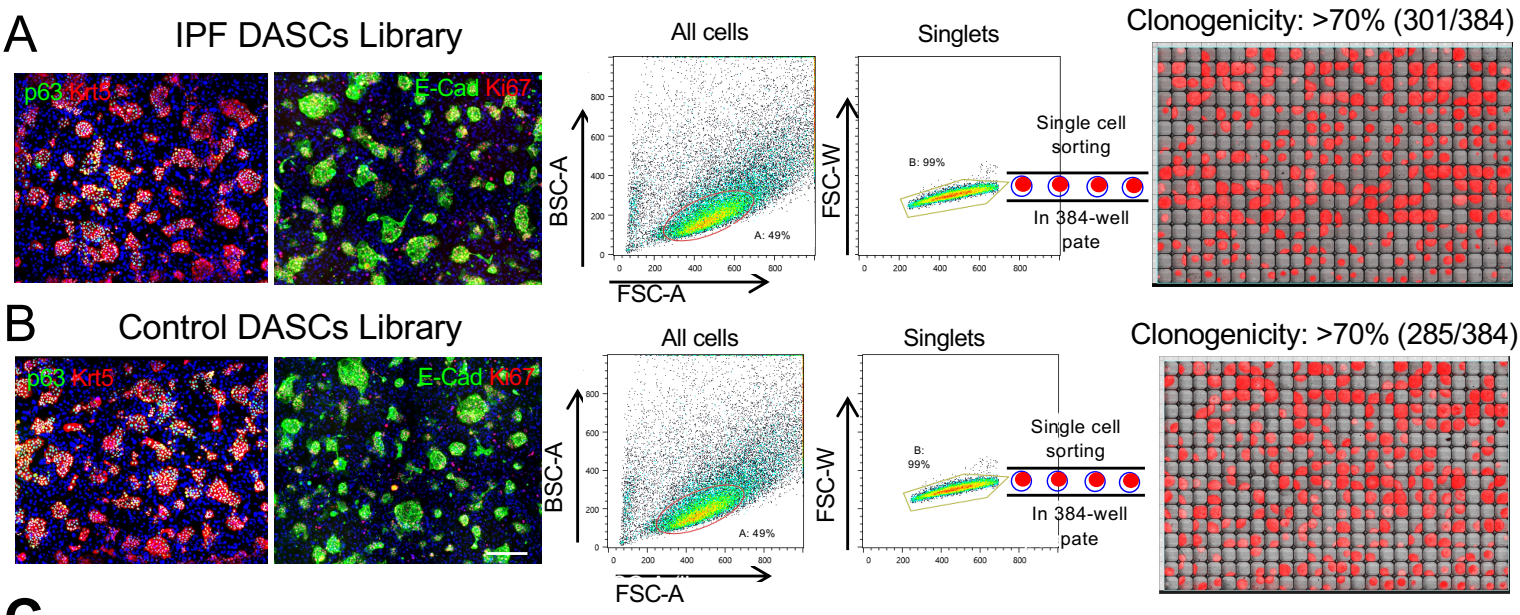
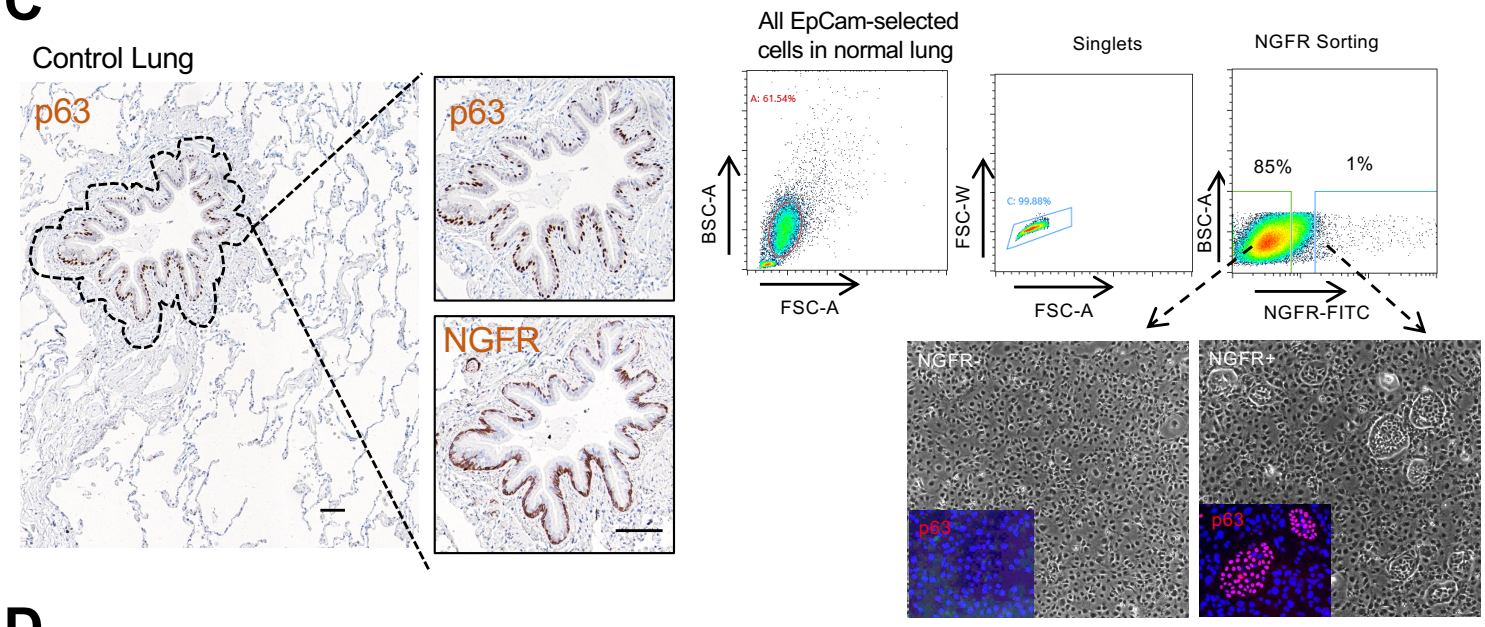
Fig. S1


A
IPF DASCs Library
p63 Krt5
E-Cad Ki67
All cells
Singlets
Clonogenicity: >70% (301/384)
BSC-A
FSC-A
FSC-W
B: 99%
A: 49%
Single cell sorting
In 384-well pate
B
Control DASCs Library
p63 Krt5
E-Cad Ki67
All cells
Singlets
Clonogenicity: >70% (285/384)
BSC-A
FSC-A
FSC-W
B: 99%
A: 49%
Single cell sorting
In 384-well pate


C
Control Lung
p63
NGFR
All EpCam-selected cells in normal lung
Singlets
NGFR Sorting
A: 61.54%
C: 99.88%
85%
1%
BSC-A
FSC-A
FSC-W
NGFR-FITC
NGFR-
NGFR+
p63

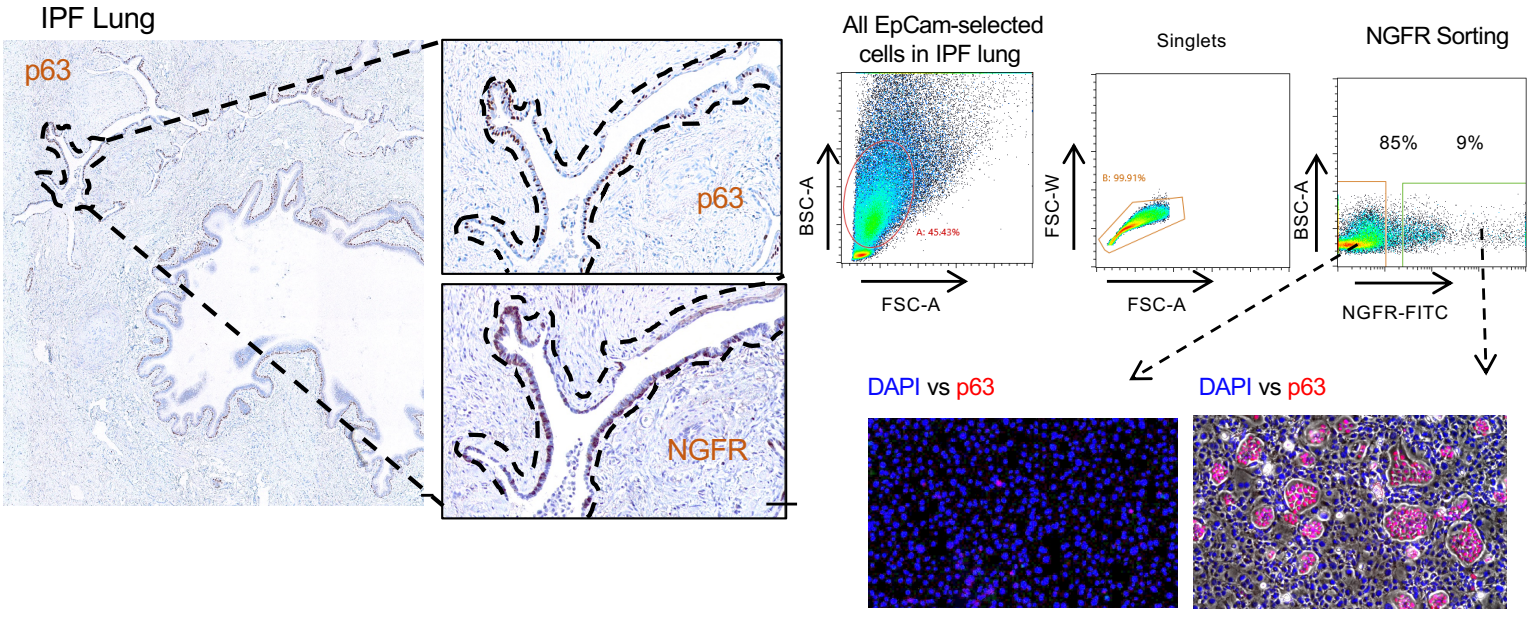
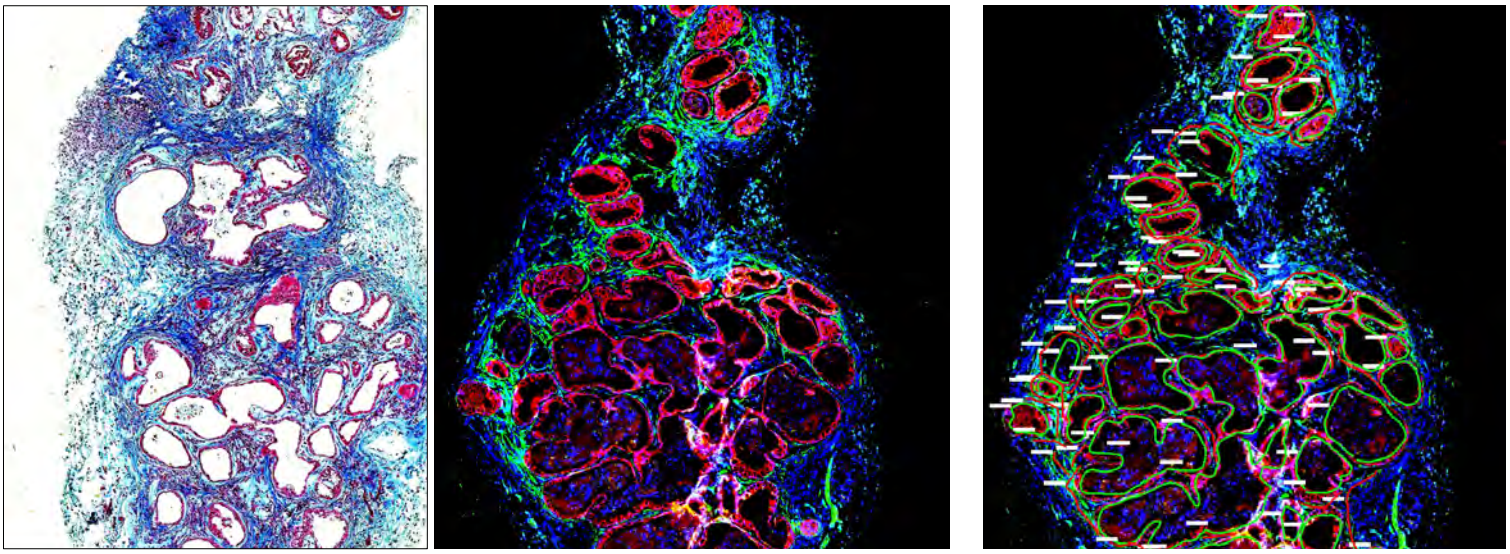

D
IPF Lung
p63
NGFR
All EpCam-selected cells in IPF lung
Singlets
NGFR Sorting
A: 45.43%
B: 99.91%
85%
9%
BSC-A
FSC-A
FSC-W
NGFR-FITC
DAPI vs p63
DAPI vs p63

Fig. S1 *A.,B. Un*iform TP63/Krt5 expression and high clonogenicity of library cells.
A. *Left,* Immunostaining of colonies in control libraries with antibodies to TP63, KRT5, ECAD, and Ki67. *Right,* FACS sorting of EpCAM selected cells from control libraries to single wells of a 384-well plate for assessment of clonogenicity. ***B.*** Immunostaining of colonies in IPF libraries with antibodies to TP63, KRT5, ECAD, and Ki67. *Right,* FACS sorting of EpCAM selected cells from control libraries to single wells of a 384-well plate for assessment of clonogenicity.
C. Isolation and cloning of TP63+/NGFR+ cells from control lung. *Left,* Immunohistochemistry of terminal bronchiole of control lung histological sections with antibodies to TP63 and NGFR. *Right*, FACS sorting of singlet NFGR+ cells from EpCam-selected (epithelial) lung cells and colony-forming assays on lawns of irradiated 3T3-J2 fibroblasts of NFGR+ and NGFR- cells, coupled with colony staining with antibodies to TP63. ***D.*** Isolation and cloning of TP63+/NGFR+ cells from IPF lung. *Left*, Immunohistochemistry of terminal bronchiole of IPF lung histological sections with antibodies to TP63 and NGFR. *Right*, FACS sorting of singlet NFGR+ cells from EpCam-selected (epithelial) lung cells and colony-forming assays on lawns of irradiated 3T3-J2 fibroblasts of NFGR+ and NGFR- cells, coupled with colony staining with antibodies to TP63.

Fig. S2

IPF xenograft F04



IPF xenograft F07

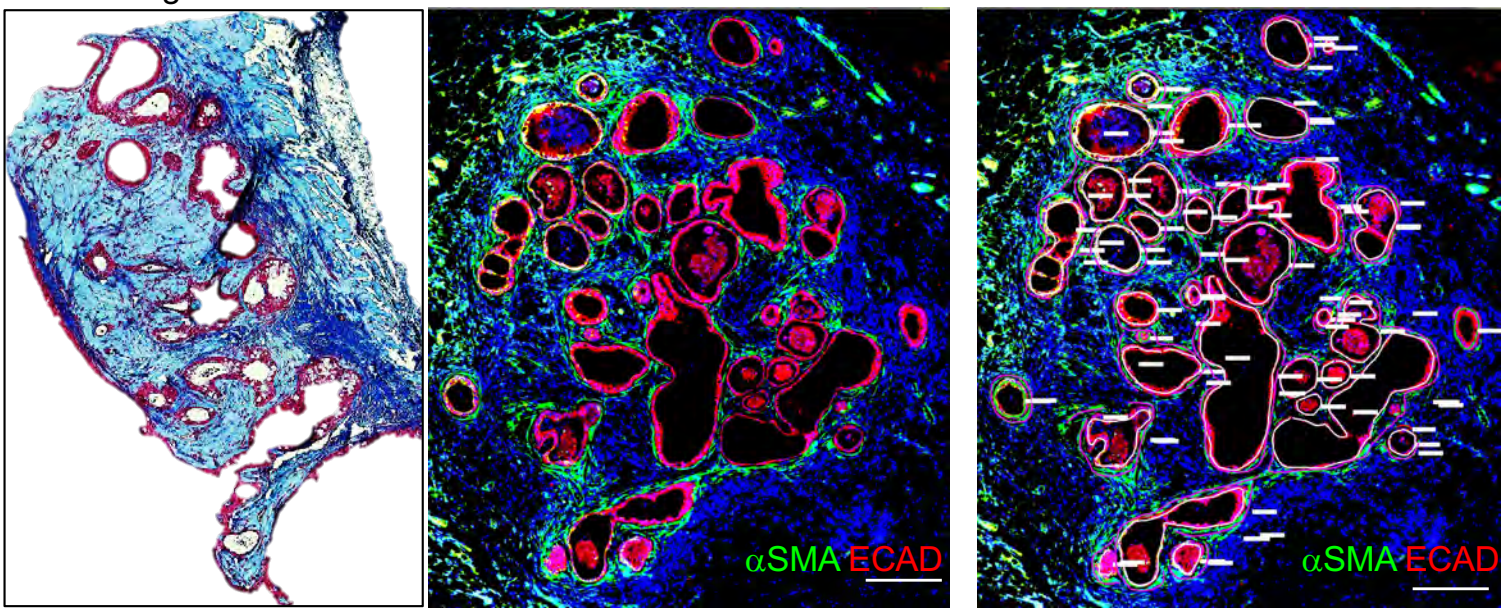
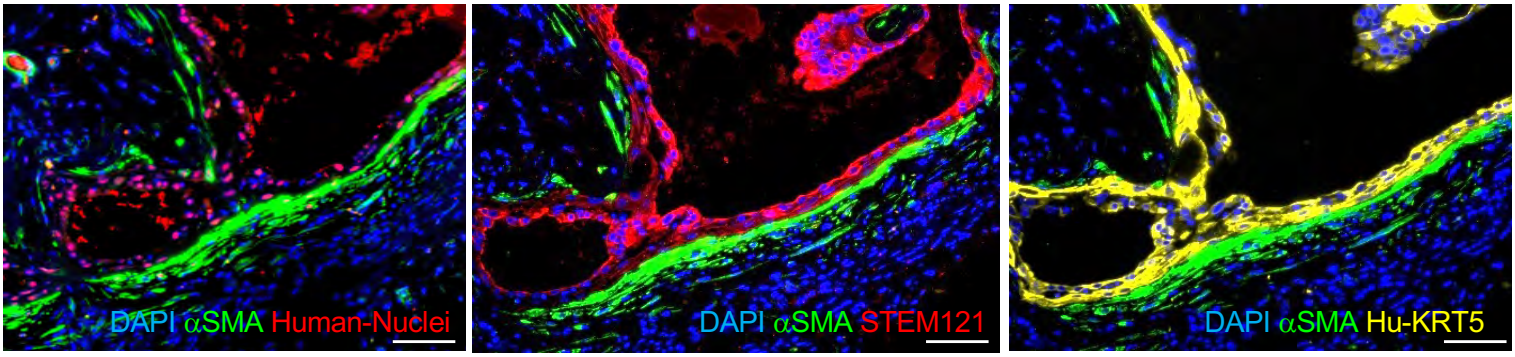


Fig. S2. Fibrosis in xenografts and morphometric analysis.
Left, Histological sections of xenograft nodules formed four weeks after transplantation of stem cell libraries from IPF (F04, F05, F07) stained Masson's trichrome (*left*). *Middle*, immunofluorescence with the indicated antibodies (E-cadherin, ECAD (red); anti-alpha smooth muscle actin, aSMA (green) (*Middle*). *Right*, Immunofluorescence micrographs were assessed by morphometric measurements on the lengths of ECAD+ epithelia (white line) and αSMA+ myofibroblasts (red line). Scale bar, 500 μm.

Fig. S3

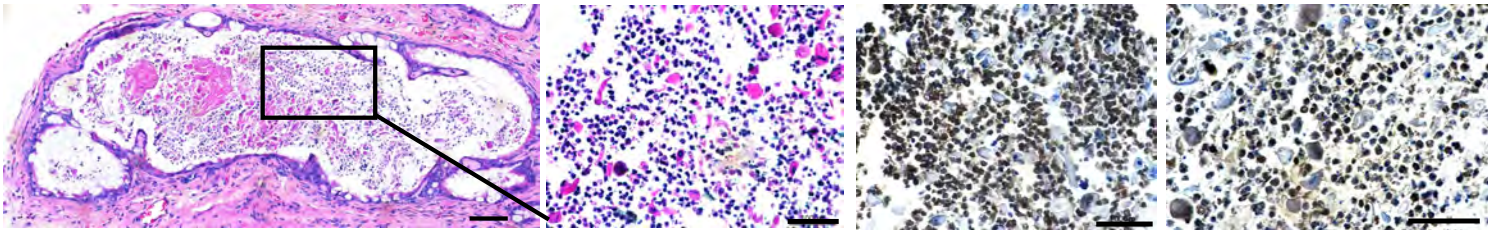
A

Murine origin of myofibroblasts in xenografts



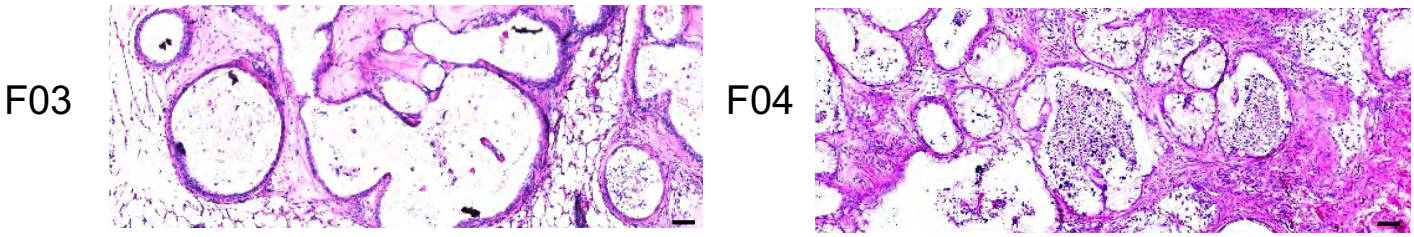
B

COPD library xenograft



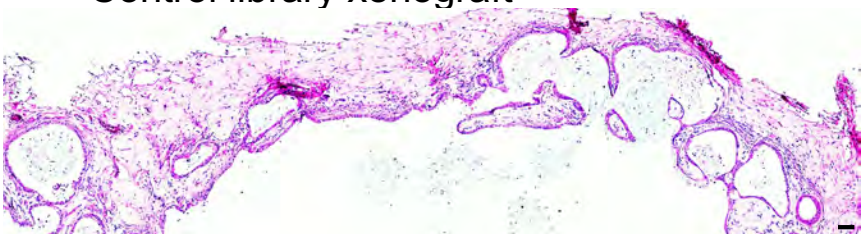
C

IPF library xenografts



D

Control library xenograft



E

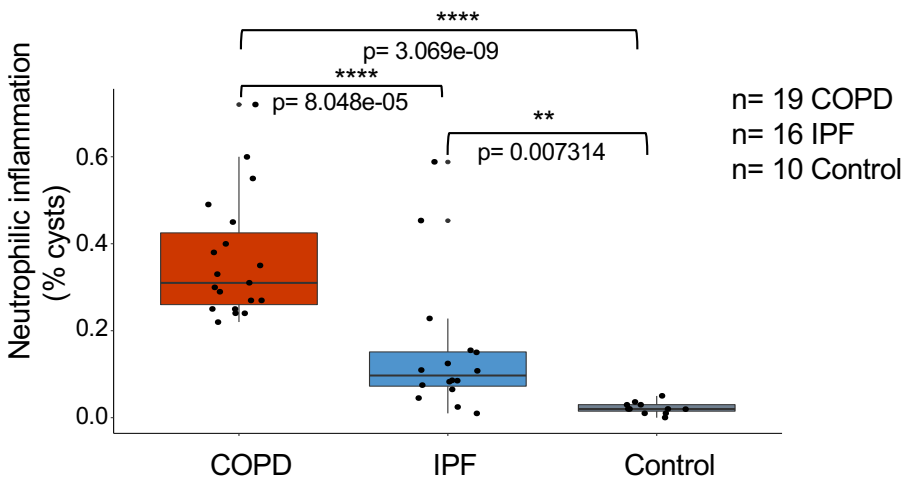


Fig. S3. Characterization of IPF library xenografts.
A. αSMA+ myofibroblasts in xenografts are of murine origin. Immunofluorescence staining of nodule sections Cluster B xenograft probed with three different, human-specific antibodies including Nucleoli Marker Antibody (NM95; red), STEM121 (red), Human-Krt5 (yellow), as well as antibodies to αSMA (green). Scale bar, 100 μm. **B.** Neutrophils in xenografts from COPD, IPF and Control cases. Histological sections through nodules formed after transplantation of COPD stem cell libraries presenting epithelial cysts marked by abundant infiltration of CD45+/Ly6G+ neutrophils. Scale bar, 100 μm. **C.** Xenograft nodules derived from transplantation of libraries from IPF cases F03 and F04, the latter of which shows neutrophil infiltration. Scale bar, 100 μm. **D.** Histological section of nodule from a control library transplant. Scale bar, 100 μm. **E.** Box-Whisker plot of neutrophilic infiltration across 19 COPD, 16 IPFs and 10 control library xenografts.

Fig. S4

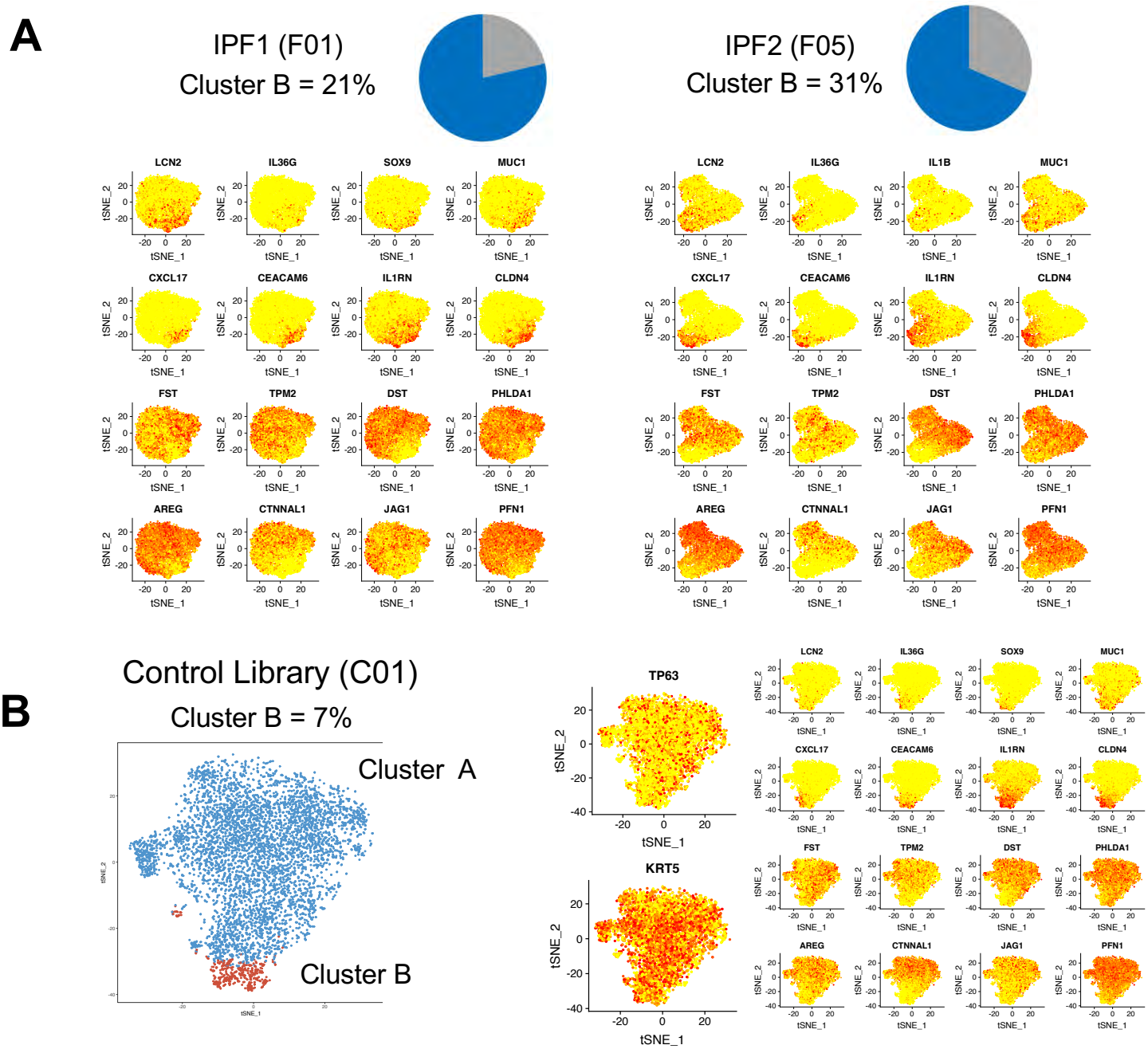
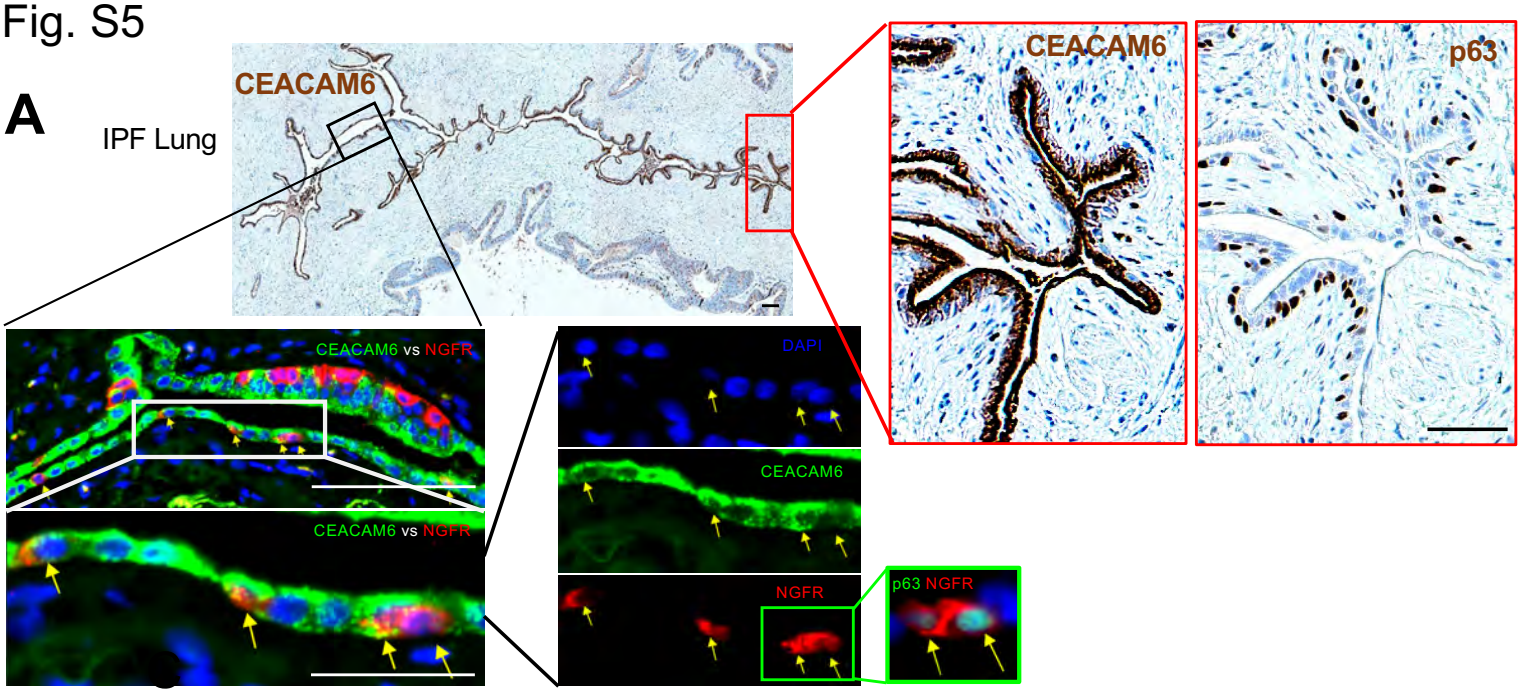
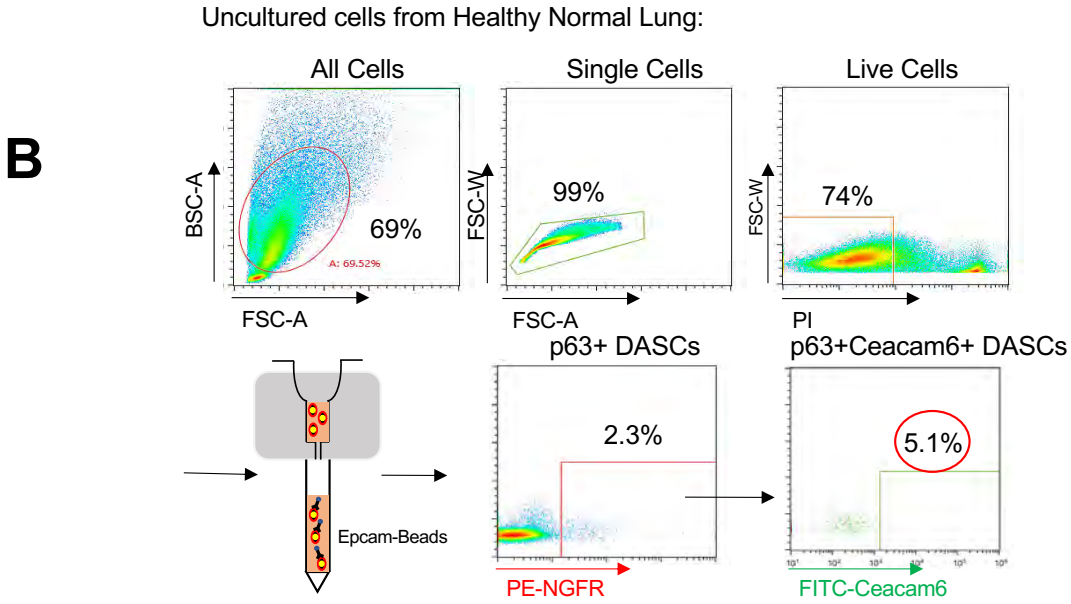

A
IPF1 (F01)
Cluster B = 21%
IPF2 (F05)
Cluster B = 31%
LCN2
IL36G
SOX9
MUC1
IL1B
CXCL17
CEACAM6
IL1RN
CLDN4
FST
TPM2
DST
PHLDA1
AREG
CTNNAL1
JAG1
PFN1
tSNE_1
tSNE_2
B
Control Library (C01)
Cluster B = 7%
Cluster A
Cluster B
TP63
KRT5
C
Controls
Cluster B = 7%
Cluster B = 10%
Cluster B = 9.6%

Fig. S4. Marker Genes in scRNAseq Profiles of IPF and control Libraries.
A. Mapping of cluster-specific expression marker genes onto tSNE profiles of scRNAseqdatasets from IPF lungs. Pie charts show the fractions of Cluster A (blue) and Cluster B (gray) cells in the respective libraries. **B.** Mapping of cluster-specific expression marker genes onto tSNE profiles of scRNAseq datasets from a control lung. Distribution of airway epithelial stem cell markers TP63 and KRT5 are shown. **C**. tSNA plots of three control lungs shows dominant normal distal lung stem cell (blue) and minor populations of variants including IPF Cluster B (gray), iSCM (red), and GCM (green).

Fig. S5


A
IPF Lung
CEACAM6
CEACAM6
p63
CEACAM6 vs NGFR
CEACAM6 vs NGFR
DAPI
CEACAM6
NGFR
p63 NGFR


B
Uncultured cells from Healthy Normal Lung:
All Cells
Single Cells
Live Cells
BSC-A
FSC-A
69%
FSC-W
FSC-A
99%
FSC-W
PI
74%
Epcam-Beads
p63+ DASCs
2.3%
PE-NGFR
p63+Ceacam6+ DASCs
5.1%
FITC-Ceacam6

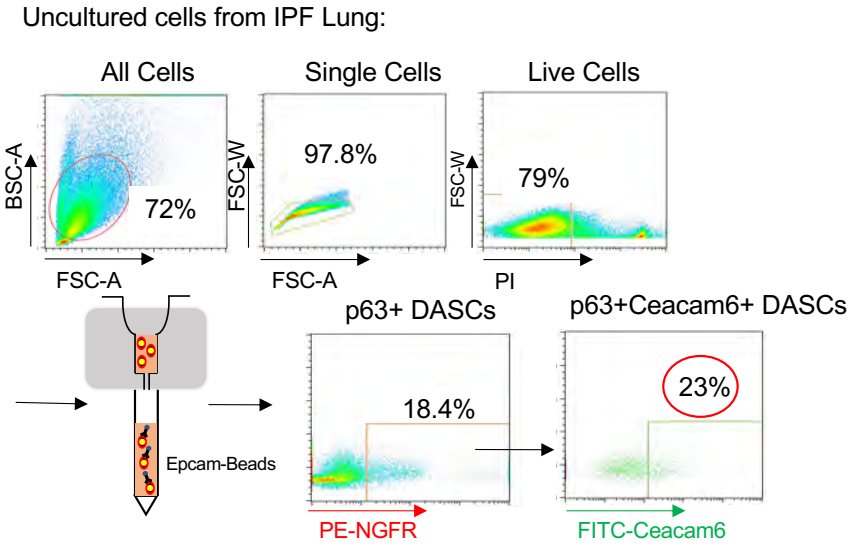
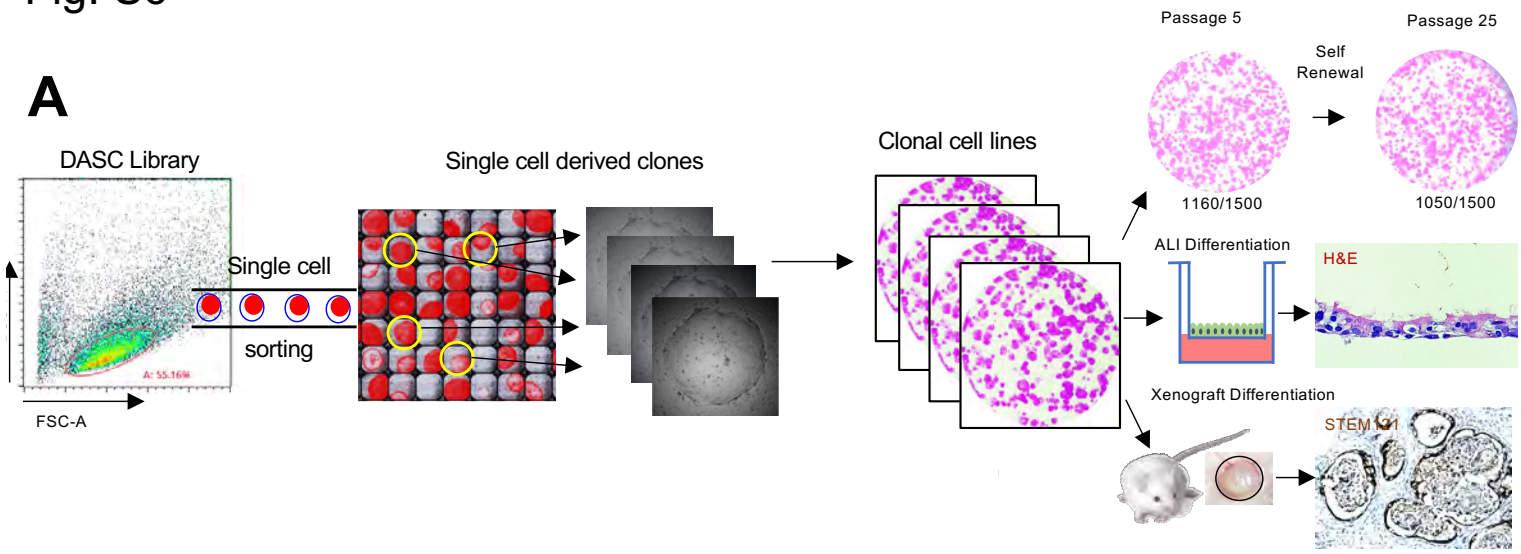
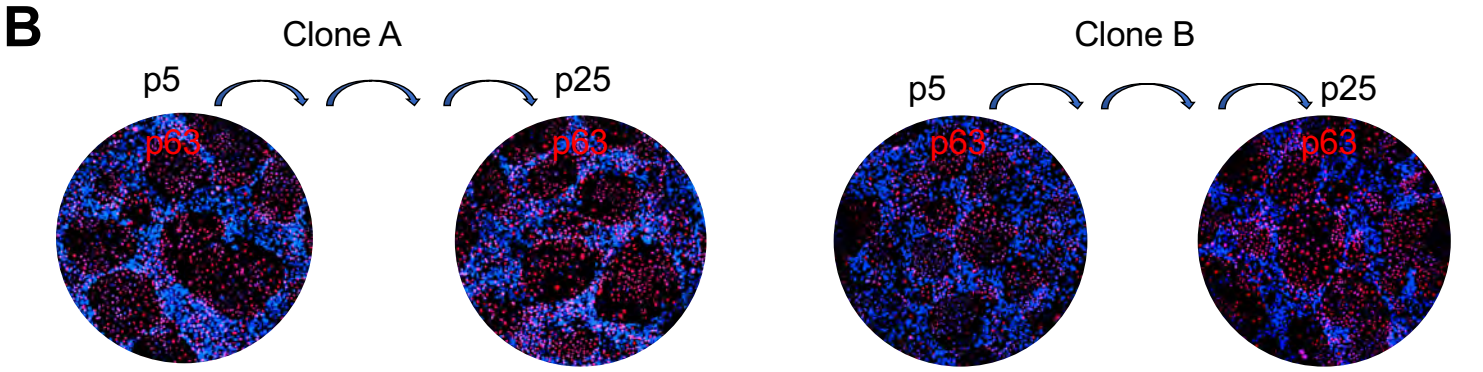
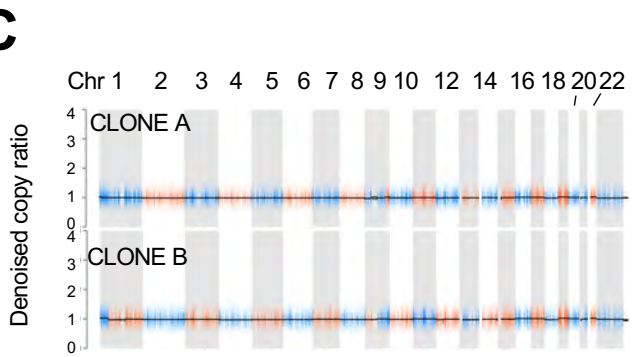

Uncultured cells from IPF Lung:
All Cells
Single Cells
Live Cells
BSC-A
FSC-A
72%
FSC-W
FSC-A
97.8%
FSC-W
PI
79%
Epcam-Beads
p63+ DASCs
18.4%
PE-NGFR
p63+Ceacam6+ DASCs
23%
FITC-Ceacam6

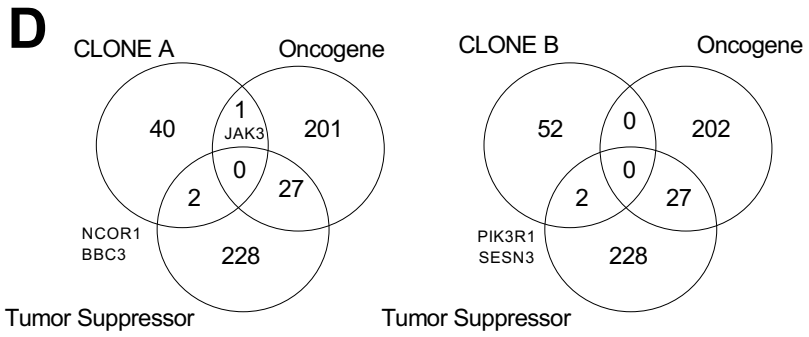
Fig. S5. Clonogenic p63+CEACAM6+ Cluster B cells in situ in IPF lung.
A. Immunohistochemical detection of CEACAM6 in the IFP lung. *Clockwise,* CEACAM6 IHC reveals differentially staining epithelia in the IPF lung. Immunofluorescence with antibodies to CEACAM6 and NGFR indicates CEACAM6 expression is in both differentiated and basal cells. **B.** Quantification of CEACAM6+/NGFR+ cells in control (Top) and IPF (Bottom). Viable singlets from total cell suspensions are isolated by FACS sorting, and selected by Epcam columns for epithelial identify. NFGR+ cells are gated and further analyzed for high CEACAM6 expression. 5.1% of NGFR+ cells show high CEACAM6 expression in normal lung, whereas the IPF lung this number is 23%.

Fig. S6


A
DASC Library
Single cell
sorting
FSC-A
Single cell derived clones
Clonal cell lines
Passage 5
Passage 25
Self
Renewal
1160/1500
1050/1500
ALI Differentiation
H&E
Xenograft Differentiation
STEM121


B
Clone A
p5
p25
p63
Clone B
p5
p25
p63


C
Chr 1 2 3 4 5 6 7 8 9 10 12 14 16 18 20 22
Denoised copy ratio
CLONE A
CLONE B


D
CLONE A
Oncogene
40
1
JAK3
201
0
2
27
NCOR1
BBC3
228
Tumor Suppressor
CLONE B
Oncogene
52
0
202
0
2
27
PIK3R1
SESN3
228
Tumor Suppressor

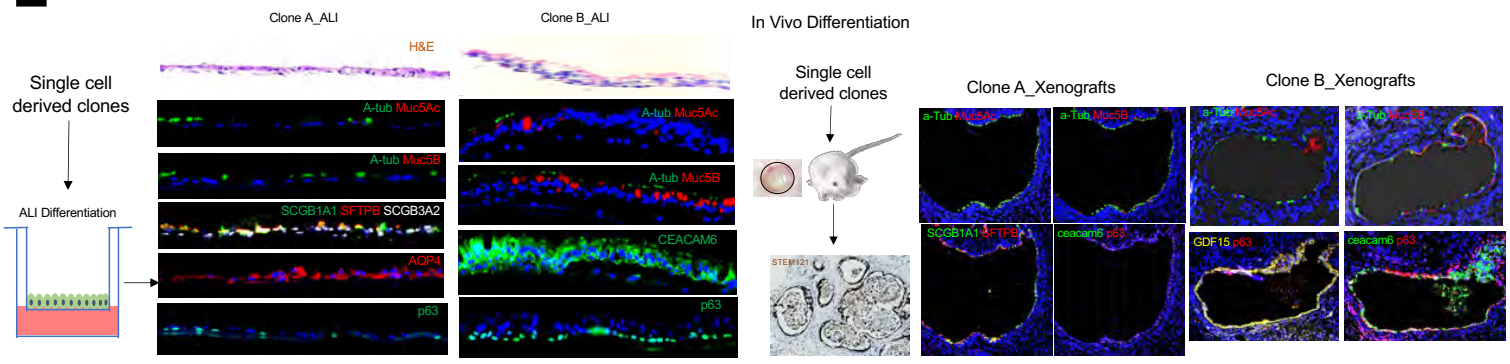
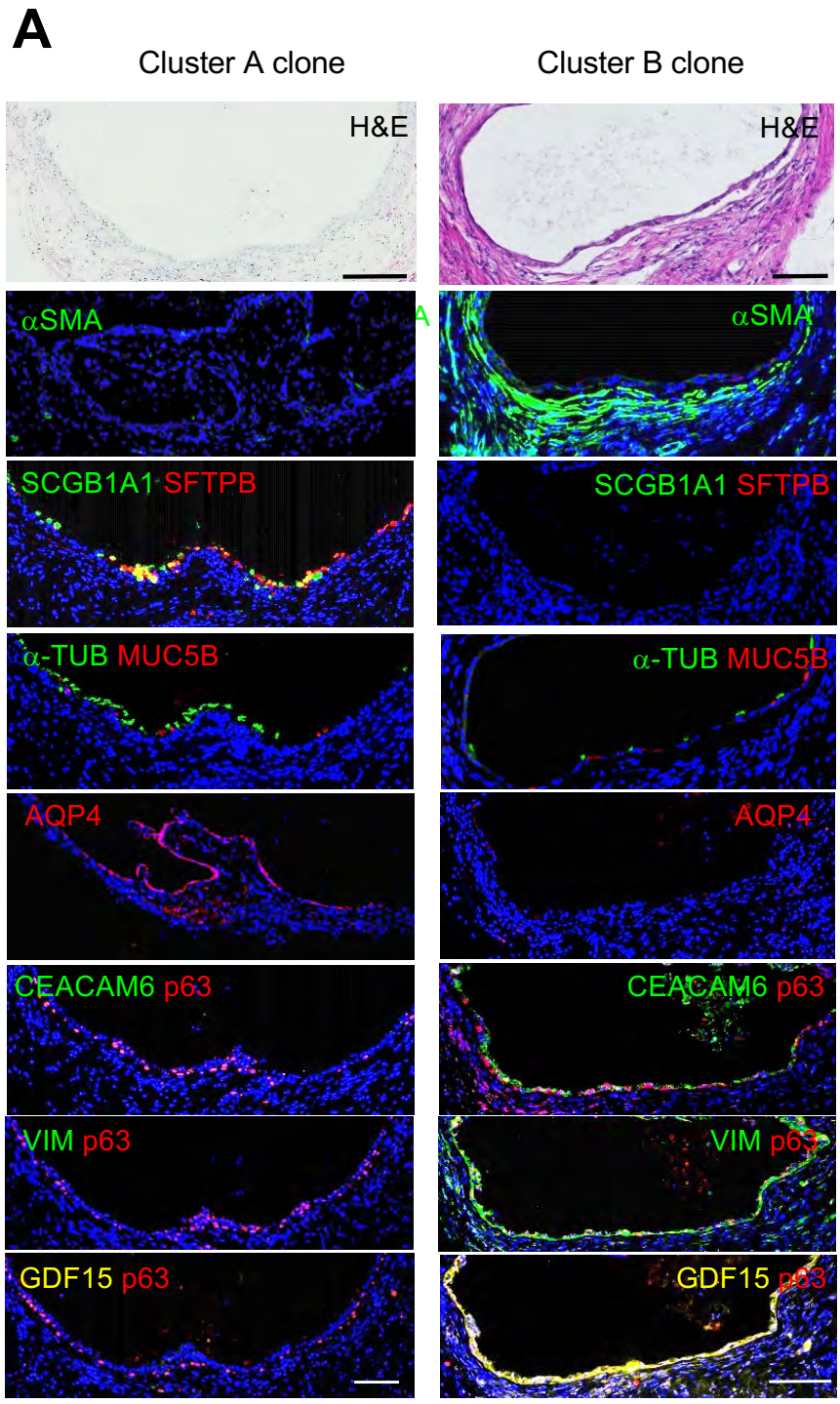
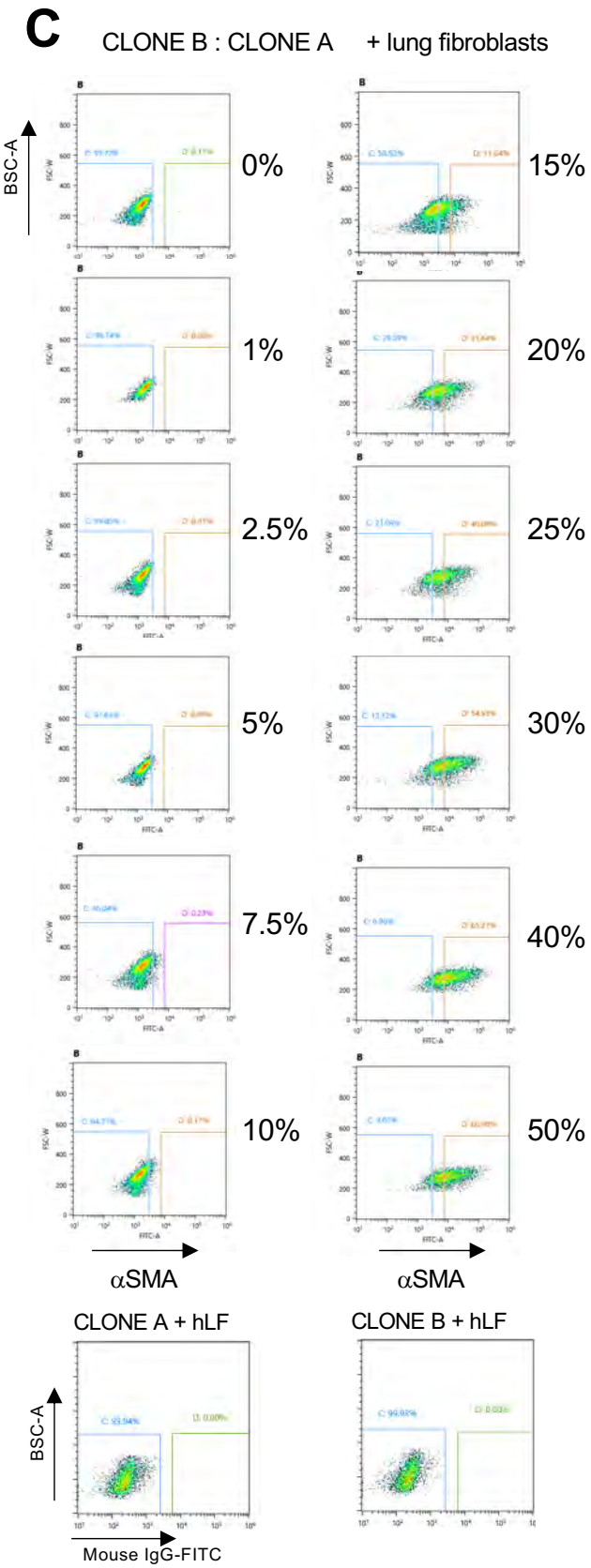

E
Clone A_ALI
Clone B_ALI
H&E
Single cell
derived clones
ALI Differentiation
A-tub Muc5Ac
A-tub Muc5B
SCGB1A1 SFTPB SCGB3A2
AQP4
p63
CEACAM6
In Vivo Differentiation
Single cell
derived clones
Clone A_Xenografts
Clone B_Xenografts
STEM121
GDF15 p63
ceacam6 p63

Fig. S6. Basal cells in IPF lung.
A. General schematic for the isolation and clonal analysis of basal cells from the patient-specific libraries for long-term self-renewal, multipotent differentiation, and *in vivo* function via xenografts. **B.** Long-term self-renewal via serial passaging *in vitro* on 3T3-J2 feeder cell cultures stained with DAPI and antibodies to TP63. **C.** Copy number variation profile derived from whole exome DNA sequencing of Cluster A and Cluster B cells at passage 25. **D.** Absence of nonsynonymous SNPs related to 201 oncogenes or 238 tumor suppressor genes in Cluster A and Cluster B cells at passage 25. **E.** Left, *In vitro* differentiation of clone A and clone B from IPF lung. Clone A cells give rise to an epithelium that is subtended by p63+ cells and suprabasal cells that show variable expression of alpha-tubulin, SCGB1A1, SCGB3A2, SFTPB, and AQP4. Clone B cells show variable expressed of MUC5AC and MUC5B, high expression of CEACAM6, and are subtended by p63+ cells. *Right*, Xenografts of Clone A and Clone B cells into immunodeficient mice yield bronchiolar (SCGB1A1+, SFTPB+) epithelia and an aberrant epithelia marked by IPF markers GDF15 and CEACAM6, respectively.

Fig. S7


A
Cluster A clone
Cluster B clone
H&E
H&E
αSMA
αSMA
SCGB1A1 SFTPB
SCGB1A1 SFTPB
α-TUB MUC5B
α-TUB MUC5B
AQP4
AQP4
CEACAM6 p63
CEACAM6 p63
VIM p63
VIM p63
GDF15 p63
GDF15 p63


C
CLONE B : CLONE A + lung fibroblasts
BSC-A
0%
15%
1%
20%
2.5%
25%
5%
30%
7.5%
40%
10%
50%
αSMA
αSMA
CLONE A + hLF
CLONE B + hLF
BSC-A
Mouse IgG-FITC

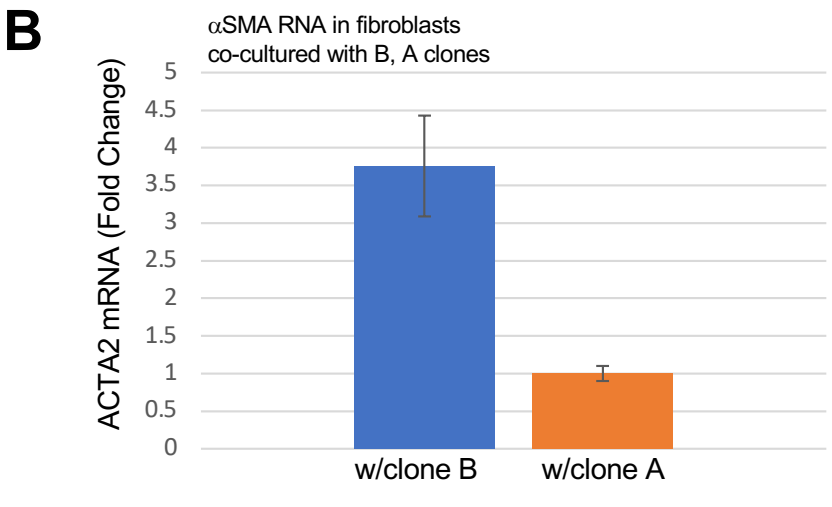

B
αSMA RNA in fibroblasts
co-cultured with B, A clones
ACTA2 mRNA (Fold Change)
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
w/clone B
w/clone A

Fig. S7. Clonal xenografts and *in vitro* fibrogenesis. A. H&E staining of histological sections of xenografts derived from patient-matched Cluster A and B clones, as well as immunofluorescence micrographs reporting marker gene expression. Scale bar, 100 μm. **B.** Quantitative PCR of the ACTA2 transcript encoding αSMA in purified fibroblasts previously co-cultured with either Cluster B or Cluster A clones. **C.** FACS profiling of αSMA expression in fibroblasts previously co-cultured with programmed ratios of Cluster B vs Cluster A cells, with Cluster B cell percentages listed.

Fig. S8

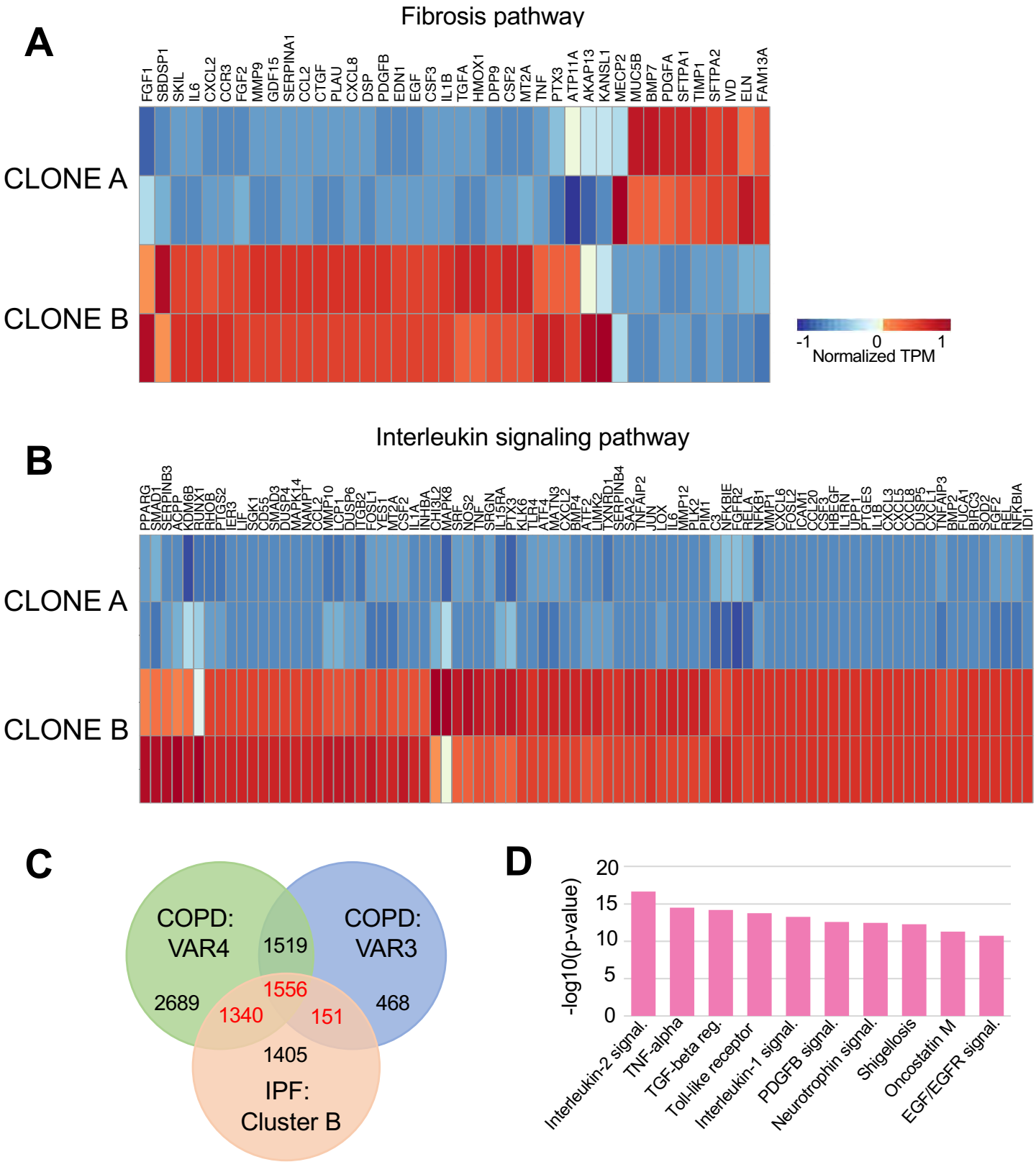

A
Fibrosis pathway
FGF1 SBDSP1 SKIL IL6 CXCL2 CCR3 FGF2 MMP9 GDF15 SERPINA1 CCL2 CTGF PLAU CXCL8 DSP PDGFB EDN1 EGF CSF3 IL1B TGFA HMOX1 DPP9 CSF2 MT2A TNF PTX3 ATP11A AKAP13 KANSL1 MECP2 MUC5B BMP7 PDGFA SFTPA1 TIMP1 SFTPA2 IVD ELN FAM13A
CLONE A
CLONE B
-1 0 1
Normalized TPM
B
Interleukin signaling pathway
PPARG SMAD1 SERPINB3 SCPP KDM6B RUNX1 RHOB PTGS2 IER3 LIF SGK1 CD55 SMAD3 DUSP4 MAPK14 NAMPT CCL2 MMP10 CP1 DUSP6 ITGB2 FOSL1 YES1 MT2A CSF2 IL1A NHBA CHI3L2 MAPK8 SRF NOS2 TNF SRGN IL15RA PTX3 KLK6 TLR4 ATF4 MATN3 CXCL2 BMP4 ATF2 LIMK2 TXNRD1 SERPINB4 SAA2 TNFAIP2 JUN LOX IL6 MMP12 PLK2 PIM1 C3 NFKBIE FGFR2 RELA NFKB1 MMP1 CXCL6 FOSL2 ICAM1 CCL20 CSF3 HBEGF IL1RN UPP1 PTGES IL1B CXCL3 CXCL5 CXCL8 DUSP5 CXCL1 TNFAIP3 BMP2 FUCA1 RPS3 SOD2 FGF2 REL NFKBIA IDI1
CLONE A
CLONE B
C
COPD: VAR4
COPD: VAR3
IPF: Cluster B
2689
1519
468
1340
1556
151
1405
D
-log10(p-value)
0 5 10 15 20
Interleukin-2 signal.
TNF-alpha
TGF-beta reg.
Toll-like receptor
Interleukin-1 signal.
PDGFB signal.
Neurotrophin signal.
Shigellosis
Oncostatin M
EGF/EGFR signal.

Fig. S8. Expression profiling of IPF clones.
A. heatmap of genes linked to *Fibrosis* (NCATS BioPlanet) in *in vitro* differentiated IPF Cluster A and Cluster B cells. **B.** Expression heatmap of genes linked to *Interleukin* (NCATS BioPlanet) in *in vitro* differentiated IPF Clone A and Clone B cells. **C.** Venn diagram showing overlap of differentially expressed genes in COPDVAR3, COPDVAR4, and IPF Clone B. **D.** Histogram detailing pathway analysis of the genes that overlap among indicated cells.

Fig. S9

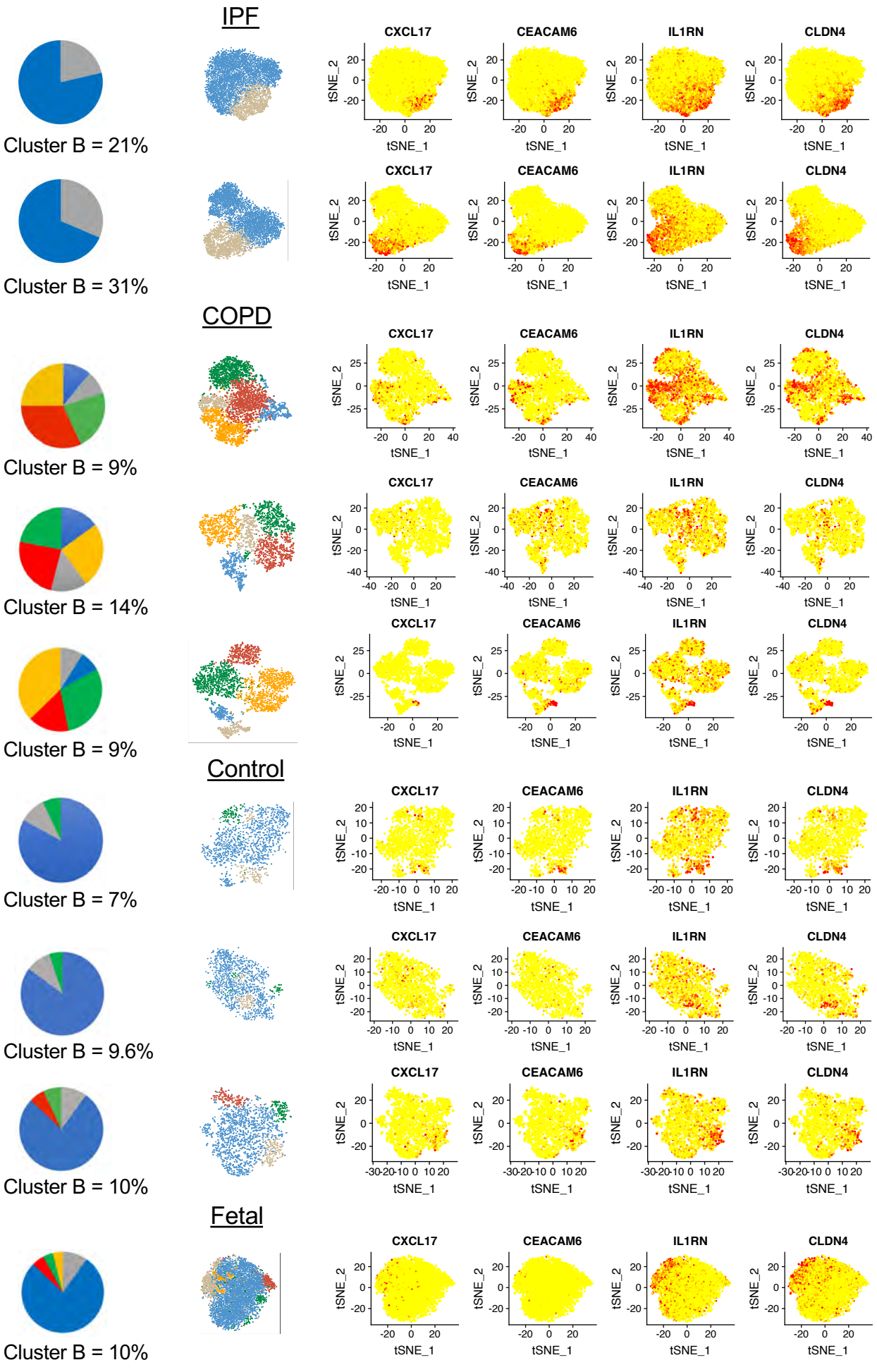

IPF
COPD
Control
Fetal
CXCL17
CEACAM6
IL1RN
CLDN4
tSNE_1
tSNE_2
Cluster B = 21%
Cluster B = 31%
Cluster B = 9%
Cluster B = 14%
Cluster B = 9%
Cluster B = 7%
Cluster B = 9.6%
Cluster B = 10%
Cluster B = 10%

Fig. S9. tSNE profiles of IPF, COPD, Control, and Fetal Lung Libraries.
From top, tSNE profiles of scRNAseq data from libraries of IPF lung, COPD lung, control lung, and human fetal lung. Associated pie charts depict the estimated percentages of the normal basal cell (Cluster A, blue), IPF Cluster B (gray), squamous cell metaplasia (SCM, orange), inflammatory squamous cell metaplasia (iSCM, red), and goblet cell metaplasia (GCM, green).

Fig. S10

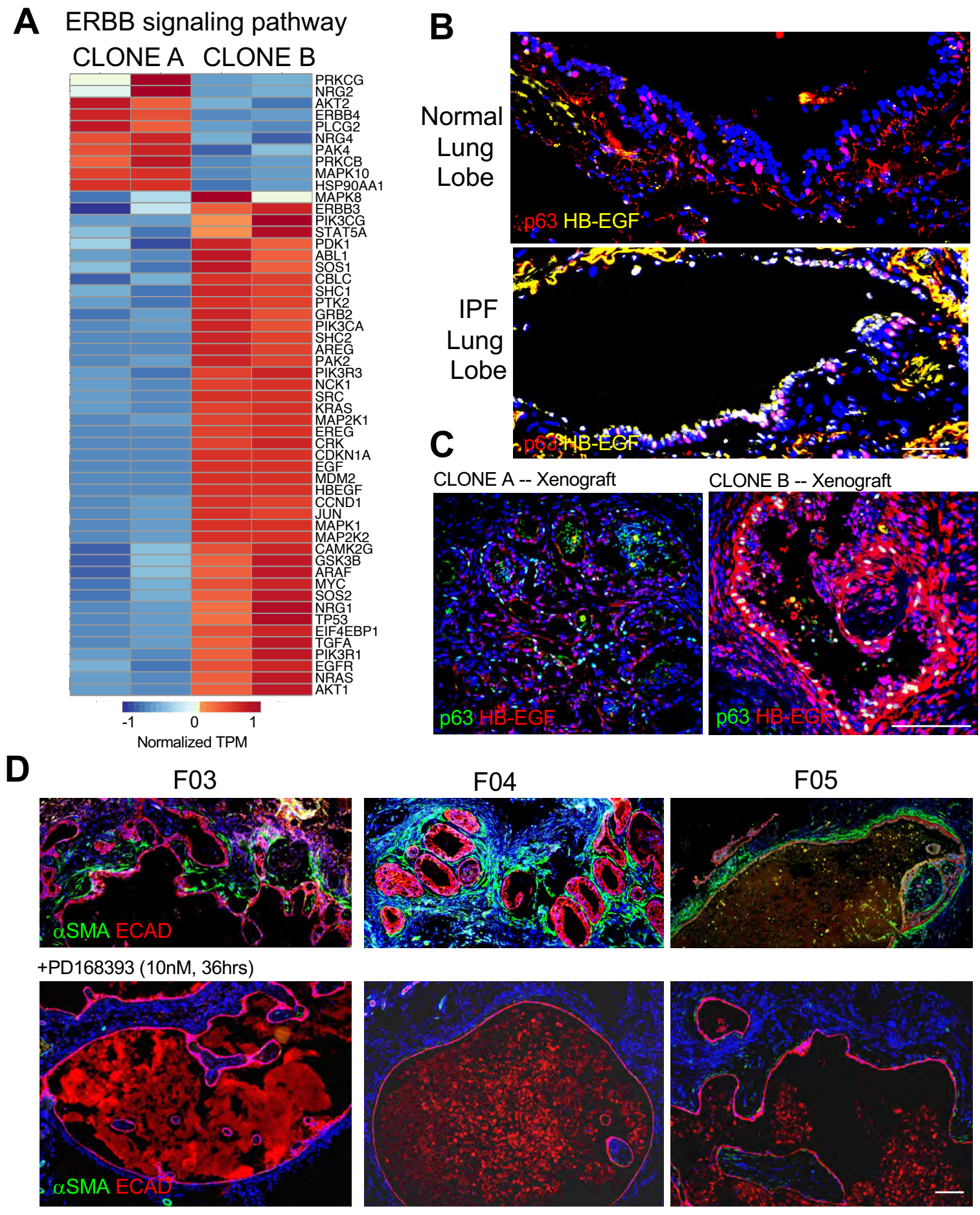

A
ERBB signaling pathway
CLONE A
CLONE B
PRKCG
NRG2
AKT2
ERBB4
PLCG2
NRG4
PAK4
PRKCB
MAPK10
HSP90AA1
MAPK8
ERBB3
PIK3CG
STAT5A
PDK1
ABL1
SOS1
CBLC
SHC1
PTK2
GRB2
PIK3CA
SHC2
AREG
PAK2
PIK3R3
NCK1
SRC
KRAS
MAP2K1
EREG
CRK
CDKN1A
EGF
MDM2
HBEGF
CCND1
JUN
MAPK1
MAP2K2
CAMK2G
GSK3B
ARAF
MYC
SOS2
NRG1
TP53
EIF4EBP1
TGFA
PIK3R1
EGFR
NRAS
AKT1
-1
0
1
Normalized TPM
B
Normal Lung Lobe
p63 HB-EGF
IPF Lung Lobe
p63 HB-EGF
C
CLONE A -- Xenograft
CLONE B -- Xenograft
p63 HB-EGF
p63 HB-EGF
D
F03
F04
F05
αSMA ECAD
+PD168393 (10nM, 36hrs)
αSMA ECAD

Fig. S10. ERBB Pathway.
A. Heatmap of genes linked to ERBB signaling pathway in *in vitro* differentiated normal Cluster A cells and IPF Cluster B cells. **B.** p63 (red) and HBEGF (yellow) immunofluorescence in histological sections of control and IPF lung lobes. Scale bar, 100 μm. **C.** Immunofluorescence staining of nodule sections Cluster A and B clone xenografts probed with p63 (green) and HBEGF (red). Scale bar, 100 μm. **D.** Immunofluorescence images of IPF libraries (F03, F04, F05) xenograft nodules using the indicated markers before (*Top*) or after (*Botttom*) *in vitro* exposure to PD168393. Scale bar, 100 μm.

Table S1. IPF and Control Patient Information.

Case I.D.	Hospital	Sex	Age	Histology	Clinical diagnosis	Tobacco History	FVC/FVC%	FEV1/FEV1 %
F01	University of Iowa Hospitals and Clinics	M	64	pathology consistent with IPF and bronchiectasis	IPF/bronchiectasis	Former Smoker 3 pack years	1.43L/35%	1.30L/43%
F02	University of Iowa Hospitals and Clinics	M	62	pathology UIP in both lungs	IPF: UIP	Non Smoker	2.76L/64%	2.39L/74%
F03	University of Iowa Hospitals and Clinics	M	49	pathology UIP in both lungs	IPF: UIP	non Smoker	2.28L/40%	2.07L/47%
F04	University of Iowa Hospitals and Clinics	M	72	pathology extensive interstitial fibrosis in both lungs	IPF	Former Smoker 15pack years	2.68L/56%	2.38L/68%
F05	UT Medical School	M	64	pathology consistent with IPF	IPF	NA	1.56L/33.9%	1.56L/44.8%
F06	UT Medical School	M	69	pathology consistent with IPF	IPF	NA	2.69L/55.9%	1.28L/34.7%
F07	UT Medical School	M	65	pathology consistent with IPF	IPF	NA	2.6L/62.8%	2.15L/69.5%
F08	UT Medical School	M	70	pathology consistent with IPF	IPF	NA	NA	NA
F09	UT Medical School	M	69	NA	IPF	NA	1.68L/39.2%	1.52L/48%
F10	UT Medical School	F	62	NA	IPF	NA	1.6L/49.2%	1.31/53.9%
F11	UT Medical School	M	53	NA	IPF/COPD	NA	1.73L/40.1%	1.59L/47.6%
F12	UT Medical School	F	56	NA	IPF	NA	NA	NA
F13	UT Medical School	F	70	NA	IPF	NA	2.02L/72%	1.61L/76%
F14	UT Medical School	M	59	NA	IPF/PH	NA	2.35L/49%	2.12L/58.2%
F15	UT Medical School	M	49	NA	IPF	NA	1.95L/51%	1.25L/40.9%
F16	UT Medical School	M	67	NA	IPF	NA	2.35L/60.5%	2.01L/69.9%
C01	UT Medical School	F	42	Normal Lung	Donor Lung	Non Smoker	NA	NA
C02	University of Iowa Hospitals and Clinics	NA	NA	Normal Lung	Donor Lung	NA	NA	NA
C03	UT Medical School	M	37	Normal Lung	CVA (stroke)	Non Smoker	NA	NA
C04	University of Iowa Hospitals and Clinics	NA	NA	Normal Lung	Donor Lung	NA	NA	NA
C05	UT Medical School	M	62	NA	Anoxia, cardiovascular	Former Smoker	NA	NA
C06	UT Medical School	M	30	NA	Anoxia, Carbon monoxide poisoning	Non Smoker	NA	NA
C07	UT Medical School	F	37	NA	CVA (stroke)	Non Smoker	NA	NA
C08	UT Medical School	M	23	NA	Blunt head trauma from MVA	Non Smoker	NA	NA
C09	UT Medical School	F	42	NA	CVA (stroke)	Non Smoker	NA	NA
C10	UT Medical School	NA	NA	NA	NA	NA	NA	NA

Table S2. Morphometrics of patient library-derived xenografts.

Case I.D.	Length of E-Cad+ epithelia in Xenografts (μm)	Length of αSMA+ myofibroblasts in Xenografts (μm)	Ratio of Fibrosis in Xenografts (%)
F01	6767	2735	40.4
F02	7052	2615	37.1
F03	16692	5298	31.7
F04	22996	8860	38.5
F05	10219	5943	58.2
F06	18421	6338	34.4
F07	19674	4293	21.8
F08	8294	2678	32.3
F09	11070	4057	36.7
F10	2329	876	37.6
F11	3887	1467	37.8
F12	1819	565	31.1
F13	5639	2235	39.6
F14	3455	1155	33.4
F15	5590	1376	24.6
F16	4429	1916	43.3
C01	12212	532	4.4
C02	11685	462	4.0
C03	9670	460	4.8
C04	3289	218	6.6
C05	5956	294	4.9
C06	6527	341	5.2
C07	4429	281	6.3
C08	10211	607	5.9
C09	3532	228	6.5
C10	2448	129	5.3