JCI insight

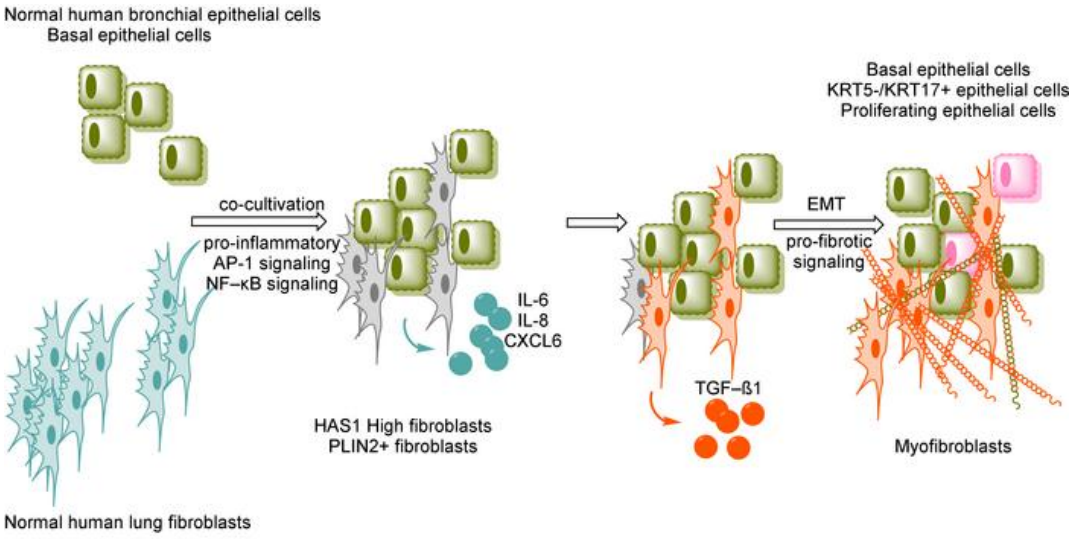
NF-κB drives epithelial-mesenchymal mechanisms of lung fibrosis in a translational lung cell model

Patrick Sieber, … , Maxime Boucher, Oliver Nayler

JCI Insight. 2022. https://doi.org/10.1172/jci.insight.154719.

Research In-Press Preview Cell biology Pulmonology

Graphical abstract



Interaction between bronchial/basal epithelial cells and lung fibroblasts leads to the appearance of epithelial and mesenchymal cell states found in IPF patient lungs.

NF-κB drives epithelial-mesenchymal mechanisms of lung fibrosis in a translational lung cell model

Patrick Sieber[1]*, Anny Schäfer [1], Raphael Lieberherr[1], Silvia L. Caimi[1], Urs Lüthi[1], Jesper Ryge[1], Jan H. Bergmann[1], François Le Goff[1], Manuel Stritt[1], Peter Blattmann[1], Bérengère Renault[1], Patrick Rammelt[1], Bruno Sempere[1], Diego Freti[1], Rolf Studer[1], Eric S. White[2, ‡], Magdalena Birker–Robaczewska[1], Maxime Boucher[1], and Oliver Nayler[1]

[1] Idorsia Pharmaceuticals Ltd., Allschwil, Switzerland

[2] Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, MI, USA

‡ Current address: Boehringer Ingelheim Pharmaceuticals, Inc., Ridgefield, CT, USA

*Corresponding author

E–mail: patrick.sieber@idorsia.com

Abstract

In the progression phase of idiopathic pulmonary fibrosis (IPF) the normal alveolar structure of the lung is lost and replaced by remodeled fibrotic tissue and by bronchiolized cystic airspaces. Although these are characteristic features of IPF, knowledge of specific interactions between these pathological processes is limited. Here, the interaction of lung epithelial and lung mesenchymal cells was investigated in a co–culture model of human primary airway epithelial cells (EC) and lung fibroblasts (FB). Single–cell RNA sequencing (sc–RNA–seq) revealed that the starting EC population was heterogenous and enriched for cells with a basal cell signature. Furthermore, fractions of the initial EC and FB cell populations adopted distinct pro–fibrotic cell differentiation states upon co-cultivation, resembling specific cell populations that were previously identified in lungs of IPF patients. Transcriptomic analysis revealed active nuclear factor NF–kappa–B (NF–κB) signaling early in the co–cultured EC and FB cells and the identified NF–κB expression signatures were also found in “HAS1 High FB” and “PLIN2+ FB” populations from IPF patient lungs. Pharmacological blockade of NF–κB signaling attenuated specific phenotypic changes of EC and prevented FB–mediated interleukin–6 (IL6), interleukin–8 (IL–8) and C–X–C motif chemokine ligand 6 (CXCL6) cytokine secretion, as well as collagen alpha–1(I) chain (COL1A1) and alpha–smooth muscle actin (α–SMA) accumulation. Thus, we identified NF–κB as a potential mediator, linking epithelial pathobiology with fibrogenesis.

Introduction

IPF, the most severe form of pulmonary fibrosis, is a chronic progressive disease leading to respiratory failure and death or lung transplantation with a median survival of 6–8 years from diagnosis. IPF is an incurable disease and currently available treatments can only slow the rate of decline of forced vital capacity (FVC) of the lung (1). The disease is characterized by pathological epithelial remodeling, accumulation of extracellular matrix (ECM) in the peripheral lung, accompanied by the destruction of functional alveolar gas exchange structures (2).

Disease progression may involve continuous proliferation of dysfunctional reprogrammed bronchoalveolar EC (1). The timing and nature of triggering and subsequent reprogramming of lung EC, however, need further exploration (1, 3). Recent sc–RNA–seq–based studies have identified specific cell populations that were enriched in IPF patients compared with control lungs (4-8). While EC play a major role in the formation of bronchiolized cystic airspaces, myofibroblasts and related mesenchymal cells are mainly responsible for the synthesis of ECM components and fibrotic distortion of the lung. To improve our understanding of the pathological epithelial–mesenchymal interaction, we set out to model the “epithelial–mesenchymal module” in vitro by co–culturing EC and FB cells. EC and mesenchymal cells have previously been cultured together, resulting in various outcomes that can be attributed to the particular cell types and applied cell culture protocol, but also to the frequently added pro–fibrotic stimuli (9-17).

Here, we identified and optimized in vitro conditions that reproducibly recapitulate, chronologically phased and in the absence of added artificial stimuli, a pro–fibrotic phenotype shift in the cultured cells. Upon direct EC/FB contact, co–culture of primary normal human lung fibroblasts (NHLF) and normal human bronchiolar epithelial cells (NHBE) led to the robust activation of a fibrogenic program in the co–cultivated FB. FB responded with a strong initial inflammatory response, followed by transforming growth factor beta (TGF–ß)–mediated synthesis of ECM components (i.e. COL1A1) and increased α–SMA, which are considered

hallmarks of FB–to–myofibroblast transition. Preceding myofibroblast differentiation, co–culture induced activation of NF–κB and activator protein 1 (AP–1), with associated downstream gene expression changes, were evident already after 3 hours in FB and EC. Using sc–RNA–seq and genomic data analysis, we show that co–cultured EC and FB progress to cell phenotypes that are also enriched in IPF patients, and we identified expression of NF–κB and AP–1 gene modules in distinct cell populations of IPF patients, highlighting the translational potential of our model.

Pharmacological inhibition of NF–κB and AP–1 signaling revealed their driving role in the early stages of fibrogenesis in the co–culture model, suggesting that NF–κB and AP–1 signaling might also contribute to the pathobiological process in IPF and could be a potential target for pharmacological intervention.

Results

EC induce FB–to–myofibroblast transformation

Previously, we identified common genomic signatures between samples from IPF patient lungs and bleomycin–treated rat lungs (18). To expand on this work, expression of selected genes was investigated on consecutive histological sections that were prepared from lung biopsies of IPF patients and non–IPF control lungs, using RNAscope in–situ hybridization (ISH). Areas with detectable gene expression exhibited distribution patterns that were generally more abundant in patients (Supplemental Figure 1, Supplemental Figure 2). Cells expressing the epithelial markers integrin subunit beta 6 *(ITGB6)* and integrin subunit alpha V *(ITGAV)* overlayed foci with high *COL1A1* expression in lung regions considered to have active disease (Figure 1, Supplemental Figure 2). In particular, *ITGB6* and *COL1A1* expression was detected in cells that were in very close proximity to each other but appeared to be mutually exclusive (Figure 1). Collectively, we found that EC and FB/myofibroblasts appeared in very close contact in areas of active disease in IPF patient lungs, which is in accordance with published data (4-6).

These observations prompted us to investigate potentially pathological interactions between EC and FB, and we co–cultivated these two cell types in vitro. To this end, NHBE cells were pre–stained with CellTracker Deep Red and co–cultured together with NHLF in a 2D tissue culture plate. As controls, NHBE and NHLF were cultivated in monocultures either in absence of any supplements (veh) or monocultured NHLF were treated with TGF–ß1, respectively, to induce myofibroblast transition. After 5 to 8 days the cells were fixed and stained with DAPI and an anti–α–SMA antibody for image analysis (Figure 2A).

Remarkably, in co–culture, NHBE cells induced pronounced transition of NHLF to α–SMA–positive myofibroblasts. The effectiveness was comparable to NHLF that were stimulated with

the known profibrotic mediator TGF–ß1 in absence of NHBE (Figure 2A).The effective induction of FB–to–myofibroblast transition in co–culture was confirmed after cell segmentation into α–SMA–negative FB or α–SMA–positive myofibroblasts, using a trained classifier (Figure 2B).

When co–cultivated with other epithelial cell types NHLF also differentiated into cells with a myofibroblast–like phenotype, albeit with different effectiveness (Supplemental Figure 3, A and B). This was not, or only weakly, observed when NHLF were co–cultured with NHLF, smooth muscle cells or endothelial cells (Supplemental Figure 3C).

Surprisingly, a very low number of NHBE cells (i.e. 200 per well) was sufficient to induce a robust induction of α–SMA expression in the maximum number of tested NHLF (i.e. 4’000 per well; Supplemental Figure 3D). The effects were comparable when IPF–patient–derived human lung fibroblasts (IHLF) instead of NHLF were used as effector cells (Supplemental Figure 3, A and C).

To complement our microscopy–based analysis results, we applied MS/MS–based protein analysis to accurately quantify cellular COL1A1 (herein abbreviated with COL1 for simplicity) and α–SMA protein levels from co–culture lysates (19) at different time–points after seeding. COL1 and α–SMA served as endpoints for ECM production and contractility, respectively. In NHLF/NHBE co–cultures a significant increase in α–SMA protein was first detected after 98 hours, whereas a significant increase in COL1 protein was already observed 50 hours after seeding, suggesting that these markers may be controlled by different regulatory mechanisms (Figure 2, C and D).

To test whether the EC–induced FB into myofibroblast transition required direct cell–to–cell contact, we performed a trans–well experiment. In this experiment NHLF and NHBE were seeded either together in the lower chamber or individually in the upper or lower chamber. This allowed co–cultivation of cells in the same chamber but without physical contact. Notably,

increased COL1 accumulation after 98 hours was only observed when NHBE and NHLF were seeded in direct contact (Figure 2E).

To confirm that secreted factors were not driving EC–induced activation of FB into myofibroblasts, conditioned monoculture supernatants from either NHBE or NHLF were added to FB and cell responses were quantified using impedance technology. This methodology detects FB to myofibroblast transformation with high sensitivity and accuracy (19). Neither the addition of supernatant of NHBE nor of vehicle–treated NHLF led to any detectable impedance change (Supplemental Figure 4A). Interestingly, the addition of supernatant from NHLF/NHBE co–culture to NHLF led to a transient increase in impedance with a peak after 4h (Supplemental Figure 4). This is indicative of a signaling event and suggested the presence of a co–culture–specific factor, which was absent in supernatants of monocultured cells. Nevertheless, the addition of this secreted co–culture factor was, in contrast to supernatant from TGF–ß1–treated NHLF, insufficient to trigger an increase of α–SMA and COL1 in NHLF (Supplemental Figure 4, B and C).

In summary, the results indicate that EC, through cell–to–cell contact with FB, trigger transformation of FB into active myofibroblast in co–cultures.

Co–cultures of EC and FB induce pro–inflammatory and pro–fibrotic signaling pathways

To follow cell–type–specific temporal changes of gene expression in co–cultured NHBE and NHLF, we used fluorescence–activated cell sorting (FACS) and quantitative real–time polymerase chain reaction (qRT–PCR) analysis. To this end, pre–labelled NHBE (CellTrace Violet) and NHLF (CellTrace Far Red) were harvested at different time points from either monocultured or co–cultured conditions and sorted by FACS. Both dyes were readily absorbed by the cells, stained the cells for the full duration of the experiment (i.e. 98 hours) and remained specific for each cell type in co–culture conditions (Figure 3A). To study gene expression

changes we first used a limited set of literature–based fibrosis–associated and cell–type marker genes (6, 18, 20).

The FACS sorted NHBE expressed the epithelial marker gene cadherin 1 *(CDH1)* (Figure 3B), which was absent from NHLF populations. The FB marker vimentin *(VIM)* was predominant in FACS sorted NHLF (Figure 3B), demonstrating that both pre–labelled cell types were successfully separated after co–culture.

NHLF, when co–cultured with NHBE, showed a temporal expression increase of genes encoding ECM proteins and proteins associated with a myofibroblast phenotype, such as actin alpha 2, smooth muscle (*ACTA2* coding for α–SMA), collagen type I alpha 1 chain *(COL1A1)*, fibronectin 1 *(FN1)*, elastin *(ELN)*, but also of *TGFB1* (coding for TGF–ß1) and cellular communication network factor 2 (*CCN2*, coding for CTGF; Figure 3C). These genes showed similar kinetics with a significant increase between t = 18h and t = 50h, which is in line with the fibrotic effector function of NHLF in this assay. NHBE in monoculture or in co–culture, expressed, in addition to *ITGB6,* a set of genes coding for secreted pro–inflammatory or pro–fibrotic proteins, tumor necrosis factor *(TNF)*, platelet derived growth factor subunit A *(PDGFA)*, and endothelin 1 *(EDN1).* Expression of these genes transiently increased, reaching a peak at t = 50h, and subsequently decreased to baseline at t = 98h (Figure 3C). Interestingly, with time, expression of *ACTA2* and *COL1A1* was also detected in co–cultured NHBE, but not monocultured NHBE (Figure 3C), suggesting EMT in co–cultured conditions. Co–cultured NHLF were the main contributing cell type to the expression of *TGFB1, CCN2* as well as to the cytokines *CXCL6, IL6* and *CXCL8*. Remarkably, *CXCL6, CXCL8* and *IL6* expression was elevated already at t = 3h in co–cultured NHLF compared to monocultured NHLF. In summary, cell–type–specific analysis of gene expression in NHLF/NHBE co–cultures revealed induction of pro–inflammatory and pro–fibrotic gene expression.

Early gene expression changes reveal NHBE–induced inflammatory responses in NHLF

To learn more about the initiating stimulus in co–cultures we performed a comprehensive transcriptional characterization of the early events of the co–culture. NHLF and NHBE cells were pre–stained (as above) and grown individually or as co–cultures. Cells were collected at the time points t = 0h, 3h and 18h, FACS–sorted to produce pure NHBE–derived and NHLF–derived cell population samples (Supplemental Figure 5) and analyzed by bulk RNA–seq.

In a first step, we identified differentially expressed genes (DEG) with a linear fold–change |LinFC| > 1,5 and a false discovery rate (FDR) < 0,05 in the co–culture and the respective monoculture at t = 3h and t = 18h, respectively. For all four conditions (i.e., NHLF, NHBE, NHLF–CC, and NHBE–CC), genes were included if expression was significantly different from expression at time t = 0h. Compared to the corresponding monoculture, 1599 (t = 3h) and 2730 (t = 18h) genes were specifically differentially regulated in the co–cultured NHLF (NHLF–CC) and 1364 (t = 3h) and 1632 genes (t = 18h), respectively, in co–cultured NHBE (NHBE–CC; Figure 4A, Supplemental Table 1). Canonical pathway analysis, as implemented in the Ingenuity Pathway Analysis (IPA) application, was performed on the DEG that were specific for the co–cultures to identify top–ranking canonical pathways (Figure 4A and Supplemental Table 2). IPA upstream regulator analysis on the co–culture specific DEG revealed that already after 3 hours both NHLF–CC and NHBE–CC mounted inflammation– and TLR–mediated signaling, i.e. via TNF, cytokines of the IL–1 family and IL–6, with the concurrent activation of the transcription factors NF–κB and STAT3 (Figure 4B; Supplemental Table 3; Supplemental Table 4).

To identify co–regulated gene clusters with similar temporal expression specifically during the early phases of co–culture, noise–robust soft clustering based on the fuzzy c–means algorithm was performed on the gene expression time series data, excluding genes with low expression (i.e. max(count) transcripts per kilobase million [TPM] < 1). 6289 DEG in NHLF–CC and 6295

DEG in NHBE–CC were subjected to hierarchical clustering and seven clusters were estimated as the most suitable number of clusters for each cell population. Hence, for both NHLF–CC and NHBE–CC, a set of seven clusters based on common gene expression kinetics were generated (Supplemental Figure 6). Next, a gene set overexpression analysis (GSOA) was performed to test whether certain biological functions or processes –, or the consensus binding sites of transcription factors (TFs), were enriched in the genes associated with the co–culture–specific time series clusters.

Querying the Hallmark (MSigDB) collection identified the gene set "TNF–α signaling via NF–κB" as highly enriched in NHLF–CC – as well as in NHBE–CC clusters that showed markedly increased gene expression between t = 0h and t = 3h (i.e. Figure 4C; Supplemental Figure 6, A, F, G, H, M and N). Query of JASPAR (21) and TRRUST (22) collections, using the R package hypeR (23) with gene sets provided by Enrichr (24), revealed an over–representation of TF consensus binding sites and TF–targets of the NF–κB – as well as AP–1 family TFs in the co–expressed genes of those clusters that were characterized by the enrichment of the gene set "TNF–α signaling via NF–κB", (i.e. Supplemental Figure 6A, G, H, and N).

We next examined whether the findings of the pathway analysis were also reflected in the gene expression of the individual NF–κB family members. Indeed, the expression of the NF–κB family members nuclear factor kappa B subunit 1 *(NFKB1),* – subunit 2 *(NFKB2),* REL proto–oncogene, NF–κB subunit *(REL),* and RELB proto–oncogene, NF–κB subunit (*RELB)* showed co–culture specific upregulation at t = 3h and t = 18h (Figure 4D). Furthermore, in NHLF–CC, the expression of the AP–1 transcription factor subunits Jun proto–oncogene *(JUN)*, JunB proto–oncogene *(JUNB)*, Fos proto–oncogene *(FOS)*, and Fos like 2 *(FOSL2)* were upregulated at t = 3h vs t = 0h (Figure 4D) and in co–culture vs monoculture. Activating transcription factor 3 *(ATF3),* – 4 *(ATF4),* and – 5 (*ATF5)* expression was increased at t = 3h vs t = 0h and increased in NHLF–CC vs NHLF at t = 3h and/or t = 18h (Figure 4D). Several AP–1 and NF–κB family subunits appeared upregulated at t = 3h in both mono– and co–

cultured NHBE, indicating that in NHBE expression changes of these genes is affected by culture rather than co–culture (Supplemental Figure 7).

In summary, our expression and GSOA analysis of the early time points of the NHBE/NHLF co–culture indicates that co–cultivation triggers a strong acute inflammatory response, particularly in the NHLF–CC, with increased expression and/or the activation of the transcription factors NF–κB, AP–1 and STAT3.

NF–κB signaling during the early phase of the co–culture is essential for the fibrogenic activation of fibroblasts

To address the translational relevance of this in vitro co–culture system to investigate anti–fibrotic treatment modalities, the two IPF standard–of–care treatments, nintedanib and pirfenidone, were assessed. To this end, compounds were added at the time of seeding of NHLF and NHBE cells and changes in protein expression were quantified 98 hours later using MS/MS.

Nintedanib showed a dose–dependent inhibition of both COL1 and α–SMA proteins, at concentrations for which no cytotoxicity was observed in a combined viability and toxicity assay on CHO cells (Figure 5; Supplemental Figure 8C; Supplemental Table 5). Pirfenidone inhibited COL1 but not α–SMA accumulation and this was only observed and statistically significant at the highest tested concentration (i.e. 10 mM; Figure 5B; Supplemental Table 5). Hence, both nintedanib and pirfenidone were able to inhibit COL1 production.

TGF–ß1 is one of the best–described inducers of pro–fibrotic signaling and its expression is upregulated in the co–culture after 18 hours (Figure 3C). The known TGF–ß–receptor inhibitor EW–7197 (25) potently and concentration–dependently inhibited co–culture–induced COL1 and α–SMA protein accumulation (Figure 5C; Supplemental Table 5).

Cell–type–dependent phosphorylation and activation of STAT3 has been attributed to IL–6 signaling via its receptors (26). The two inhibitors of IL–6/STAT3 signaling, LLL12 (27) and

255 stattic (26), inhibited the accumulation of COL1 and α–SMA (Figure 5, D and E; Supplemental
256 Table 5) at concentrations that led to a reduction in the metabolic activity, but this was not
257 associated with cytotoxicity (Supplemental Figure 8, F and G).

258 AP–1 and NF–κB signatures were identified as very early upregulated signaling pathways in
259 the co–culture model. BAY 11–7082 is a non–selective inhibitor of NF–kappa–B inhibitor alpha
260 (IκBα]) kinase IKK1 and IKK2, blocks the TNF–inducible phosphorylation of IκBα and, thus, is
261 expected to inhibit all forms of NF–κB (28), whereas T–5224 is a known inhibitor of protein c–
262 FOS (c–FOS)/AP–1 signaling (29). T–5224 and BAY 11–7082 effectively attenuated co–
263 culture–induced COL1 and α–SMA accumulation (Figure 5, F and G; Supplemental Table 5)
264 at concentrations previously shown to inhibit NF–κB and c–FOS/AP–1 in vitro (28, 29). The
265 effects were not associated with cytotoxicity (Supplemental Figure 8). When compound
266 exposure was restricted to the first 18h of co–culture and was then removed, both, BAY 11–
267 7082 and T–5224, completely and concentration–dependently abolished COL1 and α–SMA
268 accumulation at t = 98h with a potency comparable to the standard protocol with continued
269 exposure (Figure 5, H and I).

270 AP–1 and NF–κB have been found to enhance the expression of numerous genes involved in
271 the inflammatory process (30). As assessed by Luminex assays, treatment with the NF–kB
272 inhibitor BAY 11–7082 applied from t = 0h to t = 18h and then removed, strongly inhibited the
273 accumulation of IL–6, IL–8 and CXCL6 in the culture supernatant at t = 98h (Supplemental
274 Figure 9B), whereas c–FOS/AP–1 inhibition with T–5224 did not influence their levels
275 (Supplemental Figure 9C). TNF was below the limit of detection.

276 In summary, the results obtained with the pharmacological inhibitors provide supportive
277 evidence for an active and causal contribution of TGF–ß–, AP–1–, NF–κB–, and STAT3–
278 mediated signal transduction with a critical involvement of c–FOS/AP–1 and NF–κB during the
279 early phase of the co–culture. Furthermore, NF–κB signaling during the early phase of the co–

culture was essential for the accumulation of the cytokines IL–6, IL–8 and CXCL6, whereas c–FOS was dispensable.

Transcriptional expression profiles of in vitro co–cultured NHLF and NHBE resemble those of IPF–specific lung cell populations

We further evaluated the fate of NHBE–CC and NHLF–CC during co–cultivation and compared the genomic profiles of these cells at different time points with genomic profiles of cells that were isolated from patients with IPF.

To this end, droplet–based sc–RNA–seq of NHBE–CC and NHLF–CC suspensions, collected at time points t = 0h, 3h, and 50h, was performed. After normalization and integration of the data sets, we performed graph–based clustering towards identifying cell states across the different experimental variables. This analysis recovered 14 cell sub–state clusters across NHBE–CC and NHLF–CC lineage cells (Figure 6A). Each of the 6 EC – and the 8 FB cell sub–states was named after one of the most significantly associated genes for each state (see Methods for details).

Annotation of the Uniform Manifold Approximation and Projection (UMAP) embedding of the vehicle–treated cells according to the time of collection showed that cells of all sub–states were collected at each time point (Figure 6B). To capture any underlying dynamic processes or to identify trajectories between cell sub–states, we applied trajectory inference methods, also called pseudo–time analyses (31). Pseudo–time analysis, performed using the Slingshot (32) package, generated branched trajectories connecting the clusters of the FB and EC populations, respectively (Figure 6C). For the EC population, Slingshot created a branched trajectory along which cells of the EC_COL6A1 and EC_COL1A2 sub–states separated from the main EC population trajectory at EC_FBXO2 (Figure 6C). These two distinct branches segregated the 6 distinct EC sub–states based on the presence or absence of *COL1A1* and

tumor protein p63 *(TP63)* expression (Figure 6E). Pseudo–time analysis using Monocle 3 (33) allowing for more refined resolution of cell states (clusters) revealed temporal gene expression changes within the sub–states of both populations, i.e., FB and EC (Figure 6D).

To evaluate the consequences of the gene expression dynamics, we determined the similarity of the co–culture system with a human IPF patient lung cell atlas (5). Of the published reference data set, we used the IPF patient–derived and control lung mesenchymal and epithelial types to generate a “reference” atlas with reduced dimensionality keeping the author’s original cell type labeling (Figure 7, A and B).

To compare our in vitro culture with human lung data, we calculated Pearson correlations of the averaged expression of shared variably expressed genes – and, used SingleR (34) to calculate, at the single–cell level, Spearman correlations of the expressed marker genes that were shared between reference and query cells. At t = 0h, some NHLF–CC query cells paired up with “Fibroblasts” but across all time points the FB population matched most closely the reference populations “HAS1 High FB”, “PLIN2+ FB”, “Myofibroblast”, and “Smooth Muscle Cells” (Figure 7, C and D; Supplemental Table 6). If only the human reference population with the highest similarity to the respective query cell is considered, the NHLFs showed the highest similarity to “Myofibroblast” at baseline (94%) and at t = 50h (98%; Table 1). At the intermediate time point, i.e. 3 hours after seeding, 50% and 21% of the query FB population showed highest similarity to the "HAS1 High FB" and the “PLIN2+ FB”, respectively (Table 1). At t = 3h, the cells resembling the “HAS1 High FB” are composed mainly of the FB_LTBP1 (49%) state, which also contributes most (37%) of the cells to the pool showing similarity to “PLIN2+ FB”. Since the classical myofibroblast marker genes showed upregulation only from t = 50h (Figure 3), the unbiased correlation revealed that the NHLFs displayed gene expression patterns of myofibroblast from the outset and thus may have adopted a “pre–myofibroblast” state during cultivation. Furthermore, our data suggest that the "HAS1 High FB" and "PLIN2+ FB” populations may represent intermediate stages in the course of transforming into fully differentiated ECM–synthesizing myofibroblast. For the joint EC query population,

the best Pearson correlations were obtained for “Basal EC”, “KRT5–/KRT17+ EC”, and “Proliferating EC” from IPF patients (Figure 7C; Supplemental Table 6). SingleR revealed similarity of the query EC population with the “Basal EC” population already from the start of the culture (Figure 7D). At t = 50h of cultivation, fractions of the EC cell query population (i.e. 8%, 6%, and 4%), adopted states most closely resembling the “Proliferating EC”, “Myofibroblast”, and the “KRT5–/KRT17+ EC” populations from patients, respectively (Table 1). The pools resembling “Proliferating EC” were mainly composed of EC_KRT18 (69%) and EC_FBXO2 (28%) and the “KRT5–/KRT17+ EC” of EC_COL1A2 (25%), EC_COL6A1 (15%), EC_FBXO2 (25%), and EC_KRT18 (30%). All correlations for the query at the level of the sub–states are provided in the Supplemental Data (Supplemental Table 7; Supplemental Table 8; Supplemental Table 9).

In summary, our analysis revealed a pre–myofibroblast profile for the NHLF cells at the beginning of culture. Subsequently, the majority of NHLFs adopted "HAS1 High FB" and "PLIN2+ FB” identity before their complete conversion into myofibroblasts. The NHBE cells consisted largely of cells with basal cell signature and the majority retained basal cell characteristics, with fractions of them assuming the character of "Proliferating EC”, "Myofibroblast”, and "KRT5–/KRT17+ EC”. Therefore, transcriptional profiles of in vitro co–cultured NHLF and NHBE resemble cell populations that are also enriched in the lungs of ILD patients.

NF–κB contributes to early molecular changes in co–cultured fibroblast and epithelial cell populations

We next investigated whether NF–κB or AP–1 would be involved in changes to the EC – or FB populations, i.e. whether inhibition of IKK/NF–κB and c–FOS/AP–1 with BAY 11–7082 and T–5224, respectively, had potential to modulate EC or FB phenotypes.

As determined by MAST (35) both the culture time of 50 hours and the two pharmacological interventions changed the relative frequencies of the EC and FB cell populations (Figure 8A). In the EC population, both treatments significantly decreased the dominant basal EC sub–states characterized by low or absent *COL1A1* and high *TP63* expression (i.e. EC_KRT18, EC_FBXO2 and EC_KLK10; Figure 8A) and, consequently, increased the proportion of the *COL1A1* expressing transitional sub–states (i.e. EC_COL1A2 and EC_COL6A2). Among the FB sub–states, both treatments reduced the proportion of FB_COL1A1, which is characterized by elevated expression of *ACTA2, CCN2* and *IL6*. Furthermore, BAY 11–7082 and T–5224, blunted the expression increase of *ACTA2, CCN2* and *IL6* within the FB_COL1A1 substate (Supplemental Figure 10A).

GSOA for the DEG of cellular sub–states did not provide significant insights, likely due to the small population sizes. GSOA on the total EC population at t = 3h, revealed overexpression of functions indicative of development, cell junction organization, cell–cell contact, adhesion, differentiation, cornification, and keratinization (Figure 8B; Supplemental Figure 11). The addition of BAY 11–7082 suppressed these functions, indicating that these biological processes are largely controlled by NF–κB signaling. Twenty four of the hundred most significantly affected EC genes by each treatment (FDR<0.01) were commonly affected by BAY 11–7082 and T–5224 (Supplemental Figure 10B).

The two *COL1A1* expressing EC sub–states (EC_COL1A2 and EC_COL6A2) showed a dynamic and pronounced expression increase of *KRT5, KRT14, KRT17, KRT18*, *KRT19,* indicative of ongoing differentiation, keratinization, and cornification. Furthermore, they showed increased expression of genes linked to development (*SOX9*), senescence (*SOX4*), Notch signaling (*HES4*), and EMT (i.e. *COL1A1, TGFB1, CCN2, ELN, ACTA2*, Figure 8C). In cluster “EC_COL1A2”, treatment with BAY 11–7082 prevented the increase of all detected keratins, whereas T–5224 blunted the expression increase of *KRT8, KRT17, KRT18*, and –*19*; Figure 8C). The dominant *TP63* expressing basal EC sub–states (e.g. “EC_KRT18”) exhibited relatively stable keratin expression (Supplemental Figure 10C) and keratin

expression was not markedly affected by the treatments. Hence, inhibition of IKK/NF–κB and c–FOS/AP–1, modulated FB and EC phenotypes, as judged by their effect on population composition, biological pathway activities, and keratin profiles. The culture and treatment effects were restricted to certain cell states and were, thus, context dependent.

IPF patient lung–derived FB populations express the AP–1 and NF–κB signatures that were observed in the early stages of NHLF and NHBE co–cultured cells

To assess whether AP–1 and NF–κB signatures that we identified in our co–culture model were also expressed in IPF patients, a module score analysis (36) was performed on the reference human IPF patient lung cell atlas (5). In a first step, the reference data set was partitioned into control lung vs. interstitial lung disease (ILD; Figure 9A). We then assembled modules of AP–1 – and NF–κB associated genes that showed similar expression kinetics and increased expression after 3 hours in co–culture. The AP–1 module consisted of the genes *JUN, JUNB, FOS, FOSL2, ATF3, ATF4*, and *ATF5*, and the NF–κB module comprised *NFKB1, NFKB2, REL, RELA*, and *RELB*. In addition, we examined a TNF module, consisting of 84 co–culture genes overlapping with the Hallmark gene set "TNF–α signaling via NF–κB" identified as enriched in the NHLF–CC after 3 hours (Figure 4C). Expression module analysis for all 3 tested gene sets showed highest expression scores for the “HAS1 High FB”, as well as for the “PLIN2+ FB”, which are both more abundant in ILD lungs compared with control lungs (Figure 9B; Supplemental Figure 12).

Interactome–based analysis identifies potential ligand–receptor binding pairs specific to EC–FB interactions

Our results obtained thus far indicate that cell–to–cell contact between EC and FB is required. To identify potential ligand–receptor binding pairs involved in the activation of NHLF we

407 performed interactome analysis of the sc–RNA–seq data of the vehicle–treated cells at t = 0h
408 using CellPhoneDB (37). To specifically identify EC–FB pairings, the dataset was cleared of
409 those combinations that also occurred between EC–EC and FB–FB. Interactions between
410 cells and matrix (integrins) were also removed since the pairing process takes place prior to
411 significant ECM deposition. The generated interaction networks for FB and EC ligand–
412 receptor pairs were visualized with Cytoscape ((38); Figure 10). As central mechanisms by
413 which FB can interact with EC our analysis implies growth factor–driven signaling through
414 fibroblast growth factor (FGF) –, bone morphogenic protein (BMP) –, activin –, and receptor
415 tyrosine-protein kinase erbB-3 (ERBB3) receptors, respectively, as well as the frizzled class
416 of receptors, and the Leucine-rich repeat-containing G-protein coupled receptor 4 (LGR4) co–
417 receptor for Wingless (Wnt) signaling proteins. Semaphorin-7A *(SEMA7A)* and *SEMA4A*,
418 besides TNF receptor superfamily member 9 *(TNFSF9)*, are plasma membrane localized and
419 could, hence, require cell–cell contact to engage their receptors (Figure 10). In addition to the
420 *SEMA7A*> Plexin-C1 *(PLXNC1)* and *SEMA4A>PLXND1* pairs that were identified at t = 0h,
421 *SEMA3A*, *SEMA3B*, and *SEMA4B* showed increased and NHBE–CC–specific expression at
422 t = 3h and could potentially interact with their corresponding plasma membrane–localized
423 plexin receptors, or neuropilin co–receptors (i.e. *NRP2*) expressed on NHLF–CC
424 (Supplemental Table 1). Therefore, EC–driven effects on FB may potentially involve, in
425 addition to TGF–ß–, Wnt–, and cytokine–mediated responses, specific semaphorin–mediated
426 signaling via plexin receptors.

Discussion

Dysregulated epithelial–mesenchymal interactions have been proposed as a disease–causing mechanism of IPF (3). However, there has been limited progress in developing an integrated understanding of the origin of the disease, as well as of central mechanisms driving pathological epithelial remodeling and ECM expansion in lung fibrosis patients. We developed an in vitro co–culture model using primary NHBE and NHLF cells that allowed us to investigate the dynamic molecular interactions between both cell types. Surprisingly, over a period of several days, co–culturing both cell types resulted in the effective differentiation of NHLF into fibrotic mesenchymal effector cells, whereas NHBE initiated functions related to development, cell–cell contact, differentiation, keratinization and eventually EMT. Notably, only epithelial cell types effectively triggered the differentiation of co–cultured FB into α–SMA–positive myofibroblast in our high–content assay and specifically lung NHBE were the most effective in this assay.

Remarkably, the fibrotic process proceeded autonomously and without addition of external pro–fibrotic stimuli, e.g TGF–ß1, which is consistent with published data using similar conditions (11). However, when EC were used in 3D co–cultures, they seemed to rather dampen and not promote the activation of co–cultured FB (10, 12, 16). We also did not observe an increase in COL1/α–SMA when EC–FB interaction was studied in “EC culture medium”, but the decreased tubulin content suggested that NHLFs did not tolerate these culture conditions well (data not shown). Hence, specific culture conditions can influence the outcome of epithelial–mesenchymal interactions, which is in line with previous reports (12-14, 16).

The fact that both cell types must be seeded in the same compartment suggests direct cell-cell contact as part of the driving mechanism. In the healthy lung, direct cell–to–cell contact between airway EC and FB is normally prevented by the basement membrane. However, in IPF, the extensive breakdown of lung architecture in areas of active disease may allow the

453 establishment of uncontrolled contacts between the EC and FB, making it plausible that these
454 two cell types directly influence each other and drive disease progression (5, 15, 39). Our
455 interactome–based analysis has identified several potential ligand–receptor pairs that may
456 mediate specific cell–to–cell signaling between EC and FB and these will be evaluated in
457 future studies. Semaphorin signaling, as one of the possible contributors, has not been
458 thoroughly explored in the context of lung fibrosis and may provide an opportunity as novel
459 drug target. Semaphorins play a critical role in angiogenesis, but have also been reported to
460 modulate tumor microenvironment and to influence the biology of cancer–associated FB (40).

461 As revealed by our sc–RNA–seq study the commercially available NHBE preparations that
462 were used in the co–culture were significantly enriched for cells with a basal EC signature
463 already at the time of seeding. Since the inflammatory and the subsequent fibrogenic response
464 of the FB was initiated with almost immediate effect, it can be speculated that the observed
465 effects were indeed promoted by EC with a basal cell signature. This would be in line with very
466 recent literature demonstrating the profibrotic propensity of airway–derived basal cells (41).
467 By applying two independent correlation analyses we found that fractions of the seeded EC
468 population adopted phenotypes resembling the “KRT5–/KRT17+ EC” – and “Proliferating EC”
469 populations, as identified in the reference IPF patient lung data set (5). However, the cells with
470 the gene expression profile resembling the “KRT5–/KRT17+ EC” cells appear to arise only
471 over a 50h cultivation period by differentiation (possibly by undergoing EMT), presumably from
472 the predominant basal cell states EC_FBXO2 and EC_KRT18 and, possibly, by expansion of
473 the EC_COL1A2 and EC_COL6A1 states. Thus, the “KRT5–/KRT17+ EC” are unlikely to drive
474 the fibrogenic response of the FB and, hence, their role remains to be determined.

475 IPF is characterized by a high temporal and spatial heterogeneity (4-7, 9, 42, 43). Lack of
476 synchrony complicates the drawing of conclusions on the ontology of the disease. In addition,
477 the mechanisms driving IPF progression remain incompletely defined. In contrast, the
478 synchronized progression of the cellular responses in our co–culture model allows a temporal
479 resolution of the underlying processes. The phenotypic shift of co–cultured NHLF towards a

myofibroblast phenotype was preceded by processes which showed inflammatory characteristics and involved the early activation of the “rapid–acting” transcriptional regulators STAT3, NF–κB and AP–1 as revealed by our GSOA and differential expression analysis. Our correlation–based analyses revealed highest similarities of the queried FB population at the three hours timepoint to fibrotic mesenchymal populations such as the “HAS1 High FB” and “PLIN2+ FB” of the reference atlas. Furthermore, our module score analysis using AP–1–, NF–κB–, and “TNF–α signaling via NF–κB” gene signatures, achieved the highest expression scores for the “HAS1 High FB”, as well as for the “PLIN2+ FB” of the IPF lung reference atlas. Thus, the signaling network in co–cultured NHLF early after the onset of culture and certain FB populations in patient lungs appear to be very similar and both are dominated by AP–1 and NF–kB activity. Building on the temporal resolution of our model, it could be hypothesized that in patients the “HAS1 High FB” and the “PLIN2+ FB” cell populations may represent transition states towards terminally differentiated ECM–generating myofibroblast.

The expression of *JUN* (coding for c–JUN) was increased in many human fibrotic diseases (44) and strong expression of *FOSL2* (encoding FRA–2) was observed in human samples of pulmonary fibrosis associated with vascular remodeling (45). Furthermore, in lung biopsy sections from IPF patients, c–JUN and c–FOS were detected in a subset of FB in fibrotic areas in lungs of IPF patients (46). The systemic induction of c–Jun in mice lead to the development of multi–organ fibrosis with strong lung involvement (44), whereas ectopic expression of *FOSL2* resulted in lung fibrosis, associated with vascular remodeling of the pulmonary artery (45). Thus, activation of both c–FOS/AP–1 and c–JUN/AP–1 causes pulmonary fibrosis. Applying the c–FOS/AP–1 inhibitor T–5224 during the inflammatory phase (from t = 0h – 18h) of our model showed that c–FOS/AP–1 also played an essential and early role in the pro–fibrotic differentiation of co–cultured NHLF in myofibroblast. In addition, inhibition of IKK/NF–κB by BAY 11–7082, during the first 18h of co–culture completely suppressed α–SMA and COL1 production in co–cultured NHLF. Thus, in our model, the activation of the transcriptional regulators NF–κB and AP–1 occurred very early and was causative of the phenotypic shift of

NHLF towards a myofibroblast phenotype. This is in line with recently published findings which revealed that TGF–ß receptor signaling and STAT3 signaling pathways are downstream of AP–1 transcription factor c–JUN (46).

Genomic analysis of sorted NHBE at different time points of co–culture identified hallmark AP–1 and NF–κB activity. Pharmacological intervention using c–FOS/AP–1 and IKK/NF–κB inhibitors, T–5224 and BAY 11–7082, respectively, modulated the phenotypes of specific epithelial clusters, as judged by their effects on keratin profiles, cell state frequency, and biological function. Particularly pronounced in the *COL1A1* expressing EC states (e.g. “EC_COL1A2”), the effects of IKK/NF–κB inhibition were observed already at t = 3h, suggesting direct and context–dependent transcriptional control of NF–κB on epithelial keratin expression, as well as functions associated with e.g. keratinization, differentiation and cell–cell interaction.

The cells in the described co-culture lung fibrosis assay system showed aspects of phenotype changes and genomic hallmark signatures that were found in cells from IPF patient lungs. However, the known limitations of 2D culture conditions, including the high stiffness of the culture vessel surface that might artificially influence the cell behavior, still apply. Furthermore, although the IPF patient-derived fibroblasts were examined without detecting major differences compared to NHLF, the same might not be true for IPF-derived lung EC. In particular patient–derived basal EC should be considered in future studies, because they were not investigated here due to the lack of availability.

In summary, we described here the development and thorough characterization of an in vitro lung fibrosis model and demonstrate that it i) works autonomously, i.e no added pro–fibrotic stimulants such as TGF–ß are required ii) depends on direct EC–FB contact iii) develops into a fibrotic condition progressively and synchronously and iv) leads to dynamic phenotype changes of EC and FB that are very similar to cells associated with progressive fibrogenic changes in IPF patients. Furthermore, using this model, we identified NF–κB as a critical

component in both the culture–induced phenotypic changes in EC and the differentiation of FB to myofibroblast. Inhibition of NF–κB, may have potential to suppress both the fibrotic foci and the appearance of aberrant EC in IPF patients.

Methods

A detailed description of all methods and material used in this study can be found in the Supplementary Material section.

Co–culture and protein quantification assay

NHLF/NHBE co–culture: NHLF were resuspended in FGM–2 growth medium (Lonza), mixed with an equal volume of NHBE in BEGM growth medium (Lonza), and seeded into a 96–well flat–bottom culture plate (Corning) at a density of 20'000 and 1'600 cells per well, respectively. Compound dilution series were added, and the co–culture was incubated for 18h at 37°C/5 % CO_2. The medium was then replaced with FBM (Lonza) containing 0.1 % faf BSA and the compounds at the indicated concentrations. The medium was supplemented with 100 units/ml penicillin, 100 µg/ml streptomycin for the entire duration of the culture. If not indicated otherwise the co–culture was incubated at 37°C/5 % CO_2 for 96 h.

Lysis of cells for MS/MS analysis: The cell culture medium was removed, and cells were lysed on ice. Samples were prepared for MS/MS and surrogate tryptic peptides were chosen for detection of COL1 (COL1A1), α–SMA (ACTA2) and tubulin (TBA1A1) as described (19). Peak areas for COL1A1 and ACTA2 were normalized by dividing by TBA1A1 peak area.

Gene expression analysis by bulk RNAseq

Library preparation, sequencing, and data processing

Cells were isolated by FACS and lysed in RL Lysis Buffer from NORGEN Single cell RNA purification kit (Norgen). Total RNA was isolated using NORGEN Single cell RNA purification kit including a DNase treatment (Norgen), according to the manufacturer's instructions. RNA–Seq libraries of poly(A)–selected RNA were isolated from 50 ng of total RNA using NuGen

Universal Plus mRNA–Seq kit (NuGen) according to the manufacturer's recommendations. For the sequencing runs the samples were randomized across two flow cells and sequenced using NextSeq 500/550high Output Kit v2.5 with 75 Cycles (Illumina). Sequenced reads were aligned with STAR (v2.5.4b) to the human reference genome (GRCh38) and gene–wise alignments were quantified with featureCounts using Ensembl gene annotations. See Supplemental Methods for further information. All samples passed quality control. Differentially expressed genes were evaluated using edgeR(47), where genes with an FDR < 0.05 and a linear fold change (linFC) > 1.5 were considered significant.

Sc–RNA–seq

Library preparation and sequencing

Cells were resuspended in ice–cold FACS buffer (autoMACS Rinsing Solution, Miltenyi Biotec) supplemented with 0.5% fatty acid–free bovine serum albumin (Calbiochem) to approximately $1*10^6$ cells/ml. Time point 0h cells were combined directly in ice cold FACS buffer. Average viability of all samples taken into single cell sequencing was 92.5% (88.4% – 94.8%). Cell suspensions were immediately loaded on a Chromium Next GEM Chip K (10X Genomics) using the Chromium Next GEM Single Cell 5' Kit v.2 reagents according to the manufacturer's recommendations. Final libraries were pooled and sequenced to an average depth of >42'000 reads per cell on an Illumina NovaSeq system by Genewiz (Germany).

Sc–RNA–seq preprocessing

Processing of raw data, filtering, normalization, and graph–based clustering are described in detail in the Supplemental Methods. For labeling of clusters, we used a Wilcoxon rank sum test implemented in Seurat's FindConservedMarkers function to determine cluster–specific marker genes independent of timepoint or treatment.

Sc–RNA–seq analysis

Augur (48) was used to predict culture time effect for each cell cluster. Pair–wise differential expression across timepoints or treatments was performed using MAST (35) implemented in Seurat after merging all EC into one group. Expression testing was performed on the log–normalized counts. Results were filtered based on a Bonferroni–adjusted p–value of less than 0.1.

Comparison with a human lung cell data set

Curated data of control and fibrotic lung cells as published by Habermann et al (5) was downloaded (GSE135893). A subset of the data was generated to include relevant cell types (mesenchymal and EC from IPF patient or control donor (smoker reject lungs). This "reference" set of cells was re–processed via the default SCTransfrom workflow in parallel to the subset of vehicle–only cells from the co–culture assay ("query"). Mappings of relevant Habermann (5) cell identity labels to proposed consensus re–annotation of integrated single–cell transcriptomic human lung cell atlas (49) are listed in Supplemental Table 11. For each, reference and query, we determined the 8'000 most variable genes and determined the union of overlapping variable genes that were mutually expressed in the reference and query data. Average per–cell log expression was calculated for each cluster and used to determine the reported Pearson correlations.

For the implementation of SingleR (34) we computed the top–100 positive marker genes for each reference cluster using a Wilcoxon rank sum test. The query cells were individually profiled against this reference model using SingleR default settings. The per–cell gene module co–expression scores were calculated using Seurat's "AddModuleScore" function using an approach implemented in (36).

Statistics

Calculations were performed using Microsoft Excel. Statistical analysis was performed using GraphPad Prism 6 or R 3.6.0 (R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.r-project.org/). Data were analyzed using ANOVA followed by post hoc tests, according to the experiment (see figure legends for details). A P value < 0.05 was considered significant.

Study approval

Formalin–fixed, paraffin–embedded sections of human lungs with IPF or controls (obtained from areas distal to lung cancer resection) were obtained from extra material from subjects undergoing surgical lung biopsy for their respective condition at the University of Michigan. Samples were de–identified prior to receipt and thus did not require patient consent to obtain.

Data and materials availability

All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. RNA–seq data were deposited in the ArrayExpress database at EMBL–EBI (www.ebi.ac.uk/arrayexpress) under accession numbers E–MTAB–11832 and E–MTAB–11847. The R command line functions and arguments that were specified for the analysis of the RNA–seq and sc–RNA–seq data are listed in Supplemental Table 12.

Author contributions

AS, RL, SLC, FLG, MS, PB, BR, JR, JHB, PR, BS, DF, RS, MBR, MB, ON, UL, and PS contributed to experimental work, data analysis and/or interpretation. PS, FLG, MS, PB, JR, JHB, RS, MBR, and UL designed and led implementation of experiments, analyzed data and supervised experimental activity. PS conceived and conceptualized the study. ESW provided human patient lung biopsy specimens for ISH analysis. The manuscript was drafted by PS

and ON and was reviewed and edited by all other authors. All authors reviewed and approved the final manuscript as submitted.

Competing financial interests

All authors, except ESW are current or former employees of Idorsia Pharmaceuticals Ltd. The authors have declared that no conflict of interest exists.

Acknowledgements

We thank Richard Welford for his contribution to the establishment of MS/MS quantification of proteins from cell lysate, Marco Caldarone for technical assistance with the MS/MS, Pijus Brasauskas for support with the RNA–seq, and John Gatfield for careful review and valuable suggestions on the manuscript. RNAscope in situ hybridizations were performed by Marie Lauigan, Madhura Nayak and Ruby Hsu from Advanced Cell Diagnostics.

References

1. Wolters PJ, Blackwell TS, Eickelberg O, Loyd JE, Kaminski N, Jenkins G, et al. Time for a change: is idiopathic pulmonary fibrosis still idiopathic and only fibrotic? *Lancet Respir Med.* 2018;6(2):154-60.
2. Blackwell TS, Tager AM, Borok Z, Moore BB, Schwartz DA, Anstrom KJ, et al. Future directions in idiopathic pulmonary fibrosis research. An NHLBI workshop report. *Am J Respir Crit Care Med.* 2014;189(2):214-22.
3. Selman M, King TE, and Pardo A. Idiopathic pulmonary fibrosis: prevailing and evolving hypotheses about its pathogenesis and implications for therapy. *Ann Intern Med.* 2001;134(2):136-51.
4. Adams TS, Schupp JC, Poli S, Ayaub EA, Neumark N, Ahangari F, et al. Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. *Sci Adv.* 2020;6(28):eaba1983.
5. Habermann AC, Gutierrez AJ, Bui LT, Yahn SL, Winters NI, Calvi CL, et al. Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. *Sci Adv.* 2020;6(28):eaba1972.
6. Xu Y, Mizuno T, Sridharan A, Du Y, Guo M, Tang J, et al. Single-cell RNA sequencing identifies diverse roles of epithelial cells in idiopathic pulmonary fibrosis. *JCI Insight.* 2016;1(20):e90558.
7. Strunz M, Simon LM, Ansari M, Kathiriya JJ, Angelidis I, Mayr CH, et al. Alveolar regeneration through a Krt8+ transitional stem cell state that persists in human lung fibrosis. *Nat Commun.* 2020;11(1):3559.
8. Huang KY, and Petretto E. Cross-species integration of single-cell RNA-seq resolved alveolar-epithelial transitional states in idiopathic pulmonary fibrosis. *Am J Physiol Lung Cell Mol Physiol.* 2021.
9. Barkauskas CE, Chung MI, Fioret B, Gao X, Katsura H, and Hogan BL. Lung organoids: current uses and future promise. *Development.* 2017;144(6):986-97.
10. Barkauskas CE, Cronce MJ, Rackley CR, Bowie EJ, Keene DR, Stripp BR, et al. Type 2 alveolar cells are stem cells in adult lung. *J Clin Invest.* 2013;123(7):3025-36.
11. Bon H, Hales P, Lumb S, Holdsworth G, Johnson T, Qureshi O, et al. Spontaneous extracellular matrix accumulation in a human in vitro model of renal fibrosis is mediated by alphaV integrins. *Nephron.* 2019;142(4):328-50.
12. Lama V, Moore BB, Christensen P, Toews GB, and Peters-Golden M. Prostaglandin E2 synthesis and suppression of fibroblast proliferation by alveolar epithelial cells is cyclooxygenase-2-dependent. *Am J Respir Cell Mol Biol.* 2002;27(6):752-8.
13. Moll S, Ebeling M, Weibel F, Farina A, Araujo Del Rosario A, Hoflack JC, et al. Epithelial cells as active player in fibrosis: findings from an in vitro model. *PLoS One.* 2013;8(2):e56575.
14. Nugraha B, Mohr MA, Ponti A, Emmert MY, Weibel F, Hoerstrup SP, et al. Monitoring and manipulating cellular crosstalk during kidney fibrosis inside a 3D in vitro co-culture. *Sci Rep.* 2017;7(1):14490.
15. Sakai N, and Tager AM. Fibrosis of two: Epithelial cell-fibroblast interactions in pulmonary fibrosis. *Biochim Biophys Acta.* 2013;1832(7):911-21.
16. Tan Q, Ma XY, Liu W, Meridew JA, Jones DL, Haak AJ, et al. Nascent lung organoids reveal epithelium- and bone morphogenetic protein-mediated suppression of fibroblast activation. *Am J Respir Cell Mol Biol.* 2019;61(5):607-19.

17. Schruf E, Schroeder V, Le HQ, Schonberger T, Raedel D, Stewart EL, et al. Recapitulating idiopathic pulmonary fibrosis related alveolar epithelial dysfunction in a human iPSC-derived air-liquid interface model. *FASEB J.* 2020;34(6):7825-46.
18. Bauer Y, Tedrow J, de Bernard S, Birker-Robaczewska M, Gibson KF, Guardela BJ, et al. A novel genomic signature with translational significance for human idiopathic pulmonary fibrosis. *Am J Respir Cell Mol Biol.* 2015;52(2):217-31.
19. Sieber P, Schafer A, Lieberherr R, Le Goff F, Stritt M, Welford RWD, et al. Novel high-throughput myofibroblast assays identify agonists with therapeutic potential in pulmonary fibrosis that act via EP2 and EP4 receptors. *PLoS One.* 2018;13(11):e0207872.
20. Rajkumar R, Konishi K, Richards TJ, Ishizawar DC, Wiechert AC, Kaminski N, et al. Genomewide RNA expression profiling in lung identifies distinct signatures in idiopathic pulmonary arterial hypertension and secondary pulmonary hypertension. *Am J Physiol Heart Circ Physiol.* 2010;298(4):H1235-48.
21. Fornes O, Castro-Mondragon JA, Khan A, van der Lee R, Zhang X, Richmond PA, et al. JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* 2020;48(D1):D87-D92.
22. Han H, Cho JW, Lee S, Yun A, Kim H, Bae D, et al. TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Res.* 2018;46(D1):D380-D6.
23. Federico A, and Monti S. hypeR: an R package for geneset enrichment workflows. *Bioinformatics.* 2020;36(4):1307-8.
24. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics.* 2013;14:128.
25. Yoon JH, Jung SM, Park SH, Kato M, Yamashita T, Lee IK, et al. Activin receptor-like kinase5 inhibition suppresses mouse melanoma by ubiquitin degradation of Smad4, thereby derepressing eomesodermin in cytotoxic T lymphocytes. *EMBO Mol Med.* 2013;5(11):1720-39.
26. Schust J, Sperl B, Hollis A, Mayer TU, and Berg T. Stattic: a small-molecule inhibitor of STAT3 activation and dimerization. *Chem Biol.* 2006;13(11):1235-42.
27. Lin L, Hutzen B, Li PK, Ball S, Zuo M, DeAngelis S, et al. A novel small molecule, LLL12, inhibits STAT3 phosphorylation and activities and exhibits potent growth-suppressive activity in human cancer cells. *Neoplasia.* 2010;12(1):39-50.
28. Pierce JW, Schoenleber R, Jesmok G, Best J, Moore SA, Collins T, et al. Novel inhibitors of cytokine-induced IkappaBalpha phosphorylation and endothelial cell adhesion molecule expression show anti-inflammatory effects in vivo. *J Biol Chem.* 1997;272(34):21096-103.
29. Aikawa Y, Morimoto K, Yamamoto T, Chaki H, Hashiramoto A, Narita H, et al. Treatment of arthritis with a selective inhibitor of c-Fos/activator protein-1. *Nat Biotechnol.* 2008;26(7):817-23.
30. Yan C, Deng C, Liu X, Chen Y, Ye J, Cai R, et al. TNF-alpha induction of IL-6 in alveolar type II epithelial cells: Contributions of JNK/c-Jun/AP-1 element, C/EBPdelta/C/EBP binding site and IKK/NF-kappaB p65/kappaB site. *Mol Immunol.* 2018;101:585-96.
31. Saelens W, Cannoodt R, Todorov H, and Saeys Y. A comparison of single-cell trajectory inference methods. *Nat Biotechnol.* 2019;37(5):547-54.
32. Street K, Risso D, Fletcher RB, Das D, Ngai J, Yosef N, et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics.* 2018;19(1):477.
33. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods.* 2017;14(10):979-82.
34. Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol.* 2019;20(2):163-72.

35. Finak G, McDavid A, Yajima M, Deng J, Gersuk V, Shalek AK, et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* 2015;16:278.
36. Tirosh I, Venteicher AS, Hebert C, Escalante LE, Patel AP, Yizhak K, et al. Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. *Nature.* 2016;539(7628):309-13.
37. Efremova M, Vento-Tormo M, Teichmann SA, and Vento-Tormo R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc.* 2020;15(4):1484-506.
38. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003;13(11):2498-504.
39. Katzen J, and Beers MF. Contributions of alveolar epithelial cell quality control to pulmonary fibrosis. *J Clin Invest.* 2020;130(10):5088-99.
40. Gu C, and Giraudo E. The role of semaphorins and their receptors in vascular development and cancer. *Exp Cell Res.* 2013;319(9):1306-16.
41. Jaeger B, Schupp JC, Plappert L, Terwolbeck O, Artysh N, Kayser G, et al. Airway basal cells show a dedifferentiated KRT17(high) phenotype and promote fibrosis in idiopathic pulmonary fibrosis. *Nat Commun.* 2022;13(1):5637.
42. Gokey JJ, Snowball J, Sridharan A, Speth JP, Black KE, Hariri LP, et al. MEG3 is increased in idiopathic pulmonary fibrosis and regulates epithelial cell differentiation. *JCI Insight.* 2018;3(17).
43. Seibold MA, Smith RW, Urbanek C, Groshong SD, Cosgrove GP, Brown KK, et al. The idiopathic pulmonary fibrosis honeycomb cyst contains a mucocilary pseudostratified epithelium. *PLoS One.* 2013;8(3):e58658.
44. Wernig G, Chen SY, Cui L, Van Neste C, Tsai JM, Kambham N, et al. Unifying mechanism for different fibrotic diseases. *Proc Natl Acad Sci U S A.* 2017;114(18):4757-62.
45. Eferl R, Hasselblatt P, Rath M, Popper H, Zenz R, Komnenovic V, et al. Development of pulmonary fibrosis through a pathway involving the transcription factor Fra-2/AP-1. *Proc Natl Acad Sci U S A.* 2008;105(30):10525-30.
46. Cui L, Chen SY, Lerbs T, Lee JW, Domizi P, Gordon S, et al. Activation of JUN in fibroblasts promotes pro-fibrotic programme and modulates protective immunity. *Nat Commun.* 2020;11(1):2795.
47. Robinson MD, McCarthy DJ, and Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 2010;26(1):139-40.
48. Skinnider MA, Squair JW, Kathe C, Anderson MA, Gautier M, Matson KJE, et al. Cell type prioritization in single-cell data. *Nat Biotechnol.* 2021;39(1):30-4.
49. Sikkema L, Strobl D, Zappia L, Madissoon E, Markov N, Zaragosi L, et al. An integrated cell atlas of the human lung in health and disease. *bioRxiv.* 2022:2022.03.10.483747.

Figure legends

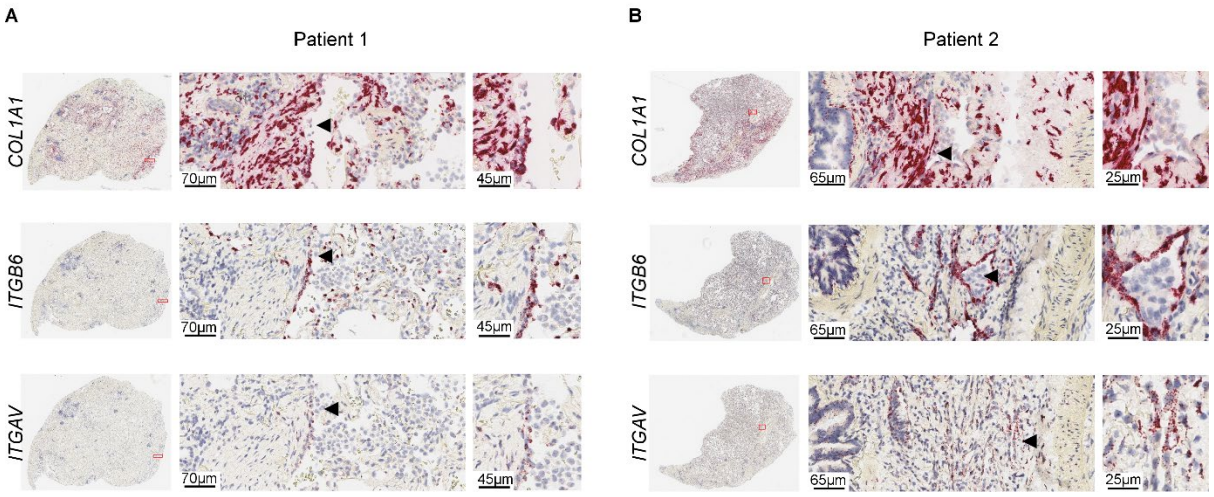


Figure 1. Epithelial and mesenchymal cells are localized in close proximity in lung sections of IPF patients. Expression of the mesenchymal marker gene *COL1A1*, as well as the epithelial marker genes *ITGB6* and *ITGAV*, as detected by RNAscope ISH on formalin–fixed, paraffin–embedded lung sections of IPF patient 1 **(A)** and 2 **(B)**. Boxed area in the low magnification lung section overviews in the left image indicates the enlarged region. The arrowheads point to the approximate same location in close consecutive tissue sections. Representative images of two out of three IPF patients are shown.

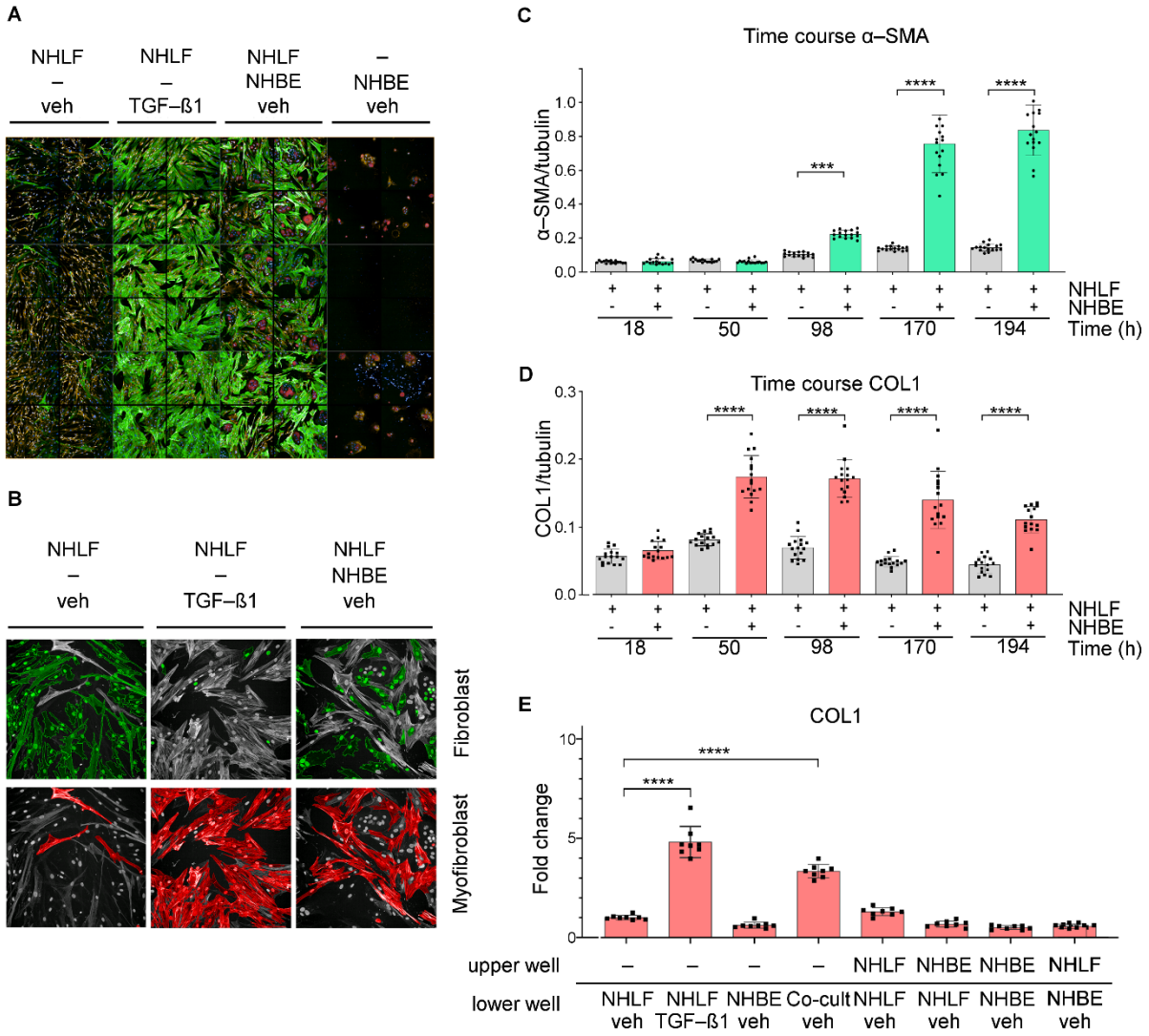


Figure 2. EC induce NHLF–to–myofibroblast transformation by cell–cell contact. (A) High–content confocal microscopy images of NHLF cells, NHLF cells stimulated with TGF–ß1 (5 ng/ml), NHLF cells co–cultured with NHBE, and NHBE cells cultured alone are shown. NHBE cells were pre–stained with CellTracker Deep Red. Nuclei were stained with DAPI, and α–SMA staining was performed using an anti–α–SMA antibody. **(B)** 5 days after seeding cells were classified into epithelial, fibroblast and myofibroblast, respectively, by a trained classifier (see Methods for details). Classified fibroblasts and myofibroblasts are colored in green and red, respectively. **(C)** Time–course of α–SMA and **(D)** COL1 accumulation. 18h, 50h, 98h, 170h and 194 h after seeding α–SMA and COL1 were quantified by MS/MS and plotted, normalized to tubulin, against time (h). A one–way ANOVA with Tukey's multiple comparison test was used. *N* = 16 for each condition. **(E)** COL1, normalized to tubulin, was detected by MS/MS after 98h in trans–well cell cultures of NHBE and NHLF seeded either together in the lower chamber or individually in the upper or lower chamber in presence of 5 ng/ml TGF–ß1 or vehicle as indicated. *N* = 8 for each condition. A one–way ANOVA with Dunnett's multiple comparison test was used. Bars indicate mean ± SD. ***p ≤ 0.001, ****p ≤ 0.0001.

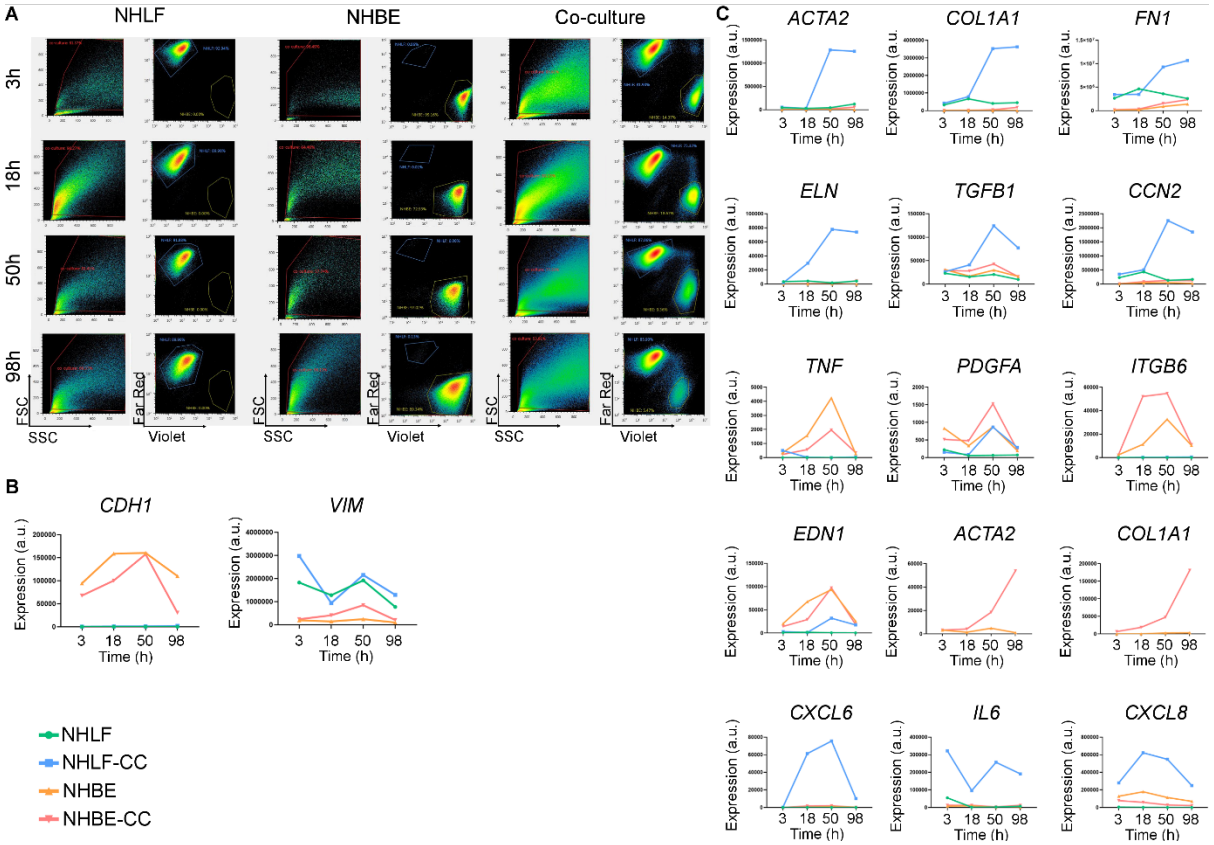


Figure 3. Pro–inflammatory and pro–fibrotic genes are induced in co–cultured NHLF and NHBE cells. (A) For each condition, NHLF cells were pre–stained with CellTrace Far Red and the NHBE cells with CellTrace Violet, seeded, either separately or in combination as depicted, at t = 0h and then FACS sorted, followed by lysis at t = 3h, 18h, 50h, and 98h. Left hand charts show scatter plots based on forward (FSC) and side scattering (SSC) profiles. Red line indicates gating threshold. Right hand panels: Gated cells were sorted on CellTrace Far Red and CellTrace Violet staining intensities, respectively. **(B)** Selected genes are shown to address identity, inflammatory and profibrotic gene expression changes of the involved cell types over time. Expression level of *CDH1 and VIM* **(C)** *ACTA2, COL1A1, FN1, ELN, TGFB1, CCN2, TNF, PDGFA, ITGB6, EDN1, CXCL6, IL6* and *CXCL8* in either monocultured NHBE and NHLF or co–cultured NHBE (NHBE–CC) and NHLF (NHLF–CC). Relative expression levels, as determined by qRT–PCR and depicted as arbitrary units (a.u.) are plotted vs time (h). The experiment was performed once with *n* = 2–3 replicates per sample.

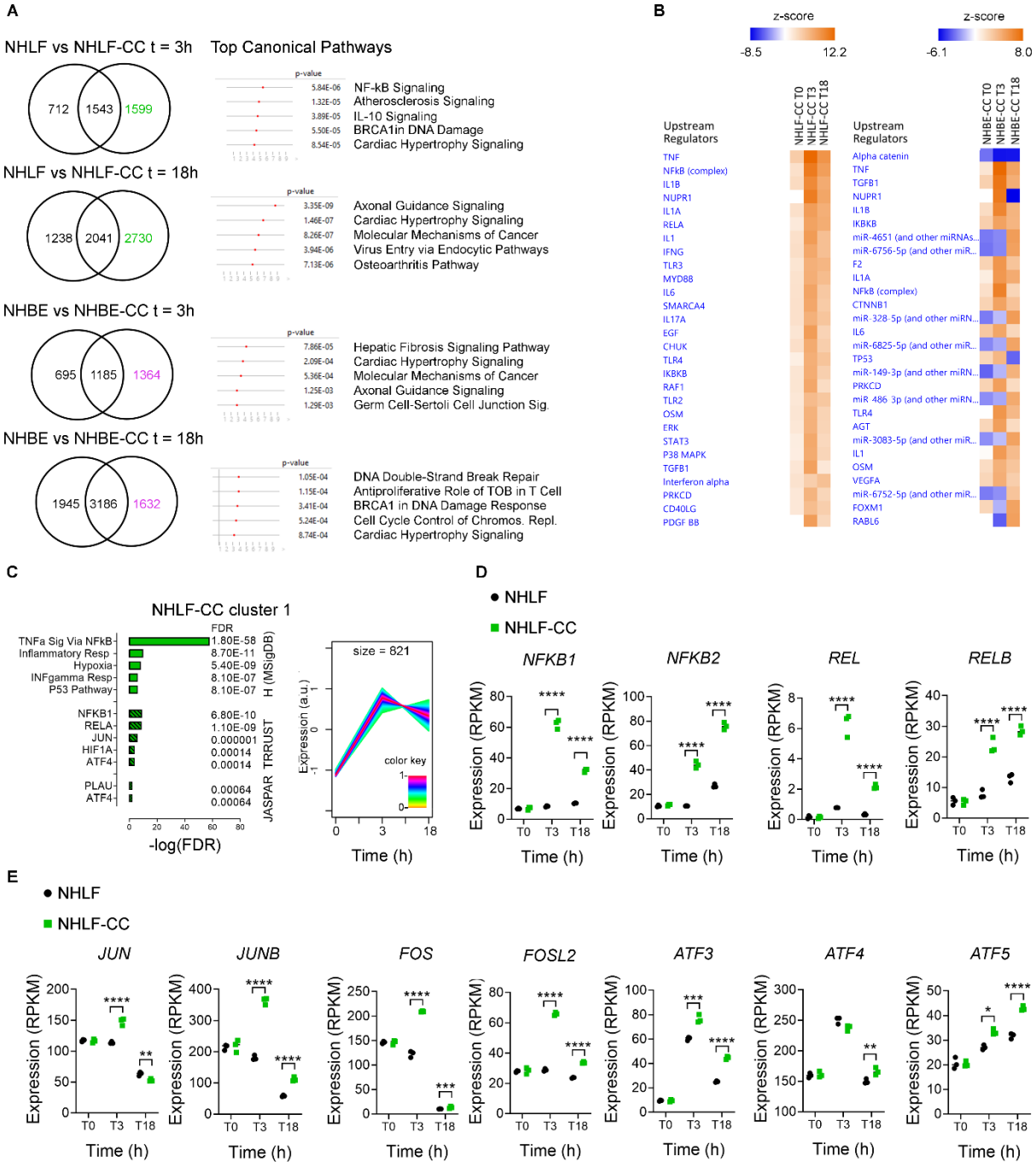


Figure 4. Early gene expression changes reveal co–culture–induced NF–κB and AP–1 gene expression signature in NHLF and NHBE. (A) Venn diagrams displaying the number of overlapping differentially expressed genes per condition compared to t = 0h (FDR < 0.05 and |linFC| > 1.5 and the top–ranked canonical pathways predicted for the co–culture specific genes, as identified by Ingenuity pathway analysis application. **(B)** Ingenuity upstream regulator analysis of the DEGs between co–cultured and monocultured NHLF and NHBE (FDR < 0.05 and |linFC| > 1.5) at t = 0h, 3h, and 18h, respectively, sorted according to the activation z–score. **(C)** Expression kinetics and results of the GSOA of NHLF–CC cluster 1. The comprised expression kinetics of the clustered genes is depicted as a z–score (from –1 to 1) across the time points t = 0h (T0), 3h (T3), and 18h (T18), respectively. The expression pattern of each gene is associated with a cluster weight between 0 and 1 (according to its match with cluster dynamics), color coded in the figure according to the inserted palette.

838 Enriched gene sets are displayed as bars representing the –log10(FDR) with the
839 corresponding FDR values. Gene expression, depicted in reads per kilobase of transcript, per
840 million mapped reads (RPKM), of selected **(D)** NF–κB and **(E)** AP–1 transcription factor family
841 members, respectively, showing gene expression increases between conditions NHLF and
842 NHLF–CC at the time points t =0h, 3h, or 18h. Differentially expressed genes in comparison
843 to monocultured control were evaluated using edgeR and are depicted as * FDR ≤ 0.05, **
844 FDR ≤ 0.01, *** FDR ≤ 0.001, and **** FDR ≤ 0.0001. The experiment was performed once
845 with *n* = 3 replicates per sample.

846

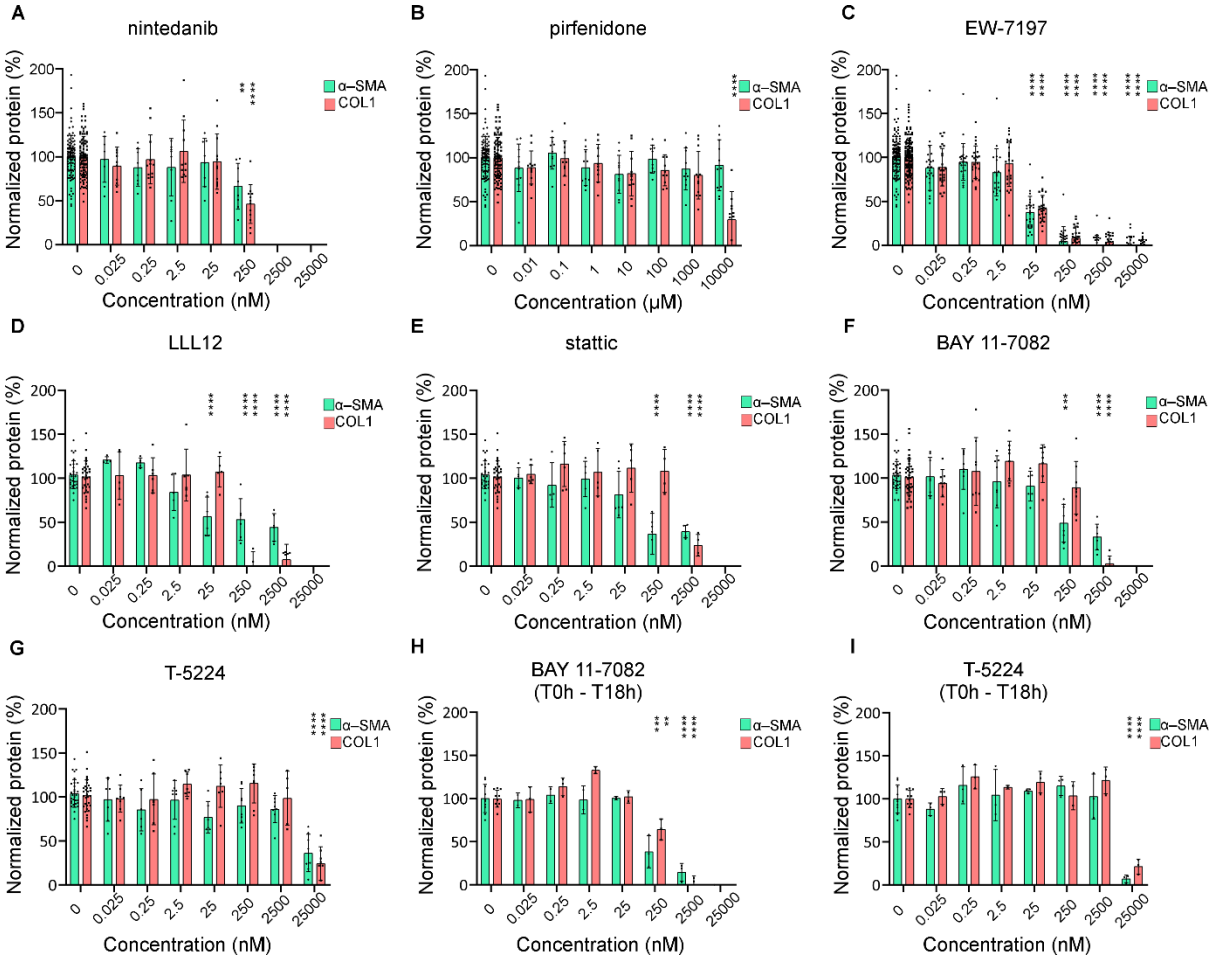


Figure 5. Pharmacological intervention attenuates α–SMA and COL1 accumulation in co–cultures of NHBE and NHLF cells. NHLF and NHBE cells were seeded in full–medium in presence of **(A)** nintedanib, **(B)** pirfenidone, **(C)** EW–7197, **(D)** LLL12, **(E)** stattic, **(F, H)** BAY 11–7082, and **(G, I)** T–5224, at the indicated concentration range (0.025 – 25000 nM), except for pirfenidone (0.01 – 10'000 µM). At t = 18h after seeding, cells were switched to starvation medium containing compounds at the indicated concentration **(A–G)** or solvent **(H, I)** for the remaining duration of the experiment. Lysis was performed 98h after seeding and α–SMA and COL1 were quantified by MS/MS. The mean value of the normalized analyte level of the solvent control is given as 100%. Bars represent protein data normalized to tubulin expressed as (%) in relation to the solvent control and show mean ± SD; *n* = 8 **(A, D–G)**, *n* = 11 **(B),** *n* = 26 **(C)** and *n* = 3 **(H, I).** Data for which cytotoxic effects are not excluded are omitted. A two–way ANOVA with Tukey's multiple comparison test was used. P–values of comparison to solvent control are depicted. **p ≤ 0.01, ***p ≤ 0.001, ****p ≤ 0.0001. 0.25% DMSO was present in all wells.

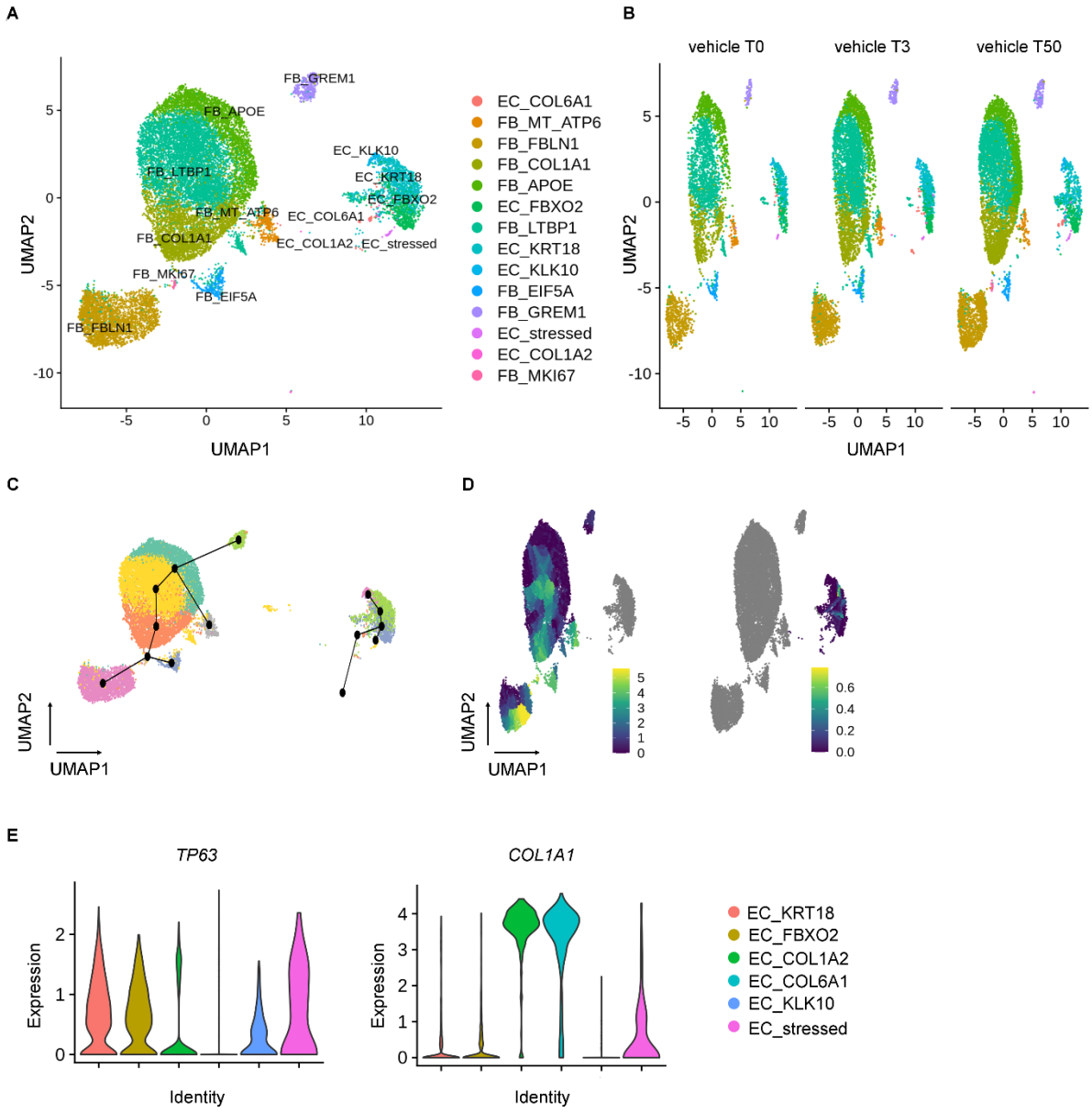


Figure 6. Single–cell landscape of NHBE and NHLF after 0h, 3h and 50h of co–cultivation. (A) Uniform Manifold Approximation and Projection (UMAP) embedding of jointly analyzed single cell transcriptomes from 18'676 cells from vehicle control. Cells were collected at three different time points (t = 0h, 3h, and 50h) in a single experiment. The fourteen identified cell states were named after one of the most significant and cell state–specific marker genes as determined by the FindConservedMarkers function implemented in Seurat. **(B)** UMAP embedding, **(C)** Slingshot–based –, and **(D)** Monocle3–based pseudo–time trajectories calculated from UMAP embeddings of jointly analyzed 2'011 vehicle–treated EC, and 16'665 FB cells, including all time points (t = 0h, 3h, and 50h). **(E)** Violin plots displaying the expression level of *TP63*, and *COL1A1*, separated by sub–state. Normalized gene expression is depicted as log(counts+1).

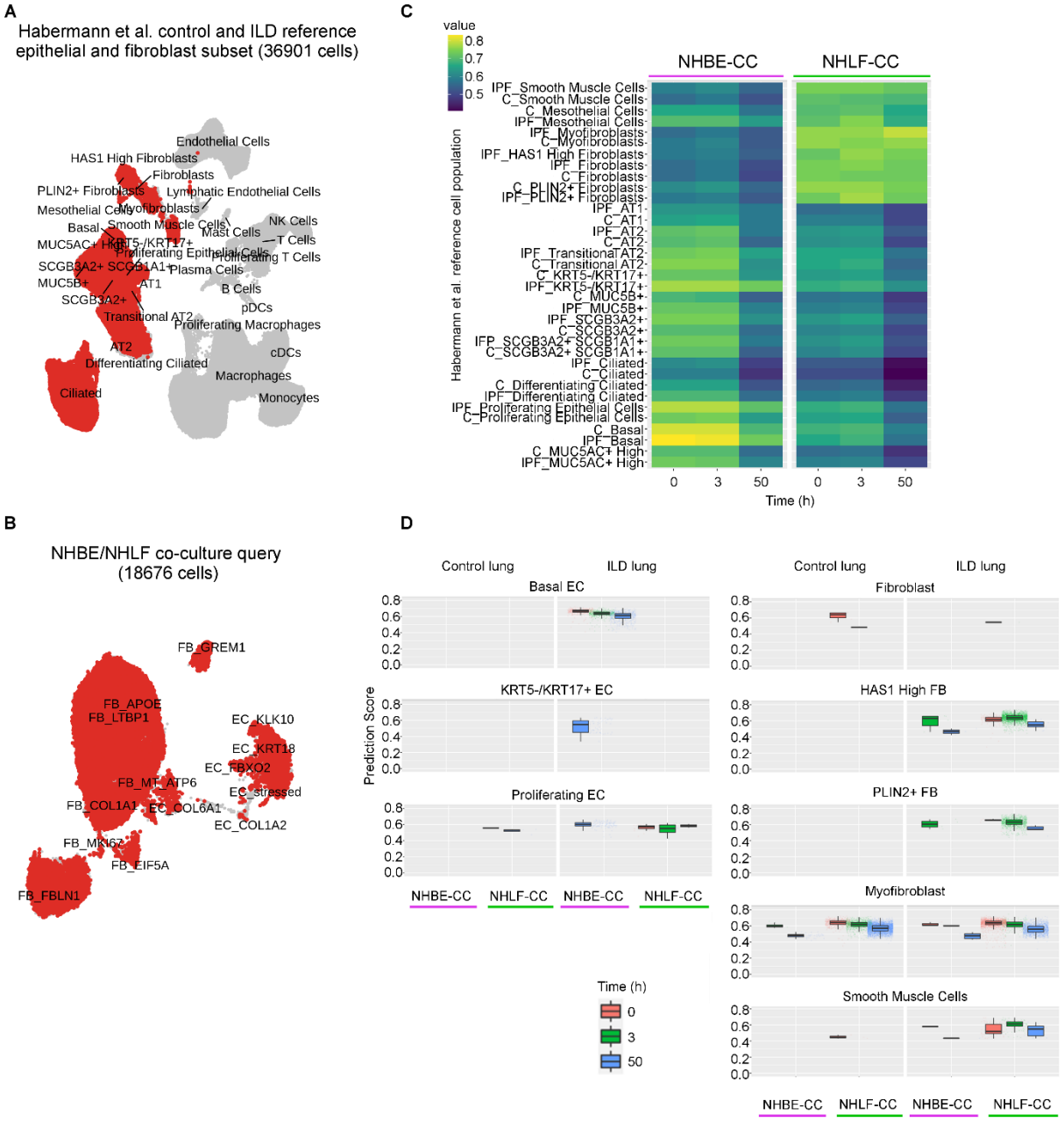


Figure 7. The genomic expression signature of NHBE and NHLF in the co–culture shows similarity to cell populations isolated from IPF patient lungs. (A) UMAP embedding of the Habermann data set. Lung mesenchymal and epithelial cell types highlighted in red were carried forward to generate a reference data set comprising 36'901 epithelial and mesenchymal cells representing control and interstitial lung disease (ILD) subsets. **(B)** All cells corresponding to the vehicle group time points t = 0h, 3h and 50h, highlighted in red, were included from the co–culture system and served as the query data to probe the reference data set. **(C)** Pearson correlations of vehicle–treated NHBE–CC and NHLF–CC query populations with human IPF (IPF) and control (C) reference populations. **(D)** Iterative Spearman correlations at single cell level between marker genes shared by both,

query — and Habermann reference population, and depicted as overall SingleR correlation prediction scores of vehicle group cell types in human lung.

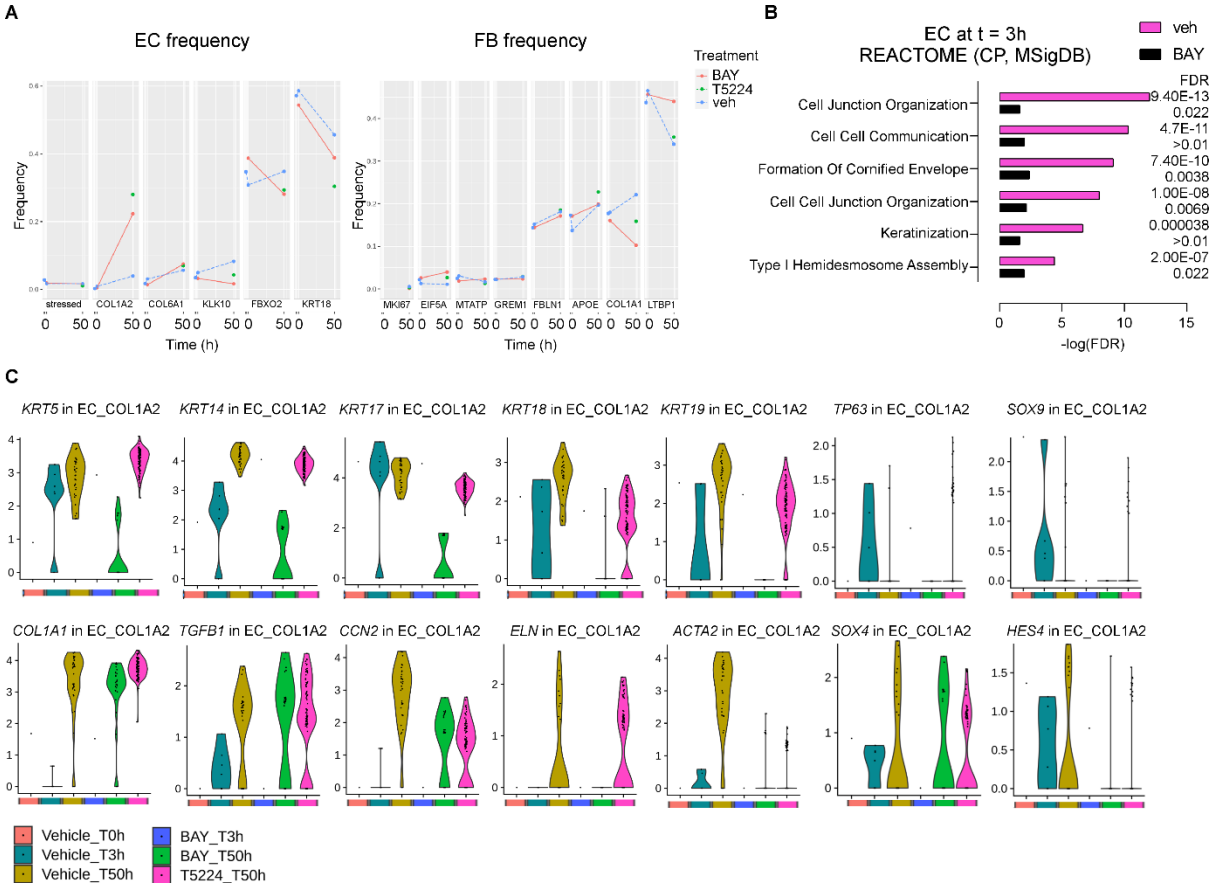


Figure 8. Inhibition of c–FOS/AP–1 and IKK/NF–κB affects co–cultured epithelial cell population frequencies and marker gene expression at sub–population resolution. (A) Effect of treatment (i.e. 2.5 μM BAY 11–7082 (BAY), 25 μM T–5224 (T5224), and vehicle (veh)) on epithelial lineage cell frequencies (y–axis) separated by EC sub–state and plotted for each time point (t = 0h, 3h, and 50h; x–axis). **(B)** Effect of the NF–κB inhibitor BAY 11–7082 (2.5 μM) applied from t = 0h, on enriched canonical pathways (CP, Reactome) in the EC DEGs at time t = 3h displayed as bars representing the –log10(FDR). **(C)** Violin plots displaying the expression level of *KRT5, KRT14, KRT17, KRT18, KRT19*, *TP63*, *SOX9*, *COL1A1*, *TGFB1*, *CCN2*, *ELN*, *ACTA2*, *SOX4*, and *HES4*, separated by sample and at the level of the sub–clustered cell state identifier "EC_COL1A2". Cells were either untreated (vehicle) or treated (i.e. 2.5 μM BAY 11–7082, 25 μM T–5224) and collected in a single experiment at the time points (t = 0h, 3h, and 50h). Normalized gene expression is depicted as log(counts+1) for a non–statistical overview of gene–of–interest expression in the dataset.

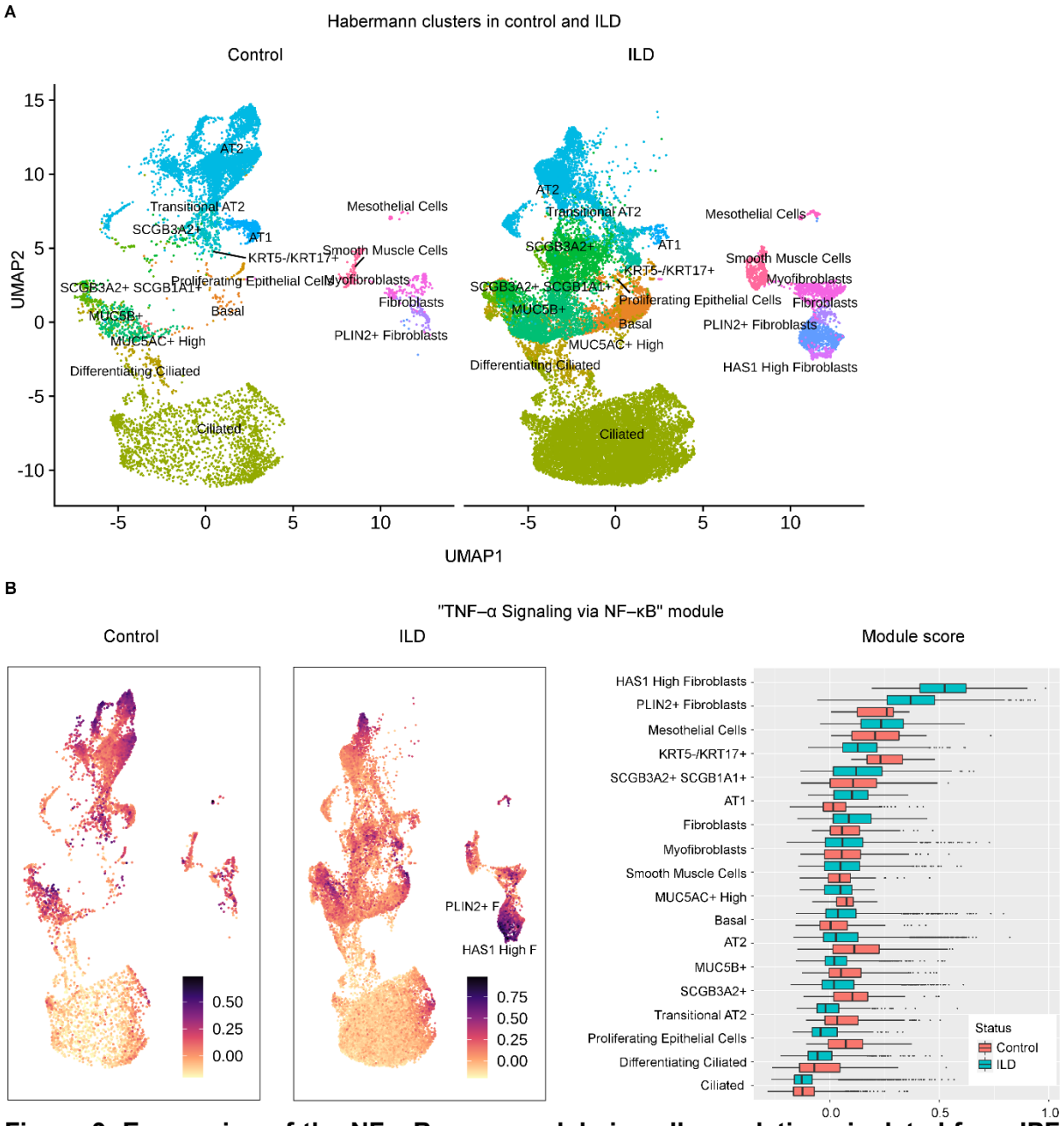


Figure 9. **Expression of the NF–κB gene module in cell populations isolated from IPF patient lungs. (A)** UMAP embedding of the Habermann reference subset comprising 36'901 epithelial and mesenchymal cells divided into control and interstitial lung disease (ILD) subsets. **(B)** Results of the module score analysis querying the Habermann reference human IPF patient lung cell atlas with the gene expression module "TNF–α Signaling via NF–κB". The UMAP space of control and ILD cells is overlaid with the obtained module scores for each individual cell. Box plots display module scores (x–axis) obtained for the respective reference cell populations (y–axis) in control (red boxes) and ILD patient–derived (turquoise boxes) cells.

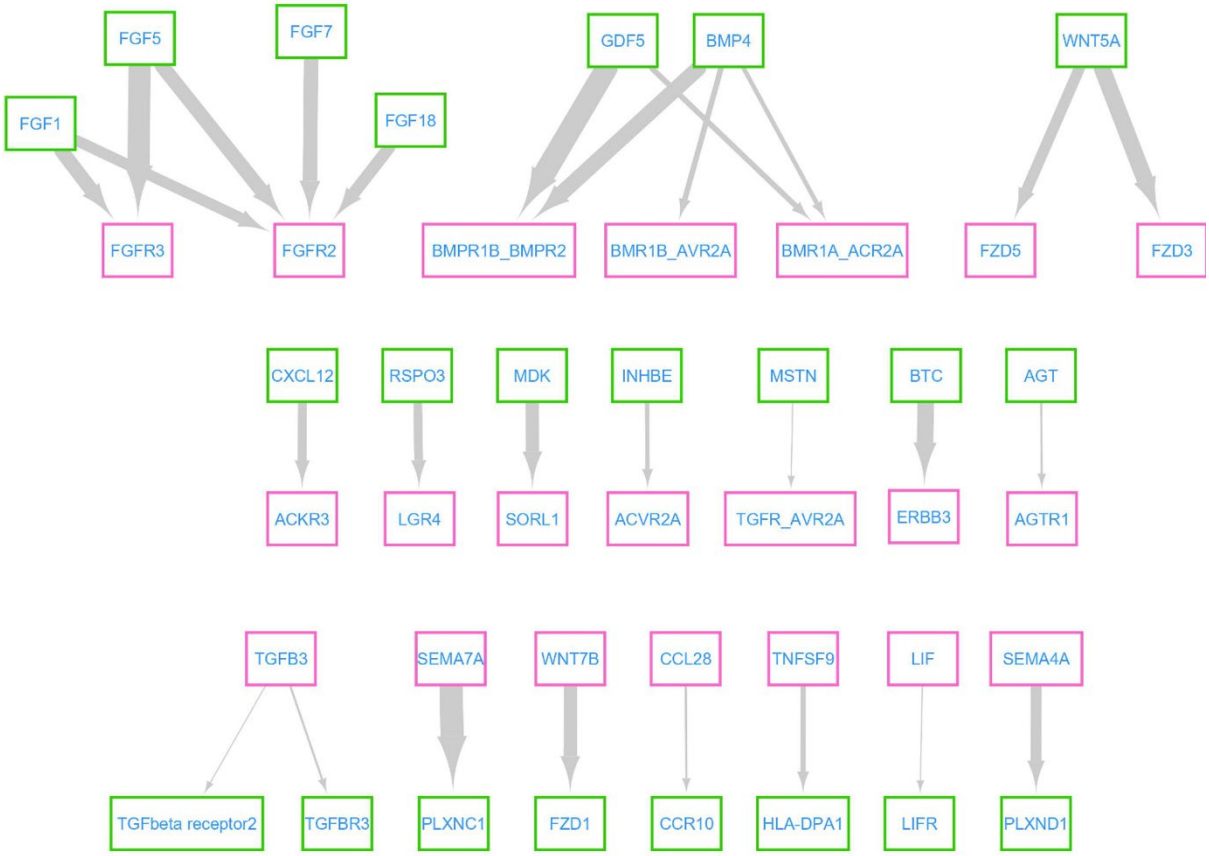


Figure 10. Predicted ligand–receptor pairs between FB and EC. Ligand–receptor pairs between FB (green box) and EC (pink box) were extracted from the sc–RNAseq data of the vehicle–treated cells at t = 0h using CellPhoneDB and visualized with Cytoscape. The arrowhead points to the receptor–expressing cell type. Arrow thickness correlates with the magnitude of the product of mean ligand expression and mean receptor expression for each cell type interaction.

928

Tables

Table 1. Fraction of the co–culture query population showing the highest SingleR correlation value to the respective reference population

	EC (vehicle)			FB (vehicle)		
	T0	**T3**	**T50**	**T0**	**T3**	**T50**
Basal	99	97	82	0	0	0
Proliferating EC	0	0	8	1	0	0
KRT5–/KRT17+	0	0	4	0	0	0
PLIN2+ FB	0	1	0	0	21	0
HAS1 High FB	0	1	0	2	50	0
Myofibroblast	1	1	6	94	21	98
Smooth Muscle Cells	0	0	0	3	7	1
Total [%]	100	100	100	100	100	100

Sc–RNA–seq on vehicle–treated query cells was performed at time points t = 0 (T0), 3 (T3), and 50h (T50) and matched with SingIR against a reference population of control and pulmonary fibrosis patients. The frequency of query cells to the human reference population with the highest similarity is expressed as a percentage of total epithelial (EC) or FB (FB) cells