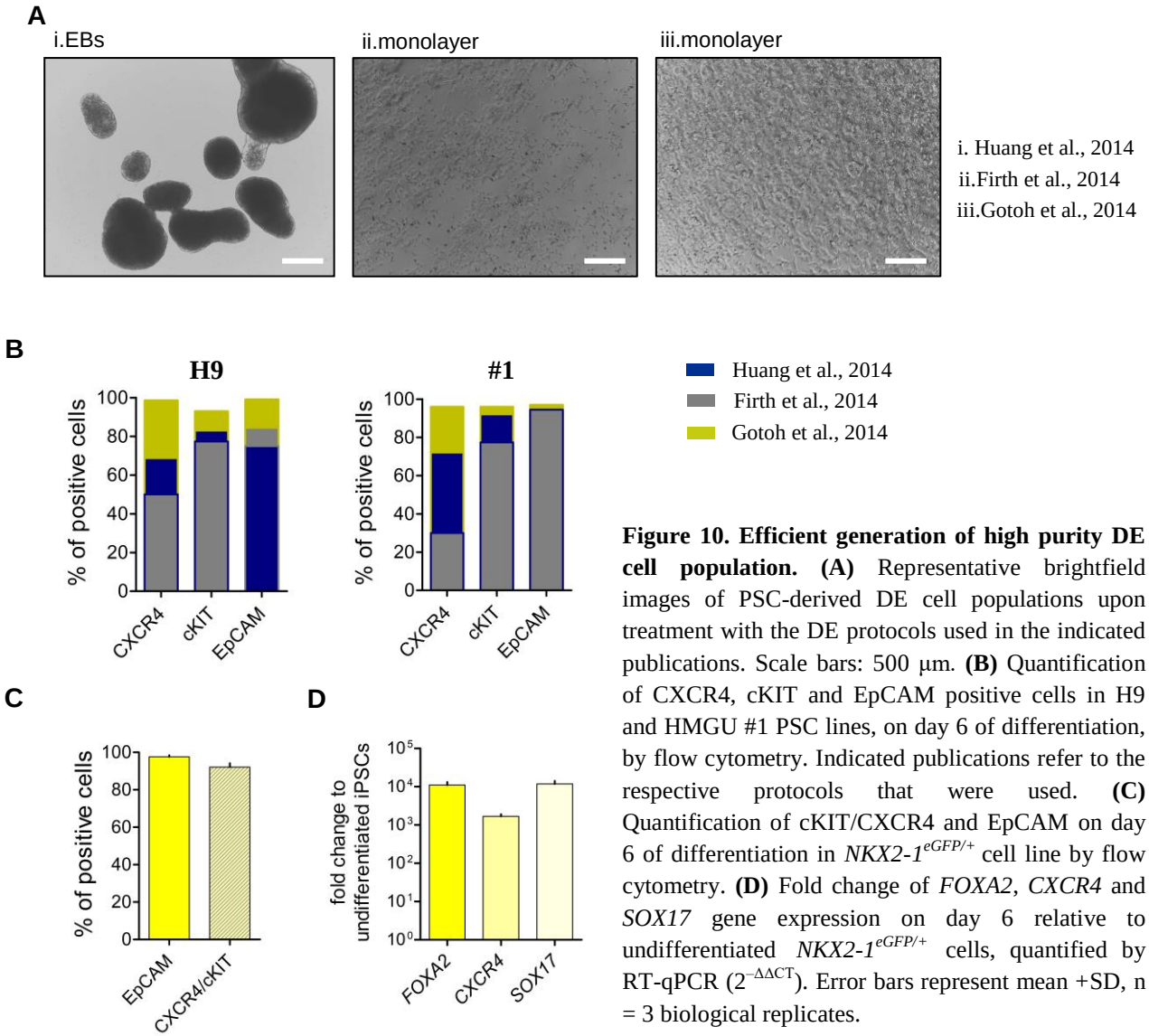


Figure 10. Efficient generation of high purity DE cell population. (A) Representative brightfield images of PSC-derived DE cell populations upon treatment with the DE protocols used in the indicated publications. Scale bars: 500 μm. **(B)** Quantification of CXCR4, cKIT and EpCAM positive cells in H9 and HMGU #1 PSC lines, on day 6 of differentiation, by flow cytometry. Indicated publications refer to the respective protocols that were used. **(C)** Quantification of cKIT/CXCR4 and EpCAM on day 6 of differentiation in *NKX2-1*$^{eGFP/+}$ cell line by flow cytometry. **(D)** Fold change of *FOXA2*, *CXCR4* and *SOX17* gene expression on day 6 relative to undifferentiated *NKX2-1*$^{eGFP/+}$ cells, quantified by RT-qPCR ($2^{-\Delta\Delta CT}$). Error bars represent mean +SD, n = 3 biological replicates.

4.1.2. Patterning into foregut endoderm and generation of eGFP+ lung progenitors

Following the induction of definitive endoderm, next step towards the respiratory cell fate was the patterning to foregut endoderm. Although it is well-established that dual inhibition of BMP and TGF-β is able to induce foregut endoderm in culture (Green et al., 2011), high cell death rate was observed upon re-plating and medium change during the transition from definitive to foregut endoderm, on day 6, followed by further complications on the next stages of differentiation and the

yields of NKX2-1+ cells. Given that sonic hedgehog signaling is important for the development of the foregut and cell survival (Fernandes-Silva et al., 2017; Litingtung et al., 1998; Motoyama et al., 1998), I hypothesized that addition of SHH from day 6 to 8 in the culture may improve the differentiation efficiency. Indeed, SHH administration improved the overall cell survival and resulted in a foregut endoderm cell population positive for the transcription factor FOXA2 (**Figure 11A-C**) (Ori et al., 2021).

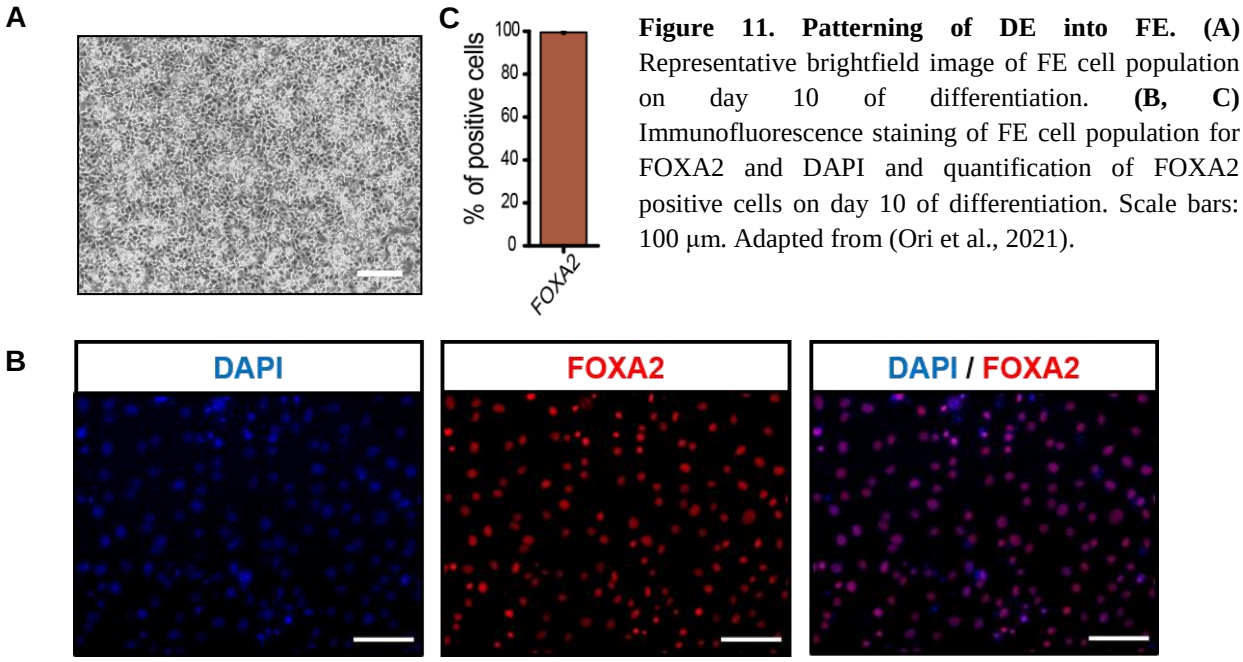


Figure 11. Patterning of DE into FE. (A) Representative brightfield image of FE cell population on day 10 of differentiation. **(B, C)** Immunofluorescence staining of FE cell population for FOXA2 and DAPI and quantification of FOXA2 positive cells on day 10 of differentiation. Scale bars: 100 µm. Adapted from (Ori et al., 2021).

Furthermore, this strategy improved the efficiency of lung differentiation by increasing the number of eGFP+ cells from day 13 on ($p = 0.07$), and combined with FGF10 at the lung progenitor stage nearly doubled the production of eGFP+ cells; from 17% to 32% on day 15 (**Figure 12A** and **B**), while the overall differentiation reproducibility reached to 80% (**Figure 12B**). Lastly, replacement of DMEM/F-12-based basal medium (FE-BM1) with the IMDM-based one (FE-BM2) led to 100% success rate, allowing the reproducibility of the modified protocol (**Figure 12B**) (Ori et al., 2021).

The transcription factor NKX2-1 is also expressed in precursors of other organs, namely, the thyroid and the forebrain (Kimura et al., 1996). Thus, it was essential to exclude the presence of those undesired lineages in the culture. Gene expression analysis confirmed the low expression of

PAX8 and *PAX6*, thyroid and brain specific markers respectively, in the modified protocol, followed by slight upregulation of *SFTPB* and *MUC5AC* (**Figure 12C**) (Ori et al., 2021). Interestingly, the eGFP signal was localized into small cell clusters, independently of the cell line used (**Figure 12D** and **E**).

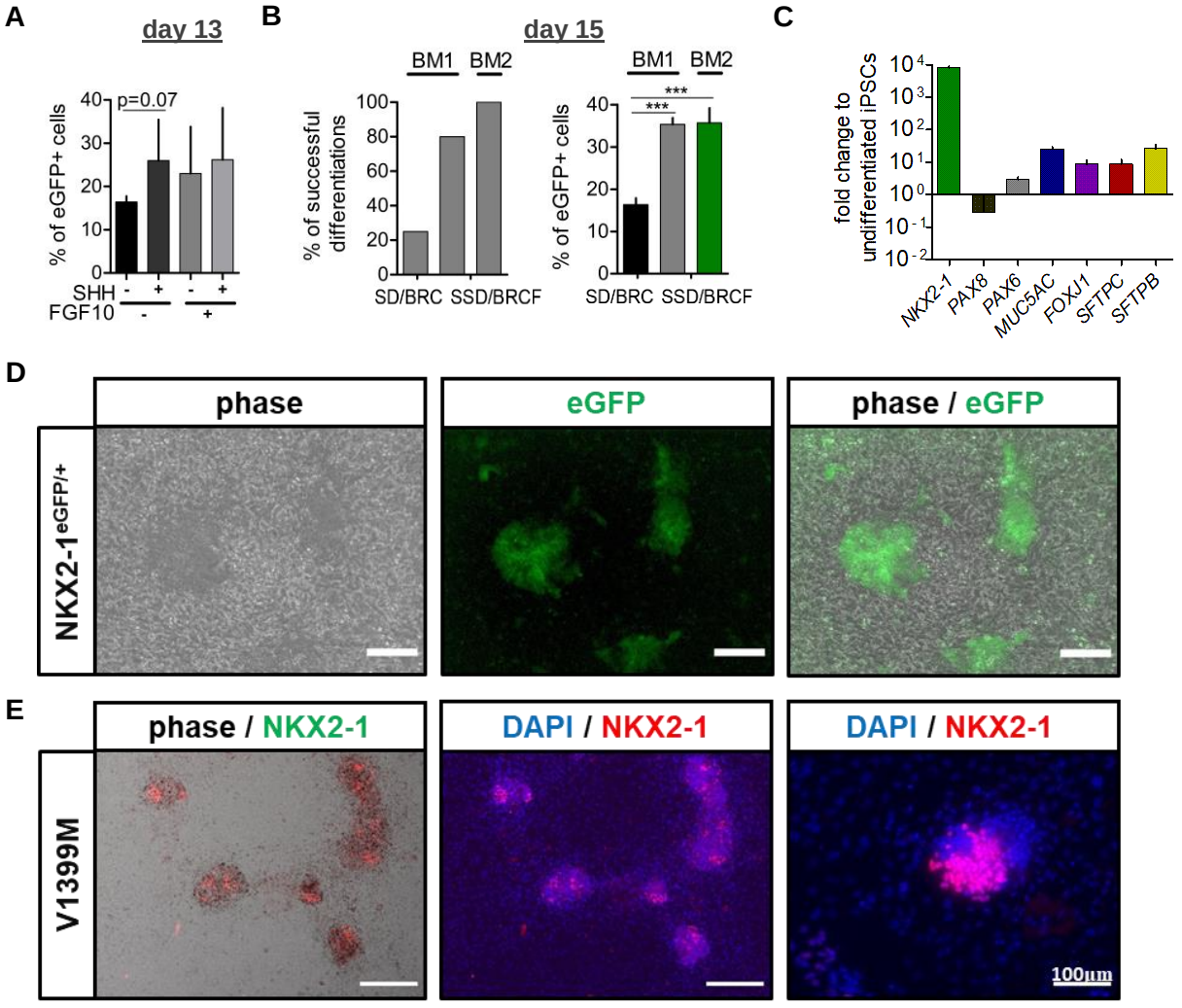
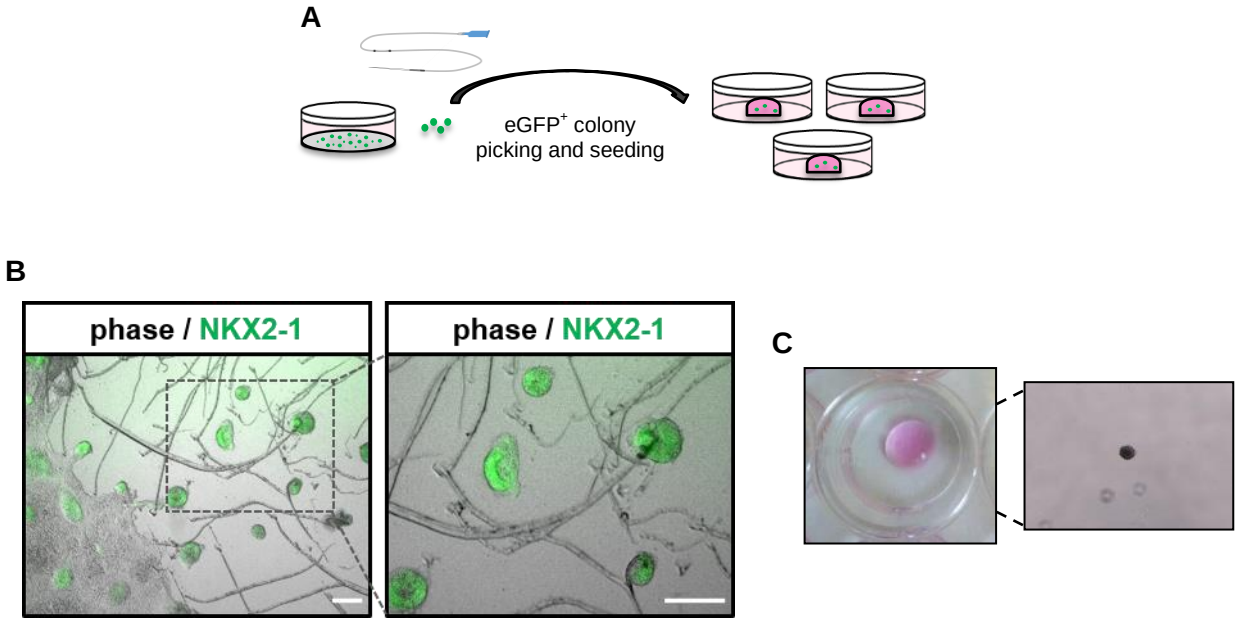


Figure 12. Differentiation of hPSCs to lung progenitors. (A, B) Quantification of eGFP+ cells analyzed on days 13 and 15 of differentiation by flow cytometry, and percentage of successful differentiation experiments upon ± SHH and ± FGF10 treatments and culture in defined basal media FE-BM1 or -BM2 respectively. Error bars represent mean +SD, n = 3 biological replicates, ***p < 0.001 unpaired, one-tailed t-test. Media composition is described in **Table 8** (Ori et al., 2021). **(C)** Fold change of *NKX2-1*, *PAX8* and *PAX6* as well as *MUC5AC*, *FOXJ1*, *SFTPC* and *SFTPB* gene expression, on day 15 relative to undifferentiated *NKX2-1*$^{eGFP/+}$ cells, quantified by RT-qPCR ($2^{-\Delta\Delta CT}$). Bars represent mean +SD, n = 3. **(D)** Fluorescence microscopy of eGFP+ colonies, on day 15 of differentiation. Scale bars: 200 µm. **(E)** Immunofluorescence staining for NKX2-1 and DAPI on day 15 of differentiation of V1399M hiPSC line differentiation. The patient-derived iPSC line for the ABCA3 mutation V1399M was generated at Helmholtz Center Munich (courtesy of Dr. med. Maria Forstner). Scale bars: 200 µm and 100 µm.

4.2. NKX2-1 POSITIVE CELLS ARE FUNCTIONAL MULTIPOTENT PROGENITORS

4.2.1. Lung organoid generation from NKX2-1+ lung progenitors

Having established a method that allows the reproducible differentiation of hPSCs to NKX2-1+ cells, the next step was to investigate the potential of the eGFP+ cells to differentiate further and give rise lung organoids (LOs) from multiple respiratory lineages, with a great interest to distal, alveolar type II (ATII) cells. As shown in representative fluorescence microscopic images of *NKX2-1*$^{eGFP/+}$ differentiated cells on day 15, the eGFP+ colonies demonstrated a convex shape (**Figure 13B**), providing an advantage for isolation and manipulation; allowing not only to easily distinguish the NKX2-1+ cells from the surroundings, but also enabling pure dissection and picking of the desired colonies manually (**Figure 13A** and **B**). On day 15 of differentiation, I re-seeded eGFP+ colonies into 3D Matrigel culture as clumps (**Figure 13A-C**), and as demonstrated in **Figure 13D**, they were able to expand and form spheroid-shaped outgrowths in serum-free basal media, supplemented with CHIR99021, KGF, and FGF10 (days 15-22). Also, the eGFP+ spheroids tripled their size in one week (days 15 to 22) maintaining their spherical shape (**Figure 13D**).



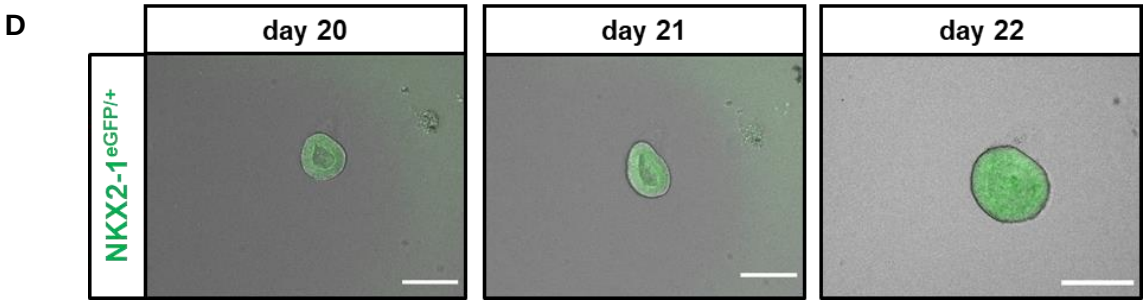
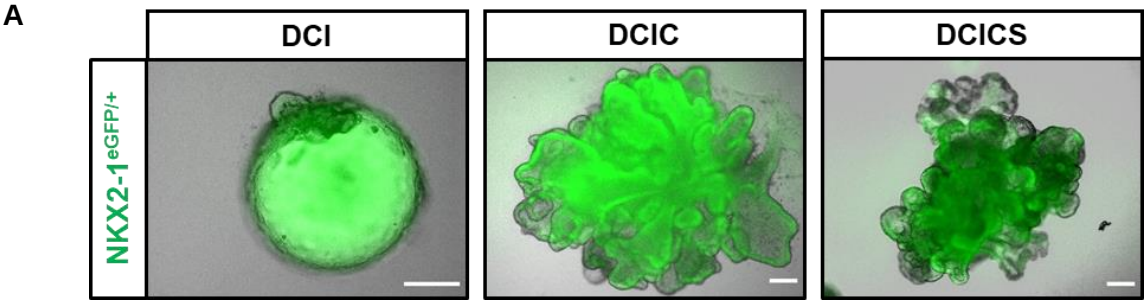


Figure 13. Expansion of eGFP+ outgrowths cultured in defined media. (A) Schematic illustration of eGFP+ colony picking and re-seeding in 3D Matrigel culture on day 15 of differentiation. **(B)** Representative fluorescence images of eGFP+ colonies upon dissection. **(C)** Representative Matrigel drop with seeded outgrowth on day 15 of differentiation. **(D)** Fluorescence images of representative outgrowths on days 20–22 of differentiation. Scale bars: 200 μm.

Aiming to induce further patterning and maturation of the LP-derived outgrowths, the culture conditions were switched on day 22 of differentiation to 1) Pneuma air-liquid interface (ALI)-culture medium or 2) serum-free basal medium, enriched with dexamethasone, cAMP and IBMX (DCI) until day 35, factors which are known to promote maturation of the fetal lungs, with or without addition of CHIR99021 (C) and SB431542 (S). The CHIR99021-treated LOs exhibited branching morphogenesis and only when combined with SB431542 (DCICS), upregulation of the alveolar type II marker *SFTPC* was induced, with the organoids reaching 1.6 ± 0.16 mm in size (**Figure 14A-C**) (Ori et al., 2021). In contrast, in the absence of these stimuli, WNT/β-catenin activation and TGF-β pathway inhibition, the DCI-induced organoids expressed genes characteristic of more proximal regions of the lung, such as *SCGB1A1* a marker of club cells, and when treated in PneumaCult-ALI conditions they produced secretions, while positive for *MUC5AC*, a marker of bronchial goblet cells (**Figure 14B** and **D**). Immunostaining for DCI ± CS supported my previous observations with cells expressing SFTPC and SFTPB proteins in DCICS conditions and SCGB1A1, MUC5AC but not SFTPC in absence of CHIR99021 and SB431542 (**Figure 14E** and **F**), demonstrating the multipotent progenitor capacity of the NKX2-1+ cells, which can differentiate into different lung cell types depending on the culture conditions (Ori et al., 2021).



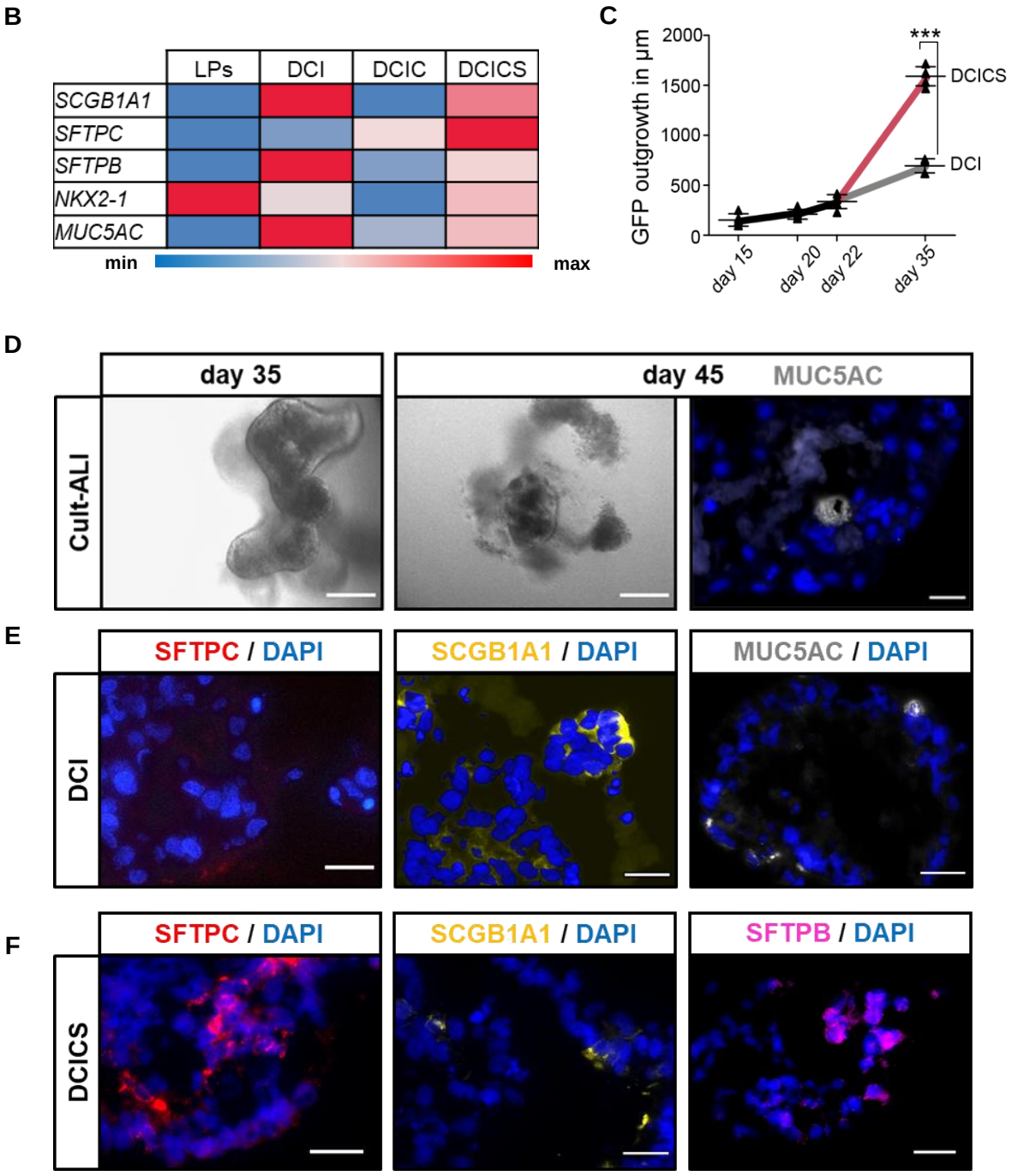


Figure 14. Characterization of hPSC-derived lung organoids. (A) Fluorescent images and **(B)** Gene expression analysis of lung markers *NKX2-1*, *SFTPB*, *SFTPC*, *SCGB1A1*, *MUC5AC* on day 35 of differentiation, upon indicated treatment (quantified by RT-qPCR $2^{-\Delta CT}$, n=5 biological replicates; DCI: dexamethasone, cAMP, IBMX, CS: CHIR99021, SB431542). Scale bars: 200 µm. **(C)** Size measuring of DCI and DCICS organoids upon seeding in 3D culture, from day 15 to day 35 in µm. Bars represent mean ± SD, n = 5 biological replicates; ***p < 0.001 by unpaired, one-tailed t-test. **(D)** Brightfield images (scale bars: 200 µm) and immunofluorescence staining for MUC5AC (scale bar: 20 µm) of Cult-ALI treated organoids on days 35 and 45 of differentiation. **(E, F)**. Immunofluorescence staining of day 35 DCI and DCICS lung organoids, for NKX2-1, SFTPC, SFTPB, SCGB1A1 and MUC5AC on day 35 of differentiation. Scale bars: 20 µm. Adapted from (Ori et al., 2021).

4.2.2. Transcriptomic characterization of lung epithelial cells

To gain further information about the molecular profile of lung progenitors and the lineages generated in differentiation cultures, I analyzed the transcriptomes of undifferentiated iPSCs (day 0), sorted eGFP+ (day 15), sorted eGFP- (day 15) and lung organoids (day 35) (**Figure 15A** and **B**) (Ori et al., 2021). Analysis revealed that eGFP+ cells expressed genes which are known to be important for the formation of the lung foregut, namely *FOXA1, FOXA2, FOXP1, IRX3, MECOM,* and *NKX2-1* (**Figure 15C**) (Ori et al., 2021). These genes were also expressed in DCICS-treated organoids, which additionally exhibited markers that are typical of early and late branching morphogenesis and cell commitment of the distal lung, including *RUNX1, MUC1, SFTPC, SFTPB, CLDN18 and NAPSA* (Beauchemin et al., 2016; Hurley et al., 2020) (**Figure 15C**), representing a more advanced developmental stage (Ori et al., 2021). Consistently, analysis of gene ontology (GO) terms over undifferentiated iPSCs revealed enrichment of pathways involved in embryonic morphogenesis, endodermal cell differentiation and development of respiratory system in eGFP+ progenitors, as well as respiratory tube and alveolus development in DCICS-treated organoids, respectively (**Figure 15D**) (Ori et al., 2021).

I next sought to compare the iPSC-derived early respiratory cells with their *in vivo* counterparts. Despite the difficulty of studying human embryonic lung organogenesis due to limited access to fetal tissue, Nikolić et al. (2017) were able to analyze human fetal lungs and identify key transcription factors that are exclusively expressed either in the tips or the stalk of the lung, or in both regions. Comparison of this data generated from human primary tissue to the iPSC-derived organoids that I produced, indicated that DCICS-organoids showed higher resemblance to alveolar tips of human fetal lungs relative to the stalk (**Figure 15E**) (Ori et al., 2021). On the other hand, consistent with my previous gene and protein expression analysis (**Figure 14E** and **F**), DCI-treated organoids resembled to a more proximal stalk-like identity (**Figure 15E**) (Ori et al., 2021). Furthermore, multipotent eGFP+ lung progenitors were highly correlated with the gene signature of the fetal lung tips compared to the stalk (**Figure 15E**). This is a very interesting finding which parallels human *in vivo* studies, implying that similar to mouse, human embryonic lung epithelial tips constitute multipotent progenitors that have self-renewal capacity and can be expanded *in vitro* (Nikolić et al., 2017).

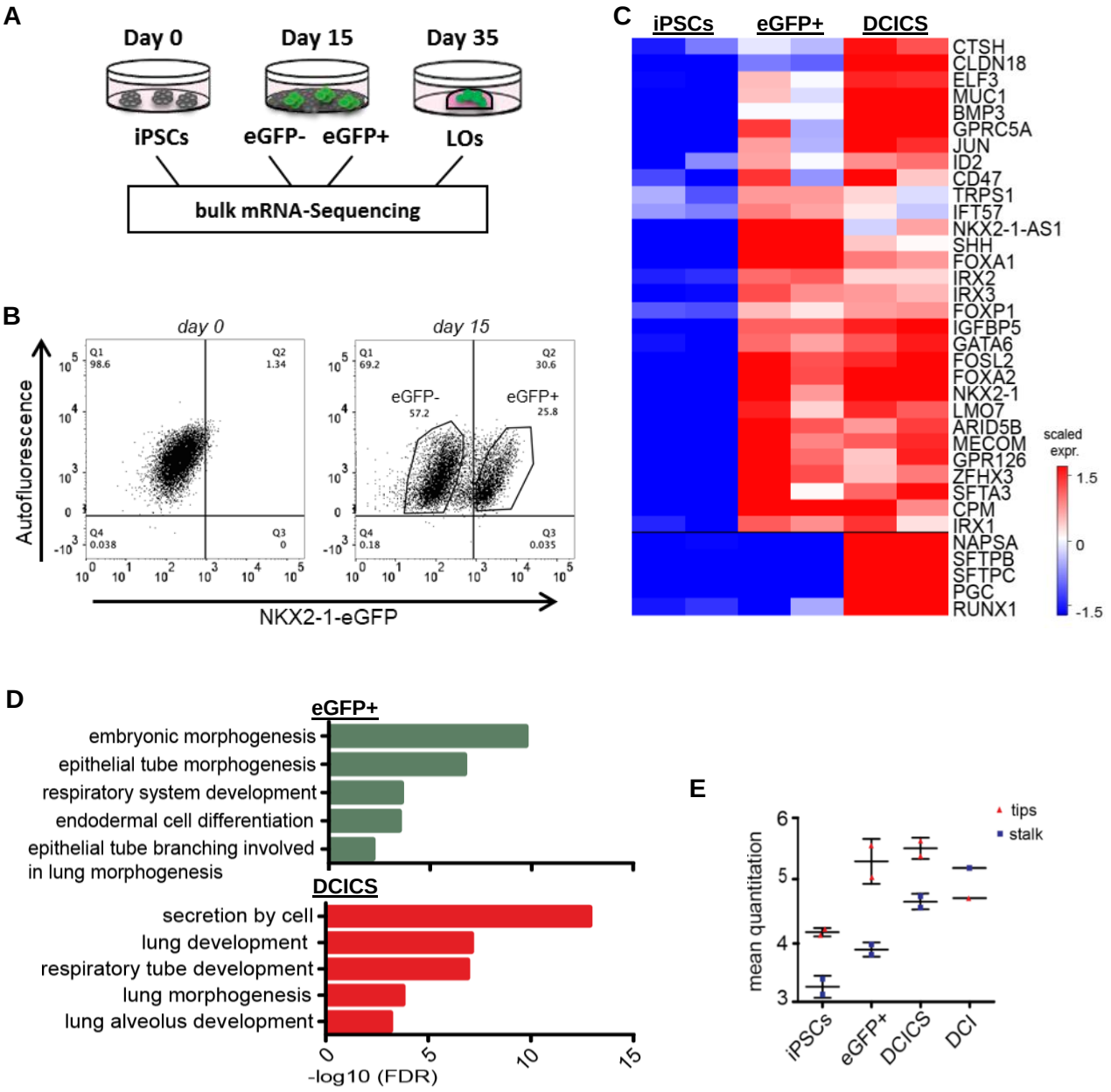


Figure 15. Transcriptomic characterization of iPSC-derived lung lineages. (A) Schematic illustration of cell sample collection for bulk mRNA-sequencing experiment. **(B)** Representative flow cytometry plots of *NKX2-1*$^{eGFP/+}$ cells on day 15 of differentiation (right panel) with gates indicating the strategy used for sorting eGFP+ and eGFP- cells, respectively (undifferentiated cells were used as negative controls, left panel). **(C)** Heatmap representing the expression of selected genes in undifferentiated cells, eGFP+ cells (day 15) and DCICS (day 35) based on bulk RNA-Seq analysis (DESeq2, p-value < 0.05, scale = row normalized log2 expression). **(D)** Gene Ontology (GO) enrichment analysis of differentially expressed genes in eGFP+ lung progenitors and DCICS organoids, relative to undifferentiated *NKX2-1*$^{eGFP/+}$ cells, as defined by DESeq2 statistical test (p-value < 0.05). **(E)** Comparison of generated LPs and LOs to human *in vivo* study. Star war plot displaying the mean quantitation for each data store correlated to tip (red) and stalk (blue) gene signatures provided by Nicolic et al., 2017. Adapted from (Ori et al., 2021).

4.2.3. Extended culture of lung organoids promotes further maturation

I also sought to examine whether the organoids were able to maintain their distal-lung like identity upon passaging and subsequent culture for a longer period. Interestingly, passaged lung DCICS-treated organoids not only maintained the gene expression signature which is representative for the distal alveoli-like regions, namely *SFTPC, SFTPB, NAPSA,* but they also expressed additional genes relevant to saccular lung developmental stage such as *HOPX*, *LAMP3*, *BHLHE40* and *CTSH* (Beauchemin et al., 2016). Furthermore, the organoids maintained the eGFP expression and self-renewal capacity (**Figure 16A-C**).

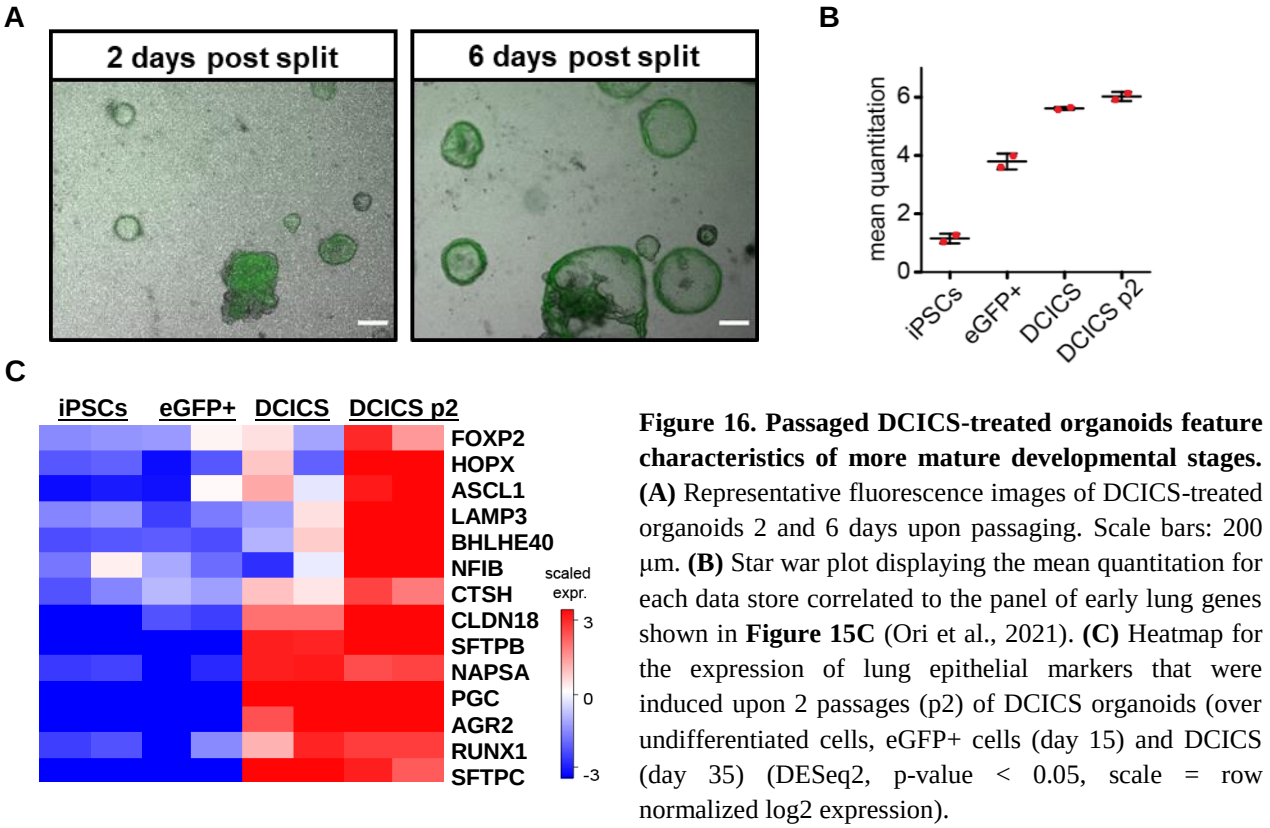


Figure 16. Passaged DCICS-treated organoids feature characteristics of more mature developmental stages. (A) Representative fluorescence images of DCICS-treated organoids 2 and 6 days upon passaging. Scale bars: 200 µm. **(B)** Star war plot displaying the mean quantitation for each data store correlated to the panel of early lung genes shown in **Figure 15C** (Ori et al., 2021). **(C)** Heatmap for the expression of lung epithelial markers that were induced upon 2 passages (p2) of DCICS organoids (over undifferentiated cells, eGFP+ cells (day 15) and DCICS (day 35) (DESeq2, p-value < 0.05, scale = row normalized log2 expression).

These observations led to the hypothesis that passaged DCICS-treated lung organoids might represent a more mature stage of the respiratory development. To gain more information, I correlated the transcriptome of the hiPSC-derived lung organoids with previously published human transcriptome datasets describing the gene signatures of the first trimester (T1) of pregnancy, the second trimester and adult (T2/adult) (Roost et al., 2015). Excitingly, expression profiles of later passage lung organoids resembled to more advanced developmental stages of the lung (**Figure 17A**). Consistently, gene signature of early developmental stages (53 days post conception, dpc) is rather

represented by the early lung progenitors at 15 day of differentiation, while the DCICS passaged organoids represent more mature lung tissues when compared to human fetal lung tissues spanning various developmental time points between 53 to 140 dpc (Kho et al., 2010) (**Figure 17B**).

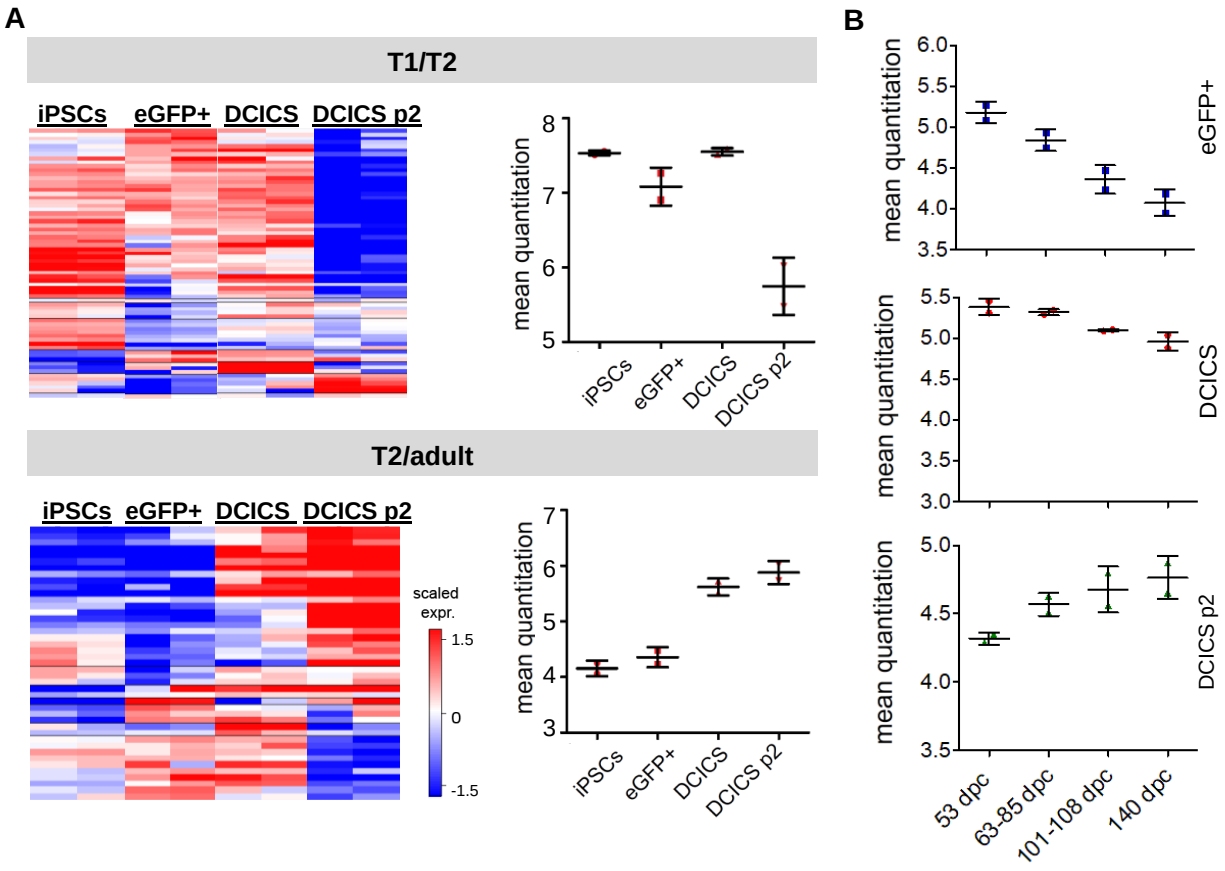


Figure 17. Correlation of PSC-derived lung organoids to maturation level of *in vivo* developed lungs. (A) Heatmaps for the expression of genes that were induced in the trimesters T1/2 and T2/adult, at respective stages of the differentiation protocol (DESeq2 was used for analysis, p-value < 0.05, T1 and T2 represent first and second trimester of pregnancy respectively (Roost et al., 2015), scale = row normalized log2 expression). Star war plots represent the mean quantitation for each data store (x-axis) correlated to the panel of genes in the respective heatmap for T1/2 and T2/adult respectively. **(B)** Star war plots show each data store against the genes expressed in respective developmental stages (with average mean expression >7.0). Data has been derived from (Kho et., 2010).

4.3. HEPATOCYTE PROGENITORS ARE GENERATED DURING IPSCS DIFFERENTIATION TOWARDS LUNG PROGENITORS

Besides, I sought to interrogate the identity of the non-lung lineages (eGFP-) emerging in the culture on day 15 of differentiation. Transcriptional analyses revealed that the eGFP- population (**Figure 15B**) featured genes that are expressed in fetal liver including apolipoproteins (e.g. *APOA1*, *APOB*), fibrinogens (*FGB)*, and plasma protein alpha fetoprotein *AFP* (Carpentier et al., 2014; Hannan et al., 2013), and GO term analysis revealed enrichment of processes associated with fetal liver development, regeneration, and metabolism (**Figure 18A** and **B**) (Ori et al., 2021). The existence of liver-like cells in the culture was also confirmed by immunostaining for the liver specific gene AFP (**Figure 18C**) (Ori et al., 2021). Importantly, mutually exclusive expression of eGFP and AFP indicated separate clusters of lung progenitors and hepatoblasts (HB), respectively (**Figure 18C**) (Ori et al., 2021). Overall, we concluded lung and liver fates were co-induced during the differentiation, raising the question of the mechanisms underlying the mutually exclusive specification of these lineages from foregut endoderm.

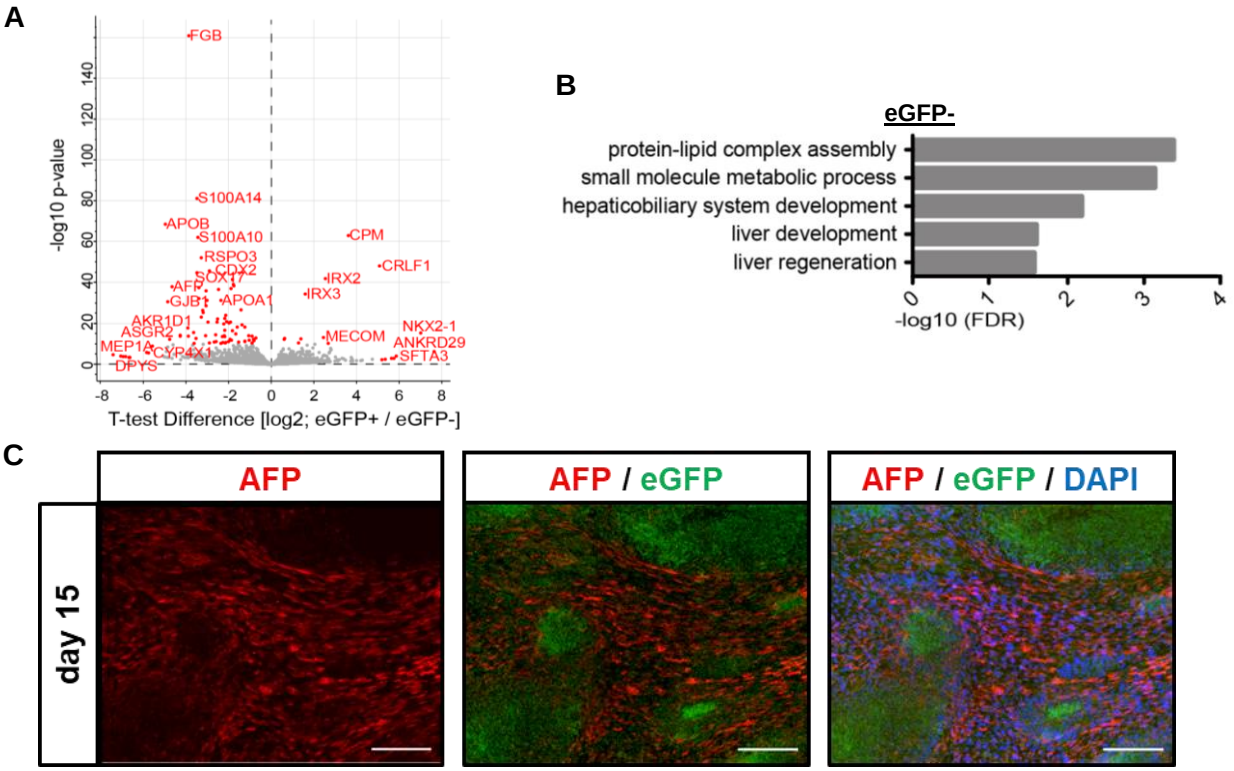


Figure 18. Characterization of eGFP- cells reveals the existence of hepatocyte progenitors in the culture. (A) Volcano plot showing differentially expressed genes between sorted eGFP+ and eGFP- populations. **(B)** GO terms for eGFP- (over eGFP+ as defined by DESeq2 statistical test, p-value < 0.05). **(C)** Representative immunofluorescence staining for AFP and DAPI, and imaging of eGFP+ cells on day 15 of differentiation. Scale bars: 100 µm. Adapted from (Ori et al., 2021).

4.4. INVESTIGATION OF THE REPROGRAMMING CAPACITY OF NKX2-1 BY ECTOPIC EXPRESSION IN PSC AND SECONDARY FIBROBLASTS

Given that NKX2-1 is established as the key transcription factor driving lung organogenesis (Kimura et al., 1996; Minoo et al., 1999), I sought to investigate whether ectopic expression of *NKX2-1* can promote respiratory specification from human pluripotent and somatic cells. To achieve ectopic expression of *NKX2-1*, I constructed PiggyBac (PB)-based tet-inducible NKX2-1 vectors and integrated them into human PSCs (Ding et al., 2005). For the ectopic expression of NKX2-1 in ESCs an eGFP reporter was used, whereas a mCherry reporter was used in *NKX2-1*$^{eGFP/+}$ iPS cell line that allows to distinguish the endogenous NKX2-1 from the ectopic NKX2-1(-mCherry) expression upon doxycycline (dox) treatment, resulting in the generation of *iNKX2-1-eGFP-H9* and *iNKX2-1-mCherry-NKX2-1*$^{eGFP/+}$ cell lines, respectively **(Figures 19** and **20)**.

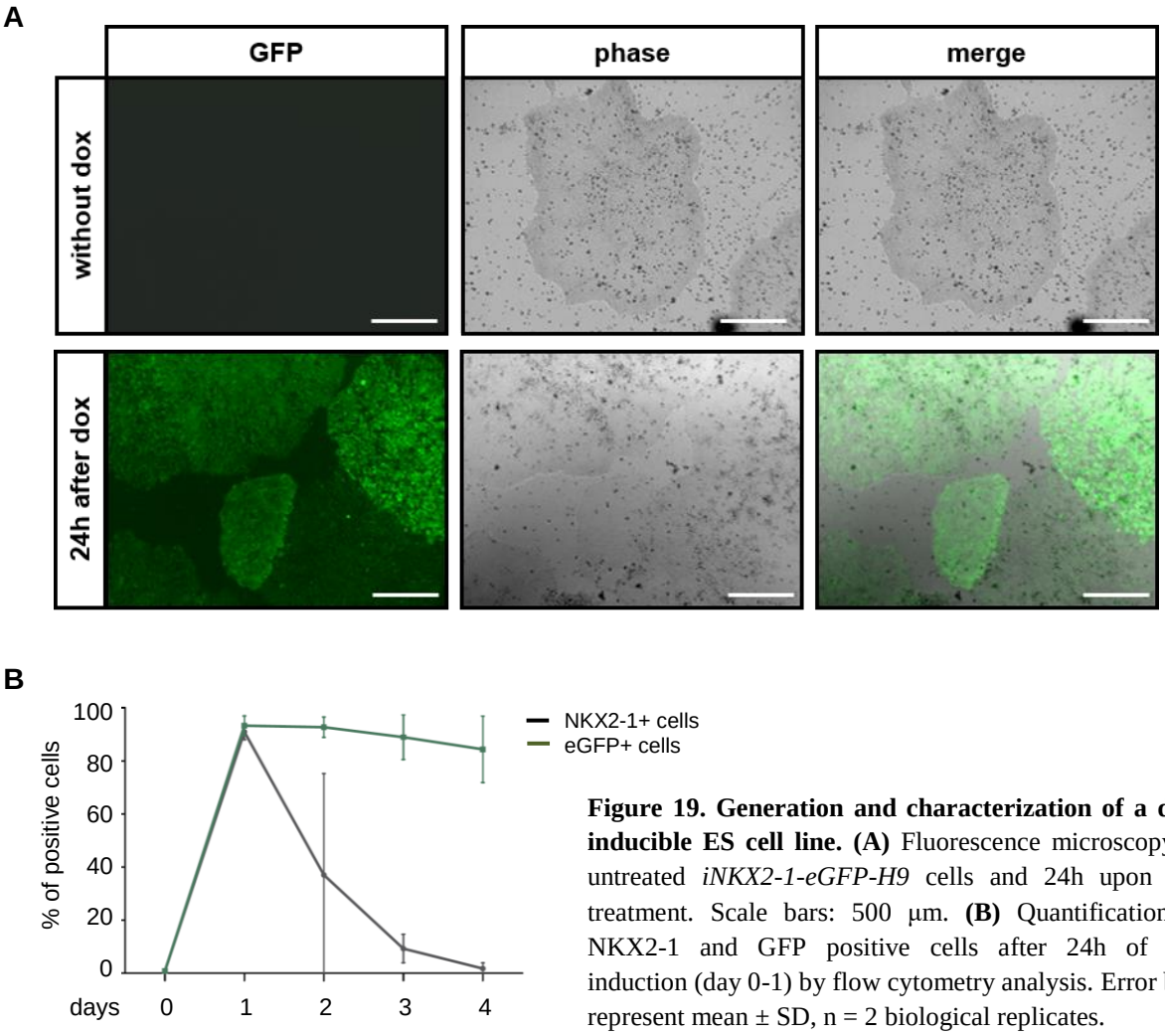


Figure 19. Generation and characterization of a dox-inducible ES cell line. (A) Fluorescence microscopy of untreated *iNKX2-1-eGFP-H9* cells and 24h upon dox treatment. Scale bars: 500 μm. **(B)** Quantification of NKX2-1 and GFP positive cells after 24h of dox induction (day 0-1) by flow cytometry analysis. Error bars represent mean ± SD, n = 2 biological replicates.

Characterization of the inducible cell lines by fluorescence microscopy, immunostaining, flow cytometry, and gene expression analysis demonstrated the efficient overexpression of NKX2-1 upon dox induction (**Figures 19A** and **20A**). Also, exogenous *NKX2-1* expression drops rapidly 24h upon dox withdrawal and after 48h it is completely diminished (**Figure 19B** and **20B**), whereas the expression of the reporters mCherry and eGFP persisted for several days (**Figure 19B** and **20B**).

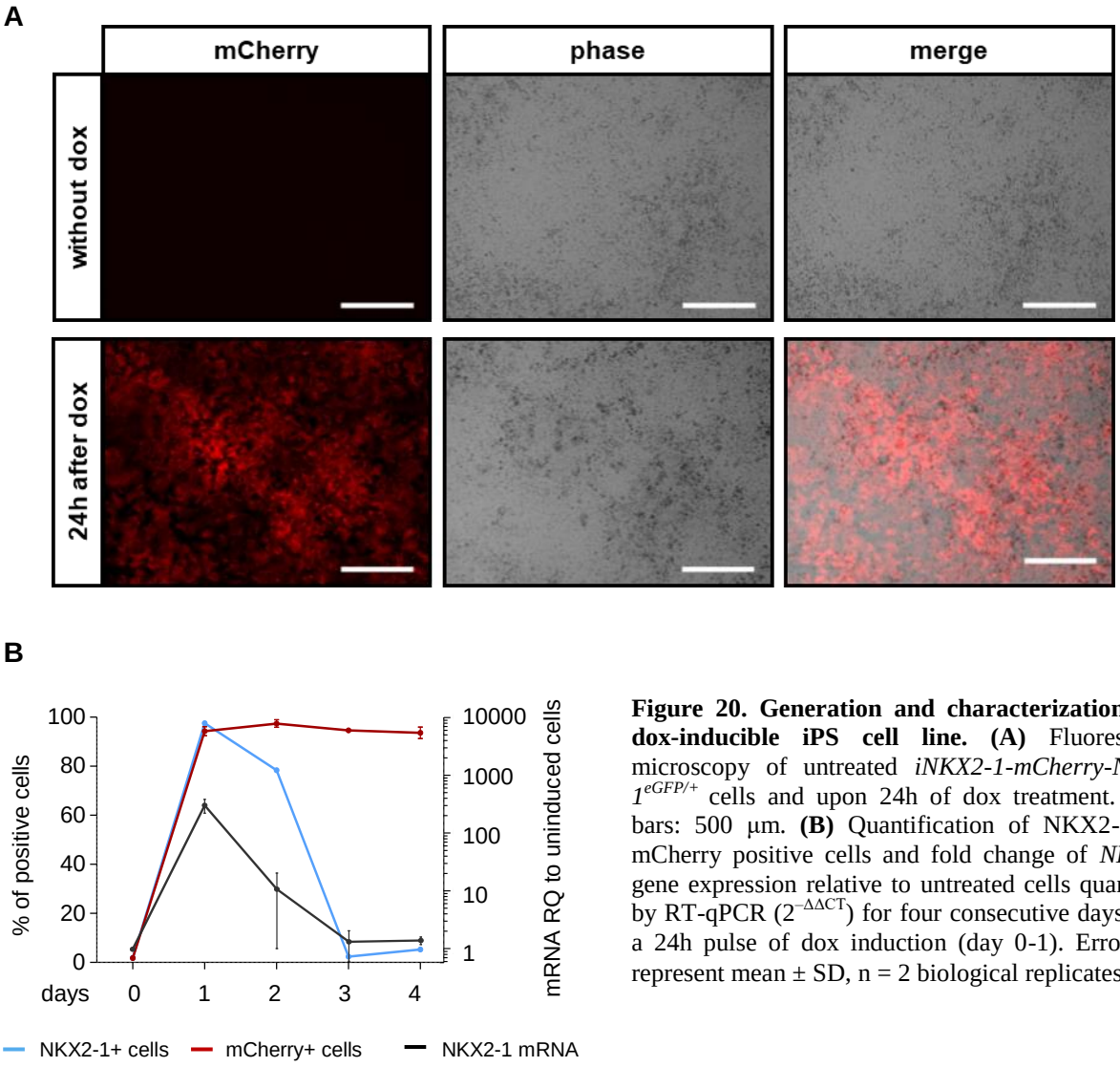


Figure 20. Generation and characterization of a dox-inducible iPS cell line. (A) Fluorescence microscopy of untreated *iNKX2-1-mCherry-NKX2-1*$^{eGFP/+}$ cells and upon 24h of dox treatment. Scale bars: 500 µm. **(B)** Quantification of NKX2-1 and mCherry positive cells and fold change of *NKX2-1* gene expression relative to untreated cells quantified by RT-qPCR ($2^{-\Delta\Delta CT}$) for four consecutive days upon a 24h pulse of dox induction (day 0-1). Error bars represent mean ± SD, n = 2 biological replicates.

Last, we verified the pluripotency of generated cell lines by immunostaining and gene expression analysis for the pluripotency factors OCT4, NANOG and SOX2 (**Figures 21A** and **B**), and observed that the differentiation potential has not be affected resulting in 33% of LPs on day 15 of differentiation, similar to the non-transfected counterparts (**Figure 21C**).

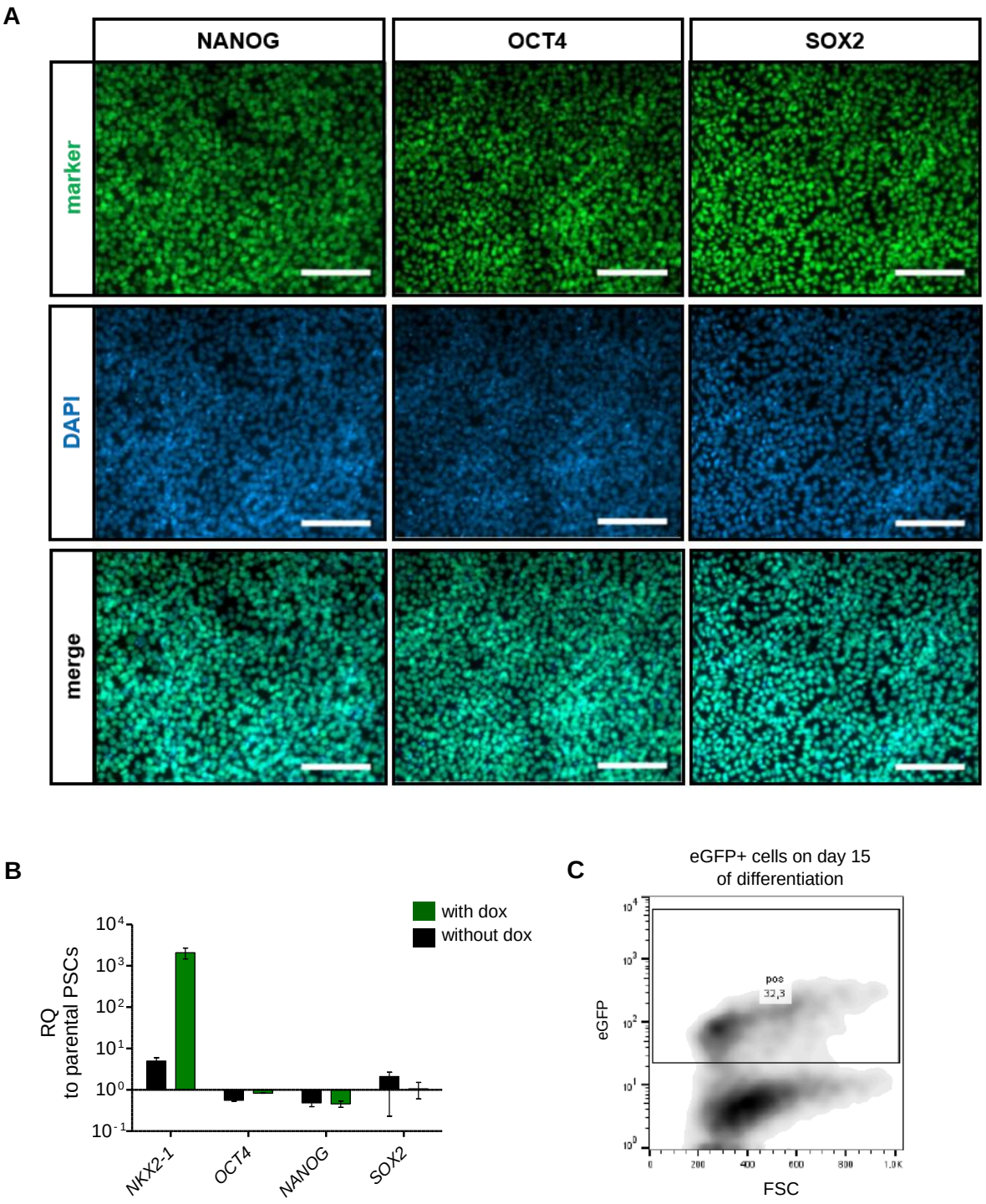


Figure 21. Characterization of dox-inducible cell lines for pluripotency and differentiation potential. (A) Representative immunofluorescence staining of *iNKX2-1-mCherry-NKX2-1*$^{eGFP/+}$ cells for the pluripotency factors NANOG, OCT4, SOX2 and DAPI overlay in untreated cells. Scale bars: 150 µm. **(B)** Fold change in gene expression of *NKX2-1*, *OCT4*, *NANOG and SOX2* relative to parental PSCs quantified by RT-qPCR ($2^{-\Delta\Delta CT}$), upon treatment with or without dox. Bars represent mean ± SD, n = 2 biological replicates. **(C)** Quantification of NKX2-1 positive cells on day 15 of lung differentiation.

4.4.1. Effect of ectopic expression of NKX2-1 in hPSCs

To accelerate the lung differentiation protocol using the overexpression cell line *iNKX2-1-mCherry-NKX2-1*$^{eGFP/+}$ as a tool, I tested two different media combinations, namely CHIR99021, KGF (FGF7), FGF10 (CKF), which was used for the maintenance and expansion of LPs (Hawkins et al., 2017), and CKF supplemented with BMP4 and RA (CKF+BR), which are required for the induction of NKX2-1 in regular PSC differentiation protocol. Furthermore, I sought to assess whether a treatment with VPA would increase the expression of the NKX2-1 target genes, namely *SFTPC*, *SFTPB*, and *MUC5AC*, by increasing chromatin accessibility to this transcription factor (Cheng et al., 2014; Huangfu et al., 2008; Kogiso et al., 2013; Thoma et al., 2014). For the accelerated derivation of lung progenitor-like spheroids via ectopic expression of NKX2-1, I started the protocol from day 15-like conditions of the step wise differentiation on, corresponding to the LP stage, when NKX2-1 is endogenously expressed; and proceeded with the 3D outgrowth cultures. NKX2-1 target genes *SFTPC* and *MUC5AC* were upregulated with two days of dox induction in CKF+BR media, whereas *SFTPB* was upregulated only upon continuous dox administration in CKF+BR (**Figure 22**). In addition, the expression of these genes was increased on average in samples treated with VPA, with the gene *SFTPB* being the most affected. Interestingly, dox induction of 8 days did not result in a significant upregulation of the indicated genes (**Figure 22**). Taken together, two days of dox induction combined with the medium composed of CKF+BR and VPA administration and was the most promising culture condition for the follow-up experiments.

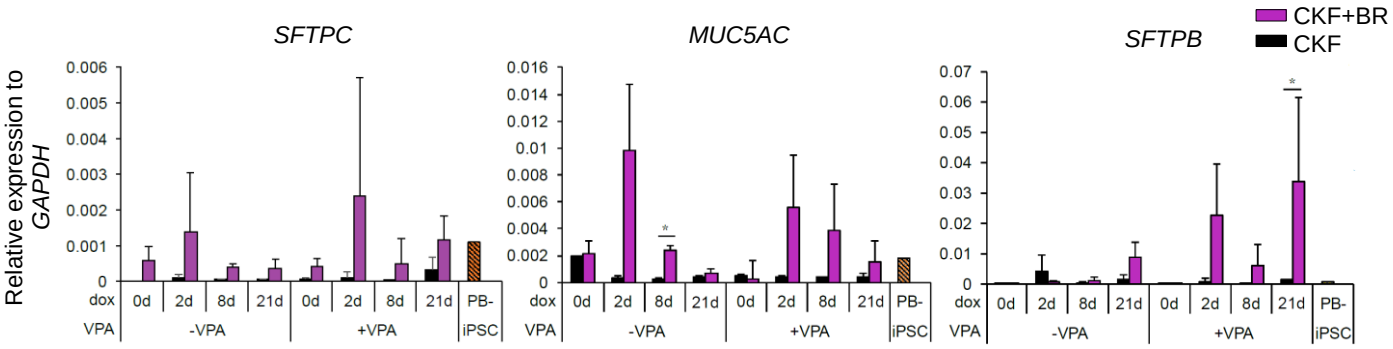
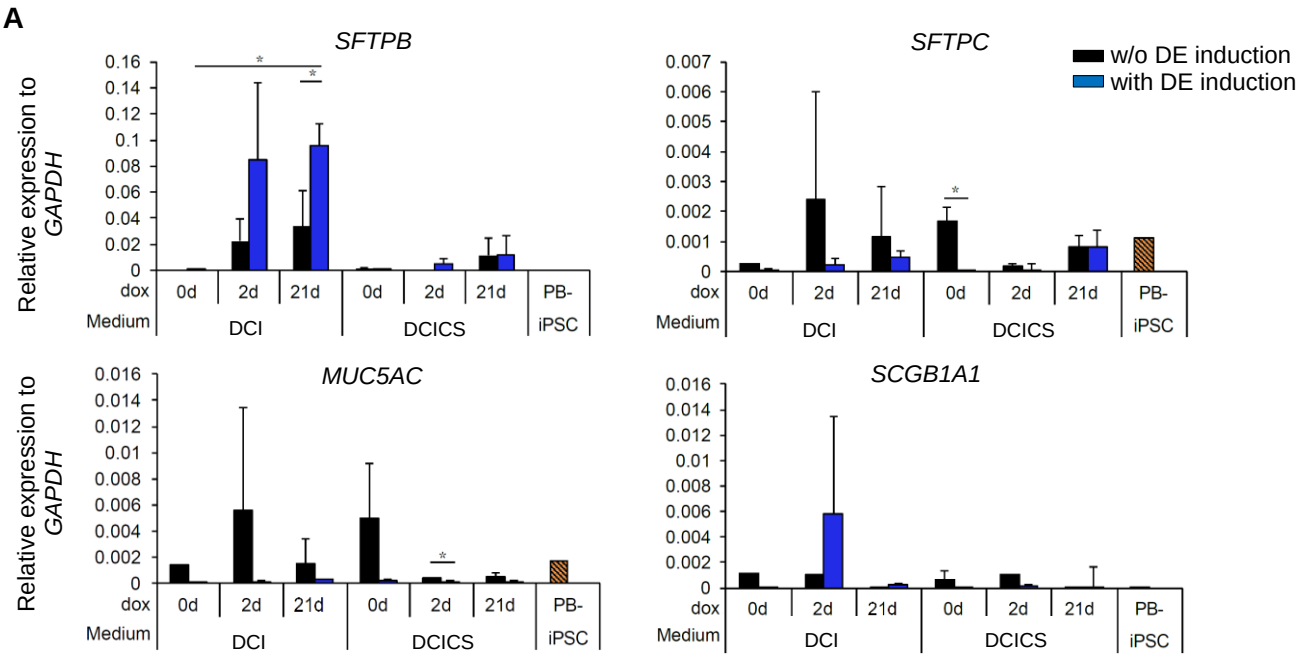


Figure 22. Characterization of spheroids generated by the ectopic expression of *NKX2*-1 and upon various dox-induction intervals. Gene expression analysis of *SFTPB, SFTPC* and *MUC5AC* upon ± 48h VPA treatment and dox induction intervals for 0, 2, 8 and 21 days, respectively, normalized to *GAPDH* and quantified by RT-qPCR ($2^{-\Delta CT}$). Error bars represent the +SD, of n=3 biological replicates for induced cells, and n=2 biological replicates for uninduced cells *$p < 0.05$.

Next, I sought to examine the effect of maturation media DCI±CS on the expression of lung epithelial markers and assess whether a combination with a short induction of DE prior to the acceleration experiment could increase the efficiency of differentiation. Intriguingly, no significant change was observed in gene expression levels, with or without DE pre-induction in DCICS treated samples. In contrast, *SFTPB* expression increased significantly when DCI medium was used in combination with DE pre-induction, and upon 2 days of dox or continuous (21 days) dox administration (**Figure 23A**). DE pre-induction had a remarkable effect also on organoid morphology, namely DE pre-treatment induced organoid expansion, resulting in a sphere-like morphology, as well as in endogenous NKX2-1 expression, a hallmark of regular differentiation when DCI maturation is used (**Figure 23B**). On the other hand, endodermal pre-induction had no substantial inductive effect on *SFTPC* and *MUC5AC* expression.



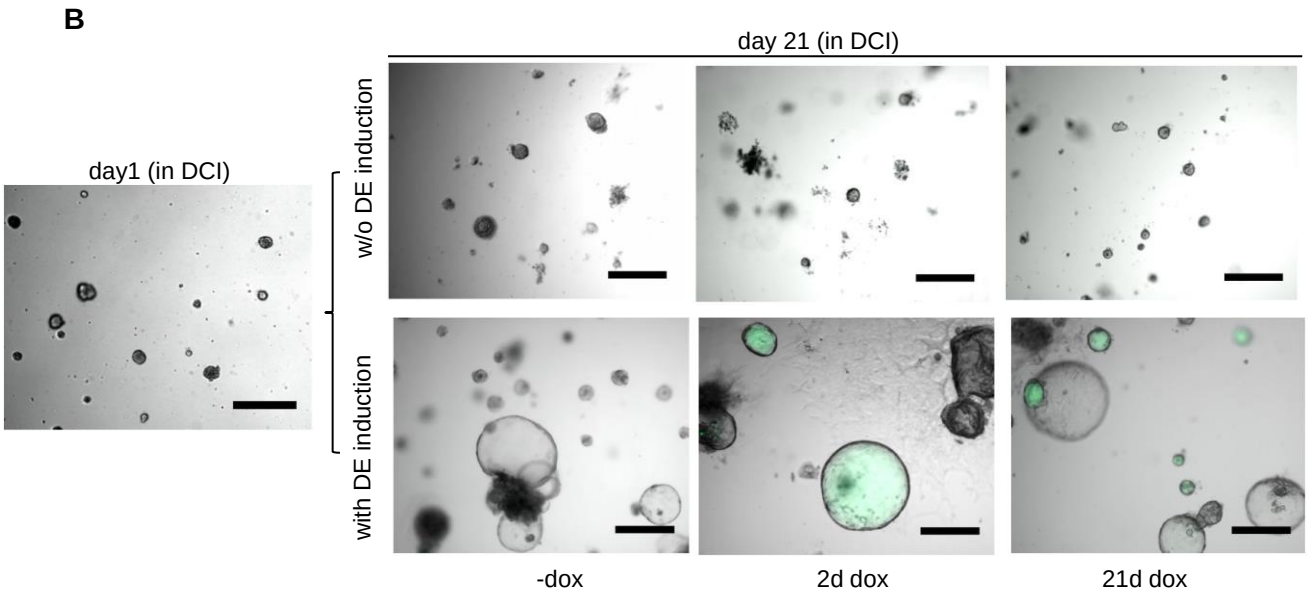
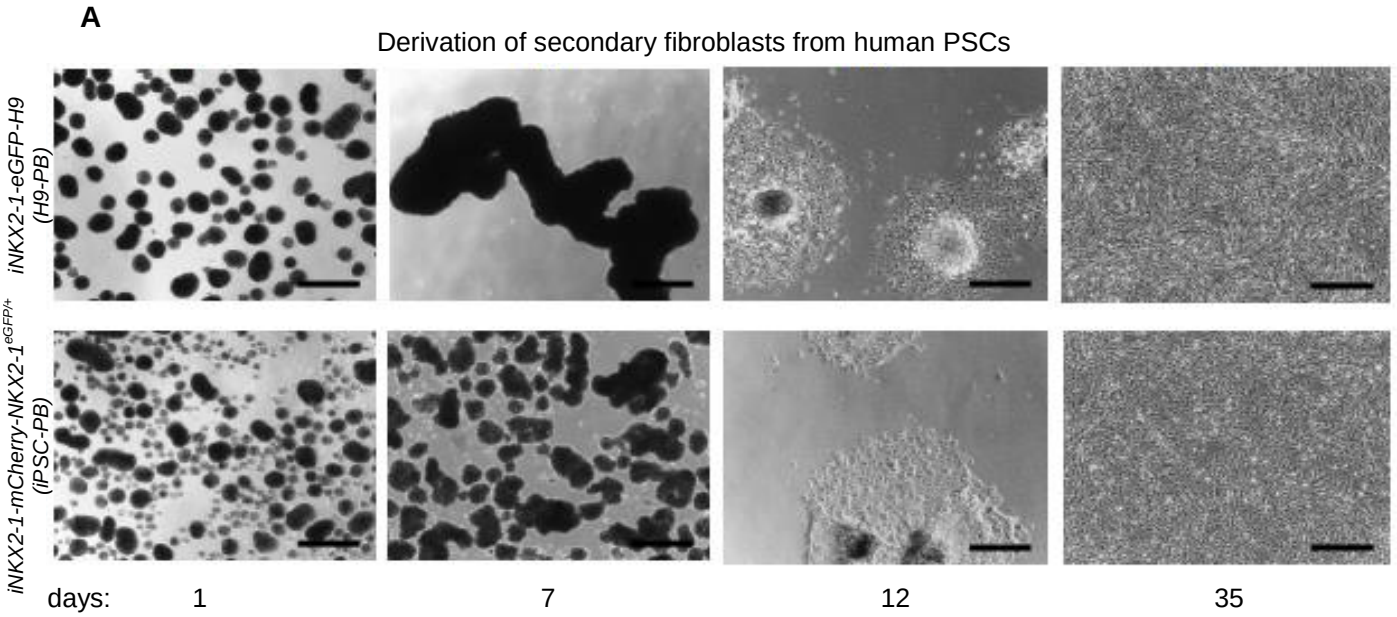
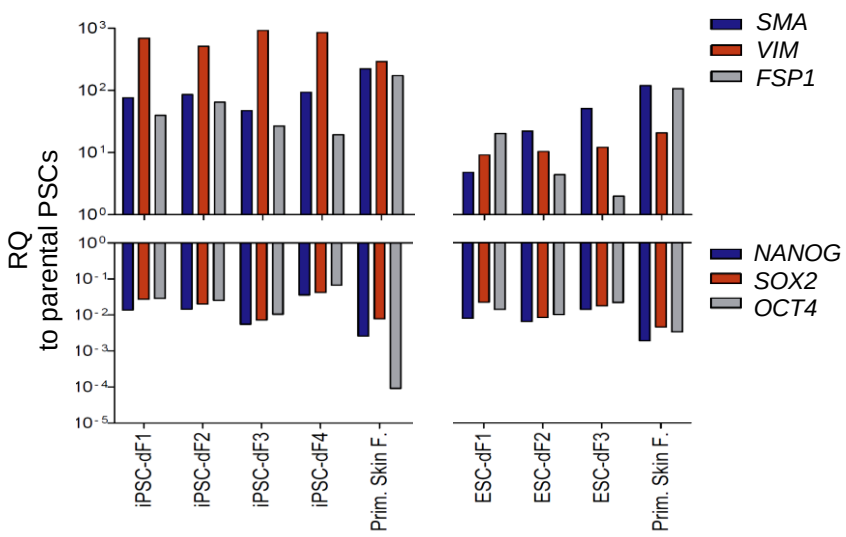


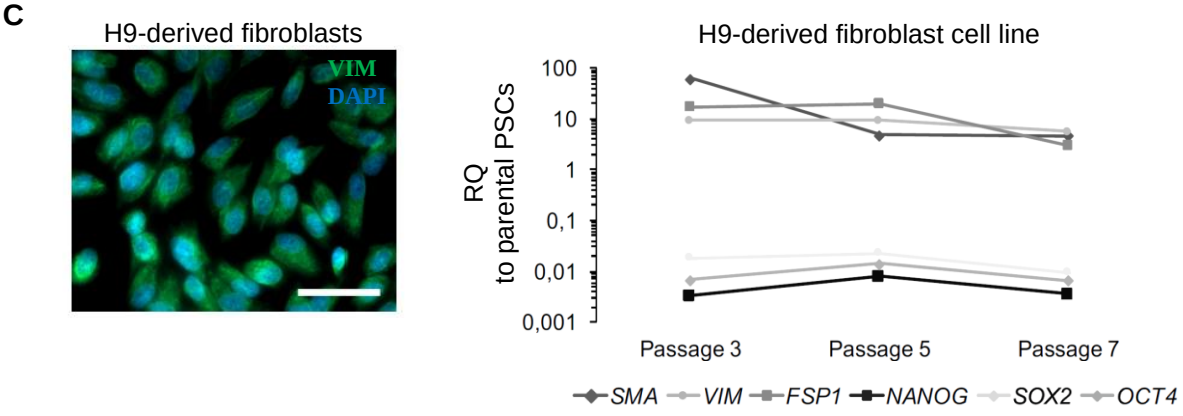
Figure 23. Effect of differentiation media and DE pre-induction on the accelerated differentiation outcome of human PSCs. (A) Gene expression analysis of *SFTPB*, *SFTPC*, *MUC5AC* and *SCGB1A1* after 21 days of differentiation, normalized to *GAPDH* and quantified by RT-qPCR ($2^{-\Delta CT}$). Cells were pre-differentiated (+DE) or not (-DE) to definitive endoderm for 3 days and were induced with dox for 0, 2 and 21 days in DCI or DCICS maturation media respectively. Error bars represent the mean +SD of n=3 replicates for induced cells, and n=2 replicates for uninduced cells (*$p < 0.05$). **(B)** Fluorescence microscopy of organoids in maturation medium DCI shows endogenous NKX2-1 expression marked by eGFP. Cells were pre-treated or not with DE induction condition for 3 days and induced with dox for 0, 2 and 21 days. Scale bars: 500 μm.

4.4.2. Successful generation of secondary fibroblasts from hPSCs

To investigate whether the ectopic expression of NKX2-1 is able and sufficient to promote trans-differentiation of fibroblasts into lung progenitor cells, we generated secondary fibroblasts from the *iNKX2-1-eGFP-H9* and *iNKX2-1-mCherry-NKX2-1*$^{eGFP/+}$ cell lines (**Figure 19** and **20**), using a protocol that was previously established with ESCs (Togo et al., 2011). Immunostaining for VIM as well as gene expression analysis for fibroblast markers *ACTA2* (smooth muscle actin alpha, *SMA*), *VIM* and *FSP1* (also called *S100A4*) and the pluripotency factors *NANOG*, *SOX2* and *OCT4* confirmed the similarity of generated fibroblasts to primary fibroblasts. They were able to stably maintain the gene expression levels along different passages and up to passage 14 no considerable altered morphologies or growth rates were observed (**Figure 24A-D**).


A
Derivation of secondary fibroblasts from human PSCs
iNKX2-1-eGFP-H9 (H9-PB)
iNKX2-1-mCherry-NKX2-1eGFP/+ (iPSC-PB)
days: 1 7 12 35


B
RQ to parental PSCs
SMA
VIM
FSP1
NANOG
SOX2
OCT4
iPSC-dF1
iPSC-dF2
iPSC-dF3
iPSC-dF4
Prim. Skin F.
ESC-dF1
ESC-dF2
ESC-dF3
Prim. Skin F.


C
H9-derived fibroblasts
VIM
DAPI
H9-derived fibroblast cell line
RQ to parental PSCs
Passage 3
Passage 5
Passage 7
SMA
VIM
FSP1
NANOG
SOX2
OCT4

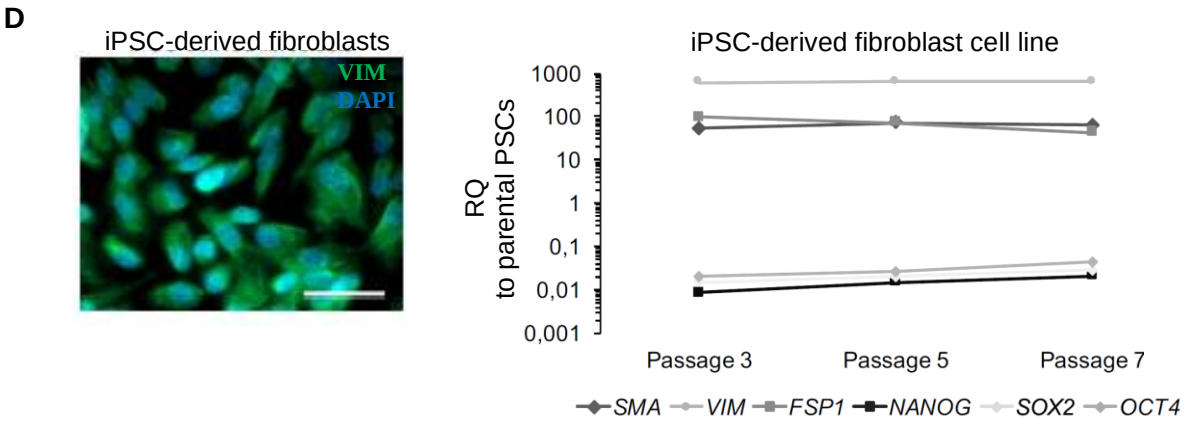


Figure 24. Generation and characterization of secondary fibroblasts from hPSCs. (A) Brightfield microscopy on days 1, 7, 12 and 35 of fibroblast differentiation. Scale bars: 500 µm. **(B)** Fold change of fibroblast markers *SMA, VIM*, and *FSP1* and pluripotency markers *NANOG, OCT4*, and *SOX2* relative to parental PSCs, normalized to *GAPDH* and quantified by RT-qPCR ($2^{-\Delta\Delta CT}$). **(C, D)** Representative fluorescence images of ESC- and iPSC-derived fibroblasts stained for VIM and DAPI and fold change of *SMA, VIM, FSP1* and *NANOG, OCT4*, and *SOX2* for 3 different passages (3, 5, 7) relative to undifferentiated PSCs, quantified by RT-qPCR ($2^{-\Delta\Delta CT}$). Scale bars: 50 µm. H9 and iPSC refer to the corresponding *iNKX2-1-eGFP-H9* and *iNKX2-1-mCherry-NKX2-1*$^{eGFP/+}$ PB cell lines, respectively.

4.4.3. Ectopic expression of NKX2-1 is not sufficient for the successful reprogramming of fibroblasts into lung progenitor cells

Aiming for the transdifferentiation of secondary fibroblasts into NKX2-1-expressing epithelial lung progenitor cells, I examined different combinations of growth factors (Morrisey and Hogan, 2010; Tsao et al., 2008). We used the mixtures of (1) CKF, which was used for the maintenance and expansion of lung progenitors in regular differentiation protocol (Hawkins et al., 2017), (2) CKF+BR that had performed the best outcome in the acceleration approach (**Figure 22**), as well as (3) CKF with DAPT (CKF+D), and two induction intervals (1 and 3 days). In consistent with the acceleration approach, CKF+BR conditions were more successful, resulting in downregulation of *FSP1* expression levels and in parallel upregulation of *GATA6* and *MUC5AC*, similarly to those of the LP. However, upregulation of *E-Cad* and *EpCAM* was not sufficient. In addition, no treatment led to an upregulation of the endodermal marker *FOXA2*. Between the two induction intervals tested, no significant difference was observed (**Figure 25A**). Treatment with VPA for 24 hours under CKF+BR conditions, combined with the two different induction intervals did not improve the transdifferentiation efficiency. Furthermore, even in uninduced conditions, some gene expression levels were changed, thereby following the similar trend of dox induced conditions. In conclusion,

expression levels of *FSP1*, *GATA6* and *MUC5AC* approached to those of lung progenitors, while *E-Cad* and *EpCAM* were again insufficiently upregulated, and substantial upregulation of *FOXA2* was lacking (**Figure 25B**).

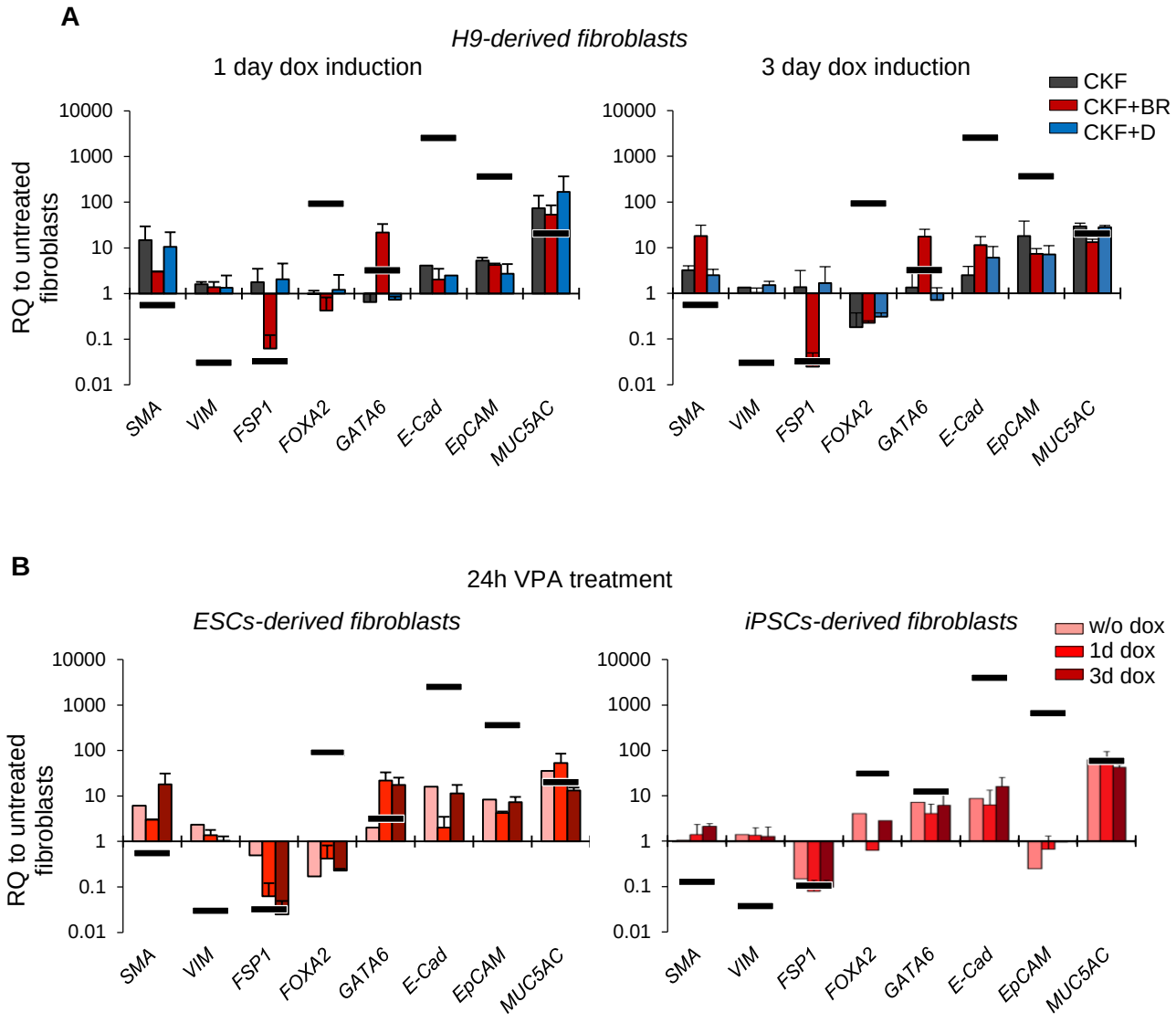


Figure 25. Effect of media composition and VPA treatment on transdifferentiation of PSC-derived fibroblasts into lung epithelial cells. (A) Relative quantification of *SMA, VIM, FSP1, FOXA2, GATA6, E-Cad, EpCAM, MUC5AC* on day 7, quantified by RT-qPCR ($2^{-\Delta\Delta CT}$). Three different media, CKF, CKF+BR, and CKF+D are compared, as well as the induction intervals of 1 and 3 days. **(B)** Relative quantification of *SMA, VIM, FSP1, FOXA2, GATA6, E-Cad, EpCAM, MUC5AC* on day 7 (in CKF+BR), relative to respective untreated H9- or iPSC-derived fibroblasts, quantified by RT-qPCR ($2^{-\Delta\Delta CT}$). Error bars represent the mean +SD, n=2 biological replicates. H9 and iPSC refer to the corresponding *iNKX2-1-eGFP-H9* and *iNKX2-1-mCherry-NKX2-1*$^{eGFP/+}$ PB cell lines, respectively.

4.5. RECONSTRUCTION OF DEVELOPMENTAL GENE TRAJECTORIES REGULATING LUNG SPECIFICATION

The heterogeneity observed during hiPSC differentiation towards lung was highlighted by the mutually exclusive expression of lung and liver hallmarks. This observation led me to investigate deeper the regulatory mechanisms resulting in the two lineages, as well as to pinpoint the divergence of the two lineages by profiling single cells across differentiation. Bioinformatic processing of sequencing data for this chapter of the thesis was performed by Meshal Ansari.

4.5.1. High-throughput single cell RNA-seq time course

In order to identify differentiation trajectories and mechanisms that specify lung progenitor cells and hepatoblasts, we conducted a 16-day time-series scRNA-seq using the Dropseq workflow (Macosko et al., 2015) during the differentiation protocol. Single cell suspensions were processed daily, resulting in 10,667 cells used for subsequent analysis (**Figure 26A-C**) (Ori et al., 2021).

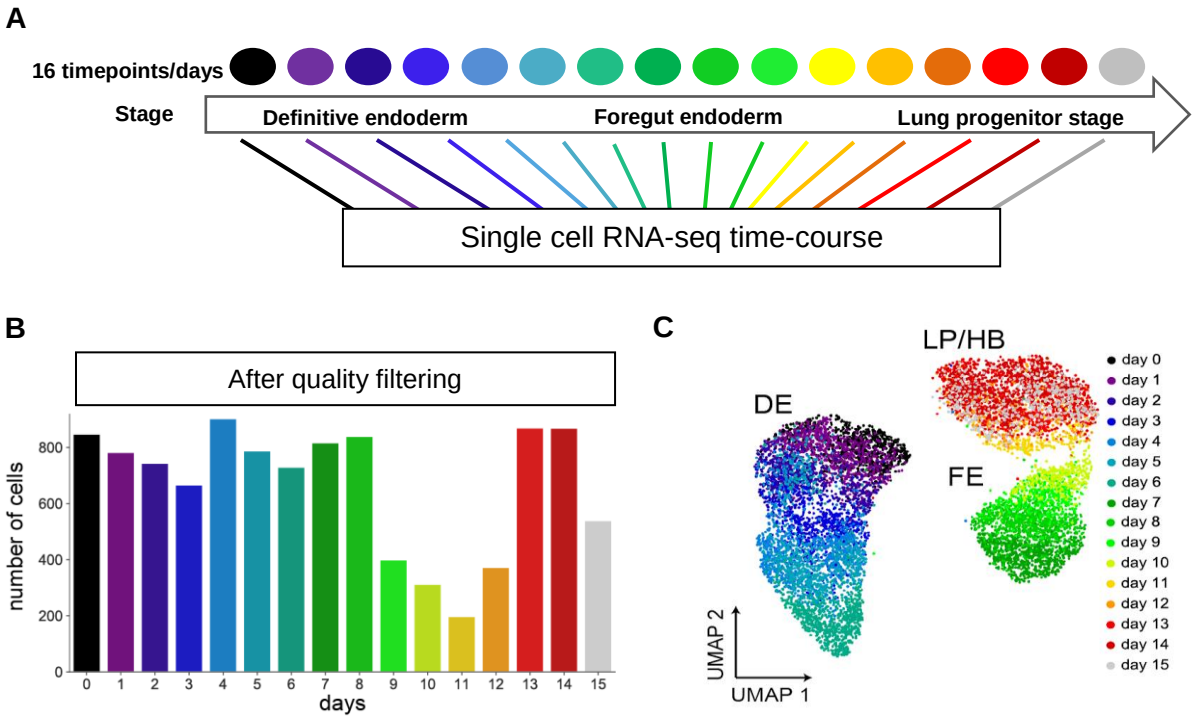


Figure 26. Single cell RNA-seq time-course as a tool to resolve the heterogeneity underlying the lung differentiation protocol. (A) Descriptive scheme of experimental strategy illustrating the daily sampling for scRNA-seq analysis during the first 15 days of differentiation. **(B)** Barplot displaying the amount of cells used for the following analysis upon quality control filtering per time point. **(C)** UMAP embedding of single cells color-coded by the day of sampling. Adapted from (Ori et al., 2021).

Initially, Uniform Manifold Approximation and Projection (UMAP) and Partition-based Graph Abstraction (PAGA) analysis was used (Wolf et al., 2019), in order to project the gene expression data into two dimensional space and assessed connectivity of the clusters (**Figure 26C, 27A** and **B**) (Ori et al., 2021). Three major domains emerged in the high dimensional gene expression manifold corresponding to the three developmental stages of the differentiation approach (DE, FE and LP/HB) (**Figures 26C** and **27A**), and sub-clusters marked by genes highlighting the progress of differentiation as well as the heterogeneity along the stages (**Figure 27C**) (Ori et al., 2021).

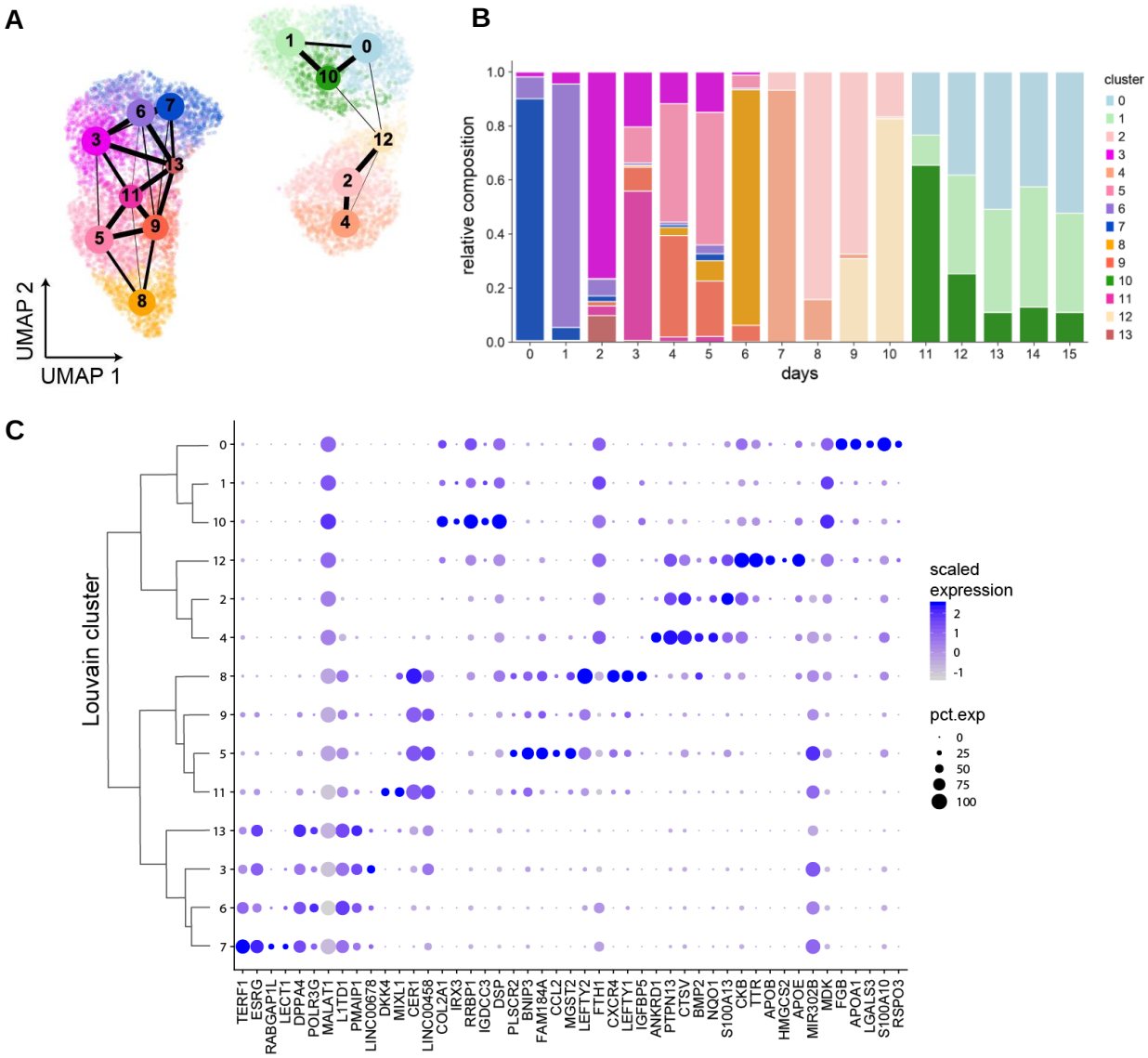


Figure 27. Overview of data set after quality filtering. (A) UMAP plot displaying the connectivity of Louvain clusters calculated by PAGA. **(B)** Barplot demonstrating the relative composition of Louvain clusters for each time point of sampling. **(C)** Dendrogram demonstrating the top 5 genes per cluster generated by unsupervised clustering, presenting the normalized expression level (across the clusters) in blue and the percentage of cells expressing the relevant gene according to dot size. Adapted from (Ori et al., 2021).

Intriguingly, in last stage of differentiation, representing days 11-15, three interconnected Louvain clusters were found, also connected to the same sub-cluster at FE progenitor stage (**Figure 27A**), with the lung and hepatocyte markers *NKX2-1* and *AFP* being mutually exclusive in clusters 10 and 0/12, respectively (**Figure 28A**) (Ori et al., 2021). Furthermore, expression of lung markers, including *MECOM* and *IRX3*, as well as putative markers e.g. *IGFBP5* and *IGDCC3* (**Figure 18A**), was associated with the two clusters that expressed *NKX2-1* and *eGFP*, and known hepatocyte markers, including *APOA1*, *APOB*, *FGB*, and *AFP*, as well as putative markers e.g. *GJB1*, *MEP1A* and *DPYS* were associated with different cluster(s) (**Figure 28A-C**) (Ori et al., 2021). Taken together, this indicated that lung and hepatocyte progenitors were also readily detected by scRNA-seq.

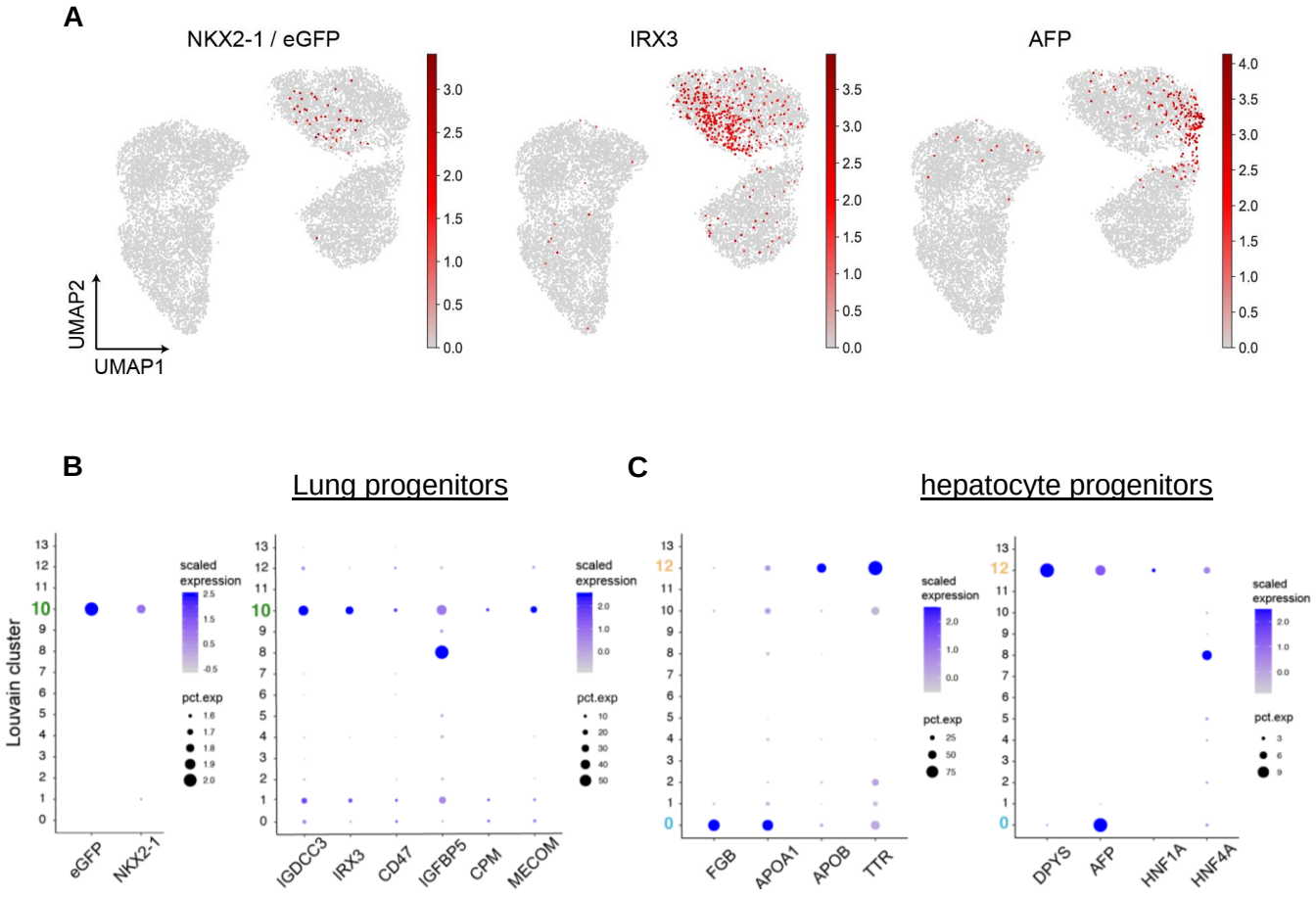
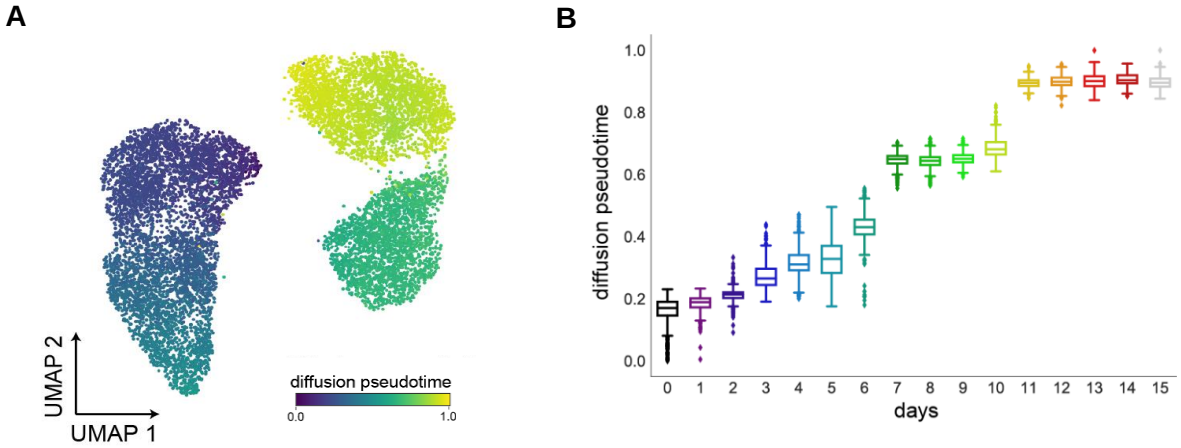


Figure 28. Hallmarks of lung and liver progenitors are exclusively present at last stage of differentiation. (A) UMAPs showing cells positive for the TFs expressed in lung progenitors *NKX2-1*, *IRX3* and the liver marker *AFP*. **(B, C)** Dot plots demonstrating that genes which were found to be expressed in lung progenitor cells (**Figures 15C** and **18A**) are higher expressed in Louvain clusters 10 and 1, while markers related to hepatocytes are mostly present in clusters 0 and 12. Adapted from (Ori et al., 2021).

In order to model the dynamics of gene expression along the entire trajectory starting from iPSCs to lung and hepatoblast progenitors, pseudotime inference was used (**Figure 29A**), which paralleled the time points of sampling (**Figure 29B**) (Ori et al., 2021). Unsupervised hierarchical clustering of the genes along the pseudotime uncovered stage specific markers including transcription factors expressed in undifferentiated iPSCs e.g. *POU5F1 (OCT4)* and *NANOG*; DE factors e.g. *CER1*, *CXCR4* and *SOX17*; FE TFs including e.g. *FOXA2* and *FOXP1*; and genes that are important for the formation of lung and hepatocyte progenitors in the last stage e.g. *NKX2-1*, *IRX3*, *MECOM*, and *AFP*, *APOB*, *APOA1*, respectively (**Figure 29C** and **D**) (Ori et al., 2021). The pseudotime was consistent with sequential expression of DE, FE and lung progenitor markers, e.g. *NODAL*, *CER1*, *FOXA2*, *IRX3* and *NKX2-1* along the time course (**Figure 29D**) (Ori et al., 2021). Significantly, genes that are characteristic to hepatoblast and hepatocyte generation, including *AFP*, *APOA2*, *FGB and APOB* were up-regulated already during the second stage of the protocol (**Figure 29C** and **D**) (Ori et al., 2021). These were found in Louvain clusters where lung markers were not detected, namely clusters 0 and 12 (**Figure 28C**) (Ori et al., 2021).



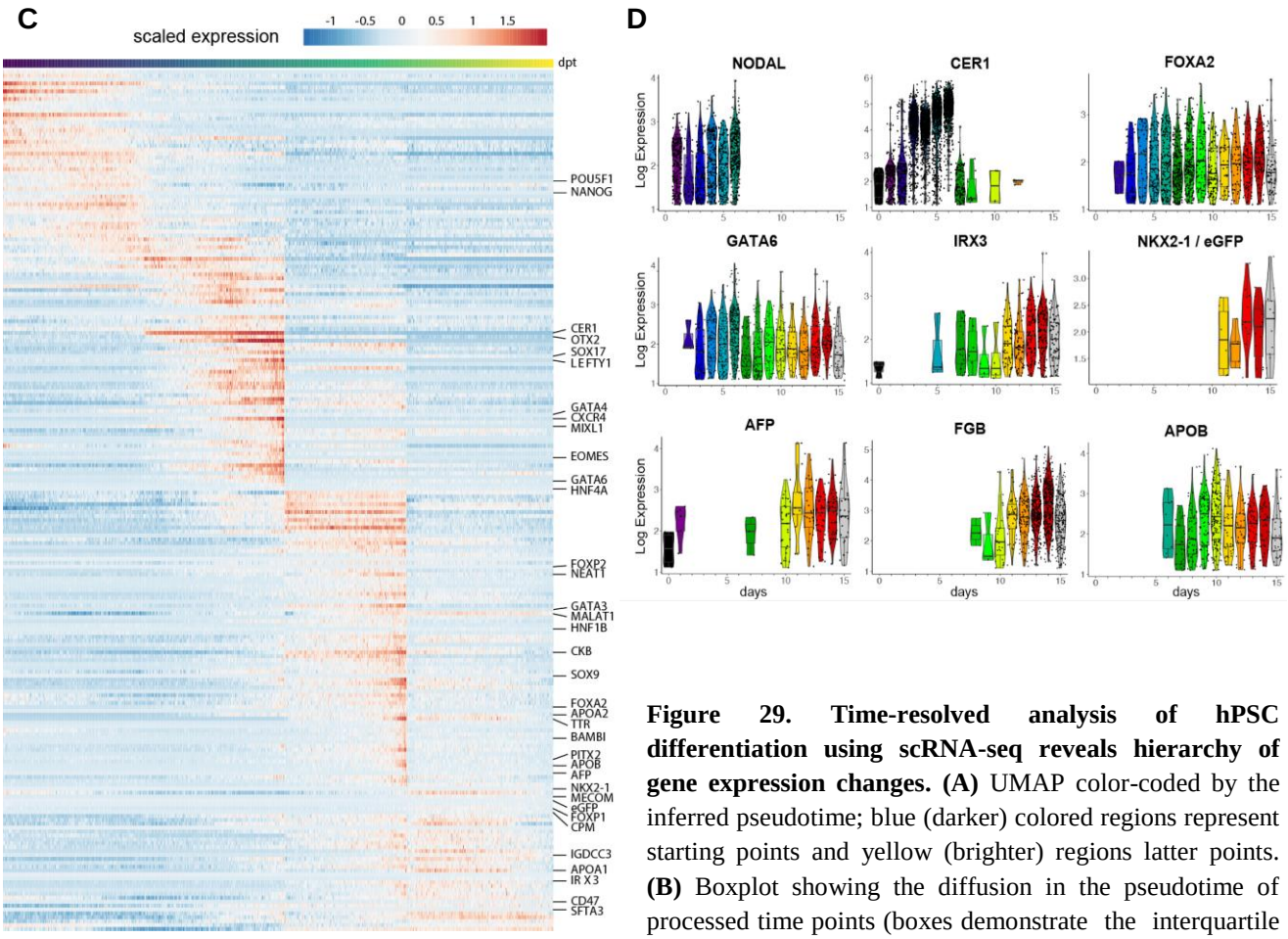


Figure 29. Time-resolved analysis of hPSC differentiation using scRNA-seq reveals hierarchy of gene expression changes. (A) UMAP color-coded by the inferred pseudotime; blue (darker) colored regions represent starting points and yellow (brighter) regions latter points. **(B)** Boxplot showing the diffusion in the pseudotime of processed time points (boxes demonstrate the interquartile range, horizontal line the median, and the whiskers standing for 1.5 times the interquartile range). **(C)** Heatmap of gene expression during differentiation towards LP/HP fate along the pseudotime, with chosen stage specific marker genes being presented. **(D)** Violin plots showing the expression levels of stage-specific gene markers during the days of differentiation protocol. Adapted from (Ori et al., 2021).

To identify candidate signaling pathways underlying the differentiation of lung progenitors from human PSCs we performed gene category enrichment analysis. As expected, cells were highly enriched on day 0 for the stem cell related terms, namely inner cell mass cell proliferation, embryonic cleavage and stem cell maintenance (**Figure 30**). Upon induction of differentiation, DE formation was associated with TGF-β/Smad and Wnt/β-catenin signaling pathways, in conjunction with endoderm formation and gastrulation (**Figure 30**). Furthermore, I noted association of Hippo signaling with the FE differentiation (**Figure 30**) (Ori et al., 2021). Hippo signaling is known to be important in lung epithelial cell development by complex interactions with SHH, β-catenin, and FGF10 signaling (Volckaert et al., 2019), ligands of which were used in the differentiation protocol.

Finally, analysis of the putative NKX2-1/eGFP lung progenitors revealed association with Notch signaling (Ori et al., 2021). Notch signaling is crucial for lung development in the mouse by regulating early patterning and alveolar morphogenesis (Tsao et al., 2016), as well as for lung regeneration (Rock et al., 2011). Moreover, associations with integrin and insulin-like growth factor signaling were also readily apparent. Intriguingly, studies have suggested the involvement of integrins in lung branching morphogenesis and in the establishment of epithelial cell polarity during early development (Coraux et al., 1998; Plosa et al., 2014). Also, IGF-1R is required for gestational maturation of the lung (Epaud et al., 2012; Wang et al., 2018). Collectively, these data indicated the involvement of a number of signaling pathways that have not been previously implicated in the differentiation of lung progenitors from human PSCs.

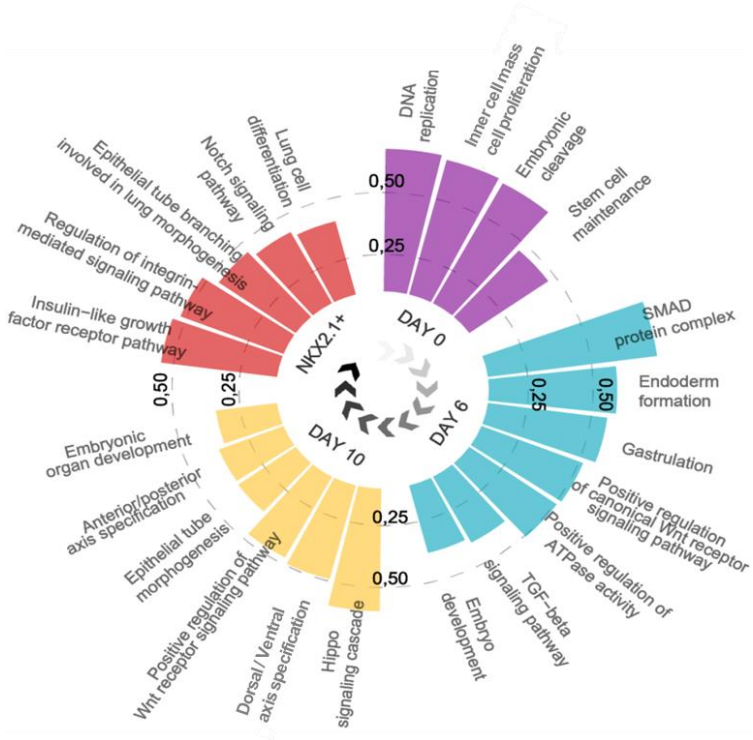


Figure 30. Identification of candidate pathways that underlie the differentiation of lung progenitors via gene category enrichment analysis. Circular bar chart demonstrating enriched gene categories derived from Uniprot Keywords, GO terms and KEGG pathways related to the early stages of lung differentiation namely, PSCs (day 0), DE (day 6), FE (day 10) and NKX2-1+ progenitors (day 15) (FDR <5%). Taken from (Ori et al., 2021).

4.5.2. Mechanistic pathway leading to DE induction

In order to identify the mechanisms promoting the emergence of lung progenitors, first I sought to investigate the fundamental programs underlying definitive endoderm specification (**Figure 31A**) (Ori et al., 2021), with a particular focus on the timing of up-regulation of particular transcription factors. Analysis showed that immediately upon Activin-A administration to the culture the expression of *NODAL* was strongly up-regulated (**Figure 31B**) and shortly upon the initiation of differentiation pluripotency factors *POU5F1*, *SOX2* and *NANOG* were down-regulated along with the simultaneous up-regulation of the DE specific genes, namely *CXCR4*, *SOX17* and *GATA4/6* (**Figure 31B**) (Ori et al., 2021). *T (Brachyury)* was detected in only few cells during the first 6 days of differentiation (**Figure 31B**) (Ori et al., 2021). This is very interesting as a fundamental question in developmental biology is whether endodermal organs develop from mesendoderm intermediates. Previous analysis of definitive endoderm induction from human PSCs by scRNA-seq had shown that cells transiently expressed the primitive streak gene *Brachyury* (*T*) following 24 hours of differentiation (Chu et al., 2016), while in the mouse embryo a minority population, which initially express T and later Foxa2, is considered mesendoderm (Burtscher and Lickert, 2009). *T* expression was detected only in few cells, and moreover *FOXA2* and *EOMES* were not co-expressed with *T* (**Figure 31B**), suggesting that *FOXA2* and *EOMES* expression marked the formation of DE rather than mesendoderm and that the origin of the liver and lung cells *in vitro* is not necessarily related to mesendoderm origin (Teo et al., 2011; Ori et al., 2021).

Furthermore, given the indications of asynchronous emergence of liver and lung progenitors in my culture (section **4.5.1.**), I aimed to identify regulators that stimulate the emergence of lung progenitors, and the formation of hepatoblast-like cells at the DE stage. In this context, Isl1 was recently identified as a key regulator of *Nkx2-1* in the early foregut (Kim et al., 2019), whereas Hhex has been linked to the activation of the *Hnf4a,* both TFs which are essential for liver development (Hunter et al., 2007). In addition, *Gata6*, *Foxa1* and *Foxa2* are well known TFs for their vital roles in liver generation (Li et al., 2009; Zhao et al., 2005). The data showed up-regulation of *GATA6*, *FOXA2* and *HHEX*, but not of *ISL1* or *NKX2-1* (**Figure 31B**) indicating that the mechanism which stimulates the specification into liver cells is being established during the early formation of definitive endoderm rather than upon the generation of foregut (Ori et al., 2021).

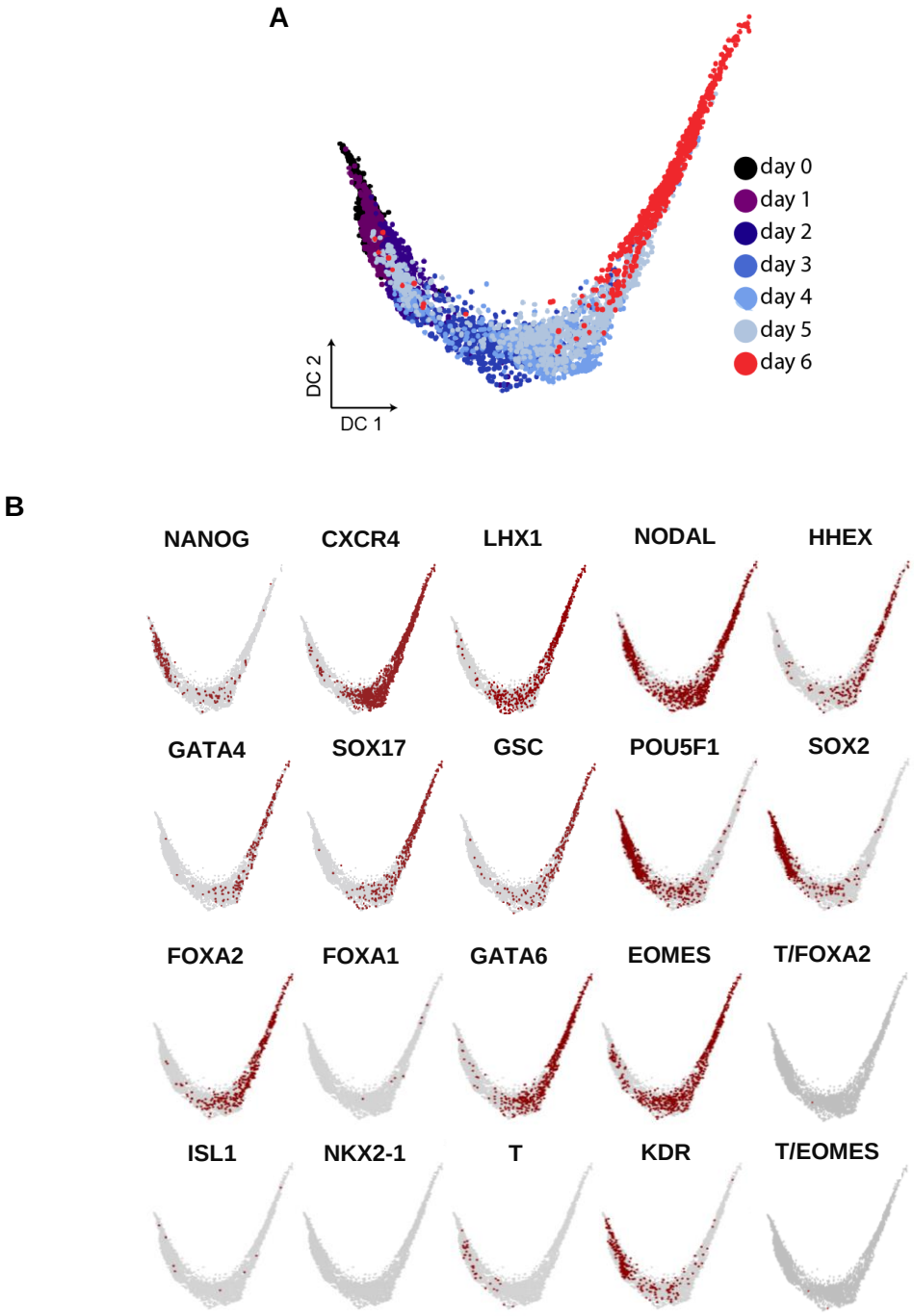


Figure 31. Analysis of DE specification. (A, B) Diffusion maps of single cells during DE differentiation (days 0-6) **(A)**, indicating cells with at least one transcript for the respective genes **(B)**. Adapted from (Ori et al., 2021).

To gain a comprehensive view of the circuitry of TFs that regulate the differentiation into definitive endoderm, we used hierarchical clustering of the genes that showed transcriptional changes during the first 6 days of differentiation. For the six major clusters that were generated (A1-

A6), representing different gene expression kinetics, we performed Gene Ontology enrichment analysis (**Figure 32A** and **B**) (Ori et al., 2021). First of all, pluripotency genes exhibited heterogeneous patterns of down-regulation. More specifically, the expression of *SOX2* (in cluster A1) persisted slightly longer and its down-regulation took place smoothly, while the TFs *POU5F1* and *NANOG* belonging in the cluster A2 were shortly induced before being down-regulated (**Figure 32B**) (Ori et al., 2021). This is in line with the role of Activin/Nodal signaling in promoting NANOG expression in human PSCs (Vallier et al., 2009), while also interestingly co-regulation of *NANOG*, *MYC* and *ID1* exhibited a similar trend in human preimplantation embryos (Blakeley et al., 2015), implying a regulatory network within different developmental stages.

Moreover, expression of key DE genes such as *EOMES, LHX1, OTX2, CXCR4, LEFTY1, SOX17* followed a similar upregulation pattern which gradually reached a peak at day 6, representing a cluster (A6) enriched in pathways involved in endoderm specification and regulation of SMADs and TGF-β signaling (**Figure 32A** and **B**), including the loop of EOMES-Activin/Nodal signaling (**Figure 33**) (Ori et al., 2021). Importantly, full activation of the EOMES gene circuit was readily apparent during early differentiation; evident by the continuous up-regulation of *FOXA2*, *LHX1* and *GSC* upon *EOMES* induction (**Figure 31B**) (Costello et al., 2015; Ori et al., 2021). Moreover, the maintenance of this circuit by a feed-forward loop was apparent as *NODAL* was constitutively expressed from the time point of 24 hours onwards (**Figure 31B**) and given that Activin/Nodal signaling is known to directly regulate EOMES (Teo et al., 2011). Furthermore, studies that have previously examined bulk DE cells have shown that a subset of KDR+ (VEGF receptor 2) progenitor cells produces hepatic cells from differentiated human ESCs, and that KDR signaling is crucial for their induction (Gordillo et al., 2015). In contrast, I observed that the up-regulation of *FOXA2*, *GATA4* and *GATA6* concurred with a downward trend of *KDR* (**Figure 32B** and cluster A3), implying that KDR plays a role either only in very early specification events or might not be necessarily involved in driving the early liver phenotype in parallel to lung cells in this case (Ori et al., 2021).

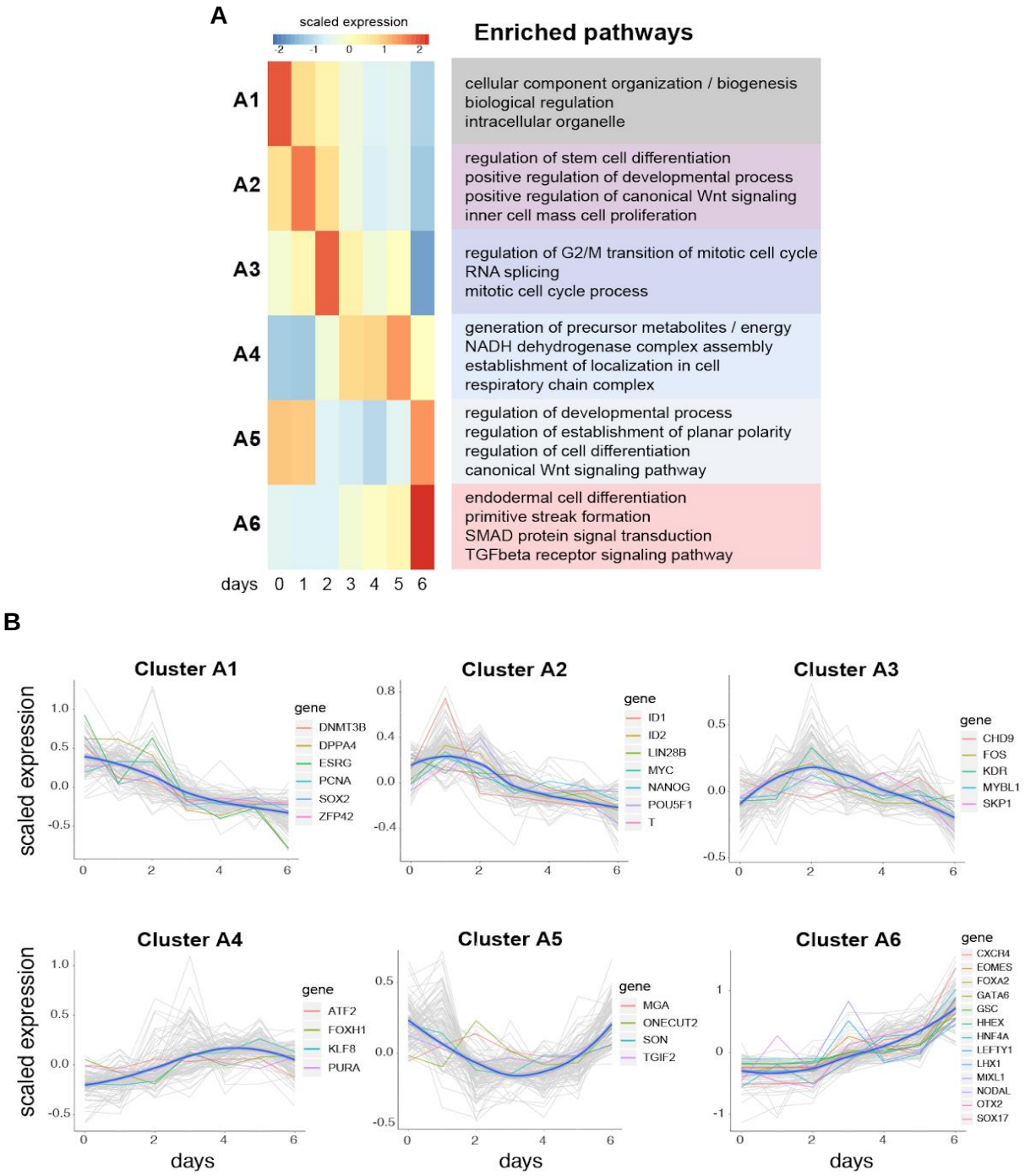


Figure 32. Analysis of the gene expression kinetic patterns shaping the PSC differentiation towards DE. (A) Heatmap showing the scaled mean expression per gene cluster and the related pathways revealed by hierarchical clustering of the genes with significant temporal association during the formation of DE (adjusted p-value < 0.0005, 6101 genes; high-to-low association colored red to blue). **(B)** Hierarchical clustering of transcriptomic data showing selected genes relevant for the transition from pluripotency to DE. The scaled expression values of the top 100 genes per cluster are indicated by the grey lines, while the blue lines represent the median expression within the corresponding cluster. Taken from (Ori et al., 2021).

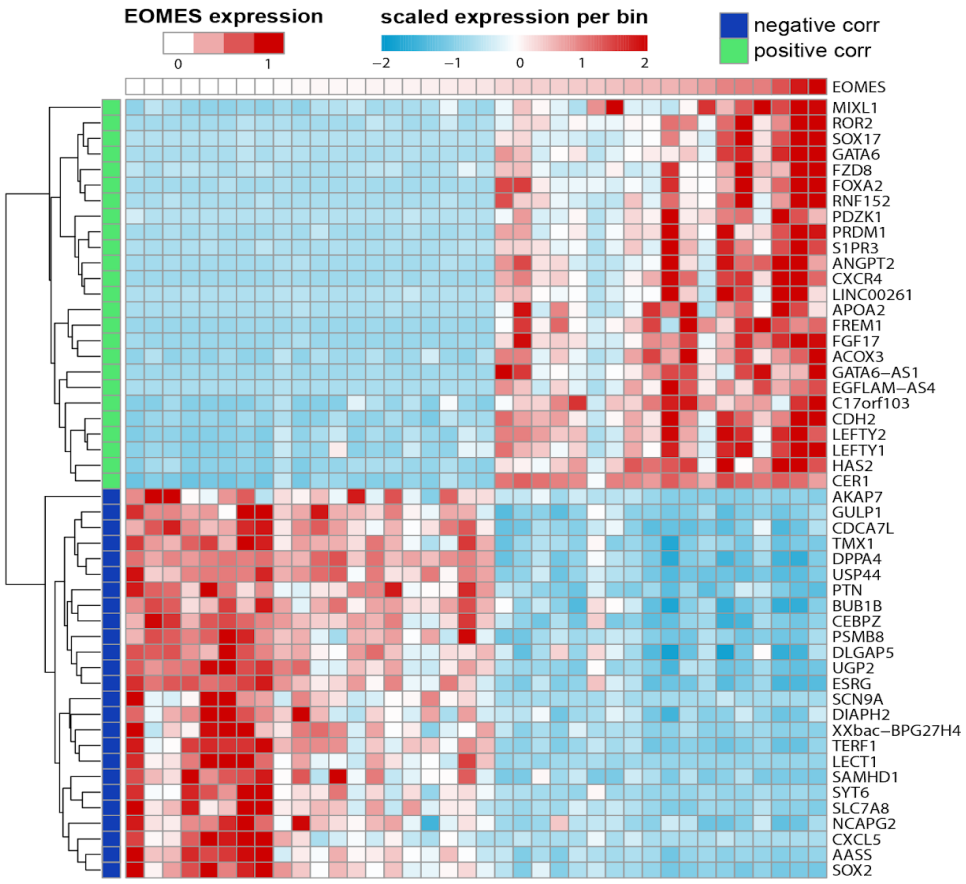


Figure 33. Gene correlation to EOMES during the differentiation of PSCs to DE. Cells were clustered into 38 bins based on transcriptomic profile; row-scaled average expression per bin is displayed ordered based on the expression levels of EOMES. Taken from (Ori et al., 2021).

4.5.3. Gene networks underlying definitive endoderm patterning to foregut endoderm

As gastrulation proceeds from definitive endoderm, the foregut endoderm emerges at the primitive gut tube (Zorn and Wells, 2009). Analysis of the second stage of differentiation (days 6-10), showed that immediately upon Activin-A withdrawal, *NODAL* expression was down-regulated, whereas expression of *FOXA2*, *FOXP1* and *HHEX*, genes that are indicative for foregut endoderm formation, was up-regulated (**Figure 34A** and **B**) (Ori et al., 2021). Moreover, cells were negative in that stage for markers characteristic of more posterior regions of the gut tube namely *PDX1* (midgut) and *CDX2* (hindgut) (**Figure 34B**) (Ori et al., 2021). In addition, genes that were previously described in foregut namely *PITX2*, *ISL1*, and *CTNNB1* (Faucourt et al., 2001; Harris-Johnson et al., 2009; Kim

et al., 2019), as well as genes that demonstrated exclusive expression in lung and hepatocyte progenitors at the next stage of differentiation, such as *IGDCC3* and *APOB* (**Figure 28B** and **C**), were up-regulated (**Figure 34B**) (Ori et al., 2021).

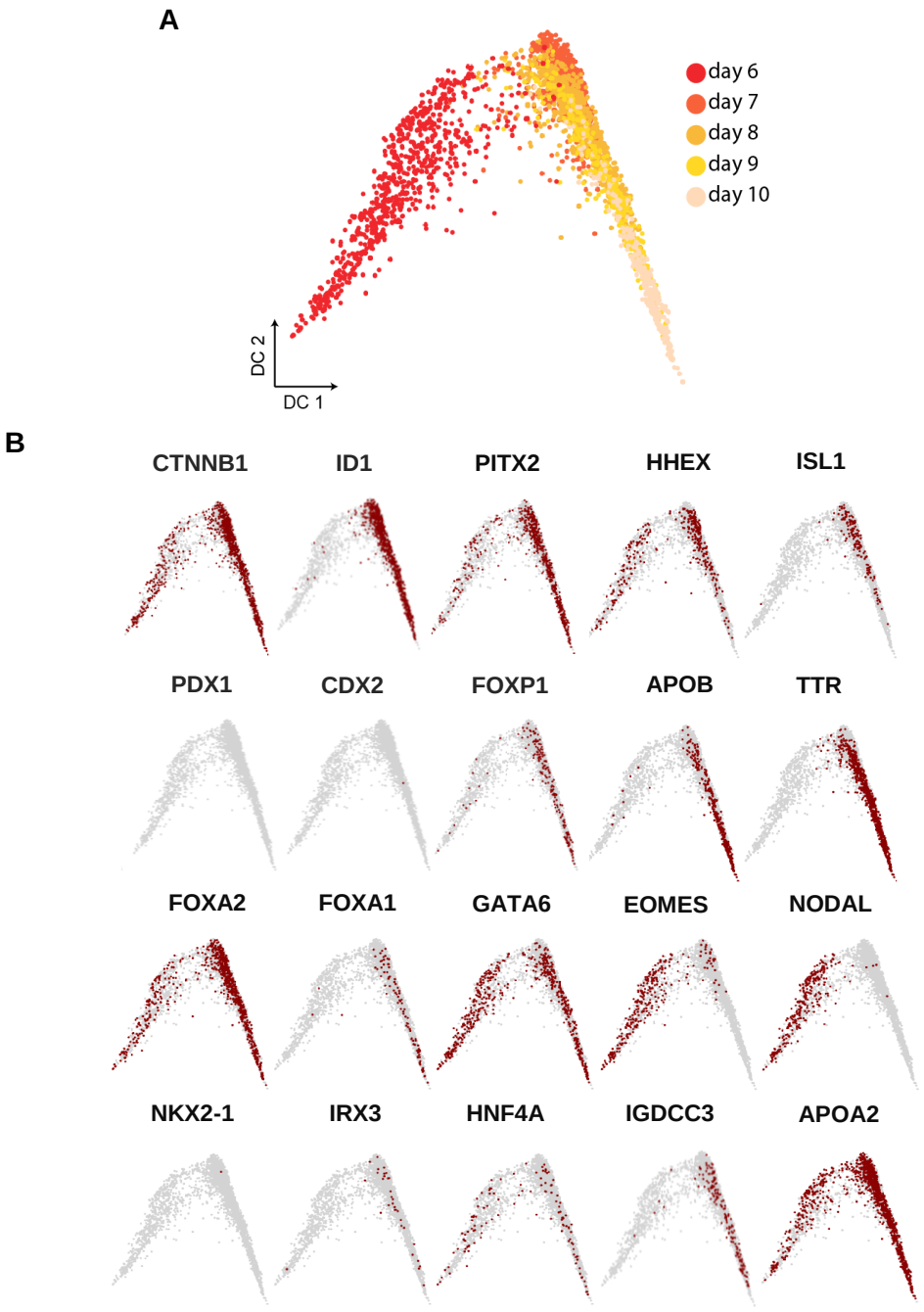


Figure 34. DE patterning to FE. (A, B) Diffusion maps of single cells (as above in **Figure 31**) during FE differentiation (days 6-10) (**A**), indicating cells expressing the displayed genes (**B**). Adapted from (Ori et al., 2021).

In order to study in detail whether lung progenitors emerge during the second stage of hiPSC differentiation, and to understand whether this process coincides with the generation of hepatoblasts, we used hierarchical clustering to generate six clusters of gene expression (B1-B6), for days 6-10 of differentiation and subsequently investigated the pathways enriched (**Figure 35A** and **B**) (Ori et al., 2021). The genes associated with DE formation, namely *EOMES*, *LHX1*, *GSC*, *OTX2*, *SOX17* and *CXCR4* (Costello et al., 2015), and GO pathways associated with mesendoderm development were down-regulated immediately upon Activin-A withdrawal and application of dual-SMAD inhibition in this stage (cluster B1, **Figure 35A** and **B**) (Ori et al., 2021). However, *FOXA2* which is crucial for lung morphogenesis and asymmetry in the mouse, but not for the early specification of the lung (Lin et al., 1999; Shu et al., 2007), was clustered together with *FOXP1* and *PITX2* (cluster B5) (**Figure 35B**) (Ori et al., 2021). Consistent with these, GO terms for lung morphogenesis, embryonic organ development and epithelial cell differentiation were enriched in cluster B5 (**Figure 35A** and **B**), although at this stage *NKX2-1* is not yet up-regulated (**Figure 34B**) (Ori et al., 2021). However, at this stage, *ISL1* expression was up-regulated in cluster B3 alongside *IRX3*. Isl1 was shown to control the development of lung lobes and trachea-esophagus tube separation by inducing the activation of *Nkx2-1* (Kim et al., 2019), and Irx3 is essential for lung generation by stimulating the proliferation of branched epithelium (van Tuyl et al., 2006). In line with that, the GO terms of cluster B3 included epithelial cell proliferation and tube closure, which together showed that the lung progenitor program has begun between days 6-10 (**Figure 35A**) (Ori et al., 2021). Intriguingly, Isl1 is also known to regulate SHH in the foregut endoderm (Lin et al., 2006), implying that differentiation of eGFP+ cells might also be further promoted by endogenous SHH (**Figure 12B**) (Ori et al., 2021). The association with SHH pathway activation at days 6-10 was evident in clusters with upward trend, including *GLI4* in B5 and *GLI3* in B6 (**Figure 35B**) (Ori et al., 2021).

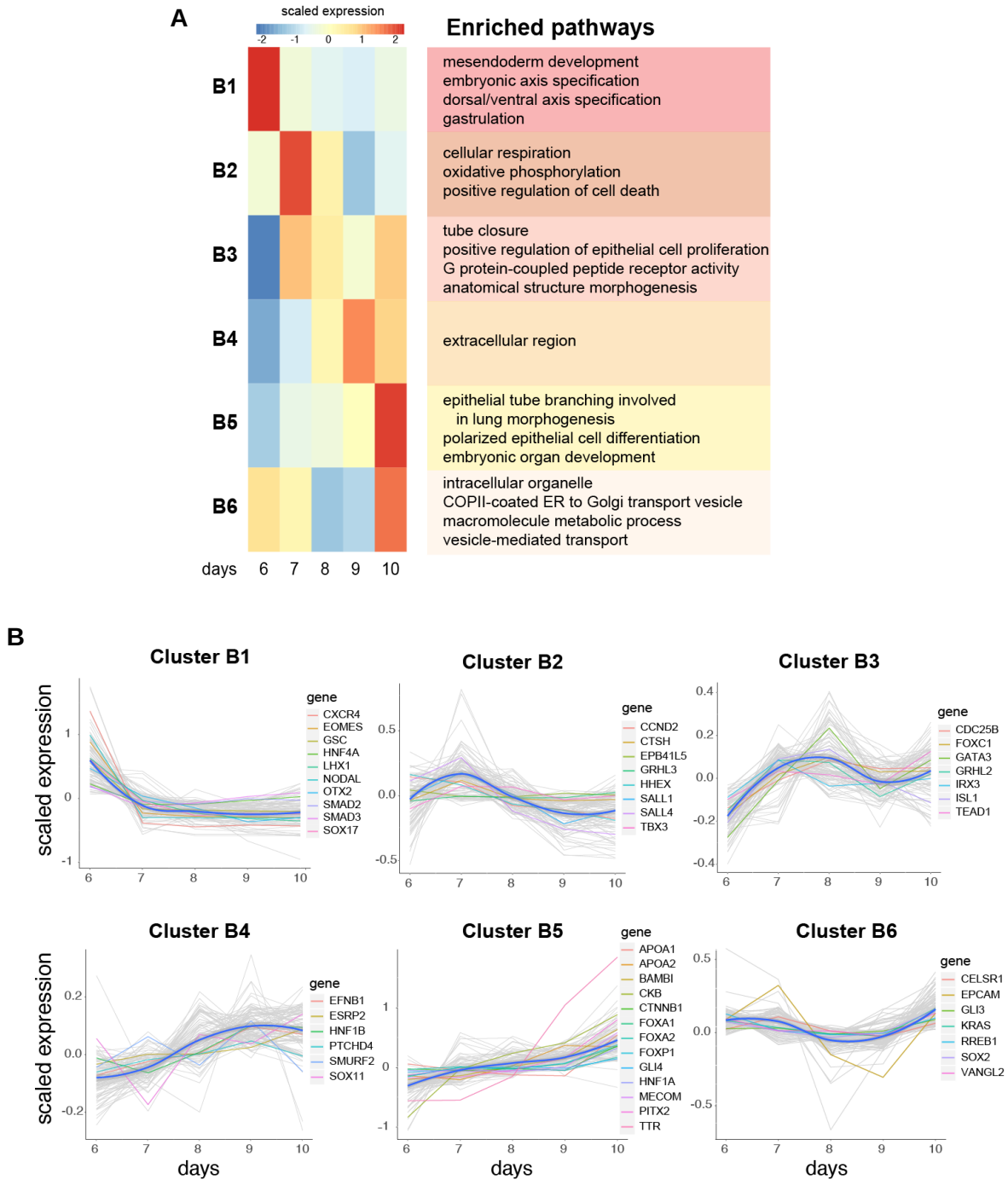
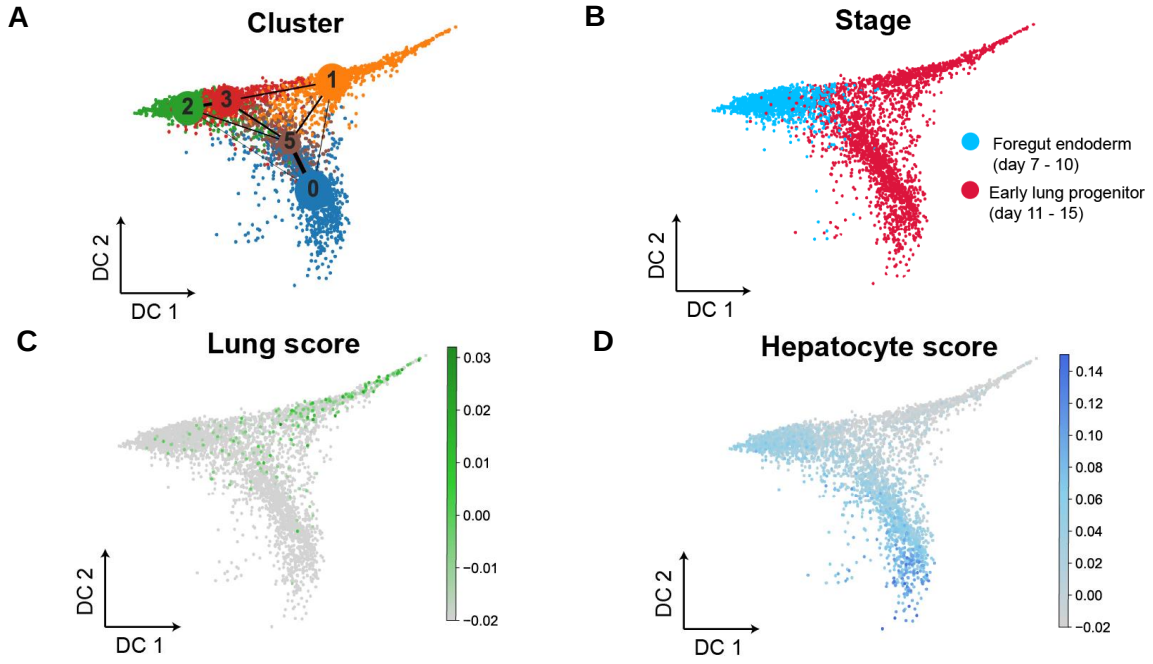


Figure 35. Kinetics and pathway analysis during DE patterning to FE. (A, B) Hierarchical cluster analysis (as shown above in **Figure 32**) during the formation of FE stage (corresponding to days 6-10 of differentiation). **(A)** Heatmap showing the scaled mean expression per gene cluster and the related pathways revealed by enrichment analysis using the established gene groups with significant associations during the formation of FE (adjusted p-value < 0.005, 9352 genes; high-to-low association). **(B)** Hierarchical clustering of transcriptomic data showing selected genes relevant for the transition from DE to FE. Scaled expression values of the top 100 genes per cluster are indicated by the grey lines. The blue lines represent the median expression within the corresponding cluster. Taken from (Ori et al., 2021).

Last, we investigated the expression of hepatocyte related genes during days 6-10 and upon activation of *HHEX*. Landmark liver transcription factors including *HNF1A*, *HNF1B*, *HNF4A* and *TBX3*, which control the expansion of the liver buds, were upregulated (Lüdtke et al., 2009), as well as the functional liver genes, namely, *APOB*, *APOA1*, and Transthyretin (*TTR*) which are translated into proteins secreted by the liver were evident (**Figure 34B** and **35B**) (Ori et al., 2021). Importantly, the expression of these genes was linked to Louvain clusters 0 and 12 that were largely mutually exclusive from the above mentioned lung factors and related Louvain cluster 10 (**Figure 27A** and **28C**) (Ori et al., 2021). Collectively, this scRNA-seq data suggests that the lung and liver lineages are settled in their respective trajectories during the second stage of the protocol leading to foregut endoderm specification (Ori et al., 2021).

4.5.4. Trajectory analysis for the reconstruction of lung-lineage specification

Next, I sought to investigate the gene trajectories distinguishing lung from liver fates. Analysis of differentiation days 11-15 combined with days 7-10 unraveled a branching event in the high dimensional single cell data manifold (**Figure 36A** and **B**) (Ori et al., 2021). At the intersection of the FE stage with the last stage in which advanced lung progenitors emerged, the single cells initiate to separate into two groups and the maturity of both fates increased over time compared to eGFP+ and eGFP- populations (**Figure 36C** and **D**). Interestingly, as mentioned above, the hepatoblast fate preceded the lung signature and was already established on days 7-10 **(Figure 36D**) (Ori et al., 2021).



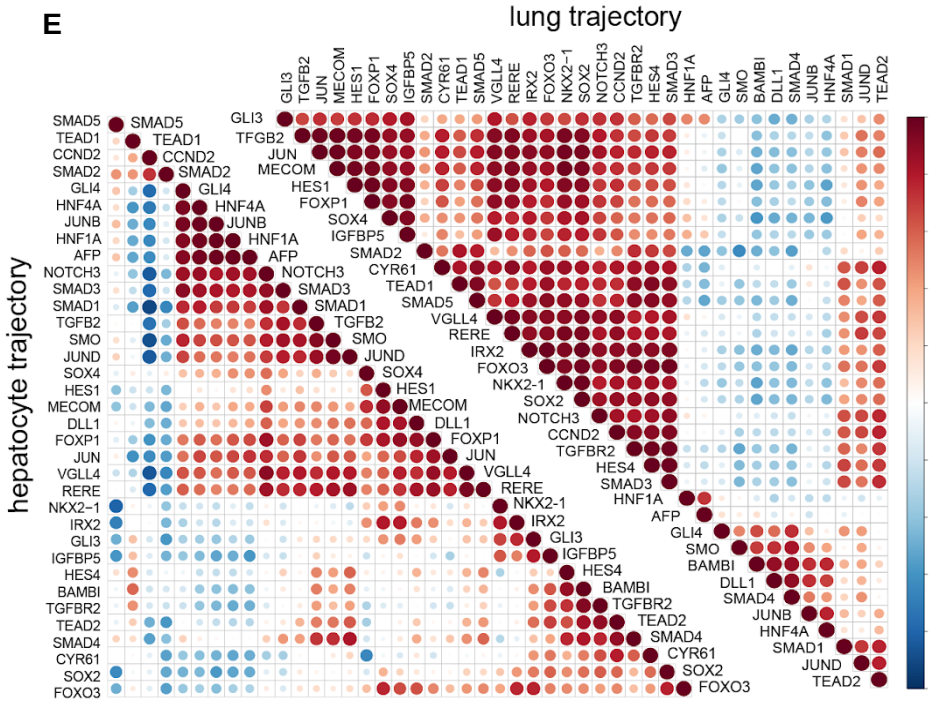
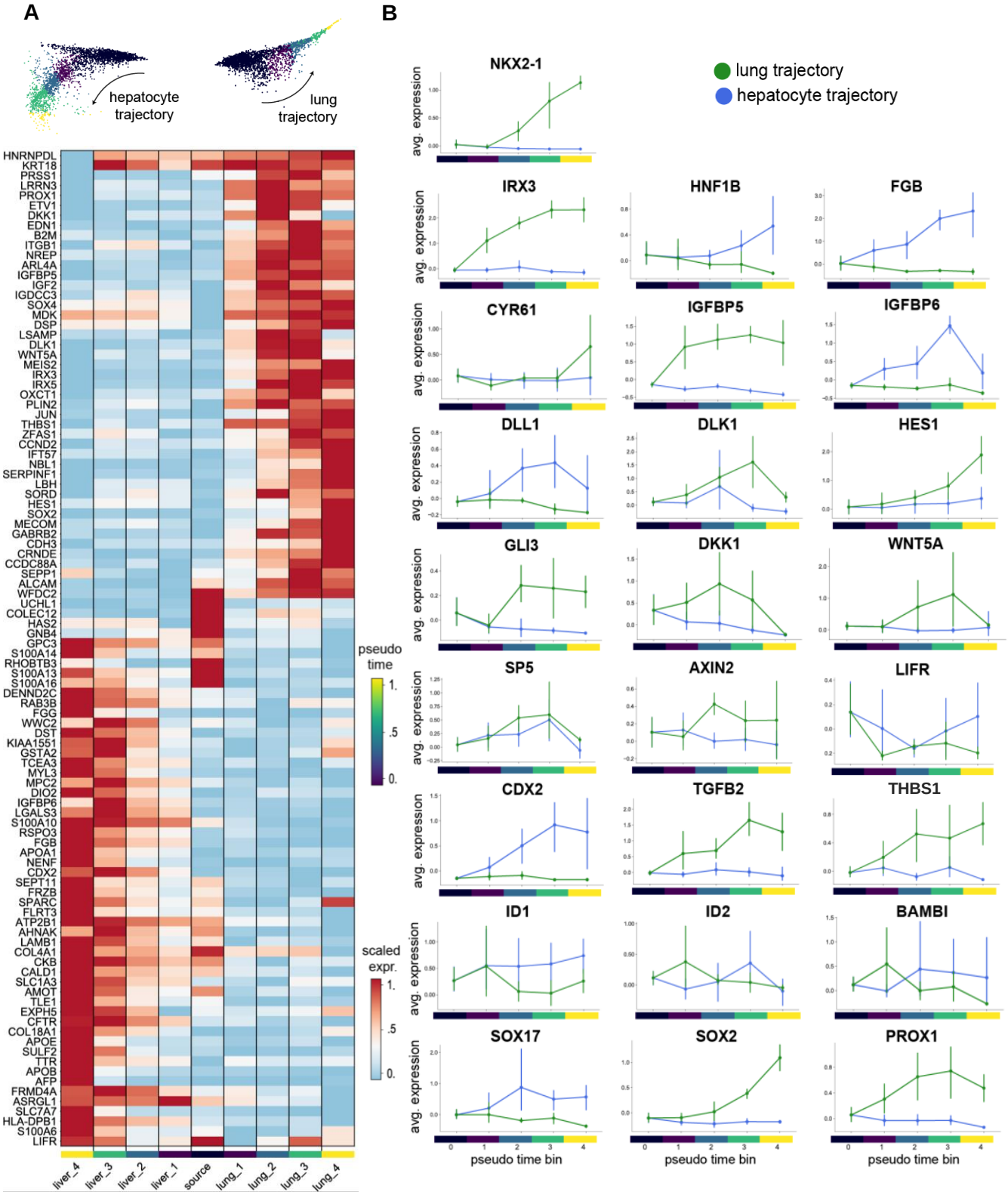


Figure 36. Trajectory analysis reveals major differences in lung and liver branches during differentiation. (A, B) Diffusion maps and implemented Louvain clustering depicting the single cell transcriptomes of FE and early LP stages (days 7-10 and 11-15, respectively) **(B)**. **(C, D)** Color-coded scores illustrating the affinity of single cell transcriptomes to the data derived from bulk mRNA-seq analysis of sorted eGFP+ (**C**) and eGFP- populations on day 15 of differentiation **(D)**. **(E)** Pairwise gene Pearson correlation profiles along the lung and liver pseudotime trajectories (courtesy of F. Uzbas). Adapted from (Ori et al., 2021).

Furthermore, the pseudotime trajectories in the last stage of my differentiation protocol enabled me to unravel major differences in lung and liver specific branches (**Figure 36E**, **37A** and **B**) (Ori et al., 2021). The expression of lineage specific markers such as *NKX2-1* and *IRX3*, and *HNF1B*, and *FGB*, was exclusively observed along the two trajectories, respectively (**Figure 37**) (Ori et al., 2021). Furthermore, *GLI3*, which is a key transcriptional repressor regulated by SHH signaling, was upregulated specifically in the lung branch. This could be explained by the exogenous and endogenous activation of the pathway by SHH and *ISL1* respectively (**Figure 35A** and **B**) (Ori et al., 2021). Key components of the Wnt/β-catenin pathway, including *DKK1*, *WNT5A*, *SP5*, as well as *AXIN2*, were considerably higher in the lung branch (**Figure 37A** and **B**) (Ori et al., 2021). The specificity of these pathways for the lung trajectory is of great interest, as exogenous treatment with SHH and CHIR99021 (**Figure 12**) did not induce the pathways in the adjacent hepatoblast cells (**Figure 18C**). Taken together with previous studies showing a central role of SHH and Wnt/β-catenin in FE-early lung development (Goss et al., 2009; Harris-Johnson et al., 2009; Motoyama et al., 1998), these findings suggest that early hepatocytes might be refractory to these pathways. Accordingly, the exclusive expression of *SOX2* in the lung trajectory is a further evidence for the lineage specific activity of Wnt/β-catenin, as SOX2, canonical Wnt signaling, and FGFs often

cooperate in the regulation of self-renewal in development (**Figure 37B**) (Turner et al., 2014; Ori et al., 2021).



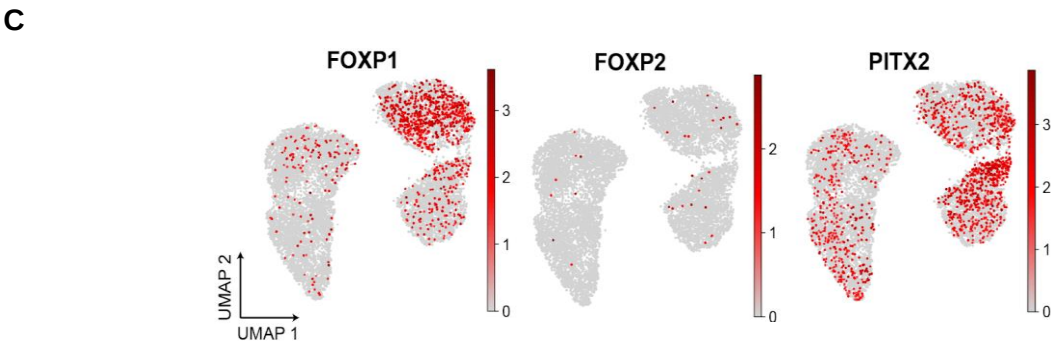


Figure 37. Reconstructing the transition from FE to lung and liver progenitors. (A) Heatmap demonstrating the gene expression of top differentially expressed genes along the "lung" and "hepatocyte" pseudotime trajectories. **(B)** Line plots for the displayed genes, exhibiting their expression patterns in the corresponding trajectories, based on the binned pseudotime ordering (vertical lines represent confidence intervals of 95%). **(C)** UMAP of all time points demonstrating the expression levels of selected genes. Adapted from (Ori et al., 2021).

Last, the pseudotime ordering of the cells revealed a surprising involvement of key developmental genes in hepatoblasts and lung progenitor cells. I found that the expression of *CDX2*, *SOX17*, and LIF/STAT3 signaling via *LIFR* was enhanced in the liver trajectory (**Figure 37B**), although studies in mice did not imply fundamental roles of these factors in liver development (Kumar et al., 2019; Onishi and Zandstra, 2015; Spence et al., 2009). In addition, *PROX1* that is a marker for the developing liver and pancreas in the mammalian foregut (Burke and Oliver, 2002), was instead expressed in the lung trajectory in my data (**Figure 37B**) (Ori et al., 2021). Last, *FOXP1* and *PITX2* expression spanned both the lung and liver trajectories, whereas *FOXP2* was not detected (**Figure 37C**) (Ori et al., 2021), in contrast to the specific expression in lung progenitor cells illustrated in the mouse embryo (Lin et al., 1999; Shu et al., 2007). This suggests differences in the regulation of key developmental genes between mouse and human, which we integrated in a comprehensive model for human lung and hepatoblast differentiation starting from the pluripotency state (**Figure 39**) (Ori et al., 2021). Taken together, these results indicate that common foregut endoderm progenitors produced by our protocol might be plastic enough to support the development of at least two foregut lineages, and draw a comprehensive roadmap for human lung and hepatocyte development from PSC.

4.6. IMPLEMENTION OF TRANSCRIPTOMIC SIGNATURE FEATURES TO THE DIFFERENTIATION STRATEGIES

4.6.1. NOTCH and TGF-β inhibition affect negatively the lung differentiation efficiency

Importantly, by investigating the gene kinetics of elements of the Notch signaling pathway, I found evidence of activation of the Notch signaling pathway specifically in the lung branch. In particular, this implementation was based on the lung branch specific expression of *HES1*, a TF of signaling Notch pathway, and the Notch pathway modulator *DLK1*, known to be involved in lung branching and morphogenesis (**Figure 36** and **37B**) (Falix et al., 2012; Ori et al., 2021). In order to test mechanistically whether Notch is involved in lung specification, I sought to inhibit the Notch by applying in my system the *γ*-secretase inhibitor DAPT from day 11 onwards (**Figure 38A** and **B**) (Ori et al., 2021). This treatment led to a significant decrease in the number of NKX2-1+ progenitors, a decline in *NKX2-1* expression, and a reciprocal increase of *AFP* expression. Strikingly, the hypothesis was further supported by the specific expression of *DLL1* in the liver branch. DLL1 is key ligand of Notch pathway, and given that NOTCH mediates cell-to-cell communication, those results indicate that hepatoblasts may promote Notch signaling in lung progenitors by paracrine signaling (**Figure 38B**). Similarly, after noting exclusive expression of *THBS1* and *TGFB2* components of TGF-β pathway in lung branch expression, I also sought to inhibit the TGF-β pathway using SB431542. A similar effect to Notch inhibition was observed, characterized by a decrease in the number of eGFP+ progenitors and expression of *NKX2-1* in treated cells (**Figure 38A** and **B**) (Ori et al., 2021). Taken together, this suggests a role of Notch and TGF-β in the specification of the foregut towards lung progenitors.

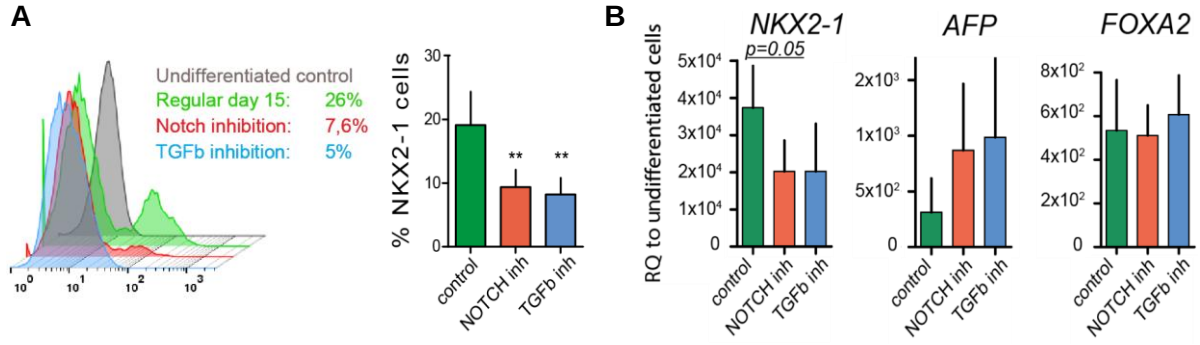


Figure 38. Inhibition of Notch and TGF-β signaling pathways on the FE induces liver specification at the expense of lung progenitors. (A) Quantification of eGFP positive cells, analyzed on day 15 of differentiation by flow cytometry. Indicated inhibition of Notch or TGF-β was performed by treatment with DAPT or SB431542, respectively (bars represent mean ± SD, n = 4 biological replicates, **p < 0.01 by unpaired, one-tailed t-test). **(B)** Fold change of *NKX2-1*, *AFP* and *FOXA2* gene expression on day 15 relative to undifferentiated *NKX2-1*$^{eGFP/+}$ cells (quantified by RT-qPCR ($2^{-\Delta\Delta CT}$), bars represent mean + SD, n = 3 biological replicates, unpaired, one-tailed t-test). Adapted from (Ori et al., 2021).

5. DISCUSSION

5.1. ANALYZING THE STEPWISE DIFFERENTIATION OF HPSCS TO LUNG PROGENITORS

Despite the progress made in stem cell differentiation approaches towards desired lineages, cellular heterogeneity remains a barrier during PSC specification, especially in long lasting differentiation protocols, such as lung differentiation. To obtain high efficiency, which in this case is translated in high yields of lung progenitors, it is crucial to begin with generation of a homogeneous definitive endoderm cellular population as basis for subsequent steps of differentiation protocol. This begins with validating the stepwise differentiation strategy.

5.1.1. Activin/Nodal and Wnt efficiently produce definitive endoderm

Efficient derivation of definitive endoderm provides the basis for many directed differentiation approaches, as it is the prerequisite for PSC differentiation to multiple mature endoderm derivatives. Generation of high purity definitive endoderm is essential for the establishment of an efficient lung differentiation protocol. However, several studies have shown that the differentiation capacity among stem cell lines, as well as their adaption to different protocols varies, with certain cell lines differentiating better into derivatives of a specific germ layer under specific basal media (Bock et al., 2011; Boulting et al., 2011; Huang et al., 2014; Osafune et al., 2008). By comparing existing protocols, whose common main drivers were Activin/Nodal and Wnt/β-catenin signaling, I was able to generate a cell population with ~95% cells committed to definitive endoderm by day 6 of differentiation (**Figure 10**) and thus, concluded that Nodal and Wnt signaling are sufficient for successful definitive endoderm generation although additional growth factors have been previously suggested to be required, such as BMP4 and/or FGF2 (Green et al., 2011; Huang et al., 2014; Sampaziotis et al., 2017).

5.1.2. Patterning of definitive endoderm to foregut endoderm

Patterning of definitive endoderm into distinct progenitor domains depends on a variety of reciprocal signaling interactions between endoderm and surrounding mesoderm (Kraus and Grapin-Botton, 2012; Rankin et al., 2016; Zaret, 2016; Zorn and Wells, 2009). Induction of foregut endoderm is crucial in PSC differentiation protocols aiming to generate DE derivatives, namely lung-, thyroid- and thymus-competent progenitors (Green et al., 2011; Longmire et al., 2012; Mou et al., 2012).

Green and colleagues successfully established the specification of foregut endoderm fate by inhibiting BMP and TGF-β signalling (Green et al., 2011). Upon testing multiple combinations of morphogens and inhibitors, they found that efficient foregut endoderm generation from DE took place upon dual SMAD inhibition. This strategy was adopted by me and others including Gotoh et al. (2014) and Hawkins et al. (2017) to induce lung epithelial cell differentiation. However, the reproducibility of the experiments was not satisfying due to the occasionally low cell survival rate during the DE to FE transition of the culture conditions. Such issues, lead to unsuccessful continuation of the protocol towards the lung progenitor stage. Sonic hedgehog signaling is required for the formation of foregut while suppressing more posterior derivatives (Deutsch et al., 2001; Litingtung et al., 1998; Motoyama et al., 1998). Consistently, supplementation of SHH during the transition from definitive to foregut endoderm lead to FOXA2+ cells (**Figure 11**). The combination of SHH with IMDM-based basal media improved the overall cell survival, led to higher reproducibility, and higher yields of NKX2-1+ cells at the lung progenitor stage (**Figure 12**).

5.1.3. Human PSCs differentiate in parallel to lung progenitors and to early liver cells

Cellular heterogeneity is a trait in lung differentiation approaches (Gotoh et al., 2014; Hawkins et al., 2017; Konishi et al., 2016; McCauley et al., 2017) and it has not been solved to date. The use of a targeted NKX2-1 fluorescent reporter cell line *NKX2-1*$^{eGFP/+}$ (Olmer et al., 2019) not only enabled the monitoring of NKX2-1 expression during the differentiation, but also facilitated the protocol optimization by permitting the quantification and isolation of NKX2-1+ and NKX2-1- populations. Reproducible induction of lung progenitors was also verified by gene expression analysis. Interestingly, analysis of the eGFP- cells revealed the emergence of hepatocytes in cell culture (**Figures 15** and **18**). Mutually exclusive expression of AFP and NKX2-1 (**Figure 18**) led me to the conclusion that co-induction of lung and liver fates had taken place, raising questions concerning the mechanisms which drive the specification of these distinct lineages from foregut endoderm. One potential explanation could implement the close proximity of hepatocyte and lung primordia in the primitive gut tube (**Figure 3**) (Han et al., 2020; Zorn and Wells, 2009). It is known that mammalian organogenesis is driven by signaling pathways that act in a gradient and time dependent manner. Deviations of lung differentiation factors either in concentration or in the induction time point may result in the presence of additional endodermal lineages, found in close proximity to the lung *in vivo*.

In conclusion, further optimization of the current protocol is required for the improvement of lung differentiation efficiency and the reciprocal suppression of unwanted endodermal lineages.

5.2. PROPERTIES OF PSC-DERIVED LUNG EPITHELIAL CELLS

5.2.1. NKX2-1+ progenitors can expand *in vitro* and give rise to lung organoids

Lung organoid model systems can greatly increase our ability to study lung development and disease. Furthermore, in differentiation approaches it is crucial to prove that the differentiated progenitor cells are functional. One potential way of studying the functionality of my lung progenitors would be to use them as a starting population for an organoid assay. The early lung progenitors would be able to differentiate and generate respiratory organoids, ensuring further possibilities for promising developmental studies. In line with these, I reliably verified the developmental potential of putative eGFP+ lung progenitors using 3D culture assay (**Figure 13** and **14**). On that count, the colony shaped organization of the eGFP+ cells that my protocol generated (**Figure 12** and **13**) provided an advantage to the subsequent organoid differentiation compared to other protocols. Currently available approaches use FACS sorting for isolating the NKX2-1+ cells for subsequent lung organoid derivation (Gotoh et al., 2014; Hawkins et al., 2017). This step requires the availability of large number of NKX2-1+ progenitor cells, in order to proceed with the 3D organoid cultures. My approach leads to colony formation, which allows the picking of several colonies even in experimental attempts, where the overall differentiation efficiency is not very high, providing material for the continuation of the differentiation. The convex shape of the colonies combined with the expression of key genes known to induce lung branching, implies a lung progenitor capacity similar to the cells in the embryonic lung where budding initiates (**Figure 3B, 13** and **15**).

Notably, as predicted by mouse genetic models and previous human PCS studies, activation of Wnt/β-catenin pathway promoted branching morphogenesis and expression of markers characteristic of alveolar type II cells (**Figure 14**). Wnt signaling has crucial role in embryonic lung formation as well as lung repair and regeneration. There is evidence that Wnt seems to impact proximodistal epithelial patterning during development. In particular, it has been observed in mouse studies that Wnt activation favours distal over proximal epithelial patterning in early post-lung specification and it is needed for alveolar proliferation (Hashimoto et al., 2012; Mucenski et al., 2003; Shu et al., 2005; Zemke et al., 2009). Furthermore, Wnt has been shown to have a similar role

during human lung development by promoting distal fate (McCauley et al., 2017). When Wnt was activated via CHIR99021, branching morphogenesis took place as well as up-regulation of the expression of *SFTPC*, a gene expressed at the distal-alveoli space. A stronger proximodistal patterning effect was observed when DCIC was combined with inhibition of TGF-β via SB431542, with the organoids demonstrating high expression of alveolar epithelial type II cell markers, namely, *SFTPC* and *SFTPB* (**Figure 14**). This parallels the previous studies showing TGF-β inhibiting the expression of ATII cell markers and promoting transdifferentiation of ATII cells into ATI cells (Zhao et al., 2013). Furthermore, TGF-β ligands have shown to negatively regulate secondary branching morphogenesis during lung growth (Serra et al., 1994). To summarize, I demonstrated that eGFP+ lung progenitor cells are plastic and exhibit multipotent progenitor capacity as they are able to differentiate towards several lung cell types depending on the provided growth conditions.

5.2.2. Transcriptomic analysis of lung progenitors and fetal distal lung epithelial cells reveals shared gene regulatory networks

Advances in RNA sequencing (RNA-seq) approaches have allowed for a more detailed characterization of lung tissue. Transcriptomic analysis on micro-dissected stalks and branching epithelial tips of human fetal lung was recently performed (Miller et al., 2018; Nikolić et al., 2017). These studies showed that human tip epithelium is analogous to the mouse population and has a highly conserved transcriptome. In particular, comparison of the human and mouse epithelial tip cell transcriptomes has demonstrated that 96% of orthologous genes expressed in human tips are also present in mouse (Laresgoiti et al., 2016; Nikolić et al., 2017). Despite the high level of transcriptome conservation between human and mouse tips, different media are required for *in vitro* long-term self-renewal of mouse and human tips (Nikolić et al., 2017). These differences reflect molecular variations between human and mouse, questioning the use of mouse lungs as a model system to study human development and indicating the need of more reliable models.

Analysis of the transcriptome data sets of lung progenitor-derived organoids and the human fetal lung data showed higher correlation of the DCICS-treated organoids to human epithelial tips, while DCI-treated organoids correlated better to stalk (**Figure 15E**), supporting the outcome of the proximodistal patterning of lung progenitors upon respective treatment (**Figures 14** and **15**). Furthermore, correlation was observed between the transcriptome of distal human tips and hiPS-derived lung progenitors (**Figure 15E**). Importantly, human epithelial tips have demonstrated self-

renewal and differentiation capacities. In particular, it was suggested that human tips are multipotent progenitors at pseudoglandular stage capable of generating bronchial and alveolar epithelium *in vitro* (Miller et al., 2018; Nikolić et al., 2017; Rawlins et al., 2009). Hence, these findings imply the existence of a progenitor population in distal lung epithelium that shares transcriptomic characteristics with human early lung progenitors, possibly allowing the distal epithelium to maintain its multipotent capacities.

Based on the results of the comparative transcriptome analysis it can be concluded that human PSC-derived lung cells provide a valid research strategy in lung stem cell field. Their primary research implementation would be the identification of signaling requirements necessary for long-term self-renewal, the culturing of human tips as differentiation-competent organoids, and the generation of mature cell lineages. Human PSCs derived lung models can aid in overcoming the limited accessibility to human fetal tissues, as well as the simultaneous need of optimizing the culture conditions. As such iPSC derived lung models can push forward studies focusing on the use of autologous primary cells overcoming the limitations in clinical practice.

5.2.3. Extended cultivation of LP-derived organoids promotes maturation of respiratory organoids

Several groups have aimed at characterizing the gene expression changes throughout lung development (Kho et al., 2010; Roost et al., 2015). The availability of such data sets allowed me to correlate the transcriptomic signature of my iPSC-derived organoids and draw conclusions on their maturation. My results suggested that long-term culture of organoids induces additional maturation. In particular, prolonged culture of hPSC-derived organoids resulted in an increased expression of alveolar type II markers and up-regulation of additional genes involved in alveolar differentiation, including *NAPSA*, *LAMP3*, *CTSH*, *SFTPB*, *SFTPC*, *CLDN18* (Beauchemin et al., 2016) (**Figure 16**). Furthermore, comparison of my data set to the transcriptome of the developing human lung demonstrated better correlation to the later developmental stages (**Figure 17**). Collectively my data demonstrates the potential of further organoid maturation *in vitro* upon organoid maintenance in culture for a longer period under suitable conditions. This phenomenon was also observed in other organoid systems, where long-term cultivation of hPSC-derived cerebral organoids induces maturation of such organoids (Matsui et al., 2018). Taking into consideration the data that has been

previously described (Kho et al., 2010; Nikolić et al., 2017; Roost et al., 2015), I conclude that lung organoids on day 35 represent early developmental stages (pseudoglandular to canalicular).

In addition, passaged lung organoids maintained the eGFP+ expression but they did not demonstrate branching morphogenesis. Instead, they were rather growing as spheroids (**Figure 16**), similar to primary alveolar spheres (Barkauskas et al., 2013). The phenotype of DCICS organoids on day 35, the transcriptomic data, as well as the accumulation of SFTPC staining on the border of the organoid branches, indicated that the DCICS organoids constitute a rather multipotent bronchioalveolar-like structure, with distal alveolar-like structures. Also, the passaged organoids seem to represent a more advanced developmental stage of lung organogenesis (**Figure 17**), albeit the high expression of distal lung markers, the correlation to human fetal lungs was diminished.

5.3. ECTOPIC EXPRESSION OF NKX2-1 IS NOT SUFFICIENT FOR THE INDUCTION OF RESPIRATORY CELL FATE

Derivation of specific cell types from PSCs involves application of extrinsic factors to guide the differentiation process toward the cell type of interest, thereby mimicking the *in vivo* development. However, this approach usually faces limitations such as low purity, inefficiency and time-consuming culture conditions. Given that the sequential administration of signaling molecules coincides with the activation of transcriptional programs, transcription-factor mediated specification of PSCs via overexpression of cell type-specific TFs represents an alternative method to guide cell fate acquisition. NKX2-1 is the earliest and most studied TF for pulmonary lineage, which makes it a good candidate for such a strategy. I attempted to promote lung specification in hPSCs as well as to directly convert somatic cells (fibroblasts) via ectopic overexpression of *NKX2-1* (**Figure 19-22**).

By means of a transcription factor-mediated lung specification strategy, I was able to induce the upregulation of lung specific genes upon ectopic overexpression of the transcription factor NKX2-1. However this strategy did not result in a functional lung epithelial population. BMP4 and RA administration seemed to favor the expression of lung specific genes *SFTPB*, *SFTPC* and *MUC5AC*, when coupled with two days of *NKX2-1* overexpression (**Figure 22**), while supplementation of VPA in the culture medium increased the expression of NKX2-1 target gene, *SFTPB* in the acceleration approach (**Figure 22**). Treatment with VPA was hypothesized to contribute to the lung lineage specification upon NKX2-1 overexpression; given that HDAC inhibitors have been shown to improve the efficiency in other reprogramming studies (Cheng et al.,

2014; Göttlicher et al., 2001; Huangfu et al., 2008; Kogiso et al., 2013; Thoma et al., 2014). Furthermore, generation of DE led to significant increase of *SFTPB* and *SCGB1A1* expression in DCI-treated organoids, resulting in endogenously NKX2-1-expressed organoid morphologies similar to the ones obtained during regular step-wise differentiation approaches (**Figure 23**). However, DE did not favor the expression of *SFTPC* suggesting that the resulting outgrowths rather contained lung progenitor cells or club cells.

The overexpression of key transcription factors represent a well-established approach to manipulate cellular identities and kickstart the transdifferentiation process, with the derivation of iPSCs from somatic cells being the most prominent example (Takahashi et al., 2007; Takahashi and Yamanaka, 2006). In my attempt to investigate whether ectopic expression of NKX2-1 can induce transdifferentiation of fibroblasts into lung progenitor cells, I generated secondary fibroblasts and subsequently utilized them for lung transdifferentiation experiments (**Figure 24**). It is known that directed differentiation of hPSCs varies in specificity and efficiency, as cell lines fluctuate in their propensity to differentiate into particular lineages due to genetic factors (Carcamo-Orive et al., 2017; DeBoever et al., 2017; Kilpinen et al., 2017), epigenetic memory of donor material (Bar-Nur et al., 2011; Kim et al., 2010), as well as activation of endogenous signaling pathways (Nazareth et al., 2013; Strano et al., 2020). In my experimental set up, the differentiation propensity of both hPSC (*iNKX2-1-eGFP-H9* and *iNKX2-1-mCherry-NKX2-1*$^{eGFP/+}$) lines was very efficient as both lines demonstrated high expression levels of fibroblast markers *SMA*, *VIM* and *FSP1*, while simultaneously down-regulating the pluripotency factors. The comparison of the expression levels of fibroblasts specific markers of the hPSC-derived secondary fibroblasts and the primary skin fibroblasts showed that all marker genes were less up-regulated in H9-derived fibroblasts compared to hiPSCs fibroblasts (**Figure 24**). Importantly, even if mRNA from primary skin fibroblasts was used as a reference, there is no evidence suggesting that the secondary fibroblasts resemble skin fibroblasts. Similarly to the acceleration approach, a short induction of exogenous *NKX2-1* combined with BMP4 and RA supplementation together with CKF resulted in a promising experimental outcome. Although *FSP1* was down-regulated for both cell lines, *VIM* and *SMA* remained upregulated when compared to LP reference samples. Furthermore, *FOXA2* exhibited a slight upregulation in CKF+BR treated iPSC-derived fibroblasts, expression of *EpCAM* as well as *E-Cad* remained low (**Figure 25**). Taken together these experimental results indicated the need to identify

additional factors, which would be essential towards overcoming the MET barrier and lead to the epithelialization of the cultures.

5.4. RECONSTRUCTION OF TRANSCRIPTIONAL NETWORKS GOVERNING THE GENERATION OF EARLY LUNG PROGENITORS BY SINGLE CELL RNA-SEQ

The heterogeneity occurring during the PSCS-driven lung differentiation (**Figures 12, 18, 27, 36** and **37**), as well as the insufficiency of NKX2-1 ectopic expression to induce a respiratory cell fate (**Figures 22-25**) created the necessity of generating the type of data set, which could be used to elucidate the transcriptional networks interacting during the step wise differentiation of the early lung progenitor cells. Such data would provide invaluable insights for the lung stem cell biology field, which has been lagging behind compared to other organs. By performing single cell sequencing, I aimed to uncover complex developmental gene trajectories governing the specification mechanisms distinguishing the lung epithelium from other closely related endodermal lineages (**Figures 26-29**). The densely sampled large-scale single-cell resolution enabled me to resolve cellular heterogeneity, identify cell-to-cell variations and uncover cell state transitions in comparison to standard methods involving the collection of large cell number for deep sequencing (bulk RNA-seq) at single or multiple time points as performed in previous studies (Hawkins et al., 2017; Yamamoto et al., 2017).

5.4.1. Single cell RNA-Sequencing identifies stage-specific regulation and signaling pathways

Although hPSCs-derived LPs have been isolated and described (**Figures 15** and **18**) (Hawkins et al., 2017), the transcriptional networks underlying the human early lung commitment and patterning remained unclear, as well as the mechanisms involved in the presence of mutually existed lineages *in vitro*. The transcriptional profiling of single cells represents a promising strategy for obtaining a global view of these developmental processes. Performing single-cell RNA-seq time-course, during the *in vitro* directed differentiation of iPSC-derived LP, I was able to reconstruct the lung-like lineage trajectory from pluripotency, identify the fundamental transcriptional networks and the signaling pathways dictating early lung specification from hPSCs (**Figures 29** and **30**).

The cellular heterogeneity observed in *in vitro* approaches is also a feature of organ development, wherein progenitor cell populations, histologically indistinguishable, undergo diverse cell fate decisions to become specified cell types. In order to comprehend the signals regulating cell

fate decisions, it is necessary to uncover the early gene expression events associated with the differentiation trajectories of individual cells. Studies that use bulk RNA samples from pooled progenitor cells can only generate a virtual average of the diverse constituent cells, without allowing the unraveling of signals that drive a progenitor down a specific lineage. In contrast, single-cell based analysis allows the molecular discrimination of various cell types within a heterogeneous population, such as a developing organ. Dissection of heterogeneity in developmental or differentiating cellular populations, previously hidden, has the potential to uncover detailed cell state transitions, while allowing to pinpoint the exact emergence of cellular state. Although single-cell time-course studies have been performed for hPSC-derived putative respiratory cells from hPSC-derived lung progenitors (Hurley et al., 2020; McCauley et al., 2017), the developmental questions concerning the commitment of early endodermal lung specification remain open. Researchers have generated proximal and distal lung epithelial cells from hPSCs and aimed to profile the kinetics of these lineages over time, as well as to illustrate that relative plasticity of various proximal and distal putative PSC-derived lung lineages can be regulated. These studies confirmed that similarly to mouse, Wnt signaling in human is involved in the regulation of proximal-distal patterning of the lung (Hashimoto et al., 2012; Mucenski et al., 2003; Shu et al., 2005; Zemke et al., 2009). This observation is consistent with my findings, as observed by the distalization of the lung epithelium upon CHIR99021 treatment, which leads to Wnt activation (**Figures 14** and **15**). However, these studies focus on the global transcriptomic profiling of cells already committed to the lung lineage (upon NKX2-1 expression) and are further differentiated into specific respiratory lineages (Hurley et al., 2020; McCauley et al., 2017). Hawkins et al. (2017) sought to model the global gene expression kinetics of early human lung specification using microarray analysis, an approach which allowed only for single point experiments, representing the key stages of iPSC lung-directed differentiation, without permitting unraveling of cell state transitions.

To chart a comprehensive map of transcriptional changes driving the differentiation of hPSCs to lung progenitors, single-cell RNA-seq time-course was performed. Analysis revealed three transcriptional domains that corresponded to the three major developmental stages (**Figures 26-29**) and allowed the identification of candidate signaling cascades, regulating the differentiation process towards lung progenitors from human PSCs (**Figure 30**). Importantly, genes characteristic to hepatoblasts were detected earlier then the LP stage (**Figure 29**). The disparities in early induction of endoderm raised questions regarding the origin and identity of the hepatic cells, which neighbor

NKX2-1+ progenitors during lung differentiation. As results, it was important to research each stage individually in order to identify genes and mechanisms promoting the emergence of lung and liver progenitor cells respectively.

5.4.1.1. Specification of foregut endoderm coincides with initiation of lung and liver hallmarks

Differentiation approaches for lung and hepatic progenitor generation are similar in their initial activation of Activin/Nodal pathway to specify DE. However, in contrast to lung differentiation protocols, hepatic induction is usually based on FGF2 and/or BMP4 treatment in the DE stage (Hannan et al., 2013). During the second stage of hepatocyte differentiation some protocols include SMAD inhibition, similar to the currently used methods for lung differentiation (Green et al., 2011), while others utilize HGF administration (Carpentier et al., 2014). Interestingly, in my protocol, hepatoblasts appeared without requiring stimulation by exogenous FGF2, BMP4, or HGF. The appearance of both lung and liver cells during the differentiation protocol raised the question whether a common progenitor give rise to the two lineages. My trajectory analysis did not find strong evidence in favor of bipotent lung-liver progenitors. It was rather observed that hepatoblast transcription networks were established earlier than the initiation of lung program, specifically at the DE stage, with the expression of *HHEX*, *HNF4A* and *HNF1B* (**Figures 31-35**) (Ori et al., 2021).

Mechanistic studies in mice have demonstrated that activation of the regulatory loop of Activin/Nodal and EOMES induces TFs involved in early hepatogenesis, including *Gata6* and *Gata4* (Zhao et al., 2005), *Foxa2* and *Foxa1* (Lee et al., 2005). These TFs have been shown to pioneer chromatin remodeling in foregut precursors (Zaret and Carroll, 2011), and have been used as drivers of hepatocyte transdifferentiation (Chen et al., 2018; Wamaitha et al., 2015). Furthermore, GATA4 and GATA6 are functionally redundant in promoting liver growth, and together with FOXA2 they promote liver specification through the induction of Hhex (Denson et al., 2000). My data suggests a capability of human DE to directly convert into hepatoblasts, as *HHEX* is already detected at DE, in contrast to *ISL1* that is not expressed (**Figure 31**) (Hunter et al., 2007; Kim et al., 2019). Remarkably, in my data, several hepatoblast TFs e.g. *HNF4A*, *TBX3*, and liver markers e.g. *APOA2*, *FGB, TTR, APOB* were induced shortly after *HHEX*, which is evidence supporting the "fast track" liver differentiation model from early DE (**Figure 39**) (Ori et al., 2021).

It remains unclear how *NKX2-1* is upregulated during foregut endoderm stage. We performed analysis to pinpoint the activity of *FOXA2* and *ISL1* during the FE stage. The expression of the TF *FOXA2* is upregulated at DE stage and persists in the FE stage, while *EOMES, MIXL1* and *GSC* are down-regulated. *FOXA2* was shown to activate *ISL1* in cardiac progenitors (Kang et al., 2009), *ISL1* was shown to activate *NKX2-1* in the early foregut as a prerequisite for lung generation (Kim et al., 2019) and it has been suggested that ISL1 regulates *Shh* expression in the foregut (Lin et al., 2006). In my study I showed that SHH treatment increased the efficiency of NKX2-1+ LP generation (**Figure 12**). In addition, the link between ISL1 and SHH is supported by developmental abnormalities in the cardiovascular system, specifically in *Isl1* and *Smo* mutant mouse embryos (Lin et al., 2006). Importantly, the Foxa2, Otx2 and Lhx1 complexes have been involved in autoregulation and pioneering the transcription of foregut genes (Costello et al., 2015). My data suggests a similar process could control the lung progenitor differentiation *in vitro*, placing *FOXA2* and *OTX2* in an autoregulatory loop during the formation of DE which induces the expression of *ISL1* and then *NKX2-1* (**Figure 31-37**) (Ori et al., 2021).

5.4.1.2. Identification of key pathways regulating lung specification

Expansion of the existent knowledge regarding the transcriptional profile and signaling pathways involved in human lung specification is important for the establishment of more efficient and robust differentiation protocols, with an emphasis on the manipulation of hPSCs cultures in a logical and dynamic manner. This can be achieved through the combinatorial activation and repression via supplementation of varying dosages of extracellular signals with the aim of systematically restricting the formation of unwanted cell types. Pathway analysis was performed to understand the developmental program towards lung specification. This revealed enrichment of Hippo, TGF-β, and Notch signaling pathway components (**Figures 30, 36** and **37**). In line with mouse genetic models, Hippo signaling was shown to be important for lung development by promoting lung epithelial lineage commitment, while lung specific *Mst1/2* double mutants died postnatally within the first 15 days (Chung et al., 2013). I revealed a potentially earlier role of Hippo signaling in FE stage during the lung specification. Furthermore, expression of components of TGF-β superfamily namely *THBS1, TGFB2* and *JUN* are shown to be players of lung lineage specification (**Figures 36** and **37**). Intriguingly, recent work from Ikonomou and colleagues sought to identify the pathways that regulate the cell-fate determination of lung primordia from the pre-specified embryonic foregut. The

authors demonstrated, in consistence to my findings, upregulation of TGF-β pathway in lung primordium upon specification. Furthermore, regarding Hippo signalling they implied a potentially transient activation of YAP/TAZ immediately prior to lung specification (Ikonomou et al., 2020). Given the limitations in investigating the respective stages during human development *in vivo*, such studies provide a great material for validating how closely do *in vitro* generated human progenitors resemble their *in vivo* counterparts.

In addition, Notch plays a major role in the embryonic lung, in regulating the amount of ciliated and secretory club cells during airway differentiation by controlling the balance of ciliated, secretory, and neuroendocrine cells in the airway epithelium. More specifically, conditional inactivation of Notch leads to airways, overpopulated with ciliated cells at the expense of club cells (Morimoto et al., 2012, 2010; Tsao et al., 2009). In the adult lung, Notch is involved in repair and regeneration of airway cell types including club cells, basal cells and lineage-negative epithelial progenitor cells (LNEP) (Rock et al., 2011; Vaughan et al., 2015; Xing et al., 2012). There is accumulating evidence that Notch is also involved in alveolar development by being activated at the distal epithelium of primary lung buds (Post et al., 2000; Tsao et al., 2008). Furthermore, it has been suggested that overexpression of Notch in epithelial cells impairs commitment to alveolar cell fate (Guseh et al., 2009). However, epithelial cell-specific Notch inhibition prenatally by Morimoto and colleagues did not generate an abnormal phenotype in the distal lung (Morimoto et al., 2010), while later disruption of its function at the alveolarization stage and in the distal lung epithelium led to a major deficit in alveolar formation, mainly as a result of the reduced proliferation and maturation of ATII cells (Tsao et al., 2016). Interestingly, those cell types are NKX2-1 expressing cells in adult lung epithelium (Ikonomou et al., 2020).

On that note, I observed that consistent with the data interpretation, inhibition of both Notch and TGF-β signaling pathways favor hepatocyte specification at the expense of lung (**Figure 38**) (Ori et al., 2021). The close proximity of respiratory and liver primordia in the embryonic foregut might demonstrate the difficulty in identifying the proper balance of signals required for the commitment of endodermal progenitors into lung cells. It has been suggested that during early development spatially adjacent cell types in the embryonic gut tube have similar expression profiles compared to distant cell types (Han et al., 2020). Therefore understanding the molecular mechanisms underlying the regionalization of developing foregut into distinct organ domains would push forward the discovery

of optimal cell culture conditions and advance the differentiation of hiPSCs into lung, which has been lagging behind compared to other lineages.

In conclusion, this study provided insights for i) the up-regulation of *SOX2* in the pulmonary branch coinciding with the activation of Notch and Wnt/β-catenin signalling pathways, ii) upregulation of the Hippo pathway components at the FE stage, iii) the decrease of NKX2-1 progenitors upon inhibition of TGF-β and Notch, which combined with iv) Yap- dependent TGF-β-mediated induction of *Sox2* expression in mammalian airway progenitors (Mahoney et al., 2014), creates a putative picture of renewal mechanisms that could contribute to the future efficient generation of human lung stem cells. The novel transcriptional transitions identified in this study which distinguish the lung progenitor generation from hepatoblast are demonstrated in **Figure 39** (Ori et al., 2021).

Understanding of early human lung specification illustrates the trajectories and potential plasticity of endodermal progenitors, reveals transcriptomic signatures unique to the lung endoderm, and advances our understanding of embryonic development in human; providing a comprehensive source for deeper investigation with potential implications in tissue repair and cancer progression.

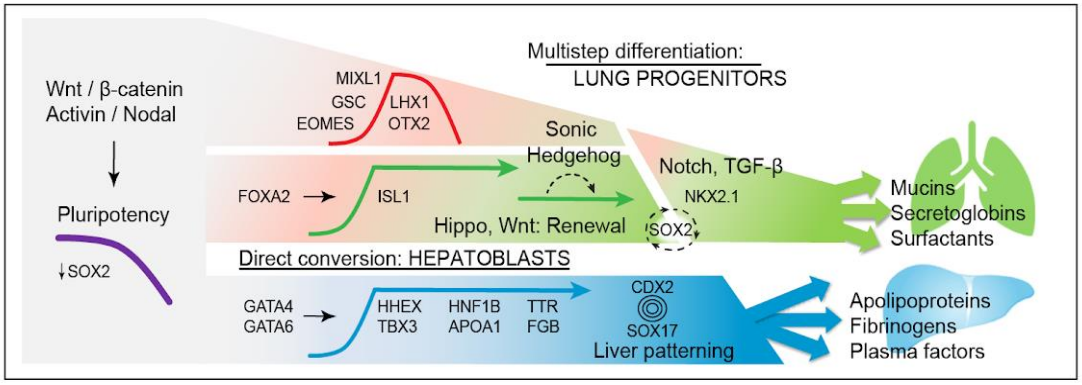


Figure 39. Graphical summary of the identified mechanisms in this study underlying the differentiation of lung and liver progenitors from pluripotency state. Taken from (Ori et al., 2021).

6. LITERATURE

Al Alam, D., Green, M., Tabatabai Irani, R., Parsa, S., Danopoulos, S., Sala, F.G., Branch, J., El Agha, E., Tiozzo, C., Voswinckel, R., Jesudason, E.C., Warburton, D., Bellusci, S., 2011. Contrasting Expression of Canonical Wnt Signaling Reporters TOPGAL, BATGAL and Axin2LacZ during Murine Lung Development and Repair. PLoS ONE 6, e23139. https://doi.org/10.1371/journal.pone.0023139

Amador, C., Weber, C., Varacallo, M., 2022. Anatomy, Thorax, Bronchial, in: StatPearls. StatPearls Publishing, Treasure Island (FL).

Andrade, C.F., Wong, A.P., Waddell, T.K., Keshavjee, S., Liu, M., 2007. Cell-based tissue engineering for lung regeneration. Am. J. Physiol. Lung Cell. Mol. Physiol. 292, L510-518. https://doi.org/10.1152/ajplung.00175.2006

Arcasoy, S.M., Kotloff, R.M., 1999. Lung transplantation. N. Engl. J. Med. 340, 1081–1091. https://doi.org/10.1056/NEJM199904083401406

Ardini-Poleske, M.E., Clark, R.F., Ansong, C., Carson, J.P., Corley, R.A., Deutsch, G.H., Hagood, J.S., Kaminski, N., Mariani, T.J., Potter, S.S., Pryhuber, G.S., Warburton, D., Whitsett, J.A., Palmer, S.M., Ambalavanan, N., The LungMAP Consortium, 2017. LungMAP: The Molecular Atlas of Lung Development Program. Am. J. Physiol.-Lung Cell. Mol. Physiol. 313, L733–L740. https://doi.org/10.1152/ajplung.00139.2017

Arman, E., Haffner-Krausz, R., Chen, Y., Heath, J.K., Lonai, P., 1998. Targeted disruption of fibroblast growth factor (FGF) receptor 2 suggests a role for FGF signaling in pregastrulation mammalian development. Proc. Natl. Acad. Sci. 95, 5082–5087. https://doi.org/10.1073/pnas.95.9.5082

Artus, J., Piliszek, A., Hadjantonakis, A.-K., 2011. The primitive endoderm lineage of the mouse blastocyst: Sequential transcription factor activation and regulation of differentiation by Sox17. Dev. Biol. 350, 393–404. https://doi.org/10.1016/j.ydbio.2010.12.007

Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., Lovell-Badge, R., 2003. Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev. 17, 126–140. https://doi.org/10.1101/gad.224503

Ballard, P.L., Gonzales, L.W., Williams, M.C., Roberts, J.M., Jacobs, M.M., 1991. Differentiation of type II cells during explant culture of human fetal lung is accelerated by endogenous prostanoids and adenosine 3',5'-monophosphate. Endocrinology 128, 2916–2924. https://doi.org/10.1210/endo-128-6-2916

Barkauskas, C.E., Cronce, M.J., Rackley, C.R., Bowie, E.J., Keene, D.R., Stripp, B.R., Randell, S.H., Noble, P.W., Hogan, B.L.M., 2013. Type 2 alveolar cells are stem cells in adult lung. J. Clin. Invest. 123, 3025–3036. https://doi.org/10.1172/JCI68782

Bar-Nur, O., Russ, H.A., Efrat, S., Benvenisty, N., 2011. Epigenetic memory and preferential lineage-specific differentiation in induced pluripotent stem cells derived from human pancreatic islet beta cells. Cell Stem Cell 9, 17–23. https://doi.org/10.1016/j.stem.2011.06.007

Bayha, E., Jørgensen, M.C., Serup, P., Grapin-Botton, A., 2009. Retinoic Acid Signaling Organizes Endodermal Organ Specification along the Entire Antero-Posterior Axis. PLoS ONE 4, e5845. https://doi.org/10.1371/journal.pone.0005845

Beauchemin, K.J., Wells, J.M., Kho, A.T., Philip, V.M., Kamir, D., Kohane, I.S., Graber, J.H., Bult, C.J., 2016. Temporal dynamics of the developing lung transcriptome in three common inbred strains of laboratory mice reveals multiple stages of postnatal alveolar development. PeerJ 4, e2318. https://doi.org/10.7717/peerj.2318

Behjati, S., Lindsay, S., Teichmann, S.A., Haniffa, M., 2018. Mapping human development at single-cell resolution. Development 145, dev152561. https://doi.org/10.1242/dev.152561

Bellusci, S., Furuta, Y., Rush, M.G., Henderson, R., Winnier, G., Hogan, B.L., 1997a. Involvement of Sonic hedgehog (Shh) in mouse embryonic lung growth and morphogenesis. Dev. Camb. Engl. 124, 53–63.

Bellusci, S., Grindley, J., Emoto, H., Itoh, N., Hogan, B.L., 1997b. Fibroblast growth factor 10 (FGF10) and branching morphogenesis in the embryonic mouse lung. Dev. Camb. Engl. 124, 4867–4878.

Bénazéraf, B., Pourquié, O., 2013. Formation and Segmentation of the Vertebrate Body Axis. Annu. Rev. Cell Dev. Biol. 29, 1–26. https://doi.org/10.1146/annurev-cellbio-101011-155703
Berg, T., Cassel, T.N., Schwarze, P.E., Nord, M., 2002. Glucocorticoids regulate the CCSP and CYP2B1 promoters via C/EBPbeta and delta in lung cells. Biochem. Biophys. Res. Commun. 293, 907–912. https://doi.org/10.1016/S0006-291X(02)00319-4
Blakeley, P., Fogarty, N.M.E., del Valle, I., Wamaitha, S.E., Hu, T.X., Elder, K., Snell, P., Christie, L., Robson, P., Niakan, K.K., 2015. Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. Development 142, 3613–3613. https://doi.org/10.1242/dev.131235
Bock, C., Kiskinis, E., Verstappen, G., Gu, H., Boulting, G., Smith, Z.D., Ziller, M., Croft, G.F., Amoroso, M.W., Oakley, D.H., Gnirke, A., Eggan, K., Meissner, A., 2011. Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. Cell 144, 439–452. https://doi.org/10.1016/j.cell.2010.12.032
Boulting, G.L., Kiskinis, E., Croft, G.F., Amoroso, M.W., Oakley, D.H., Wainger, B.J., Williams, D.J., Kahler, D.J., Yamaki, M., Davidow, L., Rodolfa, C.T., Dimos, J.T., Mikkilineni, S., MacDermott, A.B., Woolf, C.J., Henderson, C.E., Wichterle, H., Eggan, K., 2011. A functionally characterized test set of human induced pluripotent stem cells. Nat. Biotechnol. 29, 279–286. https://doi.org/10.1038/nbt.1783
Brauer, P.R., 2003. Human embryology: the ultimate USMLE step 1 review. Hanley & Belfus, Philadelphia, Pa.
Burke, Z., Oliver, G., 2002. Prox1 is an early specific marker for the developing liver and pancreas in the mammalian foregut endoderm. Mech Dev 118, 147–155.
Burri, P.H., 1984. Fetal and Postnatal Development of the Lung. Annu. Rev. Physiol. 46, 617–628. https://doi.org/10.1146/annurev.ph.46.030184.003153
Burtscher, I., Lickert, H., 2009. Foxa2 regulates polarity and epithelialization in the endoderm germ layer of the mouse embryo. Dev. Camb. Engl. 136, 1029–1038. https://doi.org/10.1242/dev.028415
Carcamo-Orive, I., Hoffman, G.E., Cundiff, P., Beckmann, N.D., D'Souza, S.L., Knowles, J.W., Patel, A., Papatsenko, D., Abbasi, F., Reaven, G.M., Whalen, S., Lee, P., Shahbazi, M., Henrion, M.Y.R., Zhu, K., Wang, S., Roussos, P., Schadt, E.E., Pandey, G., Chang, R., Quertermous, T., Lemischka, I., 2017. Analysis of Transcriptional Variability in a Large Human iPSC Library Reveals Genetic and Non-genetic Determinants of Heterogeneity. Cell Stem Cell 20, 518-532.e9. https://doi.org/10.1016/j.stem.2016.11.005
Cardoso, W.V., 2006. Regulation of early lung morphogenesis: questions, facts and controversies. Development 133, 1611–1624. https://doi.org/10.1242/dev.02310
Carlson, B.M., 2014. Human embryology and developmental biology, 5th ed. ed. Elsevier/Saunders, Philadelphia, PA.
Carpentier, A., Tesfaye, A., Chu, V., Nimgaonkar, I., Zhang, F., Lee, S.B., Thorgeirsson, S.S., Feinstone, S.M., Liang, T.J., 2014. Engrafted human stem cell-derived hepatocytes establish an infectious HCV murine model. J Clin Invest 124, 4953–4964.
Chakraborty, S., Ji, H., Kabadi, A.M., Gersbach, C.A., Christoforou, N., Leong, K.W., 2014. A CRISPR/Cas9-Based System for Reprogramming Cell Lineage Specification. Stem Cell Rep. 3, 940–947. https://doi.org/10.1016/j.stemcr.2014.09.013
Chazaud, C., Yamanaka, Y., Pawson, T., Rossant, J., 2006. Early Lineage Segregation between Epiblast and Primitive Endoderm in Mouse Blastocysts through the Grb2-MAPK Pathway. Dev. Cell 10, 615–624. https://doi.org/10.1016/j.devcel.2006.02.020
Chen, C., Pla-Palacín, I., Baptista, P.M., Shang, P., Oosterhoff, L.A., van Wolferen, M.E., Penning, L.C., Geijsen, N., Spee, B., 2018. Hepatocyte-like cells generated by direct reprogramming from murine somatic cells can repopulate decellularized livers. Biotechnol Bioeng 115, 2807–2816.
Chen, F., Desai, T.J., Qian, J., Niederreither, K., Lü, J., Cardoso, W.V., 2007. Inhibition of Tgf beta signaling by endogenous retinoic acid is essential for primary lung bud induction. Dev. Camb. Engl. 134, 2969–2979. https://doi.org/10.1242/dev.006221

Chen, Y.-W., Huang, S.X., de Carvalho, A.L.R.T., Ho, S.-H., Islam, M.N., Volpi, S., Notarangelo, L.D., Ciancanelli, M., Casanova, J.-L., Bhattacharya, J., Liang, A.F., Palermo, L.M., Porotto, M., Moscona, A., Snoeck, H.-W., 2017. A three-dimensional model of human lung development and disease from pluripotent stem cells. Nat. Cell Biol. 19, 542–549. https://doi.org/10.1038/ncb3510

Chen, Z., Li, S., Subramaniam, S., Shyy, J.Y.-J., Chien, S., 2017. Epigenetic Regulation: A New Frontier for Biomedical Engineers. Annu. Rev. Biomed. Eng. 19, 195–219. https://doi.org/10.1146/annurev-bioeng-071516-044720

Cheng, L., Hu, W., Qiu, B., Zhao, J., Yu, Y., Guan, W., Wang, M., Yang, W., Pei, G., 2014. Generation of neural progenitor cells by chemical cocktails and hypoxia. Cell Res. 24, 665–679. https://doi.org/10.1038/cr.2014.32

Chu, L.-F., Leng, N., Zhang, J., Hou, Z., Mamott, D., Vereide, D.T., Choi, J., Kendziorski, C., Stewart, R., Thomson, J.A., 2016. Single-cell RNA-seq reveals novel regulators of human embryonic stem cell differentiation to definitive endoderm. Genome Biol. 17. https://doi.org/10.1186/s13059-016-1033-x

Chuang, P.-T., Kawcak, T., McMahon, A.P., 2003. Feedback control of mammalian Hedgehog signaling by the Hedgehog-binding protein, Hip1, modulates Fgf signaling during branching morphogenesis of the lung. Genes Dev. 17, 342–347. https://doi.org/10.1101/gad.1026303

Chung, C., Kim, T., Kim, Miju, Kim, Minchul, Song, H., Kim, T.-S., Seo, E., Lee, S.-H., Kim, H., Kim, S.K., Yoo, G., Lee, D.-H., Hwang, D.-S., Kinashi, T., Kim, J.-M., Lim, D.-S., 2013. Hippo-Foxa2 signaling pathway plays a role in peripheral lung maturation and surfactant homeostasis. Proc. Natl. Acad. Sci. U. S. A. 110, 7732–7737. https://doi.org/10.1073/pnas.1220603110

Conlon, F.L., Lyons, K.M., Takaesu, N., Barth, K.S., Kispert, A., Herrmann, B., Robertson, E.J., 1994. A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. Dev. Camb. Engl. 120, 1919–1928.

Coraux, C., Delplanque, A., Hinnrasky, J., Peault, B., Puchelle, E., Gaillard, D., 1998. Distribution of Integrins During Human Fetal Lung Development. J. Histochem. Cytochem. 46, 803–810. https://doi.org/10.1177/002215549804600703

Costello, I., Nowotschin, S., Sun, X., Mould, A.W., Hadjantonakis, A.-K., Bikoff, E.K., Robertson, E.J., 2015. Lhx1 functions together with Otx2, Foxa2, and Ldb1 to govern anterior mesendoderm, node, and midline development. Genes Dev. 29, 2108–2122. https://doi.org/10.1101/gad.268979.115

Craft, J., Gordon, C.J., Huether, S.E., McCance, K.L., Brashers, V.L., Rote, N.S., 2019. Understanding pathophysiology.

Crane, A.M., Kramer, P., Bui, J.H., Chung, W.J., Li, X.S., Gonzalez-Garay, M.L., Hawkins, F., Liao, W., Mora, D., Choi, S., Wang, J., Sun, H.C., Paschon, D.E., Guschin, D.Y., Gregory, P.D., Kotton, D.N., Holmes, M.C., Sorscher, E.J., Davis, B.R., 2015. Targeted Correction and Restored Function of the CFTR Gene in Cystic Fibrosis Induced Pluripotent Stem Cells. Stem Cell Rep. 4, 569–577. https://doi.org/10.1016/j.stemcr.2015.02.005

D'Amour, K.A., Agulnick, A.D., Eliazer, S., Kelly, O.G., Kroon, E., Baetge, E.E., 2005. Efficient differentiation of human embryonic stem cells to definitive endoderm. Nat. Biotechnol. 23, 1534–1541. https://doi.org/10.1038/nbt1163

Danahay, H., Pessotti, A.D., Coote, J., Montgomery, B.E., Xia, D., Wilson, A., Yang, H., Wang, Z., Bevan, L., Thomas, C., Petit, S., London, A., LeMotte, P., Doelemeyer, A., Vélez-Reyes, G.L., Bernasconi, P., Fryer, C.J., Edwards, M., Capodieci, P., Chen, A., Hild, M., Jaffe, A.B., 2015. Notch2 Is Required for Inflammatory Cytokine-Driven Goblet Cell Metaplasia in the Lung. Cell Rep. 10, 239–252. https://doi.org/10.1016/j.celrep.2014.12.017

Daniely, Y., Liao, G., Dixon, D., Linnoila, R.I., Lori, A., Randell, S.H., Oren, M., Jetten, A.M., 2004. Critical role of p63 in the development of a normal esophageal and tracheobronchial epithelium. Am. J. Physiol.-Cell Physiol. 287, C171–C181. https://doi.org/10.1152/ajpcell.00226.2003

Danopoulos, S., Alonso, I., Thornton, M.E., Grubbs, B.H., Bellusci, S., Warburton, D., Al Alam, D., 2018. Human lung branching morphogenesis is orchestrated by the spatiotemporal distribution of ACTA2, SOX2, and SOX9. Am. J. Physiol.-Lung Cell. Mol. Physiol. 314, L144–L149. https://doi.org/10.1152/ajplung.00379.2017

De Moerlooze, L., Spencer-Dene, B., Revest, J.M., Hajihosseini, M., Rosewell, I., Dickson, C., 2000. An important role for the IIIb isoform of fibroblast growth factor receptor 2 (FGFR2) in mesenchymal-epithelial signalling during mouse organogenesis. Dev. Camb. Engl. 127, 483–492.

DeBoever, C., Li, H., Jakubosky, D., Benaglio, P., Reyna, J., Olson, K.M., Huang, H., Biggs, W., Sandoval, E., D'Antonio, M., Jepsen, K., Matsui, H., Arias, A., Ren, B., Nariai, N., Smith, E.N., D'Antonio-Chronowska, A., Farley, E.K., Frazer, K.A., 2017. Large-Scale Profiling Reveals the Influence of Genetic Variation on Gene Expression in Human Induced Pluripotent Stem Cells. Cell Stem Cell 20, 533-546.e7. https://doi.org/10.1016/j.stem.2017.03.009

Denson, L.A., McClure, M.H., Bogue, C.W., Karpen, S.J., Jacobs, H.C., 2000. HNF3beta and GATA-4 transactivate the liver-enriched homeobox gene, Hex. Gene 246, 311–320.

Desai, T.J., Brownfield, D.G., Krasnow, M.A., 2014. Alveolar progenitor and stem cells in lung development, renewal and cancer. Nature 507, 190–194. https://doi.org/10.1038/nature12930

Desai, T.J., Chen, F., Lü, J., Qian, J., Niederreither, K., Dollé, P., Chambon, P., Cardoso, W.V., 2006. Distinct roles for retinoic acid receptors alpha and beta in early lung morphogenesis. Dev. Biol. 291, 12–24. https://doi.org/10.1016/j.ydbio.2005.10.045

Desai, T.J., Malpel, S., Flentke, G.R., Smith, S.M., Cardoso, W.V., 2004. Retinoic acid selectively regulates Fgf10 expression and maintains cell identity in the prospective lung field of the developing foregut. Dev. Biol. 273, 402–415. https://doi.org/10.1016/j.ydbio.2004.04.039

Dessimoz, J., Opoka, R., Kordich, J.J., Grapin-Botton, A., Wells, J.M., 2006. FGF signaling is necessary for establishing gut tube domains alongthe anterior–posterior axis in vivo. Mech. Dev. 123, 42–55. https://doi.org/10.1016/j.mod.2005.10.001

Deutsch, G., Jung, J., Zheng, M., Lóra, J., Zaret, K.S., 2001. A bipotential precursor population for pancreas and liver within the embryonic endoderm. Dev. Camb. Engl. 128, 871–881.

Ding, S., Wu, X., Li, G., Han, M., Zhuang, Y., Xu, T., 2005. Efficient Transposition of the piggyBac (PB) Transposon in Mammalian Cells and Mice. Cell 122, 473–483. https://doi.org/10.1016/j.cell.2005.07.013

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., Gingeras, T.R., 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. https://doi.org/10.1093/bioinformatics/bts635

Domyan, E.T., Ferretti, E., Throckmorton, K., Mishina, Y., Nicolis, S.K., Sun, X., 2011. Signaling through BMP receptors promotes respiratory identity in the foregut via repression of Sox2. Development 138, 971–981. https://doi.org/10.1242/dev.053694

Driscoll, K.E., Carter, J.M., Iype, P.T., Kumari, H.L., Crosby, L.L., Aardema, M.J., Isfort, R.J., Cody, D., Chestnut, M.H., Burns, J.L., LeBoeuf, R.A., 1995. Establishment of immortalized alveolar type II epithelial cell lines from adult rats. Vitro Cell. Dev. Biol. - Anim. 31, 516–527. https://doi.org/10.1007/BF02634029

Dye, B.R., Hill, D.R., Ferguson, M.A., Tsai, Y.-H., Nagy, M.S., Dyal, R., Wells, J.M., Mayhew, C.N., Nattiv, R., Klein, O.D., White, E.S., Deutsch, G.H., Spence, J.R., 2015. In vitro generation of human pluripotent stem cell derived lung organoids. eLife 4. https://doi.org/10.7554/eLife.05098

Epaud, R., Aubey, F., Xu, J., Chaker, Z., Clemessy, M., Dautin, A., Ahamed, K., Bonora, M., Hoyeau, N., Fléjou, J.-F., Mailleux, A., Clement, A., Henrion-Caude, A., Holzenberger, M., 2012. Knockout of Insulin-Like Growth Factor-1 Receptor Impairs Distal Lung Morphogenesis. PLoS ONE 7, e48071. https://doi.org/10.1371/journal.pone.0048071

Evans, M.J., Kaufman, M.H., 1981. Establishment in culture of pluripotential cells from mouse embryos. Nature 292, 154–156. https://doi.org/10.1038/292154a0

Falix, F.A., Aronson, D.C., Lamers, W.H., Gaemers, I.C., 2012. Possible roles of DLK1 in the Notch pathway during development and disease. Biochim Biophys Acta 1822, 988–995.

Faucourt, M., Houliston, E., Besnardeau, L., Kimelman, D., Lepage, T., 2001. The pitx2 homeobox protein is required early for endoderm formation and nodal signaling. Dev. Biol. 229, 287–306. https://doi.org/10.1006/dbio.2000.9950

Feldman, B., Poueymirou, W., Papaioannou, V., DeChiara, T., Goldfarb, M., 1995. Requirement of FGF-4 for postimplantation mouse development. Science 267, 246–249. https://doi.org/10.1126/science.7809630
Fernandes-Silva, H., Correia-Pinto, J., Moura, R.S., 2017. Canonical Sonic Hedgehog Signaling in Early Lung Development. J. Dev. Biol. 5. https://doi.org/10.3390/jdb5010003
Firth, A.L., Dargitz, C.T., Qualls, S.J., Menon, T., Wright, R., Singer, O., Gage, F.H., Khanna, A., Verma, I.M., 2014. Generation of multiciliated cells in functional airway epithelia from human induced pluripotent stem cells. Proc. Natl. Acad. Sci. 111, E1723–E1730. https://doi.org/10.1073/pnas.1403470111
Firth, A.L., Menon, T., Parker, G.S., Qualls, S.J., Lewis, B.M., Ke, E., Dargitz, C.T., Wright, R., Khanna, A., Gage, F.H., Verma, I.M., 2015. Functional Gene Correction for Cystic Fibrosis in Lung Epithelial Cells Generated from Patient iPSCs. Cell Rep. 12, 1385–1390. https://doi.org/10.1016/j.celrep.2015.07.062
Flozak, A.S., Lam, A.P., Russell, S., Jain, M., Peled, O.N., Sheppard, K.A., Beri, R., Mutlu, G.M., Budinger, G.R.S., Gottardi, C.J., 2010. Beta-catenin/T-cell factor signaling is activated during lung injury and promotes the survival and migration of alveolar epithelial cells. J. Biol. Chem. 285, 3157–3167. https://doi.org/10.1074/jbc.M109.070326
Frank, D.B., Penkala, I.J., Zepp, J.A., Sivakumar, A., Linares-Saldana, R., Zacharias, W.J., Stolz, K.G., Pankin, J., Lu, M., Wang, Q., Babu, A., Li, L., Zhou, S., Morley, M.P., Jain, R., Morrisey, E.E., 2019. Early lineage specification defines alveolar epithelial ontogeny in the murine lung. Proc. Natl. Acad. Sci. 116, 4362–4371. https://doi.org/10.1073/pnas.1813952116
Gamer, L.W., Wright, C.V.E., 1995. Autonomous Endodermal Determination in Xenopus: Regulation of Expression of the Pancreatic Gene XlHbox 8. Dev. Biol. 171, 240–251. https://doi.org/10.1006/dbio.1995.1275
Ghaedi, M., Calle, E.A., Mendez, J.J., Gard, A.L., Balestrini, J., Booth, A., Bove, P.F., Gui, L., White, E.S., Niklason, L.E., 2013. Human iPS cell-derived alveolar epithelium repopulates lung extracellular matrix. J. Clin. Invest. 123, 4950–4962. https://doi.org/10.1172/JCI68793
Giangreco, A., Reynolds, S.D., Stripp, B.R., 2002. Terminal bronchioles harbor a unique airway stem cell population that localizes to the bronchoalveolar duct junction. Am. J. Pathol. 161, 173–182. https://doi.org/10.1016/S0002-9440(10)64169-7
Gilbert, S.F., 2000. Developmental biology, 6th ed. ed. Sinauer Associates, Sunderland, Mass.
Gilpin, S.E., Charest, J.M., Ren, X., Ott, H.C., 2016. Bioengineering Lungs for Transplantation. Thorac. Surg. Clin. 26, 163–171. https://doi.org/10.1016/j.thorsurg.2015.12.004
Gonzales, L.W., Guttentag, S.H., Wade, K.C., Postle, A.D., Ballard, P.L., 2002. Differentiation of human pulmonary type II cells in vitro by glucocorticoid plus cAMP. Am. J. Physiol.-Lung Cell. Mol. Physiol. 283, L940–L951. https://doi.org/10.1152/ajplung.00127.2002
Gordillo, M., Evans, T., Gouon-Evans, V., 2015. Orchestrating liver development. Development 142, 2094–2108.
Goss, A.M., Tian, Y., Tsukiyama, T., Cohen, E.D., Zhou, D., Lu, M.M., Yamaguchi, T.P., Morrisey, E.E., 2009. Wnt2/2b and β-Catenin Signaling Are Necessary and Sufficient to Specify Lung Progenitors in the Foregut. Dev. Cell 17, 290–298. https://doi.org/10.1016/j.devcel.2009.06.005
Gotoh, S., Ito, I., Nagasaki, T., Yamamoto, Y., Konishi, S., Korogi, Y., Matsumoto, H., Muro, S., Hirai, T., Funato, M., Mae, S.-I., Toyoda, T., Sato-Otsubo, A., Ogawa, S., Osafune, K., Mishima, M., 2014. Generation of alveolar epithelial spheroids via isolated progenitor cells from human pluripotent stem cells. Stem Cell Rep. 3, 394–403. https://doi.org/10.1016/j.stemcr.2014.07.005
Göttlicher, M., Minucci, S., Zhu, P., Krämer, O.H., Schimpf, A., Giavara, S., Sleeman, J.P., Lo Coco, F., Nervi, C., Pelicci, P.G., Heinzel, T., 2001. Valproic acid defines a novel class of HDAC inhibitors inducing differentiation of transformed cells. EMBO J. 20, 6969–6978. https://doi.org/10.1093/emboj/20.24.6969
Gray, P.C., Harrison, C.A., Vale, W., 2003. Cripto forms a complex with activin and type II activin receptors and can block activin signaling. Proc. Natl. Acad. Sci. 100, 5193–5198. https://doi.org/10.1073/pnas.0531290100

Green, M.D., Chen, A., Nostro, M.-C., d'Souza, S.L., Schaniel, C., Lemischka, I.R., Gouon-Evans, V., Keller, G., Snoeck, H.-W., 2011. Generation of anterior foregut endoderm from human embryonic and induced pluripotent stem cells. Nat. Biotechnol. 29, 267–272. https://doi.org/10.1038/nbt.1788

Guseh, J.S., Bores, S.A., Stanger, B.Z., Zhou, Q., Anderson, W.J., Melton, D.A., Rajagopal, J., 2009. Notch signaling promotes airway mucous metaplasia and inhibits alveolar development. Dev. Camb. Engl. 136, 1751–1759. https://doi.org/10.1242/dev.029249

Haghverdi, L., Büttner, M., Wolf, F.A., Buettner, F., Theis, F.J., 2016. Diffusion pseudotime robustly reconstructs lineage branching. Nat. Methods 13, 845–848. https://doi.org/10.1038/nmeth.3971

Han, L., Chaturvedi, P., Kishimoto, K., Koike, H., Nasr, T., Iwasawa, K., Giesbrecht, K., Witcher, P.C., Eicher, A., Haines, L., Lee, Y., Shannon, J.M., Morimoto, M., Wells, J.M., Takebe, T., Zorn, A.M., 2020. Single cell transcriptomics identifies a signaling network coordinating endoderm and mesoderm diversification during foregut organogenesis. Nat. Commun. 11, 4158. https://doi.org/10.1038/s41467-020-17968-x

Hannan, N.R.F., Segeritz, C.-P., Touboul, T., Vallier, L., 2013. Production of hepatocyte-like cells from human pluripotent stem cells. Nat Protoc 8, 430–437.

Harris-Johnson, K.S., Domyan, E.T., Vezina, C.M., Sun, X., 2009. beta-Catenin promotes respiratory progenitor identity in mouse foregut. Proc. Natl. Acad. Sci. U. S. A. 106, 16287–16292. https://doi.org/10.1073/pnas.0902274106

Hart, A.H., Hartley, L., Ibrahim, M., Robb, L., 2004. Identification, cloning and expression analysis of the pluripotency promoting Nanog genes in mouse and human. Dev. Dyn. Off. Publ. Am. Assoc. Anat. 230, 187–198. https://doi.org/10.1002/dvdy.20034

Hashimoto, K., Nakatsuji, N., 1989. Formation of the Primitive Streak and Mesoderm Cells in Mouse Embryos-Detailed Scanning Electron Microscopical Study. (primitive streak/cell migration/extracellular matrix/mouse gastrulation/scanning electron microscopy). Dev. Growth Differ. 31, 209–218. https://doi.org/10.1111/j.1440-169X.1989.00209.x

Hashimoto, S., Chen, H., Que, J., Brockway, B.L., Drake, J.A., Snyder, J.C., Randell, S.H., Stripp, B.R., 2012. β-Catenin-SOX2 signaling regulates the fate of developing airway epithelium. J. Cell Sci. 125, 932–942. https://doi.org/10.1242/jcs.092734

Hawkins, F., Kramer, P., Jacob, A., Driver, I., Thomas, D.C., McCauley, K.B., Skvir, N., Crane, A.M., Kurmann, A.A., Hollenberg, A.N., Nguyen, S., Wong, B.G., Khalil, A.S., Huang, S.X., Guttentag, S., Rock, J.R., Shannon, J.M., Davis, B.R., Kotton, D.N., 2017. Prospective isolation of NKX2-1-expressing human lung progenitors derived from pluripotent stem cells. J. Clin. Invest. 127, 2277–2294. https://doi.org/10.1172/JCI89950

Hebrok, M., Kim, S.K., Melton, D.A., 1998. Notochord repression of endodermal Sonic hedgehog permits pancreas development. Genes Dev. 12, 1705–1713. https://doi.org/10.1101/gad.12.11.1705

Henry, G.L., Brivanlou, I.H., Kessler, D.S., Hemmati-Brivanlou, A., Melton, D.A., 1996. TGF-beta signals and a pattern in Xenopus laevis endodermal development. Dev. Camb. Engl. 122, 1007–1015.

Herrmann, B.G., 1991. Expression pattern of the Brachyury gene in whole-mount TWis/TWis mutant embryos. Dev. Camb. Engl. 113, 913–917.

Hild, M., Jaffe, A.B., 2016. Production of 3-D Airway Organoids From Primary Human Airway Basal Cells and Their Use in High-Throughput Screening. Curr. Protoc. Stem Cell Biol. 37. https://doi.org/10.1002/cpsc.1

Hogan, B.L.M., Barkauskas, C.E., Chapman, H.A., Epstein, J.A., Jain, R., Hsia, C.C.W., Niklason, L., Calle, E., Le, A., Randell, S.H., Rock, J., Snitow, M., Krummel, M., Stripp, B.R., Vu, T., White, E.S., Whitsett, J.A., Morrisey, E.E., 2014. Repair and regeneration of the respiratory system: complexity, plasticity, and mechanisms of lung stem cell function. Cell Stem Cell 15, 123–138. https://doi.org/10.1016/j.stem.2014.07.012

Hong, K.U., Reynolds, S.D., Giangreco, A., Hurley, C.M., Stripp, B.R., 2001. Clara cell secretory protein-expressing cells of the airway neuroepithelial body microenvironment include a label-retaining subset and are critical for epithelial renewal after progenitor cell depletion. Am. J. Respir. Cell Mol. Biol. 24, 671–681. https://doi.org/10.1165/ajrcmb.24.6.4498

Hong, K.U., Reynolds, S.D., Watkins, S., Fuchs, E., Stripp, B.R., 2004. Basal Cells Are a Multipotent Progenitor Capable of Renewing the Bronchial Epithelium. Am. J. Pathol. 164, 577–588. https://doi.org/10.1016/S0002-9440(10)63147-1

Hovenberg, H.W., Davies, J.R., Carlstedt, I., 1996. Different mucins are produced by the surface epithelium and the submucosa in human trachea: identification of MUC5AC as a major mucin from the goblet cells. Biochem. J. 318, 319–324. https://doi.org/10.1042/bj3180319

Huang, S.X.L., Islam, M.N., O'Neill, J., Hu, Z., Yang, Y.-G., Chen, Y.-W., Mumau, M., Green, M.D., Vunjak-Novakovic, G., Bhattacharya, J., Snoeck, H.-W., 2014. Efficient generation of lung and airway epithelial cells from human pluripotent stem cells. Nat. Biotechnol. 32, 84–91. https://doi.org/10.1038/nbt.2754

Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A.E., Melton, D.A., 2008. Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. Nat. Biotechnol. 26, 795–797. https://doi.org/10.1038/nbt1418

Hunter, M.P., Wilson, C.M., Jiang, X., Cong, R., Vasavada, H., Kaestner, K.H., Bogue, C.W., 2007. The homeobox gene Hhex is essential for proper hepatoblast differentiation and bile duct morphogenesis. Dev Biol 308, 355–367.

Hurley, K., Ding, J., Villacorta-Martin, C., Herriges, M.J., Jacob, A., Vedaie, M., Alysandratos, K.D., Sun, Y.L., Lin, C., Werder, R.B., Huang, J., Wilson, A.A., Mithal, A., Mostoslavsky, G., Oglesby, I., Caballero, I.S., Guttentag, S.H., Ahangari, F., Kaminski, N., Rodriguez-Fraticelli, A., Camargo, F., Bar-Joseph, Z., Kotton, D.N., 2020. Reconstructed Single-Cell Fate Trajectories Define Lineage Plasticity Windows during Differentiation of Human PSC-Derived Distal Lung Progenitors. Cell Stem Cell 26, 593-608.e8. https://doi.org/10.1016/j.stem.2019.12.009

Ikonomou, L., Herriges, M.J., Lewandowski, S.L., Marsland, R., Villacorta-Martin, C., Caballero, I.S., Frank, D.B., Sanghrajka, R.M., Dame, K., Kańduła, M.M., Hicks-Berthet, J., Lawton, M.L., Christodoulou, C., Fabian, A.J., Kolaczyk, E., Varelas, X., Morrisey, E.E., Shannon, J.M., Mehta, P., Kotton, D.N., 2020. The in vivo genetic program of murine primordial lung epithelial progenitors. Nat. Commun. 11, 635. https://doi.org/10.1038/s41467-020-14348-3

Kang, J., Nathan, E., Xu, S.-M., Tzahor, E., Black, B.L., 2009. Isl1 is a direct transcriptional target of Forkhead transcription factors in second-heart-field-derived mesoderm. Dev. Biol. 334, 513–522. https://doi.org/10.1016/j.ydbio.2009.06.041

Katsirntaki, K., Mauritz, C., Olmer, R., Schmeckebier, S., Sgodda, M., Puppe, V., Eggenschwiler, R., Duerr, J., Schubert, S.C., Schmiedl, A., Ochs, M., Cantz, T., Salwig, I., Szibor, M., Braun, T., Rathert, C., Martens, A., Mall, M.A., Martin, U., 2015. Bronchoalveolar Sublineage Specification of Pluripotent Stem Cells: Effect of Dexamethasone Plus cAMP-Elevating Agents and Keratinocyte Growth Factor. Tissue Eng. Part A 21, 669–682. https://doi.org/10.1089/ten.tea.2014.0097

Kho, A.T., Bhattacharya, S., Tantisira, K.G., Carey, V.J., Gaedigk, R., Leeder, J.S., Kohane, I.S., Weiss, S.T., Mariani, T.J., 2010. Transcriptomic Analysis of Human Lung Development. Am. J. Respir. Crit. Care Med. 181, 54–63. https://doi.org/10.1164/rccm.200907-1063OC

Khoor, A., Stahlman, M.T., Gray, M.E., Whitsett, J.A., 1994. Temporal-spatial distribution of SP-B and SP-C proteins and mRNAs in developing respiratory epithelium of human lung. J. Histochem. Cytochem. 42, 1187–1199. https://doi.org/10.1177/42.9.8064126

Kilpinen, H., Goncalves, A., Leha, A., Afzal, V., Alasoo, K., Ashford, S., Bala, S., Bensaddek, D., Casale, F.P., Culley, O.J., Danecek, P., Faulconbridge, A., Harrison, P.W., Kathuria, A., McCarthy, D., McCarthy, S.A., Meleckyte, R., Memari, Y., Moens, N., Soares, F., Mann, A., Streeter, I., Agu, C.A., Alderton, A., Nelson, R., Harper, S., Patel, M., White, A., Patel, S.R., Clarke, L., Halai, R., Kirton, C.M., Kolb-Kokocinski, A., Beales, P., Birney, E., Danovi, D., Lamond, A.I., Ouwehand, W.H., Vallier, L., Watt, F.M., Durbin, R., Stegle, O., Gaffney, D.J., 2017. Common genetic variation drives molecular heterogeneity in human iPSCs. Nature 546, 370–375. https://doi.org/10.1038/nature22403

Kim, C.F.B., Jackson, E.L., Woolfenden, A.E., Lawrence, S., Babar, I., Vogel, S., Crowley, D., Bronson, R.T., Jacks, T., 2005. Identification of Bronchioalveolar Stem Cells in Normal Lung and Lung Cancer. Cell 121, 823–835. https://doi.org/10.1016/j.cell.2005.03.032

Kim, E., Jiang, M., Huang, H., Zhang, Y., Robert, J., Gilmore, N., Gan, L., Que, J., 2019. Isl1 Regulation of Nkx2.1 in the Early Foregut Epithelium Is Required for Trachea-Esophageal Separation and Lung Lobation. SSRN Electron. J. https://doi.org/10.2139/ssrn.3387653

Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M.J., Ji, H., Ehrlich, L.I.R., Yabuuchi, A., Takeuchi, A., Cunniff, K.C., Hongguang, H., McKinney-Freeman, S., Naveiras, O., Yoon, T.J., Irizarry, R.A., Jung, N., Seita, J., Hanna, J., Murakami, P., Jaenisch, R., Weissleder, R., Orkin, S.H., Weissman, I.L., Feinberg, A.P., Daley, G.Q., 2010. Epigenetic memory in induced pluripotent stem cells. Nature 467, 285–290. https://doi.org/10.1038/nature09342

Kim, S.K., Melton, D.A., 1998. Pancreas development is promoted by cyclopamine, a hedgehog signaling inhibitor. Proc. Natl. Acad. Sci. U. S. A. 95, 13036–13041. https://doi.org/10.1073/pnas.95.22.13036

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

Kinder, S.J., Tsang, T.E., Wakamiya, M., Sasaki, H., Behringer, R.R., Nagy, A., Tam, P.P., 2001. The organizer of the mouse gastrula is composed of a dynamic population of progenitor cells for the axial mesoderm. Dev. Camb. Engl. 128, 3623–3634.

Kogiso, T., Nagahara, H., Otsuka, M., Shiratori, K., Dowdy, S.F., 2013. Transdifferentiation of human fibroblasts into hepatocyte-like cells by defined transcriptional factors. Hepatol. Int. 7, 937–944. https://doi.org/10.1007/s12072-013-9432-5

Konishi, S., Gotoh, S., Tateishi, K., Yamamoto, Y., Korogi, Y., Nagasaki, T., Matsumoto, H., Muro, S., Hirai, T., Ito, I., Tsukita, S., Mishima, M., 2016. Directed Induction of Functional Multi-ciliated Cells in Proximal Airway Epithelial Spheroids from Human Pluripotent Stem Cells. Stem Cell Rep. 6, 18–25. https://doi.org/10.1016/j.stemcr.2015.11.010

Koressaar, T., Remm, M., 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics 23, 1289–1291. https://doi.org/10.1093/bioinformatics/btm091

Korogi, Y., Gotoh, S., Ikeo, S., Yamamoto, Y., Sone, N., Tamai, K., Konishi, S., Nagasaki, T., Matsumoto, H., Ito, I., Chen-Yoshikawa, T.F., Date, H., Hagiwara, M., Asaka, I., Hotta, A., Mishima, M., Hirai, T., 2019. In Vitro Disease Modeling of Hermansky-Pudlak Syndrome Type 2 Using Human Induced Pluripotent Stem Cell-Derived Alveolar Organoids. Stem Cell Rep. 12, 431–440. https://doi.org/10.1016/j.stemcr.2019.01.014

Kraus, M.R.C., Grapin-Botton, A., 2012. Patterning and shaping the endoderm in vivo and in culture. Curr. Opin. Genet. Dev. 22, 347–353. https://doi.org/10.1016/j.gde.2012.05.002

Kreda, S.M., Gynn, M.C., Fenstermacher, D.A., Boucher, R.C., Gabriel, S.E., 2001. Expression and localization of epithelial aquaporins in the adult human lung. Am. J. Respir. Cell Mol. Biol. 24, 224–234. https://doi.org/10.1165/ajrcmb.24.3.4367

Kubo, A., 2004. Development of definitive endoderm from embryonic stem cells in culture. Development 131, 1651–1662. https://doi.org/10.1242/dev.01044

Kumar, N., Tsai, Y.-H., Chen, L., Zhou, A., Banerjee, K.K., Saxena, M., Huang, S., Toke, N.H., Xing, J., Shivdasani, R.A., Spence, J.R., Verzi, M.P., 2019. The lineage-specific transcription factor CDX2 navigates dynamic chromatin to control distinct stages of intestine development. Development 146.

Lancaster, M.A., Knoblich, J.A., 2014. Organogenesis in a dish: Modeling development and disease using organoid technologies. Science 345, 1247125–1247125. https://doi.org/10.1126/science.1247125

Lange, A.W., Sridharan, A., Xu, Y., Stripp, B.R., Perl, A.-K., Whitsett, J.A., 2015. Hippo/Yap signaling controls epithelial progenitor cell proliferation and differentiation in the embryonic and adult lung. J. Mol. Cell Biol. 7, 35–47. https://doi.org/10.1093/jmcb/mju046

Laresgoiti, U., Nikolić, M.Z., Rao, C., Brady, J.L., Richardson, R.V., Batchen, E.J., Chapman, K.E., Rawlins, E.L., 2016. Lung epithelial tip progenitors integrate glucocorticoid- and STAT3-mediated signals to control progeny fate. Dev. Camb. Engl. 143, 3686–3699. https://doi.org/10.1242/dev.134023

Lazzaro, D., Price, M., de Felice, M., Di Lauro, R., 1991. The transcription factor TTF-1 is expressed at the onset of thyroid and lung morphogenesis and in restricted regions of the foetal brain. Dev. Camb. Engl. 113, 1093–1104.

Lee, C.S., Friedman, J.R., Fulmer, J.T., Kaestner, K.H., 2005. The initiation of liver development is dependent on Foxa transcription factors. Nature 435, 944–947.
Li, C., Xiao, J., Hormi, K., Borok, Z., Minoo, P., 2002. Wnt5a Participates in Distal Lung Morphogenesis. Dev. Biol. 248, 68–81. https://doi.org/10.1006/dbio.2002.0729
Li, Z., White, P., Tuteja, G., Rubins, N., Sackett, S., Kaestner, K.H., 2009. Foxa1 and Foxa2 regulate bile duct development in mice. J Clin Invest 119, 1537–1545.
Lickert, H., Kutsch, S., Kanzler, B., Tamai, Y., Taketo, M.M., Kemler, R., 2002. Formation of Multiple Hearts in Mice following Deletion of β-catenin in the Embryonic Endoderm. Dev. Cell 3, 171–181. https://doi.org/10.1016/S1534-5807(02)00206-X
Lin, C.R., Kioussi, C., O'Connell, S., Briata, P., Szeto, D., Liu, F., Izpisúa-Belmonte, J.C., Rosenfeld, M.G., 1999. Pitx2 regulates lung asymmetry, cardiac positioning and pituitary and tooth morphogenesis. Nature 401, 279–282.
Lin, L., Bu, L., Cai, C.-L., Zhang, X., Evans, S., 2006. Isl1 is upstream of sonic hedgehog in a pathway required for cardiac morphogenesis. Dev Biol 295, 756–763.
Ling, T.-Y., Liu, Y.-L., Huang, Y.-K., Gu, S.-Y., Chen, H.-K., Ho, C.-C., Tsao, P.-N., Tung, Y.-C., Chen, H.-W., Cheng, C.-H., Lin, K.-H., Lin, F.-H., 2014. Differentiation of lung stem/progenitor cells into alveolar pneumocytes and induction of angiogenesis within a 3D gelatin--microbubble scaffold. Biomaterials 35, 5660–5669. https://doi.org/10.1016/j.biomaterials.2014.03.074
Litingtung, Y., Lei, L., Westphal, H., Chiang, C., 1998. Sonic hedgehog is essential to foregut development. Nat. Genet. 20, 58–61. https://doi.org/10.1038/1717
Liu, P., Wakamiya, M., Shea, M.J., Albrecht, U., Behringer, R.R., Bradley, A., 1999. Requirement for Wnt3 in vertebrate axis formation. Nat. Genet. 22, 361–365. https://doi.org/10.1038/11932
Loh, K.M., Ang, L.T., Zhang, J., Kumar, V., Ang, J., Auyeong, J.Q., Lee, K.L., Choo, S.H., Lim, C.Y.Y., Nichane, M., Tan, J., Noghabi, M.S., Azzola, L., Ng, E.S., Durruthy-Durruthy, J., Sebastiano, V., Poellinger, L., Elefanty, A.G., Stanley, E.G., Chen, Q., Prabhakar, S., Weissman, I.L., Lim, B., 2014. Efficient endoderm induction from human pluripotent stem cells by logically directing signals controlling lineage bifurcations. Cell Stem Cell 14, 237–252. https://doi.org/10.1016/j.stem.2013.12.007
Longmire, T.A., Ikonomou, L., Hawkins, F., Christodoulou, C., Cao, Y., Jean, J.C., Kwok, L.W., Mou, H., Rajagopal, J., Shen, S.S., Dowton, A.A., Serra, M., Weiss, D.J., Green, M.D., Snoeck, H.-W., Ramirez, M.I., Kotton, D.N., 2012. Efficient derivation of purified lung and thyroid progenitors from embryonic stem cells. Cell Stem Cell 10, 398–411.
Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. https://doi.org/10.1186/s13059-014-0550-8
Lüdtke, T.H.-W., Christoffels, V.M., Petry, M., Kispert, A., 2009. Tbx3 promotes liver bud expansion during mouse development by suppression of cholangiocyte differentiation. Hepatology 49, 969–978.
Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., Trombetta, J.J., Weitz, D.A., Sanes, J.R., Shalek, A.K., Regev, A., McCarroll, S.A., 2015. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell 161, 1202–1214. https://doi.org/10.1016/j.cell.2015.05.002
Mahoney, J.E., Mori, M., Szymaniak, A.D., Varelas, X., Cardoso, W.V., 2014. The Hippo Pathway Effector Yap Controls Patterning and Differentiation of Airway Epithelial Progenitors. Dev. Cell 30, 137–150. https://doi.org/10.1016/j.devcel.2014.06.003
Mailleux, A.A., Tefft, D., Ndiaye, D., Itoh, N., Thiery, J.P., Warburton, D., Bellusci, S., 2001. Evidence that SPROUTY2 functions as an inhibitor of mouse embryonic lung growth and morphogenesis. Mech. Dev. 102, 81–94. https://doi.org/10.1016/S0925-4773(01)00286-6
Malpel, S., Mendelsohn, C., Cardoso, W.V., 2000. Regulation of retinoic acid signaling during lung morphogenesis. Dev. Camb. Engl. 127, 3057–3067.
Martin, G.R., 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc. Natl. Acad. Sci. 78, 7634–7638. https://doi.org/10.1073/pnas.78.12.7634

Matsui, T.K., Matsubayashi, M., Sakaguchi, Y.M., Hayashi, R.K., Zheng, C., Sugie, K., Hasegawa, M., Nakagawa, T., Mori, E., 2018. Six-month cultured cerebral organoids from human ES cells contain matured neural cells. Neurosci. Lett. 670, 75–82. https://doi.org/10.1016/j.neulet.2018.01.040

McCauley, K.B., Alysandratos, K.-D., Jacob, A., Hawkins, F., Caballero, I.S., Vedaie, M., Yang, W., Slovik, K.J., Morley, M., Carraro, G., Kook, S., Guttentag, S.H., Stripp, B.R., Morrisey, E.E., Kotton, D.N., 2018. Single-Cell Transcriptomic Profiling of Pluripotent Stem Cell-Derived SCGB3A2+ Airway Epithelium. Stem Cell Rep. 10, 1579–1595. https://doi.org/10.1016/j.stemcr.2018.03.013

McCauley, K.B., Hawkins, F., Serra, M., Thomas, D.C., Jacob, A., Kotton, D.N., 2017. Efficient Derivation of Functional Human Airway Epithelium from Pluripotent Stem Cells via Temporal Regulation of Wnt Signaling. Cell Stem Cell 20, 844-857.e6. https://doi.org/10.1016/j.stem.2017.03.001

Miller, A.J., Hill, D.R., Nagy, M.S., Aoki, Y., Dye, B.R., Chin, A.M., Huang, S., Zhu, F., White, E.S., Lama, V., Spence, J.R., 2018. In Vitro Induction and In Vivo Engraftment of Lung Bud Tip Progenitor Cells Derived from Human Pluripotent Stem Cells. Stem Cell Rep. 10, 101–119. https://doi.org/10.1016/j.stemcr.2017.11.012

Minoo, P., Su, G., Drum, H., Bringas, P., Kimura, S., 1999. Defects in Tracheoesophageal and Lung Morphogenesis inNkx2.1(−/−) Mouse Embryos. Dev. Biol. 209, 60–71. https://doi.org/10.1006/dbio.1999.9234

Mollard, R., Ghyselinck, N.B., Wendling, O., Chambon, P., Mark, M., 2000. Stage-dependent responses of the developing lung to retinoic acid signaling. Int. J. Dev. Biol. 44, 457–462.

Morimoto, M., Liu, Z., Cheng, H.-T., Winters, N., Bader, D., Kopan, R., 2010. Canonical Notch signaling in the developing lung is required for determination of arterial smooth muscle cells and selection of Clara versus ciliated cell fate. J. Cell Sci. 123, 213–224. https://doi.org/10.1242/jcs.058669

Morimoto, M., Nishinakamura, R., Saga, Y., Kopan, R., 2012. Different assemblies of Notch receptors coordinate the distribution of the major bronchial Clara, ciliated and neuroendocrine cells. Development 139, 4365–4373. https://doi.org/10.1242/dev.083840

Morita, R., Suzuki, M., Kasahara, H., Shimizu, N., Shichita, T., Sekiya, T., Kimura, A., Sasaki, K., Yasukawa, H., Yoshimura, A., 2015. ETS transcription factor ETV2 directly converts human fibroblasts into functional endothelial cells. Proc. Natl. Acad. Sci. 112, 160–165. https://doi.org/10.1073/pnas.1413234112

Morrisey, E.E., Hogan, B.L.M., 2010. Preparing for the First Breath: Genetic and Cellular Mechanisms in Lung Development. Dev. Cell 18, 8–23. https://doi.org/10.1016/j.devcel.2009.12.010

Motoyama, J., Liu, J., Mo, R., Ding, Q., Post, M., Hui, C., 1998. Essential function of Gli2 and Gli3 in the formation of lung, trachea and oesophagus. Nat. Genet. 20, 54–57. https://doi.org/10.1038/1711

Mou, H., Zhao, R., Sherwood, R., Ahfeldt, T., Lapey, A., Wain, J., Sicilian, L., Izvolsky, K., Lau, F.H., Musunuru, K., Cowan, C., Rajagopal, J., 2012. Generation of Multipotent Lung and Airway Progenitors from Mouse ESCs and Patient-Specific Cystic Fibrosis iPSCs. Cell Stem Cell 10, 385–397. https://doi.org/10.1016/j.stem.2012.01.018

Mucenski, M.L., Wert, S.E., Nation, J.M., Loudy, D.E., Huelsken, J., Birchmeier, W., Morrisey, E.E., Whitsett, J.A., 2003. β-Catenin Is Required for Specification of Proximal/Distal Cell Fate during Lung Morphogenesis. J. Biol. Chem. 278, 40231–40238. https://doi.org/10.1074/jbc.M305892200

Narayanan, M., Owers-Bradley, J., Beardsmore, C.S., Mada, M., Ball, I., Garipov, R., Panesar, K.S., Kuehni, C.E., Spycher, B.D., Williams, S.E., Silverman, M., 2012. Alveolarization Continues during Childhood and Adolescence: New Evidence from Helium-3 Magnetic Resonance. Am. J. Respir. Crit. Care Med. 185, 186–191. https://doi.org/10.1164/rccm.201107-1348OC

Nazareth, E.J.P., Ostblom, J.E.E., Lücker, P.B., Shukla, S., Alvarez, M.M., Oh, S.K.W., Yin, T., Zandstra, P.W., 2013. High-throughput fingerprinting of human pluripotent stem cell fate responses and lineage bias. Nat. Methods 10, 1225–1231. https://doi.org/10.1038/nmeth.2684

Nikolić, M.Z., Caritg, O., Jeng, Q., Johnson, J.-A., Sun, D., Howell, K.J., Brady, J.L., Laresgoiti, U., Allen, G., Butler, R., Zilbauer, M., Giangreco, A., Rawlins, E.L., 2017. Human embryonic lung epithelial tips are multipotent progenitors that can be expanded in vitro as long-term self-renewing organoids. eLife 6. https://doi.org/10.7554/eLife.26575

Nishioka, N., Inoue, K., Adachi, K., Kiyonari, H., Ota, M., Ralston, A., Yabuta, N., Hirahara, S., Stephenson, R.O., Ogonuki, N., Makita, R., Kurihara, H., Morin-Kensicki, E.M., Nojima, H., Rossant, J., Nakao, K., Niwa, H., Sasaki, H., 2009. The Hippo Signaling Pathway Components Lats and Yap Pattern Tead4 Activity to Distinguish Mouse Trophectoderm from Inner Cell Mass. Dev. Cell 16, 398–410. https://doi.org/10.1016/j.devcel.2009.02.003

Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., Rossant, J., 2005. Interaction between Oct3/4 and Cdx2 Determines Trophectoderm Differentiation. Cell 123, 917–929. https://doi.org/10.1016/j.cell.2005.08.040

Ohuchi, H., Hori, Y., Yamasaki, M., Harada, H., Sekine, K., Kato, S., Itoh, N., 2000. FGF10 Acts as a Major Ligand for FGF Receptor 2 IIIb in Mouse Multi-Organ Development. Biochem. Biophys. Res. Commun. 277, 643–649. https://doi.org/10.1006/bbrc.2000.3721

Olmer, R., Dahlmann, J., Merkert, S., Baus, S., Göhring, G., Martin, U., 2019. Generation of a NKX2.1 knock-in reporter cell line from human induced pluripotent stem cells (MHHi006-A-2). Stem Cell Res. 39, 101492. https://doi.org/10.1016/j.scr.2019.101492

Onishi, K., Zandstra, P.W., 2015. LIF signaling in stem cells and development. Development 142, 2230–2236.

Ori, C., Ansari, M., Angelidis, I., Theis, F.J., Schiller, H.B., Drukker, M., 2021. Single cell trajectory analysis of human pluripotent stem cells differentiating towards lung and hepatocyte progenitors (preprint). Developmental Biology. https://doi.org/10.1101/2021.02.23.432413

Osafune, K., Caron, L., Borowiak, M., Martinez, R.J., Fitz-Gerald, C.S., Sato, Y., Cowan, C.A., Chien, K.R., Melton, D.A., 2008. Marked differences in differentiation propensity among human embryonic stem cell lines. Nat. Biotechnol. 26, 313–315. https://doi.org/10.1038/nbt1383

Oshika, E., Liu, S., Ung, L.P., Singh, G., Shinozuka, H., Michalopoulos, G.K., Katyal, S.L., 1998. Glucocorticoid-induced effects on pattern formation and epithelial cell differentiation in early embryonic rat lungs. Pediatr. Res. 43, 305–314. https://doi.org/10.1203/00006450-199803000-00001

Ott, H.C., Clippinger, B., Conrad, C., Schuetz, C., Pomerantseva, I., Ikonomou, L., Kotton, D., Vacanti, J.P., 2010. Regeneration and orthotopic transplantation of a bioartificial lung. Nat. Med. 16, 927–933. https://doi.org/10.1038/nm.2193

Parker, M.W., Rossi, D., Peterson, M., Smith, K., Sikström, K., White, E.S., Connett, J.E., Henke, C.A., Larsson, O., Bitterman, P.B., 2014. Fibrotic extracellular matrix activates a profibrotic positive feedback loop. J. Clin. Invest. 124, 1622–1635. https://doi.org/10.1172/JCI71386

Peng, T., Frank, D.B., Kadzik, R.S., Morley, M.P., Rathi, K.S., Wang, T., Zhou, S., Cheng, L., Lu, M.M., Morrisey, E.E., 2015. Hedgehog actively maintains adult lung quiescence and regulates repair and regeneration. Nature 526, 578–582. https://doi.org/10.1038/nature14984

Petersen, T.H., Calle, E.A., Zhao, L., Lee, E.J., Gui, L., Raredon, M.B., Gavrilov, K., Yi, T., Zhuang, Z.W., Breuer, C., Herzog, E., Niklason, L.E., 2010. Tissue-Engineered Lungs for in Vivo Implantation. Science 329, 538–541. https://doi.org/10.1126/science.1189345

Phelps, D.S., Floros, J., 1991. Localization of pulmonary surfactant proteins using immunohistochemistry and tissue in situ hybridization. Exp. Lung Res. 17, 985–995. https://doi.org/10.3109/01902149109064330

Plosa, E.J., Young, L.R., Gulleman, P.M., Polosukhin, V.V., Zaynagetdinov, R., Benjamin, J.T., Im, A.M., van der Meer, R., Gleaves, L.A., Bulus, N., Han, W., Prince, L.S., Blackwell, T.S., Zent, R., 2014. Epithelial 1 integrin is required for lung branching morphogenesis and alveolarization. Development 141, 4751–4762. https://doi.org/10.1242/dev.117200

Porotto, M., Ferren, M., Chen, Y.-W., Siu, Y., Makhsous, N., Rima, B., Briese, T., Greninger, A.L., Snoeck, H.-W., Moscona, A., 2019. Authentic Modeling of Human Respiratory Virus Infection in Human Pluripotent Stem Cell-Derived Lung Organoids. mBio 10. https://doi.org/10.1128/mBio.00723-19

Post, L.C., Ternet, M., Hogan, B.L.M., 2000. Notch/Delta expression in the developing mouse lung. Mech. Dev. 98, 95–98. https://doi.org/10.1016/S0925-4773(00)00432-9

Potter, S.S., 2018. Single-cell RNA sequencing for the study of development, physiology and disease. Nat. Rev. Nephrol. 14, 479–492. https://doi.org/10.1038/s41581-018-0021-7

Que, J., Choi, M., Ziel, J.W., Klingensmith, J., Hogan, B.L.M., 2006. Morphogenesis of the trachea and esophagus: current players and new roles for noggin and Bmps. Differentiation 74, 422–437. https://doi.org/10.1111/j.1432-0436.2006.00096.x

Que, J., Okubo, T., Goldenring, J.R., Nam, K.-T., Kurotani, R., Morrisey, E.E., Taranova, O., Pevny, L.H., Hogan, B.L.M., 2007. Multiple dose-dependent roles for Sox2 in the patterning and differentiation of anterior foregut endoderm. Dev. Camb. Engl. 134, 2521–2531. https://doi.org/10.1242/dev.003855

Rankin, S.A., Han, L., McCracken, K.W., Kenny, A.P., Anglin, C.T., Grigg, E.A., Crawford, C.M., Wells, J.M., Shannon, J.M., Zorn, A.M., 2016. A Retinoic Acid-Hedgehog Cascade Coordinates Mesoderm-Inducing Signals and Endoderm Competence during Lung Specification. Cell Rep. 16, 66–78. https://doi.org/10.1016/j.celrep.2016.05.060

Rawlins, E.L., Clark, C.P., Xue, Y., Hogan, B.L.M., 2009. The Id2+ distal tip lung epithelium contains individual multipotent embryonic progenitor cells. Dev. Camb. Engl. 136, 3741–3745. https://doi.org/10.1242/dev.037317

Regev, A., Teichmann, S., Rozenblatt-Rosen, O., Stubbington, M., Ardlie, K., Amit, I., Arlotta, P., Bader, G., Benoist, C., Biton, M., Bodenmiller, B., Bruneau, B., Campbell, P., Carmichael, M., Carninci, P., Castelo-Soccio, L., Clatworthy, M., Clevers, H., Conrad, C., Eils, R., Freeman, J., Fugger, L., Goettgens, B., Graham, D., Greka, A., Hacohen, N., Haniffa, M., Helbig, I., Heuckeroth, R., Kathiresan, S., Kim, S., Klein, A., Knoppers, B., Kriegstein, A., Lander, E., Lee, J., Lein, E., Linnarsson, S., Macosko, E., MacParland, S., Majovski, R., Majumder, P., Marioni, J., McGilvray, I., Merad, M., Mhlanga, M., Naik, S., Nawijn, M., Nolan, G., Paten, B., Pe'er, D., Philippakis, A., Ponting, C., Quake, S., Rajagopal, J., Rajewsky, N., Reik, W., Rood, J., Saeb-Parsy, K., Schiller, H., Scott, S., Shalek, A., Shapiro, E., Shin, J., Skeldon, K., Stratton, M., Streicher, J., Stunnenberg, H., Tan, K., Taylor, D., Thorogood, A., Vallier, L., van Oudenaarden, A., Watt, F., Weicher, W., Weissman, J., Wells, A., Wold, B., Xavier, R., Zhuang, X., Committee, H.C.A.O., 2018. The Human Cell Atlas White Paper. ArXiv181005192 Q-Bio.

Reynolds, S.D., Giangreco, A., Power, J.H.T., Stripp, B.R., 2000. Neuroepithelial Bodies of Pulmonary Airways Serve as a Reservoir of Progenitor Cells Capable of Epithelial Regeneration. Am. J. Pathol. 156, 269–278. https://doi.org/10.1016/S0002-9440(10)64727-X

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., Smyth, G.K., 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43, e47.

Robb, L., Tam, P.P.L., 2004. Gastrula organiser and embryonic patterning in the mouse. Semin. Cell Dev. Biol. 15, 543–554. https://doi.org/10.1016/j.semcdb.2004.04.005

Rock, J.R., Gao, X., Xue, Y., Randell, S.H., Kong, Y.-Y., Hogan, B.L.M., 2011. Notch-Dependent Differentiation of Adult Airway Basal Stem Cells. Cell Stem Cell 8, 639–648. https://doi.org/10.1016/j.stem.2011.04.003

Rock, J.R., Onaitis, M.W., Rawlins, E.L., Lu, Y., Clark, C.P., Xue, Y., Randell, S.H., Hogan, B.L.M., 2009. Basal cells as stem cells of the mouse trachea and human airway epithelium. Proc. Natl. Acad. Sci. 106, 12771–12775. https://doi.org/10.1073/pnas.0906850106

Roost, M.S., van Iperen, L., Ariyurek, Y., Buermans, H.P., Arindrarto, W., Devalla, H.D., Passier, R., Mummery, C.L., Carlotti, F., de Koning, E.J.P., van Zwet, E.W., Goeman, J.J., Chuva de Sousa Lopes, S.M., 2015. KeyGenes, a Tool to Probe Tissue Differentiation Using a Human Fetal Transcriptional Atlas. Stem Cell Rep. 4, 1112–1124. https://doi.org/10.1016/j.stemcr.2015.05.002

Rubio, A., Luoni, M., Giannelli, S.G., Radice, I., Iannielli, A., Cancellieri, C., Di Berardino, C., Regalia, G., Lazzari, G., Menegon, A., Taverna, S., Broccoli, V., 2016. Rapid and efficient CRISPR/Cas9 gene inactivation in human neurons during human pluripotent stem cell differentiation and direct reprogramming. Sci. Rep. 6, 37540. https://doi.org/10.1038/srep37540

Sampaziotis, F., de Brito, M.C., Geti, I., Bertero, A., Hannan, N.R., Vallier, L., 2017. Directed differentiation of human induced pluripotent stem cells into functional cholangiocyte-like cells. Nat. Protoc. 12, 814–827. https://doi.org/10.1038/nprot.2017.011

Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., Regev, A., 2015. Spatial reconstruction of single-cell gene expression data. Nat Biotechnol 33, 495–502.
Schittny, J.C., 2017. Development of the lung. Cell Tissue Res. 367, 427–444. https://doi.org/10.1007/s00441-016-2545-0
Schoch, K.G., Lori, A., Burns, K.A., Eldred, T., Olsen, J.C., Randell, S.H., 2004. A subset of mouse tracheal epithelial basal cells generates large colonies in vitro. Am. J. Physiol.-Lung Cell. Mol. Physiol. 286, L631–L642. https://doi.org/10.1152/ajplung.00112.2003
Sekine, K., Ohuchi, H., Fujiwara, M., Yamasaki, M., Yoshizawa, T., Sato, T., Yagishita, N., Matsui, D., Koga, Y., Itoh, N., Kato, S., 1999. Fgf10 is essential for limb and lung formation. Nat. Genet. 21, 138–141. https://doi.org/10.1038/5096
Serls, A.E., 2004. Different thresholds of fibroblast growth factors pattern the ventral foregut into liver and lung. Development 132, 35–47. https://doi.org/10.1242/dev.01570
Serra, R., Pelton, R.W., Moses, H.L., 1994. TGF beta 1 inhibits branching morphogenesis and N-myc expression in lung bud organ cultures. Dev. Camb. Engl. 120, 2153–2161.
Sherwood, R.I., Maehr, R., Mazzoni, E.O., Melton, D.A., 2011. Wnt signaling specifies and patterns intestinal endoderm. Mech. Dev. 128, 387–400. https://doi.org/10.1016/j.mod.2011.07.005
Shirasawa, M., Fujiwara, N., Hirabayashi, S., Ohno, H., Iida, J., Makita, K., Hata, Y., 2004. Receptor for advanced glycation end-products is a marker of type I lung alveolar cells: RAGE is a marker of ATI cells. Genes Cells 9, 165–174. https://doi.org/10.1111/j.1356-9597.2004.00712.x
Shu, W., Guttentag, S., Wang, Z., Andl, T., Ballard, P., Lu, M.M., Piccolo, S., Birchmeier, W., Whitsett, J.A., Millar, S.E., Morrisey, E.E., 2005. Wnt/β-catenin signaling acts upstream of N-myc, BMP4, and FGF signaling to regulate proximal–distal patterning in the lung. Dev. Biol. 283, 226–239. https://doi.org/10.1016/j.ydbio.2005.04.014
Shu, W., Lu, M.M., Zhang, Y., Tucker, P.W., Zhou, D., Morrisey, E.E., 2007. Foxp2 and Foxp1 cooperatively regulate lung and esophagus development. Development 134, 1991–2000.
Singh, D., Zo, S.M., Kumar, A., Han, S.S., 2013. Engineering three-dimensional macroporous hydroxyethyl methacrylate-alginate-gelatin cryogel for growth and proliferation of lung epithelial cells. J. Biomater. Sci. Polym. Ed. 24, 1343–1359. https://doi.org/10.1080/09205063.2012.759505
Singh, G., Singh, J., Katyal, S.L., Brown, W.E., Kramps, J.A., Paradis, I.L., Dauber, J.H., Macpherson, T.A., Squeglia, N., 1988. Identification, cellular localization, isolation, and characterization of human Clara cell-specific 10 KD protein. J. Histochem. Cytochem. Off. J. Histochem. Soc. 36, 73–80. https://doi.org/10.1177/36.1.3275712
Smith, A.G., 2001. Embryo-Derived Stem Cells: Of Mice and Men. Annu. Rev. Cell Dev. Biol. 17, 435–462. https://doi.org/10.1146/annurev.cellbio.17.1.435
Smith, J.C., Price, B.M.J., Nimmen, K.V., Huylebroeck, D., 1990. Identification of a potent Xenopus mesoderm-inducing factor as a homologue of activin A. Nature 345, 729–731. https://doi.org/10.1038/345729a0
Sokocevic, D., Bonenfant, N.R., Wagner, D.E., Borg, Z.D., Lathrop, M.J., Lam, Y.W., Deng, B., Desarno, M.J., Ashikaga, T., Loi, R., Hoffman, A.M., Weiss, D.J., 2013. The effect of age and emphysematous and fibrotic injury on the re-cellularization of de-cellularized lungs. Biomaterials 34, 3256–3269. https://doi.org/10.1016/j.biomaterials.2013.01.028
Song, H., Yao, E., Lin, C., Gacayan, R., Chen, M.-H., Chuang, P.-T., 2012. Functional characterization of pulmonary neuroendocrine cells in lung development, injury, and tumorigenesis. Proc. Natl. Acad. Sci. 109, 17531–17536. https://doi.org/10.1073/pnas.1207238109
Spence, J.R., Lange, A.W., Lin, S.-C.J., Kaestner, K.H., Lowy, A.M., Kim, I., Whitsett, J.A., Wells, J.M., 2009. Sox17 regulates organ lineage segregation of ventral foregut progenitor cells. Dev Cell 17, 62–74.
Strano, A., Tuck, E., Stubbs, V.E., Livesey, F.J., 2020. Variable Outcomes in Neural Differentiation of Human PSCs Arise from Intrinsic Differences in Developmental Signaling Pathways. Cell Rep. 31, 107732. https://doi.org/10.1016/j.celrep.2020.107732

Strikoudis, A., Cieślak, A., Loffredo, L., Chen, Y.-W., Patel, N., Saqi, A., Lederer, D.J., Snoeck, H.-W., 2019. Modeling of Fibrotic Lung Disease Using 3D Organoids Derived from Human Pluripotent Stem Cells. Cell Rep. 27, 3709-3723.e5. https://doi.org/10.1016/j.celrep.2019.05.077

Sun, H., Zhu, Y., Pan, H., Chen, X., Balestrini, J.L., Lam, T.T., Kanyo, J.E., Eichmann, A., Gulati, M., Fares, W.H., Bai, H., Feghali-Bostwick, C.A., Gan, Y., Peng, X., Moore, M.W., White, E.S., Sava, P., Gonzalez, A.L., Cheng, Y., Niklason, L.E., Herzog, E.L., 2016. Netrin-1 Regulates Fibrocyte Accumulation in the Decellularized Fibrotic Sclerodermatous Lung Microenvironment and in Bleomycin-Induced Pulmonary Fibrosis. Arthritis Rheumatol. Hoboken NJ 68, 1251–1261. https://doi.org/10.1002/art.39575

Swarr, D.T., Morrisey, E.E., 2015. Lung Endoderm Morphogenesis: Gasping for Form and Function. Annu. Rev. Cell Dev. Biol. 31, 553–573. https://doi.org/10.1146/annurev-cellbio-100814-125249

Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., Jensen, L.J., Mering, C. von, 2019. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 47, D607–D613. https://doi.org/10.1093/nar/gky1131

Tada, S., 2005. Characterization of mesendoderm: a diverging point of the definitive endoderm and mesoderm in embryonic stem cell differentiation culture. Development 132, 4363–4374. https://doi.org/10.1242/dev.02005

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., Yamanaka, S., 2007. Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. Cell 131, 861–872. https://doi.org/10.1016/j.cell.2007.11.019

Takahashi, K., Yamanaka, S., 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126, 663–676. https://doi.org/10.1016/j.cell.2006.07.024

Tata, P.R., Mou, H., Pardo-Saganta, A., Zhao, R., Prabhu, M., Law, B.M., Vinarsky, V., Cho, J.L., Breton, S., Sahay, A., Medoff, B.D., Rajagopal, J., 2013. Dedifferentiation of committed epithelial cells into stem cells in vivo. Nature 503, 218–223. https://doi.org/10.1038/nature12777

Tefft, J.D., Lee, M., Smith, S., Leinwand, M., Zhao, J., Bringas, P., Crowe, D.L., Warburton, D., 1999. Conserved function of mSpry-2, a murine homolog of Drosophila sprouty, which negatively modulates respiratory organogenesis. Curr. Biol. 9, 219–222. https://doi.org/10.1016/S0960-9822(99)80094-3

Teixeira, V.H., Nadarajan, P., Graham, T.A., Pipinikas, C.P., Brown, J.M., Falzon, M., Nye, E., Poulsom, R., Lawrence, D., Wright, N.A., McDonald, S., Giangreco, A., Simons, B.D., Janes, S.M., 2013. Stochastic homeostasis in human airway epithelium is achieved by neutral competition of basal cell progenitors. eLife 2, e00966. https://doi.org/10.7554/eLife.00966

Teo, A.K.K., Arnold, S.J., Trotter, M.W.B., Brown, S., Ang, L.T., Chng, Z., Robertson, E.J., Dunn, N.R., Vallier, L., 2011. Pluripotency factors regulate definitive endoderm specification through eomesodermin. Genes Dev. 25, 238–250. https://doi.org/10.1101/gad.607311

Thoma, E.C., Merkl, C., Heckel, T., Haab, R., Knoflach, F., Nowaczyk, C., Flint, N., Jagasia, R., Jensen Zoffmann, S., Truong, H.H., Petitjean, P., Jessberger, S., Graf, M., Iacone, R., 2014. Chemical conversion of human fibroblasts into functional Schwann cells. Stem Cell Rep. 3, 539–547. https://doi.org/10.1016/j.stemcr.2014.07.014

Thomas, P.D., Campbell, M.J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., Diemer, K., Muruganujan, A., Narechania, A., 2003. PANTHER: a library of protein families and subfamilies indexed by function. Genome Res 13, 2129–2141.

Thomsen, G., Woolf, T., Whitman, M., Sokol, S., Vaughan, J., Vale, W., Melton, D.A., 1990. Activins are expressed early in Xenopus embryogenesis and can induce axial mesoderm and anterior structures. Cell 63, 485–493. https://doi.org/10.1016/0092-8674(90)90445-K

Thomson, J.A., 1998. Embryonic Stem Cell Lines Derived from Human Blastocysts. Science 282, 1145–1147. https://doi.org/10.1126/science.282.5391.1145

Tichelaar, J.W., Wert, S.E., Costa, R.H., Kimura, S., Whitsett, J.A., 1999. HNF-3/Forkhead Homologue-4 (HFH-4) Is Expressed in Ciliated Epithelial Cells in the Developing Mouse Lung. J. Histochem. Cytochem. 47, 823–831. https://doi.org/10.1177/002215549904700612

Togo, S., Sato, T., Sugiura, H., Wang, X., Basma, H., Nelson, A., Liu, X., Bargar, T.W., Sharp, J.G., Rennard, S.I., 2011. Differentiation of embryonic stem cells into fibroblast-like cells in three-dimensional type I collagen gel cultures. Vitro Cell. Dev. Biol. - Anim. 47, 114–124. https://doi.org/10.1007/s11626-010-9367-2

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., Rinn, J.L., 2014. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat. Biotechnol. 32, 381–386. https://doi.org/10.1038/nbt.2859

Treutlein, B., Brownfield, D.G., Wu, A.R., Neff, N.F., Mantalas, G.L., Espinoza, F.H., Desai, T.J., Krasnow, M.A., Quake, S.R., 2014. Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature 509, 371–375. https://doi.org/10.1038/nature13173

Treutlein, B., Lee, Q.Y., Camp, J.G., Mall, M., Koh, W., Shariati, S.A.M., Sim, S., Neff, N.F., Skotheim, J.M., Wernig, M., Quake, S.R., 2016. Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq. Nature 534, 391–395. https://doi.org/10.1038/nature18323

Tsao, P.-N., Chen, F., Izvolsky, K.I., Walker, J., Kukuruzinska, M.A., Lu, J., Cardoso, W.V., 2008. γ-Secretase Activation of Notch Signaling Regulates the Balance of Proximal and Distal Fates in Progenitor Cells of the Developing Lung. J. Biol. Chem. 283, 29532–29544. https://doi.org/10.1074/jbc.M801565200

Tsao, P.-N., Matsuoka, C., Wei, S.-C., Sato, A., Sato, S., Hasegawa, K., Chen, H., Ling, T.-Y., Mori, M., Cardoso, W.V., Morimoto, M., 2016. Epithelial Notch signaling regulates lung alveolar morphogenesis and airway epithelial integrity. Proc. Natl. Acad. Sci. 113, 8242–8247. https://doi.org/10.1073/pnas.1511236113

Tsao, P.-N., Vasconcelos, M., Izvolsky, K.I., Qian, J., Lu, J., Cardoso, W.V., 2009. Notch signaling controls the balance of ciliated and secretory cell fates in developing airways. Dev. Camb. Engl. 136, 2297–2307. https://doi.org/10.1242/dev.034884

Tsuchiya, T., Sivarapatna, A., Rocco, K., Nanashima, A., Nagayasu, T., Niklason, L.E., 2014. Future prospects for tissue engineered lung transplantation: Decellularization and recellularization-based whole lung regeneration. Organogenesis 10, 196–207. https://doi.org/10.4161/org.27846

Turner, D.A., Hayward, P.C., Baillie-Johnson, P., Rué, P., Broome, R., Faunes, F., Martinez Arias, A., 2014. Wnt/β-catenin and FGF signalling direct the specification and maintenance of a neuromesodermal axial progenitor in ensembles of mouse embryonic stem cells. Development 141, 4243–4253.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., Cox, J., 2016. The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods 13, 731–740. https://doi.org/10.1038/nmeth.3901

Uhl, F.E., Wagner, D.E., Weiss, D.J., 2017. Preparation of Decellularized Lung Matrices for Cell Culture and Protein Analysis. Methods Mol. Biol. Clifton NJ 1627, 253–283. https://doi.org/10.1007/978-1-4939-7113-8_18

Vallier, L., Mendjan, S., Brown, S., Chng, Z., Teo, A., Smithers, L.E., Trotter, M.W.B., Cho, C.H.-H., Martinez, A., Rugg-Gunn, P., Brons, G., Pedersen, R.A., 2009. Activin/Nodal signalling maintains pluripotency by controlling Nanog expression. Development 136, 1339–1349. https://doi.org/10.1242/dev.033951

Van Raemdonck, D., Neyrinck, A., Verleden, G.M., Dupont, L., Coosemans, W., Decaluwé, H., Decker, G., De Leyn, P., Nafteux, P., Lerut, T., 2009. Lung donor selection and management. Proc. Am. Thorac. Soc. 6, 28–38. https://doi.org/10.1513/pats.200808-098GO

van Tuyl, M., Liu, J., Groenman, F., Ridsdale, R., Han, R.N.N., Venkatesh, V., Tibboel, D., Post, M., 2006. Iroquois genes influence proximo-distal morphogenesis during rat lung development. Am J Physiol Lung Cell Mol Physiol 290, L777–L789.

Vaughan, A.E., Brumwell, A.N., Xi, Y., Gotts, J.E., Brownfield, D.G., Treutlein, B., Tan, K., Tan, V., Liu, F.C., Looney, M.R., Matthay, M.A., Rock, J.R., Chapman, H.A., 2015. Lineage-negative progenitors mobilize to regenerate lung epithelium after major injury. Nature 517, 621–625. https://doi.org/10.1038/nature14112

Vaughan, M.B., Ramirez, R.D., Wright, W.E., Minna, J.D., Shay, J.W., 2006. A three-dimensional model of differentiation of immortalized human bronchial epithelial cells. Differentiation 74, 141–148. https://doi.org/10.1111/j.1432-0436.2006.00069.x

Volckaert, T., Yuan, T., Yuan, J., Boateng, E., Hopkins, S., Zhang, J.-S., Thannickal, V.J., Fässler, R., De Langhe, S.P., 2019. Hippo signaling promotes lung epithelial lineage commitment by curbing Fgf10 and β-catenin signaling. Dev. Camb. Engl. 146. https://doi.org/10.1242/dev.166454

Wagner, D.E., Bonvillain, R.W., Jensen, T., Girard, E.D., Bunnell, B.A., Finck, C.M., Hoffman, A.M., Weiss, D.J., 2013. Can stem cells be used to generate new lungs? *Ex vivo* lung bioengineering with decellularized whole lung scaffolds: Ex vivo lung bioengineering. Respirology 18, 895–911. https://doi.org/10.1111/resp.12102

Wamaitha, S.E., del Valle, I., Cho, L.T.Y., Wei, Y., Fogarty, N.M.E., Blakeley, P., Sherwood, R.I., Ji, H., Niakan, K.K., 2015. Gata6 potently initiates reprograming of pluripotent and differentiated cells to extraembryonic endoderm stem cells. Genes Dev 29, 1239–1255.

Wang, J.H., Deimling, S.J., D'Alessandro, N.E., Zhao, L., Possmayer, F., Drysdale, T.A., 2011. Retinoic acid is a key regulatory switch determining the difference between lung and thyroid fates in Xenopus laevis. BMC Dev. Biol. 11, 75. https://doi.org/10.1186/1471-213X-11-75

Wang, Z., Dollé, P., Cardoso, W.V., Niederreither, K., 2006. Retinoic acid regulates morphogenesis and patterning of posterior foregut derivatives. Dev. Biol. 297, 433–445. https://doi.org/10.1016/j.ydbio.2006.05.019

Wang, Z., Li, W., Guo, Q., Wang, Y., Ma, L., Zhang, X., 2018. Insulin-Like Growth Factor-1 Signaling in Lung Development and Inflammatory Lung Diseases. BioMed Res. Int. 2018, 1–27. https://doi.org/10.1155/2018/6057589

Wansleeben, C., Barkauskas, C.E., Rock, J.R., Hogan, B.L.M., 2013. Stem cells of the adult lung: their development and role in homeostasis, regeneration, and disease. Wiley Interdiscip. Rev. Dev. Biol. 2, 131–148. https://doi.org/10.1002/wdev.58

Warburton, D., Schwarz, M., Tefft, D., Flores-Delgado, G., Anderson, K.D., Cardoso, W.V., 2000. The molecular basis of lung morphogenesis. Mech. Dev. 92, 55–81. https://doi.org/10.1016/S0925-4773(99)00325-1

Weaver, M., Yingling, J.M., Dunn, N.R., Bellusci, S., Hogan, B.L., 1999. Bmp signaling regulates proximal-distal differentiation of endoderm in mouse lung development. Dev. Camb. Engl. 126, 4005–4015.

Weibel, E.R., 2013. Morphometry of the Human Lung. Elsevier Science, Burlington.

Wilkinson, D.G., Bhatt, S., Herrmann, B.G., 1990. Expression pattern of the mouse T gene and its role in mesoderm formation. Nature 343, 657–659. https://doi.org/10.1038/343657a0

Wolf, F.A., Angerer, P., Theis, F.J., 2018. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol 19, 15.

Wolf, F.A., Hamey, F.K., Plass, M., Solana, J., Dahlin, J.S., Göttgens, B., Rajewsky, N., Simon, L., Theis, F.J., 2019. PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. Genome Biol. 20, 59. https://doi.org/10.1186/s13059-019-1663-x

Wong, A.P., Chin, S., Xia, S., Garner, J., Bear, C.E., Rossant, J., 2015. Efficient generation of functional CFTR-expressing airway epithelial cells from human pluripotent stem cells. Nat. Protoc. 10, 363–381. https://doi.org/10.1038/nprot.2015.021

Wong, A.P., Shojaie, S., Liang, Q., Xia, S., Di Paola, M., Ahmadi, S., Bilodeau, C., Garner, J., Post, M., Duchesneau, P., Waddell, T.K., Bear, C.E., Nagy, A., Rossant, J., 2019. Conversion of human and mouse fibroblasts into lung-like epithelial cells. Sci. Rep. 9, 9027. https://doi.org/10.1038/s41598-019-45195-y

Wu, T., Economopoulos, K.P., Ott, H.C., 2017. Engineering Bioartificial Lungs for Transplantation. Curr. Stem Cell Rep. 3, 55–67. https://doi.org/10.1007/s40778-017-0082-8

Xing, Y., Li, A., Borok, Z., Li, C., Minoo, P., 2012. NOTCH1 is required for regeneration of Clara cells during repair of airway injury. Stem Cells Dayt. Ohio 30, 946–955. https://doi.org/10.1002/stem.1059

Yamamoto, Y., Gotoh, S., Korogi, Y., Seki, M., Konishi, S., Ikeo, S., Sone, N., Nagasaki, T., Matsumoto, H., Muro, S., Ito, I., Hirai, T., Kohno, T., Suzuki, Y., Mishima, M., 2017. Long-term expansion of alveolar stem cells derived from human iPS cells in organoids. Nat. Methods 14, 1097–1106. https://doi.org/10.1038/nmeth.4448

Yeom, Y.I., Fuhrmann, G., Ovitt, C.E., Brehm, A., Ohbo, K., Gross, M., Hübner, K., Schöler, H.R., 1996. Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. Dev. Camb. Engl. 122, 881–894.

Yusen, R.D., Edwards, L.B., Kucheryavaya, A.Y., Benden, C., Dipchand, A.I., Goldfarb, S.B., Levvey, B.J., Lund, L.H., Meiser, B., Rossano, J.W., Stehlik, J., 2015. The Registry of the International Society for Heart and Lung Transplantation: Thirty-second Official Adult Lung and Heart-Lung Transplantation Report—2015; Focus Theme: Early Graft Failure. J. Heart Lung Transplant. 34, 1264–1277. https://doi.org/10.1016/j.healun.2015.08.014

Zaret, K.S., 2016. From Endoderm to Liver Bud: Paradigms of Cell Type Specification and Tissue Morphogenesis. Curr. Top. Dev. Biol. 117, 647–669. https://doi.org/10.1016/bs.ctdb.2015.12.015

Zaret, K.S., Carroll, J.S., 2011. Pioneer transcription factors: establishing competence for gene expression. Genes Dev 25, 2227–2241.

Zemke, A.C., Teisanu, R.M., Giangreco, A., Drake, J.A., Brockway, B.L., Reynolds, S.D., Stripp, B.R., 2009. beta-Catenin is not necessary for maintenance or repair of the bronchiolar epithelium. Am. J. Respir. Cell Mol. Biol. 41, 535–543. https://doi.org/10.1165/rcmb.2008-0407OC

Zhao, L., Yee, M., O'Reilly, M.A., 2013. Transdifferentiation of alveolar epithelial type II to type I cells is controlled by opposing TGF-β and BMP signaling. Am. J. Physiol. Lung Cell. Mol. Physiol. 305, L409-418. https://doi.org/10.1152/ajplung.00032.2013

Zhao, R., Fallon, T.R., Saladi, S.V., Pardo-Saganta, A., Villoria, J., Mou, H., Vinarsky, V., Gonzalez-Celeiro, M., Nunna, N., Hariri, L.P., Camargo, F., Ellisen, L.W., Rajagopal, J., 2014. Yap Tunes Airway Epithelial Size and Architecture by Regulating the Identity, Maintenance, and Self-Renewal of Stem Cells. Dev. Cell 30, 151–165. https://doi.org/10.1016/j.devcel.2014.06.004

Zhao, R., Watt, A.J., Li, J., Luebke-Wheeler, J., Morrisey, E.E., Duncan, S.A., 2005. GATA6 is essential for embryonic development of the liver but dispensable for early heart formation. Mol Cell Biol 25, 2622–2631.

Ziegenhain, C., Vieth, B., Parekh, S., Reinius, B., Guillaumet-Adkins, A., Smets, M., Leonhardt, H., Heyn, H., Hellmann, I., Enard, W., 2017. Comparative Analysis of Single-Cell RNA Sequencing Methods. Mol. Cell 65, 631-643.e4. https://doi.org/10.1016/j.molcel.2017.01.023

Zorn, A.M., Wells, J.M., 2009. Vertebrate endoderm development and organ formation. Annu. Rev. Cell Dev. Biol. 25, 221–251. https://doi.org/10.1146/annurev.cellbio.042308.113344