Efficient generation and transcriptomic profiling of human iPSC-derived pulmonary neuroendocrine cells

Pooja Hor
Vasu Punj
Ben A Calvert
Alessandra Castaldi
Alyssa J Miller

See next page for additional authors

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs
Authors
Pooja Hor, Vasu Punj, Ben A Calvert, Alessandra Castaldi, Alyssa J Miller, Gianni Carraro, Barry R Stripp, Steven L Brody, Jason R Spence, Justin K Ichida, Amy L Ryan Firth, and Zea Borok
Efficient Generation and Transcriptomic Profiling of Human iPSC-Derived Pulmonary Neuroendocrine Cells

iPSCs Day 0
Submerged culture

Lung progenitors Day 17
Switched to air-liquid interface

Mixed airway cells Day 90
Continuous Notch inhibition

Secreted products

HIGHLIGHTS
PNECs can be efficiently generated from induced pluripotent stem cells (iPSCs)
Induced PNECs express key PNEC markers and express and secrete all major neuropeptides
Induced PNECs resemble the transcriptomic profile of human primary fetal PNECs
Basal cells are able to differentiate into PNECs

DATA AND CODE AVAILABILITY
GSE146990

Pooja Hor, Vasu Punj, Ben A. Calvert, ..., Justin K. Ichida, Amy L. Ryan (Firth), Zea Borok
ichida@usc.edu (J.K.I.)
amy.firth@med.usc.edu (A.L.R.)
zborok@med.usc.edu (Z.B.)

HIGHLIGHTS
PNECs can be efficiently generated from induced pluripotent stem cells (iPSCs)
Induced PNECs express key PNEC markers and express and secrete all major neuropeptides
Induced PNECs resemble the transcriptomic profile of human primary fetal PNECs
Basal cells are able to differentiate into PNECs

DATA AND CODE AVAILABILITY
GSE146990

Hor et al., iScience 23, 101083
May 22, 2020 © 2020 The Authors.
https://doi.org/10.1016/j.isci.2020.101083
Efficient Generation and Transcriptomic Profiling of Human iPSC-Derived Pulmonary Neuroendocrine Cells

Pooja Hor,1,2 Vasu Punj,1 Ben A. Calvert,1 Alessandra Castaldi,1 Alyssa J. Miller,3 Gianni Carraro,4 Barry R. Stripp,4 Steven L. Brody,5 Jason R. Spence,3,6,7,8 Justin K. Ichida,2,* Amy L. Ryan (Firth),1,2,12,* and Zea Borok1,9,10,*

SUMMARY
Expansion of pulmonary neuroendocrine cells (PNECs) is a pathological feature of many human lung diseases. Human PNECs are inherently difficult to study due to their rarity (<1% of total lung cells) and a lack of established protocols for their isolation. We used induced pluripotent stem cells (iPSCs) to generate induced PNECs (iPNECs), which express core PNEC markers, including ROBO receptors, and secrete major neuropeptides, recapitulating known functions of primary PNECs. Furthermore, we demonstrate that differentiation efficiency is increased in the presence of an air-liquid interface and inhibition of Notch signaling. Single-cell RNA sequencing (scRNA-seq) revealed a PNEC-associated gene expression profile that is concordant between iPNECs and human fetal PNECs. In addition, pseudotime analysis of scRNA-seq results suggests a basal cell origin of human iPNECs. In conclusion, our model has the potential to provide an unlimited source of human iPNECs to explore PNEC pathophysiology associated with several lung diseases.

INTRODUCTION
Pulmonary neuroendocrine cells (PNECs) are endogenous chemosensory cells that form around gestational week 8 in humans, making them the first differentiated cell type to appear in fetal lung (Cutz et al., 1985). Innervated clusters of PNECs in the human airway epithelium are referred to as "neuro-epithelial bodies or NEBs" (Lauweryns and Peuskins, 1972; Lauweryns and Van Ranst, 1988). PNECs are multifunctional cells with known roles in oxygen sensing, redistribution and regulation of blood flow especially during hypoxia, modulation of immune responses, and stimulation of intraepithelial vagosensory nerve fibers (Linnoila, 2006; Sunday, 1996). In addition, PNECs have been reported to serve as a stem cell niche during severe lung injury, with the potential to give rise to club and ciliated cells (Ouadah et al., 2019; Song et al., 2012). PNEC ontogeny, however, is not entirely clear. A recent lineage tracing study in mice showed that a small population of cells expressing the basal cell marker, tumor protein 63+ (TP63+), and NOTCH1+ during severe lung injury, with the potential to give rise to club and ciliated cells (Ouadah et al., 2019; Song et al., 2012). PNEC ontogeny, however, is not entirely clear. A recent lineage tracing study in mice showed that a small population of cells expressing the basal cell marker, tumor protein 63+ (TP63+), and labeled at embryonic day (E) 9.5, gave rise to calcitonin gene-related polypeptide alpha (CGRP)-expressing PNECs, that can be detected at E18.5 by immunostaining (Yang et al., 2018). However, whether basal cells have the capacity to give rise to PNECs in the human fetal or adult lung remains to be determined.

A spectrum of lung disorders including neuroendocrine hyperplasia of infancy (NEHI), chronic obstructive pulmonary disease (COPD), cystic fibrosis, asthma, bronchopulmonary dysplasia (BPD), and pulmonary hypertension are associated with expansion of PNECs in the airway epithelium (Cutz, 2015; Cutz et al., 2007; Gu et al., 2014; Johnson et al., 1985; Johnson and Georgieff, 1989; Pan et al., 2006; Schindler et al., 1995). The pathophysiological consequences of increased PNEC abundance in these diseases is not clear. As PNECs comprise <1% of all respiratory cell types, primary PNECs are inherently difficult to isolate in sufficient quantities for detailed characterization. We, therefore, addressed an unmet need to develop a model system for efficiently generating human PNECs to study their physiological and pathological functions using induced pluripotent stem cells (iPSCs).

Within the lung milieu, activation of pro-neural transcription factor ASCL1 (Achaete-Scute Family BHLH Transcription Factor 1) is required for cells to form the pulmonary neuroendocrine lineage (Linnoila, 2006; Sunday, 1996).
2006). The Notch-HES1/HEY1 (Hairy/Enhancer-Of-Split related BHLH transcription factor family) pathway regulates the non-neuroendocrine fate of lung endoderm by repressing pro-neural genes like ASCL1 (Henke et al., 2009; Nelson et al., 2009). Recently, it has been shown that inhibition of Notch can increase PNEC production (Chen et al., 2019). These studies, however, focused specifically on modeling small cell lung carcinoma (SCLC) using human embryonic stem cell-derived PNECs, or on generating proximal airway epithelial spheroids from human pluripotent cells (Chen et al., 2019; Konishi et al., 2016). In-depth characterization of iPNECs or comparison at the transcriptional level with primary PNECs was not performed. In this article, we report the differentiation of iPSCs to human iPNECs with a gene expression profile similar to that of primary fetal PNECs that could be used in future studies of pathophysiological changes in diseases such as NEHI or BPD.

RESULTS

Directed Differentiation of iPSCs to iPNECs

We adapted our previously published differentiation protocol to create airway epithelium from human iPSCs, recapitulating the key stages of embryonic lung development (Firth et al., 2014). iPSCs were differentiated in culture without sorting, resulting in a mixed population of mesoderm and epithelium comprising the proximal airways. To validate the presence of PNECs in our directed differentiation, cultures were stained for a panel of genes known to be expressed in primary human PNECs (Linnoila, 2006; Sunday, 1996), including synaptophysin (SYP), chromogranin A (CHGA), PGP9.5 (expressed by the gene UCHL1), and neuron-specific enolase (ENO2). After 31 days of differentiation, including culture at an air-liquid interface (ALI) from Day 17, cells expressing these PNEC markers were observed in cultures generated from three independent iPSC lines (Figures 1A and S1A). Expression of SYP, CHGA, PGP9.5, and ENO2 was quantified as a percentage of the mean fluorescence intensity (MFI) of the respective marker normalized to the expression of nuclear marker 4’,6-diamidino-2-phenylindole (DAPI). The distribution of these markers ranged from 31.64% ± 3.01% MFI for SYP to 61.20% ± 4.78% MFI for ENO2 (Figure 1B). Orthogonal projection of the differentiated cultures revealed that iPNECs are present in closer proximity to the apical-air interface (Figure S1B). Analysis of key PNEC markers, including ENO2 and SYP, throughout the differentiation time course from iPSC indicated that ENO2 protein was robustly expressed from Day 13 (Figure S2A), whereas SYP expression was increased gradually after exposure at the ALI from Day 17 onward (Figure S2B). Expression of these PNEC markers was additionally validated in explanted human adult lung tissue sections (Figure 1C). Cells expressing SYP, CHGA, and PGP9.5 were identified within the epithelium of the distal airways. In support of observations in our iPNEC cultures, ENO2 expression was also more widespread than the other PNEC markers in the human lung sections (Figure 1C), consistent with ENO2 being expressed in other lung cells in addition to PNECs. Indeed, our data indicated that KRT5+ basal cells co-express ENO2 in adult human lung tissues (Figure S2C). This is also supported by data deposited in the LungMAP database (www.lungmap.net). Gastrin-releasing protein (GRP), also known as bombesin, is another key marker of PNECs whose expression peaks during human fetal lung development around week 20–23 (Linnoila, 2006; Sunday, 1996; Sunday et al., 1991). GRP mRNA expression peaked during specification of the primordial lung progenitor cells from ventralized anterior foregut endoderm, Days 13–17 of our iPSC differentiation protocol (Figure S2D), and then decreased.

Induction of PNECs Is Augmented by Air-Liquid Interface Culture during Directed Differentiation

Based on our observations that SYP expression increased after exposure of the differentiation cultures to an ALI, we evaluated whether the air interface was an important factor in PNEC induction. We continued the culture of primordial lung progenitors at Day 17 of differentiation (Firth et al., 2014), in both submerged and ALI conditions. Owing to consistency in staining across experimental replicates, the relative MFI of SYP-DAPI was compared (Figure 1B). After 14 additional days of differentiation through to experimental Day 31, SYP+ iPNECs were only detected in the presence of an ALI (Figure 1D, bottom panel). No SYP-expressing cells were observed in continually submerged cultures (Figure 1D, top panel), indicating that air exposure provides a microenvironment suited to the continued development of PNECs.

Characterization of Human iPNECs

To comprehensively characterize and validate iPNECs, we co-labeled the differentiating cultures with NKX2.1, a primordial lung endoderm transcription factor; EPCAM, a pan-epithelial marker; and vimentin, VIM, a mesenchymal marker, in combination with the PNEC marker SYP. SYP+ iPNECs co-express NKX2.1 and VIM, with high expression levels, indicating that these cells are consistent with PNECs. This is supported by the observation that NKX2.1+ cells co-express ENO2 in adult human lung tissues (Figure S2C). This is also supported by data deposited in the LungMAP database (www.lungmap.net). Gastrin-releasing protein (GRP), also known as bombesin, is another key marker of PNECs whose expression peaks during human fetal lung development around week 20–23 (Linnoila, 2006; Sunday, 1996; Sunday et al., 1991). GRP mRNA expression peaked during specification of the primordial lung progenitor cells from ventralized anterior foregut endoderm, Days 13–17 of our iPSC differentiation protocol (Figure S2D), and then decreased.
supporting a lung endodermal origin, but do not express VIM or EPCAM (Figures 2A and 2B). Human lung tissue sections were used to validate this co-staining pattern: primary PNECs within these sections also co-express NKX2.1, but not EPCAM or VIM, supporting our iPNEC data (Figure 2B). NKX2.1 also identifies neuronal cells developing in the forebrain (Germain et al., 2013; Goulburn et al., 2011); therefore, to exclude the possibility that our differentiation was creating forebrain neurons we stained for additional neuronal-specific marker, microtubule-associated protein 2 (MAP2), at Days 31 and 60 of differentiation. No MAP2 expression was evident at either of these time points (data not shown).
NOTCH Inhibition Augments iPNEC Specification and Phenotypic Complexity

A transcriptional program regulated by transcription factor ASCL1 is required for activating the neuroendocrine lineage in developing lung to generate PNECs (Linnoila, 2006). ASCL1 expression is known to be repressed by NOTCH signaling, which supports expansion and differentiation of lung basal and secretory cells, respectively. During iPSC differentiation, ASCL1 mRNA is detectable from Day 10 (Figure S2E), preceding the appearance of SYP+ cells from Day 13 (Figure S2B). We also observe that expression of HEY1/HES1, downstream targets of the NOTCH pathway, peaks between Days 10 and 17 of differentiation (Figures S2F and S2G), whereas ASCL1 peaks at Day 31 (Figure S2E). This suggests that in our differentiation protocol, much like during development, there is an inverse correlation between activity of the Notch signaling pathway and ASCL1 expression. To evaluate the impact of Notch inhibition on neuroendocrine differentiation and expansion during our differentiation protocol, we performed a dose-response to

Figure 2. Characterization of Human iPNECs

(A) Representative immunofluorescence (IF) images of ALI cultures at Day 31 of directed differentiation: SYP+ iPNECs (red), NKX2.1, EPCAM, and vimentin (VIM) (all green). Left panels (20X) represent low-power images (scale bars, 200 μm), and right panels (40X) show magnified views of the marked areas (scale bars, 10 μm).

(B) Representative IF images of adult human distal lung sections: SYP+ iPNECs (green), NKX2.1, EPCAM, and VIM (all red). Yellow dotted lines demarcate the airways. Left panels (20X) represent low-power images, and right panels (40X) show magnified views of the marked areas (scale bars, 50 μm). Nuclei are counterstained with DAPI (blue).
γ-secretase/Notch inhibitor, 3tert-Butyl(2S)-2-[(2S)-2-[[2-(3,5-difluorophenyl) acetyl] amino] propanoyl] amino]-2-phenylacacetate (DAPT). As before, marker expression was quantified as a percentage of the MFI for the respective marker normalized to the MFI of nuclear marker DAPI. Continuous addition of 1, 10, and 20 μM DAPT to cultures from Day 17 onward resulted in a dose-dependent increase in the relative MFI of SYP at Day 31 (Figures S3A, S3A, and S3B). The effect of Notch inhibition was validated using a second Notch signaling inhibitor, dibenzazepine (DBZ), at 0.5, 2, and 5 μM. A 2-fold increase in relative MFI of SYP was observed when increasing the concentration of DBZ from 0.5 to 2 μM (Figures 3C and S3B). However, no further increase in the MFI ratio of SYP was observed with an increase to 5 μM DBZ treatment. These findings support inhibition of Notch signaling as an important mechanism for augmenting PNEC differentiation and expansion. To deduce the optimal concentrations of Notch inhibitors, DAPT and DBZ, required for iPNEC differentiation, we also co-labeled Day 31 cultures with cleaved caspase 3, an apoptotic marker to evaluate whether high concentrations of Notch inhibitors induced toxicity and cell death. Based on immunofluorescent staining, an increase in the caspase 3’ relative MFI percentage was observed at the highest concentration (20 μM DAPT and 5 μM DBZ) of the Notch inhibitors (Figures S3A and S3B), indicating that 10 μM DAPT or 2 μM DBZ were optimal for induction of iPNECs. These concentrations were used for all subsequent experiments. Extended culture of iPNECs with continuous addition of DAPT/DBZ for up to 90 days resulted in a significant increase in mRNA expression for representative PNEC genes ASCL1, DLL3, GAD1/GAD67, ENO2, ROBO2, and SYP (Figure 3D). In addition, there was a significant increase in SYP, ROBO2, and CHGA staining (Figures 3E, 3F, and 3H). ENO2 was robustly expressed by Day 31 with no further increase at Day 91 (Figure 3G). These data support a role for continuous repression of Notch signaling in augmenting iPNEC expansion and homeostasis. Interestingly, SYP, ROBO2, and ENO2 staining patterns show an intricate arborized network formation over time in culture. Using Sholl analysis (Figure S4A), we quantified this arborization as a change in the number of intersections (Figures S4C, S4E, and S4G) and branching density (Figures S4C, S4E, and S4G) of SYP+, ROBO2+, and ENO2+ iPNECs (Binley et al., 2014; Bird and Cuntz, 2019; Sholl, 1953; Stanko and Fenton, 2017; Ferreira et al., 2014). Such changes in the cellular complexity of the iPNECs likely reflect phenotypic maturation similar to that observed in the NEBs of adult lungs (Pan et al., 2004, 2006).

Transcriptomic Characterization of iPNECs

To evaluate the transcriptomic profile of iPNECs, we performed single-cell RNA sequencing (scRNA-seq) of cultures maintained with continuous treatment with 10 μM DAPT at Day 91 of iPSC differentiation (for schematic see Figure 5A). Two independent biological replicates were sequenced and analyzed. Based on the quality parameters as detailed in Transparent Methods, we selected a set of 1,705 and 1,502 cells from replicates 1 and 2, respectively, for analysis. The two replicates were corrected for any batch effect before further analysis. A subset of 1,171 most variable genes that showed a high level of concordance between the two replicates was selected for further downstream analysis (Pearson correlation coefficient, r = 0.94, Figure S5B). A flowchart describing the bioinformatics analysis pipeline for the scRNA-seq data is shown in Figure S5C. Unsupervised graph-based clustering (false discovery rate, FDR p < 0.05) and non-linear dimensionality reduction distributes the cells into six independent cell populations on Day 91 of differentiation (Figure 4A). Two of these clusters, labeled iPNEC1 and iPNEC2, representing >30% of the cells captured, were defined by a transcriptomic signature to reflect a neuroendocrine lineage determined by the expression of known PNEC genes including, but not limited to, SYP, CHGA, PGP9.5, ENO2, and ROBO2. Further in-depth analysis revealed that these two clusters are separated from the other cell clusters by differential expression (fold change, FC) of key PNEC marker genes SYP, ENO2, UCHL1, and CHGA, and Notch pathway genes, HES1 and NOTCH2 (Figure 4B). The larger cluster, iPNEC1 (yellow dots), has a higher degree of Notch repression and a higher FC of SYP, ENO2, UCHL1, and CHGA than the smaller cluster, iPNEC2 (blue dots). For all subsequent analyses the two clusters, iPNEC1 and iPNEC2 were grouped together as iPNECs. A set of 255 differentially enriched genes (FDR p < 0.05; FC ≥ 1.5) was identified using gene set enrichment analysis against a compendium of 350 different human tissues and cell types; we refer to this set of genes as iPNEC-associated genes. Figure S5C shows a flowchart describing the process of identifying 255 iPNEC-associated genes. In a genome-wide comparison, these 255 genes are significantly enriched (FDR p < 0.5; FC ≥ 1.5) in the putative iPNEC cell population compared with other populations (Figure 4C) and includes established PNEC marker genes including SYP, CHGA, and ROBO2 (Figure 4D). UCHL1, although highly enriched within the iPNEC cluster, is also expressed in the non-iPNEC population. According to the LungMAP database, UCHL1 is also expressed to some extent by lung mesenchyme (www.lungmap.net). We observed co-differentiation of mesenchymal-like clusters in our Day 91 cultures (VIM+ fibroblast-like and ACTA2+ smooth muscle cells in Figure 4A), which could contribute to UCHL1 expression within the non-iPNEC population. PNECs receive stimulation from neurotransmitters/neurosecretions innervating...
the lung and respond to such stimuli by releasing hormones (neuropeptides and amines). PNECs, therefore, have important functions in assimilating characteristics of both the nervous and endocrine systems. Accordingly, analysis of the pathways most enriched in the 255 gene set includes neuronal system development, synaptic...

Figure 3. Extended Notch Inhibition Increases iPNEC Specification and Phenotypic Complexity

(A–C) (A) Representative immunofluorescence (IF) images and quantification of SYP⁺ iPNECs (red) at Day 31 of differentiation in the presence of DAPT (1 μM and 10 μM). Quantification of SYP expression (MFI expressed as a percentage of total DAPI fluorescence) at Day 31 of differentiation in the presence of (B) DAPT (1 μM and 10 μM) (paired t test) and (C) DBZ (0.5 μM, 2 μM and 5 μM) (one-way ANOVA). Data are expressed as mean ± SEM from 3 independent experiments. **p < 0.01, ****p < 0.0001.

(D–L) (D) Quantitative RT-PCR showing PNEC marker expression after 91 days of differentiation. Data are normalized to internal control β-actin and expressed as comparisons among Days 31, 60, and 91 with respect to Day 0. N = 3 experimental replicates each with 3 technical replicates. One-way ANOVA. *p < 0.05, **p < 0.01, ****p < 0.0001. MFI expressed as percentage of DAPI fluorescence for cells expressing SYP⁺ (E), ROBO2⁺ (F), ENO2⁺ (G), and CHGA⁺ (H) at Days 31, 60, and 91 of differentiation. One-way ANOVA. **p < 0.01, ****p < 0.0001. Scale bars, 100 μm. Nuclei are counterstained with DAPI (blue). (E) Representative IF images show SYP⁺ (I), ROBO2⁺ (J), ENO2⁺ (K), and CHGA⁺ (L) cells at Days 31, 60, and 91 of differentiation (all red).

See also Figures S3 and S4.
Figure 4. Transcriptomic Characterization of iPNECs

(A) tSNE plot identifies six cell clusters in two replicates of 91-day old iPSC differentiated cells.

(B) Expression of representative genes between iPNECs1 and iPNECs2. Fold change represents differential expression between iPNEC1/2 clusters versus other cells.

(C) Box whisker plot comparing 255 iPNEC-associated genes within iPNEC and non-iPNEC populations. Key PNEC markers are shown (CHGA, SYP, UCHL1, ENO2, and NEUROD1).

(D) Violin plots showing expression of individual PNEC marker genes (ROBO2, CHGA, SYP, and UCHL1) within iPNEC and non-iPNEC populations.

(E) Signaling pathways altered within 255 iPNEC-associated genes.

(F) Scatterplot showing concordance of expression of 151 common genes (Pearson correlation coefficient, \( r = 0.75 \)) in Day 91 iPNECs and human fetal lung-derived primary PNECs pooled from gestational weeks 15, 18, and 21.

(G) Heatmap comparing the gene expression signature between iPNECs and human fetal lung-derived primary PNECs pooled from gestational weeks 15, 18, and 21.

(H) Venn diagram showing the number of genes overlapping between iPNECs and human fetal lung-derived primary PNECs and the known key PNEC markers that are common between the two samples.

See also Figures S5 and S6.
vesicle biogenesis, and bioactive peptide signaling (Figure 4E), consistent with some of the known functions of PNECs. We also observe Notch signaling as one of the most differentially altered pathways (Figure 4G), reflecting the fact that continued Notch inhibition was administered to induce a neuroendocrine fate.

Given the limited number of PNECs in human lung, performing RNA sequencing on primary PNECs is challenging. We, therefore, procured the only available scRNA-seq dataset of human fetal lung explants pooled from gestational weeks 15, 18, and 21 (EMBL-EBI ArrayExpress Database: E-MTAB-8221) (Miller et al., 2018) and compared the neuroendocrine cluster (220 cells) with the iPNEC cluster within our scRNA-seq dataset (GEO accession number: GSE146990). Comparative analysis revealed a significant correlation coefficient (r = 0.75) with a high concordance of 151 genes found to be commonly expressed between iPNECs and fetal PNECs (Figure 4F). One-way hierarchical clustering indicates an overall similar gene expression pattern between iPNECs and human primary fetal PNECs (Figure 4G) including expression of all known major PNEC genes such as SYP, CHGA, PGP9.5, ROBO2, ASCL1, NCAM1, ENO2, DLL3, and NEUROD1 (Figure 4H and Tables S1, S2, and S3), indicating considerable overlap between the transcriptomes of iPNECs and primary human fetal PNECs.

PNECs are considered neuroepithelial in nature; however, we did not observe co-staining of iPNECs or primary PNECs in human lung tissue sections stained with epithelial marker EPCAM (Figures 2A and 2B). To further clarify whether EPCAM is expressed in PNEC we further analyzed the scRNA-seq data. In fetal PNECs, Miller et al., identified a neuroendocrine cell population within a computationally extracted epithelial cluster based on high EPCAM expression (Miller et al., 2018). Detailed analysis of the human fetal lung scRNA-seq data enabled a differentiation of two subsets of fetal PNECs based on EPCAM expression: of the 220 fetal PNECs 77 were EPCAM+ and 143 EPCAM−. Analysis of our 255 PNEC-associated genes in EPCAM− fetal PNECs indicated a significant concordance of 243 genes between both datasets (read count ≥2, r = 0.6) (Figure S5D). We, therefore, demonstrate that (1) a majority of human fetal PNECs do not express EPCAM and (2) iPNECs have significant gene expression concordance with human fetal PNECs.

**iPNECs Express and Secrete Major Neuropeptides**

PNECs contain vesicles filled with neuropeptides, amines, and neurotransmitters that are secreted upon stimulation. Recently, neuropeptides proenkephalin (PENK), somatostatin (SST), and neurotransmitter synaptic vesicle protein 2 (SV2) have been shown to be secreted by primary PNECs (Branchfield et al., 2016). Differentiation cultures at days 31, 60, and 91 containing greater than 30% iPNECs (Figures 3 and 4) express PENK, SST, and SV2 at the mRNA level (Figure 5A). ENO2 secretion has been widely used as a biochemical marker to monitor the progression of PNEC-associated disorders, namely SCLC and pulmonary neuroendocrine tumors (Anastasiades et al., 1987; Kasprzak et al., 2007). Immunostaining of our cultures at Day 91 confirms that the majority of SYP+ cells co-express ENO2 (Figure S6B). We, therefore, proceeded to

---

**Figure 5. Neuropeptide and Marker Expression Profile of iPNECs**

(A) qRT-PCR showing the neuropeptide profile of cultures at Day 0, 31, 60, and 91 of differentiation. Data are normalized to internal control b-actin and expressed as comparisons among Days 31, 60, and 91 with respect to Day 0. N = 3 experimental replicates each with 3 technical replicates. One-way ANOVA. *p < 0.05, ****p < 0.0001.

(B) Enzyme-linked immunosorbent assay (ELISA) data showing secreted levels of enzyme ENO2 in the culture media. N = 4 experimental replicates. One-way ANOVA, ****p < 0.0001.

See also Figure S6.
perform an ELISA to evaluate whether SYP+ iPNECs also secrete ENO2. Media were collected from the basolateral compartment of Day 91 ALI cultures used for the scRNA-seq experiment using basal media and iPSC-conditioned media as negative controls and recombinant ENO2 as a positive control. Day 91 differentiation cultures, containing >30% iPNECs (Figures 3 and 4), expressed ENO2 at mRNA (Figures 3D and S6A) and protein levels (Figure 5B) in the absence of external stimulation, showing similarity in terms of their gene expression profile to primary PNECs as previously shown (Cutz et al., 2008). It is important to note that scRNA-seq analysis indicates that ENO2 expression is only found in the defined iPNEC populations (Figure S6A) supporting selective release of ENO2 from iPNEC in our Day 91 cultures.

Human Basal Cells Are Capable of Differentiating into PNECs

Lineage tracing studies in mice have indicated that TP63+ basal cells can give rise to PNECs (Yang et al., 2018). To determine whether human basal cells are also able to differentiate into PNECs, we performed two experiments: (1) pseudotime analysis of our human differentiation Day 91 iPNEC scRNA-seq data and (2) ALI differentiation of purified human tracheal epithelial cells (HTECs). We analyzed the scRNA-seq data of Day 91 differentiation cultures to track cell population changes across the differentiation trajectory (Figure 4). A stream plot was generated using Stream and Monocle3 for easy visualization of the different cell types in our cultures projected onto the differentiation trajectory “tree” according to their pseudotime locations and the distances from their assigned branches (Figure 6A). Based on the cell state during the reprogramming each cell type/population is assigned to a distinct “branch” of the “tree” representing a specific cell fate choice/developmental lineage that proceeds down a distinct path. Cells on the same or nearby branches are predicted to be hierarchically related. The stream plot (Figure 6A) indicates that several different cell types are present in our Day 91 differentiation cultures including basal cells, respiratory smooth muscle-like cells, fibroblast-like cells, and iPNECs. We demonstrate that the key genes expressed in the S3 – S0 branch overlap with the published human lung basal cell database (GSE24337), confirming that iBasal cells in this culture possess a transcriptomic signature resembling that of primary human basal cells (Figure 6B). We tracked six key PNEC marker genes, INS1M, UCHL1, NCAM1, SYP, NEUROD1, and ENO2 across the branches where a significant increase in their expression was detected as the cells transition from the iBasal cell state into the two iPNEC trajectories S0 – S1 or S0 – S2 (p < 0.001, Figures S7A–S7F). To further confirm the differentiation of iBasal cells to iPNECs, we used the reverse graph embedding tool in Monocole3 to reconstruct an unbiased lineage trajectory and assigned iBasal cells as the starting cell lineage (Figure 6C). Cells start at the root (State 1) and progress to one of the two alternative reprogramming outcomes, denoted by State 2 or State 3. Monocle3 trajectories support that iBasal cells differentiate into two states, State 2 and State 3, which correspond to the two different states of iPNECs, iPNEC1 and iPNEC2 respectively. Therefore, it is possible that the two populations observed in graph-based clustering in Figure 4A represent two different physiological states of iPNECs, as observed by Monocle3 (Trapnell, 2020; Trapnell et al., 2014a, 2014b). Analysis of the differentially expressed genes between State 2 (iPNEC1) and State 3 (iPNEC2) indicates a difference in the level of the PNEC marker gene expression. UCHL1, NCAM1, and INS1M are high in iPNEC1, whereas SYP and ENO2 are high in iPNEC2 (Figure 6D). Overall, the pseudotime trajectory analysis supports an iBasal cell predating iPNECs in the differentiation protocol described in this article. Purification and expansion of iPSC-derived basal cells is currently challenging. We therefore used HTECs to corroborate our in silico finding. HTECs, consisting almost exclusively of a self-renewing population of KRT5+/TP63+ basal cells (Figure S8A), were cultured at the ALI and stained for SYP. After more than 75 days of differentiation at the ALI, the distribution of SYP staining in the cultures was estimated to be 17.71% ± 2.49% based on the relative MFI percentage of SYP:DAPI (Figures 6E, 6F, and S8B). Using two independent approaches our data strongly suggest that human basal cells, both iPSC-derived and primary, are capable of giving rise to PNECs.

DISCUSSION

We have established a directed differentiation strategy for deriving iPNECs from human iPSCs that are transcriptomically similar to human fetal primary PNECs. Generation of these cells from a renewable and patient-specific cellular source overcomes limitations with regard to accessibility to sufficient primary human PNECs and enhances our capacity to model human diseases, like NEHI. Recent reports have demonstrated the generation of PNEC-like cells from human pluripotent stem cells; these studies, however, focused primarily on their sporadic presence in spheroid cultures or on their potential as lung cancer models (Chen et al., 2019; Konishi et al., 2016) (17, 18). To date, there has been no in-depth phenotypic and transcriptional characterization of these cells to validate their similarity to their primary counterparts. The iPNEC differentiation protocol described in this article is adapted from our previously established airway epithelial differentiation protocol to specifically promote differentiation toward the neuroendocrine lineage (Firth et al., 2014). Using this model we are able to generate iPNECs expressing all major PNEC
marker genes including SYP, CHGA, PGP9.5, ROBO2, and ENO2, consistent with primary human and rodent PNECs (Branchfield et al., 2016; Linnoila, 2006; Sunday, 1996). By specifically adapting the culture conditions to favor expansion of PNECs we are able to generate cultures with an increased efficiency of generating iPNECs over currently published protocols (Chen et al., 2019). Our differentiation is robust across...
multiple iPSC lines despite inherent variability between stem cells lines/clones with regard to their differentiation efficiency. Furthermore, differentiation protocols, such as those developed in this study, rely on the sequential application of growth factors and cytokines directing differentiation toward lung epithelium: at any given point cells may be captured at varied stages of differentiation and maturation, again contributing to heterogeneity. It should be noted that our iPNEC protocol does not utilize any purification steps and thus an endpoint culture composed of mesenchymal (niche) cells and epithelium is generated with the iPNECs. Isolation of iPSC-derived lung basal cells, akin to primary HTEC, and differentiation toward PNECs may reduce this variability; however, loss of putative niche cells may also impact efficiency. Such factors will be evaluated in a subsequent study (Ryan laboratory, unpublished data).

Interestingly, we found that exposure to an air interface was important for the continued development of iPNECs. This condition typically mimics the microenvironment of the postnatal and adult lung (Van Haute et al., 2009; Wong et al., 2015, 2012; Yamaya et al., 1992). This observation has not previously been reported, and although PNECs first appear in the embryonic lung when the lung is filled with amniotic fluid (Hooper and Harding, 1995; Morrissey and Hogan, 2010), it is in the adult lung, when the lung is exposed to air/oxygen, that PNECs are known to function as airway chemosensors (Garg et al., 2019; Linnoila, 2006; Nikolić et al., 2018; Sunday, 1996). Our protocol mimics this trajectory, with iPNECs being detected for the first time around Day 13 of differentiation, correlating with the early expression of NKX2.1-expressing anterior foregut endoderm cells (Figure S1B); at this stage the cultures are still submerged. Persistent culture in submerged growth conditions after specification of the primordial lung progenitor results in complete loss of the SYP-expressing cells suggesting that this microenvironment is not conducive to their continued growth and proliferation. There are a number of factors that may influence this including exposure to a more hypoxic environment upregulating hypoxia response genes such as HIF-1α and HIF-2α causing an upregulation of Notch signaling, as previously reported (Gerovac et al., 2014). Notch upregulation, in turn, is known to be antagonistic toward PNEC fate induction (Yao et al., 2018). Existing literature has not specifically indicated a requirement for an ALI during differentiation to generate PNEC-like cells (Chen et al., 2019; Konishi et al., 2016). Other iPSC-derived PNECs have been generated in submerged cultures by formation of spheroids/organoids perhaps accounting for differences in iPNEC yield observed between these protocols and our current protocol. The continued and strong suppression of Notch signaling favors PNEC specification in our protocol. We did not observe significant numbers of ciliated cells under these conditions in contrast to our original protocol. We speculate that this may be due to continuous and increased (10-fold higher) levels of Notch suppression. These cultures were observed for over 3 months, and a substantial increase in both the number and the phenotypic complexity of iPNECs was evident. Although it is currently not possible to determine the precise maturity of iPNECs due to a lack of adult human isolated PNECs for direct comparison, our data suggest that the increase in branched network formation in iPNECs may correlate with their cellular maturation over time.

PNECs are considered neuroepithelial in nature (Cutz et al., 2013; Verckist et al., 2017) as they arise from the undepleted epithelial progenitors located at the Id2+ (inhibitor of DNA binding 2) distal tips of the branching epithelium along with all other epithelial cell lineages within the lung (Rawlins et al., 2009; Song et al., 2012). However, neither primary nor induced SYP+ PNECs were observed to co-express the epithelial marker, EPCAM. This finding is in contrast to a recent study in mouse (Montoro et al., 2018), where analysis of scRNA-seq on purified EPCAM+ tracheal epithelial cells identified a specific cell cluster as neuroendocrine. It should be noted that no analysis of EPCAM-negative cells was performed. There are also no published studies that show co-staining of human PNECs with EPCAM. Furthermore, we found that 65% of human fetal PNECs analyzed in the dataset from Miller et al. is EPCAM− and shows a high level of gene expression concordance with our iPNECs (Miller et al., 2018). Interestingly, we did not observe CGRP expression in our iPNECs. CGRP is often used as a PNEC marker; however, its expression and distribution in PNECs appears to vary among species. CGRP expression in humans, for example, has been observed to correlate to a cancer phenotype (Edbrooke et al., 1985) and observed to only be present in ~20% of primary lung PNECs (Weichselbaum et al., 2005). In human lungs, bombesin/GRP is more robust and widely accepted as a marker of fetal-stage PNECs (Linnoila, 2006; Sunday, 1996; Sunday et al., 1991). Consistent with the literature, we also observe GRP expression peaking during the earlier stages of iPNEC differentiation (Days 13 and 17).

Primary PNECs are known to contain neuropeptides that are stored in dense core vesicles in the cytoplasm and are secreted in response to external stimuli such as hypoxia, nicotine, or tobacco smoke (Branchfield et al., 2016; Cutz et al., 2013; Linnoila, 2006). The iPNECs generated in this study exhibit similar secretion of...
ENO2 and expression of neuropeptides. ROBO genes that help form sensory organoid clusters in vivo via a Slit-ligand-dependent process are also a known feature of mouse PNECs (Branchfield et al., 2016).

Due to a lack of gene expression data focusing on PNECs only a small list of genes is available to define and study PNEC. We hypothesize that our 255 iPNEC-associated genes could serve as a more comprehensive gene set for PNECs and acknowledge that it will be important to validate the iPNEC-associated gene signature in adult human PNECs. Interestingly, our scRNA-seq data indicated two distinct iPNEC clusters from both the t-distributed Stochastic Neighbor Embedding (tSNE) and stream plot analyses, with additional analysis of cell trajectory predicting that the two clusters correspond to two individual physiological states. In a recent publication, rare Notch2+ NEstem and intermediate transit amplifying Hes1+ Notch-active NE cells that contribute to epithelial repair during lung injury in mice were identified (Ouadah et al., 2019). Despite maintaining our 90-day iPSC PNEC differentiation in the continued presence of Notch inhibition we are able to observe two individual iPNEC populations with differential levels of Notch2 and Hes1 expression. We hypothesize that these two distinct subpopulations, iPNEC1 (low Notch) and iPNEC2 (high Notch), may resemble the PNEC heterogeneity observed by Ouadah and colleagues (Ouadah et al., 2019); however, this will require further validation and development of a similar injury model using human iPSC-derived PNECs.

In conclusion, we have successfully derived PNECs from human iPSCs that resemble the known phenotype of human fetal and adult primary PNECs at the cellular and molecular levels. This represents an important step toward creating human cellular models for investigation of lung diseases characterized by PNEC hyperplasia, such as NEHI.

Limitations of the Study
In this study, we undertook the directed differentiation of human iPSCs into PNECs and performed robust characterization of these iPNECs at the cellular and molecular levels. Although we were able to demonstrate significant gene expression concordance with human fetal lung PNECs, it will be important to also make comparisons to the transcriptomes of human adult lung PNECs as these data become available to determine the maturation state of our iPNECs. With <1% of the cellular composition of the adult lung being PNECs, it remains challenging to capture sufficient cells to profile. Currently, methods for the specific isolation/enrichment of PNECs do not exist, making such comparisons extremely challenging. Although we have included experiments that support a functional profile for the iPNECs such as the expression and secretion of major neuropeptides (Figure 5), additional experiments will be important to further establish their functional competency. Such experiments might include measurement of neuropeptide release under hypoxic conditions and mechanical stress or in response to allergic stimulants, which would also be beneficial for disease modeling. Although analysis of our scRNA-seq datasets indicates the presence of two PNEC populations, in the current article the functional differences between these two populations were not addressed. In future experiments it would be interesting to determine whether these are two independent PNEC populations or whether they are cells at different stages along the differentiation trajectory. It also remains unclear whether primary or induced PNECs can be expanded ex vivo. This would require an extensive series of experiments to optimize both the purification and culture of both PNEC and iPNECs; however, this would be incredibly valuable to further study the biology of PNECs. Finally, to definitively confirm a basal cell origin of PNECs, lineage tracing will need to be performed in both primary HTECs and iPSC-derived cultures requiring the generation of iPSC reporter lines.

METHODS
All methods can be found in the accompanying Transparent Methods supplemental file.

DATA AND CODE AVAILABILITY
We have uploaded and made available our single-cell RNA sequencing data generated from our Day 91 differentiation cultures on Gene Expression Omnibus (GEO) database under accession number: GSE146990.

SUPPLEMENTAL INFORMATION
Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2020.101083.
ACKNOWLEDGMENTS

We acknowledge the use of the services at the USC Optical Imaging Facility, Genomics Core in the Department of Biomedical Sciences at the Cedars-Sinai Medical Center, and the LungMAP Consortium. This work was supported by NIH grants R35HL135747, R00NS077435, and R01NS097850; the New York Stem Cell Foundation; the USC Broad Innovation Award; and the Hastings Foundation. J.K.I. is a New York Stem Cell Foundation-Robertson Investigator. Z.B. is Ralph Edgington Chair in Medicine and Hastings Professor of Medicine. The LungMAP Consortium (U01HL122642) and the LungMAP Data Coordinating Center (1U01HL122638) are funded by the National Heart, Lung, and Blood Institute (NHLBI).

AUTHOR CONTRIBUTIONS


DECLARATION OF INTERESTS

J.K.I. is a co-founder of AcuraStem Incorporated. J.K.I. declares that he is bound by confidentiality agreements that prevent him from disclosing details of his financial interests in this work. All data needed to evaluate the conclusions in the article are present in the article paper and/or the Supplemental Information. Additional data related to this article may be requested from the authors.

Received: October 22, 2019
Revised: February 13, 2020
Accepted: April 15, 2020
Published: May 22, 2020

REFERENCES


Supplemental Information

Efficient Generation and Transcriptomic Profiling of Human iPSC-Derived Pulmonary Neuroendocrine Cells

Pooja Hor, Vasu Punj, Ben A. Calvert, Alessandra Castaldi, Alyssa J. Miller, Gianni Carraro, Barry R. Stripp, Steven L. Brody, Jason R. Spence, Justin K. Ichida, Amy L. Ryan (Firth), and Zea Borok
SUPPLEMENTAL INFORMATION

Supplemental Figure Legends

Figure S1. Directed differentiation of additional human iPSC lines into iPNECs related to Figure 1. (A) Representative immunofluorescence (IF) images of iPNECs at Day 31 of differentiation derived from two additional human iPSC lines show expression of SYP, CHGA, PGP9.5 and ENO2 (all red). Scale bars = 50 µm. (B) Orthogonal projection of Day 31 of differentiation showing SYP+ iPNEC (red) are present on the top layer of the 3D cultures. Nuclei are counterstained with DAPI (blue). Scale bar = 200 µm.

Figure S2. Expression of key PNEC markers during in vitro differentiation and in adult human lung related to Figure 1. Representative IF images of (A) ENO2 (scale bar = 100 µm) and (B) SYP (in red) over 31 days of iPSC differentiation and (C) adult human bronchial lung sections stained for ENO2 (red) and KRT5 (green) (scale bar = 50 µm). Nuclei are counterstained with DAPI (blue). Relative mRNA expression of (D) GRP (E) ASCL1 (F) HEY1 and (G) HES1 over the time course of differentiation; data are normalized to Day 5 of differentiation and internal control β-actin and expressed as mean ± SEM from 3 independent experiments.

Figure S3. Notch inhibition dose response to induce optimal iPNEC differentiation related to Figure 3. Representative IF images of SYP+ iPNECs (red) and cleaved Caspase 3 (green) at Day 31 of differentiation. (A) In the presence of DAPT (1 µM, 10 µM and 20 µM). N=3 experimental replicates. Scale bars =100 µm. (B) In the presence of DBZ (0.5 µM, 2 µM and 5 µM). N=2 experimental replicates. Scale bars = 50 µm.

Figure S4. Extended Notch inhibition increases iPNEC phenotypic complexity related to Figure 3. Depiction of Sholl analysis to quantify arbors in an individual image. Quantification of number of intersections and branching density for SYP+ (B, C), ROBO2+ (D, E) and ENO2+ (F, G) cells at Days 31, 60 and 91 of differentiation. N=3 experimental replicates. One-way ANOVA. *p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001. Scale bars = 100 µm. Nuclei are counterstained with DAPI (blue).

Figure S5. Transcriptomic analysis of Day 91 iPNECs related to Figure 4. (A) Schematic representation of iPSCs cultured at ALI with 10 µM DAPT for 91 days that were used for single-cell RNA sequencing. (B)
Correlation of selected 1171 most variable genes among two replicates (Pearson correlation coefficient, r=0.94).
(C) Flowchart describing the analytical approach taken to derive an iPNEC-associated gene list comprised of 255 genes. (D) Scatter plot showing concordance of expression of Day 91 iPNECs with respect to human fetal lung-derived primary EpCAM-negative PNECs pooled from gestational weeks 15, 18 and 21.

**Figure S6. ENO2 and ROBO2 expression in Day 91 iPNECs related to Figure 3, Figure 4 and Figure 5.** (A) t-SNE plot showing expression of ENO2 among single cells at Day 91 of iPSC differentiation cultures. Representative IF images showing (B) ENO2 (red) co-localization with SYP (green) and (C) ROBO2 (red) co-localization with SYP (yellow) in Day 91 iPSC differentiation cultures. Scale bars = 100 µm. Nuclei are counterstained with DAPI (blue).

**Figure S7. PNEC marker gene expression in a lineage-specific manner across the trajectories related to Figure 6.** (A) INSM1 (B) UCHL1 (C) NCAM1 (D) SYP (E) NEUROD1 (F) ENO2.

**Figure S8. Human tracheal epithelial basal cells give rise to PNECs related to Figure 6.** (A) Representative IF images indicate that human adult tracheal epithelium-derived primary (HTEC) cultures are comprised of TP63+ and KRT5+ basal cells. Scale bars = 200µm. (B) Representative IF images indicating SYP+ PNECs (red) in >78 day ALI cultures from human adult tracheal epithelium-derived primary basal cells (HTEC) from 3 individual donors. Scale bars = 25 µm.
Supplemental Figures

A.

<table>
<thead>
<tr>
<th></th>
<th>iPSC 1</th>
<th>Overview</th>
<th>iPSC 2</th>
<th>Overview</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYP</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
</tr>
<tr>
<td>CHGA</td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
</tr>
<tr>
<td>PGP9.5</td>
<td><img src="image9.png" alt="Image" /></td>
<td><img src="image10.png" alt="Image" /></td>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
</tr>
<tr>
<td>ENO2</td>
<td><img src="image13.png" alt="Image" /></td>
<td><img src="image14.png" alt="Image" /></td>
<td><img src="image15.png" alt="Image" /></td>
<td><img src="image16.png" alt="Image" /></td>
</tr>
</tbody>
</table>

B.

![Image](image17.png)

Figure S1
Figure S2
Figure S3
Figure S4
Figure S5

A. iPSCs
   Day 0
   →
   Day 17
   ALI
   →
   10μM
   DAPT
   →
   Day 91
   IPNECs

B. Correlation of expression of common
   1171 genes in two replicates

C. Raw data from 10X genomics
   platform (2 replicates)
   ↓
   Alignment with STAR with GRCh38
   ↓
   Consensus 1171 most variable genes
   ↓
   Graph based clustering
   DE (ANOVA) FDR p<0.05
   ↓
   Non-linear dimensionality reduction
   tSNE plots
   ↓
   Identification of cell biomarkers
   ↓
   Gene set enrichment with marker genes
   with 350 tissues and cell types
   ↓
   Identification of 255 iPNEC-associated
   genes

D. Correlation coefficient between expression of 243 common iPNEC-associated
   genes and EPCAM-negative fetal PNEC genes

$r = 0.94$

$r = 0.6$
Figure S6
Figure S8
Supplemental Tables:

Table S4. Primary antibodies related to Figures 1, 2, 3, 4 and Supplemental Figures S1, S2, S3, S4, S6 and S8.

<table>
<thead>
<tr>
<th>Name</th>
<th>Brand</th>
<th>Catalog number</th>
<th>Dilutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYP</td>
<td>Abcam</td>
<td>ab32127</td>
<td>1:300</td>
</tr>
<tr>
<td>SYP</td>
<td>Santa Cruz Biotechnology</td>
<td>sc-17750: SYP (D-4)</td>
<td>1:100</td>
</tr>
<tr>
<td>ENO2</td>
<td>Santa Cruz Biotechnology</td>
<td>sc-271384: Enolase (A-5)</td>
<td>1:100</td>
</tr>
<tr>
<td>CHGA</td>
<td>Proteintech</td>
<td>60135-1-Ig</td>
<td>1:100</td>
</tr>
<tr>
<td>ROBO2</td>
<td>R&amp;D Systems</td>
<td>AF3147</td>
<td>1:100</td>
</tr>
<tr>
<td>PGP9.5</td>
<td>Santa Cruz Biotechnology</td>
<td>sc-271639: UCH-L1 (C-4)</td>
<td>1:100</td>
</tr>
<tr>
<td>EPCAM</td>
<td>R&amp;D Systems</td>
<td>AF960</td>
<td>1:100</td>
</tr>
<tr>
<td>VIM</td>
<td>Sigma-Aldrich</td>
<td>V6630</td>
<td>1:100</td>
</tr>
<tr>
<td>NKX2.1</td>
<td>Abcam</td>
<td>ab76013</td>
<td>1:100</td>
</tr>
<tr>
<td>KRT5</td>
<td>Biolegend</td>
<td>905901</td>
<td>1:100</td>
</tr>
<tr>
<td>TP63</td>
<td>Cell Signaling Technology</td>
<td>4892S</td>
<td>1:100</td>
</tr>
</tbody>
</table>
Table S5. Quantitative RT-PCR primers related to Figure 5 and Supplemental Figure S2.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>β-Actin</td>
<td>5’-CATGTACGTTGGTTATCGAGGC-3’</td>
<td>5’-CTCCTTAATGTCACGCACGAT-3’</td>
</tr>
<tr>
<td>HES1</td>
<td>5’-ACGTGCGAGGGCGTTAATAC-3’</td>
<td>5’-GGGGTAGGTACATGGCAATTG-3’</td>
</tr>
<tr>
<td>HEY1</td>
<td>5’-AGAGTGCGGAGGAGATGGAACC-3’</td>
<td>5’-CGTCGCGCCTTCTCAATTATTC-3’</td>
</tr>
<tr>
<td>ASCL1</td>
<td>5’-CCCAAGCAAGTCAACCGACA-3’</td>
<td>5’-AAGCCGCTGAAGTGGAGC-3’</td>
</tr>
<tr>
<td>GRP</td>
<td>5’-AAAGAGCAAGGGAGGTTATGCA-3’</td>
<td>5’-TCCTTTGCTTCTAGAGACCCA-3’</td>
</tr>
<tr>
<td>DLL3</td>
<td>5’-CGTCGCTAGATGGAAATGC-3’</td>
<td>5’-TCCCGAGGCATAGGAAAGG-3’</td>
</tr>
<tr>
<td>GAD67</td>
<td>5’-GCTTCCGGCTAAAGGCAAGTG-3’</td>
<td>5’-TTCGGGACATAGGAGGAGT-3’</td>
</tr>
<tr>
<td>ENO2</td>
<td>5’-CTGATGGCATCGGATGGATG-3’</td>
<td>5’-CCATGATCGGATGGAGG-3’</td>
</tr>
<tr>
<td>ROBO2</td>
<td>5’-GTTTGTGGTGAGGAGGACTATCT-3’</td>
<td>5’-GTGTGTGGTGAGGAGGACTATCT-3’</td>
</tr>
<tr>
<td>SYP</td>
<td>5’-CTCAGCATCGAGGTCGAGTTC-3’</td>
<td>5’-GAGGAGTAGTCCCAAATGAAGA-3’</td>
</tr>
</tbody>
</table>

Transparent Methods

**iPSC reprogramming:** Human fibroblasts were reprogrammed into iPSCs using lentiviral encoded human reprogramming factors (Oct4, Lin28, cMyc, Klf4, Sox2) along with an shRNA targeting human p53 (Firth et al., 2014). Briefly, 1x 10^5 cells were transduced with recombinant viral particles for 24-48 hours and plated on feeder cells. Twenty-four hours later the media was changed to iPSC media (KO DMEM, 20% KO serum replacement, 0.01mM β Mercaptoethanol, 1% NEAA, 1% Glutamax (all Invitrogen) and 10 ng/ml FGF-2) and subsequently replaced every other day. For titer calculation, 293T cells were transduced by different reprogramming viruses for 2 days and percent of transduced cells was determined by staining for OCT4. Colonies displaying good stem cell morphology were individually picked and expanded, usually around day 15-21. Two additional human lymphoblastoid cell lines from healthy subjects (ND00184 and ND03719) were obtained from the National Institute of Neurological Disorders and Stroke (NINDS) Biorepository at the Coriell Institute. These lines were reprogrammed into iPSCs and served as biological replicates (labeled ‘iPSC1’ and ‘iPSC2’ respectively in Fig. S1A). These lines were created using episomal plasmids, encoding 5 factors POU5F1, SOX2, KLF4, MYCL, and
LIN28A and an shRNA to p53 (Okita et al., 2013) and maintained on Geltrex (Life Technologies) or Matrigel (BD Bioscience) in mTeSR1 medium (STEMCELL Technologies) supplemented with penicillin/streptomycin.

**Differentiation to airway epithelium:** Transwell inserts were coated with a combination of fibronectin (5 μg/mL; BD Biosciences), laminin (5 μg/mL; Sigma-Aldrich), and collagen IV (60 μg/mL; Sigma-Aldrich), or six-well plates were coated with Geltrex at a 1:200 dilution from a 15 mg/ml stock (Thermo Fisher Scientific, A1413302). iPSCs were dissociated to a single-cell suspension using Accutase (Stemcell Technologies, 7920). Then, 300,000 iPSCs were plated per 30-mm insert or 120,000 per 12-mm insert in mTeSR (Stemcell Technologies, 85850) with 10 μM ROCK inhibitor Y-27632 (Selleckchem, Houston, TX, S1049) for 24 hours. iPSCs were then differentiated through definitive endoderm, anterior foregut endoderm and lung-specific endoderm stages (Firth et al., 2014). All media compositions used for the differentiation are the same as in the original protocol. For some experiments, small molecule γ-secretase inhibitors DAPT (Cayman Chemical, Ann Arbor, Michigan, 208255-80-5) or dibenzazepine (DBZ, Cayman Chemical, Ann Arbor, Michigan, 209984-56-5) were added at concentrations of 1, 10 and 20 µM (for DAPT) or 0.5, 2 and 5 µM (for DBZ), starting at Day 17 of differentiation until the experimental end point, with media changes every other day.

**Immunocytochemistry:** Cells were fixed in 3.2% (vol/vol) paraformaldehyde (PFA) for 1 hour at 4°C, washed 3X with phosphate-buffered saline (PBS), and permeabilized in 0.5% Triton X-100 for 1 hour at 4°C. Cells were blocked for 1 hour with commercial CAS block (Thermo Fisher Scientific, 008120) and then incubated in primary antibody (Supplemental Table S2) at 4°C overnight. Cells were washed 3X in 0.1% Triton X-100 in PBS and incubated for 1 hour in secondary antibodies. Stained cells were then incubated with DAPI for 5 min before washing and mounting on glass microscope slides in Immunomount (Thermo Fisher Scientific, 9990402). Alternately, cells were washed after secondary antibody incubation and directly mounted on the glass microscope slide using Vectashield with DAPI (Vector Laboratories, Burlingame, CA, H-1500). Images were captured at either 20X, 40X or 63X on a Zeiss LSM 800 confocal microscope at the USC Optical Imaging Facility (Los Angeles, CA) and processed/counted using Image J (Schneider et al., 2012).

**Immunohistochemistry:** Explanted human lung tissues were obtained in compliance with an Institutional Review Board approved protocol (HS-18-00273) for using human source material in research and the research was
considered exempt. Tissues were fixed for 24 hours in formalin, embedded in paraffin and sectioned at 10 µm thickness. Paraffin-embedded sections were deparaffinized by washing 3 x 4 min in 100% toluene, 1 x 3 min in 95% EtOH (vol/vol), 1x 3 min in 70% EtOH (vol/vol), 1 x 3 min in 50% EtOH (vol/vol) and finally with distilled water. Antigen retrieval buffer (Sigma) was pre-heated to 95°C, and slides were incubated for 30 min and allowed to cool for 20 min at room temperature. Slides were blocked in CAS block for a minimum of 1 hour before staining in primary antibody (Supplemental Table S2) overnight at 4°C. Sections were washed 3x in 0.1% Triton X-100 in PBS and incubated for 1 hour in secondary antibody. Slides were mounted using Vectashield with DAPI (Vector Laboratories). Images were captured at 20X, 40X or 63X on a Zeiss LSM 800 confocal microscope at the USC Optical Imaging Facility.

Quantitative RT-PCR: Total RNA was isolated using the Qiagen RNA Easy Mini Kit (74004, 74134) as per the manufacturer’s protocol. Then, 500 ng of RNA was DNase-treated (Worthington Biochemical Corporation, Lakewood, NJ, LS003172) per the manufacturer’s protocol and reverse-transcribed to cDNA using a Protoscript first strand synthesis kit (New England Biolabs, E6560S). Quantitative PCR was performed using SYBR Green (Biorad, 1725125) in a 5 µl reaction volume at 50 °C for 2 min, 95 °C for 10 min, 95 °C for 15 s, and 60 °C for 1 min for 40 cycles and analyzed using 7900HT SDS software (Biorad). Each run consisted of triplicate technical repeats per experiment. Beta-actin was used as an internal control. The data were normalized to Day 0 for Fig 3D to observe the overall increase/decrease in key PNEC marker expression over 90 days of differentiation. For Supplemental Fig S1D-G, the data were normalized to Day 5 in order to compare the expression of individual markers with an earlier time point during differentiation rather than starting iPSCs at Day 0. Since Day 0 iPSCs tend to have some expression of some of these markers it makes it harder to track smaller gene expression changes over first 30 days of differentiation. Data are expressed as normalized cycle threshold (Ct) ± SEM. Primers used for qRT-PCR are included in Supplemental Table S3.

Single cell RNA sequencing: Day 91 differentiation cultures were dissociated for 30 mins at 37°C using Accutase to obtain single-cell suspensions. Cells were stained with DAPI and sorted for live single cells followed by cDNA library preparation per the Chromium Single-Cell 30 Library Kit (10x Genomics) manufacturer’s guide at the Genomics Core of the Department of Biomedical Sciences at Cedars-Sinai Medical Center and sequenced using
the 10X Genomics platform on a Next Seq 500. Two experimental replicates of 2113 cells and 2039 cells were sequenced at 260-280 million reads per sample.

Single cell RNA-seq data analysis to identify iPNEC-associated genes: Raw reads from the 10X Genomics platform were de-multiplexed using barcode processing and then aligned to the human reference genome (GRCh 38) using STAR aligner. Mitochondrial genes and low abundance genes were filtered out. Post-normalization expression counts for each gene were collapsed and normalized to unique molecular identifiers (UMI) to generate an expression matrix. The count data were analyzed based on the methodology reported by Lun et al. (Lun et al., 2016). A total of 1705 and 1502 cells were selected respectively from each replicate post quality control filtering for further analysis. Based on the log transformed counts, a total of 1171 differentially expressed (DE) genes were identified to be common between both replicates within the iPNEC cluster in comparison to other clusters (Kolodziejczyk et al., 2015). All further analyses were done using this selection of 1171 DE genes. Graph-based clustering followed by statistical ANOVA test was used to identify biomarkers of each cluster. The clustered data was visualized using t-SNE and PCA plots. The PNEC clusters were identified by using a gene set enrichment analysis with various human tissues and cells as reference using the K-nearest neighbor algorithm (R-package Weighted k-Nearest Neighbors for Classification).

Pseudotemporal reconstruction of differentiation trajectories: Cell fate decisions and differentiation trajectories were reconstructed with the single-cell trajectories reconstruction, exploration and mapping v4.1 (H. Chen et al., 2019). It incorporates a Modified Local Linear Embedding (MLLE) along with ElPiGraph (Lever et al., 2016; Zhang and Wang, 2006) and is not dependent on the predefined number of branches thereby allowing an unbiased study of lineage potential across a heterogeneous population (Qiu et al., 2017). To define the developmental trajectories, we used cluster-based MST (minimum spanning tree) analysis to infer the lineage structure. We used normalized expression values to generate a pseudotime plot that can account for both branched and linear differentiation processes. We have further analyzed the potential marker genes associated with each trajectory and its branching. Additionally, we have used the Monocle package that uses reverse graph embedding to reconstruct lineage trajectories (Trapnell, 2020; Trapnell et al., 2014a, 2014b). We selected cells, that were projected onto specific lineages based on the branches termed as cell “states”. Subsequently, we
reran the Monocle3 algorithm and assigned basal cells as starting cell lineage until only a single/ simple trajectory was identified.

**Enzyme-linked immunosorbent assay (ELISA):** ELISA was performed using Human Enolase 2/Neuron-specific Enolase Quantikine ELISA Kit (R&D systems), as per the manufacturer’s instructions. The experiment was repeated four times with basal media and iPSC-conditioned media as controls.

**Differentiation of HTEC:** Human donor lung specimens were dissected into smaller segments for dissociating and harvesting the airway cells (Gu et al., 2014). Total epithelial cell yield will vary depending upon tissue viability and size. Airway epithelial cells were seeded on to collagen IV-coated transwell membranes. A collagen stock (Corning, 3470) of 60 µg/mL was used as the working solution for coating membrane surfaces. The day after seeding the media was removed from the apical surface (air-interfacing) and from there on the cultures were maintained under air-liquid-interface (ALI) conditions. The experiment was repeated with HTEC preparations from seven different donor lungs.

**Image analysis:** Images captured on an LSM 800 confocal microscope were quantified and analyzed using ImageJ (Schneider et al., 2012).

**iPNEC quantification:** To quantify iPNECs in a given culture, the mean fluorescence intensity of the antibody stain and DAPI were quantified for 10–12 images (which covered almost the entire filter) at 20X magnification per time point/condition. For all quantifications, unless otherwise mentioned, percentage of MFI was calculated for each marker fluorescence relative to total DAPI fluorescence, followed by statistical analysis.

**iPNEC branching analysis:** To measure the phenotypic complexity of iPNECs, we used the Sholl analysis plugin on ImageJ, commonly used for analyzing neuronal arbors ([https://imagej.net/Sholl_Analysis](https://imagej.net/Sholl_Analysis)), to measure the extent of iPNEC intersections. 10–12 images at 20X magnification per time point/condition were used for the analysis of each marker. For Sholl analysis, the outputs measured were number of intersections (I) and branching
density (l/area of analyzed) (Binley et al., 2014; Bird and Cuntz, 2019; “Sholl Analysis,” n.d.; Sholl, 1953; Stanko and Fenton, 2017).

**Statistical analysis:** All experiments were conducted using at least three technical replicates (e.g., three 6-wells or Transwells) from the same differentiation. All experiments were performed (independent differentiations) at least three times except where otherwise indicated. Unless otherwise noted, all analyses were performed in GraphPad Prism v.6. All data are presented as mean ± SEM. Statistical significance was determined by Student’s t test (two-tailed) between two groups. Three or more groups were analyzed by one-way analysis of variance (ANOVA). P < 0.05 was considered statistically significant.

**Accession numbers:** Raw scRNAseq data of primary human fetal lung-derived PNECs have been deposited in the EMBL-EBI ArrayExpress database (Accession number: E-MTAB-8221).

**Supplemental References**


