Article

Integrated Single-Cell RNA-Sequencing Analysis of Aquaporin 5-Expressing Mouse Lung Epithelial Cells Identifies GPRC5A as a Novel Validated Type I Cell Surface Marker

Masafumi Horie [1,†], **Alessandra Castaldi** [1,†], **Mitsuhiro Sunohara** [1,2,3], **Hongjun Wang** [1], **Yanbin Ji** [1], **Yixin Liu** [1], **Fan Li** [4], **Thomas A. Wilkinson** [4], **Long Hung** [4], **Hua Shen** [1], **Hidenori Kage** [2], **Ite A. Offringa** [5,6,7], **Crystal N. Marconett** [5,6,7], **Per Flodby** [1], **Beiyun Zhou** [1,6] **and Zea Borok** [1,6,7,*]

[1] Hastings Center for Pulmonary Research and Division of Pulmonary, Critical Care and Sleep Medicine, Department of Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA; mhorie-tky@umin.ac.jp (M.H.); acastald@usc.edu (A.C.); mitsuhiro.sunohara@nifty.ne.jp (M.S.); hjwang@chla.usc.edu (H.W.); yanbinji@usc.edu (Y.J.); yixinliu@usc.edu (Y.L.); huashen@usc.edu (H.S.); pflodby@hotmail.com (P.F.); bzhou@usc.edu (B.Z.)

[2] Department of Respiratory Medicine, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan; kageh-tky@umin.ac.jp

[3] Division for Health Service Promotion, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

[4] Single-Cell, Sequencing, and CyTOF Core (SC2), Children's Hospital Los Angeles, Los Angeles, CA 90027, USA; fanli.gcb@gmail.com (F.L.); tomwilkinson2010@gmail.com (T.A.W.); lhung@chla.usc.edu (L.H.)

[5] Department of Surgery, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA; ilaird@usc.edu (I.A.O.); cmarcone@usc.edu (C.N.M.)

[6] USC Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA

[7] Department of Biochemistry and Molecular Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA

* Correspondence: zborok@usc.edu; Tel.: +323-409-7184; Fax: +323-226-2738

† These authors contributed equally to this work.

check for updates

Abstract: Molecular and functional characterization of alveolar epithelial type I (AT1) cells has been challenging due to difficulty in isolating sufficient numbers of viable cells. Here we performed single-cell RNA-sequencing (scRNA-seq) of tdTomato[+] cells from lungs of AT1 cell-specific Aqp5-Cre-IRES-DsRed (ACID);R26tdTomato reporter mice. Following enzymatic digestion, CD31[-]CD45[-]E-cadherin[+]tdTomato[+] cells were subjected to fluorescence-activated cell sorting (FACS) followed by scRNA-seq. Cell identity was confirmed by immunofluorescence using cell type-specific antibodies. After quality control, 92 cells were analyzed. Most cells expressed 'conventional' AT1 cell markers (*Aqp5*, *Pdpn*, *Hopx*, *Ager*), with heterogeneous expression within this population. The remaining cells expressed AT2, club, basal or ciliated cell markers. Integration with public datasets identified three robust AT1 cell- and lung-enriched genes, *Ager*, *Rtkn2* and *Gprc5a*, that were conserved across species. GPRC5A co-localized with HOPX and was not expressed in AT2 or airway cells in mouse, rat and human lung. GPRC5A co-localized with AQP5 but not pro-SPC or CC10 in mouse lung epithelial cell cytospins. We enriched mouse AT1 cells to perform molecular phenotyping using scRNA-seq. Further characterization of putative AT1 cell-enriched genes revealed GPRC5A as a conserved AT1 cell surface marker that may be useful for AT1 cell isolation.

1. Introduction

The lung alveolar epithelium comprises two morphologically and functionally distinct cell types: type I (AT1) and type II (AT2) cells. AT2 cells cover ~5% of the alveolar surface and have important functions including the production of surfactant proteins and both self-renewal and differentiation to AT1 cells during homeostasis and repair following injury [1–4]. AT1 cells are large, flat cells with long cytoplasmic processes that cover more than 95% of the alveolar surface and play a key role in gas exchange and ion transport [5,6]. In contrast to AT2 cells, which have been extensively studied, far less is known about AT1 cell contributions to alveolar homeostasis due to difficulty in isolating viable AT1 cell populations of sufficient yield and purity for detailed molecular and functional characterization.

An in vitro culture model which recapitulates aspects of AT2 to AT1 cell transdifferentiation has provided important insights into functional and phenotypic properties of AT1 cells [7–10]. While transdifferentiated 'AT1-like' cells express many of the phenotypic markers characteristic of AT1 cells in situ, and their transcriptomic profile overlaps considerably with that of freshly isolated AT1 cells, they are likely not entirely identical to AT1 cells due to alterations occurring in culture [11,12], making it important to validate in vitro findings using freshly isolated AT1 cells. Isolation of viable AT1 cells has been challenging due to their fragility based on their topological complexity, while lack of specific AT1 cell surface markers for fluorescence-activated cell sorting (FACS) or magnetic-activated cell sorting (MACS) has limited the ability to enrich for them [5,12,13].

Aquaporin 5 (AQP5), podoplanin (PDPN), homeodomain-only protein (HOPX) and more recently advanced glycation end products receptor (AGER) have been viewed as 'classical' AT1 cell markers [4,14]. However, all have limitations with regard to specificity for AT1 cells within the lung and are also expressed at other extrapulmonary sites, limiting their utility for isolation of pure populations of AT1 cells and generation of AT1 cell-specific Cre driver lines. In this regard, the use of podoplanin (PDPN) for AT1 cell isolation has been reported, but PDPN is also expressed in bronchial and lymphatic cells in the lung and failed to separate ATI cells from bronchiolar epithelial cells [15]. We recently identified by transcriptome profiling of rat AT1 and rat and human AT1-like cells a number of putative novel AT1 cell markers, including the GRAM domain 2 (GRAMD2) gene [12]. However, further validation of a panel of robust murine AT1 cell markers is needed for the development of novel tools for characterization of AT1 cell molecular phenotype.

Single-cell RNA-sequencing (scRNA-seq) is a powerful technique enabling characterization of the molecular states of individual cells by analyzing their transcriptional profiles. Implementation of this technique has advanced our understanding of cellular heterogeneity in various tissues and disease states [16–20]. In scRNA-seq analyses of whole lung, AT1 cells have been captured to only a limited extent [21–23], and the proportion of AT1 cells recovered for analysis has been quite low compared to that of other cell types. We therefore considered enriching for AT1 cells prior to scRNA-seq as a potential approach to facilitate further characterization of AT1 cell transcriptomes.

AQP5 is a water channel protein that is expressed on the apical surface of AT1 cells of the alveolar epithelium [24–26]. It is also expressed on some cells in the large airways, and as previously reported, depending on the strain of mice, on a small subset of AT2 cells [25]. We previously generated *Aqp5-Cre-IRES-DsRed* (ACID) mice in which a *Cre-IRES-DsRed* cassette was inserted into exon 1 of the endogenous *Aqp5* gene, which is abundantly expressed in AT1 cells. However, ACID mice did not show a specific DsRed signal, likely due to low levels of protein expression. Thus, in the present study ACID mice were crossed with *ROSA26-stop^{flox}-tdTomato* conditional knockin mice, resulting in *ACID;R26tdTomato* mice that express tdTomato in *Aqp5*-expressing cells. The tdTomato$^+$ population included cells identified as AT1 cells as well as other lung cell types; nevertheless, using these *Aqp5*-Cre reporter mice, we obtained sufficient enrichment to successfully perform scRNA-seq of a population of

cells identified by Uniform Manifold Approximation and Projection (UMAP) as AT1 cells. Consistent with our previous study [27], analysis of the expression pattern of 'classical' AT1 cell markers within this AT1 cell population at a single-cell level revealed considerable heterogeneity. By integrating our data with public datasets [28–31] and a previous dataset derived from our in vitro AT2 to AT1 cell differentiation model [12], we identified three conserved lung- and AT1 cell-enriched genes: *Ager*, rhotekin 2 (*Rtkn2*) and G protein-coupled receptor class C group 5 member A (*Gprc5a*). One of these, *Ager*, has previously been identified as enriched in AT1 cells, although it is also expressed in other lung cells [14]. Finally, we evaluated cell-specific protein expression in the airways and distal lung by immunofluorescence staining and validated GPRC5A as a novel AT1 cell surface marker.

2. Materials and Methods

2.1. Generation of Aqp5-Reporter Mice

To generate AT1 cell-specific reporter mice, *Aqp5-Cre-IRES-DsRed* (ACID) mice in which a *Cre-IRES-DsRed* cassette is knocked into exon 1 of the endogenous *Aqp5* gene were crossed with mice with a *ROSA26-stop^flox^-tdTomato* conditional knockin allele [25,32]. Because our ACID knockin construct which includes the DsRed gene coding for red fluorescent protein showed no specific DsRed signal, likely due to low levels of protein expression, we used *ROSA26-stop^flox^-tdTomato* mice as reporter mice [25]. Double-heterozygous mice (termed *ACID;R26tdTomato*) which express the red fluorescent protein Tomato after Cre/loxP recombination were generated (Supplemental Figure S1). Mice were on an 129S6/SvEvTac background. Three male mice were used for optimization of the digestion and sorting strategy. A four-month old male mouse was used for scRNA-seq analysis.

2.2. Lung Digestion

Mouse lungs were surgically harvested following perfusion with phosphate-buffered saline (PBS) via the right ventricle. After digestion with 0.2% pronase (Roche, Indianapolis, IN, USA) and 0.08% collagenase/dispase (Roche), lungs were minced, suspended in Dulbecco's modified Eagle's medium (DMEM)/F12 (Sigma, St. Louis, MO, USA) with 0.01% DNase I (Roche), and sequentially filtered through cell strainers of 100 μm and 40 μm pore size (BD Falcon, Franklin Lakes, NJ, USA). Cells were resuspended in Red Blood Cell Lysis Buffer (Roche) for 1 min on ice and washed with PBS. The resulting cell preparation was suspended in cold buffer consisting of PBS with 2% fetal bovine serum (FBS, HyClone/Thermo Fisher Scientific, Tustin, CA, USA), 2 g/L glucose (Sigma) and 20 mM HEPES (Sigma) at a density of 10^6 cells/100 μL, and subsequently incubated with antibodies (Abs) as described below for cell sorting.

2.3. Flow Cytometry

After pre-incubation with Ultra-LEAF purified anti-mouse CD16/32 Ab (Biolegend, San Diego, CA, USA) for 5 min on ice to block of Fc receptors, cells were incubated with primary Abs on ice for 45 min as follows: CD31-biotin (1:50, clone: 390, eBioscience, San Diego, CA, USA), CD45-biotin (1:100, clone: 30-F11, eBioscience) and CD324 (E-Cadherin)-PerCP-eFluor 710 (1:20, Thermo Fisher Scientific). After washing 3 times, cells were incubated with Streptavidin Allophycocyanin (1:200, APC, BD Biosciences, San Jose, CA, USA) secondary Ab for 30 min on ice. After washing 3 times, APC-negative, PerCP-eFluor 710-positive, and tdTomato-positive cells were sorted with a FACSAria cell sorter (BD Biosciences) in preparation for scRNA-seq. Cell viability following FACS was measured using trypan blue vital dye staining, and >95% of cells were confirmed to be viable. To confirm tdTomato expression in sorted cells, 1×10^5 sorted cells were fixed with 4% paraformaldehyde (PFA), and cytospins were prepared by applying cells to EZ Single Cytofunnels (Thermo Fisher Scientific). Cytospins were counterstained with 4',6-diamidino-2-phenylindole (DAPI, Vector Laboratories, Burlingame, CA, USA), and tdTomato expression was confirmed using a Nikon Eclipse 80i microscope (Nikon Instruments, Inc., Melville, NY, USA).

2.4. scRNA-Seq

A single-cell suspension was loaded onto a C1 mRNA Seq HT IFC (High Throughput Integrated Fluidics Circuit, Fluidigm, South San Francisco, CA, USA) designed to capture 800 single cells with diameters of 10–17 μm. The IFC was run on a C1 Single Cell Auto Prep System (Fluidigm) using the C1 Single-Cell mRNA Seq HT Reagent Kit v2 (Fluidigm). After initial cell capture, the IFC was scanned on a Leica DMI6000B microscope (Leica Microsystems, Wetzlar, Germany) to verify cell capture. Cell lysis, reverse transcription, PCR amplification, cell bar code tagging and harvesting were done onboard the C1 System. Harvested sample cleanup was performed using AMPure XP beads (Beckman Coulter, Brea, CA, USA). Sequencing libraries were generated using the Nextera XT DNA Library Preparation Kit (Illumina, San Diego, CA, USA), with indexes from Nextera XT Index Primer Set A and Set B. Libraries were pooled, and library quality was assessed on an Agilent High Sensitivity DNA Chip (Agilent Technologies, Santa Clara, CA, USA). The pooled library, combined with 30% PhiX, was loaded on a Nextseq 500/550 High Output Kit v2 Flowcell (Illumina) and sequenced on a Nextseq 500 Sequencing System (Illumina) using paired end reads (R1–26bp, R2–120bp).

2.5. Data Processing

Quality control of sequence reads was performed by Trim Galore and FastQC/multiQC [33]. Mapping was performed by HISAT2 against GRCm38 [34], and quantification of mapped sequence reads was calculated by featureCounts [35]. Quality control of cells was performed using the following parameters: (1) total read counts (>90,000 counts), (2) detected genes (>500 genes) and (3) proportion of sequence reads from mitochondrial genes (<0.15) [36]. Expression level of cell population markers was evaluated as absolute read count. Data clustering, including Uniform Manifold Approximation and Projection (UMAP) and identification of cluster-specific marker genes, was done using R package Seurat v3.0 [37]. Pathway analysis was performed using Ingenuity Pathway Analysis (IPA, QIAGEN, Redwood City, CA, USA). For the analysis of the bulk RNA-seq data obtained from GTEx portal site [31], counts per million were used for normalization between samples. The dataset has been deposited in the GEO database (GSE120285).

2.6. Comparison with Public Datasets

For validation of AT1 cell-enriched markers, (1) the scRNA-seq dataset of normal adult mouse lung was obtained from GSE108097 (Mouse Cell Atlas) [28], (2) the scRNA-seq dataset of postnatal day 1 mouse lung was obtained from the LungMAP portal site [29], (3) the scRNA-seq dataset of AT1 cells from postnatal day 60 mouse lung was obtained from GSE106960 [30], (4) bulk RNA-seq data from the human AT2-AT1 cell in vitro differentiation model were obtained from GSE66627 [12] and (5) bulk RNA-seq data of 334 normal lung tissues and 8564 other tissues from 29 different organs were obtained from the GTEx portal site [31]. Localization of protein expression was validated using data from the Human Protein Atlas (https://www.proteinatlas.org/) [38]. For evaluation of *Aqp5* expression in airway cells, the scRNA-seq dataset of airway epithelium was obtained from the Single Cell Portal site (https://portals.broadinstitute.org/single_cell/) [39].

2.7. Immunofluorescence Staining of Frozen Lung Sections and Cytospins

For mouse and rat, lungs were cleared of blood by perfusing in PBS, fixed in 4% PFA, incubated in sucrose solution, filled with Optimal Cutting Temperature Compound (OCT; VWR, Radnor, PA, USA))/50% PBS and frozen in OCT. Paraffin-embedded samples were prepared from human lungs that were deemed not suitable for transplantation. Lung cryosections (5 μm) were prepared as described [40]. Following antigen retrieval in Antigen Unmasking Solution at low pH (Vector Laboratories), slides of lung sections were incubated for 30 min in 0.2% Triton-X in PBS to permeabilize cells. After incubation in CAS block (Invitrogen/Zymed, San Diego, CA, USA), slides were incubated with primary Abs overnight at 4 °C. Goat anti-pro-SPC (sc-7706; Santa Cruz Biotechnology, Inc., Santa Cruz, CA, USA),

mouse anti-ABCA3 (17-H5-24; Seven Hills, Cincinnati, OH, USA), mouse anti-HOPX (sc-398703; Santa Cruz), goat anti-AQP5 (sc-9890; Santa Cruz), goat anti-CC10 (sc-9772; Santa Cruz) or rabbit anti-GPRC5A (abx005719; Abbexa, Cambridge, UK) were used as primary Abs. Slides were then incubated with biotinylated anti-goat (Millipore, Temecula, CA, USA), biotinylated anti-rabbit (Millipore) or Alexa Fluor 594 anti-mouse (Thermo Fisher Scientific) secondary Abs, followed by streptavidin-FITC (Vector). Normal goat IgG (Vector), normal rabbit IgG (Vector) or normal mouse IgG (Vector) were used as negative controls. Finally, slides were mounted with Vectashield mounting medium including DAPI (Vector). Confocal images were captured using a Leica SP8 confocal system (Leica Microsystems), and negative controls were used to set the laser intensity (Supplemental Figures S7 and S8 and Figures 3C,D, 5A, 6A, and 7). Cytospins of crude cells harvested from digestion of mouse lungs were prepared as described for sorted cells above. For immunostaining, cytospins were post-fixed in 4% PFA, and antigen retrieval and staining were performed as described above for frozen lung sections. The percentage of cells that were double positive for tdTomato and pro-SPC or CC10 in immunofluorescence staining of lung sections was calculated by manually counting 10 random fields, and the percentage of cells that expressed tdTomato in immunofluorescence staining of cytospins was calculated by manually counting 5 random fields. Microscopic evaluation was independently validated by two blinded operators.

2.8. Statistics

R package Seurat v3.0 was used for statistical analysis. Spearman's correlation coefficient (R) was calculated for correlation analysis, and $p < 0.05$ was regarded as significant. For violin plots, the smoothed curves were generated by Seurat using the kernel density estimator.

3. Results

3.1. Isolation of AT1 Cells from ACID;R26tdTomato Mice

To optimize the lung digestion and sorting strategy, we initially isolated AT1 cells from three *ACID;R26tdTomato* mice that expressed the red fluorescent protein tdTomato in Aqp5-expressing cells (Supplemental Figure S1). Single-cell RNAseq was subsequently performed on sorted cells from one four-month-old male mouse (Figure 1). After enzymatic digestion of *ACID;R26tdTomato* mouse lungs, cells were incubated with CD45/CD31/E-cadherin Abs, and CD31⁻CD45⁻E-cadherin⁺tdTomato⁺ cells were sorted by FACS (Figure 1A, Supplemental Figure S2). FACS analysis of cells isolated from three lungs showed that 31 ± 10% of epithelial cells were positive for tdTomato (data not shown). By fluorescence microscopy, we confirmed that most sorted cells (80%) examined in cytospins were positive for tdTomato (Figure 1B). The remaining cells (~20%) were negative for tdTomato, perhaps due to loss of their cytoplasm as a result of the FACS procedure. A sorted cell suspension was loaded onto the C1 Fluidigm system, and manual checking by microscopy revealed that 392 cells (49.0%) were singlets while the remaining cells were doublets (6.75%) or empty wells. Only single cells were subsequently analyzed. After extensive quality checking of cells based on the number of detected genes, total read counts and expression of mitochondrial genes, 92 cells were selected for final analysis (Supplemental Figure S3).
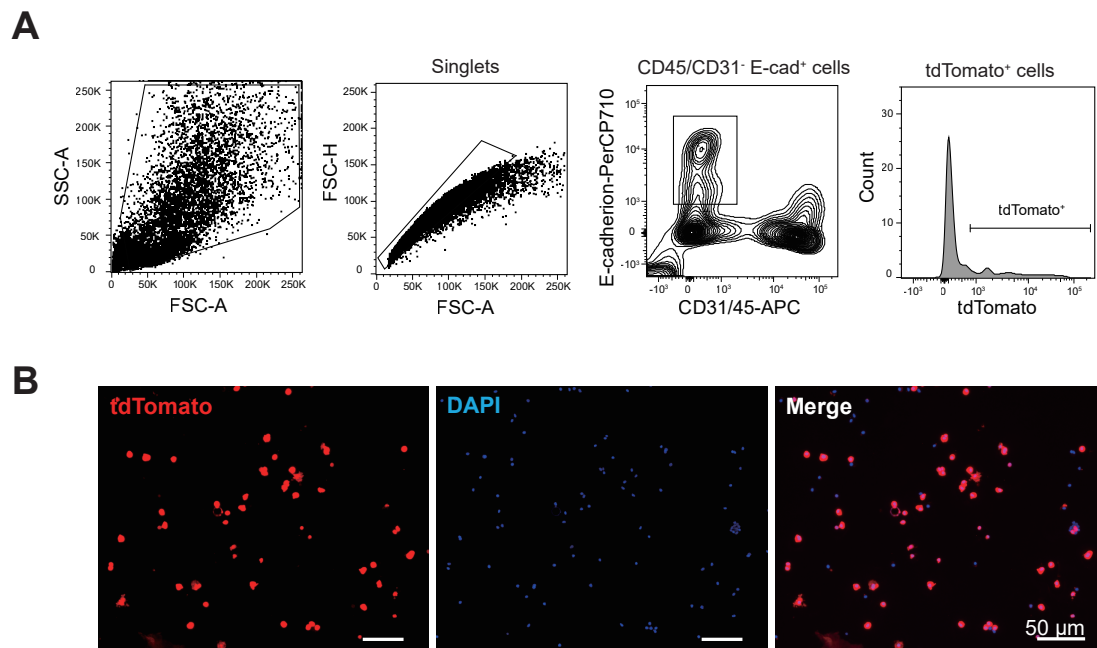
A

B

Figure 1. Sorting of tdTomato⁺ cells from *ACID;R26tdTomato* mice. (**A**) Flow cytometry strategy for sorting AT1 cells. CD45/CD31 negative, E-cadherin positive, tdTomato positive cells were sorted; 31% ± 10% (*n* = 3) of epithelial cells from test mice were positive for tdTomato (panel 4). The box in the third panel delineates the CD45/CD31-negative and E-cadherin-positive populations. (**B**) Immunofluorescence images of cytospins after fluorescence-activated cell sorting (FACS). Approximately 80% of cells are positive for tdTomato.

3.2. Unsupervised UMAP Divided Cells into Six Clusters

Unsupervised UMAP divided these 92 cells into six clusters (C1–C6, Figure 2A), and enriched genes in each cluster were identified (Supplementary Table S1). A heatmap with 10–15 markers representative of fibroblasts, club, basal, AT2, AT1 and ciliated cells showed that gene expression patterns of these six clusters correspond to specific cell types (Figure 2B). For C6, the number of fibroblasts was small (*N* = 2) so that no C6-enriched genes met statistical significance; however, the cluster was distinct from the others as shown by the heat map (Figure 2B). As expected, AT1 cells represented the biggest cluster (Figure 2A, cluster C4). The expression of various cell-specific markers confirmed that C1 represented club cells (*Scgb1a1*), C2 basal cells (*Krt5*), C3 AT2 cells (*Sftpc*), C4 AT1 cells (*Ager*, *Hopx*, *Pdpn*, *Cav1*, *Cav2*, *Igfbp2* and *Gramd2*) and C5 ciliated cells (*Foxj1*), as shown in the UMAP (Figure 2C) and violin plots (Figure 2D). Of note, the expression of endogenous markers (e.g., beta-actin, *Actb*) was stable among cell clusters (Supplemental Figure S4). The full list of enriched genes in each cluster is reported in Supplemental Table S1. The expression of *tdTomato* most closely resembled that of *Aqp5* (Figure 3A), and both *tdTomato*⁺ and *Aqp5*⁺ cells were highly enriched in the AT1 cell population. A moderate correlation between *Aqp5* and *tdTomato* expression was observed for epithelial cells (R = 0.32, *p* < 0.05, Figure 3B). Immunofluorescence staining of mouse lung tissue showed that ~10% of tdTomato⁺ cells were also labeled with CC10 (Figure 3C) and about 12% with pro-SPC (Figure 3D), indicating tdTomato expression in some club and AT2 cells in the *ACID;R26tdTomato* mouse, although correlation between *Aqp5* and *Scgb1a1* and *Sftpc* expression was very weak (R = −0.06 and −0.15, respectively). Nevertheless, the largest cluster was identified as AT1 cells, indicating that tdTomato sorting of lung cells isolated from *ACID;R26tdTomato* mice successfully enriched for the AT1 cell population.
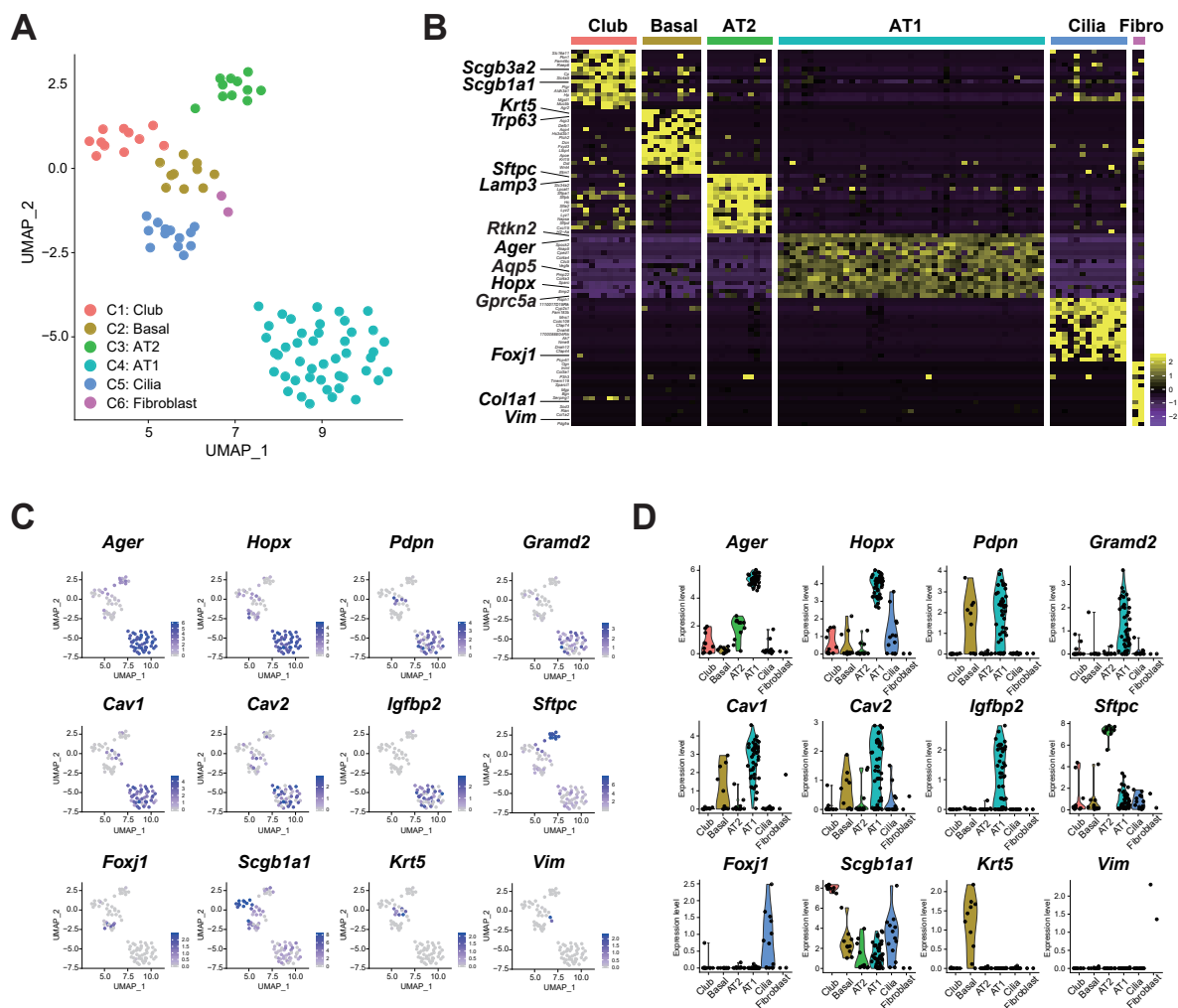
Figure 2. Clustering of 92 tdTomato-positive cells in *ACID;R26tdTomato* mice. (**A**) Unsupervised UMAP divided 92 cells into 6 clusters (C1–C6). Expression level (C and D) represents absolute read count. (**B**) Heatmap with 10–15 markers representative of each cell type. Yellow and purple colors indicate high and low expression, respectively. (**C**) UMAP plots with known cell type-specific marker genes. AT1 cells (*Ager*, *Hopx*, *Pdpn*, *Igfbp2*, *Cav1*, *Cav2* and *Gramd2*), AT2 cells (*Sftpc*), club cells (*Scgb1a1*), basal cells (*Krt5*), ciliated cells (*Foxj1*) and fibroblasts (*Vim*). Blue and grey indicate high and low expression, respectively. (**D**) Violin plot with the same marker genes as in C.

3.3. Identification of Gprc5a as a Novel Candidate AT1 Cell Marker

Our scRNA-seq data indicated that some known AT1 cell genes such as *Pdpn*, *Cav1* and *Cav2* are also variably expressed at low levels in airway epithelial cells and AT2 cells (Figure 2C,D, Supplemental Figure S5A), reinforcing the notion that specificity of currently available AT1 cell markers is not optimal. Furthermore, three of these classical markers (*Pdpn*, *Aqp5* and *Cav2*), as well as other more recently described markers (e.g., *Scnn1g*, *Gramd2* and *Igfbp2*), were found to be heterogeneously expressed in subsets of AT1 cells at the single-cell level (Supplemental Figure S5B).
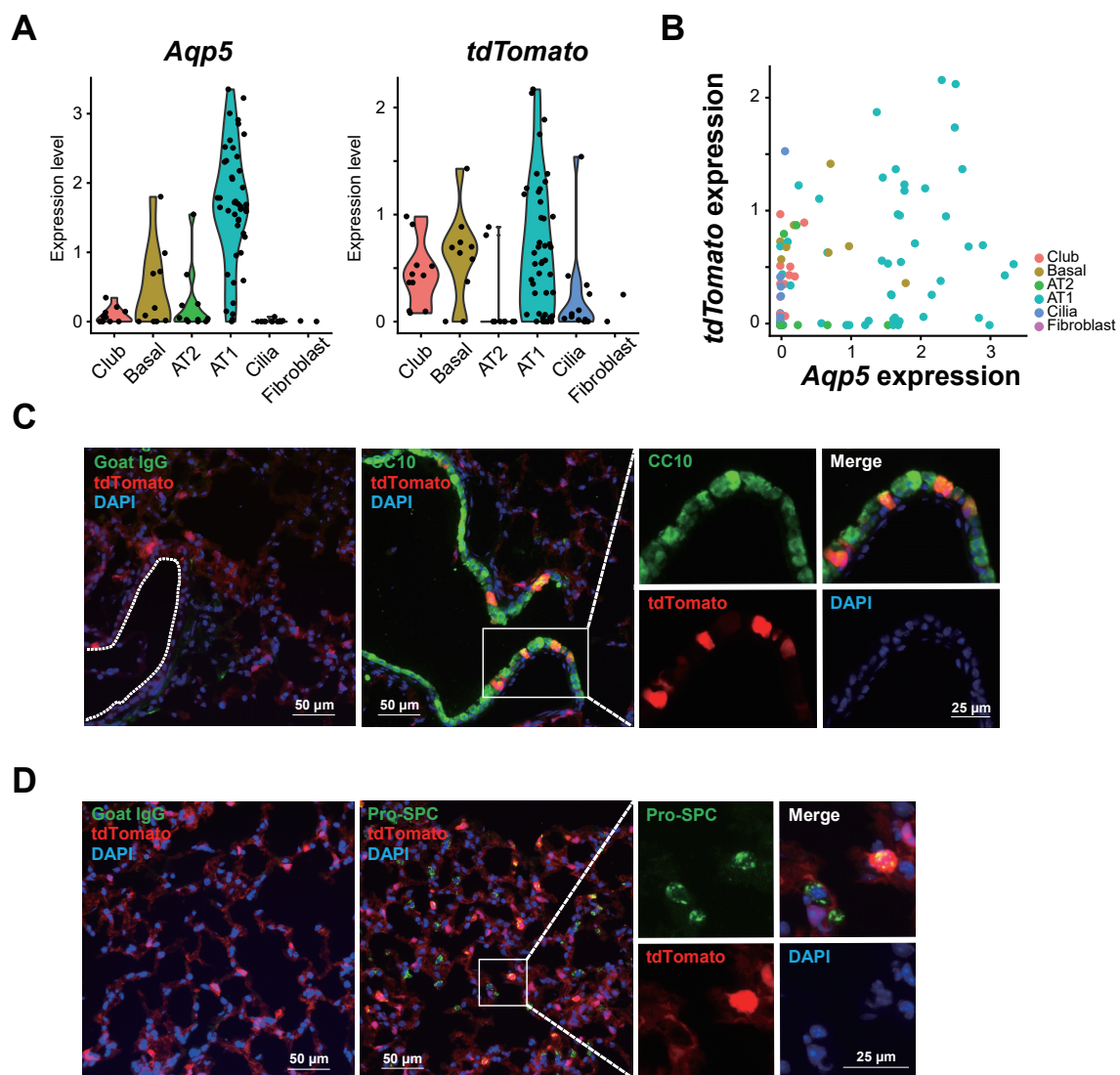
Figure 3. TdTomato expression in *ACID;R26tdTomato* mouse lung. (**A**) Violin plot for *Aqp5* (left) and *tdTomato* (right). (**B**) Scatter plot between *Aqp5* and *tdTomato* expression. (**C**) Immunofluorescence staining shows CC10 (green) and tdTomato (red) expression in *ACID;R26tdTomato* mouse lung. Goat IgG (left panel) was used as negative control. White dotted line indicates outline of airway. (**D**) Immunofluorescence staining shows pro-SP-C (green) and tdTomato (red) expression in *ACID;R26tdTomato* mouse lung. Goat IgG (left panel) was used as negative control.

Our scRNA-seq was purposely biased by the use of Aqp5-lineage traced mice, limiting the analysis to only the Aqp5-positive AT1 cell population, and was used as a way to enrich for these cells. Subsequently, in order to identify a novel candidate AT1 cell surface marker, we undertook an unbiased approach, shown schematically in Figure 4A, by integrating our data with existing datasets reporting gene expression in AT1 cells. Because our goal was to identify a surface marker conserved between murine and human species, we included in our analysis available RNA-seq datasets originating from both mouse and human samples. First, the top 100 genes with the lowest false discovery rate (FDR) values and highest fold-change in AT1 cells were extracted from our dataset (Supplemental Table S2). We then integrated these 100 genes with three independent scRNA-seq datasets of mouse lung in which AT1 cell markers were included [28,30,41], further reducing the number of AT1 cell-enriched genes to 34 ("AT1 cell-enriched among all datasets" column in Supplemental Table S2). Next, we compared these data with RNA-seq data from our in vitro human AT2-AT1 cell differentiation model [12]. Among 34 genes, 25 were also enriched in AT1-like cells. Of note, *Gramd2* which was recently

identified by cross-species transcriptome profiling between human and rat [12] is included in this list (Supplemental Table S2). Furthermore, *Gprc5a*, which was suggested as a possible AT1 cell marker in the same study but not further investigated, is in the list as well. Subsequently, to ensure that the candidate AT1 cell marker was not expressed in other tissues, we extracted lung-enriched genes using GTEx, which contains RNA-seq data of human tissues from various organs. Among the 25 genes, *Rtkn2*, *Ager*, and *Gprc5a* were identified as both AT1 cell- and lung-tissue enriched genes. UMAP plot from the Mouse Cell Atlas shows that these three genes are enriched in AT1 cells, although *Ager* is highly expressed in AT2 cells and *Gprc5a* is slightly expressed in some AT2 cells (Figure 4B). UMAP plots for 'classical' AT1 cell markers are shown in Supplemental Figure S6. On the other hand, *Rtkn2* appears to be expressed in only a subset of AT1 cells (Figure 4B). RNA-seq data from GTEx show that expression of these three genes in lung is clearly higher than in other organs (Figure 4C). Finally, integrating the comprehensive annotation of protein localization from the Human Protein Atlas, GPRC5A and AGER, but not RTKN2, were found to be plasma membrane proteins already validated at the protein level, and specifically GPRC5A protein was reported as 'high in pneumocytes' (https://www.proteinatlas.org/). We selected GPRC5A for further validation as it strictly met the following criteria: (1) mouse/human conservation, (2) lung specificity, (3) AT1 cell enrichment and 4) transmembrane localization criteria. This made it the best novel AT1 cell marker candidate with high potential for future isolation of AT1 cells and generation of specific Cre reporter mice.
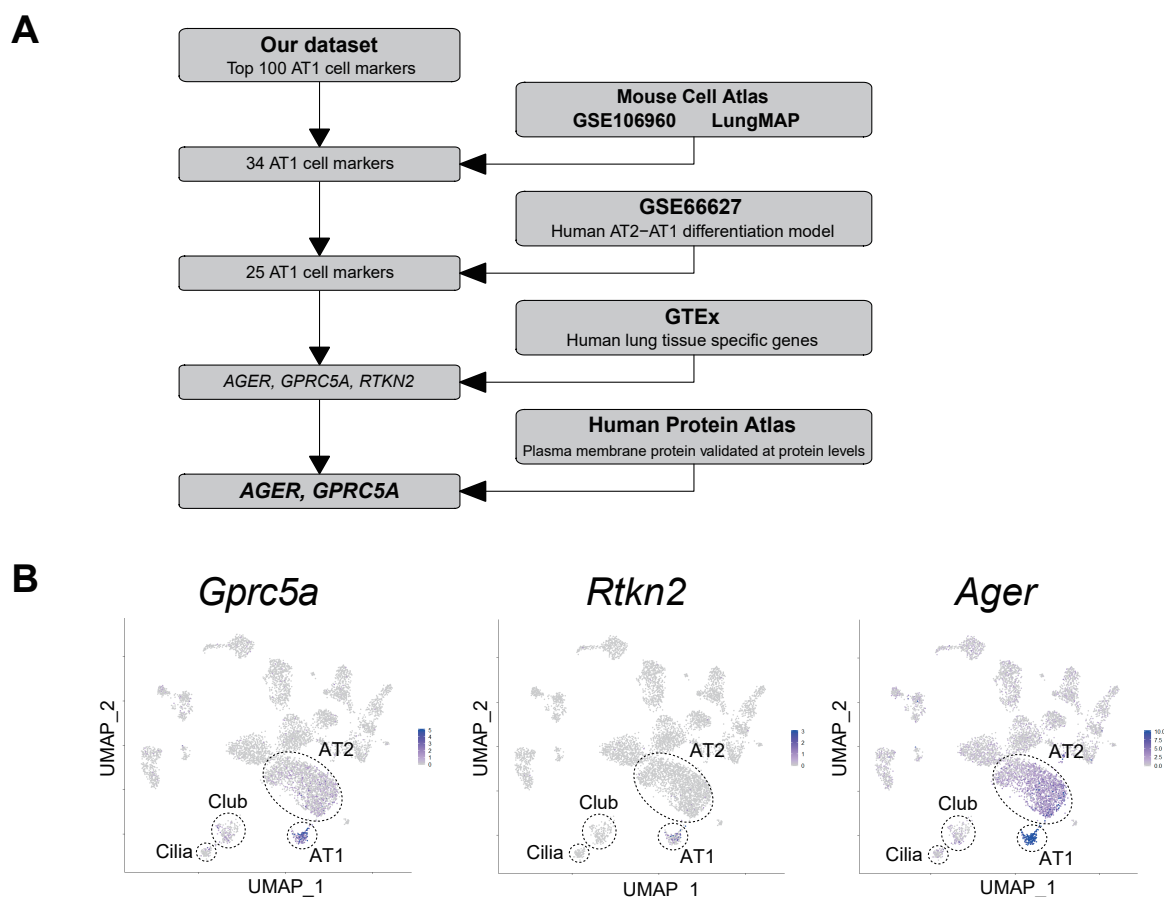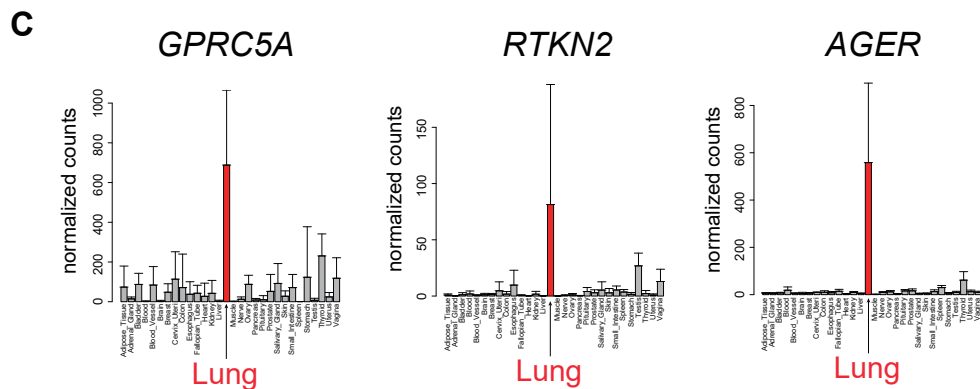

Figure 4. *Cont.*

C

Figure 4. High expression of *Gprc5a*, *Rtkn2* and *Ager* in AT1 cells and lung. (**A**) Flowchart for identifying AT1 cell- and lung-enriched genes. (**B**) UMAP of scRNA-seq data from Mouse Cell Atlas [28] with *Gprc5a*, *Rtkn2* and *Ager*. Blue and grey indicate high and low expression (determined by absolute read count), respectively. (**C**) Gene expression of *GPRC5A*, *RTKN2* and *AGER* in lung (red) and 29 other organs (grey) from GTEx database. Counts per million was used for normalization between samples.

3.4. Immunofluorescence Staining Revealed that GPRC5A is Specifically Expressed in Mouse, Rat and Human AT1 Cells

To confirm GPRC5A protein expression in AT1 cells, immunofluorescence staining was performed. GPRC5A was localized at the surface of AT1 cells in mouse lung (Figure 5A), and double staining with HOPX showed clear co-localization of GPRC5A and HOPX (Figure 5B). Negative controls are shown in Supplemental Figure S7A. Furthermore, double staining with pro-SPC showed that AT2 cells do not express GPRC5A (Figure 5C), indicating that expression of GPRC5A protein is AT1 cell-specific in mouse alveoli. Negative controls are shown in Supplemental Figure S7B. Next, we performed immunofluorescence with human lung tissue. Similar to findings in mouse lung, GPRC5A is clearly localized at the surface of AT1 cells (Figure 6A) while human AT2 cells, which are positive for ABCA3, do not express GPRC5A (Figure 6B). Negative controls are shown in Supplemental Figure S7C. Proximal and distal mouse, human and rat airways did not express GPRC5A above background (Figure 6C and Supplemental Figure S8). AT1 cell-specific staining of GPRC5A was also confirmed by immunofluorescence co-staining with AQP5, pro-SPC and CC10 in cytospins of digested mouse lung (Figure 7). Collectively, these results confirm that GPRC5A protein is specifically expressed on the surface of AT1 cells in both mouse and human lungs.
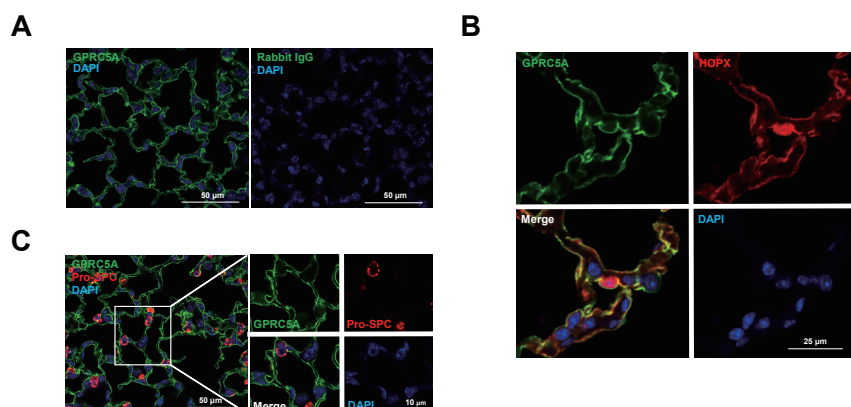
A   **B**

C

Figure 5. GPRC5A expression in mouse distal lung. (**A**) Representative single staining for GPRC5A (green) in mouse distal lung. Blue = DAPI. Rabbit IgG (right panel) was used as negative control. (**B**) Representative double staining for GPRC5A (green) and HOPX (red) in mouse distal lung. Blue = 4′,6-diamidino-2-phenylindole (DAPI). Negative controls are shown in Supplemental Figure S7A. (**C**) Representative double staining for GPRC5A (green) and pro-SPC (red) in mouse distal lung. Blue = DAPI. Negative controls are shown in Supplemental Figure S7B.
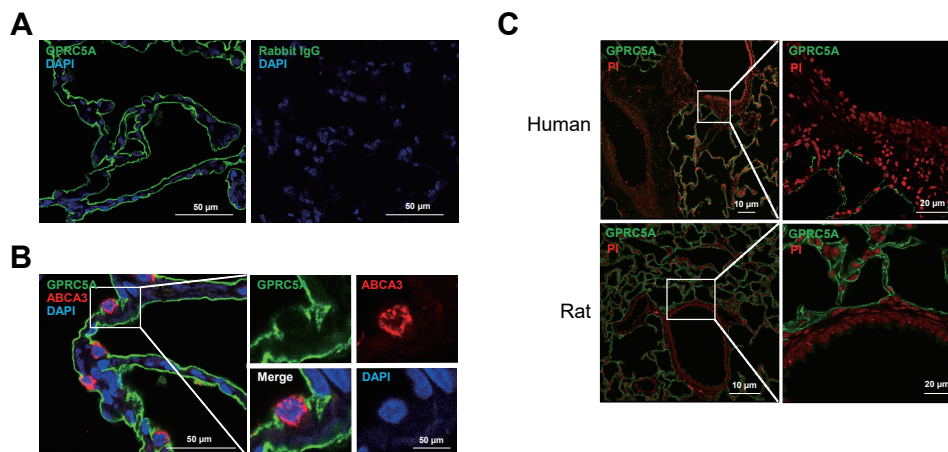
Figure 6. GPRC5A expression in human and rat lung. (**A**) Representative single staining for GPRC5A (green) in human distal lung. Blue = DAPI. Rabbit IgG (right panel) is the negative control. (**B**) Representative double staining for GPRC5A (green) and ABCA3 (red) in human distal lung. Blue = DAPI. Negative controls are showed in Supplemental Figure S7C. (**C**) Representative single staining for GPRC5A (green) in human (upper panel) and rat (lower panel) airway. White box indicates airway. Red = propidium iodide (PI).
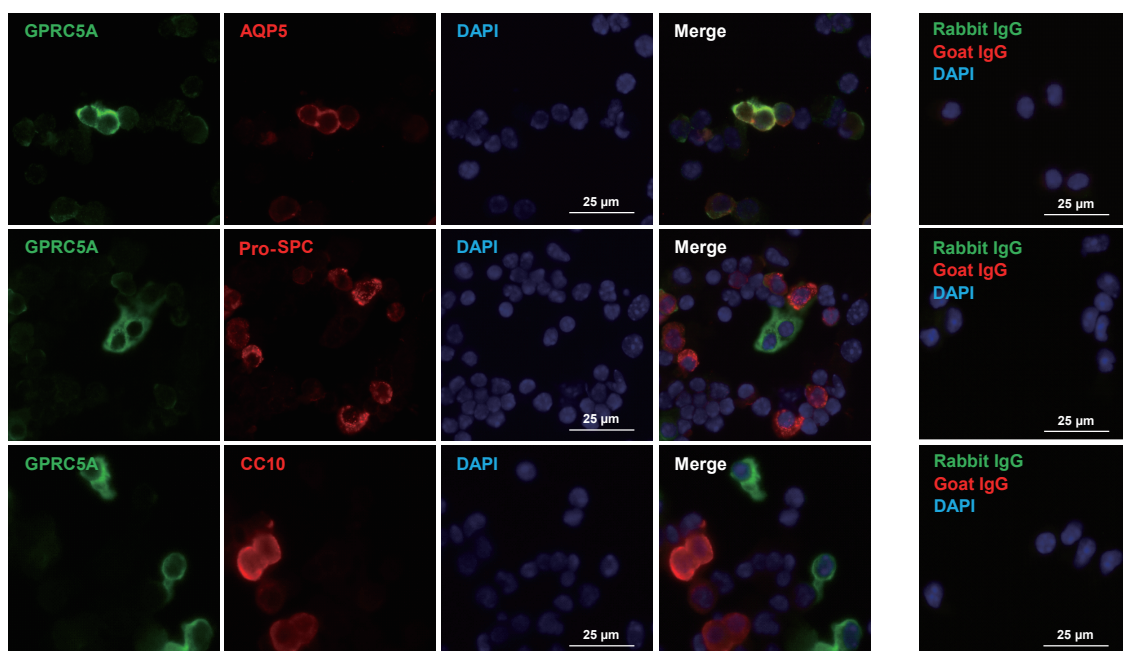
Figure 7. Co-staining of GPRC5A with AQP5, pro-SPC and CC10 in crude cytospin preparations from mouse lung. Representative double staining for GPRC5A (green) and AQP5 (upper panel, red), pro-SPC (middle panel, red) or CC10 (lower panel, red) in cytospins of crude preparations from digested mouse lung tissue. Blue = DAPI. Rabbit and goat IgG were negative controls.

4. Discussion

We successfully performed scRNA-seq of FACS-sorted tdTomato⁺ AT1 cells from *ACID;R26tdTomato* reporter mice and integrated these results with publicly available mouse and human datasets, with the goal of identifying and further characterizing novel AT1 cell genes that could also be used as surface markers to facilitate cell sorting. In our scRNA-seq results, approximately 50% of the 92 cells analyzed expressed conventional AT1 cell markers, while smaller subsets expressed markers of other cell types such as those for AT2, club, basal or ciliated cells, indicating that this approach allowed considerable enrichment for AT1 cells. Among the frequently used 'classical' AT1 cell

markers, we found that *Pdpn*, *Cav1*, *Cav2* and *Aqp5* were also expressed in some cells in other epithelial clusters, and that *Pdpn*, *Cav2*, *Scnn1g*, *Gramd2*, *Igfbp2* and *Aqp5* were heterogeneously expressed within the AT1 cell population at a single-cell level. By integrating our scRNA-seq data with public datasets and our previous in vitro AT2 to AT1 cell differentiation model dataset [10,12,28,30], we overcame the bias of analyzing only Aqp5-positive cells and identified three robust putative AT1 cell-enriched and lung-specific genes: *Rtkn2*, *Ager* and *Gprc5a*. Based on strict selection to fulfill criteria of mouse/human conservation, AT1 cell-enriched expression and membrane localization, we selected *Gprc5a* as the best candidate for further characterization. Consistent with our selection, immunofluorescence staining of human, mouse and rat lung sections validated GPRC5A as a candidate AT1 cell surface marker. It is important to note that neither GPRC5A nor previously identified AT1 markers are uniquely expressed in AT1 cells, and further studies are needed to determine whether this reflects the biological function or origin of AT1 cells. Nevertheless, the identification of a conserved surface marker enriched in the AT1 cell population, such as GPRC5A, offers a novel tool for further study and characterization of AT1 cells.

Within our dataset, airway cell clusters in the *ACID;R26tdTomato* mouse were also positive for *tdTomato* (Figure 3A), with immunofluorescence staining showing that ~10% of CC10[+] cells are tdTomato[+] (Figure 3C). This is consistent with our previous study which revealed that *Aqp5* is also expressed in trachea and in a subset of conducting airway epithelial cells, as shown in *ACID;R26LacZ*[+/+] mice [25]. Accordingly, we found moderate, but not strong, correlation between *Aqp5* and *tdTomato* expression (Figure 3B), suggesting that some *tdTomato*-positive cells do not express *Aqp5* and conversely that not all *Aqp5*-expressing cells express *tdTomato*. Given that Cre is constitutively expressed, expression of *tdTomato* in *Aqp5*[−] cells may be the result of recombination during development with subsequent loss of *Aqp5*.

Regardless, these findings indicate that *ACID;R26tdTomato* mice can be used to enrich for AT1 cells for subsequent downstream analyses by sorting for tdTomato[+] cells, which we will pursue in future studies. Here, we also found that some *tdTomato*[+] cells express AT2 cell markers, and some cells expressing AT2 cell markers also express *Aqp5* (Figure 3A). This could reflect true expression of *Aqp5* in AT2 cells, or the presence of cells intermediate between AT1 and AT2 cells as we and others have previously shown [22,27]. However, in the present study, the number of potential intermediate cells was too small for further characterization.

Despite limitations of the Fluidigm system with regard to capturing only small numbers of cells, a major strength is the greater depth of sequencing compared to other methods. For example, a recent study identified a novel subpopulation of pulmonary neuroendocrine cells showing that a small number of cells can lead to significant insights [42]. This enabled us to demonstrate heterogeneity among AT1 cells with regard to the expression of AT1 cell markers (Supplementary Figure S5B). Interestingly, we found that some well-known AT1 cell markers, such as *Pdpn*, *Igfbp2*, *Cav2* or *Gramd2*, were not uniformly expressed in all AT1 cells (Supplemental Figure S5B), indicating considerable heterogeneity among AT1 cells within this cluster. This is consistent with other published findings of AT1 cell sub-populations and may also reflect the fact that AT1 cell marker expression can change over time [27,43], so cells that are positive for one marker might later become positive/negative for another marker. Further studies and a collaborative effort to analyze and compare datasets generated from different research groups in the field is needed to clarify and fully understand the basis for marker heterogeneity in AT1 cells. Our observation that *Igfbp2* is expressed in only half of the *Aqp5*[+] AT1 cells (Supplementary Figure S5B) is in conflict with previous findings showing that most (95%) *Hopx*[+] AT1 cells isolated from adult (postnatal day 60) *Sftpc-CreER;Rosa26-Zsgreen;Hopx-tdTomato* mice express *Igfbp2* [30]. This discrepancy may be due to a different sorting strategy that led to isolation of different cell populations or strain differences.

In our scRNA-seq data, the percentage of mitochondrial genes is relatively high (Supplementary Figure S3), suggesting that enzymatic digestion and the FACS procedure are still quite harsh for the relatively fragile AT1 cells. Other approaches such as single-nuclei RNA-seq (snRNA-seq), in which

mild digestion can be done to reduce damage to AT1 cell nuclei, may be an alternative option to study gene expression of AT1 cells, as recently reported [44].

The isolation and characterization of AT1 cells represents a major challenge within the pulmonary field, due to both the heterogeneity and lack of specificity of expression of AT1 cell markers, as well as the lack of robust antibodies for their detection. By integrating several datasets, we identified three putative AT1 cell-enriched and lung-specific genes: *Gprc5a*, *Rtkn2* and *Ager*. Among these three putative AT1 markers, *Rtkn2* expression appears as the most 'AT1-specific'. However, while *Gprc5a* and *Ager* are expressed in almost all AT1 cells (Figure 4B), *Rtkn2 is* restricted to only a subset of AT1 cells, and it would not be a useful surface marker for AT1 cell isolation as it encodes for a nuclear protein. Because of its potential usefulness as a surface marker for isolation of AT1 cells, we chose the transmembrane protein GPRC5A for further validation. Although there is low-level expression of *Gprc5a* in mouse (Figure 4B) and human AT2 cells [21,23] at the mRNA level, as well as in bronchial airway epithelial cells [23], we found that at the protein level GPRC5A is very specific for AT1 cells in mouse, rat and human lungs (Figures 5 and 6 and Supplemental Figure S8), despite a previous report of expression in airway epithelial cells [45]. We show that it is localized to the membrane of AT1, but not AT2 or airway cells by immunostaining, supporting its value as a highly promising surface marker for sorting AT1 cells specifically from distal lung cell populations.

Gprc5a was first cloned as a retinoic acid-inducible gene 1 (RAIG1) which has a sequence containing seven transmembrane domains, characteristic of G protein–coupled receptors [46]. *Gprc5a* knockout mice do not show developmental abnormalities or phenotypic changes in the lung, but develop spontaneous lung adenocarcinomas [47]. Deletion of this gene confers susceptibility to endotoxin-induced pulmonary edema and injury [48], indicating that GPRC5A is critical for lung homeostasis and functions as a tumor suppressor. Consistent with this, GPRC5A expression was found to be decreased in patients with chronic obstructive pulmonary disease (COPD) and non-small-cell lung cancer [45]. GPRC5A was shown to transduce NF-κB and STAT3 signaling [49], but the molecular function of GPRC5A, especially in AT1 cells, has not been elucidated to date. Interestingly, we previously showed a role for retinoid X receptor (RXR) signaling in AT2 to AT1 cell differentiation in vitro [10], suggesting a mechanism whereby GPRC5A might be upregulated during this phenotypic transition. Further characterization of the role of GPRC5A in AT1 cells is needed.

In summary, using *ACID;R26tdTomato* mice, we enriched for AT1 cells and successfully performed scRNA-seq, an analysis which has previously been limited by the difficulty in isolating adequate numbers of viable AT1 cells. We show that at the single-cell level, many frequently used AT1 cell markers lack specificity. Heterogeneity in expression of putative AT1 cell markers was observed, emphasizing the need for development of additional AT1 cell (and especially surface) markers for more complete characterization and sorting of the entire population of AT1 cells. Our approach allowed sufficient enrichment of AT1 cells to enable identification and characterization of GPRC5A as a potential novel validated surface marker for AT1 cells. Analysis of greater numbers of AT1 cells is warranted to validate and further characterize subgroups of AT1 cells as well as novel AT1 cell functions.

Author Contributions: Conceptualization, M.H., A.C., M.S., H.K., P.F. and Z.B.; Data curation, M.H., A.C., M.S., B.Z. and Z.B; Formal analysis, M.H., A.K., H.K. and Z.B.; Funding acquisition, Z.B.; Investigation, M.H.; Methodology, M.H., M.S., H.W., Y.L., F.L, T.A.W. and L.H.; Project administration, M.H, Z.B..; Resources, Y.J., H.S., P.F. and Z.B.; Software, F.L.; Supervision, Z.B.; Validation, M.H. and A.C.; Writing—original draft, M.H., A.C. and Z.B.; Writing—review & editing, M.H., A.C., H.K., I.A.O., C.N.M., P.F., B.Z. and Z.B. All authors have read and agreed to the published version of the manuscript.

References

1. Adamson, I.Y.; Bowden, D.H. The type 2 cell as progenitor of alveolar epithelial regeneration. A cytodynamic study in mice after exposure to oxygen. *Lab. Investig.* **1974**, *30*, 35–42. [PubMed]

2. Evans, M.J.; Cabral, L.J.; Stephens, R.J.; Freeman, G. Transformation of alveolar type 2 cells to type 1 cells following exposure to NO_2. *Exp. Mol. Pathol.* **1975**, *22*, 142–150. [CrossRef]

3. Desai, T.J.; Brownfield, D.G.; Krasnow, M.A. Alveolar progenitor and stem cells in lung development, renewal and cancer. *Nature* **2014**, *507*, 190–194. [CrossRef] [PubMed]

4. Barkauskas, C.E.; Cronce, M.J.; Rackley, C.R.; Bowie, E.J.; Keene, D.R.; Stripp, B.R.; Randell, S.H.; Noble, P.W.; Hogan, B.L. Type 2 alveolar cells are stem cells in adult lung. *J. Clin. Investig.* **2013**, *123*, 3025–3036. [CrossRef] [PubMed]

5. Borok, Z.; Liebler, J.M.; Lubman, R.L.; Foster, M.J.; Zhou, B.; Li, X.; Zabski, S.M.; Kim, K.J.; Crandall, E.D. Na transport proteins are expressed by rat alveolar epithelial type I cells. *Am. J. Physiol. Lung Cell Mol. Physiol.* **2002**, *282*, L599–L608. [CrossRef] [PubMed]

6. Flodby, P.; Kim, Y.H.; Beard, L.L.; Gao, D.; Ji, Y.; Kage, H.; Liebler, J.M.; Minoo, P.; Kim, K.J.; Borok, Z.; et al. Knockout mice reveal a major role for alveolar epithelial type I cells in alveolar fluid clearance. *Am. J. Respir. Cell Mol. Biol.* **2016**, *55*, 395–406. [CrossRef]

7. Borok, Z.; Danto, S.I.; Lubman, R.L.; Cao, Y.; Williams, M.C.; Crandall, E.D. Modulation of t1alpha expression with alveolar epithelial cell phenotype in vitro. *Am. J. Physiol.* **1998**, *275*, L155–L164. [CrossRef]

8. Borok, Z.; Lubman, R.L.; Danto, S.I.; Zhang, X.L.; Zabski, S.M.; King, L.S.; Lee, D.M.; Agre, P.; Crandall, E.D. Keratinocyte growth factor modulates alveolar epithelial cell phenotype *in vitro*: Expression of aquaporin 5. *Am. J. Respir. Cell Mol. Biol.* **1998**, *18*, 554–561. [CrossRef]

9. Danto, S.I.; Shannon, J.M.; Borok, Z.; Zabski, S.M.; Crandall, E.D. Reversible transdifferentiation of alveolar epithelial cells. *Am. J. Respir. Cell Mol. Biol.* **1995**, *12*, 497–502. [CrossRef]

10. Marconett, C.N.; Zhou, B.; Rieger, M.E.; Selamat, S.A.; Dubourd, M.; Fang, X.; Lynch, S.K.; Stueve, T.R.; Siegmund, K.D.; Berman, B.P.; et al. Integrated transcriptomic and epigenomic analysis of primary human lung epithelial cell differentiation. *PLoS Genet.* **2013**, *9*, e1003513. [CrossRef]

11. Gonzalez, R.F.; Dobbs, L.G. Isolation and culture of alveolar epithelial Type I and Type II cells from rat lungs. *Methods Mol. Biol.* **2013**, *945*, 145–159. [CrossRef] [PubMed]

12. Marconett, C.N.; Zhou, B.; Sunohara, M.; Pouldar, T.M.; Wang, H.; Liu, Y.; Rieger, M.E.; Tran, E.; Flodby, P.; Siegmund, K.D.; et al. Cross-species transcriptome profiling identifies new alveolar epithelial type I cell-specific genes. *Am. J. Respir. Cell Mol. Biol.* **2017**, *56*, 310–321. [CrossRef] [PubMed]

13. Qiao, R.; Zhou, B.; Liebler, J.M.; Li, X.; Crandall, E.D.; Borok, Z. Identification of three genes of known function expressed by alveolar epithelial type I cells. *Am. J. Respir. Cell Mol. Biol.* **2003**, *29*, 98–105. [CrossRef] [PubMed]

14. Chung, M.I.; Hogan, B.L.M. Ager-CreER(T2): A new genetic tool for studying lung alveolar development, homeostasis, and repair. *Am. J. Respir. Cell Mol. Biol.* **2018**, *59*, 706–712. [CrossRef] [PubMed]

15. Fujino, N.; Kubo, H.; Ota, C.; Suzuki, T.; Suzuki, S.; Yamada, M.; Takahashi, T.; He, M.; Kondo, T.; Yamaya, M. A novel method for isolating individual cellular components from the adult human distal lung. *Am. J. Respir. Cell Mol. Biol.* **2012**, *46*, 422–430. [CrossRef]

16. Zeisel, A.; Munoz-Manchado, A.B.; Codeluppi, S.; Lonnerberg, P.; La Manno, G.; Jureus, A.; Marques, S.; Munguba, H.; He, L.; Betsholtz, C.; et al. Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **2015**, *347*, 1138–1142. [CrossRef]

17. Treutlein, B.; Brownfield, D.G.; Wu, A.R.; Neff, N.F.; Mantalas, G.L.; Espinoza, F.H.; Desai, T.J.; Krasnow, M.A.; Quake, S.R. Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature* **2014**, *509*, 371–375. [CrossRef]

18. Xin, Y.; Kim, J.; Ni, M.; Wei, Y.; Okamoto, H.; Lee, J.; Adler, C.; Cavino, K.; Murphy, A.J.; Yancopoulos, G.D.; et al. Use of the Fluidigm C1 platform for RNA sequencing of single mouse pancreatic islet cells. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 3293–3298. [CrossRef]

19. Macosko, E.Z.; Basu, A.; Satija, R.; Nemesh, J.; Shekhar, K.; Goldman, M.; Tirosh, I.; Bialas, A.R.; Kamitaki, N.; Martersteck, E.M.; et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **2015**, *161*, 1202–1214. [CrossRef]

20. Xu, Y.; Mizuno, T.; Sridharan, A.; Du, Y.; Guo, M.; Tang, J.; Wikenheiser-Brokamp, K.A.; Perl, A.T.; Funari, V.A.; Gokey, J.J.; et al. Single-cell RNA sequencing identifies diverse roles of epithelial cells in idiopathic pulmonary fibrosis. *JCI Insight* **2016**, *1*, e90558. [CrossRef]

21. Reyfman, P.A.; Walter, J.M.; Joshi, N.; Anekalla, K.R.; McQuattie-Pimentel, A.C.; Chiu, S.; Fernandez, R.; Akbarpour, M.; Chen, C.I.; Ren, Z.; et al. Single-cell transcriptomic analysis of human lung provides insights into the pathobiology of pulmonary fibrosis. *Am. J. Respir. Crit. Care Med.* **2019**, *199*, 1517–1536. [CrossRef] [PubMed]

22. Riemondy, K.A.; Jansing, N.L.; Jiang, P.; Redente, E.F.; Gillen, A.E.; Fu, R.; Miller, A.J.; Spence, J.R.; Gerber, A.N.; Hesselberth, J.R.; et al. Single cell RNA sequencing identifies TGFbeta as a key regenerative cue following LPS-induced lung injury. *JCI Insight* **2019**, *5*. [CrossRef]

23. Raredon, M.S.B.; Adams, T.S.; Suhail, Y.; Schupp, J.C.; Poli, S.; Neumark, N.; Leiby, K.L.; Greaney, A.M.; Yuan, Y.; Horien, C.; et al. Single-cell connectomic analysis of adult mammalian lungs. *Sci. Adv.* **2019**, *5*, eaaw3851. [CrossRef] [PubMed]

24. Borok, Z.; Li, X.; Fernandes, V.F.; Zhou, B.; Ann, D.K.; Crandall, E.D. Differential regulation of rat aquaporin-5 promoter/enhancer activities in lung and salivary epithelial cells. *J. Biol. Chem.* **2000**, *275*, 26507–26514. [CrossRef]

25. Flodby, P.; Borok, Z.; Banfalvi, A.; Zhou, B.; Gao, D.; Minoo, P.; Ann, D.K.; Morrisey, E.E.; Crandall, E.D. Directed expression of Cre in alveolar epithelial type 1 cells. *Am. J. Respir. Cell Mol. Biol.* **2010**, *43*, 173–178. [CrossRef]

26. Krane, C.M.; Fortner, C.N.; Hand, A.R.; McGraw, D.W.; Lorenz, J.N.; Wert, S.E.; Towne, J.E.; Paul, R.J.; Whitsett, J.A.; Menon, A.G. Aquaporin 5-deficient mouse lungs are hyperresponsive to cholinergic stimulation. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 14114–14119. [CrossRef]

27. Liebler, J.M.; Marconett, C.N.; Juul, N.; Wang, H.; Liu, Y.; Flodby, P.; Laird-Offringa, I.A.; Minoo, P.; Zhou, B. Combinations of differentiation markers distinguish subpopulations of alveolar epithelial cells in adult lung. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **2016**, *310*, L114–L120. [CrossRef]

28. Han, X.; Wang, R.; Zhou, Y.; Fei, L.; Sun, H.; Lai, S.; Saadatpour, A.; Zhou, Z.; Chen, H.; Ye, F.; et al. Mapping the Mouse Cell Atlas by microwell-Seq. *Cell* **2018**, *172*, 1091–1107.e17. [CrossRef]

29. LungMAP Consortium [U01HL122642]. Single-cell RNA-seq Data. Available online: https://www.lungmap.net/ (accessed on 22 January 2019).

30. Wang, Y.; Tang, Z.; Huang, H.; Li, J.; Wang, Z.; Yu, Y.; Zhang, C.; Dai, H.; Wang, F.; Cai, T.; et al. Pulmonary alveolar type I cell population consists of two distinct subtypes that differ in cell fate. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 2407–2412. [CrossRef]

31. GTExPortal. Available online: https://gtexportal.org/home/ (accessed on 22 January 2019).

32. Arenkiel, B.R.; Hasegawa, H.; Yi, J.J.; Larsen, R.S.; Wallace, M.L.; Philpot, B.D.; Wang, F.; Ehlers, M.D. Activity-induced remodeling of olfactory bulb microcircuits revealed by monosynaptic tracing. *PLoS ONE* **2011**, *6*, e29423. [CrossRef]

33. Ewels, P.; Magnusson, M.; Lundin, S.; Kaller, M. MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **2016**, *32*, 3047–3048. [CrossRef] [PubMed]

34. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **2015**, *12*, 357–360. [CrossRef] [PubMed]

35. Liao, Y.; Smyth, G.K.; Shi, W. featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **2014**, *30*, 923–930. [CrossRef] [PubMed]

36. Satija, R.; Farrell, J.A.; Gennert, D.; Schier, A.F.; Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **2015**, *33*, 495–502. [CrossRef] [PubMed]

37. Butler, A.; Hoffman, P.; Smibert, P.; Papalexi, E.; Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **2018**, *36*, 411–420. [CrossRef] [PubMed]

38. Uhlen, M.; Fagerberg, L.; Hallstrom, B.M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; Sivertsson, A.; Kampf, C.; Sjostedt, E.; Asplund, A.; et al. Proteomics. Tissue-based map of the human proteome. *Science* **2015**, *347*, 1260419. [CrossRef] [PubMed]

39. Plasschaert, L.W.; Zilionis, R.; Choo-Wing, R.; Savova, V.; Knehr, J.; Roma, G.; Klein, A.M.; Jaffe, A.B. A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. *Nature* **2018**, *560*, 377–381. [CrossRef]

40. Muzumdar, M.D.; Tasic, B.; Miyamichi, K.; Li, L.; Luo, L. A global double-fluorescent Cre reporter mouse. *Genesis* **2007**, *45*, 593–605. [CrossRef]

41. Ardini-Poleske, M.E.; Clark, R.F.; Ansong, C.; Carson, J.P.; Corley, R.A.; Deutsch, G.H.; Hagood, J.S.; Kaminski, N.; Mariani, T.J.; Potter, S.S.; et al. LungMAP: The molecular atlas of lung development program. *Am. J. Physiol. Lung Cell Mol. Physiol.* **2017**, *313*, L733–L740. [CrossRef]

42. Ouadah, Y.; Rojas, E.R.; Riordan, D.P.; Capostagno, S.; Kuo, C.S.; Krasnow, M.A. Rare pulmonary neuroendocrine cells are stem cells regulated by Rb, p53, and Notch. *Cell* **2019**, *179*, 403–416.e23. [CrossRef]

43. Jansing, N.L.; McClendon, J.; Henson, P.M.; Tuder, R.M.; Hyde, D.M.; Zemans, R.L. Unbiased quantitation of alveolar Type II to alveolar type I cell transdifferentiation during repair after lung injury in mice. *Am. J. Respir. Cell Mol. Biol.* **2017**, *57*, 519–526. [CrossRef]

44. Koenitzer, J.R.; Wu, H.; Atkinson, J.J.; Brody, S.L.; Humphreys, B.D. Single nucleus RNASeq profiling of mouse lung: Reduced dissociation bias and improved rare cell type detection compared with single cell RNASeq. *Am. J. Respir. Cell Mol. Biol.* **2020**. [CrossRef] [PubMed]

45. Fujimoto, J.; Kadara, H.; Garcia, M.M.; Kabbout, M.; Behrens, C.; Liu, D.D.; Lee, J.J.; Solis, L.M.; Kim, E.S.; Kalhor, N.; et al. G-protein coupled receptor family C, group 5, member A (GPRC5A) expression is decreased in the adjacent field and normal bronchial epithelia of patients with chronic obstructive pulmonary disease and non-small-cell lung cancer. *J. Thorac. Oncol.* **2012**, *7*, 1747–1754. [CrossRef] [PubMed]

46. Cheng, Y.; Lotan, R. Molecular cloning and characterization of a novel retinoic acid-inducible gene that encodes a putative G protein-coupled receptor. *J. Biol. Chem.* **1998**, *273*, 35008–35015. [CrossRef] [PubMed]

47. Tao, Q.; Fujimoto, J.; Men, T.; Ye, X.; Deng, J.; Lacroix, L.; Clifford, J.L.; Mao, L.; Van Pelt, C.S.; Lee, J.J.; et al. Identification of the retinoic acid-inducible Gprc5a as a new lung tumor suppressor gene. *J. Natl. Cancer Inst.* **2007**, *99*, 1668–1682. [CrossRef] [PubMed]

48. Liao, Y.; Song, H.; Xu, D.; Jiao, H.; Yao, F.; Liu, J.; Wu, Y.; Zhong, S.; Yin, H.; Liu, S.; et al. Gprc5a-deficiency confers susceptibility to endotoxin-induced acute lung injury via NF-kappaB pathway. *Cell Cycle* **2015**, *14*, 1403–1412. [CrossRef]

49. Chen, Y.; Deng, J.; Fujimoto, J.; Kadara, H.; Men, T.; Lotan, D.; Lotan, R. Gprc5a deletion enhances the transformed phenotype in normal and malignant lung epithelial cells by eliciting persistent Stat3 signaling induced by autocrine leukemia inhibitory factor. *Cancer Res.* **2010**, *70*, 8917–8926. [CrossRef]