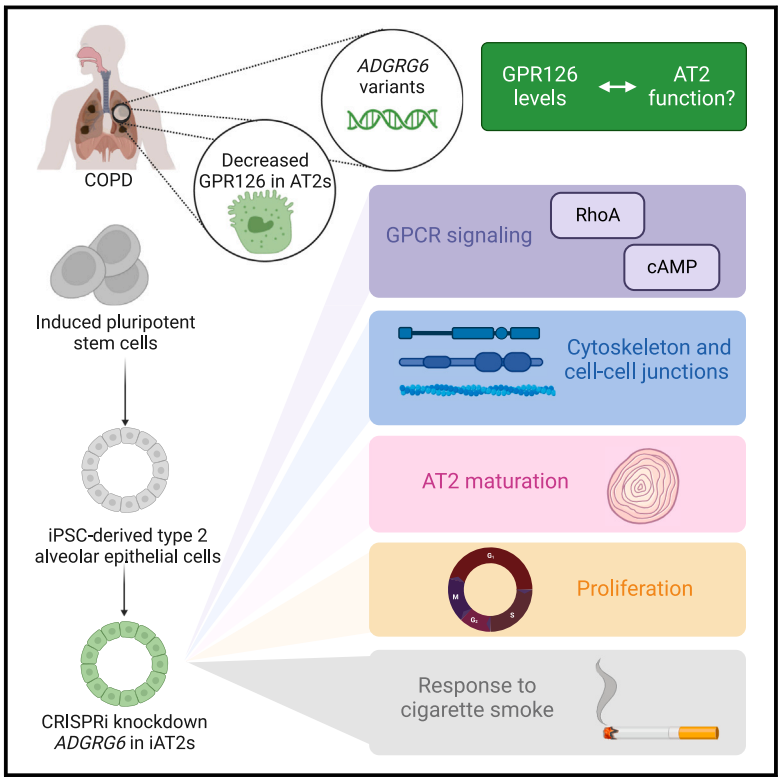
The COPD GWAS gene *ADGRG6* instructs function and injury response in human iPSC-derived type II alveolar epithelial cells

Graphical abstract



Authors

Rhiannon B. Werder, Kayleigh A. Berthiaume, Carly Merritt, ..., Xiaobo Zhou, Michael H. Cho, Andrew A. Wilson

Correspondence

awilson@bu.edu

***ADGRG6* variants are implicated by GWASs in chronic obstructive pulmonary disease pathogenesis. Here, we investigated the function of *ADGRG6* using a CRISPR-interference system in induced pluripotent stem cell-derived type 2 alveolar epithelial cells. *ADGRG6* exerts pleiotropic effects on cell adhesion, cytoskeletal arrangement, proliferation, and response to cigarette smoke injury.**

Werder et al., 2023, The American Journal of Human Genetics *110*, 1735–1749
October 5, 2023
https://doi.org/10.1016/j.ajhg.2023.08.017

CellPress

The COPD GWAS gene *ADGRG6* instructs function and injury response in human iPSC-derived type II alveolar epithelial cells

Rhiannon B. Werder,[1,2,3] Kayleigh A. Berthiaume,[1] Carly Merritt,[1,2] Marissa Gallagher,[1] Carlos Villacorta-Martin,[1] Feiya Wang,[1] Pushpinder Bawa,[1] Vidhi Malik,[4] Shawn M. Lyons,[5] Maria C. Basil,[6] Edward E. Morrisey,[6] Darrell N. Kotton,[1,2] Xiaobo Zhou,[4] Michael H. Cho,[4] and Andrew A. Wilson[1,2,*]

Summary

Emphysema and chronic obstructive pulmonary disease (COPD) most commonly result from the effects of environmental exposures in genetically susceptible individuals. Genome-wide association studies have implicated *ADGRG6* in COPD and reduced lung function, and a limited number of studies have examined the role of *ADGRG6* in cells representative of the airway. However, the *ADGRG6* locus is also associated with DLCO/VA, an indicator of gas exchange efficiency and alveolar function. Here, we sought to evaluate the mechanistic contributions of *ADGRG6* to homeostatic function and disease in type 2 alveolar epithelial cells. We applied an inducible CRISPR interference (CRISPRi) human induced pluripotent stem cell (iPSC) platform to explore *ADGRG6* function in iPSC-derived AT2s (iAT2s). We demonstrate that *ADGRG6* exerts pleiotropic effects on iAT2s including regulation of focal adhesions, cytoskeleton, tight junctions, and proliferation. Moreover, we find that *ADGRG6* knockdown in cigarette smoke-exposed iAT2s alters cellular responses to injury, downregulating apical complexes in favor of proliferation. Our work functionally characterizes the COPD GWAS gene *ADGRG6* in human alveolar epithelium.

Introduction

Chronic obstructive pulmonary disease (COPD) is the third-leading cause of death globally and manifests as both airways disease and emphysema resulting from damage to the airways and alveoli, respectively.[1] Emphysema results from loss of total alveolar surface area with associated impairment of gas exchange which can be functionally quantified through measurement of diffusing capacity of the lung for carbon monoxide (DLCO).[2] The contribution of environmental exposures, such as cigarette smoke, to emphysema pathogenesis is well established. While the role of genetics in determining susceptibility to injury from those exposures has been less well characterized, mounting evidence demonstrates that COPD is heritable[3,4] and in some cases occurs in the absence of known environmental exposures.[5,6] Genome-wide association studies (GWASs) have implicated a number of genetic loci involved in gas exchange and COPD which overlap with loci associated with population-based lung function[7–13] and lung development,[7,9,14] further supporting a genetic basis for COPD susceptibility. Using fine mapping, we and others recently identified variants (both coding and non-coding variants, summarized in Table S1) in *ADGRG6* that are associated with COPD, reduced lung function, and impaired alveolar gas exchange (measured by DLCO per alveolar volume [VA], i.e., DLCO/VA).[10–13]

Adhesion G-protein coupled receptor (aGPCR) 126 (GPR126), encoded by *ADGRG6* and activated by a number of ligands including laminin-211 and collagen IV,[15,16] plays vital roles in peripheral nerve development and angiogenesis.[16–19] When GPR126 ligands interact with the receptor, the extracellular region of GPR126 is autoproteolysed to expose the GAIN domain which in turn binds to the 7 transmembrane domain of GPR126 to activate the G protein-signaling cascade.[20] Previous work has shown that GPR126 activation in human smooth muscle cells or bronchial epithelial cells modulates intracellular cAMP.[21] However, the specific cell type through which GPR126 affects COPD pathogenesis is unknown. COPD GWAS signals are enriched in regulatory regions for human type 2 alveolar epithelial cells (AT2)[22] but the role of *ADGRG6* in AT2s, the facultative progenitor cells of the lung compartment primarily injured in emphysema, has yet to be fully explored.

Advances in understanding the biology of AT2s have been limited by challenges in accessing them, maintaining them in culture, and manipulating their genome. To address these hurdles, we and others have applied induced pluripotent stem cell (iPSC)-derived type 2 alveolar epithelial cells (iAT2s) that provide an inexhaustible supply of

[1]Center for Regenerative Medicine of Boston University and Boston Medical Center, Boston, MA 02118, USA; [2]The Pulmonary Center and Department of Medicine, Boston University School of Medicine, Boston, MA 02118, USA; [3]Murdoch Children's Research Institute, Melbourne, VIC 3052, Australia; [4]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA; [5]Biochemistry Department, Boston University School of Medicine, Boston, MA 02118, USA; [6]School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

*Correspondence: awilson@bu.edu

https://doi.org/10.1016/j.ajhg.2023.08.017.

disease-relevant cells.[23–27] iAT2s express lamellar bodies, overlap transcriptomically with primary human AT2s, recapitulate disease-specific phenotypes, and can be used to model environmental exposures.[23–26,28] We recently developed an inducible CRISPR interference (CRISPRi) iPSC system to interrogate the contribution of COPD GWAS genes to iAT2 function.[27] In this study, we applied CRISPRi to characterize the functional relevance of *ADGRG6* to iAT2 phenotypes. We find that GPR126 regulates iAT2 cytoskeletal organization, canonical GPCR signaling pathways, proliferation, and the cellular response to cigarette smoke exposure.

Material and methods

Maintenance of PSC lines

PSCs were maintained in feeder-free conditions in mTeSR1 medium (StemCell Technologies) on growth factor reduced Matrigel (Corning)-coated plates. Gentle Cell dissociation reagent (StemCell Technologies) was used for routine passaging. CRISPRi-iPSCs were generated as previously described,[27] targeting a doxycycline-inducible CRISPRi construct to the AAVS1 locus. All iPSCs displayed a normal karyotype when analyzed by G-banding (Cell Line Genetics). Detailed protocols relating to iPSC derivation and culture are available to download at https://crem.bu.edu/cores-protocols/#protocols and iPSCs used in this manuscript are available upon request from the CReM iPSC repository at https://stemcellbank.bu.edu. All experiments involving human iPSC lines were conducted with informed consent and approval from the Institutional Review Board of Boston University (protocol H33122).

Directed differentiation of iAT2s and CRISPR-mediated knockdown

Human iPSCs underwent directed differentiation to generate type 2 alveolar epithelial cells (iAT2s), as we have previously described.[24] iPSCs were differentiated to definitive endoderm ($CXCR4^+cKit^+$) using the STEMdif Definitive Endoderm Kit (StemCell Technologies), then dissociated with Gentle Cell dissociation reagent and replated on growth factor reduced Matrigel (Corning)-coated plates. Cells were then cultured in anteriorization media consisting of complete serum-free differentiation medium (cSFDM) base media, supplemented with 2 mM Dorsomorphin (Stemgent) and 10 mM SB431542 (Tocris), and for the first 24 h with 10 mM Y-27632 ("Y"; Tocris). After three days, cells were then cultured in cSFDM supplemented with 3 mM CHIR99021 (Tocris), 10 ng/mL recombinant human BMP4 (R&D Systems), and 100 nM retinoic acid (Sigma). On day 14–15 of the differentiation, cells were dissociated with 0.05% trypsin and sorted for $NKX2\text{-}1^+$ lung progenitors (based on NKX2-1-GFP) on a MoFlo Astrios (described in FACS methods). After sorting, $NKX2\text{-}1^+$ cells were resuspended in growth factor reduced Matrigel (Corning) to generate droplets which when solidified were covered with 3 mm CHIR99021, 10 ng/mL rhKGF (R&D Systems), 50 nM dexamethasone (Sigma), 0.1 mM 8-Bromoadenosine 3′,5′ cyclic monophosphate sodium salt (Sigma), and 0.1 mM 3-Isobutyl-1-methylxanthine (IBMX; Sigma) in cSFDM (called CK+DCI media). iAT2s were then maintained in CK+DCI, feeding every two days, and in serially passaged every two weeks, as described.[29] Following single cell dissociation, iAT2s were maintained in CK+DCI+Y-27632 for 3 days, then switched to CK+DCI media. Cell counts were taken from triplicate wells at each passage to calculate cell yield. Unless specified, experiments characterizing dissociated iAT2s or alveolospheres were performed at 7 days post passage. To maintain purity of the cultures, iAT2s were also sorted for NKX2-1-GFP and/or SFTPC-tdTomato as necessary.

To knockdown *ADGRG6* in iAT2s, gRNA were designed to target the TSS (within −50 to +200 bp) using the online gRNA design tool crispor.org, or control, non-targeting gRNA.[27,30] gRNA were cloned into a lentiviral vector (pLV-GG-hUbC-EBFP2) (deposited at Addgene).[27] iAT2s were transduced with lentivirus at an MOI 20, then sorted for EBFP2, as described.[27] To initiate gene knockdown, iAT2s were cultured in CK+DCI containing 2 μM doxycycline (Sigma). Functional knockdown (20%–50% reduction in mRNA) was assessed by qPCR.

In some experiments, iAT2s were plated on Transwells to establish air-liquid interface (ALI) cultures, as described.[26,29] In brief, iAT2s were dissociated to single cells and 200,000 cells were then plated on growth-factor reduced Matrigel coated Costar 6.5 mm Clear Transwells (Millipore-Sigma) in CK+DCI+Y. After two days, the apical surface was aspirated ("air-lift"). Basolateral media was changed to CK+DCI three days after initial plating and refreshed every two days after. *Trans*-epithelial electrical resistance (TEER) was measured using a Millicell ERS-2 Voltohmmeter (Millipore, MERS00002).

Reverse transcription quantitative PCR (qRT-PCR)

iAT2s were collected and stored in Qiazol (Qiagen). RNA extraction was performed using the RNeasy Plus Mini Kit (Qiagen). To generate complementary DNA (cDNA), the MultiScribe Reverse Transcriptase kit was used (Applied Biosystems). qRT-PCR was performed with predesigned Taqman probes (Applied Biosystems) for 40 cycles. The Ct value for technical triplicate wells was calculated and normalized to internal 18S, then fold change was calculated relative to control cells using $2^{-\Delta\ \Delta Ct}$. The following Taqman probes (Thermo Fisher) were used in this study: *ACTB,* Hs01060665_g1; *ADGRG6,* Hs01089210_m1; *AURKB,* Hs00945858_g1; *CDK2,* Hs01548894_m1; *CCNA2,* Hs00996788_m1; *CDC20,* Hs00426680_mH; *CDT1,* Hs00925491_g1; *CDC6,* Hs00154374_m1; *E2F1,* Hs00153451_m1; *ITGA2,* Hs00158127_m1; *ITGB1,* Hs01127536_m1; *MYC,* Hs00153408_m1; *TOP2A,* Hs01032137_m1.

Flow cytometry and sorting

To sort $NKX2\text{-}1\text{-}GFP^+$ lung progenitors, single cells were resuspended in sort buffer (Hank's Balanced Salt Solution [ThermoFisher], 2% FBS, 10 μm Y-27632, and 10 μM calcein blue AM [Life Technologies]). To purify iAT2s, single cells were resuspended in sort buffer and sorted based on NKX2-1-GFP and SFTPC-tdTomato. Cells were sorted on a MoFlo Astrios EQ (Beckman Coulter) at the Boston University Flow Cytometry Core Facility (FCCF). Proliferation was measured by flow cytometry using the Click-iT EdU Alexa Fluor 647 Flow Cytometry Assay Kit (C10424, Thermo Fisher), as per manufacturer's instructions.

Immunohistochemistry

COPD and control primary human lung tissue analyzed in this study were collected through an established protocol ("PROPEL") with informed consent and approval by the University of Pennsylvania IRB.[31] Healthy (47-year-old female) or COPD (66-year-old female) human lung samples were fixed and embedded in paraffin. iAT2 alveolospheres were fixed in 4% PFA for 20 min at 37°C. Cells

were permeabilized with 0.3% Triton and blocked with 4% normal donkey serum (NDS). Samples were incubated overnight with primary antibodies, diluted in 4% NDS, and incubated overnight at 4°C. The following primary antibodies were used in this study: GPR126 (ThermoFisher, BS-12027R), pro-SFTPC (Santa Cruz, sc518029), ZO-1 (Invitrogen, C#61-7300), ITGB1 (Invitrogen, C#PA5-29606), and ITGA2 (eBioscience, eBioY418). Samples were then washed with 0.05% Tween and incubated with fluorescently conjugated secondary antibodies for 1 h at room temperature. The following secondary antibodies were used in this study: Goat anti-Rabbit AF488, anti-rabbit AF647, and anti-Mouse AF488 (Invitrogen #A-11008, #A-21121, #A-21245). Cells were counterstained with Hoechst 33342. To stain F-actin, phalloidin (Invitrogen, C#A22287) was added to iAT2s, then immediately counterstained with Hoechst. Cells were imaged with a Leica SP5 confocal microscope. Images were processed and staining intensity quantified using ImageJ Fiji software (https://imagej.net/software/fiji).

Western blot

Cells were lysed in 20 mM HEPES-KOH (pH 7.4), 8 M Urea. Lysates were quantified using Bradford reagent (Bio-Rad). SDS-PAGE was performed using 10 μg of cell lysate and transferred to nitrocellulose membrane using TransBlot Turbo system (Bio-Rad). Membranes were blocked in 5% milk/TBST. Blots were incubated with primary antibodies overnight. The following antibodies were used in this study: Total 4EBP1 (Cell Signaling Technology #9644), Non-Phospho-4EBP1 (Thr46) (Cell Signaling Technology #4923), total Rb (Cell Signaling Technology #9309), Phospho-Rb (Cell Signaling Technology #8516), and GAPDH (Invitrogen AM4300). Blots were washed 3× with TBST and incubated for 1 h with appropriate secondary antibodies. The following secondaries were used: Immunstar anti-Rabbit-HRP (Bio-Rad #1705046), Immunstar anti-Mouse-HRP (Bio-Rad #1705047). Blots were washed 3× with TBST and visualized on ChemiDoc (Bio-Rad) using Clarity ECL detection reagent. The Chemiluminescent signal was quantified using Imagelab (Bio-Rad).

Single-cell RNA sequencing

iAT2 alveolospheres (D184, 7 days post passage) were dissociated to single cells and incubated with Fc block (Biolegend, #422301) prior to staining with hashing antibodies (Biolegend, #394631 and #394633). Live cells were sorted on a MoFlo Astrios EQ (Beckman Coulter) by Zombie NIR exclusion (Biolegend, #423106) and hashed cells were pooled 1:1 prior to capture. Libraries were prepared as per the 10× Genomics scRNA-Seq 3′v3.1 and hashtag oligonucleotide (HTO) protocols and sequenced using an Illumina NextSeq 2000 instrument, pooled 50:1. Sequencing generated reads with 94% $\geq$ Q30. Cellranger 3.0.2 pipeline was used to generate fastq files and the count matrices (combining gene expression and antibody capture libraries for each sample). Seurat (v.4.0.1) was used for further data processing and analysis.[32] To demultiplex the samples based on hashing antibody capture, the HTODemux function was used.[33] Each cell was classified as positive or negative for each HTO. Cells negative for an HTO, cells that were positive for more than one HTO (annotated as doublets), cells with more than 15% of Unique Molecular Identifiers mapping to mitochondrial genes, or cells with fewer than 800 genes detected were filtered out. Data were normalized using the regularized negative binomial regression method (SCTransform function), and cell degradation (mitochondrial percentage) was regressed out.[34] Principal components analysis (PCA) on the sparse expression matrix was performed and uniform manifold projection (UMAP) on the top 20 principal components. Louvain algorithm at different resolutions (0.05–1.5) was used to compute clustering. Cell cycle stage was calculated using the scoring method from Kowalczyk et al.,[35] as implemented in Seurat. Differential gene expression was determined by a log fold change of 0.25 with a Wilcoxon rank-sum test and gene set enrichment analysis (GSEA) was performed using hypeR.[36] Data are deposited at GEO: GSE223078.

Oxygen consumption rate

Oxygen consumption rate (OCR) was measured using the Extracellular Flux Analyzer (Seahorse Bioscience). Single cells were plated on XFe96 Seahorse plate (coated with 2D Matrigel) at a density of 50,000 cells per well in CKDCI+Y ± dox. 24 h later, media was changed to Agilent Seahorse XF Base Medium (Agilent Technologies) supplemented with 25 mM glucose and 10 mM pyruvate. Oligomycin (2 μM; Agilent), carbonyl cyanide-4-(trifluoromethoxy) phenylhydrazone (FCCP; 1.5 μM; Agilent), and antimycin A plus rotenone (both 1 μM; Agilent) were added over the course of the assay. Upon completion of the assay, cells were lysed and total protein was measured by BCA assay (Pierce; #23225), and then OCR and Extracellular Acidification Rate (ECAR) were normalized to total protein.

Cigarette smoke exposure

3R4F reference cigarettes (University of Kentucky) were preconditioned in a temperature- and humidity-controlled chamber for 48 h (as per ISO 3402). iAT2s (D183) were plated at air-liquid interface (ALI) for 6 days then exposed to gas-phase combustible cigarette smoke using a Vitrocell VC1 exposure system (Vitrocell Systems), using the ISO 3308 protocol, as we have previously described.[26,27] Cigarette smoke was diluted with humidified room air to expose cultures to 5% (v/v) smoke. Every 60 s, 35-mL puffs were drawn for 2 s, with a total of 32 puffs per exposure. For bulk RNA sequencing, samples were collected 8 h post exposure. RNA was collected directly in Trizol. To assess colony-forming efficiency (CFE) after smoke exposure, ALIs were dissociated with Accutase (A6964, Millipore-Sigma), replated as single cells in Matrigel (400 cells/uL) and imaged and quantified 14 days later.

Bulk RNA sequencing

RNA was isolated using the RNeasy Mini kit (Qiagen), including on-column DNA digestion with the RNAse-free DNAse set (Qiagen). Total mRNA was isolated using magnetic bead-based poly(A) selection then fragmented, prior to cDNA generation by reverse transcription. cDNA was end-paired, ligated to Illumina sequencing adapters, and amplified to generate libraries. Pooled libraries were sequenced using a NextSeq 2000 (Illumina). Raw data quality was assessed using FastQC (https://github.com/s-andrews/FastQC). Sequence reads were aligned to the human reference genome (GRCh38) and counts per gene were summarized using the featureCounts function from the subread package. The edgeR package was used to import and filter the counts (https://github.com/OliverVoogd/edgeR). After exploratory analysis with principal component analysis (PCA), differential expression testing was conducted to compare the gene expression between different sample sets. The limma package with its voom method, namely, linear modeling and empirical Bayes moderation, was used to

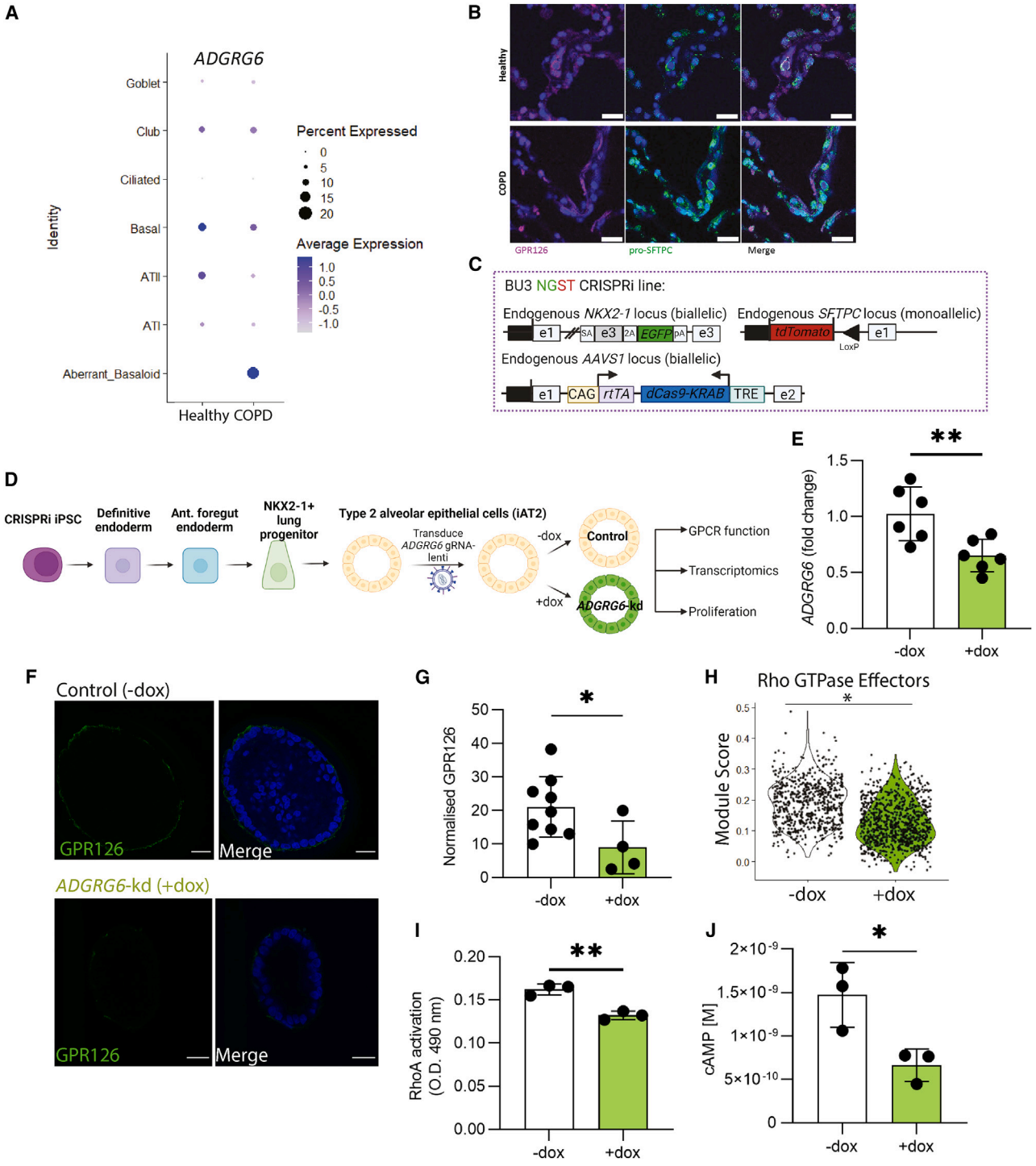


Figure 1. GPR126 knockdown attenuates GPCR signaling in iAT2s

(A) *ADGRG6* expression in healthy and COPD lungs.

(B) GPR126 levels in healthy and COPD lungs. Nuclei, blue; GPR126, magenta; pro-SFTPC, green; scale bar, 50 μm.

(C) BU3 NGST CRISPRi iPSCs contain a GFP reporter knocked into the *NKX2-1* locus, tdTomato reporter knocked into the *SFTPC* locus, and doxycycline-inducible dCas9-KRAB knocked into the *AAVS1* locus.

(D) Directed differentiation of BU3 NGST CRISPRi-targeted iPSCs to type 2 alveolar epithelial cells (iAT2s). CRISPRi-iAT2s are transduced with gRNA targeting the transcriptional start site of *ADGRG6* and sorted for transduced cells, prior to treatment with doxycycline to initiate *ADGRG6* knockdown.

(E) *ADGRG6* expression in iAT2s, measured by qRT-PCR.

(F) GPR126 (green) production in iAT2s. Nuclei (blue), scale bar, 20 μm.

(legend continued on next page)

test differential expression (moderate t test) (https://github.com/cran/limma). p values were adjusted for multiple testing using Benjamini-Hochberg correction (false discovery rate-adjusted p value; FDR). Differentially expressed genes for each comparison were visualized using Glimma package (https://github.com/Shians/Glimma), and FDR < 0.05 was set as the threshold for determining significant differential gene expression. Enrichment analyses were carried out using Fgsea package (https://github.com/ctlab/fgsea). Gene sets used for enrichment analysis were Hallmark enrichment pathways.[37] Data are deposited at GEO: GSE223077.

Measurement of GPCR function

RhoA and cAMP were measured in iAT2s, 7 days post passage. RhoA activity was measured using the RhoA G-LISA Activation Assay Kit (BK124, Cytoskeleton), with protein normalized prior to running the assay, as per manufacturer's instructions. Intracellular cAMP was measured using the LANCE Ultra cAMP Detection Kit (TRF0262, PerkinElmer). iAT2s were dissociated to single cells, then 5,000 cells per well were assayed, as per manufacturer's instructions.

CRISPR editing of rs17280293

To disrupt the rs17280293 locus (GenBank: NM_020455.6, c.367A>G [p.Ser123Gly]), we used a ribonucleoprotein (RNP) complex, consisting of HiFi Cas9 Nuclease V3 (104 pmol) and gRNA (120 pmol, 5′-TCGCACTTGAGTTAAATGAT-3′) (IDT). iAT2s were dissociated to single cells and nucleofected (program EA104) with RNP complex and Cas9 electroporation enhancer (0.2 μM) in P3 solution (Lonza). Cells were immediately replated in Matrigel and allowed to form alveolospheres (over 2 weeks). Clonal alveolospheres were then individually isolated manually and allowed to expand over a further 4 weeks, prior to analysis by Sanger Sequencing and downstream analyses.

Statistics

Data are represented as means, with error bars representing standard deviation (SD). To compare two groups, unpaired two-tailed Student's t tests were used. To compare three or more groups, one-way analysis of variance (ANOVA) with Tukey multiple comparisons test were used. Details of tests used are provided in the figure legends. A p value of <0.05 was determined to be statistically significant. p value annotations on graphs were as follows: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

Results

ADGRG6 expression in type 2 alveolar epithelial cells is decreased in COPD

Previous studies of whole lung tissue have reported discrepant expression levels of *ADGRG6* in COPD-affected individuals.[12,21] To profile epithelial cell-specific expression of *ADGRG6,* we analyzed previously published single-cell RNA sequencing (scRNA-seq) of 17 individuals with advanced COPD and 15 age-matched control subjects (GEO: GSE136831).[22] *ADGRG6* was broadly expressed in airway and alveolar epithelial cells, including club cells, basal cells, aberrant basaloid cells, AT1s, and AT2s (Figure 1A). *ADGRG6* expression was decreased in basal cells and AT2s in people with COPD but was expressed in aberrant basaloid cells, a population present in a subset of COPD donors but absent in healthy tissue (Figure 1A). To validate these findings at the protein level, we analyzed explanted lung tissue from COPD-affected lungs or age-matched control lungs by immunostaining. Relative to controls, we observed reduced staining for GPR126 among pro-SFTPC$^+$ AT2s in COPD lung tissue, but no change in airway epithelial cells (Figures 1B, S1A, and S1B). Thus, in established COPD, *ADGRG6*/GPR126 expression is diminished in AT2s.

Knockdown of *ADGRG6* in iAT2 modulates canonical GPCR signaling pathways

To interrogate the function of GPR126 in human type 2 alveolar epithelial cells, we made use of our previously developed inducible CRISPR interference (CRISPRi) iPSC platform (Figure 1C).[27] In this system, dCas9-KRAB protein is induced in iPSC-derived cell types by the addition of doxycycline to culture medium while guide (g)RNA are delivered using a lentiviral vector, allowing us to explore the effects of *ADGRG6* expression levels on canonical GPCR signaling together with associated downstream consequences in differentiated iAT2s (Figure 1D). Inducible CRISPRi-mediated *ADGRG6* knockdown in iAT2s significantly decreased both mRNA expression (20%–50% knockdown, Figure 1E) and GPR126 (Figures 1F, 1G, and S1C) levels. To assess the effect of *ADGRG6* knockdown (kd) on the iAT2 transcriptome, we performed single-cell RNA-sequencing (scRNA-seq) on control (−dox) and knockdown (+dox) iAT2s (Figures S1D–S1G). Gene set enrichment analysis (GSEA) of genes contained in the annotated GPCR Reactome database[38] identified Rho GTPase effectors and signaling via Rho GTPases pathways to be significantly downregulated in *ADGRG6*-kd iAT2s (Figures 1H and S1H). GPR126 couples to $G_{\alpha s}$, $G_{\alpha i}$, and/or $G_{\alpha 12/13}$ in a range of tissues.[39,40] Hence, to determine whether *ADGRG6*-kd alters downstream GPCR signaling in iAT2s, we measured intracellular active RhoA (downstream of $G_{\alpha 12/13}$) and cAMP (downstream of $G_{\alpha s}$ or $G_{\alpha i}$) in control versus knockdown cells and found that both were significantly attenuated in *ADGRG6*-kd iAT2s (Figures 1I and 1J). GPR126 is activated by extracellular matrix (ECM) components, including laminin 211 and collagen IV[15,16] that are found in Matrigel, the matrix in which iAT2s are embedded for 3D culture. While iAT2

(G) Normalized GPR126 levels.
(H) Module score of Reactome Rho GTPase effectors from single-cell RNA sequencing of iAT2s.
(I) Active RhoA.
(J) Intracellular cAMP. n = 3–6 experimental replicates of independent wells of a differentiation; error bars represent SD. Statistical significance was determined by unpaired, two-tailed Student's t test; *$p < 0.05$ and **$p < 0.005$.

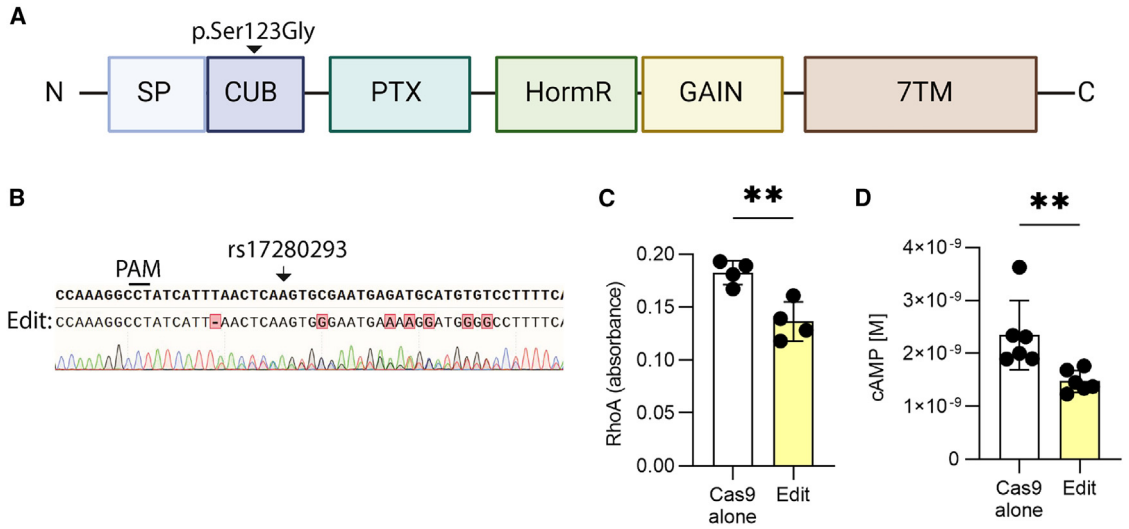


Figure 2. CRISPR-mediated mutagenesis of the region surrounding rs17280293 impairs GPR126 signaling
(A) Schematic of p.Ser123Gly variant in CUB domain of GPR126.
(B) CRISPR edit in the rs17280293 region.
(C and D) Active RhoA (C) and intracellular cAMP (D) in rs17280293-edited iAT2s. n = 4–6 experimental replicates of independent wells of a differentiation; error bars represent SD. Statistical significance was determined by unpaired, two-tailed Student's t test; **$p < 0.005$.

culture conditions could thus activate GPR126, we also considered the possibility that iAT2s may manufacture and secrete these ligands themselves, resulting in autocrine and/or paracrine receptor activation. We observed that iAT2s express *LAMA2* and *COL4A1* and that expression of both decreased following *ADGRG6*-kd (Figure S1II). Collectively, these data suggest that decreased GPR126 levels in iAT2s perturbs canonical GPCR signaling, leading to diminished cAMP and RhoA signaling.

rs17280293 region regulates GPCR function in iAT2s

The CUB domain of GPR126 is essential for ligand binding and receptor activation.[20] Genome-wide association studies (GWASs) of DLCO/VA and COPD have identified genetic variants at 6q24.1 in or near *ADGRG6* at genome-wide significance.[8,12] To further examine variants in this region, we identified previous lead variants associated with COPD, lung function, or DLCO/VA, and examined them with GTEx and eQTLGen (Table S1; Figure S2A). COPD risk alleles at this locus are associated with increased *ADGRG6* expression,[21] though these associations have not been identified in lung tissue. One of the top causal candidates in this region, rs17280293, is a missense variant (p.Ser123Gly) associated with both diffusion capacity and COPD.[10,12] rs17280293 is located in the region encoding the CUB domain of GPR126; however, the importance of this region to receptor activation in the lung remains unknown (Figure 2A). To understand the significance of this region to GPR126 activation and function in iAT2s, we delivered Cas9 together with gRNAs (versus Cas9 alone to control cells) to generate indels in close proximity to rs17280293 (Figure 2B). Disruption of the region surrounding rs17280293 did not affect GPR126 antibody staining (Figure S2B) consistent with continued GPR126 production rather than nonsense-mediated decay; however, we observed significantly reduced intracellular cAMP and active RhoA in *ADGRG6* mutant iAT2s (Figures 2C and 2D), consistent with reduced signaling through GPR126. These data suggest that the CUB domain contributes to GPR126 activation in AT2s that can be modified by disruption of exon 3 close to rs17280293.

Knockdown of *ADGRG6* increases an AT2 gene expression program

We have previously shown that COPD GWAS genes regulate the transcriptome and cellular phenotype of iAT2s.[27] To determine whether knockdown of *ADGRG6* modulates AT2 programs, we analyzed our scRNA-seq comparing control (−dox) to *ADGRG6*-kd (+dox) iAT2s. Uniform manifold projection (UMAP) visualization and Louvain clustering revealed that *ADGRG6*-kd iAT2s clustered separately from control (−dox) iAT2s (Figures 3A, S3A, and S3B). While expression of the master lung transcription factor *NKX2-1* was not altered between groups (Figure 3B), expression of genes associated with AT2 maturation[22,23] was significantly elevated in *ADGRG6*-kd iAT2s (Figures 3C and S3C). Moreover, *ADGRG6*-kd increased the percent of iAT2s producing high levels of SPC as indicated by *SFTPC*-driven expression of the tdTomato reporter (Figures 3D and 3E) and as validated by qPCR (Figure 3F). These data demonstrate that reduced GPR126 production induces the maturation of iAT2s.

ADGRG6 expression regulates the cytoskeleton and cell-cell junctions

To determine the effects of *ADGRG6*-kd on the iAT2 global transcriptome, we performed gene set enrichment analysis (GSEA) comparing differentially expressed genes between

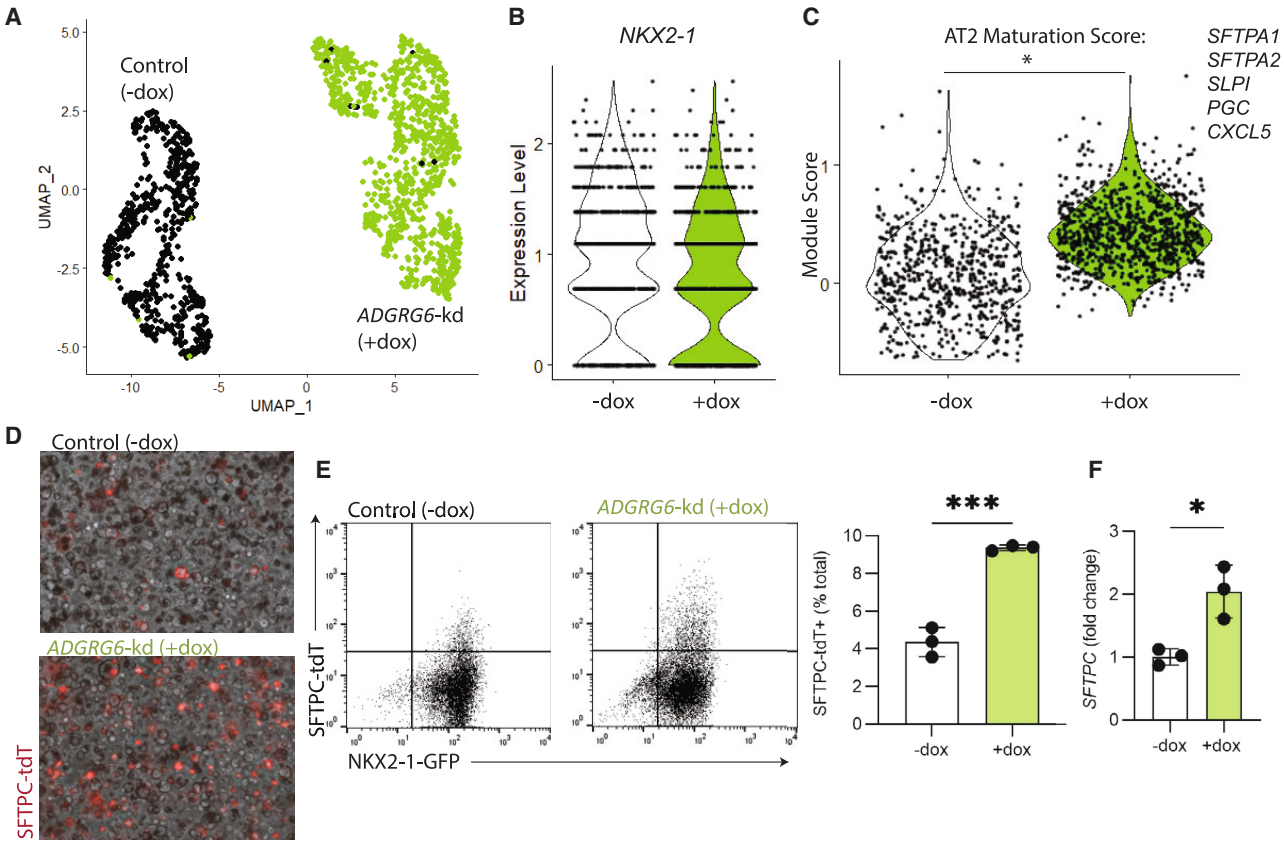


Figure 3. *ADGRG6*-kd in iAT2s upregulates an AT2 transcriptional program

(A) Uniform manifold projection (UMAP) of control (−dox, black) and *ADGRG6*-kd (+dox, green) iAT2s.

(B) Normalized *NKX2-1* single cell expression.

(C) Module score of AT2 maturation (*SFTPA1, SFTPA2, SLPI, PGC, CXCL5*).

(D) SFTPC-tdTomato (red) in control (−dox) and *ADGRG6*-kd (+dox) iAT2s.

(E) SFTPC-tdTomato and NKX2-1-GFP in control (−dox) and *ADGRG6*-kd (+dox) iAT2s.

(F) *SFTPC* expression in iAT2s, measured by qRT-PCR. n = 3 experimental replicates of independent wells of a differentiation; error bars represent SD. Statistical significance was determined by unpaired, two-tailed Student's t test; ***$p < 0.001$.

control (−dox) and *ADGRG6*-kd (+dox) iAT2s (Figure 4A). Pathway analysis revealed that focal adhesion, regulation of actin cytoskeleton, and tight junction pathways, known to be regulated by aGPCRs,[41–43] were downregulated in *ADGRG6*-kd iAT2s (Figure 4B). The ECM is linked to the actin cytoskeleton through focal adhesions. We found that expression of focal adhesion genes and proteins such as integrin-β1 (ITGB1) and integrin-α2 (ITGA2) were significantly downregulated following *ADGRG6*-kd (Figures 4C, 4D, and S4A–S4C), findings that were confirmed in a second, distinct pluripotent stem cell line (Figures S4F–S4H). Expression of these genes was unaffected in iAT2s transduced with non-targeting gRNA (Figures S4I–S4L). Furthermore, expression of genes associated with actin cytoskeleton regulation were decreased (Figures 4E and S4D), as were F-actin filament levels (Figure 4F) in *ADGRG6*-kd iAT2s. Lastly, GSEA identified tight junction pathways as downregulated in *ADGRG6*-kd (Figures 4G and S4E). As exposure of the apical cell surface to air induces the formation of tight junctions, we plated the cells on transwells and cultured them at an air-liquid interface (ALI).[26] After four days in ALI conditions, we measured transepithelial electrical resistance (TEER) to quantify barrier function and found a significant reduction in *ADGRG6*-kd iAT2s relative to controls (Figure 4H). Immunostaining for the tight junction protein ZO-1 in ALI cultures confirmed lower protein levels in *ADGRG6*-kd iAT2s (Figure 4I), consistent with this finding. Taken together, our data demonstrate that *ADGRG6* expression regulates cytoskeletal and cell junction proteins in alveolar epithelial cells.

Knockdown of *ADGRG6* increases iAT2 proliferation

GSEA revealed that multiple pathways were upregulated in *ADGRG6*-kd iAT2s, particularly pathways involved in cell cycling, such as mTOR signaling, Myc, and E2F targets (Figures 5A and 5B). Consistent with these results, we observed that *ADGRG6*-kd iAT2s were substantially more proliferative than control iAT2s, evidenced by an increased proportion of cells in S or G2M phase in scRNA-seq (Figures 5C and S5A) and significantly elevated cell yield across 5 passages (Figure 5D). We have previously found that iAT2s upregulate oxygen consumption rate (OCR) to accommodate increased proliferative demand.[27] We

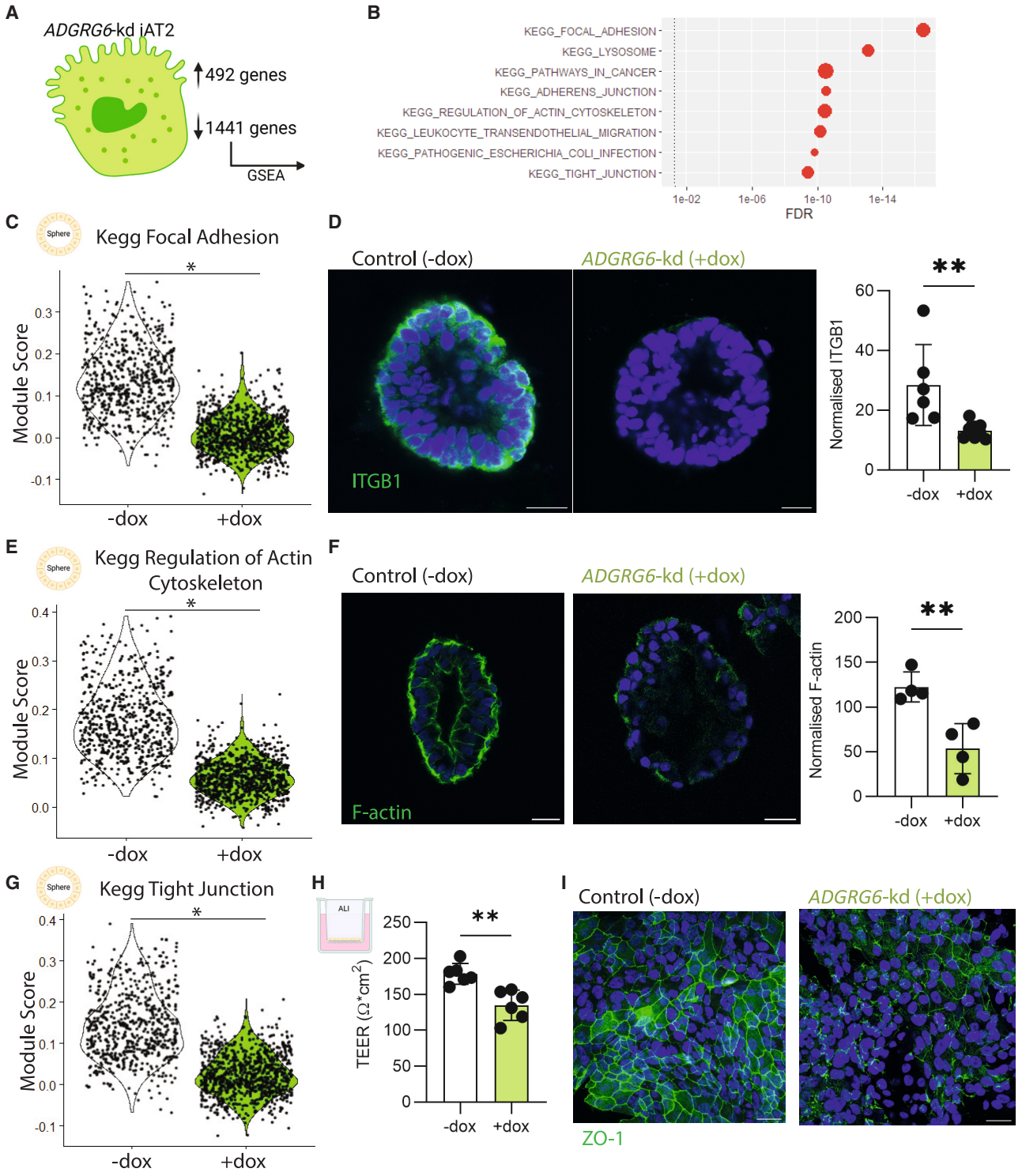


Figure 4. GPR126 knockdown regulates focal adhesion, actin cytoskeleton, and tight junctions
(A and B) Gene set enrichment analysis (GSEA) of downregulated genes following *ADGRG6*-kd in iAT2s.
(C) Module score of Kegg focal adhesion of scRNA-seq data.
(D) ITGB1 production.
(E) Module score of Kegg regulation of actin cytoskeleton of scRNA-seq data.
(F) F-actin (phalloidin) levels.
(G) Module score of Kegg tight junction of scRNA-seq data.
(H) *Trans*-epithelial electrical resistance (TEER) of control (−dox) and *ADGRG6*-kd (+dox) plated at air-liquid interface (ALI).
(I) ZO-1 (green) localization. n = 4–6 experimental replicates of independent wells of a differentiation; error bars represent SD. Scale bar for all images represents 20 μm. Statistical significance was determined by unpaired, two-tailed Student's t test; **$p < 0.005$.

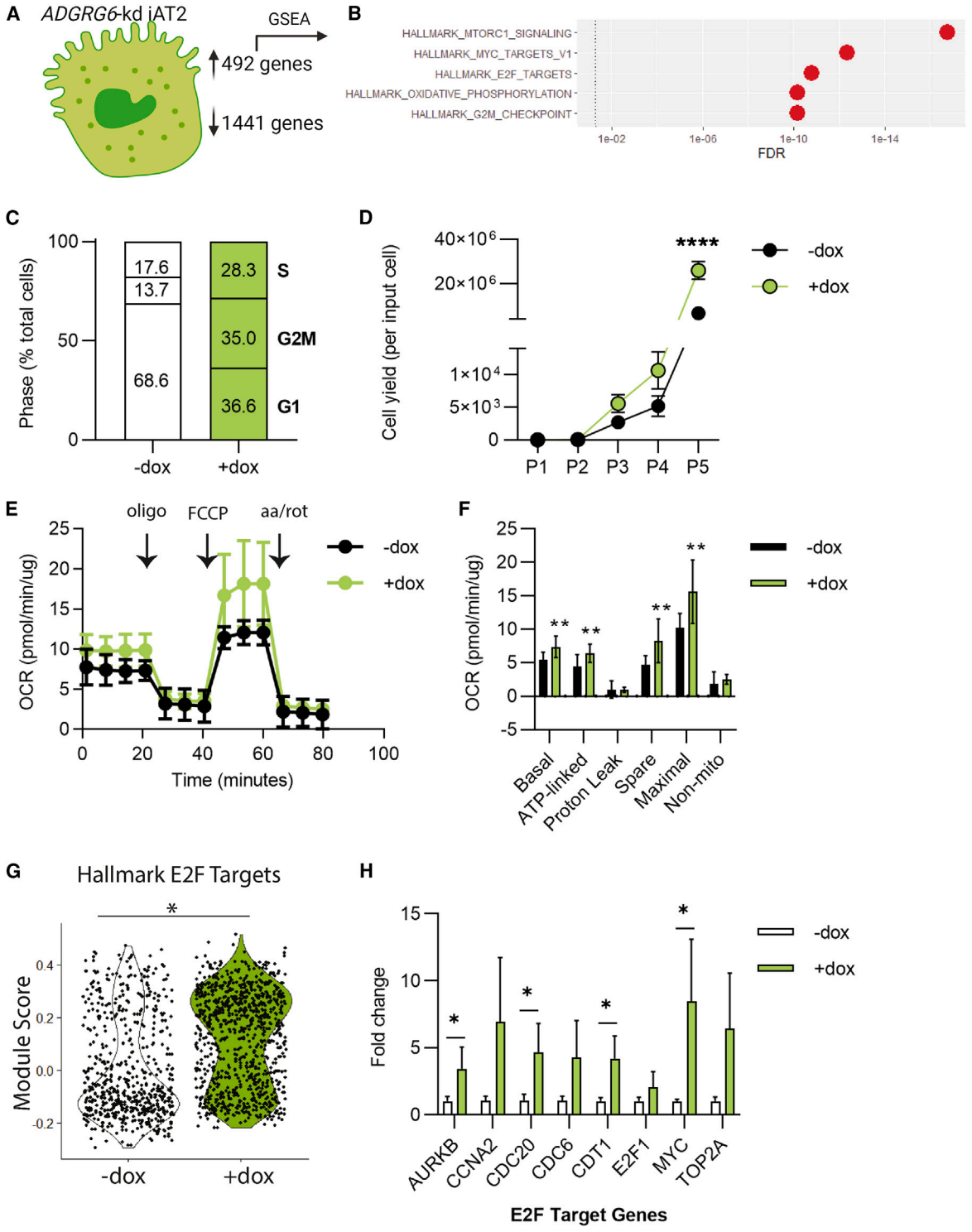


Figure 5. *ADGRG6*-kd iAT2s are hyperproliferative

(A and B) GSEA of upregulated genes following *ADGRG6*-kd in iAT2s.

(C) Cell cycle phase, assessed by scRNA-seq of control (−dox, white) and *ADGRG6*-kd (+dox, green).

(D) Cell yield per input iAT2 transduced with *ADGRG6* gRNA and treated without (−dox) or with (+dox) doxycycline across five passages.

(E and F) Basal oxygen consumption rate (OCR) was measured, followed by injection of oligomycin (oligo), FCCP, and rotenone + antimycin A (aa/rot), as shown. Data were normalized to total protein after the assay was complete. n = 6 technical replicates of a differentiation. iAT2s were D210 on the day of OCR experiments.

(G) Module score of hallmark E2F targets from scRNA-seq data.

(H) qRT-PCR of E2F target genes. n = 3 experimental replicates of independent wells of a differentiation; error bars represent SD. Statistical significance was determined by unpaired, two-tailed Student's t test; *$p < 0.05$ and **$p < 0.005$.

therefore measured mitochondrial capacity and indeed found that *ADGRG6*-kd iAT2s demonstrated heightened basal, ATP-linked, spare, and maximal respiratory capacity relative to controls (Figures 5E and 5F). This phenotype was not due to increased mitochondrial biogenesis, as expression of the master regulator *PPARGC1α* and total mitochondrial numbers were similar between control and *ADGRG6*-kd iAT2s (Figures S5B and S5C). To identify the mechanism through which *ADGRG6*-kd promotes iAT2 proliferation, we interrogated the mTOR and E2F signaling pathways. While the gene signature for mTORC1 signaling was upregulated in *ADGRG6*-kd iAT2s, we observed decreased production of proteins associated with mTOR signaling (Figures S5D–S5G) suggesting that hyperproliferation in *ADGRG6*-kd iAT2s is likely driven by an alternative process. E2F transcription factors are essential regulators of cell proliferation and expression of E2F targets was enriched in *ADGRG6*-kd iAT2s by scRNA-seq (Figures 5B and 5G) and elevated in confirmatory qPCR analysis (*AURKB, CDC20, CDT1, MYC*); however, phosphorylated retinoblastoma (Rb) protein was not significantly altered (Figures 5H, S5H, and S5I). Together, these results demonstrate that *ADGRG6* expression levels modulate iAT2 proliferation through a mechanism that remains uncertain.

ADGRG6 expression regulates iAT2 response to cigarette smoke

Emphysema is known to result from environmental exposures, particularly cigarette smoke, in genetically susceptible individuals. To elucidate the role of *ADGRG6* in the cellular response to cigarette smoke injury, we plated control vs. *ADGRG6*-kd iAT2s at an ALI and exposed them to cigarette smoke or air before harvest 8 h later for bulk RNA sequencing (RNA-seq) (Figure 6A). Principal component analysis (PCA) found that replicates from each treatment clustered together. *ADGRG6*-kd samples segregated from control samples along principal component 1, whereas samples separated along principal component 2 based on exposure condition (Figures 6B and S6A). Gene expression analysis identified 1,841 differentially expressed genes between control and *ADGRG6*-kd iAT2s exposed to air, and 4,650 differentially expressed genes between groups of iAT2s exposed to cigarette smoke (Figure 6C). By GSEA, we found that both control and *ADGRG6*-kd iAT2s shared canonical responses to cigarette smoke exposure including upregulation of genes encoding P450 enzymes (*CYP1A1* and *CYP1B1*), aldo-keto reductases (*AKR1C1, AKR1C2*, and *AKR1C3*), regulators of oxidative stress (*HMOX1*), and pathways such as Unfolded Protein Response (Figures S6B and S6C), consistent with results we have observed previously.[26] Next, we explored differential responses to cigarette smoke in control compared with *ADGRG6*-kd iAT2s. GSEA revealed that apical junction and inflammatory responses were downregulated by cigarette smoke exposure in *ADGRG6*-kd iAT2s compared to controls (Figures 6D, 6E, S6D, and S6E) while pathways involved in cell cycle (E2F Targets, G2M checkpoint) were upregulated (Figures 6D–6F, S6F, and S6G), suggesting that smoke exposure augments the effects of *ADGRG6* knockdown. To determine whether observed transcriptional differences altered cellular function, we next measured colony-forming efficiency (CFE) in smoke-exposed iAT2s. Consistent with the transcriptional data indicating increased proliferation in *ADGRG6*-kd iAT2s, we observed that *ADGRG6*-kd iAT2s formed a greater number of colonies than control iAT2s following cigarette smoke injury. These findings reveal that GPR126 levels modulate the alveolar epithelial response to cigarette smoke exposure.

Discussion

Variants in *ADGRG6* are associated with diffusing capacity and lung function in COPD. In this study, we explored the function of GPR126 in AT2s, the facultative progenitor cells of the distal lung. Using an inducible CRISPRi iPSC platform, we demonstrated that GPR126 activates intracellular cAMP and RhoA in iAT2s. Moreover, we found that *ADGRG6* expression regulates iAT2 focal adhesion, cytoskeletal organization, tight junction formation, and proliferation. Finally, we elucidated the consequences of reduced *ADGRG6* expression in cigarette smoke-exposed iAT2s.

The *ADGRG6* locus appears to harbor multiple independent signals, and the causal variant, cell type, and direction of effect are not clearly established. Our analysis of human scRNA-seq data, in combination with previous work,[12] suggests that *ADGRG6* expression is decreased in the end-stage COPD lung, particularly in AT2s and basal cells, though whether expression levels in these cell types are causal or instead the consequence of injury has not been determined. Interestingly, *ADGRG6* was highly expressed among a small number of aberrant basaloid cells cells present in a subset of the COPD lungs analyzed (33 cells identified among 9 of 17 COPD lung samples analyzed). This cell population, originally described in the IPF lung, is present in significantly smaller numbers and is of uncertain significance in COPD.[22,44] Based on its demonstrated association with diffusing capacity, we investigated the region surrounding rs17280293, which causes a missense variant (p.Ser123Gly) in the CUB domain of GPR126. In contrast to recombinant cell models which found no effect of p.Ser123Gly on receptor activation,[21] our results suggest the possibility that this region is important for GPR126 activation in iAT2s. These discrepant findings (1) could reflect the indirect nature of our approach, which did not recapitulate the amino acid change, (2) could result from differences in the experimental platforms utilized, or (3) could indicate that the effects of rs17280293 on receptor activation are cell type specific.

Adhesion GPCRs modulate cellular migration, adhesion, and polarity.[41] We observed that GPR126 production alters

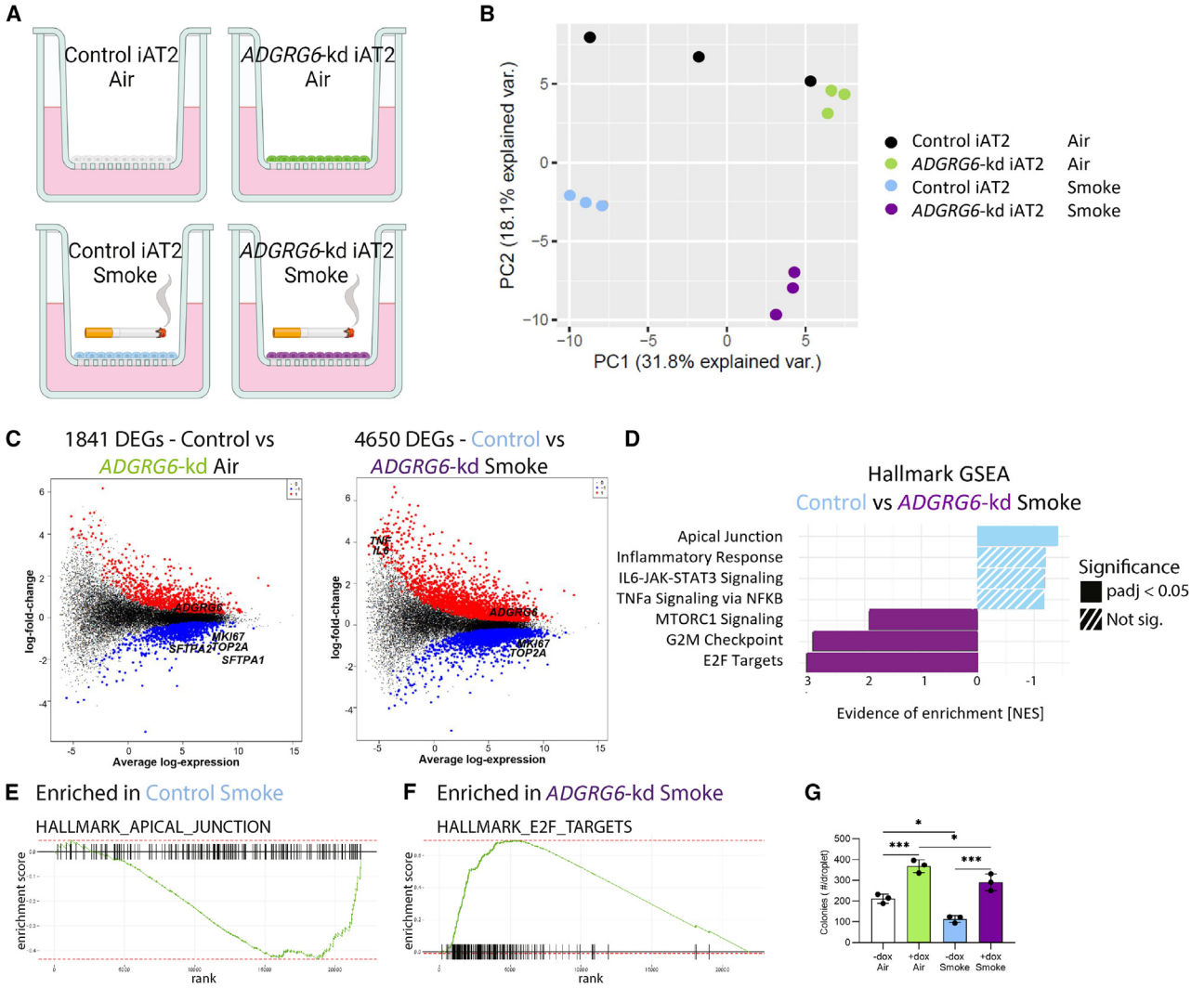


Figure 6. GPR126 knockdown modulates iAT2 response to cigarette smoke

(A) Control (−dox) and *ADGRG6*-kd (+dox) iAT2s were plated at air-liquid interface (ALI) prior to exposure to cigarette smoke gas phase. Eight hours following exposure, cells were harvested for bulk RNA sequencing.

(B) Principal component analysis (PCA) of control air-exposed (black), *ADGRG6*-kd air-exposed (green), control smoke-exposed (blue), and *ADGRG6*-kd smoke-exposed (purple) iAT2s.

(C) Differentially expressed genes (DEGs) in control air-exposed versus *ADGRG6*-kd air-exposed (left), or control smoke-exposed versus *ADGRG6*-kd smoke-exposed (right).

(D) Gene set enrichment analysis (GSEA) of control smoke-exposed versus *ADGRG6*-kd smoke-exposed.

(E) Enrichment plot of hallmark apical junction, downregulated in *ADGRG6*-kd smoke-exposed iAT2s.

(F) Enrichment plot of hallmark E2F targets, upregulated in *ADGRG6*-kd smoke-exposed.

(G) iAT2s were exposed to cigarette smoke, then dissociated and replated in matrigel to reform alveolospheres. Colony-forming efficiency was then calculated after 14 days in culture. n = 3 experimental replicates of independent wells of a differentiation; error bars represent SD. Statistical significance was determined by unpaired, two-tailed Student's t test; *p < 0.05 and **p < 0.005.

focal adhesions, cytoskeletal arrangement, and tight junction formation. ECM components, such as laminin-211 and collagen IV, bind to integrins and are ligands for GPR126.[15,16] Because our system relies on native levels of both the ligands and receptor, it is likely more analogous to *in vivo* activation of GPR126 in the alveolus than would be overexpression of the receptor and/or addition of exogenous Stachel ligand. Our findings in iAT2s indicate that GPR126 production regulates integrins, which are known to exert downstream effects on the actin cytoskeleton. Interestingly, *ADGRG6*-kd iAT2s also express lower levels of the genes encoding laminin-211 and collagen IV. This raises the possibility that GPR126 directly modulates integrin production and/or exerts these effects indirectly by altering the production of ECM ligands.[45] Moreover, *ADGRG6*-kd resulted in perturbated canonical GPCR signaling which diminished intracellular RhoA and cAMP, both of which play integral direct roles in actin cytoskeletal organization. G proteins, including $G_{\alpha 12}$ and RhoA, regulate tight junction formation,[42,46] potentially explaining reduced tight junction formation observed in *ADGRG6*-kd iAT2s.

Previous findings suggest that GPR126 regulates the proliferation of pulmonary arterial smooth muscle cells and airway smooth muscle cells,[21,47] and here we extend these observations to include iAT2s. We find that *ADGRG6*-kd induces expression of cell cycle genes, including target genes of the E2F family of transcription factors. This finding is in accord with studies in airway smooth muscle cells demonstrating that GPR126 activation downregulates E2F targets.[21] To compensate for increased proliferation, we found that *ADGRG6*-kd iAT2s upregulated mitochondrial respiration without increased mitochondrial biogenesis. When exploring pathways that may be responsible for hyperproliferation in *ADGRG6*-kd iAT2s, we found that the mTOR pathway, although transcriptionally upregulated, was decreased at the protein level, potentially representing a compensatory effect resulting from reduced mTOR activation in *ADGRG6*-kd. Moreover, while E2F transcription factors were transcriptionally upregulated, p-Rb was not significantly different. In light of these data it is likely that this phenotype is regulated by multiple intersecting pro-proliferative pathways. Finally, AT2 hyperproliferation in *ADGRG6*-kd may also occur through downregulation of integrins,[48] although there remains a significant gap in understanding the intracellular events linking aGPCRs with transcriptional regulation of cell cycle genes.

We have previously examined the effects of knocking down the desmosome/junctional protein, desmoplakin (encoded by *DSP*) in AT2s.[27] Intriguingly, pathways dysregulated in *DSP*-kd iAT2s, including AT2 maturation, cell-cell junctions, cytoskeletal organization, and hyperproliferation, were shared in common with *ADGRG6*-kd. This overlap suggests the possibility of shared mechanisms through which modulation of these genes by GWAS variants might confer risk for COPD/emphysema and provide potential insight into disease pathogenesis. Moreover, given that the *DSP* variant (rs2076295) associated with COPD risk increases DSP expression, we may be able to infer the direction of expression exerted by ADGRG6 variants and associated with risk in disease.

A recent study described variants in *ADGRG6* with gene-by-smoking interactions.[49] In this study, we found that *ADGRG6*-kd substantially alters alveolar epithelial responses to cigarette smoke injury: where control cells upregulate apical complexes and immune signaling, *ADGRG6*-kd iAT2s instead adopt a hyperproliferative state. Emerging work has identified the involvement of aGPCRs in immune cell development and behavior[50]; this study extends these findings to test the responsiveness of an aGPCR to a proinflammatory stimulus (i.e., cigarette smoke exposure) in epithelial cells. Interestingly, our results suggest that GPR126 is involved in mounting an epithelial inflammatory response to smoke exposure. In future work, it will be interesting to determine whether GPR126 coordinates epithelial responses to other proinflammatory insults, such as lung infections. Cigarette smoke alters the composition of the ECM,[51] raising the possibility that GPR126 activation may be modulated by the effects of smoke inhalation on activating ligands. Finally, our findings demonstrate that *ADGRG6*-kd iAT2s are more proliferative and better able to form colonies following cigarette smoke injury. We speculate that this response might support repair in the setting of injury and that increased *ADGRG6* expression might thus impair a regenerative response, increasing risk of injury. Future studies are needed to understand the downstream consequences of this observation with respect to lung regeneration and whether GPR126 levels alter the alveolar response to injury *in vivo*.

Our study sought to determine the function of *ADGRG6* in type 2 alveolar epithelial cells, a distal lung cell type central to maintaining alveolar homeostasis, immune responses, and regeneration. A limitation of our reductionist platform is that iAT2s were studied in isolation in the absence of AT1s, stromal and immune cells colocated with them in the lung. We could not extend our findings to an *in vivo* mouse model, as *Adgrg6* is not highly expressed in murine lung epithelial cells,[52] consistent with divergent evolutionary patterns observed for other GWAS genes.[53] Further studies to validate these findings in orthologous systems might be thus performed in large animal models (e.g., pigs or ferrets) and/or human precision cut lung slices.

In conclusion, we provide functional characterization of the COPD GWAS gene, *ADGRG6*, in AT2-like cells. We find that GPR126 exerts pleiotropic effects on iAT2s including GPCR signaling, focal adhesions, and cytoskeletal arrangement, maturation, proliferation, and response to cigarette smoke injury. Emphysema occurs in genetically susceptible individuals in response to injury, and our findings show that a COPD GWAS gene, *ADGRG6*, regulates many processes associated with response to injury (e.g., cell adhesion, proliferation, and proinflammatory responses), providing further insight into gene-environment interactions in COPD pathogenesis. Given that GPCRs represent the largest family of druggable targets, future studies to explore GPR126 as a novel therapeutic target for COPD may be warranted.

Data and code availability

The single-cell RNA-sequencing data generated during this study are available at GEO (GSE223078). The bulk RNA-sequencing data generated in this study are available at GEO (GSE223077).

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.ajhg.2023.08.017.

Acknowledgments

We are grateful to members of the Wilson, Kotton, Hawkins, and Murphy laboratories in the Center for Regenerative Medicine (CReM) for advice; Greg Miller (CReM Laboratory Manager); and

Marianne James (CReM iPSC Core Manager). Thank you to Mikel Garcia-Marcos for helpful discussions. We thank Brian Tilton and the BUMC Flow Cytometry Core for cell sorting services (supported by NIH grant #1S10OD021587-01A1); Yuriy Aleksyeyev, Ashley LeClerc, and the BUMC Sequencing Core Facility for performing single-cell RNA sequencing and bulk RNA sequencing; and Michael T. Kirber of the Cellular Imaging Core. This work was supported by a CJ Martin Early Career Fellowship from the Australian National Health and Medical Research Council and a Boston University Clinical & Translational Science Institute sequencing grant (EPAR5926) awarded to R.B.W.; NIH grants U01TR001810, R01DK101501, and R01DK117940 awarded to A.A.W.; and NIH grants R01HL153248 and R01HL147148 to M.H.C.

Author contributions

R.B.W. and A.A.W. conceptualized the project and designed experiments; R.B.W., K.A.B., C.M., M.G., and S.M.L. performed experiments; R.B.W., C.V.-M., F.W., V.M., and P.B. performed bioinformatics analyses; M.C.B. and E.E.M. provided human lung sections; S.M.L., D.N.K., X.Z., and M.H.C. provided expert input on experimental design and data interpretation; R.B.W. and A.A.W. wrote the manuscript.

Declaration of interests

M.H.C. has received grant support from GSK and Bayer and speaking/consulting fees from Illumina, AstraZeneca, and Genentech. X.Z. has received grant support from Bayer.

Received: March 31, 2023
Accepted: August 31, 2023
Published: September 20, 2023

References

1. Adeloye, D., Song, P., Zhu, Y., Campbell, H., Sheikh, A., Rudan, I.; and NIHR RESPIRE Global Respiratory Health Unit (2022). Global, regional, and national prevalence of, and risk factors for, chronic obstructive pulmonary disease (COPD) in 2019: a systematic review and modelling analysis. Lancet Respir. Med. *10*, 447–458. https://doi.org/10.1016/S2213-2600(21)00511-7.
2. Hegewald, M.J. (2009). Diffusing Capacity. Clin. Rev. Allergy Immunol. *37*, 159–166. https://doi.org/10.1007/s12016-009-8125-2.
3. Regan, E.A., Lynch, D.A., Curran-Everett, D., Curtis, J.L., Austin, J.H.M., Grenier, P.A., Kauczor, H.U., Bailey, W.C., DeMeo, D.L., Casaburi, R.H., et al. (2015). Clinical and Radiologic Disease in Smokers With Normal Spirometry. JAMA Intern. Med. *175*, 1539–1549. https://doi.org/10.1001/jamainternmed.2015.2735.
4. Zhou, J.J., Cho, M.H., Castaldi, P.J., Hersh, C.P., Silverman, E.K., and Laird, N.M. (2013). Heritability of chronic obstructive pulmonary disease and related phenotypes in smokers. Am. J. Respir. Crit. Care Med. *188*, 941–947. https://doi.org/10.1164/rccm.201302-0263OC.
5. Bakke, P.S., Baste, V., Hanoa, R., and Gulsvik, A. (1991). Prevalence of obstructive lung disease in a general population: relation to occupational title and exposure to some airborne agents. Thorax *46*, 863–870. https://doi.org/10.1136/thx.46.12.863.
6. Tan, W.C., Sin, D.D., Bourbeau, J., Hernandez, P., Chapman, K.R., Cowie, R., FitzGerald, J.M., Marciniuk, D.D., Maltais, F., Buist, A.S., et al. (2015). Characteristics of COPD in never-smokers and ever-smokers in the general population: results from the CanCOLD study. Thorax *70*, 822–829. https://doi.org/10.1136/thoraxjnl-2015-206938.
7. Hobbs, B.D., de Jong, K., Lamontagne, M., Bossé, Y., Shrine, N., Artigas, M.S., Wain, L.V., Hall, I.P., Jackson, V.E., Wyss, A.B., et al. (2017). Genetic loci associated with chronic obstructive pulmonary disease overlap with loci for lung function and pulmonary fibrosis. Nat. Genet. *49*, 426–432. https://doi.org/10.1038/ng.3752.
8. Soler Artigas, M., Loth, D.W., Wain, L.V., Gharib, S.A., Obeidat, M., Tang, W., Zhai, G., Zhao, J.H., Smith, A.V., Huffman, J.E., et al. (2011). Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. Nat. Genet. *43*, 1082–1090. https://doi.org/10.1038/ng.941.
9. Wain, L.V., Shrine, N., Artigas, M.S., Erzurumluoglu, A.M., Noyvert, B., Bossini-Castillo, L., Obeidat, M., Henry, A.P., Portelli, M.A., Hall, R.J., et al. (2017). Genome-wide association analyses for lung function and chronic obstructive pulmonary disease identify new loci and potential druggable targets. Nat. Genet. *49*, 416–425. https://doi.org/10.1038/ng.3787.
10. Shrine, N., Guyatt, A.L., Erzurumluoglu, A.M., Jackson, V.E., Hobbs, B.D., Melbourne, C.A., Batini, C., Fawcett, K.A., Song, K., Sakornsakolpat, P., et al. (2019). New genetic signals for lung function highlight pathways and chronic obstructive pulmonary disease associations across multiple ancestries. Nat. Genet. *51*, 481–493. https://doi.org/10.1038/s41588-018-0321-7.
11. Sakornsakolpat, P., Prokopenko, D., Lamontagne, M., Reeve, N.F., Guyatt, A.L., Jackson, V.E., Shrine, N., Qiao, D., Bartz, T.M., Kim, D.K., et al. (2019). Genetic landscape of chronic obstructive pulmonary disease identifies heterogeneous cell-type and phenotype associations. Nat. Genet. *51*, 494–505. https://doi.org/10.1038/s41588-018-0342-2.
12. Terzikhan, N., Sun, F., Verhamme, F.M., Adams, H.H.H., Loth, D., Bracke, K.R., Stricker, B.H.C., Lahousse, L., Dupuis, J., Brusselle, G.G., and O'Connor, G.T. (2018). Heritability and genome-wide association study of diffusing capacity of the lung. Eur. Respir. J. *52*, 1800647. https://doi.org/10.1183/13993003.00647-2018.
13. Shrine, N., Izquierdo, A.G., Chen, J., Packer, R., Hall, R.J., Guyatt, A.L., Batini, C., Thompson, R.J., Pavuluri, C., Malik, V., et al. (2023). Multi-ancestry genome-wide association analyses improve resolution of genes and pathways influencing lung function and chronic obstructive pulmonary disease risk. Nat. Genet. *55*, 410–422. https://doi.org/10.1038/s41588-023-01314-0.
14. Miller, S., Melén, E., Merid, S.K., Hall, I.P., and Sayers, I. (2016). Genes associated with polymorphic variants predicting lung function are differentially expressed during human lung development. Respir. Res. *17*, 95. https://doi.org/10.1186/s12931-016-0410-z.
15. Paavola, K.J., Sidik, H., Zuchero, J.B., Eckart, M., and Talbot, W.S. (2014). Type IV collagen is an activating ligand for the adhesion G protein-coupled receptor GPR126. Sci. Signal. *7*, ra76. https://doi.org/10.1126/scisignal.2005347.
16. Petersen, S.C., Luo, R., Liebscher, I., Giera, S., Jeong, S.J., Mogha, A., Ghidinelli, M., Feltri, M.L., Schöneberg, T., Piao,

X., and Monk, K.R. (2015). The adhesion GPCR GPR126 has distinct, domain-dependent functions in Schwann cell development mediated by interaction with laminin-211. Neuron *85*, 755–769. https://doi.org/10.1016/j.neuron.2014.12.057.

17. Waller-Evans, H., Prömel, S., Langenhan, T., Dixon, J., Zahn, D., Colledge, W.H., Doran, J., Carlton, M.B.L., Davies, B., Aparicio, S.A.J.R., et al. (2010). The Orphan Adhesion-GPCR GPR126 Is Required for Embryonic Development in the Mouse. PLoS One *5*, e14047. https://doi.org/10.1371/journal.pone.0014047.
18. Patra, C., van Amerongen, M.J., Ghosh, S., Ricciardi, F., Sajjad, A., Novoyatleva, T., Mogha, A., Monk, K.R., Mühlfeld, C., and Engel, F.B. (2013). Organ-specific function of adhesion G protein-coupled receptor GPR126 is domain-dependent. Proc. Natl. Acad. Sci. USA *110*, 16898–16903. https://doi.org/10.1073/pnas.1304837110.
19. Monk, K.R., Naylor, S.G., Glenn, T.D., Mercurio, S., Perlin, J.R., Dominguez, C., Moens, C.B., and Talbot, W.S. (2009). A G protein-coupled receptor is essential for Schwann cells to initiate myelination. Science (New York, N.Y.) *325*, 1402–1405. https://doi.org/10.1126/science.1173474.
20. Leon, K., Cunningham, R.L., Riback, J.A., Feldman, E., Li, J., Sosnick, T.R., Zhao, M., Monk, K.R., and Araç, D. (2020). Structural basis for adhesion G protein-coupled receptor Gpr126 function. Nat. Commun. *11*, 194. https://doi.org/10.1038/s41467-019-14040-1.
21. Hall, R.J., O'Loughlin, J., Billington, C.K., Thakker, D., Hall, I.P., and Sayers, I. (2021). Functional genomics of GPR126 in airway smooth muscle and bronchial epithelial cells. Faseb. J. *35*, e21300. https://doi.org/10.1096/fj.202002073R.
22. Sauler, M., McDonough, J.E., Adams, T.S., Kothapalli, N., Barnthaler, T., Werder, R.B., Schupp, J.C., Nouws, J., Robertson, M.J., Coarfa, C., et al. (2022). Characterization of the COPD alveolar niche using single-cell RNA sequencing. Nat. Commun. *13*, 494. https://doi.org/10.1038/s41467-022-28062-9.
23. Hurley, K., Ding, J., Villacorta-Martin, C., Herriges, M.J., Jacob, A., Vedaie, M., Alysandratos, K.D., Sun, Y.L., Lin, C., Werder, R.B., et al. (2020). Reconstructed Single-Cell Fate Trajectories Define Lineage Plasticity Windows during Differentiation of Human PSC-Derived Distal Lung Progenitors. Cell Stem Cell *26*, 593–608.e8. https://doi.org/10.1016/j.stem.2019.12.009.
24. Jacob, A., Morley, M., Hawkins, F., McCauley, K.B., Jean, J.C., Heins, H., Na, C.L., Weaver, T.E., Vedaie, M., Hurley, K., et al. (2017). Differentiation of Human Pluripotent Stem Cells into Functional Lung Alveolar Epithelial Cells. Cell Stem Cell *21*, 472–488.e10. https://doi.org/10.1016/j.stem.2017.08.014.
25. Alysandratos, K.D., Russo, S.J., Petcherski, A., Taddeo, E.P., Acín-Pérez, R., Villacorta-Martin, C., Jean, J.C., Mulugeta, S., Rodriguez, L.R., Blum, B.C., et al. (2021). Patient-specific iPSCs carrying an SFTPC mutation reveal the intrinsic alveolar epithelial dysfunction at the inception of interstitial lung disease. Cell Rep. *36*, 109636. https://doi.org/10.1016/j.celrep.2021.109636.
26. Abo, K.M., Sainz de Aja, J., Lindstrom-Vautrin, J., Alysandratos, K.D., Richards, A., Garcia-de-Alba, C., Huang, J., Hix, O.T., Werder, R.B., Bullitt, E., et al. (2022). Air-liquid interface culture promotes maturation and allows environmental exposure of pluripotent stem cell-derived alveolar epithelium. JCI insight *7*, e155589. https://doi.org/10.1172/jci.insight.155589.
27. Werder, R.B., Liu, T., Abo, K.M., Lindstrom-Vautrin, J., Villacorta-Martin, C., Huang, J., Hinds, A., Boyer, N., Bullitt, E., Liesa, M., et al. (2022). CRISPR interference interrogation of COPD GWAS genes reveals the functional significance of desmoplakin in iPSC-derived alveolar epithelial cells. Sci. Adv. *8*, eabo6566. https://doi.org/10.1126/sciadv.abo6566.
28. Alysandratos, K.D., Garcia de Alba Rivas, C., Yao, C., Pessina, P., Villacorta-Martin, C., Huang, J., Hix, O.T., Minakin, K., Konda, B., Stripp, B.R., et al. (2022). Culture impact on the transcriptomic programs of primary and iPSC-derived human alveolar type 2 cells. JCI Insight *8*.
29. Werder, R.B., Huang, J., Abo, K.M., Hix, O.T., Minakin, K., Alysandratos, K.-D., Merritt, C., Berthiaume, K., Alber, A.B., Burgess, C.L., et al. (2022). Generating 3D Spheres and 2D Air-Liquid Interface Cultures of Human Induced Pluripotent Stem Cell-Derived Type 2 Alveolar Epithelial Cells. J. Vis. Exp.e63875 https://doi.org/10.3791/63875.
30. Mandegar, M.A., Huebsch, N., Frolov, E.B., Shin, E., Truong, A., Olvera, M.P., Chan, A.H., Miyaoka, Y., Holmes, K., Spencer, C.I., et al. (2016). CRISPR Interference Efficiently Induces Specific and Reversible Gene Silencing in Human iPSCs. Cell Stem Cell *18*, 541–553. https://doi.org/10.1016/j.stem.2016.01.022.
31. Basil, M.C., Cardenas-Diaz, F.L., Kathiriya, J.J., Morley, M.P., Carl, J., Brumwell, A.N., Katzen, J., Slovik, K.J., Babu, A., Zhou, S., et al. (2022). Human distal airways contain a multipotent secretory cell that can regenerate alveoli. Nature *604*, 120–126. https://doi.org/10.1038/s41586-022-04552-0.
32. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., 3rd, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell *177*, 1888–1902.e21. https://doi.org/10.1016/j.cell.2019.05.031.
33. Stoeckius, M., Zheng, S., Houck-Loomis, B., Hao, S., Yeung, B.Z., Mauck, W.M., Smibert, P., and Satija, R. (2018). Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. Genome Biol. *19*, 224. https://doi.org/10.1186/s13059-018-1603-1.
34. Hafemeister, C., and Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol. *20*, 296. https://doi.org/10.1186/s13059-019-1874-1.
35. Kowalczyk, M.S., Tirosh, I., Heckl, D., Rao, T.N., Dixit, A., Haas, B.J., Schneider, R.K., Wagers, A.J., Ebert, B.L., and Regev, A. (2015). Single-cell RNA-seq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells. Genome Res. *25*, 1860–1872. https://doi.org/10.1101/gr.192237.115.
36. Federico, A., and Monti, S. (2020). hypeR: an R package for geneset enrichment workflows. Bioinformatics *36*, 1307–1308. https://doi.org/10.1093/bioinformatics/btz700.
37. Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. *1*, 417–425. https://doi.org/10.1016/j.cels.2015.12.004.
38. Jassal, B., Jupe, S., Caudy, M., Birney, E., Stein, L., Hermjakob, H., and D'Eustachio, P. (2010). The systematic annotation of the three main GPCR families in Reactome. Database *2010*, baq018. https://doi.org/10.1093/database/baq018.
39. Mogha, A., Harty, B.L., Carlin, D., Joseph, J., Sanchez, N.E., Suter, U., Piao, X., Cavalli, V., and Monk, K.R. (2016). Gpr126/Adgrg6 Has Schwann Cell Autonomous and

Nonautonomous Functions in Peripheral Nerve Injury and Repair. J. Neurosci. *36*, 12351–12367. https://doi.org/10.1523/JNEUROSCI.3854-15.2016.

40. Lizano, E., Hayes, J.L., and Willard, F.S. (2021). A synthetic method to assay adhesion-family G-protein coupled receptors. Determination of the G-protein coupling profile of ADGRG6(GPR126). Biochem. Biophys. Res. Commun. *534*, 317–322. https://doi.org/10.1016/j.bbrc.2020.11.086.
41. Hamann, J., Aust, G., Araç, D., Engel, F.B., Formstone, C., Fredriksson, R., Hall, R.A., Harty, B.L., Kirchhoff, C., Knapp, B., et al. (2015). International Union of Basic and Clinical Pharmacology. XCIV. Adhesion G protein-coupled receptors. Pharmacol. Rev. *67*, 338–367. https://doi.org/10.1124/pr.114.009647.
42. Chumki, S.A., van den Goor, L.M., Hall, B.N., and Miller, A.L. (2022). p115RhoGEF activates RhoA to support tight junction maintenance and remodeling. Mol. Biol. Cell *33*, ar136. https://doi.org/10.1091/mbc.E22-06-0205.
43. Sakurai, T., Kamakura, S., Hayase, J., Kohda, A., Nakamura, M., and Sumimoto, H. (2022). GPR125 (ADGRA3) is an autocleavable adhesion GPCR that traffics with Dlg1 to the basolateral membrane and regulates epithelial apicobasal polarity. J. Biol. Chem. *298*, 102475. https://doi.org/10.1016/j.jbc.2022.102475.
44. Adams, T.S., Schupp, J.C., Poli, S., Ayaub, E.A., Neumark, N., Ahangari, F., Chu, S.G., Raby, B.A., DeIuliis, G., Januszyk, M., et al. (2020). Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. Sci. Adv. *6*, eaba1983. https://doi.org/10.1126/sciadv.aba1983.
45. Delcommenne, M., and Streuli, C.H. (1995). Control of Integrin Expression by Extracellular Matrix (*). J. Biol. Chem. *270*, 26794–26801. https://doi.org/10.1074/jbc.270.45.26794.
46. Meyer, T.N., Schwesinger, C., and Denker, B.M. (2002). Zonula occludens-1 is a scaffolding protein for signaling molecules. Galpha(12) directly binds to the Src homology 3 domain and regulates paracellular permeability in epithelial cells. J. Biol. Chem. *277*, 24855–24858. https://doi.org/10.1074/jbc.C200240200.
47. Gorr, M.W., Sriram, K., Muthusamy, A., and Insel, P.A. (2020). Transcriptomic analysis of pulmonary artery smooth muscle cells identifies new potential therapeutic targets for idiopathic pulmonary arterial hypertension. Br. J. Pharmacol. *177*, 3505–3518. https://doi.org/10.1111/bph.15074.
48. Plosa, E.J., Benjamin, J.T., Sucre, J.M., Gulleman, P.M., Gleaves, L.A., Han, W., Kook, S., Polosukhin, V.V., Haake, S.M., Guttentag, S.H., et al. (2020). β1 Integrin regulates adult lung alveolar epithelial cell inflammation. JCI insight *5*, e129259. https://doi.org/10.1172/jci.insight.129259.
49. Yang, T., Jackson, V.E., Smith, A.V., Chen, H., Bartz, T.M., Sitlani, C.M., Psaty, B.M., Gharib, S.A., O'Connor, G.T., Dupuis, J., et al. (2021). Rare and low-frequency exonic variants and gene-by-smoking interactions in pulmonary function. Sci. Rep. *11*, 19365. https://doi.org/10.1038/s41598-021-98120-7.
50. Lala, T., and Hall, R.A. (2022). Adhesion G protein-coupled receptors: structure, signaling, physiology, and pathophysiology. Physiol. Rev. *102*, 1587–1624. https://doi.org/10.1152/physrev.00027.2021.
51. Obernolte, H., Niehof, M., Braubach, P., Fieguth, H.-G., Jonigk, D., Pfennig, O., Tschernig, T., Warnecke, G., Braun, A., and Sewald, K. (2022). Cigarette smoke alters inflammatory genes and the extracellular matrix — investigations on viable sections of peripheral human lungs. Cell Tissue Res. *387*, 249–260. https://doi.org/10.1007/s00441-021-03553-1.
52. 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. (2019). Single-cell connectomic analysis of adult mammalian lungs. Sci. Adv. *5*, eaaw3851. https://doi.org/10.1126/sciadv.aaw3851.
53. Travaglini, K.J., Nabhan, A.N., Penland, L., Sinha, R., Gillich, A., Sit, R.V., Chang, S., Conley, S.D., Mori, Y., Seita, J., et al. (2020). A molecular cell atlas of the human lung from single-cell RNA sequencing. Nature *587*, 619–625. https://doi.org/10.1038/s41586-020-2922-4.