

OPEN

# Reversal of Surfactant Protein B Deficiency in Patient Specific Human Induced Pluripotent Stem Cell Derived Lung Organoids by Gene Therapy

Sandra Lawrynowicz Leibel<sup>1,2,3</sup>, Alicia Winqvist<sup>2,3</sup>, Irene Tseu<sup>4</sup>, Jinxia Wang<sup>4</sup>, Daochun Luo<sup>4</sup>, Sharareh Shojaie<sup>4</sup>, Neal Nathan<sup>2,3</sup>, Evan Snyder<sup>1,2,3</sup> & Martin Post<sup>4,5</sup>

Surfactant protein B (SFTPB) deficiency is a fatal disease affecting newborn infants. Surfactant is produced by alveolar type II cells which can be differentiated *in vitro* from patient specific induced pluripotent stem cell (iPSC)-derived lung organoids. Here we show the differentiation of patient specific iPSCs derived from a patient with SFTPB deficiency into lung organoids with mesenchymal and epithelial cell populations from both the proximal and distal portions of the human lung. We alter the deficiency by infecting the SFTPB deficient iPSCs with a lentivirus carrying the wild type SFTPB gene. After differentiating the mutant and corrected cells into lung organoids, we show expression of SFTPB mRNA during endodermal and organoid differentiation but the protein product only after organoid differentiation. We also show the presence of normal lamellar bodies and the secretion of surfactant into the cell culture medium in the organoids of lentiviral infected cells. These findings suggest that a lethal lung disease can be targeted and corrected in a human lung organoid model *in vitro*.

Lung development is a complicated process that involves the specialization of multiple bronchiolar and alveolar epithelial populations. The primordial lung bud expresses Nkx2 homeobox 1 (NKX2-1), an important transcription factor in lung development<sup>1</sup> that also activates surfactant related genes<sup>2</sup>. Alveolar type II (ATII) epithelial cells synthesize, secrete and recycle all components of surfactant and dysfunction in surfactant metabolism can result in a variety of pediatric lung diseases including respiratory distress syndrome and interstitial lung disease<sup>3</sup>. To understand the various mechanisms driving disease, it is important to have a good model system *in vitro* that can recapitulate the *in vivo* disease process.

Lung organoids have been successfully used to study lung development and disease<sup>4–8</sup>. They are 3-dimensional, complex, multicellular structures that can be derived from patient specific human induced pluripotent stem cells (hiPSCs). hiPSC derived lung organoids represent early stages (pseudoglandular and canalicular) of lung development<sup>9,10</sup> and can contain multicellular structures<sup>9,11</sup> or be pure alveolar spheroids<sup>12,13</sup>. They can be used to study cellular and metabolic biology without the use of an animal model or fetal tissue. Although many differentiation protocols in the literature have been successful in mimicking lung development from stem cells, there has not been an examination of how a specific mutation impacts the differentiation process including its effects on the early endoderm, as well as the proximal and distal lung epithelial cell populations in the lung organoids.

Inherited deficiency of surfactant protein B (SFTPB) is a rare genetic cause of lethal respiratory distress syndrome in full-term newborn infants<sup>14,15</sup>. It is most commonly caused by a homozygous, frameshift, loss of function mutation of the surfactant protein B gene (p.Pro133GlnfsTer95, previously known as 121ins2)<sup>16–18</sup> making it a good target for gene therapy. Gene therapy has been utilized successfully to repair or inactivate mutations

<sup>1</sup>Department of Pediatrics, University of California, San Diego, Rady Children's Hospital, San Diego, La Jolla, CA, USA.

<sup>2</sup>Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA, USA. <sup>3</sup>Sanford Consortium for Regenerative Medicine, La Jolla, CA, USA. <sup>4</sup>Translational Medicine Program, Peter Gilgan Centre for Research and Learning, Hospital for Sick Children, Toronto, Ontario, Canada. <sup>5</sup>Department of Physiology, University of Toronto, Toronto, Ontario, Canada. Correspondence and requests for materials should be addressed to S.L.L. (email: [saleibel@ucsd.edu](mailto:saleibel@ucsd.edu))

in animal models of monogenic human diseases<sup>19</sup> as well as human *in vitro* cells<sup>12</sup>. There have been successful attempts at gene editing SFTPB deficiency including nuclease encoding mRNA, electroporation-mediated gene delivery and CRISPR<sup>12,19–21</sup>, but these methods have a low efficiency of mutation correction and are time intensive. Lentiviral (LV) vectors of the Retroviridae family show interesting properties for monogenic gene therapy, since they integrate into the host genome and allow long-lasting gene expression<sup>22</sup>. Lentiviral correction of monogenic mutations has been used in human disease<sup>22,23</sup>, is a highly efficient process and can be used to induce over-expression of proteins of interest after intratracheal delivery in animal models<sup>24</sup>. Infection of iPSCs with lentiviral inserts is a highly efficient process since stem cells grow quickly, remain undifferentiated in specific cell culture conditions and can establish fully infected clones within 2–3 passages. Lentiviral infection has also been targeted in lung epithelial cells<sup>24</sup> but these grow much slower *in vitro* and are prone to dedifferentiation.

In this study we generate SFTPB-deficient (p.Pro133GlnfsTer95) hiPSCs from patient specific fibroblasts and use a modified differentiation technique to transform them into 3-dimensional (3D) human lung organoids containing both epithelial and mesenchymal cell populations of the proximal and distal airways to closely replicate the human fetal lung. We employ highly efficient lentiviral infection of the wild type SFTPB gene into the mutated hiPSC line and show successful transcription and translation of SFTPB at the organoid level, the presence of lamellar bodies in ATII cells as well as the secretion of surfactant bioactive lipids by functional ATII cells.

## Results

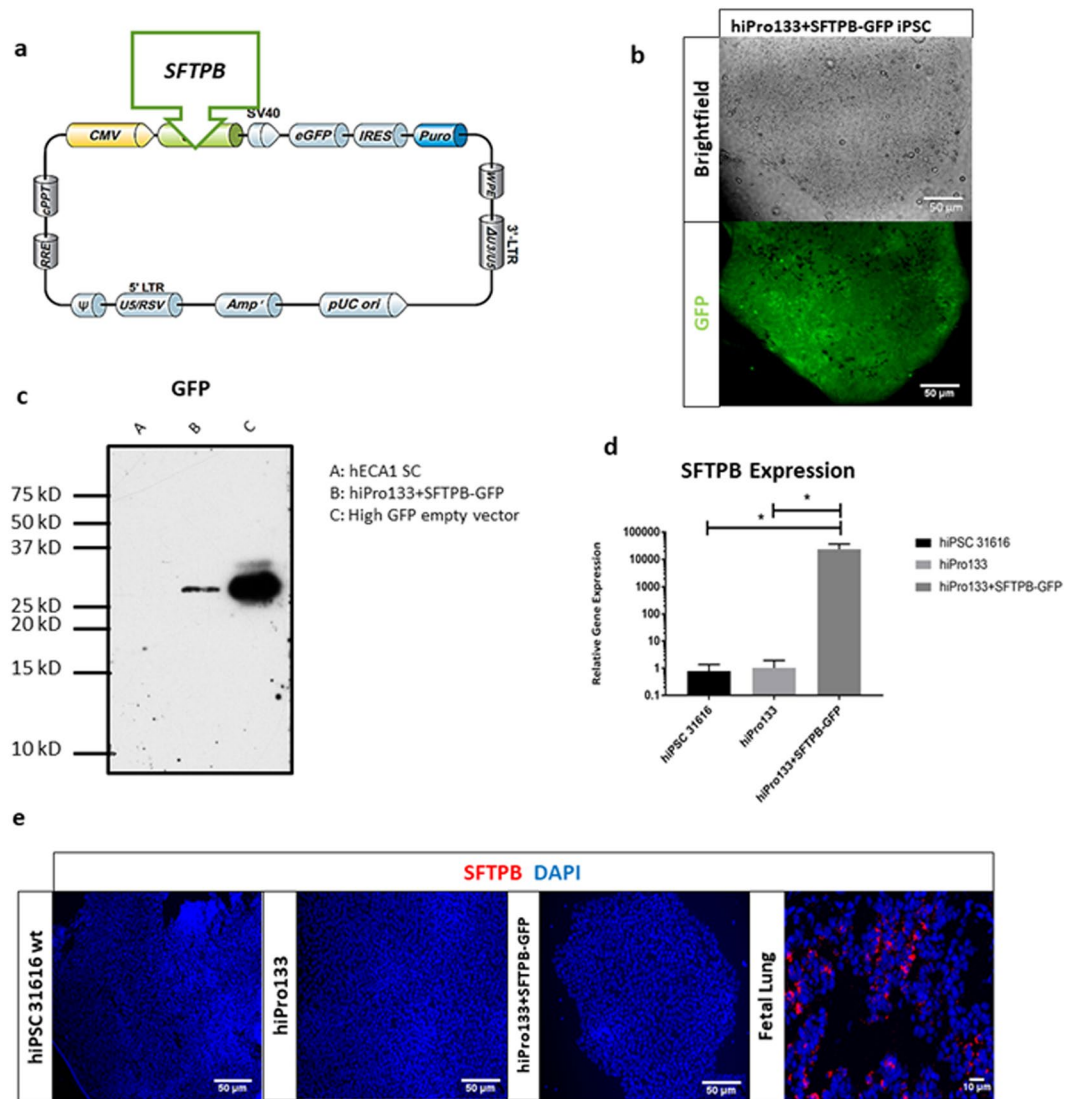
**Generation of patient specific iPSC cells containing the SFTPB deficiency.** With parental consent, a skin biopsy of a patient with the p.Pro133GlnfsTer95 (hereafter known as Pro133) SFTPB mutation<sup>18</sup> was obtained and expanded in culture. The fibroblasts were then reprogrammed into induced pluripotent stem cells (iPSCs) using a highly efficient Sendai-virus (RNA virus) based system<sup>25</sup> carrying the transcription factors hOct4, hSox2, hKlf4 and hMyc. The resulting hiPSCs (hiPro133 cells) expressed the pluripotency markers OCT4, NANOG, TRA-1-60 and SSEA4 (Supplementary Fig. S1) and the karyotype was normal (Supplementary Fig. S2). Genomic sequencing of the hiPSCs confirmed the presence of the frameshift mutation caused by the substitution of GAA for the nucleotide C (Supplementary Fig. S2). The Pro133 mutation introduces a Sfu I restriction enzyme site that permits detection of the mutation with restriction enzyme digestion<sup>26</sup>. The homozygous mutation was confirmed with Sfu I digestion revealing two distinct bands (210 and 103 bp) in the SFTPB deficient iPSCs and fibroblasts, one band (311 bp) in the wild type fibroblasts and three bands in the heterozygous fibroblasts (Supplementary Fig. S2).

**Infection of hPro133 iPSC cells with a lentivirus carrying the SFTPB cDNA sequence.** We used a commercially available lentiviral vector construct containing a CMV promoter which constitutively drove the open reading frame of homo sapiens SFTPB gene (Fig. 1a). This was followed by a SV40 promoter driving eGFP-IRES-puromycin to aid with selection after infection. We observed GFP-positive iPSC colonies that were successfully infected by the lentiviral vector (hereafter called hiPro133 + SFTPB-GFP) (Fig. 1b) and the GFP protein was confirmed through immunoblotting of the transduced cells (Fig. 1c). We also measured SFTPB gene expression in the transduced iPSC cells in comparison to the SFTPB-deficient iPSC cells (hiPro133) and a normal control iPSC line, 31616 (Fig. 1d). Expression levels of SFTPB were measured by qPCR and normalized to the housekeeping gene ACTB ( $\beta$ -actin). SFTPB mRNA expression was significantly higher in the hiPro133 + SFTPB-GFP iPSCs compared to hiPro133 and hi31616 cells. Finally, we examined whether the translation of the SFTPB gene product occurred at the iPSC level in the same cell lines through immunofluorescence using a SFTPB antibody that detects both the pro- and mature forms of SFTPB. As compared to our positive control human fetal lung, there was no protein expression of SFTPB in the hi31616, hiPro133 and hiPro133 + SFTPB-GFP iPSC lines (Fig. 1e).

**Differentiation of hPro133 + SFTPB-GFP and hPro133 iPSC cells into lung progenitor cells.** The mutant, wild type and corrected iPSC clones underwent successful endodermal differentiation into NKX2-1 expressing lung progenitor cells<sup>27,28</sup>. Briefly, mimicking fetal lung development, monolayer cultures were exposed to a cocktail of growth factors and small molecules to become definitive endoderm (DE)<sup>29</sup>, anterior foregut endoderm (AFE)<sup>30</sup> and lung progenitor cells (LPC)<sup>31,32</sup> (Fig. 2a). Throughout the differentiation process, gene expression analyses of the human embryonic stem cell line CA1 (heCA1) and/or the iPSC line hi31616 served as normative controls. Relative gene expression throughout endodermal differentiation was analyzed using qPCR normalized to ACTB. Gene expression profiles during the differentiation from stem cell to lung progenitor cell between the heCA1, hiPro133 and hiPro133 + SFTPB-GFP cell lines, showed no difference during DE and LPC induction, but in AFE the expression of SOX2 and FOXA2 was statistically greater in the heCA1 cell line compared to the hiPro133 and hiPro133 + SFTPB-GFP cell lines (Fig. 2b).

To ensure robustness of transgene expression in the hiPro133 + SFTPB-GFP cell line, the cells were exposed to puromycin at 0.25  $\mu$ g/mL for 7 days prior to differentiation. GFP expression was checked by FACS at each stage and a robust signal of >95% FITC (GFP) positive cells was consistently detected in the hiPro133 + SFTPB-GFP cell line through DE and AFE with a slight decrease in signal noted in LPC (Fig. 2e). In conjunction with the GFP signal, gene expression of SFTPB was analyzed throughout the differentiation from hiPSC to LPC. Independent of the state of differentiation, SFTPB mRNA expression remained high in the hiPro133 + SFTPB-GFP cell line compared to the wild type hi31616 and mutant hiPro133 lines in line with the constitutive nature of the CMV promoter in the hiPro133 + SFTPB-GFP cell line (Fig. 2c).

Differentiation efficiency for each cell line at each stage of endodermal differentiation was evaluated. After exposure to elevated levels of human Activin A for 3–4 days, the hiPSC derived DE cells expressed the cell surface markers CXCR4 and EPCAM through fluorescence-activated cell sorting (FACS) and SOX17 through immunofluorescence (Fig. 2d). Our differentiations consistently yielded an 80–95% pure population of hiPro133 and

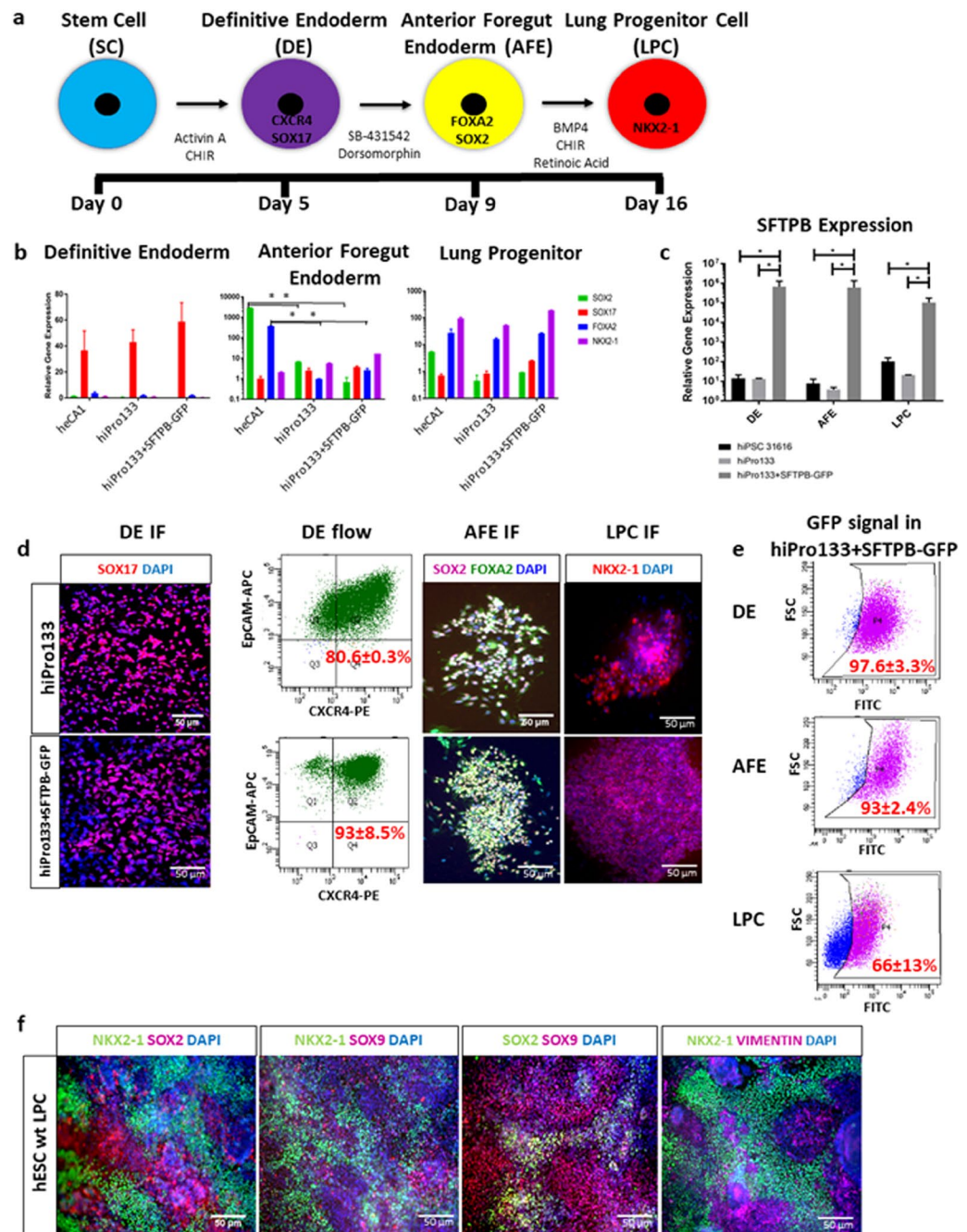


**Figure 1.** Human hiPro133 iPSC cells were successfully infected with a lentivirus carrying the pre-form SFTPb cDNA sequence. **(a)** A schematic of the lentiviral vector containing a CMV promoter driving SFTPb followed by SV40-eGFP-IRES-puromycin for selection. The lentiviral infected iPSCs hereafter known as hiPro133 + SFTPb-GFP **(b)** A phase contrast (top) and GFP live cell (bottom) image of hiPro133 + SFTPb-GFP cells (Scale bar = 50  $\mu$ m). **(c)** Western blot of GFP expression comparing a wild type non-infected heCA1 stem cell, the hiPro133 + SFTPb-GFP cell and an empty vector control (exposure time 2 minutes). **(d)** Gene expression of SFTPb in the hiPro133 + SFTPb-GFP cell line compared to an untransduced hiPro133 and wt 31616 iPSC cell lines. \* represents significant p-value <0.05 (n = 3 iPSC colonies). **(e)** Immunofluorescence examining protein expression of SFTPb in the iPSC cell lines hi31616, hiPro133 and hiPro133 + SFTPb-GFP and fetal lung tissue. No expression was detected on 3 different staining procedures compared to the positive control. (Scale bar = 50  $\mu$ m).

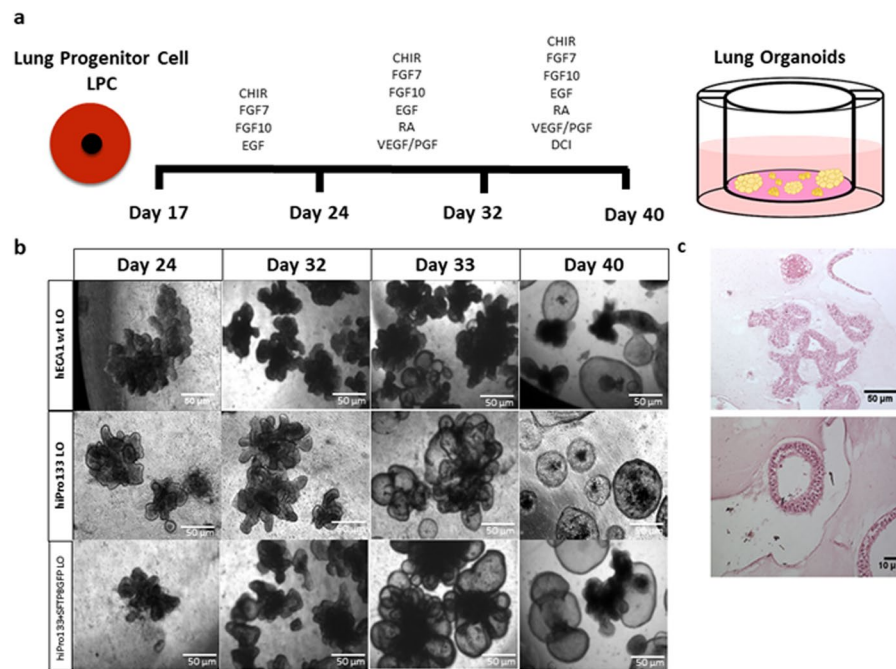
hiPro133 + SFTPb-GFP iPSC derived DE cells. Induction into AFE cells was confirmed by the co-expression of FOXA2 and SOX2<sup>33</sup> using immunofluorescence (Fig. 2d) with an efficiency of 70–90% in hiPro133 and hiPro133 + SFTPb-GFP iPSC derived AFE cells. Finally, LPC induction was confirmed with immunofluorescence staining using an antibody against the transcription factor NKX2-1 (Fig. 2d) with an efficiency of 30–50% in hiPro133 and hiPro133 + SFTPb-GFP iPSC derived LPC cells.

At the end of endodermal differentiation, prior to passing into 3D Matrigel, the LPCs were evaluated through immunostaining for NKX2-1, SOX2, SOX9 and vimentin. SOX2 is a marker of early lung and esophageal development and is expressed in the dorsal endoderm with NKX2-1 expression restricted to the ventral endoderm<sup>34</sup>. SOX9 expression is restricted to the cartilage of the trachea and progenitor cells important for branching morphogenesis in early lung development<sup>35,36</sup>. Vimentin is a marker for mesenchymal cells that are important in early lung development. Mesenchyme is derived from the mesodermal germ layer and is an indicator of the purity of the endodermal differentiation. The antibodies to the markers described above revealed distinct populations of epithelial (NKX2-1<sup>+</sup>) and mesenchymal (VIM<sup>+</sup>) cells at the lung progenitor stage (Fig. 2f). We noted that





**Figure 2.** Endodermal differentiation protocol and results of human CA1 ESCs, and 31616 wt, hiPro133 and hiPro133 + SFTPB-GFP iPS cells to lung progenitor cells. **(a)** Timeline showing culture steps from embryonic stem/human pluripotent stem cells to definitive endoderm (DE), anterior foregut endoderm (AFE) and lung progenitor (LPC) cells, including the markers expressed by the desired cell type and the combination of growth factors and small molecules used for the differentiation. **(b)** Gene expression profiles of common markers of endodermal differentiation of the wt heCA1, hiPro133 and hiPro133 + SFTPB-GFP cells at DE, AFE, and LPC. \* represents significant p-value < 0.05. (n = 3–5 separate differentiations). **(c)** Gene expression of SFTPB during the differentiation of wt hi31616, mutant hiPro133, and rescued hiPro133 + SFTPB-GFP cells at DE, AFE and LPC. \* represents significant p-value < 0.05. (n = 3 separate differentiations). **(d)** Immunofluorescence for DE marker SOX17, AFE markers FOXA2 and SOX2 and LPC marker NKX2-1 in hiPro133 and hiPro133 + SFTPB-GFP differentiated cells. (n = 3–5 separate differentiations; Scale bar = 50  $\mu$ m). Second column shows flow cytometry results for definitive endoderm markers CXCR4 and EPCAM of hiPro133 and hiPro133 + SFTPB-GFP cells. Double positive CXCR4/EPCAM cells represent DE cell differentiation efficiency as a mean  $\pm$  SEM. Representative of > 4 experiments. **(e)** Scatter plot from a flow representing the endogenous GFP expression of hiPro133 + SFTPB-GFP cells at DE, AFE and LPC from 3 separate differentiations, represented as mean  $\pm$  SEM. Unstained hiPro133 DE cells served as negative control. **(f)** Immunofluorescence of wt hESC LPC monolayers for NKX2-1, SOX2, SOX9 and vimentin. DAPI is represented in blue. Fluorophores of antibodies as labelled on image. (n = 4 differentiations; Scale bar = 50  $\mu$ m).



**Figure 3.** Timeline and histology of the differentiation of human ES/iPS-derived lung progenitor cells to 3D lung organoids. **(a)** Timeline showing monolayer culture LPCs passed to 3D Matrigel in transwell inserts with corresponding growth factors and small molecules and days of differentiation. **(b)** Representative live phase contrast images of developing 3D lung organoids at different time points from wt heCA1, hiPro133 and hiPro133 + SFTPB-GFP cells. (Scale bar = 50  $\mu$ m). **(c)** Hematoxylin and eosin images of wt heCA1 day 40 lung organoids. (Scale bar = 50  $\mu$ m).

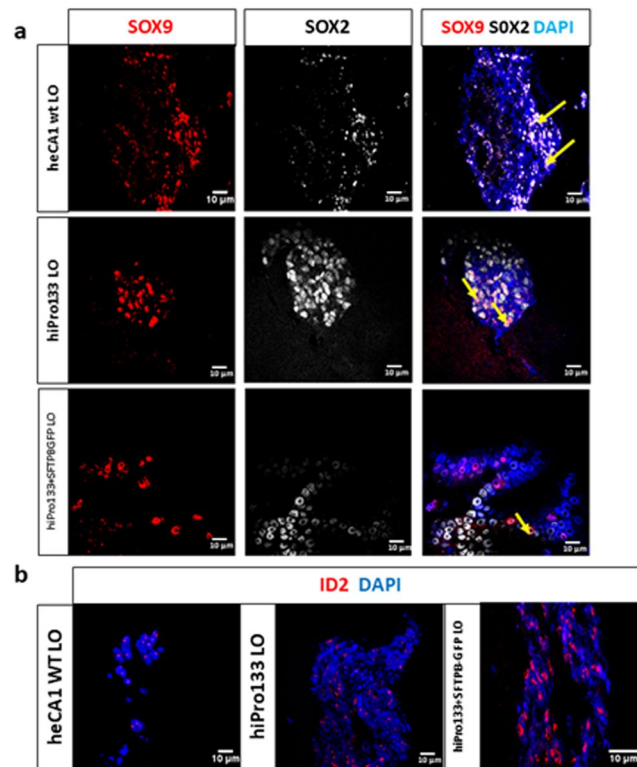
an efficient differentiation contained less VIM<sup>+</sup> cells (36  $\pm$  4%) while a non-efficient differentiation contained more VIM<sup>+</sup> cells (64  $\pm$  8%) (mean  $\pm$  s.e.m., n = 3 independent differentiations). Many investigators sort out the NKX2-1 positive cells using cell surface antigens in order to purify the endodermal population prior to passage into 3D Matrigel to make lung organoids<sup>27,32</sup>. We were interested in modeling lung development with a combination of endodermal and mesodermal elements to replicate epithelial-mesenchymal cross-talk and, therefore, passed the VIM<sup>+</sup> cells along with the NKX2-1<sup>+</sup> cells into 3D Matrigel from efficient LPC differentiations (NKX2-1<sup>+</sup> >49  $\pm$  14%). We also discovered the presence of single populations of SOX2<sup>+</sup> (31  $\pm$  4%), SOX9<sup>+</sup> (31  $\pm$  1%), and NKX2-1<sup>+</sup> (49  $\pm$  14%) cells and occasionally double positive SOX2/SOX9 cells (15  $\pm$  8.9%) in hESC LPCs (mean  $\pm$  s.e.m., n > 4 independent differentiations) (Fig. 2f).

#### Generation of 3D lung organoids from human hESC wt, hiPro133 + SFTPB-GFP and hiPro133 lung progenitor cells.

For the induction of 3D lung organoids, we passed the LPC monolayer as small aggregates containing both endodermal and mesodermal cells. The timeline for passaging LPC cells into matrigel-containing transwell inserts, along with the addition of growth factors and small molecules at specific time points during the differentiation is shown in Fig. 3a. Live phase contrast images of hESC wt, hiPro133 and hiPro133 + SFTPB-GFP LPC aggregates as they differentiate into lung organoids are shown in Fig. 3b. For our 3D lung organoid differentiation, we used chemical signals important in embryonic lung development to differentiate the LPCs into 3D lung organoids<sup>37-40</sup>. After passaging into 3D matrigel, the growth factors FGF7, FGF10 and EGF were used due to their importance in branching morphogenesis, growth and differentiation<sup>41,42</sup>. Since Wnt signaling is also important the glycogen synthase kinase (GSK-3) inhibitor CHIR was also added<sup>43</sup>. In the second week, RA and VEGF/PGF were added due to their importance in alveolar cell formation<sup>44,45</sup>. Finally, for epithelial maturation, a cocktail of dexamethasone, cAMP and isobutylxanthine was added in the final week<sup>40,46</sup>. Morphologically, one to two weeks after passaging LPC aggregates into 3D matrigel, there were multiple 3D organoids of mainly branching morphologies, but once DCI was added, the branching organoids were altered into more spherical forms by day 40 (Fig. 3b). At day 40, H&E staining was performed on the organoids which showed multiple cellular types including columnar and squamous epithelium (Fig. 3c).

#### Lung organoids from human hESC, hiPro133 and hiPro133 + SFTPB-GFP cells express branching tip and airway markers.

At the LPC stage, we identified epithelial cells positive for NKX2-1 and mesenchymal cells positive for vimentin (Fig. 2f). Therefore, by passaging them into 3D matrigel, we were expecting to more closely replicate fetal lung development and the epithelial-mesenchymal interactions important to branching morphogenesis. Our organoids expressed airway transcription factors such as SOX2 (proximal and branching marker) and SOX9 (branching and distal marker) and the bipotential branching tip progenitor ID2<sup>47,48</sup>. Occasionally, co-expression of both transcription factors was observed in the organoids (heCA1 wt 14  $\pm$  7%,



**Figure 4.** Branching tip and airway markers in lung organoids derived from wt heCA1, hiPro133 and hiPro133 + SFTPB-GFP stem cells. **(a)** Immunofluorescent expression of SOX2 and SOX9 in day 40 lung organoids from wt heCA1, hiPro133 and hiPro133 + SFTPB-GFP cells. Yellow arrows indicate double positive SOX9/SOX2 cells. Representative of >6–15 independent experiments. (Scale bar = 10  $\mu$ m). **(b)** Immunofluorescent expression of ID2 in day 40 lung organoids from wt heCA1, hiPro133 and hiPro133 + SFTPB-GFP cells. Representative of >6–15 independent experiments. (Scale bar = 10  $\mu$ m).

hiPro133  $3 \pm 2\%$  and hiPro133 + SFTPB-GFP  $8 \pm 5\%$  (mean  $\pm$  s.e.m.,  $n = 4$  independent differentiations) (Fig. 4a). The organoids were also positive for the transcription factor ID2, a specific marker found at the tips of the lung during branching morphogenesis (Fig. 4b). The ID2 staining is consistent with previous publications<sup>49</sup>.

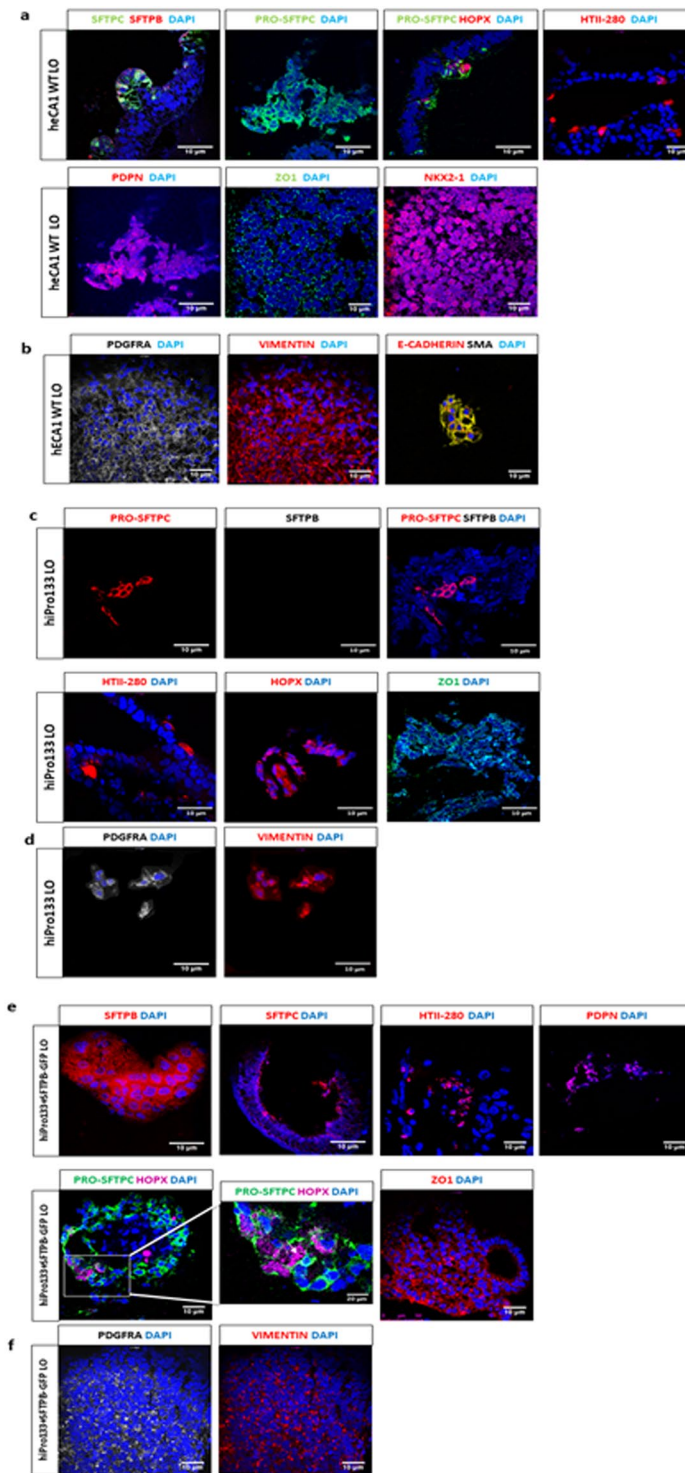
#### Lung organoids from human hESC, hiPro133 and hiPro133 + SFTPB-GFP cells express epithelial and mesenchymal markers of the proximal and distal lung.

After 6 weeks, lung organoids from heCA1 cells expressed a variety of immature and mature alveolar cell markers including cytosolic pro-SFTPC co-expressed with the bipotential alveolar progenitor transcription co-factor HOPX, podoplanin (PDPN an alveolar type I (ATI) marker), mature SFTPC and SFTPB (co-expressed ATII markers), and surface ATII marker HTII-280 (Fig. 5a). Of the total cells in the distal organoids,  $22 \pm 5.4\%$  stained positive for SFTPC and  $14.5 \pm 4.3\%$  SFTPB (mean  $\pm$  s.e.m.,  $n = 4$  independent differentiations). The transcription factor NKX2-1 that activates surfactant expression was abundantly expressed as was ZO-1 which is located on the intercellular tight junctions of alveolar cells (Fig. 5a). Furthermore, the organoids contained PDGFRA- and vimentin-positive mesenchymal cells and co-expressed the epithelial marker E-cadherin and the mesenchymal marker SMA (Fig. 5b).

We then applied our organoid differentiation protocol to the hiPro133 and hiPro133 + SFTPB-GFP iPSC-derived lung progenitor cells. ATII-specific surfactant protein B (SFTPB) was not detected in the hiPro133 iPSC-derived organoids but there was expression of HOPX and the tight junctional protein ZO-1 (Fig. 5c). There were a small number of cells in the organoids that expressed PRO-SFTPC ( $3.3 \pm 1.3\%$ , mean  $\pm$  s.e.m.,  $n = 3$  independent differentiations) and HTII-280. PDGFRA- and vimentin-positive mesenchymal cells were also present in the hiPro133 organoids (Fig. 5d). In contrast, the hiPro133 + SFTPB-GFP iPSC-derived distal lung organoids abundantly expressed SFTPB and mature SFTPC ( $36 \pm 8.6\%$  and  $18 \pm 4.4\%$  respectively, mean  $\pm$  s.e.m.,  $n = 3$  independent experiments). The organoids also stained positively for the ATII cell marker HTII-280 and tight junctional protein ZO1. The bipotential distal progenitor cell marker HOPX was also expressed, sometimes in isolated HOPX<sup>+</sup> cells or co-expressed with pro-SFTPC in the same cell<sup>50</sup> (Fig. 5e). PDPN, an ATI cell marker, was also detected in the hiPro133 + SFTPB-GFP organoids. Like the control heCA1-derived lung organoids, PDGFRA<sup>+</sup>/VIM<sup>+</sup> cells were present in the hiPro133 + SFTPB-GFP organoids (Fig. 5f).

The lung organoids derived from heCA1 wt, hiPro133, and hiPro133 + SFTPB-GFP cells also expressed gene transcripts from proximal airway epithelial cells such as SOX2, FOXJ1, KRT5, MUC5AC and SCGB1A1. Significant differences in KRT5 and FOXJ1 mRNA expression was noted between the wt and mutant cells





**Figure 5.** Expression of distal epithelial and mesenchymal lung markers in day 40 lung organoids derived from heCA1 cells, hiPro133 and hiPro133 + SFTPb-GFP stem cells. **(a)** Immunofluorescent expression of distal lung cell markers SFTPC, SFTPb, PRO-SFTPC, HOPX, HTII-280, PDPN, ZO-1 and NKX2-1 in day 40 wt heCA1 lung organoids. White arrow indicates the double positive HOPX/PRO-SFTPC cells. Data are representative of >15 independent experiments. (Scale bar = 10  $\mu$ m). **(b)** Immunofluorescent expression of mesenchymal markers PDGFRA and vimentin in day 40 wt heCA1 lung organoids. Data are representative of >15 independent experiments. (Scale bar = 10  $\mu$ m). **(c)** Immunofluorescent expression of distal lung cell markers PRO-SFTPC, SFTPb, HTII-280, HOPX, and ZO-1 in day 40 hiPro133 lung organoids. Data are representative of >15 independent experiments. Scale bars 10  $\mu$ m **(d)** Expression of mesenchymal markers PDGFRA and vimentin in day 40 hiPro133 lung organoids. Representative of >6 independent experiments. Scale bars 10  $\mu$ m. **(e)** Immunofluorescent expression of distal lung cell markers SFTPb, SFTPC, HTII-280, PDPN, PRO-SFTPC, HOPX, and ZO1 in day 40 hiPro133 + SFTPb-GFP lung organoids. White box represents area that is highlighted.

White arrows indicate the double positive HOPX/PRO-SFTPC cells. Representative of >6 independent experiments. (Scale bar = 10  $\mu$ m). (d) Expression of mesenchymal markers PDGFRA and vimentin in day 40 hiPro133 + SFTPB-GFP lung organoids. Representative of >6 independent experiments. (Scale bar = 10  $\mu$ m).

(Supplementary Fig. S3). Independent of cell line, all lung organoids stained positively for FOXJ1, a transcription factor found in ciliated cells, as well as basal cell and possible lung stem cell markers P63 and KRT5, the secretory cell marker SCGB3A2 and airway epithelial cell marker MUC5AC (Supplementary Fig. S3).

**Alveolar type II cells within the lung organoids have functional lamellar bodies.** Our human lung organoids derived from heCA1 and wild type and mutant hiPS cells expressed lamellar body specific transcripts SFTPB, SFTPC, ABCA3 and LAMP3 (Fig. 6a). However, expression of SFTPB and SFTPC were significantly down regulated in the hiPro133 organoids. Other lamellar body specific transcripts (ABCA3, LAMP3) were also decreased in the hiPro133 organoids. SFTPB is a highly processed protein which starts off as a 42 kDa pre-pro-SFTPB product then gets processed into a 23–25 kDa pro-SFTPB and finally is present as the functional 8 kDa product in lamellar bodies<sup>51</sup>. Using Western blotting, we confirmed the presence of the functional 8 kDa SFTPB protein in the lung organoids derived from all cell lines except hiPro133 (Fig. 6b). The expression of PRO-SFTPC protein was also decreased in the hiPro133 lung organoids. The lamellar bodies of the ATII cells were examined using transmission electron microscopy. We detected ATII cells in the hiPro133 + SFTPB-GFP iPSC-derived organoids containing well-organized cytosolic lamellar bodies, microvilli at the apical cell membranes, tubular myelin external to the cells and tight junctions located between them (Fig. 6c). We also examined the lamellar bodies of the ATII cells in hiPro133 organoids. Consistent with published literature reports on infants affected with SFTPB deficiency<sup>52,53</sup>, we observed no mature lamellar bodies or tubular myelin although the ATII cells had intact microvilli and tight junctions (Fig. 6c). Pulmonary surfactant is made up of 10% protein and 90% lipids, with the most abundant surface-active lipid being dipalmitoyl phosphatidylcholine (DPPC)<sup>54</sup>. To assess the functionality of the ATII cells in the lung organoids, we analyzed the ability to secrete surfactant material into the cell culture medium. We analyzed the presence of surface-active DPPC in the medium of organoids from hiPro133 and hiPro133 + SFTPB-GFP cell lines using tandem mass spectrometry (MS/MS). The organoids that over-expressed SFTPB showed a significant increase in the percentage of surfactant-specific DPPC secreted into the medium (Fig. 6d). These MS/MS results suggest functionality of the ATII cells in the lung organoids.

## Discussion

Our data demonstrate successful genotypic and phenotypic duplication of a fatal SFTPB mutation *in vitro* using lung organoids. We developed a modified lung organoid differentiation protocol for a mutant iPSC cell line derived from a patient with SFTPB deficiency (Pro133 iPSC cells) and showed expression of progenitor, proximal and distal lung cell markers as well as lung mesenchyme. We confirmed the absence of SFTPB genetic and protein material in the ATII cells of these hiPro133 iPSC-derived lung organoids and showed abnormal lamellar body formation using TEM, consistent with both murine and human data.

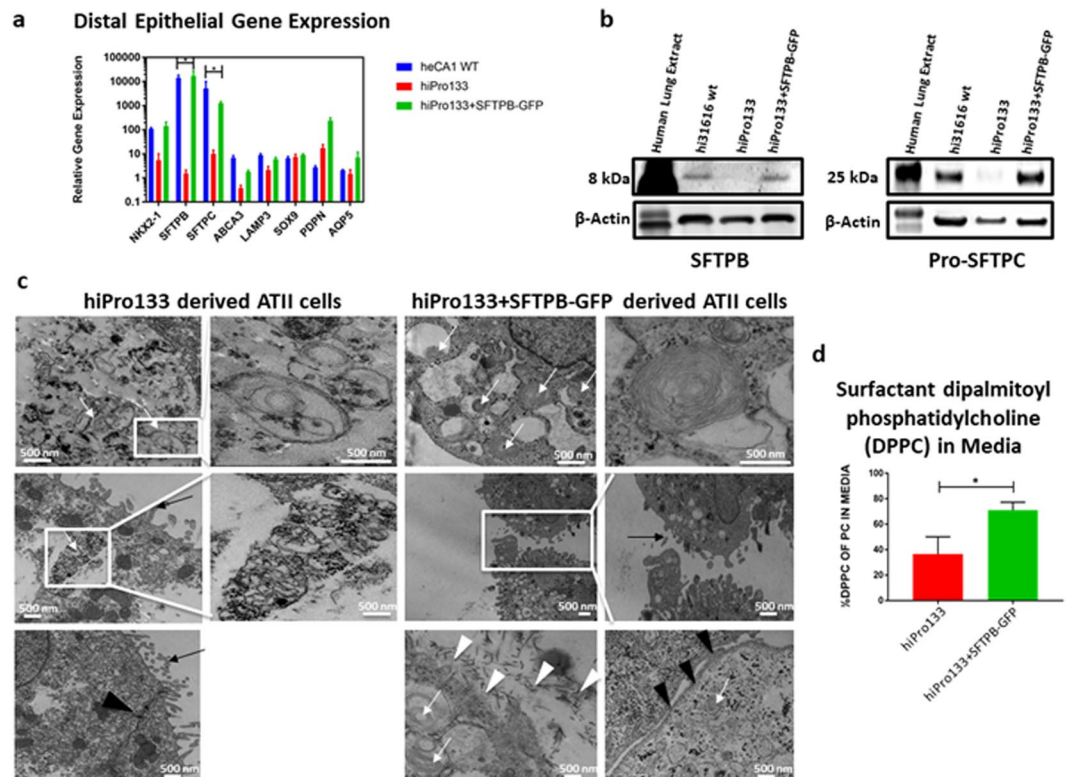
We decided to focus on the expression of established marker proteins for lung progenitors and distal lung epithelial cells using flow, immunostaining and western blotting. Proteins are the final product of gene expression and, therefore, an excellent tool to track lung epithelial differentiation. However, in some instances we performed qPCR of genes of interest to confirm and strengthen the protein analyses.

We also successfully infected a lentivirus carrying the wild-type SFTPB coding region into patient specific SFTPB deficient (Pro133) iPSC cells. We then differentiated the lentiviral infected (hiPro133 + SFTPB-GFP) iPSC cells into lung organoids and showed the presence of SFTPB genetic material without protein expression at the iPSC and early endodermal phases of differentiation and the presence of both genetic and protein material in the lung organoid. We showed normal lamellar bodies on TEM and the functionality of ATII cells in the hiPro133 + SFTPB-GFP lung organoids as demonstrated by the secretion of tubular myelin and surface-active DPPC into the culture medium. Together, our data show that lentiviral correction of surfactant protein B deficiency *in vitro* is feasible but whether this is achievable *in vivo* remains to be determined.

We chose to use lentiviral infection for many reasons, including evidence for its usefulness in other pulmonary diseases. Cystic fibrosis is a monogenic lung disease that has been a target of gene therapy. *In vitro* and *in vivo* studies have suggested that only 5–10% of normal CFTR function is required to reverse the chloride channel defect<sup>55,56</sup>, yet nearly every cell in the airway must be corrected to reverse the excess activity of the epithelial sodium channel (ENaC)<sup>57,58</sup> which is important in the reabsorption of sodium ions. Clinically, it may be more practical to have a minority of cells expressing much higher levels of the target protein to restore normal function. This would also apply to the surfactant protein B mutation. Some children with SFTPB mutations have a relatively milder phenotype<sup>59</sup> and this has led to the consideration of whether a small amount of mature SFTPB (approximately 8% to 10% of control levels) is sufficient for the milder phenotype. This hypothesis is supported by studies in which mice were genetically engineered to manipulate their production of SFTPB and developed respiratory failure when their SFTPB levels fell to 20–30% of those observed in controls<sup>60</sup>. *In vivo* patient specific cellular transplantation of corrected ATII cells would be an ideal therapeutic option in SFTPB deficiency, but studies have shown that after a relatively efficient initial deposition of cells through the trachea, long-term uptake, retention and differentiation of transplanted cells is relatively inefficient in the absence of ongoing injury<sup>61</sup>. This may support genetically engineering a small number of cells to secrete higher levels of protein.

Our transgene was inserted in the hiPro133 iPSC cells and expression was confirmed through the accompanying GFP signal as well as the expression of SFTPB transcripts throughout the differentiation. We also examined protein expression of SFTPB and did not detect it during the pluripotent state or through the stages of endodermal





**Figure 6.** Expression and function of distal lung cells in stem cell derived day 40 wild type, hiPro133 and hiPro133 + SFTP-B-GFP lung organoids. **(a)** Gene expression of alveolar epithelium and lamellar body related genes in day 40 heCA1 wt, hiPro133 and hiPro133 + SFTP-B-GFP lung organoids ( $n = 3-5$  separate differentiations). **(b)** Representative western blot of 8 kDa mature SFTP-B and 25 kDa PRO-SFTP-C expression in day 40 hi31616 wt, hiPro133 and hiPro133 + SFTP-B-GFP lung organoids ( $n = 3$  separate differentiations). **(c)** TEM images of day 40 hiPro133 and hiPro133 + SFTP-B-GFP lung organoids. White arrows indicate lamellar bodies, black arrows indicate microvilli, white arrowheads indicate tubular myelin and black arrowheads indicate tight junctions. White boxes highlight the areas that have been magnified. (Scale bar = 500 nm). **(d)** Phosphatidylcholine mass spectrometric analysis of culture medium of hiPro133 and hiPro133 + SFTP-B-GFP lung organoids ( $n = 3$  independent differentiations, \* represents significant  $p$ -value  $< 0.05$ ).

differentiation but did not detect SFTP-B protein after lung organoid differentiation. This may be explained by multiple factors including the lack of translation of the SFTP-B gene early on in development when the protein product is not required, the lack of activators and/or inducers to aid in translation at such an early stage or the complex processing of the SFTP-B mRNA product which may only occur in a specific cell type.

Our lung organoid differentiation protocol is unique in many ways. Compared to the other mixed epithelial/mesenchymal lung organoid protocols, we don't collect anterior foregut spheroids as the starting population for 3D organoids<sup>9,11</sup> or sort for surface markers of NKX2-1 positive cells<sup>12,32</sup>, but wait until the lung progenitor stage to pass the monolayer cells into matrigel. We don't use an external mesenchymal cell population to interact with the epithelial cell population<sup>32</sup> and our differentiation takes only 6 weeks as opposed to many months<sup>9</sup>. Our differentiation protocol is also free of any animal serum (FBS) which may introduce unknown growth factors into the differentiation cocktails<sup>11</sup>. We have shown the presence of lung cells at different stages of lung development including progenitor branching tip cells, proximal airway cells and distal progenitor and mature alveolar lung cells. Our efficiency of approximately 22% SFTP-C<sup>+</sup> cells is also higher than other mixed organoids quoted in the literature (eg. 5% in Dye *et al.*<sup>11</sup>) but less than that of pure alveolospheres (51% in Yamamoto *et al.*<sup>13</sup> and 69% in Jacob *et al.*<sup>12</sup>). These organoids can be used to model human lung disease such as bronchopulmonary dysplasia and viral infections since these affect multiple cell types of the lung, not just a single epithelial cell population. Interactions between mesenchyme and epithelium can also be examined since some of the supporting cells in the organoids stained positive for lung mesenchymal markers<sup>7</sup> including vimentin, PDGFRA and SMA.

There was a change in the morphology of the organoids in the third week of the 3D differentiation in matrigel after the addition of DCI. Twenty four hours after exposure to DCI, some of the dense branching organoids began "ballooning" out and forming transparent spheres. The type of iPSC line determined the number of organoids that became round and transparent. The hiPro133 (SPB deficient) derived organoids all became fully transparent spheres, the hiPro133 + SFTP-B-GFP derived organoids contained 75% transparent spheres and the wt heCA1 cells contained 50% transparent spheres at day 40 with the rest remaining dense branching organoids. We believe this morphological change is due to the maturation of some specific cell types in response to DCI but have not

elucidated which types of cell are affected. This cocktail of small molecules and hormones may be targeting the epithelial and/or the mesenchymal cell populations and activating a signaling pathway which causes the morphological changes. We aim to explore this phenomenon with bulk RNA seq after separating out the opaque and transparent organoids to determine the gene expression changes that occur after the introduction of DCI.

Our organoids did not express CD31, a well-known marker for endothelial cells (data not shown), suggesting that we did not have a vascularized organ.

Another interesting phenomenon we noted in our differentiation is the presence of progenitor branching tip cells in organoids at the same point in time of the differentiation as distal alveolar organoids. During the pseudoglandular stage, distal epithelial progenitors at the tips co-express SOX2 and SOX9<sup>48</sup>. This double-positive population is no longer present by the canalicular stages of development<sup>62</sup> and instead SOX9 becomes isolated to the distal alveolar population<sup>63</sup>. These two different populations of organoids in a single transwell may indicate that upon passing the LPC cells into 3D matrigel, organoids may not uniformly differentiate. By passaging the LPCs as cell aggregates, we may be creating organoids at various stages of lung development in a single transwell. This may be due to the LPCs containing a variety of SOX2<sup>+</sup> and SOX9<sup>+</sup> cells, either singly expressed or co-expressed, representing different cell types entering the early phases of lung development. To confirm this phenomenon, we would need to assess the transcriptomes of the LPCs as well as of the developing organoids and compare their gene expression patterns at each stage of lung development.

Finally, despite the importance of SFTPB in distal alveolar functions such as surfactant stability and SFTPB post translational modification, SFTPB-deficient iPSC derived organoids also showed changes in proximal airway gene and protein expression. SFTPB is expressed in Club cells in its pro-peptide form and may affect proximal airway cell differentiation if absent. Post translational modification of SFTPB has an important role in maintaining lung homeostasis in the mouse lung, yielding two different protein products, SP-B<sup>M</sup> and SP-B<sup>N</sup>, which are involved in lowering surface tension and promoting host defense, respectively<sup>64</sup>. Studies have shown that the balance of these two pro-peptides may alter susceptibility to lung infections<sup>65</sup> but how they impact lung development is unknown.

Taken together our work shows that human lung development and disease can be studied *in vitro* to discover new mechanisms of disease, patient specific therapeutic options and be the first step towards a cure for human lung disease.

## Methods

**Derivation of hiPS cells from fibroblasts.** Fibroblasts from a patient with the p.Pro133GlnfsTer95 SFTPB deficiency (generously donated by Drs. Cole and Hamvas, Washington University, St. Louis and Northwestern University, Chicago) were given to the Center for Commercialization of Regenerative Medicine (CCRM, Toronto) for hiPS cell derivation as well as fibroblasts from an individual with a normal surfactant profile. Reprogramming factors hOct4, hSox2, hKlf4 and hMyc (CytoTune iPSC Sendai Reprogramming Kit) were added to the fibroblasts and incubated for 24 hours. After 24 hours, the viral solution was removed and replaced with fresh fibroblast expansion medium and changed daily for a week. After day 7, cells were re-seeded onto 10-cm tissue culture dishes coated with human foreskin fibroblasts. Human ESC complete media made up of DMEM/F12, 20% (v/v) knockout serum replacement (Gibco (Life Technologies)), 0.1 mM  $\beta$ -mercaptoethanol (Sigma-Aldrich) and 10 ng/ml FGF2 (R&D) was exchanged daily until approximately day 28, or until colonies were ready for picking. Authentic colonies were identified, cleaned and expanded on human foreskin fibroblast-coated dishes in hESC complete media. The iPSC cells were characterized through karyotyping, immunostaining and FACS for OCT4, NANOG, SSEA4, TRA160, TRA181, and RT-PCR for pluripotency markers and cell line authentication (STR analysis) to confirm parental origin.

**Sfu1 digestion.** We performed PCR amplification in 10- $\mu$ L reaction volumes in each well of a 96-well plate as previously described<sup>26</sup>. Five  $\mu$ L of supernatant containing DNA template; 1.6  $\mu$ L deoxynucleotides (final concentration 200  $\mu$ M); 0.5  $\mu$ L of 5% DMSO; 0.75 U (0.1  $\mu$ L) *Taq* polymerase (Sigma Chemical Co., St. Louis, MO, U.S.A.), and 2  $\mu$ L of Buffer J [300 mM Tris HCl, 75 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, and 10 mM MgCl<sub>2</sub>, Invitrogen, Carlsbad, CA, USA]. To increase the specificity of amplification of the target sequence, we used 36-mer primers (sense: 5' TAA CTC CTT GGC ACT CGT GAA CTC CAG CAC CCT G 3'; antisense: 5' GCT GGC TGG GGT GCT GTG TGT GTG GCT CCC CCA TG 3') to amplify a 354-bp fragment of the SFTPB gene (nucleotides 1320–1674) that contains the 121ins2 mutation. We amplified extracted template in a thermocycler (MJ Research, Watertown, MA, U.S.A.) as follows: 3.5 min at 95 °C for initial denaturation, 30 cycles denaturation at 95 °C for 30 s each, annealing at 68 °C for 45 s, elongation at 72 °C for 80 s, and a final elongation at 72 °C for 3 min. We performed restriction enzyme digestion with *Sfu* I as previously described<sup>16</sup>. Briefly, we digested amplicons in each well with 2.5 U (5  $\mu$ L) *Sfu* I in H buffer (Roche Molecular Biochemicals, Indianapolis, IN, USA) by incubating at 37 °C for 90 min. We added 5  $\mu$ L of loading buffer and subjected each digest to metaphor/agarose-ethidium bromide gel electrophoresis. Each gel was photographed under UV light and scanned into a database.

**Lentiviral infection of human iPSC cells with SFTPB<sub>wt</sub>-GFP insert.** GeneCopoeia's EX-M0587-Lv201 lenti-vector was used (M0587). The vector contained a CMV promoter driving the open reading frame (ORF) of homo sapiens surfactant protein B (Accession: BC032785.1) followed by SV40-eGFP-IRES-puromycin. To generate the lentivirus, the HIV-based EX-M0587-Lv201 lenti-vector, in conjunction with GeneCopoeia's Lenti-Pac™ HIV Expression Packaging vectors were co-transfected into HEK293T cells using GeneCopoeia's EndoFectin™ Lenti Transfection Reagent. Cells were incubated in the presence of 5% CO<sub>2</sub> at 37 °C overnight. Growth medium was changed to Opti-MEM containing 3% (v/v) FBS with the addition of GeneCopoeia's Titerboost. Conditioned medium was collected after 24, 48 and 72 hours of incubation and centrifuged at 2,000  $\times$  g for 30 minutes. The supernatant was transferred to a new tube and PEG 6000 solution was then added to make the final PEG 6000

concentration to be 8.5% (w/v) and the final NaCl concentration to be 0.3 M. The mixture was incubated on ice for 3 to 6 hours and then centrifuged at  $2,000 \times g$  for 30 minutes. The viral particle containing pellet was resuspended by pipetting in 1/20 of the original harvest volume of Opti-MEM. Infection of the lentivirus was performed on 24-well or 6-well plates. Wt iPS and Pro133 iPS cells were plated one day before infection at 60–70% confluence. Five million transduction units (TU) per milliliter of infectious particle was used to infect  $10^6$  cells together with polybrene (final concentration 80  $\mu\text{g}$  per milliliter). Growth medium was changed one day after infection, with the addition of puromycin (0.25  $\mu\text{g}$  per milliliter). Cells were selected under puromycin for more than two weeks and positive cells were verified by GFP expression in iPS cells using fluorescence microscopy. For the verification of SFTP expression *in vitro*, total RNA was extracted from iPS cells using the RNeasy Mini Kit (Qiagen). Purified total RNA was reverse transcribed to cDNA using the miRNA reverse transcription kit with SFTP specific primers (Life Technologies). qRT-PCR was performed and analyzed with StepOne Software (Applied Biosystems). For sequences of primers, see Supplementary Table S1.

**Maintenance of hES and iPS cells.** Sendai generated human dermal fibroblasts iPSC lines (hi31616, hiPro133 and hiPro133 + SFTP-GFP cells) were cultured on matrigel coated plates in a MTESR 1 medium (STEMCELL Tech 85850). Medium was changed daily, and cells were passaged using ReLeSR (STEMCELL Tech 05872) every 5–7 days. Cultures were maintained in an undifferentiated state in a 5%  $\text{CO}_2$ /air environment. Human iPSC differentiations were carried out in a 5%  $\text{CO}_2$ /21%  $\text{O}_2$ /90%  $\text{N}_2$  environment.

**Directed differentiation of iPS cells to lung progenitor cells.** When the human iPSCs reached 70% confluence (Day 0), the cells were incubated in 10  $\mu\text{M}$  of Y-27632 (Wako) for one hour and then in accutase (Innovative Cell Technologies) for 20 minutes at 37°C<sup>32</sup>. The detached iPSCs were then dissociated into single cells via pipetting, and then seeded on matrigel-coated plates (BD Biosciences) at a density of  $1.75 \times 10^5$  cells/cm<sup>2</sup> in DE induction medium made up of RPMI1640 (Life Technologies), 1x B27 supplement, 1% HEPES, 1% glutamax and 50 U/ml of penicillin/streptomycin. This basal medium was supplemented with 100 ng/ml of human activin A (R&D systems), 1  $\mu\text{M}$  of CHIR99021 (Stemgent), and 10  $\mu\text{M}$  of Y-27632. On days 2–4, only activin was added. On Day 5, the medium was changed to AFE induction medium. Depending on cell line and confluence, some DE wells were passed in ratios of 1:1–1:3, while others underwent AFE induction without passage. AFE and LPC induction basal mediums consisted of 3:1 IMDM:F12 (Life Technologies), 1x B27 and N2 supplements (Life Technologies), 50 U/ml of penicillin/streptomycin, 0.25% BSA, 0.05 mg/ml of L-ascorbic acid (Sigma-Aldrich), and 0.4 mM of monothioglycerol (Wako)<sup>30</sup>. This medium was supplemented with 10  $\mu\text{M}$  SB431542 (R&D Systems) and 2  $\mu\text{M}$  Dorsomorphin (Stemgent) for AFE induction for 3 days.

On Day 8, the medium was changed to LPC induction medium, containing the basal medium with 10 ng/ml of human recombinant BMP4 (R&D Systems), 0.1  $\mu\text{M}$  of all-trans retinoic acid (RA) (Sigma-Aldrich) and 3  $\mu\text{M}$  of CHIR99021. Media was changed every 2 days for 8–11 days.

**3D culture of LPC cells to lung organoids.** The protocol for the 3D culture was modified from previous reports<sup>12,32</sup>. Monolayer LPC cells were incubated in 10  $\mu\text{M}$  of Y-27632 for one hour then placed in accutase for 10 min.  $10\text{--}40 \times 10^4$  cells were resuspended in 200  $\mu\text{l}$  of undiluted growth factor reduced matrigel and seeded onto a 12-well cell culture insert (BD Biosciences). One ml of 3D step 1 media (with FGF7 (10 ng/ml), FGF10 (10 ng/ml), CHIR (3  $\mu\text{M}$ ) and EGF (10 ng/ml) was added to the lower chamber and changed every other day for 8 days. D basal media was identical to the AFE/LPC induction basal media. On day 25, the media was changed to 3D step 2 medium consisting of FGF7 (10 ng/ml), FGF10 (10 ng/ml), CHIR (3  $\mu\text{M}$ ), RA (0.1  $\mu\text{M}$ ), EGF (10 ng/ml) and VEGF/PIGF (10 ng/ml). Media was changed every 2 days for 8 days. Finally, 3D step 3 media was added which was the same as step 2 with the addition of Dexamethasone (50 nM), cAMP (100  $\mu\text{M}$ ) and IBMX (100  $\mu\text{M}$ ). Media was changed every 2 days for 7 days.

**Flow cytometry.** The cells grown in 3D matrigel were dissociated with dispase (17105041 ThermoFisher) at 37°C for 1 hour centrifuged at  $300 g \times 5$  min and washed in 1x cold PBS. To generate single cell suspensions, cell pellets were incubated in TrypLE for 10–15 minutes. The reaction was diluted in DMEM/F12 (Life Technologies) with 2% (v/v) FBS and centrifuged at 300 g at room temperature. The cell pellets were immersed with sorting buffer 3% (v/v) FBS in PBS, cell aggregates were removed using a cell strainer with a 70- $\mu\text{m}$  pore size (BD falcon) and single cells were collected. For surface antigens, the cells were incubated with primary antibodies for 30 minutes with gentle shaking, washed twice with 3% (v/v) FBS/PBS, and if necessary, incubated with the secondary antibodies for 30 minutes followed by two more rinses with 3% (v/v) FBS/PBS. For intracellular antigens, the Flow Cytometry Fixation & Permeabilization Buffer Kit I, Flow Cytometry Permeabilization/Wash Buffer I and Flow Cytometry Staining Buffer kits (R&D). The cells were analyzed using a BD FACS Aria II flow cytometer (BD Biosciences). Unstained controls were used for gating conjugated antibodies and secondary only controls were used for gating unconjugated antibodies. The list of antibodies used for FACS is shown in Supplementary Table S2.

**Immunofluorescence staining.** The cells grown on cover slips (2D culture) were fixed with 4% (v/v) paraformaldehyde (PFA) in PBS (R&D Systems) for 15 minutes at RT. After washing three times with PBS, the cells were immersed in 0.5% (v/v) Triton X-100 in PBS for 10 minutes at RT, followed by incubation with blocking solution consisting of 5% (v/v) normal donkey serum (Millipore), 3% (w/v) BSA (Sigma-Aldrich) and 0.5% (v/v) Triton X-100 in PBS for 60 minutes at RT. The cells were then incubated in the primary antibody solution for 60–120 minutes at RT or overnight at 4 degrees Celsius, followed by washing three times with 0.1% (v/v) tween in HBSS. The cells were incubated in the secondary antibody solution for 60 minutes at RT and washed three times with 0.1% (v/v) tween in HBSS. All primary and secondary antibodies used in the present study were diluted in



blocking solution as indicated in Supplementary Table S2. Nuclei were counterstained with Hoechst-33342 or DAPI (Thermo Fisher) at 1:1,000 for 5 minutes. Controls were stained with secondary antibody only. For the 3D cultures, after fixation with PFA for one hour, the 3D matrigel disc in the transwell containing the organoids was embedded in paraffin, sectioned at 5  $\mu\text{m}$  thickness, dewaxed, rehydrated then stained with hematoxylin and eosin (H&E) to visualize the spheroids under a light microscope. For immunofluorescence staining of the 3D cultures, antigen retrieval on the fixed sections was performed with 10 mM sodium citrate, pH 6.0. The sections were then blocked and stained with antibodies as described above. Images were collected on a Leica confocal SPE microscope.

**Western blotting.** Cells were harvested by washing cell three times with PBS (Cellgro, 21-031-CV), dislodged using cell scrapers into 1x RIPA buffer (Thermo, 89901) with 1x protease and phosphatase inhibitor cocktail (“Halt” Thermo Scientific Cat, 1861280). Samples were subjected to BCA assay (Pierce, 23228, 23224) for protein quantitation. 25  $\mu\text{l}$  of 4x LDS loading dye (Life Tech, NP0007) and 10  $\mu\text{l}$  of DTT (Acros, 32719) are added to 65  $\mu\text{l}$  of sample. Typically, 6–10  $\mu\text{g}$  of protein were loaded per well of a 4–12% Bolt gradient gel (Life Tech, NW04127BOX). Gels were run at 160 V for 70 minutes and transferred at 30 V for 80 minutes using the BOLT Western Blotting system by Life Technologies (Running Buffer: NP0002; Transfer Buffer: NP0006-01). Blots were blocked in TBS (Tirzma Acid (Sigma, T3253) plus Trizma base (Fisher, BP152-1), with 5% milk (Apex, 20–241)) for 1 hour or overnight at 4 C. Blots were then incubated with primary antibody at 1:500 in TBST (TBS with Tween 20 (Acros, 23336–2500)) and 5% BSA (Fisher, BP1600-100) either overnight at 4 °C or room temperature for 2 hours. Antibodies used: Rabbit anti-SPB (7 Hills, 48604) and Rabbit anti-SPC (LS Biosciences, LS-B9161). Blots were then washed with TBST and incubated with appropriate secondaries: Goat anti-Rabbit-680 (Life Tech, A21109), Goat anti-Mouse-800 (Li-Cor, 926-32210), at a 1:4,000 dilution. Blots were then washed with TBST three times for 10 mins and then imaged on the Li-Cor Odyssey Imager. Lanes were quantified using the Li-Cor Image Studio software version 2.0.

**RNA isolation, cDNA preparation and real-time PCR.** Total RNA was isolated using the Arcturus PicoPure RNA isolation Kit (Invitrogen) according to the manufacturer’s manual. First-strand cDNA was synthesized from 80 ng of total RNA using the SuperScript IV First-Strand Synthesis System (Invitrogen). The cDNA samples were amplified using 2X SYBRSelect Master Mix (Invitrogen) with ABI7300 Real-Time PCR System (Life Technologies). All reactions were started at a cycle of 50 °C for 2 minutes, 95 °C for 10 minutes, followed by 40 cycles of 95 °C for 15 seconds, 60 °C for 1 minute. The PCR reactions were performed in triplicate for each sample. The level of expression of each gene was calibrated to that of the housekeeping gene,  $\beta$ -actin (ACTB), and compared to the level of the expression of each gene in the fetal human lungs. All primer sets are shown in Supplementary Table S1.

**Electron microscopy.** Samples were fixed in a solution containing 4% (v/v) formaldehyde and 1% (v/v) glutaraldehyde in 0.1 M phosphate buffer (pH 7.4) and then post fixed in 1% (w/v) osmium tetroxide. The specimens were then dehydrated in a graded series of acetone from 50% to 100% and subsequently infiltrated and embedded in Epon-Araldite epoxy resin. Ultrathin sections were cut with a diamond knife on the Reichert Ultracut E (Leica Inc). Sections were stained with uranyl acetate (2–3%) and lead citrate (0.1–0.4%) before being examined in a JEM-1011 (JEOL USA Corp) microscope.

**Measurement of phosphatidylcholines.** Medium (500  $\mu\text{L}$ ) conditioned by the of organoids was transferred to siliconized glass tubes containing 500  $\mu\text{L}$  ultra-pure water and spiked with an internal standard (1 mg PC-17:0/17:0). After addition of 2 mL of methanol/chloroform (1:1) samples were vortexed, centrifuged and the chloroform layer collected. After drying, samples were reconstituted in 100  $\mu\text{L}$  ethanol that was acidified with 0.2 mL of formic acid and transferred to siliconized mini vials for analysis. Liquid chromatography tandem mass spectrometry was performed on an Agilent 1290 Series binary pump (Agilent Technologies Inc., Santa Clara, CA, USA) coupled to an API5500 triple-quadrupole mass spectrometer (SCIEX, Concord, ON, Canada). Prior to analysis, Multiple Reaction Monitoring (MRM) mass transition parameters were optimized by infusion of pure PC standards (5  $\mu\text{l}/\text{min}$  of 1  $\mu\text{g}/\text{mL}$ ). MS analysis was performed in positive electrospray ionization mode. For quantitative analysis a separate standard curve was generated for each analyte measured using MRM area ratios (Analyte Peak Area/IS Peak Area). Results were then calculated by plotting the sample area ratios against their respective analyte standard curve.

**Statistical analysis.** Statistical analysis was performed using Student’s t-test or 2-way Anova for gene expression and Western blot pixel analysis (SPSS 12.0 software; SPSS) with significance set at  $P < 0.05$ . All data are expressed as mean  $\pm$  standard error of the mean (s.e.m.) of three or more experimental replicates. For immunofluorescence efficiency calculation, cells were defined by the presence of a nucleus (DAPI) in tissue sections. Cell marker positive cells were calculated as a percentage of total DAPI cells counted per field and analyzed by Student’s t test using SPSS 12.0 software. Full organoids were counted at 10x magnification to determine efficiency of specific protein expression. We performed 3 technical and at least 3 biological repeats of organoids from each cell line and experiment.

### Data Availability

The authors declare that the data supporting the findings of this study are available within the paper and its Supplementary Information Files.



## References

- Kimura, S. *et al.* The T/ebp null mouse: thyroid-specific enhancer-binding protein is essential for the organogenesis of the thyroid, lung, ventral forebrain, and pituitary. *Genes & development* **10**, 60–69 (1996).
- Xu, Y. *et al.* Transcriptional programs controlling perinatal lung maturation. *PLoS one* **7**, e37046, <https://doi.org/10.1371/journal.pone.0037046> (2012).
- Whitsett, J. A., Wert, S. E. & Weaver, T. E. Diseases of pulmonary surfactant homeostasis. *Annual review of pathology* **10**, 371–393, <https://doi.org/10.1146/annurev-pathol-012513-104644> (2015).
- Aurora, M. & Spence, J. R. hPSC-derived lung and intestinal organoids as models of human fetal tissue. *Developmental biology* **420**, 230–238, <https://doi.org/10.1016/j.ydbio.2016.06.006> (2016).
- Dye, B. R., Miller, A. J. & Spence, J. R. How to Grow a Lung: Applying Principles of Developmental Biology to Generate Lung Lineages from Human Pluripotent Stem Cells. *Current pathobiology reports* **4**, 47–57, <https://doi.org/10.1007/s40139-016-0102-x> (2016).
- Fatehullah, A., Tan, S. H. & Barker, N. Organoids as an *in vitro* model of human development and disease. *Nature cell biology* **18**, 246–254, <https://doi.org/10.1038/ncb3312> (2016).
- Nadkarni, R. R., Abed, S. & Draper, J. S. Organoids as a model system for studying human lung development and disease. *Biochemical and biophysical research communications* **473**, 675–682, <https://doi.org/10.1016/j.bbrc.2015.12.091> (2016).
- Hawkins, F. & Kotton, D. N. Embryonic and induced pluripotent stem cells for lung regeneration. *Annals of the American Thoracic Society* **12**(Suppl 1), S50–53, <https://doi.org/10.1513/AnnalsATS.201410-457MG> (2015).
- Chen, Y. W. *et al.* A three-dimensional model of human lung development and disease from pluripotent stem cells. *Nature cell biology* **19**, 542–549, <https://doi.org/10.1038/ncb3510> (2017).
- Miller, A. J. *et al.* In Vitro Induction and In Vivo Engraftment of Lung Bud Tip Progenitor Cells Derived from Human Pluripotent Stem Cells. *Stem cell reports* **10**, 101–119, <https://doi.org/10.1016/j.stemcr.2017.11.012> (2018).
- Dye, B. R. *et al.* In vitro generation of human pluripotent stem cell derived lung organoids. *eLife* **4**, <https://doi.org/10.7554/eLife.05098> (2015).
- Jacob, A. *et al.* Differentiation of Human Pluripotent Stem Cells into Functional Lung Alveolar Epithelial Cells. *Cell stem cell* **21**, 472–488.e410, <https://doi.org/10.1016/j.stem.2017.08.014> (2017).
- Yamamoto, Y. *et al.* Long-term expansion of alveolar stem cells derived from human iPS cells in organoids. *Nature methods* **14**, 1097–1106, <https://doi.org/10.1038/nmeth.4448> (2017).
- Hamvas, A. Surfactant protein B deficiency: insights into inherited disorders of lung cell metabolism. *Current problems in pediatrics* **27**, 325–345 (1997).
- Nogee, L. M., de Mello, D. E., Dehner, L. P. & Colten, H. R. Brief report: deficiency of pulmonary surfactant protein B in congenital alveolar proteinosis. *The New England journal of medicine* **328**, 406–410, <https://doi.org/10.1056/nejm199302113280606> (1993).
- Nogee, L. M. *et al.* A mutation in the surfactant protein B gene responsible for fatal neonatal respiratory disease in multiple kindreds. *The Journal of clinical investigation* **93**, 1860–1863, <https://doi.org/10.1172/jci117173> (1994).
- Nogee, L. M., Wert, S. E., Proffitt, S. A., Hull, W. M. & Whitsett, J. A. Allelic heterogeneity in hereditary surfactant protein B (SP-B) deficiency. *American journal of respiratory and critical care medicine* **161**, 973–981, <https://doi.org/10.1164/ajrccm.161.3.9903153> (2000).
- Nogee, L. M. In *Fetal and Neonatal Physiology* (Fifth Edition) (eds Polin, R. A. *et al.*) 843–854.e842 (Elsevier, 2017).
- Alapati, D. & Morrissey, E. E. G. Editing and Genetic Lung Disease. Basic Research Meets Therapeutic Application. *American journal of respiratory cell and molecular biology* **56**, 283–290, <https://doi.org/10.1165/rcmb.2016-0301PS> (2017).
- Mahiny, A. J. *et al.* In vivo genome editing using nuclease-encoding mRNA corrects SP-B deficiency. *Nat Biotech* **33**, 584–586, <https://doi.org/10.1038/nbt.3241> (2015).
- Barnett, R. C. *et al.* Featured Article: Electroporation-mediated gene delivery of surfactant protein B (SP-B) restores expression and improves survival in mouse model of SP-B deficiency. *Experimental biology and medicine* (Maywood, N.J.) **242**, 1345–1354, <https://doi.org/10.1177/1535370217713000> (2017).
- Castellani, S. & Conese, M. Lentiviral vectors and cystic fibrosis gene therapy. *Viruses* **2**, 395–412, <https://doi.org/10.3390/v2020395> (2010).
- Mamcarz, E. *et al.* Lentiviral Gene Therapy Combined with Low-Dose Busulfan in Infants with SCID-X1. *The New England journal of medicine* **380**, 1525–1534, <https://doi.org/10.1056/NEJMoa1815408> (2019).
- Gui, L., Qian, H., Rocco, K. A., Grecu, L. & Niklason, L. E. Efficient intratracheal delivery of airway epithelial cells in mice and pigs. *American journal of physiology. Lung cellular and molecular physiology* **308**, L221–228, <https://doi.org/10.1152/ajplung.00147.2014> (2015).
- Choi, I. Y., Lim, H. & Lee, G. Efficient generation human induced pluripotent stem cells from human somatic cells with Sendai-virus. *Journal of visualized experiments: JoVE*. <https://doi.org/10.3791/51406> (2014).
- Hamvas, A. *et al.* Population-Based Screening for Rare Mutations: High-Throughput DNA Extraction and Molecular Amplification from Guthrie Cards. *Pediatric research* **50**, 666, <https://doi.org/10.1203/00006450-200111000-00021> (2001).
- Hawkins, F. *et al.* Prospective isolation of NKX2-1-expressing human lung progenitors derived from pluripotent stem cells. *The Journal of clinical investigation* **127**, 2277–2294, <https://doi.org/10.1172/JCI89950> (2017).
- Rankin, S. A. *et al.* Timing is everything: Reiterative Wnt, BMP and RA signaling regulate developmental competence during endoderm organogenesis. *Developmental biology* **434**, 121–132, <https://doi.org/10.1016/j.ydbio.2017.11.018> (2018).
- Kubo, A. *et al.* Development of definitive endoderm from embryonic stem cells in culture. *Development (Cambridge, England)* **131**, 1651–1662, <https://doi.org/10.1242/dev.01044> (2004).
- Green, M. D. *et al.* Generation of anterior foregut endoderm from human embryonic and induced pluripotent stem cells. *Nature biotechnology* **29**, 267–272, <https://doi.org/10.1038/nbt.1788> (2011).
- Rankin, S. A. *et al.* A Retinoic Acid-Hedgehog Cascade Coordinates Mesoderm-Inducing Signals and Endoderm Competence during Lung Specification. *Cell reports* **16**, 66–78, <https://doi.org/10.1016/j.celrep.2016.05.060> (2016).
- Gotoh, S. *et al.* Generation of alveolar epithelial spheroids via isolated progenitor cells from human pluripotent stem cells. *Stem cell reports* **3**, 394–403, <https://doi.org/10.1016/j.stemcr.2014.07.005> (2014).
- Huang, S. X. *et al.* Efficient generation of lung and airway epithelial cells from human pluripotent stem cells. *Nature biotechnology* **32**, 84–91, <https://doi.org/10.1038/nbt.2754> (2014).
- Morrissey, E. E. & Rustgi, A. K. The Lung and Esophagus: Developmental and Regenerative Overlap. *Trends in cell biology* **28**, 738–748, <https://doi.org/10.1016/j.tcb.2018.04.007> (2018).
- Danopoulos, S., Bellusci, S., Warburton, D. & Alam, D. A. Identification of a SOX2/SOX9 Double Positive Cells Progenitor Cell Population Required for Branching Morphogenesis in Human Lung. *The FASEB Journal* **31**, 872.871–872.871, [https://doi.org/10.1096/fasebj.31.1\\_supplement.872.1](https://doi.org/10.1096/fasebj.31.1_supplement.872.1) (2017).
- Turcatel, G. *et al.* Lung mesenchymal expression of Sox9 plays a critical role in tracheal development. *BMC biology* **11**, 117–117, <https://doi.org/10.1186/1741-7007-11-117> (2013).
- Cardoso, W. V. & Williams, M. C. Basic mechanisms of lung development: Eighth Woods Hole Conference on Lung Cell Biology 2000. *American journal of respiratory cell and molecular biology* **25**, 137–140, <https://doi.org/10.1165/ajrcmb.25.2.4537> (2001).
- Desai, T. J. & Cardoso, W. V. Growth factors in lung development and disease: friends or foe? *Respiratory Research* **3**, 2, <https://doi.org/10.1186/rr169> (2001).

39. Joshi, S. & Kotecha, S. Lung growth and development. *Early Human Development* **83**, 789–794, <https://doi.org/10.1016/j.earlhumdev.2007.09.007> (2007).
40. Wade, K. C. *et al.* Gene induction during differentiation of human pulmonary type II cells *in vitro*. *American journal of respiratory cell and molecular biology* **34**, 727–737, <https://doi.org/10.1165/rcmb.2004-0389OC> (2006).
41. Hegab, A. E. *et al.* Mimicking the niche of lung epithelial stem cells and characterization of several effectors of their *in vitro* behavior. *Stem cell research* **15**, 109–121, <https://doi.org/10.1016/j.scr.2015.05.005> (2015).
42. Herriges, J. C. *et al.* FGF-Regulated ETV Transcription Factors Control FGF-SHH Feedback Loop in Lung Branching. *Developmental cell* **35**, 322–332, <https://doi.org/10.1016/j.devcel.2015.10.006> (2015).
43. de Carvalho, A. *et al.* Glycogen synthase kinase 3 induces multilineage maturation of human pluripotent stem cell-derived lung progenitors in 3D culture. *Development (Cambridge, England)* **146**, <https://doi.org/10.1242/dev.171652> (2019).
44. Malpel, S., Mendelsohn, C. & Cardoso, W. V. Regulation of retinoic acid signaling during lung morphogenesis. *Development (Cambridge, England)* **127**, 3057–3067 (2000).
45. Voelkel, N. F., Vandivier, R. W. & Tuder, R. M. Vascular endothelial growth factor in the lung. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **290**, L209–L221, <https://doi.org/10.1152/ajplung.00185.2005> (2006).
46. Gonzales, L. W., Guttentag, S. H., Wade, K. C., Postle, A. D. & Ballard, P. L. Differentiation of human pulmonary type II cells *in vitro* by glucocorticoid plus cAMP. *American journal of physiology. Lung cellular and molecular physiology* **283**, L940–951, <https://doi.org/10.1152/ajplung.00127.2002> (2002).
47. Nikolic, M. Z. *et al.* 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> (2017).
48. Rawlins, E. L., Clark, C. P., Xue, Y. & Hogan, B. L. The Id2+ distal tip lung epithelium contains individual multipotent embryonic progenitor cells. *Development (Cambridge, England)* **136**, 3741–3745, <https://doi.org/10.1242/dev.037317> (2009).
49. Havrda, M. *et al.* Id2 Mediates Oligodendrocyte Precursor Cell Maturation Arrest and Is Tumorigenic in a PDGF-Rich Microenvironment. Vol. 74 (2014).
50. Liebler, J. M. *et al.* Combinations of differentiation markers distinguish subpopulations of alveolar epithelial cells in adult lung. *American journal of physiology. Lung cellular and molecular physiology* **310**, L114–L120, <https://doi.org/10.1152/ajplung.00337.2015> (2016).
51. Guttentag, S. H., Beers, M. F., Bieler, B. M. & Ballard, P. L. Surfactant protein B processing in human fetal lung. *The American journal of physiology* **275**, L559–566 (1998).
52. deMello, D. E. *et al.* Ultrastructure of lung in surfactant protein B deficiency. *American journal of respiratory cell and molecular biology* **11**, 230–239, <https://doi.org/10.1165/ajrcmb.11.2.8049084> (1994).
53. Edwards, V., Cutz, E., Viero, S., Moore, A. M. & Nogee, L. Ultrastructure of lamellar bodies in congenital surfactant deficiency. *Ultrastructural pathology* **29**, 503–509, <https://doi.org/10.1080/01913120500323480> (2005).
54. Cruz, A. *et al.* Microstructure and dynamic surface properties of surfactant protein SP-B/dipalmitoylphosphatidylcholine interfacial films spread from lipid-protein bilayers. Vol. 29 (2000).
55. Farman, S. L. *et al.* Gene transfer of CFTR to airway epithelia: low levels of expression are sufficient to correct Cl<sup>-</sup> transport and overexpression can generate basolateral CFTR. *American journal of physiology. Lung cellular and molecular physiology* **289**, L1123–L1130, <https://doi.org/10.1152/ajplung.00049.2005> (2005).
56. Dorin, J. R. *et al.* A demonstration using mouse models that successful gene therapy for cystic fibrosis requires only partial gene correction. *Gene therapy* **3**, 797–801 (1996).
57. Johnson, L. G., Boyles, S. E., Wilson, J. & Boucher, R. C. Normalization of raised sodium absorption and raised calcium-mediated chloride secretion by adenovirus-mediated expression of cystic fibrosis transmembrane conductance regulator in primary human cystic fibrosis airway epithelial cells. *The Journal of clinical investigation* **95**, 1377–1382, <https://doi.org/10.1172/jci117789> (1995).
58. Johnson, L. G. *et al.* Efficiency of gene transfer for restoration of normal airway epithelial function in cystic fibrosis. *Nature genetics* **2**, 21–25, <https://doi.org/10.1038/ng0992-21> (1992).
59. Dunbar, A. E. 3rd *et al.* Prolonged survival in hereditary surfactant protein B (SP-B) deficiency associated with a novel splicing mutation. *Pediatric research* **48**, 275–282, <https://doi.org/10.1203/00006450-200009000-00003> (2000).
60. Melton, K. R. *et al.* SP-B deficiency causes respiratory failure in adult mice. *American journal of physiology. Lung cellular and molecular physiology* **285**, L543–549, <https://doi.org/10.1152/ajplung.00011.2003> (2003).
61. Garcia, O. *et al.* Cell-based therapies for lung disease. *British medical bulletin* **101**, 147–161, <https://doi.org/10.1093/bmb/ldr051> (2012).
62. Danopoulos, S. *et al.* Human lung branching morphogenesis is orchestrated by the spatiotemporal distribution of ACTA2, SOX2, and SOX9. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **314**, L144–L149, <https://doi.org/10.1152/ajplung.00379.2017> (2018).
63. Chen, Y. *et al.* Long-term Engraftment Promotes Differentiation of Alveolar Epithelial Cells from Human Embryonic Stem Cells Derived Lung Organoids. *Stem Cells and Development* **0**, null, <https://doi.org/10.1089/scd.2018.0042>.
64. Yang, L. *et al.* Surfactant protein B propeptide contains a saposin-like protein domain with antimicrobial activity at low pH. *Journal of immunology (Baltimore, Md.: 1950)* **184**, 975–983, <https://doi.org/10.4049/jimmunol.0900650> (2010).
65. Xu, Y. *et al.* Differential susceptibility of human sp-b genetic variants on lung injury caused by bacterial pneumonia and the effect of a chemically modified curcumin. *Shock (Augusta, Ga.)* **45**, 375–384, <https://doi.org/10.1097/shk.0000000000000535> (2016).

## Acknowledgements

The authors would like to thank Lance Prince and Xin Sun (University of California, San Diego) for review of the manuscript and Ping Yang for performing the Sfu1 digestion. The Sanford Burnham Prebys Institute Histology core embedded and sectioned the organoids. Retrogen (San Diego) performed the genomic sequencing. The UCSD Electron Microscopy core performed the electron microscopy. This work was supported by grants from the Canadian Institutes of Health Research (FDN-143309 to M.P.), Canadian Foundation of Innovation (CFI-12156 to M.P.) and the UCSD Senate grant (RS282R-LEIBEL to S.L.L.). Martin Post is the holder of a Canadian Research Chair in Fetal, Neonatal, and Maternal Health.

## Author Contributions

S.L.L. and M.P. designed the study and wrote the manuscript. S.L.L., A.W., I.T., J.W., D.L., N.N. and S.S. performed experiments and analyzed data. E.S. provided conceptual advice and edited the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-49696-8>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019