Paracrine signalling during ZEB1-mediated epithelial mesenchymal transition augments local myofibroblast differentiation in lung fibrosis

Liudi Yao[1,10], Franco Conforti[3,4,10], Charlotte Hill[1,10], Joseph Bell[3], Leena Drawater[1], Juanjuan Li[1], Dian Liu[2], Hua Xiong[2], Aiman Alzetani[4,5], Serena J. Chee[5,6], Ben G. Marshall[4,5], Sophie V. Fletcher[4,5], David Hancock[7], Mark Coldwell[1], Xianglin Yuan[2], Christian H. Ottensmeier[6], Julian Downward[7], Jane E. Collins[3], Rob M. Ewing[1], Luca Richeldi[3,4,8], Paul Skipp[1,9], Mark G. Jones[3,4], Donna E. Davies[3,4,11] and Yihua Wang[1,2,11]

[1]Biological Sciences, Faculty of Natural and Environmental Sciences, University of Southampton SO17 1BJ, UK. [2]Department of Oncology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China. [3]Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton SO16 6YD, UK. [4]NIHR Respiratory Biomedical Research Centre, University Hospital Southampton SO16 6YD, UK. [5]University Hospital Southampton, Southampton SO16 6YD, UK. [6]Cancer Sciences & NIHR and CRUK Experimental Cancer Sciences Unit, University of Southampton, Southampton SO16 6YD, UK. [7]Oncogene Biology, The Francis Crick Institute, London NW1 1AT, UK. [8]Unità Operativa Complessa di Pneumologia, Università Cattolica del Sacro Cuore, Fondazione Policlinico A. Gemelli, Rome, Italy. [9]Centre for Proteomic Research, Institute for Life Sciences, University of Southampton, Southampton, SO17 1BJ, UK. [10]These authors contributed equally to this work. [11]Correspondence should be addressed to D.E.D. (email: D.E.Davies@soton.ac.uk) or Y.W. (e-mail: yihua.wang@soton.ac.uk).

Abstract

The contribution of epithelial-mesenchymal transition (EMT) to human lung fibrogenesis is controversial. Here we provide evidence that ZEB1-mediated EMT in human alveolar epithelial type II (ATII) cells contributes to development of lung fibrosis by paracrine signalling to underlying fibroblasts. Activation of EGFR-RAS-ERK signalling in ATII cells induced EMT via ZEB1. ATII cells had extremely low extracellular matrix gene expression even after induction of EMT, however conditioned media from ATII cells undergoing RAS-induced EMT augmented TGFβ-induced profibrogenic responses in lung fibroblasts. This epithelial-mesenchymal crosstalk was controlled by ZEB1 via the expression of tissue plasminogen activator (tPA). In human fibrotic lung tissue, nuclear ZEB1 expression was detected in alveolar epithelium adjacent to sites of extracellular matrix (ECM) deposition, suggesting that ZEB1-mediated paracrine signalling has the potential to contribute to early fibrotic changes in the lung interstitium. Targeting this novel ZEB1 regulatory axis may be a viable strategy for the treatment of pulmonary fibrosis.

Introduction

Epithelial-mesenchymal transition (EMT), a dynamic and reversible biological process by which epithelial cells lose their cell polarity and down-regulate cadherin-mediated cell-cell adhesion to gain migratory properties, is involved in embryonic development, wound healing, fibrosis, and cancer metastasis[1]. EMT is executed in response to pleiotropic signalling factors, including the transforming growth factor β (TGFβ) superfamily, Sonic Hedgehog (Shh), Wnt/β-catenin, fibroblast growth factor (FGF) and epidermal growth factor (EGF). These factors regulate the expression of specific transcription factors (TFs) called EMT-TFs (e.g., Snail, ZEB, Twist, and others) that promote repression of epithelial features and induction of mesenchymal characteristics[2, 3]. Unlike EMT in cancer, which is detrimental, wound-healing-driven EMT induced in response to injury is beneficial, but exaggerated healing responses can lead to fibrosis, or tissue scarring.

Fibrosis is a hallmark of many chronic degenerative disorders and is associated with reduced organ function and eventual organ failure. Fibrotic disease is on the increase; for example, idiopathic pulmonary fibrosis (IPF), the most common type of idiopathic interstitial pneumonia, occurs with similar frequency to that of stomach, brain, and testicular cancer[4]. IPF is now generally regarded as a consequence of multiple interacting genetic and environmental risk factors, with repetitive local micro-injuries to ageing alveolar epithelium playing a central role[5]. These micro-injuries initiate the progressive accumulation of extracellular matrix (ECM) deposited by myofibroblasts. The origin of these myofibroblasts has been debated for many years, with EMT being considered as a potential source by driving the transformation of epithelial cells into ECM producing myofibroblasts[6, 7, 8, 9, 10]. However, lineage tracing in transgenic mice indicates that the contribution of those cells to the

population of myofibroblasts is negligible[11, 12, 13, 14].

In this study, we identify a novel regulatory axis involved in lung fibrosis whereby EMT contributes to the fibrotic process via paracrine activation of fibroblasts. We demonstrate that epidermal growth factor receptor (EGFR)-RAS-extracellular signal-regulated kinase (ERK) signalling induces the transcription factor ZEB1, which not only controls EMT but also regulates the production of locally-acting mediators. Specifically we identified tissue plasminogen activator (tPA) as a downstream effector of ZEB1 transcriptional activity that contributes to paracrine signalling by enhancing TGFβ-induced profibrogenic responses in fibroblasts. Consistent with this, increased ZEB1 nuclear expression was detected in alveolar epithelium adjacent to sites of ECM deposition in IPF lung tissue. Thus, rather than contributing directly to the mesenchymal population, our data suggest that ZEB1-dependent EMT of ATII cells contributes to fibrosis via epithelial-fibroblast crosstalk. The occurrence of ZEB1 activation at sites of local ECM deposition in IPF lung tissue is consistent with the concept that ZEB1-regulated paracrine signalling contributes to development of a profibrogenic microenvironment leading to interstitial lung fibrosis.

Results

Activation of EGFR signalling induces EMT in alveolar epithelial cells.

To investigate IPF associated signalling pathways, we analysed differentially expressed genes in IPF and control lung tissue from a publicly available microarray dataset (GSE24206)[15]. Using a false discovery rate (FDR) corrected *P* value of 0.05, we identified 7668 genes to be differentially expressed out of a total of 54675 probe sets. Gene network analysis using the Consensus Pathways Database[16] identified a number of pathways. Of these the EGFR-ERK pathway was the top ranked pathway with 150 of 458 pathway candidates being significantly (Q-value < 0.05) over represented in the dataset (Supplementary Fig. S1a).

Based on the transcriptomic data, we hypothesised an important role of EGFR signalling in IPF. Identification of pathological mechanisms of IPF has been challenging; however, dysregulation of alveolar type 2 (ATII) epithelial cells is thought to be central [5]. We therefore treated a human ATII cell line ($ATII^{ER:KRASV12}$)[17, 18] with EGF (Fig. 1b-d; Supplementary Fig. S1b) or transforming growth factor α (TGFα) (Supplementary Fig. S1b) to activate EGFR signalling. The human ATII cell line grows in continuous culture and expresses the ATII cell marker, pro-surfactant protein C (ProSP-C) (Fig. 1a; Fig. 2f). Our results showed that treatment of $ATII^{ER:KRASV12}$ cells with EGF for 24 hrs induced EMT, reflected by a change in their morphology from typical cuboidal epithelial cells to a more elongated mesenchymal cell phenotype with a reorganization of the actin cytoskeleton as demonstrated using Phalloidin staining of filamentous actin (F-actin) (Fig. 1b). This phenotypic switch was accompanied by a significant increase in mRNA expression of *ZEB1* and *VIM* (Vimentin), and a reduction in *CDH1* (E-cadherin); mRNA levels of other EMT-TFs, such as *SNAI1*, *SNAI2*, *TWIST* and *ZEB2* were not increased by activation of EGFR

signalling (Fig. 1c). The changes in ZEB1 and E-cadherin were further confirmed by Western blot analysis (Fig. 1d; Supplementary Fig. S1b).

Similar results were obtained using primary human ATII cells treated with EGF where an increase in ZEB1 expression was associated with down-regulation of E-cadherin (Fig. 1e). Under the same conditions, however, TGFβ was not able to induce EMT in the primary human ATII cells (Fig. 1e). Together, these results demonstrate that activation of EGFR signalling is able to activate the EMT programme in ATII cells, which is supported by a morphology change, the induction of the EMT-TF ZEB1 and a mesenchymal marker Vimentin as well as a reduction in E-cadherin expression.

Activation of the RAS pathway drives EMT via ERK-ZEB1 in ATII cells.

RAS signalling is one of the most important pathways downstream of EGFR activation and is involved in a variety of physiological and pathological responses, including EMT[19, 20, 21]. To investigate whether the RAS pathway is important for EMT in ATII cells, we utilised a RAS-inducible ATII cell model. KRASV12 (containing a single amino acid mutation in *KRAS*, glycine to valine at position 12) fused to the oestrogen receptor (ER) ligand-binding domain[22] was introduced into ATII cells to generate ATII$^{ER:KRASV12}$, in which KRASV12 expression is induced by 4-hydroxytamoxifen (4-OHT)[17, 18]. Like EGF, direct activation of the RAS pathway in ATII$^{ER:KRASV12}$ cells by treatment with 4-OHT induced EMT, reflected by a reduction in E-cadherin levels and an increase in ZEB1 and Vimentin expression (Fig. 2a and b). Time-course analysis further demonstrated that the induction of ZEB1 by RAS

activation preceded the down-regulation of E-cadherin (Fig. 2c). Consistently, an EMT morphology change with an increase in ZEB1 expression was observed upon RAS activation (Fig. 2d). When grown on a thick layer of Matrigel, ATII cells form spheres (a 3D culture model)[23]. We adopted this experimental system and used ATII$^{ER:KRASV12}$ cells to investigate whether RAS activation induces EMT in 3D cultures. Control ATII$^{ER:KRASV12}$ cells formed single round spheres. Induction of oncogenic *KRAS* by 4-OHT resulted in spheres invading into the Matrigel with protrusions (Fig. 2e; Supplementary Fig. S2). We recovered these cells from the Matrigel, and examined the protein expression. We confirmed that RAS activation induced EMT in 3D cultures, demonstrated by a reduction in E-cadherin, and an increase in ZEB1 and Vimentin expression (Fig. 2f). These observations suggest that EGFR signalling and the downstream RAS pathway are able to induce EMT in ATII cells.

Since RAS activity regulates both the RAF-ERK and phosphoinositide 3-kinase (PI3K)-protein kinase B (AKT) signalling pathways, we next investigated which one is required for EMT in the ATII cells using inhibitors for these pathways. Treatment with the ERK inhibitor U0126 in ATII$^{ER:KRASV12}$ cells was sufficient to inhibit RAS-induced ZEB1 and Vimentin expression, as well as to restore the expression of E-cadherin and the epithelial morphology; in contrast, the AKT inhibitor AKT VIII failed to do so (Fig. 3a and b; Supplementary Fig. S3a).

We next investigated which EMT-TFs are important for RAS-induced EMT in ATII cells. *ZEB1* RNA interference (RNAi), but not Snail1 or Snail2 RNAi, was able to restore E-cadherin expression and the epithelial morphology in 4-OHT-treated ATII$^{ER:KRASV12}$ cells (Fig. 3c and d; Supplementary Fig. S3b), in line with the fact the ZEB1 was the only EMT-TF induced by EGFR-RAS signalling (Fig. 1 and 2). Taken

together, our results identify that RAS activation in human ATII cells drives EMT via ERK-ZEB1 pathway.

ZEB1 is highly expressed in IPF alveolar epithelium and is critical for transcriptional regulation of secreted factors that mediate crosstalk between ATII cells and fibroblasts.

Given our *in vitro* findings, we compared ZEB1 expression in IPF and control lung tissue. In IPF tissue, we detected strong nuclear expression of ZEB1 not only in fibroblastic foci (Fig. 4a) but also in epithelial cells of thickened alveoli septae where collagen deposition in the interstitium was evident (Fig. 4b); in contrast, little ZEB1 staining or collagen deposition was observed in alveoli of control lung tissue (Fig. 4c). The presence of nuclear ZEB1 staining in alveolar epithelial cells within IPF lung tissue suggests that these cells are undergoing EMT; furthermore, the presence of ECM suggests induction of mesenchymal responses, either directly via the repairing epithelial cells undergoing EMT or by crosstalk with underlying fibroblasts.

Comparison of the relative expression of ECM components in RAS-activated $ATII^{ER:KRASV12}$ cells and fibroblasts highlights that ATII cells produce extremely low levels of ECM genes even after the induction of EMT (Supplementary Fig. S4a), suggesting that ECM production in fibrosis is more likely to be a consequence of fibroblast activation than direct deposition by epithelial cells undergoing EMT. Therefore we investigated whether ATII cells undergoing RAS-induced EMT produce paracrine factors that activate fibroblasts. For these experiments, we took advantage of the ability of 4-OHT to induce RAS pathway activation in $ATII^{ER:KRASV12}$ cells, as this was not dependent on exogenous growth factors that might directly affect fibroblast responses. We treated the MRC5 or primary human parenchymal lung

fibroblasts with conditioned media (CM) from control or 4-OHT-treated ATII$^{ER:KRASV12}$ cells in the absence or presence of TGFβ, and evaluated the fibroblast responses by measuring the expression of α-smooth muscle actin (α-SMA, a myofibroblast marker) and other ECM genes, including *COL1A1*, *COL3A1* and *FN1*. On its own, CM from RAS-activated ATII$^{ER:KRASV12}$ cells (4-OHT-treated ATII CM) had little effect on the activation of fibroblasts (Fig. 5). However, 4-OHT-treated ATII CM together with TGFβ achieved a synergistic effect in activating fibroblasts, reflected by a larger increase in α-SMA (*ACTA2*), *COL1A1* and *FN1* levels (Fig. 5a and b). Of note, 4-OHT-treated ATII CM did not augment Smad2 phosphorylation suggesting a Smad2-independent response (Fig. 5a and c). Similar results were obtained using primary human lung fibroblasts from IPF patients (IPF fibroblasts, IPFFs) and control donors (normal human lung fibroblasts, NHLFs) (Fig. 5c; Supplementary Fig. S4b).

Given the important role of ZEB1 in mediating RAS-induced EMT and the fact that ZEB1 is highly expressed in the alveolar epithelium of IPF patients, we hypothesised that ZEB1 may determine the paracrine signalling produced by ATII cells undergoing RAS-induced EMT. *ZEB1* RNAi (Fig. 6a) in ATII cells completely abolished the effects of CM from RAS-activated ATII cells on TGFβ-induced activation of fibroblasts (Fig. 6b; Supplementary Fig. S5), highlighting ZEB1 as a key regulator of EMT as well as the paracrine signalling between ATII cells and fibroblasts.

ZEB1 regulates the expression of tissue plasminogen activator (tPA), which acts as a paracrine regulator of TGFβ-induced fibroblast activation.

By performing quantitative proteomic analysis of the CM from control or 4-OHT-treated ATII$^{ER:KRASV12}$ cells, we identified ~ 430 secreted proteins whose levels

changed during RAS-induced EMT. We then checked their expression in pulmonary epithelial cells from control and IPF lung tissue using a publicly available dataset[24], and identified a total number of 25 genes/proteins that were elevated in IPF lung epithelial cells as well as in CM from 4-OHT-treated ATII$^{ER:KRASV12}$ cells (Supplementary Table S1). Of these, *PLAT*, which encodes tissue plasminogen activator (tPA) was most up-regulated in IPF epithelial cells (Fig. 7a; Supplementary Table S1) and we confirmed enhanced secretion of tPA in the CM from 4-OHT-treated ATII$^{ER:KRASV12}$ cells by Western blotting (Fig. 7b). As we had identified ZEB1 as the key regulator of epithelial-mesenchymal crosstalk, we scanned the promoter of *PLAT* for the presence of ZEB1 binding motifs (5'-CANNTG-3') and found a ZEB1 binding site -419 bp upstream of the transcriptional start site (TSS) (Supplementary Fig. S6a). Further experiments showed that the mRNA expression of *PLAT* was increased upon RAS-activation in ATII cells and this was repressed by *ZEB1* RNAi (Fig. 7c).

To validate the ZEB1 binding site in the *PLAT* promoter, we first performed a chromatin immunoprecipitation (ChIP) assay. An anti-ZEB1 antibody was used to precipitate formaldehyde cross-linked ZEB1-DNA complexes in ATII$^{ER:KRASV12}$ cells treated without or with 4-OHT. The presence of *PLAT* promoter DNA sequences in the immunoprecipitate was verified by PCR using primers amplifying the region between -547 and -345 upstream of the TSS, and we found RAS activation in ATII cells increased ZEB1 occupancy on the *PLAT* promoter (Fig. 7d; Supplementary Fig. S6b). We next generated two *PLAT* promoter constructs (-689 to -1 upstream of the TSS) which were cloned into a pGL3 basic luciferase reporter plasmid and transfected into ATII cells; the pGL3 basic-*PLAT* (−689 to −1) construct contained the ZEB1 motif whereas this was deleted in the second construct (delta −419 to −414 upstream

of the TSS) (pGL3 basic-Δ ZEB1 motif). RAS activation by 4-OHT in ATII$^{ER:KRASV12}$ cells resulted in a significant increase in pGL3 basic-*PLAT* (−689 to −1) luciferase activity. Under the same conditions, luciferase activity was not increased using pGL3 basic-Δ ZEB1 motif (Fig. 7e). These data confirm that *PLAT* (tPA) is a transcriptional target of ZEB1 in response to RAS activation in ATII cells.

Consistent with a previous report[25], tPA synergistically promoted TGFβ-induced α-SMA expression in human lung fibroblasts (Supplementary Fig. S6c). Like ZEB1, *PLAT* RNAi (Supplementary Fig. S6d) in ATII cells completely abolished the effects of CM from RAS-activated ATII cells on TGFβ-induced α-SMA expression in fibroblasts (Fig. 7f), demonstrating tPA as a key paracrine factor secreted by ATII cells undergoing RAS-induced EMT. These results provide clear evidence that a ZEB1-tPA axis is involved in the paracrine signalling between ATII cells undergoing RAS-induced EMT and fibroblasts to augment their differentiation into myofibroblasts caused by TGFβ.

Finally, in view of the requirement for exogenous TGFβ to demonstrate an effect of the 4-OHT-treated ATII CM on fibroblasts, we investigated whether ATII cells in fibrotic tissue *in vivo* or those undergoing injury/repair *in vitro* expressed endogenous TGFβ. Using a publicly available dataset[24], we found that the major *TGFB* isoform expressed by alveolar epithelial cells *in vivo* was *TGFB2* and that this was expressed at significantly higher levels in IPF compared with control lung tissue (Supplementary Fig. S7a). In contrast with the study in kidney[14], the data also revealed that Snail2 is up-regulated in IPF *vs.* control lung epithelial cells, but not Snail1 or Twist (Supplementary Fig. S7b). As we have previously shown that scrape-wounding of bronchial epithelial cells stimulates release of TGFβ2 independently of EGFR activation[26], we examined whether damage of ATII cells similarly affected

TGFB2 expression. This showed that scrape-wounded ATII cells expressed more *TGFB2* and this increased in proportion to the extent of injury (Supplementary Fig. S7c). These data suggest that damaged ATII cells are a potential a source of TGFβ *in vivo*.

Discussion

Fibrotic diseases are a major cause of morbidity and mortality worldwide and their prevalence is increasing with an ageing population. Abnormal wound healing responses appear to make major contributions to the scarring process, but the underlying pathological mechanisms are unclear, especially the role of EMT. In this study, we have used a variety of approaches to show that activation of EGFR-RAS-ERK signalling in ATII cells induces EMT via the transcriptional regulator ZEB1. Importantly, beyond its effects on the epithelial cell phenotype, we have identified that ZEB1 is a regulator of paracrine signalling between lung epithelial cells and fibroblasts, as ATII cells undergoing RAS-induced EMT secrete tPA to augment TGFβ-induced myofibroblast differentiation (Fig. 8). This may be an important profibrotic event as, relative to epithelial cells, the ability of fibroblasts to synthesise ECM is orders of magnitude greater.

Consistent with previous findings[27, 28], we found strong expression of ZEB1 in the epithelium in proximity to fibroblastic foci in IPF lung tissue. However, we also found ZEB1 was expressed in epithelial cells of thickened alveolar septae where ECM deposition was evident. This suggests that ZEB1 is induced as an early response to alveolar epithelial injury and that, by regulating expression of factors involved in paracrine signalling, ZEB1 may promote TGFβ-induced fibroblast activation in IPF. While this may be a normal physiological response to injury, persistent epithelial injury and/or failure to resolve the lesion may sensitise the underlying fibroblasts to drive a pathologic profibrogenic response. In line with this, exposure of human lung cells to nickel (Ni), an environmental and occupational pollutant linked to lung fibrosis[29], caused ZEB1-dependent EMT, which was irreversible even after the termination of Ni exposure[30]. Thus, it is conceivable that repetitive environmental

exposures to metals such as Ni could lead to deregulation of ZEB1 to cause persistent EMT and exaggerated profibrogenic crosstalk during the initiation of IPF.

EMT in the ATII cells was strongly induced by EGFR activation. The EGFR is a transmembrane receptor tyrosine kinase activated by members of the EGF family, including EGF and TGFα[31]. EGFR dimerisation activates one or more downstream effectors including the ERK, PI3K/AKT, STAT (signal transducer and activator of transcription), and mTOR (mammalian target of rapamycin) pathways through receptor autophosphorylation and cytoplasmic protein binding[32, 33, 34]. These in turn act as critical mediators of airway and alveolar homeostasis, with aberrant activation within one or more pathway components capable of driving a variety of respiratory pathologies, including lung fibrosis[33, 35]. The EGFR pathway has been implicated in lung fibrosis through studies in which transgenic mice that constitutively express TGFα in epithelial cells develop progressive lung fibrosis[36, 37]. Conversely, mice deficient in TGFα that lack normal EGFR signalling or that are treated with EGFR pathway inhibitors exhibit resistance to bleomycin-induced lung fibrosis[38]. In IPF patients, *EGFR* mutations[39] or increased expression of the EGFR[40] have been reported. Our evidence that an EGFR-RAS-ERK-ZEB1 axis may contribute to the early stages of lung fibrosis suggests that inhibiting EGFR signalling may be of clinical relevance for regulating human fibrotic lung disease.

A key finding of our study was identification that ZEB1 controls tPA expression and that this affects the sensitivity of fibroblast activation induced by TGFβ. While tPA is a key activator of fibrinolysis, it also has direct cellular effects by virtue of its ability to bind to the low density lipoprotein (LDL) receptor-related protein-1 (LRP-1), triggering LRP-1 tyrosine phosphorylation, and recruitment of β1-integrin signalling involving integrin-linked kinase (ILK)[25]. In this context, tPA acts as a

survival factor that protects fibroblasts/myofibroblasts from apoptosis, and it has previously been implicated in kidney fibrosis[25], but no other fibrotic conditions. In keeping with a role of tPA in human lung fibrosis was the identification in a publicly available dataset[24] that there is an increased expression of *PLAT* (tPA) by IPF epithelial cells.

EMT converts epithelial cells into migratory and/or invasive mesenchymal cells, and is well established during development and carcinogenesis, however, its role in fibrosis has been more controversial[1, 3, 41]. Although human IPF tissue studies have demonstrated co-localization of epithelial and mesenchymal markers[27, 28, 42], the number of fibroblasts arising from epithelial cells was small in some mouse lineage tracing studies[11, 12], suggesting effects of EMT beyond direct phenotypic conversion into matrix-producing cells. Recent studies in renal fibrosis have reported that tubular epithelial cells undergoing EMT relay signals to the interstitium, which promote myofibroblast differentiation and fibrogenesis, without directly contributing to the myofibroblast population[14, 43, 44, 45]. In tubulointerstitial renal fibrosis, TGFβ induces EMT via Snail1, and then Snail1 induces TGFβ expression generating an autocrine loop that sustains the progression of the disease by influencing the differentiation of fibroblasts into myofibroblats[14]. In contrast with studies in kidney, our analysis of publicly available transcriptomic datasets of IPF lung tissue, identified the EGFR-ERK pathway as the top ranked pathway with 150 of 458 pathway candidates being significantly over represented in the IPF dataset, highlighting the potential importance of this pathway in IPF pathogenesis. Building on these observations, our *in vitro* studies of EGFR-RAS-ERK-induced EMT, identified ZEB1 as the main transcription factor that controlled EMT as well as paracrine signalling through regulation of tPA expression, which potentiated fibroblast differentiation in the presence of TGFβ.

However, as these paracrine effects required exogenous TGFβ, this raised the question of the source of TGFβ in lung fibrosis *in vivo*. While many cell types produce TGFβ isoforms which can also be stored as latent growth factor bound in the ECM[46, 47, 48], we focussed on the epithelium and found increased expression of *TGFB2*, as well as *SNAI2* in IPF epithelial cells using publicly available datasets, with scrape wounding of ATII cells also inducing *TGFB2* expression *in vitro*. The increased epithelial *TGFB2* signature highlights the potential for EGF (ZEB1) and TGFβ (Snail2) to synergize in paracrine activation of the underlying fibroblasts.

Together with previous findings in kidney fibrosis[14, 43], our study helps to provide a unified concept for the role of EMT in fibrosis: persistent EMT of epithelial cells may dysregulate paracrine signalling between epithelial and mesenchymal cells, so creating a profibrogenic microenvironment which leads to the development of fibrosis. Based on the relative low levels of ECM biosynthesis by epithelial cells and the relatively small numbers of mesenchymal cells identified in lineage tracing studies[11, 12, 14, 43], this mechanism may be more important than direct conversion of epithelial cells into mesenchymal cells. While the finer details of these paracrine mechanisms may vary according to disease and tissue location, our identification of ZEB1 as a key regulator of EGF/RAS-induced EMT and an enhancer of paracrine signalling mediating the crosstalk between ATII cells and lung fibroblasts may help to find drug targets or biomarkers to intervene or predict the progression of pulmonary fibrosis.

Methods

Lung tissue sampling. All human lung experiments were approved by the Southampton and South West Hampshire and the Mid and South Buckinghamshire Local Research Ethics Committees, and all subjects gave written informed consent. Clinically indicated IPF lung biopsy tissue samples and non-fibrotic control tissue samples (macroscopically normal lung sampled remote from a cancer site) were deemed surplus to clinical diagnostic requirements. All IPF samples were from patients subsequently receiving a multidisciplinary diagnosis of IPF according to international consensus guidelines[49].

Cell culture, reagents and transfections. Primary parenchymal lung fibroblast cultures were established from IPF or control lung tissue as described previously[50]. Fibroblasts were cultured in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% foetal bovine serum (FBS), 50 units/ml penicillin, 50μg/ml streptomycin, 2mM L-glutamine, 1mM sodium pyruvate and 1x non-essential amino acids (DMEM/FBS) (all from Life Technologies).

Primary human ATII cells were isolated from macroscopically normal regions of surgically resected lung parenchyma as described previously[50, 51]. The alveolar epithelial cells were resuspended in fresh DCCM-1 (Biological Industries Ltd) supplemented with 10% new-born calf serum (NBCS) (Life Technologies), 1% penicillin, 1% streptomycin and 1% L-glutamine (all from Sigma Aldrich) and plated on collagen 1 (PureCol 5005-b, Advanced BioMatrix Inc.) coated 96 well plates at 60% density; purity of the cultures was determined by staining for alkaline phosphatase.

ATII$^{ER:KRASV12}$ cells[17, 18] were cultured in DCCM-1 (Biological Industries Ltd) supplemented with 10% NBCS (Life Technologies), 1% penicillin, 1% streptomycin and 1% L-glutamine (all from Sigma Aldrich). To induce RAS activation in ATII$^{ER:KRASV12}$ cells, 250 nM 4-OHT (Sigma-Aldrich) was added[17, 18]. MRC5 cells were obtained from the European Collection of Authenticated Cell Cultures (ECACC) and were cultured in DMEM (Thermo Fisher Scientific). Both cell culture media were supplemented with 10% FBS (Thermo Fisher Scientific), 1% Penicillin/Streptomycin and glutamine (Thermo Fisher Scientific). All cells were kept at 37 °C and 5% CO2. For 3D culture, ATII$^{ER:KRASV12}$ cells were cultured as previously described[23] in Matrigel (BD Biosciences). TGFα was from Fisher Scientific UK Ltd. TGFβ1 was from PeproTech. EGF and recombinant human tPA protein were from Bio-Techne. AKT VIII and U0126 were from Sigma Aldrich. No mycoplasma contamination was detected in the cell lines used.

Short interfering RNA (siRNA) oligos against *ZEB1* (MU-006564-02-0002), *SNAI1* (Snail1) (MU-010847-00-0002), *SNAI2* (Snail2) (MU-017386-00-0002) and *PLAT* (tPA) (MU-005999-01-0002) were purchased from Dharmacon. Sequences are available from Dharmacon, or on request. As a negative control we used siGENOME RISC-Free siRNA (Dharmacon). ATII$^{ER:KRASV12}$ cells were transfected with the indicated siRNA oligos at a final concentration of 35 nM using DharmaFECT 2 reagent (Dharmacon).

Western blot analysis. Western blot analysis was performed with lysates from cells with urea buffer (8 M Urea, 1 M Thiourea, 0.5% CHAPS, 50 mM DTT, and 24 mM Spermine). Primary antibodies were from: Santa Cruz (β-actin, sc-47778; ZEB1, sc-25388; ZEB2, sc-48789; E-cadherin, sc-21791; Snail2, sc-10436), Abcam (β-tubulin,

ab6046), Cell Signalling Technology (α-SMA, 14968; phospho-AKT, 9271; phospho-ERK, 9101; Snail1, 3879; Snail2, 9585; TWIST, 46702; Phospho-Smad2, 3104; β-tubulin, 86298), BD Transduction Laboratories (E-cadherin, 610405; Vimentin, 550513) and Millipore (proSP-C, AB3786; tPA, 05-883). Signals were detected using an Odyssey imaging system (LI-COR), and evaluated by ImageJ 1.42q software (National Institutes of Health).

qRT-PCR. Total RNA was isolated using RNeasy mini kit (Qiagen) according to manufacturer's instructions and quantified using a Nanodrop Spectophotometer 2000c (Thermo Fisher Scientific). Real-time quantitative RT-PCR was carried out using gene-specific primers (QuantiTect Primer Assays, Qiagen) for *CDH1* (E-cadherin) (QT00080143), *SNAI1* (Snail1) (QT00010010), *SNAI2* (Snail2) (QT00044128), *ZEB1* (QT00008555), *ZEB2* (QT00008554), *TWIST* (QT00011956), *VIM* (QT00095795), *COL1A1* (QT00037793), *COL3A1* (QT00058233), *FN1* (QT00038024), *ACTA2* (α-SMA) (QT00088102), *PLAT* (tPA) (QT00075761), *TGFB1* (QT00000728), *TGFB2* (QT00025718), *GAPDH* (QT01192646) or *ACTB* (β-actin) (QT01680476) with QuantiNova SYBR Green RT-PCR kits (Qiagen). Relative transcript levels of target genes were normalised to *GAPDH* or *ACTB* (β-actin).

Immunofluorescence microscopy. Cells were fixed in 4% PBS-paraformaldehyde for 15 min, incubated in 0.1% Triton X-100 for 5 min on ice, then in 0.2% fish skin gelatin in PBS for 1 h and stained for 1 h with an anti-Prosurfactant Protein C (proSP-C) antibody (1:100, Millipore AB3786, rabbit polyclonal) or anti-ZEB1 (1:100, Santa

Cruz sc-25388, rabbit polyclonal). Protein expression was detected using Alexa Fluor (1:400, Molecular Probes) for 20 min. DAPI (Invitrogen) was used to stain nuclei (1:1,000). Rhodamine phalloidin was used to visualize filamentous actin (F-actin) (Molecular Probes). For immunofluorescence staining of 3D cultures from $ATII^{ER:KRASV12}$ cells, spheres were fixed with 4% PBS-paraformaldehyde for 40 min, permeabilised in 0.5% Triton X-100 for 10 min on ice and stained with rhodamine phalloidin for 1 h at room temperature. Spheres were counterstained with DAPI. Samples were observed using a confocal microscope system (Leica SP8). Acquired images were analysed using Photoshop (Adobe Systems) according to the guidelines of the journal.

Immunohistochemistry, haematoxylin and eosin (H/E) and tinctorial stains. Control or IPF lung tissues (n = 3 donors) were fixed and embedded in paraffin wax; tissue sections (4μm) were processed and stained as previously described [20]. Briefly, the tissue sections were de-waxed, rehydrated and incubated with 3% hydrogen peroxide in methanol for 10 min to block endogenous peroxidase activity. Sections were then blocked with normal goat serum and incubated at room temperature with a primary antibody against ZEB1 (1:500, Sigma), followed by a biotinylated secondary antibody (1:500, Vector Laboratories Ltd, UK); antibody binding was detected using streptavidin-conjugated horse-radish peroxidase and visualised using DAB (DAKO) before counterstaining with Mayer's Haematoxylin. For H/E stain, Shandon Varistain 24-4 automatic slide stainer (Thermo Fisher Scientific) was used. For tinctorial stain, Trichrome stain (Abcam ab150686) was used according to the manufacturers' instructions. Images were acquired using an Olympus Dotslide Scanner VS110.

Chromatin immunoprecipitation (ChIP). ChIP assays were carried out using SimpleChIP enzymatic chromatin IP kits (Cell Signalling Technology) as per the manufacturer's instructions. Briefly, ATII$^{ER:KRASV12}$ cells with indicated treatments were incubated for 10 min with 1% formaldehyde solution at room temperature, followed by incubation with 125 mM glycine. Antibodies used for ChIP were as follows: rabbit anti-ZEB1 (PA5-28221, Invitrogen, rabbit polyclonal, 5µg per IP sample), normal rabbit IgG (2729, Cell Signalling Technology, 5µg per IP sample). For the ZEB1 binding site at position -419 of the human *PLAT* (tPA) promoter, the primers amplifying the region between -547 and -345 were as follows: forward 5'-GGAAAGTCCCCGGAGGCCACCTA-3' and reverse 5'-TGGAACACTTTGTGTGGTGGC-3'. DNA fragments were quantified by qPCR. PCR products were analysed in a 1.5% agarose gel by ethidium bromide staining.

Luciferase constructs and luciferase reporter assays. The human *PLAT* (tPA) promoter (sequence -689 to -1 upstream of the TSS) was amplified from human genomic DNA by PCR, and was subsequently cloned into pGL3 basic vector (Promega), termed pGL3 basic-*PLAT* (−689 to −1). The putative ZEB1 binding site, positioned −419 to −414 on the human *PLAT* promoter, was removed from pGL3 basic-*PLAT* (−689 to −1) construct to create the pGL3 basic-Δ ZEB1 motif construct.

For the luciferase reporter assays, ATII$^{ER:KRASV12}$ cells were transfected using Lipofectamine 3000 (Invitrogen) with 80 ng of phRL-CMV (Promega), which constitutively expresses the *Renilla* luciferase reporter, plus 600 ng of pGL3 basic-*PLAT* (−689 to −1) or pGL3 basic-Δ ZEB1 motif per well in the presence or absence of 4-OHT. Finally, the transcriptional assay was carried out using the Dual-Luciferase reporter assay system (Promega) following the manufacturer's protocol.

Quantitative proteomic analysis of the secretome and the subsequent data analysis. Serum-free conditioned media (CM) from ATII$^{ER:KRASV12}$ cells treated without or with 4-OHT (250 nM, 24 hrs) were analysed using an enrichment strategy based upon Strataclean resin (Agilent) in combination with the quantitative label-free approach, LC-MS^{E}, to provide in-depth proteome coverage and estimates of protein concentration in absolute amounts[52] (Details provided in Supplementary Methods).

Raw data were processed and collated into a single .csv document. Values were then normalised to total fmol of each sample multiplied by 10,000. Pseudo-counts were applied to the normalised values to replace missing ones, to allow for full statistical analysis to be completed[53]. We first sorted the normalised values in each column in order of abundance, in ascending order, then the minimum value of each sample identified. This minimum was used to replace all missing values in the data set. A two-tailed, unpaired Student's *t*-test was used to compare two groups for independent samples. $P < 0.05$ was considered statistically significant.

In order to highlight their implications in IPF, differentially expressed proteins/genes identified in the quantitative secretome analysis were searched in LGEA web portal (https://research.cchmc.org/pbge/lunggens/mainportal.html) for their levels in pulmonary epithelial cells from control and IPF lung tissue.

Bioinformatics. IPF transcriptomic data was downloaded from the NCBI's Gene Expression Omnibus (GEO). We used data from GSE24206[15], a microarray study comparing samples from 11 IPF patients undergoing lung transplantation or diagnostic biopsy to 6 normal lung samples taken from lung transplantation donors. Microarray series matrix files were imported into R, and differential expression

analysis comparing normal to IPF samples performed using the R package limma[54]. Data were log-transformed before analysis. To correct for multiple testing, a Benjamini-Hochberg false discovery rate (FDR) of 5% was applied to the data, and a Q-value cut-off of 0.02 was used to determine significance. Differentially expressed gene lists were input into the human Consensus Pathways Database, which determined pathways with differentially expressed genes overrepresented in this database. A 5% FDR was used as above.

Statistical analysis and repeatability of experiments. Each experiment was repeated at least twice. Unless otherwise noted, data are presented as mean and s.d., and a two-tailed, unpaired Student's *t*-test was used to compare two groups for independent samples. $P < 0.05$ was considered statistically significant.

Acknowledgements

This project was supported by the Academy of Medical Sciences/the Wellcome Trust Springboard Award [SBF002\1038], Wessex Medical Trust and AAIR Charity. LY was supported by China Scholarship Council. FC was supported by Medical Research Foundation [MRF-091-0003-RG-CONFO]. CH was supported by Gerald Kerkut Charitable Trust and University of Southampton Central VC Scholarship Scheme. MGJ was supported by the Wellcome Trust [100638/Z/12/Z]. DED was supported by British Lung Foundation [BLF-RG14-14]. YW was supported by the Ministry of Science and Technology of China National Key Research and Development Projects [2016YFC0904701] and the National Natural Science Foundation of China [81772827]. Instrumentation in the Centre for Proteomic Research is supported by the BBSRC [BM/M012387/1] and the Wessex Medical Trust. We thank Carine Fixmer, Maria Lane, Benjamin Johnson and the nurses of the Southampton Biomedical Research Unit for their help in the collection of human samples, supported by the Wessex Clinical Research Network and the National Institute of Health Research, UK.

References

1. Nieto MA, Huang RY, Jackson RA, Thiery JP. Emt: 2016. *Cell* 2016, **166**(1)**:** 21-45.

2. Peinado H, Olmeda D, Cano A. Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nature reviews Cancer* 2007, **7**(6)**:** 415-428.

3. Nieto MA. The ins and outs of the epithelial to mesenchymal transition in health and disease. *Annual review of cell and developmental biology* 2011, **27:** 347-376.

4. Hutchinson J, Fogarty A, Hubbard R, McKeever T. Global incidence and mortality of idiopathic pulmonary fibrosis: a systematic review. *The European respiratory journal* 2015, **46**(3)**:** 795-806.

5. Richeldi L, Collard HR, Jones MG. Idiopathic pulmonary fibrosis. *Lancet* 2017, **389**(10082)**:** 1941-1952.

6. DeMaio L, Buckley ST, Krishnaveni MS, Flodby P, Dubourd M, Banfalvi A*, et al.* Ligand-independent transforming growth factor-beta type I receptor signalling mediates type I collagen-induced epithelial-mesenchymal transition. *The Journal of pathology* 2012, **226**(4)**:** 633-644.

7. Kim KK, Wei Y, Szekeres C, Kugler MC, Wolters PJ, Hill ML*, et al.* Epithelial cell alpha3beta1 integrin links beta-catenin and Smad signaling to promote myofibroblast formation and pulmonary fibrosis. *The Journal of clinical investigation* 2009, **119**(1)**:** 213-224.

8. Kim KK, Kugler MC, Wolters PJ, Robillard L, Galvez MG, Brumwell AN*, et al.* Alveolar epithelial cell mesenchymal transition develops in vivo during pulmonary fibrosis and is regulated by the extracellular matrix. *Proceedings of the National Academy of Sciences of the United States of America* 2006, **103**(35)**:** 13180-13185.

9. Tanjore H, Xu XC, Polosukhin VV, Degryse AL, Li B, Han W*, et al.* Contribution of epithelial-derived fibroblasts to bleomycin-induced lung fibrosis. *American journal of respiratory and critical care medicine* 2009, **180**(7)**:** 657-665.

10. Degryse AL, Tanjore H, Xu XC, Polosukhin VV, Jones BR, Boomershine CS*, et al.* TGFbeta signaling in lung epithelium regulates bleomycin-induced alveolar injury and fibroblast recruitment. *American journal of physiology Lung cellular and molecular physiology* 2011, **300**(6)**:** L887-897.

11. Humphreys BD, Lin SL, Kobayashi A, Hudson TE, Nowlin BT, Bonventre JV*, et al.* Fate tracing reveals the pericyte and not epithelial origin of

myofibroblasts in kidney fibrosis. *The American journal of pathology* 2010, **176**(1)**:** 85-97.

12. Rock JR, Barkauskas CE, Cronce MJ, Xue Y, Harris JR, Liang J*, et al.* Multiple stromal populations contribute to pulmonary fibrosis without evidence for epithelial to mesenchymal transition. *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**(52)**:** E1475-1483.

13. LeBleu VS, Taduri G, O'Connell J, Teng Y, Cooke VG, Woda C*, et al.* Origin and function of myofibroblasts in kidney fibrosis. *Nature medicine* 2013, **19**(8)**:** 1047-1053.

14. Grande MT, Sanchez-Laorden B, Lopez-Blau C, De Frutos CA, Boutet A, Arevalo M*, et al.* Snail1-induced partial epithelial-to-mesenchymal transition drives renal fibrosis in mice and can be targeted to reverse established disease. *Nature medicine* 2015, **21**(9)**:** 989-997.

15. Meltzer EB, Barry WT, D'Amico TA, Davis RD, Lin SS, Onaitis MW*, et al.* Bayesian probit regression model for the diagnosis of pulmonary fibrosis: proof-of-principle. *BMC medical genomics* 2011, **4:** 70.

16. Kamburov A, Stelzl U, Lehrach H, Herwig R. The ConsensusPathDB interaction database: 2013 update. *Nucleic acids research* 2013, **41**(Database issue)**:** D793-800.

17. Molina-Arcas M, Hancock DC, Sheridan C, Kumar MS, Downward J. Coordinate direct input of both KRAS and IGF1 receptor to activation of PI3 kinase in KRAS-mutant lung cancer. *Cancer discovery* 2013, **3**(5)**:** 548-563.

18. Coelho MA, de Carne Trecesson S, Rana S, Zecchin D, Moore C, Molina-Arcas M*, et al.* Oncogenic RAS Signaling Promotes Tumor Immunoresistance by Stabilizing PD-L1 mRNA. *Immunity* 2017, **47**(6)**:** 1083-1099 e1086.

19. Downward J. Targeting RAS signalling pathways in cancer therapy. *Nature reviews Cancer* 2003, **3**(1)**:** 11-22.

20. Wang Y, Bu F, Royer C, Serres S, Larkin JR, Soto MS*, et al.* ASPP2 controls epithelial plasticity and inhibits metastasis through beta-catenin-dependent regulation of ZEB1. *Nature cell biology* 2014, **16**(11)**:** 1092-1104.

21. Wang Y, Ngo VN, Marani M, Yang Y, Wright G, Staudt LM*, et al.* Critical role for transcriptional repressor Snail2 in transformation by oncogenic RAS in colorectal carcinoma cells. *Oncogene* 2010, **29**(33)**:** 4658-4670.

22. Dajee M, Tarutani M, Deng H, Cai T, Khavari PA. Epidermal Ras blockade demonstrates spatially localized Ras promotion of proliferation and inhibition of differentiation. *Oncogene* 2002, **21**(10)**:** 1527-1538.

23. Yu W, Fang X, Ewald A, Wong K, Hunt CA, Werb Z*, et al.* Formation of cysts by alveolar type II cells in three-dimensional culture reveals a novel mechanism for epithelial morphogenesis. *Molecular biology of the cell* 2007, **18**(5)**:** 1693-1700.

24. 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.

25. Hu K, Wu C, Mars WM, Liu Y. Tissue-type plasminogen activator promotes murine myofibroblast activation through LDL receptor-related protein 1-mediated integrin signaling. *The Journal of clinical investigation* 2007, **117**(12)**:** 3821-3832.

26. Puddicombe SM, Polosa R, Richter A, Krishna MT, Howarth PH, Holgate ST*, et al.* Involvement of the epidermal growth factor receptor in epithelial repair in asthma. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 2000, **14**(10)**:** 1362-1374.

27. Chilosi M, Calio A, Rossi A, Gilioli E, Pedica F, Montagna L*, et al.* Epithelial to mesenchymal transition-related proteins ZEB1, beta-catenin, and beta-tubulin-III in idiopathic pulmonary fibrosis. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* 2017, **30**(1)**:** 26-38.

28. Park JS, Park HJ, Park YS, Lee SM, Yim JJ, Yoo CG*, et al.* Clinical significance of mTOR, ZEB1, ROCK1 expression in lung tissues of pulmonary fibrosis patients. *BMC pulmonary medicine* 2014, **14:** 168.

29. Kitamura H, Ichinose S, Hosoya T, Ando T, Ikushima S, Oritsu M*, et al.* Inhalation of inorganic particles as a risk factor for idiopathic pulmonary fibrosis--elemental microanalysis of pulmonary lymph nodes obtained at autopsy cases. *Pathology, research and practice* 2007, **203**(8)**:** 575-585.

30. Jose CC, Jagannathan L, Tanwar VS, Zhang X, Zang C, Cuddapah S. Nickel exposure induces persistent mesenchymal phenotype in human lung epithelial cells through epigenetic activation of ZEB1. *Molecular carcinogenesis* 2018.

31. Linggi B, Carpenter G. ErbB receptors: new insights on mechanisms and biology. *Trends in cell biology* 2006, **16**(12)**:** 649-656.

32. Jorissen RN, Walker F, Pouliot N, Garrett TP, Ward CW, Burgess AW. Epidermal growth factor receptor: mechanisms of activation and signalling. *Experimental cell research* 2003, **284**(1)**:** 31-53.

33. Vallath S, Hynds RE, Succony L, Janes SM, Giangreco A. Targeting EGFR signalling in chronic lung disease: therapeutic challenges and opportunities. *The European respiratory journal* 2014, **44**(2)**:** 513-522.

34. Scaltriti M, Baselga J. The epidermal growth factor receptor pathway: a model for targeted therapy. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2006, **12**(18)**:** 5268-5272.

35. Burgel PR, Nadel JA. Epidermal growth factor receptor-mediated innate immune responses and their roles in airway diseases. *The European respiratory journal* 2008, **32**(4)**:** 1068-1081.

36. Korfhagen TR, Swantz RJ, Wert SE, McCarty JM, Kerlakian CB, Glasser SW*, et al.* Respiratory epithelial cell expression of human transforming growth factor-alpha induces lung fibrosis in transgenic mice. *The Journal of clinical investigation* 1994, **93**(4)**:** 1691-1699.

37. Hardie WD, Bruno MD, Huelsman KM, Iwamoto HS, Carrigan PE, Leikauf GD*, et al.* Postnatal lung function and morphology in transgenic mice expressing transforming growth factor-alpha. *The American journal of pathology* 1997, **151**(4)**:** 1075-1083.

38. Madtes DK, Elston AL, Hackman RC, Dunn AR, Clark JG. Transforming growth factor-alpha deficiency reduces pulmonary fibrosis in transgenic mice. *American journal of respiratory cell and molecular biology* 1999, **20**(5)**:** 924-934.

39. Stella GM, Inghilleri S, Pignochino Y, Zorzetto M, Oggionni T, Morbini P*, et al.* Activation of oncogenic pathways in idiopathic pulmonary fibrosis. *Translational oncology* 2014, **7**(5)**:** 650-655.

40. Tzouvelekis A, Ntolios P, Karameris A, Vilaras G, Boglou P, Koulelidis A*, et al.* Increased expression of epidermal growth factor receptor (EGF-R) in patients with different forms of lung fibrosis. *BioMed research international* 2013, **2013:** 654354.

41. Kage H, Borok Z. EMT and interstitial lung disease: a mysterious relationship. *Current opinion in pulmonary medicine* 2012, **18**(5)**:** 517-523.

42. Lomas NJ, Watts KL, Akram KM, Forsyth NR, Spiteri MA. Idiopathic pulmonary fibrosis: immunohistochemical analysis provides fresh insights into lung tissue remodelling with implications for novel prognostic markers. *International journal of clinical and experimental pathology* 2012, **5**(1)**:** 58-71.

43. Lovisa S, LeBleu VS, Tampe B, Sugimoto H, Vadnagara K, Carstens JL*, et al.* Epithelial-to-mesenchymal transition induces cell cycle arrest and parenchymal damage in renal fibrosis. *Nature medicine* 2015, **21**(9)**:** 998-1009.

44. Ovadya Y, Krizhanovsky V. A new Twist in kidney fibrosis. *Nature medicine* 2015, **21**(9)**:** 975-977.

45. Huang S, Susztak K. Epithelial Plasticity versus EMT in Kidney Fibrosis. *Trends in molecular medicine* 2016, **22**(1)**:** 4-6.

46. Kelley J. Cytokines of the lung. *The American review of respiratory disease* 1990, **141**(3)**:** 765-788.

47. Moore B, Murphy RF, Agrawal DK. Interaction of tgf-beta with immune cells in airway disease. *Current molecular medicine* 2008, **8**(5)**:** 427-436.

48. Robertson IB, Horiguchi M, Zilberberg L, Dabovic B, Hadjiolova K, Rifkin DB. Latent TGF-beta-binding proteins. *Matrix biology : journal of the International Society for Matrix Biology* 2015, **47:** 44-53.

49. Raghu G, Collard HR, Egan JJ, Martinez FJ, Behr J, Brown KK*, et al.* An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *American journal of respiratory and critical care medicine* 2011, **183**(6)**:** 788-824.

50. Conforti F, Davies ER, Calderwood CJ, Thatcher TH, Jones MG, Smart DE*, et al.* The histone deacetylase inhibitor, romidepsin, as a potential treatment for pulmonary fibrosis. *Oncotarget* 2017, **8**(30)**:** 48737-48754.

51. Witherden IR, Vanden Bon EJ, Goldstraw P, Ratcliffe C, Pastorino U, Tetley TD. Primary human alveolar type II epithelial cell chemokine release: effects of cigarette smoke and neutrophil elastase. *American journal of respiratory cell and molecular biology* 2004, **30**(4)**:** 500-509.

52. Silva JC, Gorenstein MV, Li GZ, Vissers JP, Geromanos SJ. Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Molecular & cellular proteomics : MCP* 2006, **5**(1)**:** 144-156.

53. Wang J, Li L, Chen T, Ma J, Zhu Y, Zhuang J*, et al.* In-depth method assessments of differentially expressed protein detection for shotgun proteomics data with missing values. *Scientific reports* 2017, **7**(1)**:** 3367.

54. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W*, et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids research* 2015, **43**(7)**:** e47.

Figure Legends

Figure 1 Activation of EGFR signalling induces EMT in alveolar epithelial cells. (**a**) Immunofluorescence staining of Pro-surfactant protein-C (Pro-SP-C) (green) in ATII$^{ER:KRASV12}$ cells. DAPI (blue) was used to stain nuclei. Scale bars: 40 μm. (**b**) Immunofluorescence staining of F-actin (red) in ATII$^{ER:KRASV12}$ cells cultured in the absence or presence of 100 ng/ml EGF for 24 hrs. Rhodamine-phalloidin was used to stain F-actin. DAPI (blue) was used to stain nuclei. Scale bars: 40 μm. (**c**) Fold change in mRNA levels of *CDH1* (E-cadherin), *VIM* (Vimentin), *SNAI1* (Snail1), *SNAI2* (Snail2), *TWIST*, *ZEB1* and *ZEB2* in ATII$^{ER:KRASV12}$ cells cultured in the absence or presence of 100 ng/ml EGF for 24 hrs. GAPDH-normalised mRNA levels in control cells were used to set the baseline value at unity. Data are mean ± s.d. $n = 3$ samples per group. ** $P < 0.01$. *** $P < 0.001$. (**d**) Protein expression of E-cadherin, ZEB1 and phospho-ERK (p-ERK) in ATII$^{ER:KRASV12}$ treated with 100 ng/ml EGF for 8 or 24 hrs. β-actin was used as a loading control. (**e**) Protein expression of E-cadherin, ZEB1, phospho-Smad2 (p-Smad2), phospho-ERK (p-ERK) in primary human ATII cells treated with 100 ng/ml EGF or 5 ng/ml TGFβ over 7 days. β-actin was used as a loading control.

Figure 2 Activation of RAS signalling induces EMT in alveolar epithelial cells. (**a**) Protein expression of E-cadherin, ZEB1 and phospho-ERK (p-ERK) in ATII$^{ER:KRASV12}$ treated with 100 ng/ml EGF or 250 nM 4-OHT for 24 hrs. β-tubulin was used as a loading control. (**b**) Fold change in mRNA levels of *CDH1* (E-cadherin), *VIM* (Vimentin), *SNAI1* (Snail1), *SNAI2* (Snail2), *TWIST*, *ZEB1* and *ZEB2* in ATII$^{ER:KRASV12}$ cells cultured in the absence or presence of 250 nM 4-OHT for 24 hrs. GAPDH-normalised mRNA levels in control cells were used to set the baseline

value at unity. Data are mean ± s.d. n = 3 samples per group. *** $P < 0.001$. (**c**) Protein expression of E-cadherin, ZEB1, ZEB2, phospho-ERK (p-ERK) and phospho-AKT (p-AKT) in ATII$^{ER:KRASV12}$ treated with 250 nM 4-OHT for the indicated period. β-actin was used as a loading control. (**d**) Immunofluorescence staining of ZEB1 (green) and F-actin (red) in ATII$^{ER:KRASV12}$ cells cultured in the absence or presence of 250 nM 4-OHT for 24 hrs. Rhodamine-phalloidin was used to stain F-actin. DAPI (blue) was used to stain nuclei. Scale bars: 40 μm. (**e**) Representative 3D confocal images of ATII$^{ER:KRASV12}$ cells cultured in Matrigel in the absence or presence of 250 nM 4-OHT for 48 hrs. Spheres were stained for F-actin with Rhodamine-phalloidin (red) and DAPI (blue). Scale bars: 40 μm. (**f**) Western blot analysis of lysates from 3D-cultured ATII$^{ER:KRASV12}$ cells in Matrigel treated without or with 250 nM 4-OHT for 48 hrs showing effects on E-cadherin, ZEB1, Vimentin, phospho-ERK (p-ERK), phospho-AKT (p-AKT) and Pro-surfactant protein-C (Pro-SP-C). β-actin was used as a loading control.

Figure 3 Activation of the RAS pathway drives EMT via ERK-ZEB1 in ATII cells. (**a**) Protein expression of E-cadherin, ZEB1, phospho-AKT (p-AKT) and phospho-ERK (p-ERK) in ATII$^{ER:KRASV12}$ treated with 250 nM 4-OHT in absence or presence of inhibitors AKT VIII (10 μM) or U0126 (10 μM) for 24 hrs. DMSO was used as a vehicle control and β-actin was used as a loading control. (**b**) Fold change in mRNA levels of *CDH1* (E-cadherin), *VIM* (Vimentin) and *ZEB1* in ATII$^{ER:KRASV12}$ treated with 250 nM 4-OHT in absence or presence of inhibitors AKT VIII (10 μM) or U0126 (10 μM) for 24 hrs. DMSO was used as a vehicle control. GAPDH-normalised mRNA levels in control cells were used to set the baseline value at unity. Data are mean ± s.d. n = 3 samples per group. ** $P < 0.01$. *** $P < 0.001$. (**c**) Protein

expression of E-cadherin, ZEB1, Snail1 and Snail2 in ATII$^{ER:KRASV12}$ cells transfected with the indicated siRNA followed by treatment of 250 nM 4-OHT for 24 hrs. β-tubulin was used as a loading control. (**d**) Fold change in the mRNA level of E-cadherin in ATII$^{ER:KRASV12}$ cells transfected with the indicated siRNA followed by treatment of 250 nM 4-OHT for 24 hrs. GAPDH-normalised mRNA levels in control cells were used to set the baseline value at unity. Data are mean ± s.d. $n = 3$ samples per group. *** $P < 0.001$.

Figure 4 ZEB1 is highly expressed in IPF fibroblastic foci and epithelial cells of thickened alveoli septae where collagen deposition in the interstitium is also evident. Serial sections of IPF (fibroblastic foci in **a**, and epithelial cells of thickened alveoli septae in **b**) or control lung tissue (**c**) were stained for ZEB1 (left panel), with H&E (middle panel) or Masson's trichrome stain (right panel, collagen shown in blue). $n = 3$. Arrows: representative positive ZEB1 staining. ** a fibroblastic focus. Scale bars: 50 μm.

Figure 5 ATII cells undergoing RAS-induced EMT induce fibroblast activation via paracrine signalling. (**a**) Protein expression of α-SMA and phospho-Smad2 (p-Smad2) in MRC5 lung fibroblasts treated without or with 5 ng/ml TGFβ in the presence of conditioned media (CM) from control or 4-OHT-treated ATII$^{ER:KRASV12}$ cells for 48 hrs. β-actin was used as a loading control. (**b**) Fold change in mRNA levels of *COL1A1*, *COL3A1, FN1* and *ACTA2* in MRC5 lung fibroblasts with indicated treatments. β-actin-normalised mRNA levels in control cells were used to set the baseline value at unity. Data are mean ± s.d. $n = 3$ samples per group. * $P < 0.05$. *** $P < 0.001$. (**c**) Protein expression of α-SMA and phospho-Smad2 (p-

Smad2), and fold change in the mRNA level of *ACTA2* (α-SMA) in primary human lung fibroblasts from IPF (IPFFs) or from normal healthy lung (NHLFs) with indicated treatments. β-tubulin was used as a loading control in Western blots. β-actin-normalised mRNA levels in control cells were used to set the baseline value at unity (indicated above bars). Data are mean ± s.d. $n = 3$ samples per group. * $P < 0.05$. ** $P < 0.01$. *** $P < 0.001$.

Figure 6 ZEB1 is a key regulator of the paracrine signalling between ATII cells and fibroblasts. (**a**) Protein expression of ZEB1, E-cadherin, phospho-ERK (p-ERK) and phospho-AKT (p-AKT) in $ATII^{ER:KRASV12}$ cells with indicated treatments. β-tubulin was used as a loading control. (**b**) Protein expression of α-SMA and phospho-Smad2 (p-Smad2) in MRC5 lung fibroblasts or primary human lung fibroblasts from IPF (IPFFs) with indicated treatments. β-tubulin was used as a loading control.

Figure 7 ZEB1 regulates the expression of tissue plasminogen activator (tPA), which acts as a paracrine regulator of TGFβ-induced fibroblast activation. (**a**) Increased expression of *PLAT* (tPA) in IPF epithelial cells is shown by an online LGEA web portal (https://research.cchmc.org/pbge/lunggens/mainportal.html). (**b**) Quantitative secretome analysis identifies an increased level of tPA in the conditioned media (CM) from 4-OHT-treated $ATII^{ER:KRASV12}$ cells and a representative tPA Western blot of CM from control or 4-OHT-treated $ATII^{ER:KRASV12}$ cells. Data are individual values with mean and s.d. $n = 3$ samples per group. Values were normalised to total fmol of each sample multiplied by 10,000. (**c**) Fold change in mRNA levels of *ZEB1* and *PLAT* (tPA) in $ATII^{ER:KRASV12}$ cells with indicated treatments. β-actin-normalised mRNA levels in control cells were used to set the

baseline value at unity. Data are mean ± s.d. $n = 3$ samples per group. *** $P < 0.001$. (**d**) ChIP assays of ZEB1's ability to bind the *PLAT* (tPA) promoter in $ATII^{ER:KRASV12}$ cells with indicated treatments. The amplified *PLAT* (tPA) promoter region (-547 to -345) contains a ZEB1 binding site at -419. Values represent relative binding in relation to input (2%), normalised against control (1.0). Data are mean ± s.d. $n = 4$ samples per group. *** $P < 0.001$. (**e**) *PLAT* promoter reporter assays in $ATII^{ER:KRASV12}$ cells with indicated treatments. Values represent relative fold of firefly luciferase in relation to *Renilla* luciferase, normalised against control (1.0). Data are mean ± s.d. $n = 3$ samples per group. ** $P < 0.01$. *** $P < 0.001$. (**f**) Protein expression of α-SMA and phospho-Smad2 (p-Smad2) in MRC5 lung fibroblasts with indicated treatments. β-tubulin was used as a loading control.

Figure 8 Diagram summarising a critical role of ZEB1-tPA axis regulated by EGFR-RAS-ERK pathway in the development of lung fibrosis (Details provided in Discussion).

Figure 1

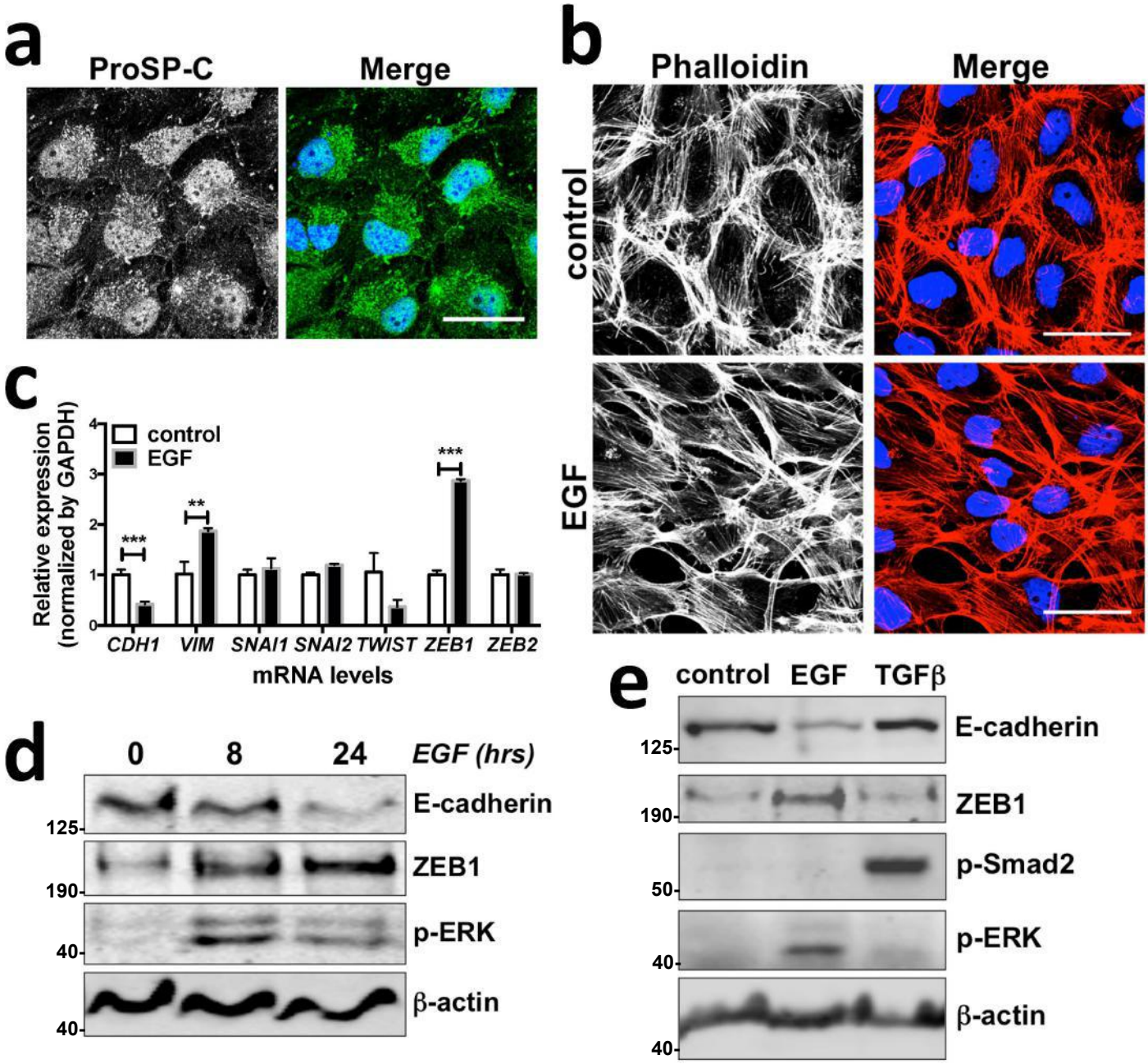

a
ProSP-C
Merge
b
Phalloidin
Merge
control
EGF
c
Relative expression (normalized by GAPDH)
control
EGF
***
**
***
CDH1
VIM
SNAI1
SNAI2
TWIST
ZEB1
ZEB2
mRNA levels
d
0
8
24
EGF (hrs)
125
E-cadherin
190
ZEB1
40
p-ERK
40
β-actin
e
control
EGF
TGFβ
125
E-cadherin
190
ZEB1
50
p-Smad2
40
p-ERK
40
β-actin

Figure 2

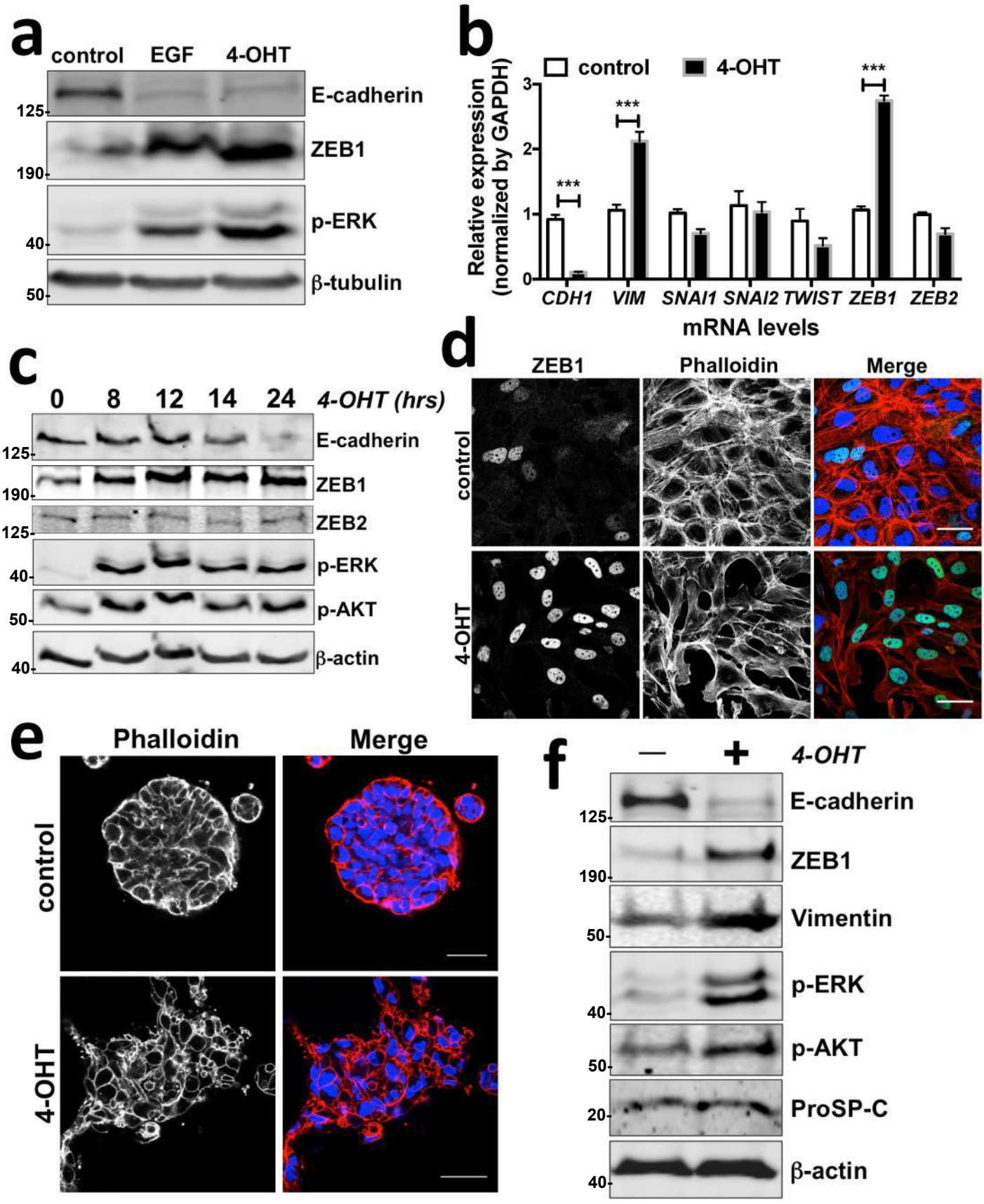

a
control
EGF
4-OHT
125
190
40
50
E-cadherin
ZEB1
p-ERK
β-tubulin
b
control
4-OHT
Relative expression (normalized by GAPDH)
3
2
1
0
***
***
***
CDH1
VIM
SNAI1
SNAI2
TWIST
ZEB1
ZEB2
mRNA levels
c
0
8
12
14
24
4-OHT (hrs)
125
190
125
40
50
40
E-cadherin
ZEB1
ZEB2
p-ERK
p-AKT
β-actin
d
ZEB1
Phalloidin
Merge
control
4-OHT
e
Phalloidin
Merge
control
4-OHT
f
—
+
4-OHT
125
190
50
40
50
20
40
E-cadherin
ZEB1
Vimentin
p-ERK
p-AKT
ProSP-C
β-actin

Figure 3

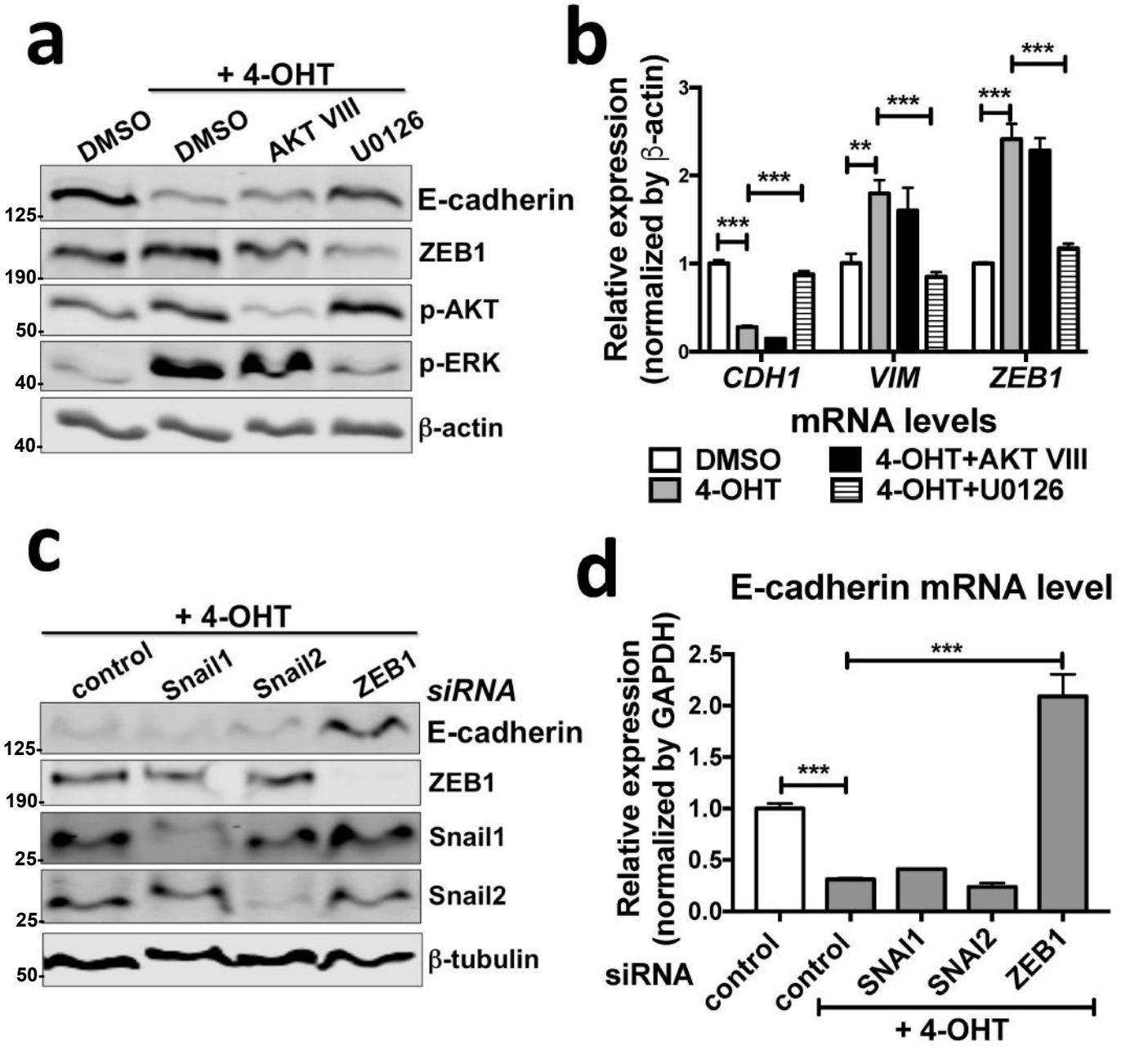

a
+ 4-OHT
DMSO
DMSO
AKT VIII
U0126
125
190
50
40
40
E-cadherin
ZEB1
p-AKT
p-ERK
β-actin
b
Relative expression
(normalized by β-actin)
3
2
1
0
***
***
**
***
***
***
CDH1
VIM
ZEB1
mRNA levels
DMSO
4-OHT
4-OHT+AKT VIII
4-OHT+U0126
c
+ 4-OHT
control
Snail1
Snail2
ZEB1
siRNA
125
190
25
25
50
E-cadherin
ZEB1
Snail1
Snail2
β-tubulin
d
E-cadherin mRNA level
Relative expression
(normalized by GAPDH)
2.5
2.0
1.5
1.0
0.5
0.0
***
***
siRNA
control
control
SNAI1
SNAI2
ZEB1
+ 4-OHT

Figure 4

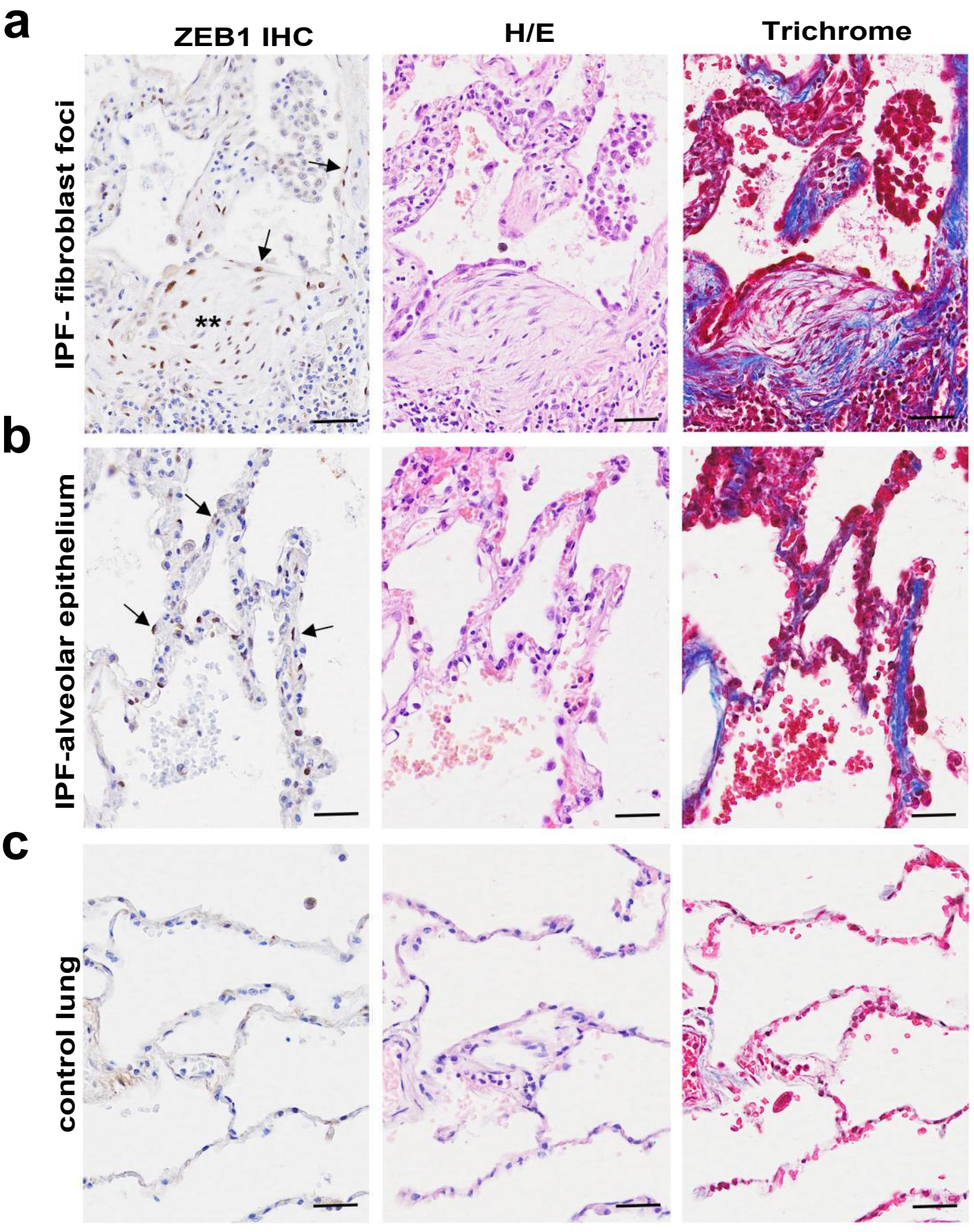

a
ZEB1 IHC
H/E
Trichrome
IPF- fibroblast foci
**
b
IPF-alveolar epithelium
c
control lung

Figure 5

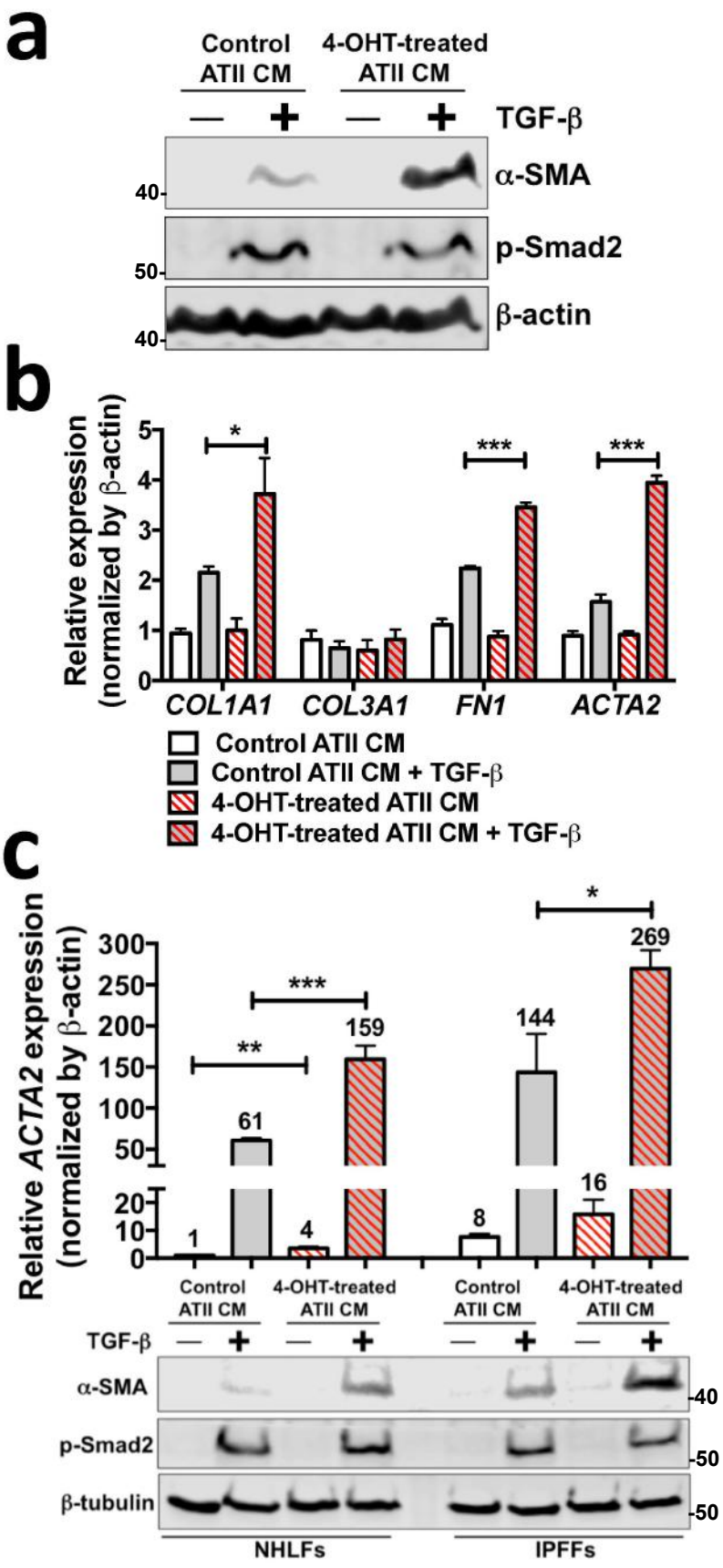

a
Control ATII CM
4-OHT-treated ATII CM
TGF-β
α-SMA
p-Smad2
β-actin
40
50
40
b
Relative expression (normalized by β-actin)
COL1A1
COL3A1
FN1
ACTA2
Control ATII CM
Control ATII CM + TGF-β
4-OHT-treated ATII CM
4-OHT-treated ATII CM + TGF-β
c
Relative ACTA2 expression (normalized by β-actin)
1
61
4
159
8
144
16
269
TGF-β
α-SMA
p-Smad2
β-tubulin
NHLFs
IPFFs

Figure 6

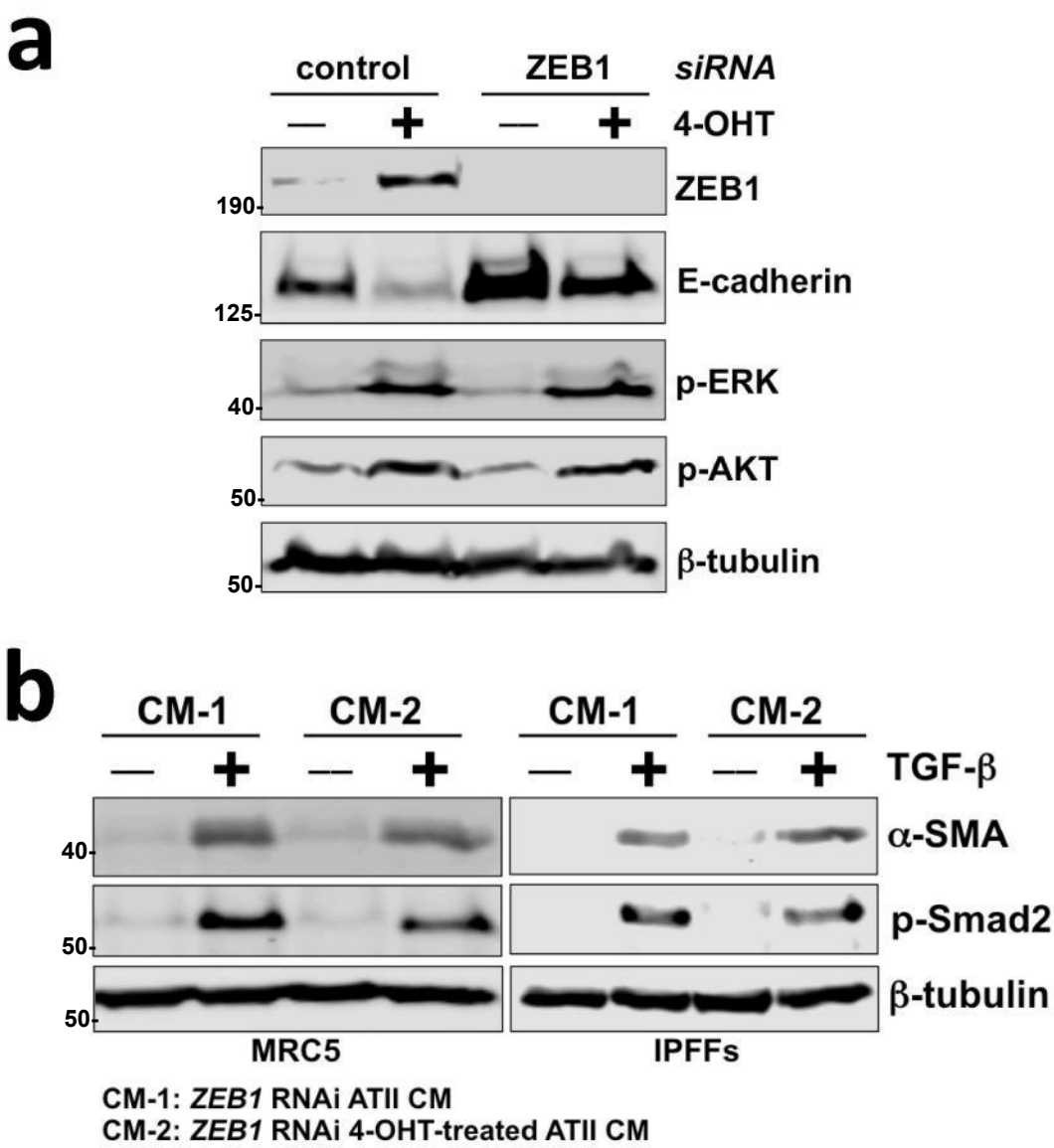

a
control
ZEB1
siRNA
— + — +
4-OHT
190
ZEB1
125
E-cadherin
40
p-ERK
50
p-AKT
50
β-tubulin
b
CM-1
CM-2
CM-1
CM-2
— + — + — + — +
TGF-β
40
α-SMA
50
p-Smad2
50
β-tubulin
MRC5
IPFFs
CM-1: ZEB1 RNAi ATII CM
CM-2: ZEB1 RNAi 4-OHT-treated ATII CM

Figure 7

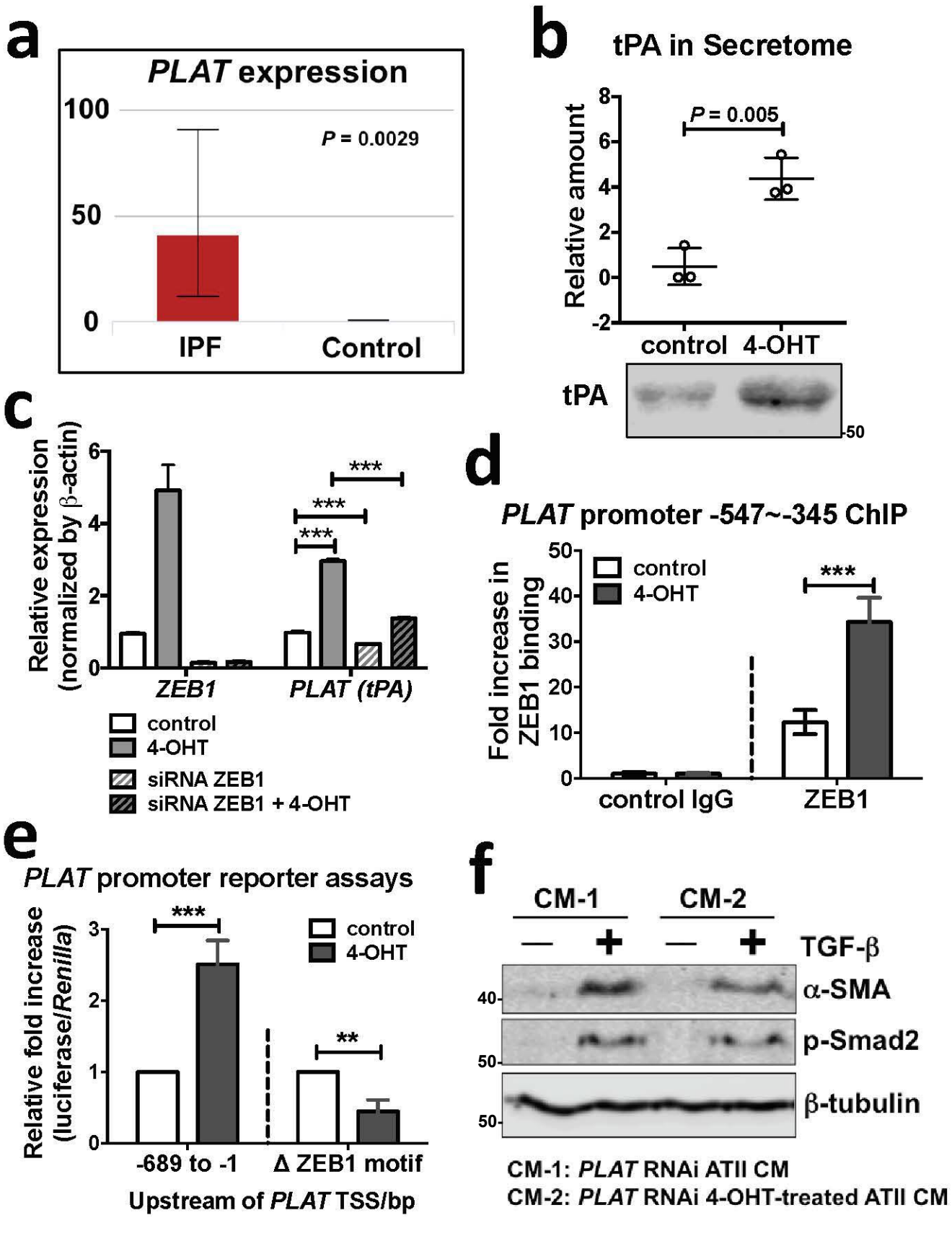
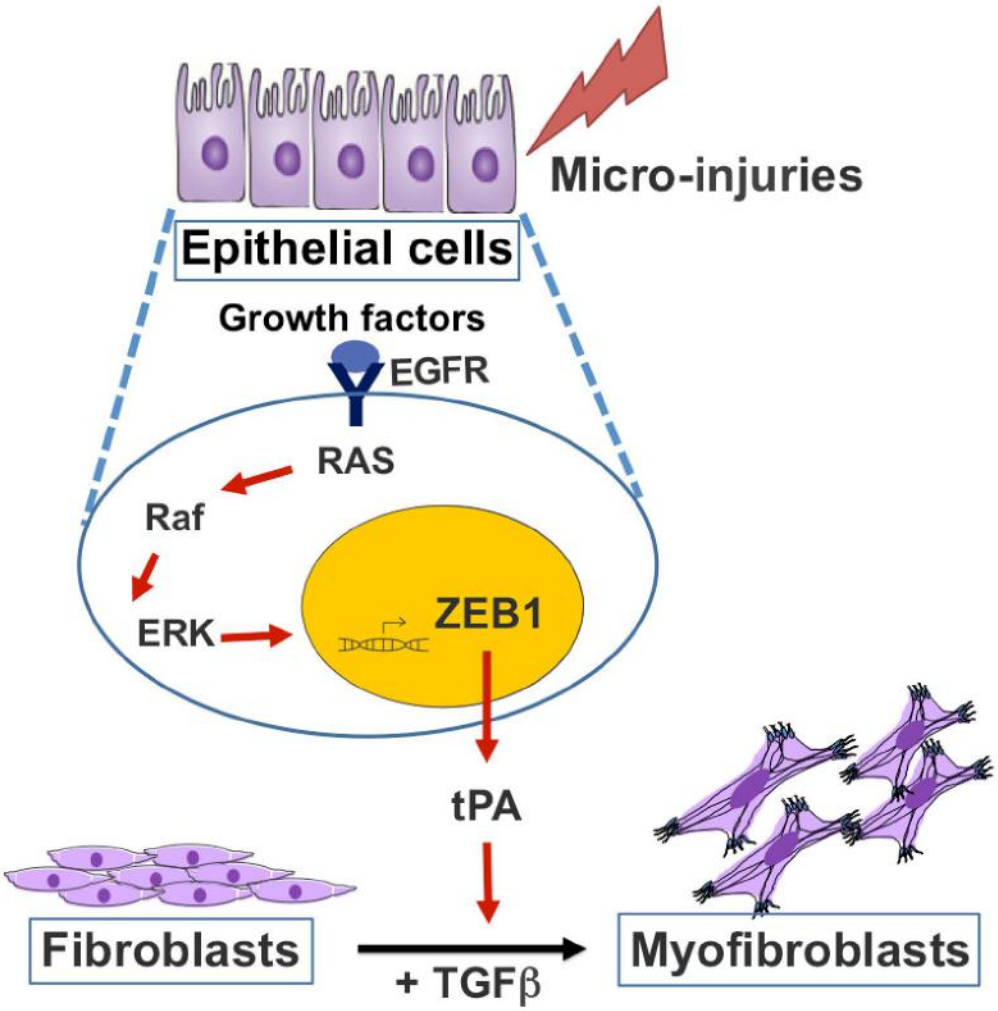

a
PLAT expression
100
50
0
P = 0.0029
IPF
Control
b
tPA in Secretome
Relative amount
8
6
4
2
0
-2
P = 0.005
control
4-OHT
tPA
50
c
Relative expression (normalized by β-actin)
6
4
2
0
***
***
***
ZEB1
PLAT (tPA)
control
4-OHT
siRNA ZEB1
siRNA ZEB1 + 4-OHT
d
PLAT promoter -547~-345 ChIP
Fold increase in ZEB1 binding
50
40
30
20
10
0
control
4-OHT
***
control IgG
ZEB1
e
PLAT promoter reporter assays
Relative fold increase (luciferase/Renilla)
3
2
1
0
***
**
control
4-OHT
-689 to -1
Δ ZEB1 motif
Upstream of PLAT TSS/bp
f
CM-1
CM-2
—
+
—
+
TGF-β
40
α-SMA
50
p-Smad2
50
β-tubulin
CM-1: PLAT RNAi ATII CM
CM-2: PLAT RNAi 4-OHT-treated ATII CM

Figure 8


Micro-injuries
Epithelial cells
Growth factors
EGFR
RAS
Raf
ERK
ZEB1
tPA
Fibroblasts
+ TGFβ
Myofibroblasts

Supplementary Methods

Quantitative proteomic analysis of the secretome.

1. Sample preparation. Serum-free conditioned media (CM) were centrifuged at 17,000 x *g* for 1 min and the supernatants were transferred to fresh tubes. StrataClean resin was vortexed and 50 µL of the slurry was added to 1 mL of the CM and incubated for 1 h at 4°C on a rotator. Samples were centrifuged at 17,000 x *g*, for 3 min and the supernatant removed and discarded. The resin/protein pellet was re-suspended in 300 µL of 100 mM ammonium bicarbonate containing 0.1% SDS. Samples were reduced using 0.5 µg of dithiothreitol and incubating for 1 h with mixing using an Eppendorf Thermomixer® set to 60°C at 1,000 RPM. After 1 h, proteins were alkylated by addition of 2.5 µg iodoacetamide and then incubated for 45 min at room temperature in the dark. Samples were subsequently digested using 0.5 µg sequencing grade modified trypsin (1/50 (w/w)) overnight at 37°C. Digested samples were centrifuged at 17,000 x *g* for 3 min and the peptide supernatant was transferred to fresh tubes. The resin was washed with a further 500 µL of 100% acetonitrile, vortexed, centrifuged at 17,000 x *g*, and the supernatant was removed. The supernatants were combined and lyophilised to dryness at 35°C in a vacuum centrifuge (Eppendorf, UK).

Peptides were reconstituted in 150 µL 3% acetonitrile + 0.1% formic acid and vortexed. Samples were acidified to pH < 3.0 with TFA and loaded onto an Empore C18 96-well solid phase extraction plate (3M, Maplewood, MN). Each acidified sample was loaded onto a conditioned C18 reverse-phase Empore Plate, and washed with 20 µL of 0.5% acetic acid. Peptides were eluted from the plate using 40 µL of 80% acetonitrile + 0.5% acetic acid. Samples were lyophilised to dryness at 35°C and

re-suspended in 25 µL of 2% acetonitrile + 0.1% formic acid containing 100 fmol/µL of the internal enolase digestion standard (Waters).

2. LC-MS^E analysis. 1 µL of digested secretome samples were loaded onto a reverse phase trap column (Symmetry C18, 5 µm, 180 µm x 20mm, Waters Corporation, Milford, MA), at a trapping rate of 5 µL/min and washed for 10 min with buffer A prior to the analytical nanoscale LC separation using a C18 reversed phase column (HSS T3, 1.8 µm, 500 mm x 75µm, Waters). The peptides were eluted over a 120 min linear gradient from 1% acetonitrile + 0.1% formic acid to 60 % acetonitrile + 0.1% formic acid, at a flow rate of 300 nL/min. Eluted samples were sprayed directly into a Synapt G2-Si mass spectrometer (Waters Corporation, Wilmslow, UK) operating in the data independent High Definition Mass Spectometry (HDMS^E) mode. Data were acquired from 50 to 2000 *m/z* using alternate low and high collision energy (CE) scans. Low CE was 5 V and elevated CE was ramped from 15 to 40 V. Ion mobility was implemented prior to fragmentation using a wave height of 650 m/s and wave velocity of 40V. The lockmass Glu[1]-Fibrinopeptide B ((M+2H)+2, *m/z* = 785.8426) was infused at a concentration of 100 fmol/µL at a flow rate of 250 nL/min and acquired every 60 s.

3. Database searches. Raw data were processed using a custom package (Regression tester) based upon executable files from ProteinLynx Global Server 3.0 (Waters). The optimal setting for peak detection across the dataset was determined using Threshold inspector (Waters) and these thresholds were applied: low energy = 100 counts; high energy = 30. Database searches were performed using regression tester and searched against the Uniprot human reference database (24/07/2017; 71,599 entries) with added sequence information for the internal standard Enolase. A maximum of two

missed cleavages was allowed for tryptic digestion with a fixed modification for carboxyamidomethylation of cysteine and a variable modification for the oxidation of methionine. Precursor and product ion mass tolerances were calculated automatically during data processing and the false discovery rate (FDR) was set at 4% and then filtered to 1% FDR. Quantity was estimated in absolute amounts (fmol) using the Top 3 method[1]. The ion accounting output files were compiled and summary information was generated from search log files using custom Python scripts. Information contained in ion accounting files was collated into a single .csv document using a custom Python script.

1. Silva JC, Gorenstein MV, Li GZ, Vissers JP, Geromanos SJ. Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Molecular & cellular proteomics : MCP* 2006, **5**(1)**:** 144-156.

Supplementary Figure Legends

Figure S1 Activation of EGFR signalling induces EMT in alveolar epithelial cells. (**a**) Differential gene expression and pathway analysis in IPF and normal lung tissue. Microarray data (GSE24206) was analysed using LIMMA software and differentially expressed genes determined using a Q value cut-off of 0.02. Gene network analysis using the Consensus Pathways Database was used to identify the pathways containing the differentially expressed genes. (**b**) Protein expression of E-cadherin, ZEB1 and phospho-ERK (p-ERK) in $ATII^{ER:KRASV12}$ treated with 100 ng/ml EGF or TGFα for 24 hrs. β-tubulin was used as a loading control. Scores under the bands are relative levels when compared with control (1.0).

Figure S2 Activation of RAS signalling induces EMT in alveolar epithelial cells. Representative phase contrast images of $ATII^{ER:KRASV12}$ cells cultured in Matrigel in the absence or presence of 250 nM 4-OHT for 48 hrs. Scale bars: 50 μm.

Figure S3 Activation of the RAS pathway drives EMT via ERK-ZEB1 in ATII cells. (**a**) Representative phase contrast images of $ATII^{ER:KRASV12}$ cells treated with 250 nM 4-OHT in absence or presence of inhibitors AKT VIII (10 μM) or U0126 (10 μM) for 24 hrs. DMSO was used as a vehicle control. 20 X magnification. (**b**) Representative phase contrast images of $ATII^{ER:KRASV12}$ cells transfected with the indicated siRNA followed by treatment of 250 nM 4-OHT for 24 hrs. 20 X magnification.

Figure S4 ATII cells undergoing RAS-induced EMT induce fibroblasts activation via paracrine signalling. (**a**) Fold change in mRNA levels of *COL1A1*, *COL3A1, FN1* and *ACTA2* in $ATII^{ER:KRASV12}$ or MRC5 cells with indicated treatments. β-actin-normalised mRNA levels in control ATII cells were used to set the baseline value at unity. Data

are mean ± s.d. n = 3 samples per group. (**b**) Fold change in mRNA levels of *COL1A1, FN1* and *ACTA2* in primary human lung fibroblasts from IPF (IPFFs) with indicated treatments. β-actin-normalised mRNA levels in control cells were used to set the baseline value at unity. Data are mean ± s.d. n = 3 samples per group. * $P < 0.05$. ** $P < 0.01$. *** $P < 0.001$.

Figure S5 ZEB1 is a key regulator of the paracrine signalling between ATII cells and fibroblasts. Fold change in mRNA levels of *COL1A1*, *COL3A1, FN1* and *ACTA2* in MRC5 cells with indicated treatments. β-actin-normalised mRNA levels in control cells were used to set the baseline value at unity. Data are mean ± s.d. n = 3 samples per group.

Figure S6 ZEB1 regulates the expression of tissue plasminogen activator (tPA), which acts as a paracrine regulator of TGFβ-induced fibroblast activation. (**a**) Diagram showing the *PLAT* promoter (-689 to -1 upstream of TSS, transcription start site) for cloning into the pGL3 basic construct. The amplified *PLAT* promoter region (-547 to -345) for ChIP, containing a ZEB1 binding site (5'-CANNTG-3') at -419 (highlighted in red, 5'-CAGGTG-3'), is underlined. (**b**) ChIP assays showing the ability of ZEB1 to bind the *PLAT* promoter in $ATII^{ER:KRASV12}$ cells in response to the indicated treatments. qPCR of fragments of the *PLAT* promoter immunoprecipitated in ChIP assays from $ATII^{ER:KRASV12}$ cells with an antibody against ZEB1 and control IgG. Amplified *PLAT* promoter region (-547~-345) contains a ZEB1 binding site at −419. (**c**) Protein expression of α-SMA and phospho-Smad2 (p-Smad2) in MRC5 lung fibroblasts treated with 10^{-8} M recombinant tPA in the presence or absence of 5 ng/ml TGFβ for 48 hrs. β-tubulin was used as a loading control. (**d**) Fold change in the mRNA level of *PLAT* (tPA) in $ATII^{ER:KRASV12}$ cells with indicated treatments. β- actin-normalised

mRNA levels in control cells were used to set the baseline value at unity. Data are mean ± s.d. $n = 3$ samples per group. *** $P < 0.001$.

Figure S7 Damaged ATII cells are a potential source of TGFβ. (**a**) Expression (*y*-axis) of *TGFB1*, *TGFB2* and *TGFB3* in IPF compared with control lung epithelial cells is shown from an online LGEA web portal (https://research.cchmc.org/pbge/lunggens/mainportal.html). (**b**) Expression (*y*-axis) of *SNAI1*, *SNAI2*, *SNAI3*, *TWIST1* and *TWIST2* in IPF *vs.* control epithelial cells is shown from an online LGEA web portal (https://research.cchmc.org/pbge/lunggens/mainportal.html). (**c**) Fold change in mRNA levels of *TGFB1* and *TGFB2* in ATII$^{ER:KRASV12}$ cells with indicated treatments. β-actin-normalised mRNA levels in control cells were used to set the baseline value at unity. Monolayers of ATII$^{ER:KRASV12}$ cells were scrape wounded with 8 or 12 scratches and allowed to repair for 24 hrs. Data are mean ± s.d. $n = 3$ samples per group. ** $P < 0.01$. *** $P < 0.001$.

Supplementary Table S1

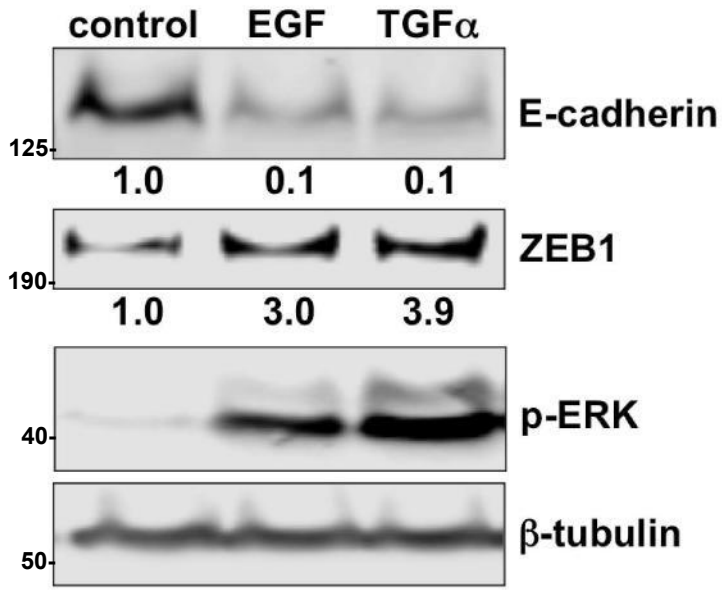
List of proteins/genes that are up-regulated in the CM from 4-OHT-treated ATII$^{ER:KRASV12}$ cells identified by quantitative proteomic analysis and in IPF epithelial cells using an online LGEA web portal (https://research.cchmc.org/pbge/lunggens/mainportal.html).

Supple Figure S1

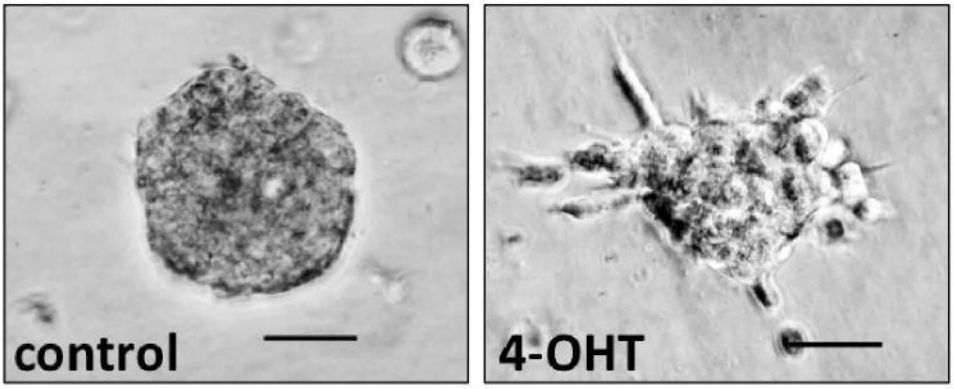
a

P value	Q value	Pathway	Source
6.76×10^{-8}	**2.13×10^{-4}**	**EGFR1**	**NetPath**
1.05×10^{-6}	1.62×10^{-3}	FoxO signaling pathway - Homo sapiens	KEGG
1.54×10^{-6}	1.62×10^{-3}	Amoebiasis - Homo sapiens	KEGG
3.11×10^{-6}	2.45×10^{-3}	Interleukin-6 signaling	Reactome
6.73×10^{-6}	3.94×10^{-3}	Hemostasis	Reactome
7.51×10^{-6}	3.94×10^{-3}	Leptin	NetPath
1.93×10^{-5}	8.69×10^{-3}	Extracellular matrix organization	Reactome
2.41×10^{-5}	9.50×10^{-3}	p53 signaling pathway - Homo sapiens	KEGG
4.55×10^{-5}	1.57×10^{-2}	Leptin signaling pathway	Wikipathways
5.00×10^{-5}	**1.57×10^{-2}**	**MAPK1 (ERK2) activation**	**Reactome**
6.00×10^{-5}	1.72×10^{-2}	Thyroid hormone signaling pathway - Homo sapiens	KEGG
7.23×10^{-5}	1.87×10^{-2}	ECM-receptor interaction - Homo sapiens	KEGG
9.32×10^{-5}	1.87×10^{-2}	Platelet activation - Homo sapiens	KEGG
9.60×10^{-5}	1.87×10^{-2}	Retinoblastoma (RB) in Cancer	Wikipathways
1.81×10^{-4}	1.87×10^{-2}	Alpha6Beta4Integrin	NetPath
1.81×10^{-4}	1.87×10^{-2}	Ectoderm Differentiation	Wikipathways
1.83×10^{-4}	**1.87×10^{-2}**	**EGF-EGFR Signaling Pathway**	**Wikipathways**
2.25×10^{-4}	1.87×10^{-2}	Toxoplasmosis	Wikipathways
2.49×10^{-4}	1.87×10^{-2}	Polycystic Kidney Disease Pathway	Wikipathways

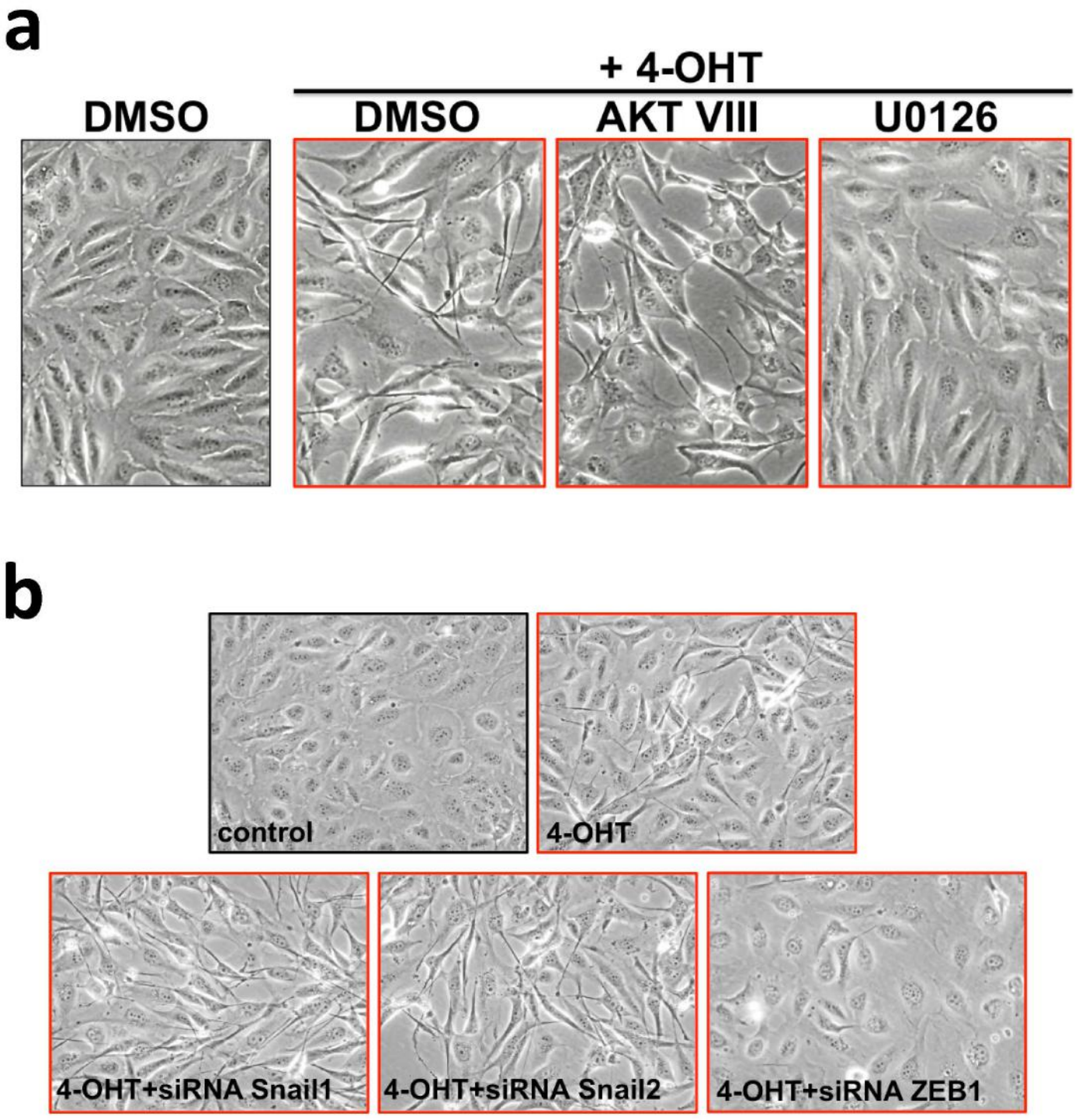
b



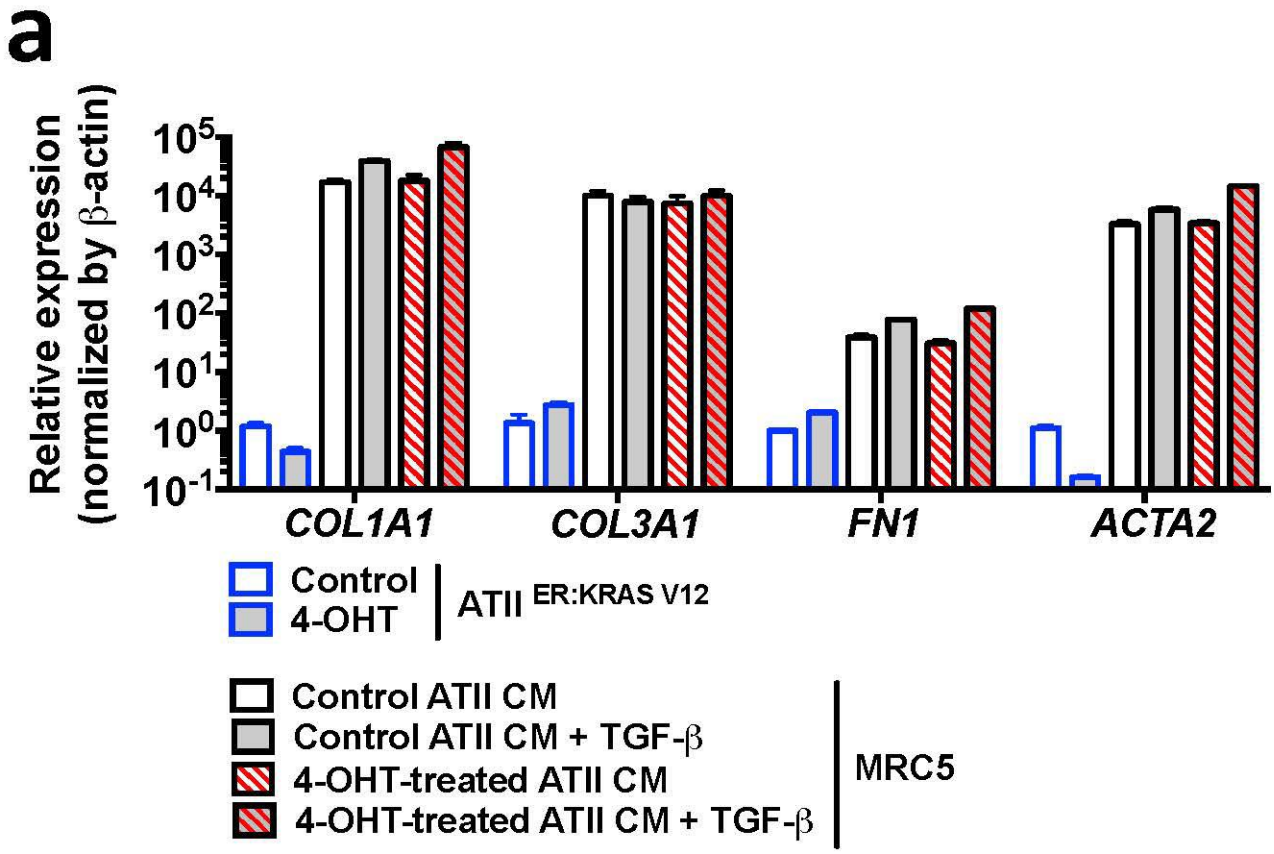
Supple Figure S2

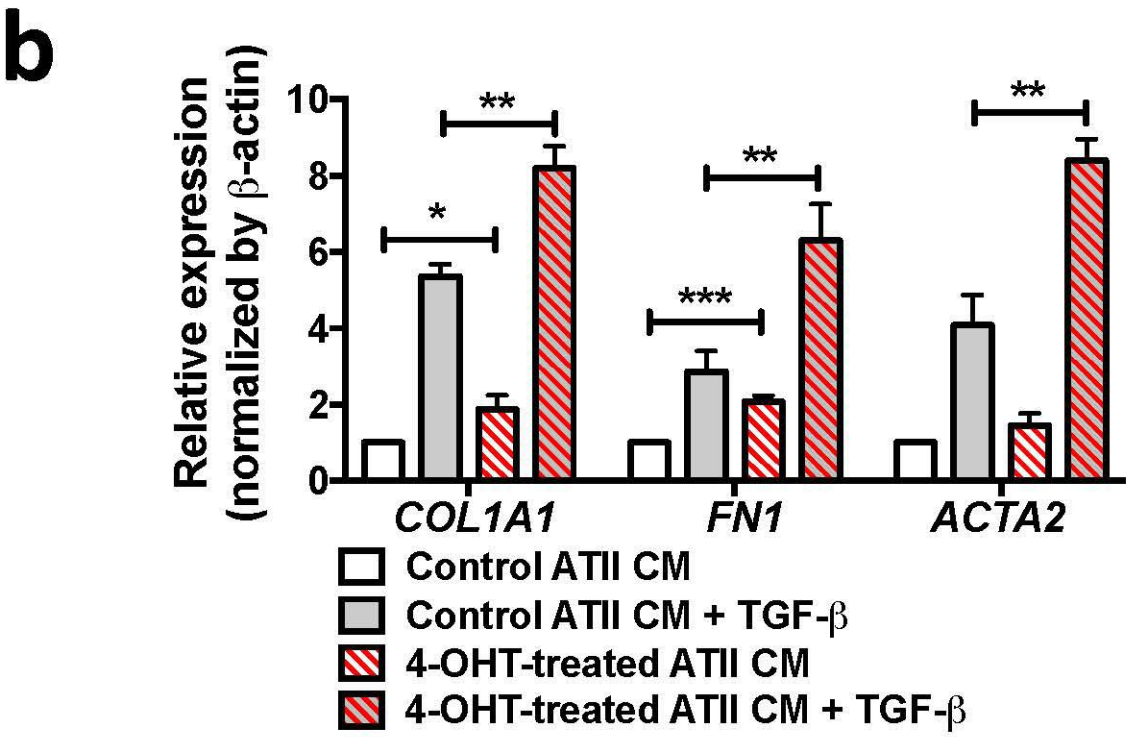


Supple Figure S3

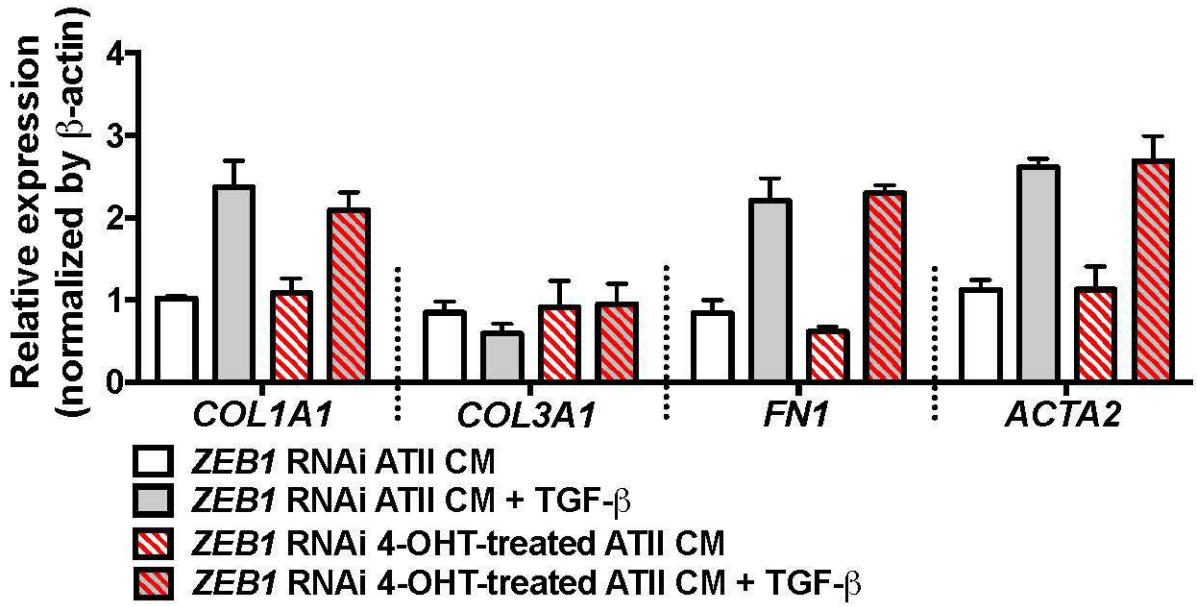


Supple Figure S4





Supple Figure S5



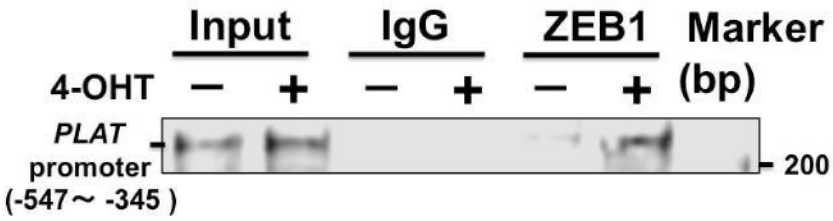
Supple Figure S6

a

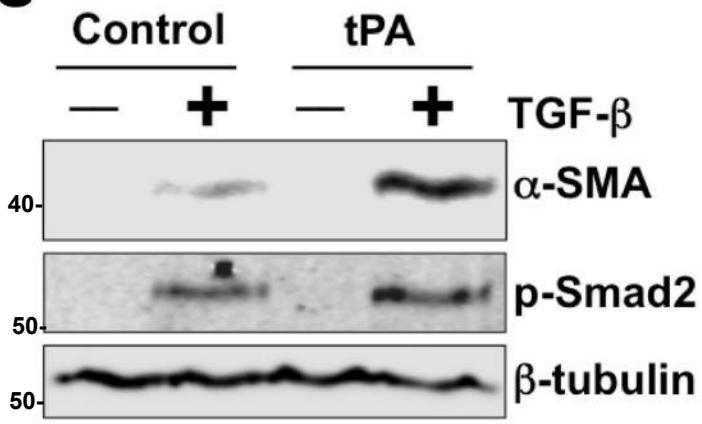
***PLAT* promoter for cloning (-689 to -1 upstream of TSS/bp)**
[ZEB1 binding site E box (5-CANNTG-3)]

5'TCTATTTCCAGGGGCATTTTACAAGCAAATACTGAAAGGCTTCGGTGGGCTTAAGGGCTGATGGCTTTG
ATCGAATTTCAGGCATGTTGGCCCCAAGGCCCTGTGTATATTCCCTGGGCCCACTCAAGGGGATGCTGG
AGCCGGAAAGTCCCCGGAGGCCACCTACTGCAGCCCTGCACTTTACAAAGAAGAGAAAGATTCTCCCTA
AAATTACAGAACAGGGCCAAAGATGCCTACCGGAGCAAACCCCCATGGGGGCACCTCCTACCGCAGGTG
AGCCCAAGGCTGGTCCTGCCTTCTCAGTGGCTACCCCCCTGAGCTCCCGCCACCACACAAAGTGTTCCA
ATCCTTGTGCATCCTCCAGTCCTTTTAACCTCTCATGTCCTGAGAGGCCAGAGCTACAGCCACAGATTCC
AGAAGACACCCCACTCCCAGCCCCAACCTGCTGCCTTTAGAATTATAAACACTTCTTGTCATCACAGGGT
CCTGAAAGTCCCTTTTAAGCCTGGGACACTAGGACTCTAAAGGAAGATGATTCTTAAGGTCCCATCCCAC
TTCCAAATTCCTGCGATTCAATGACATCACGGCTGTGAATAATCAGCCTGGCCCGAAGCCAGGATGGGCT
GTGCTGCTTCCACCGTGAACTTCCTCCCCCTGCTTTATAAAAACAGGCCTGCCTCAGCTCCCTC3'

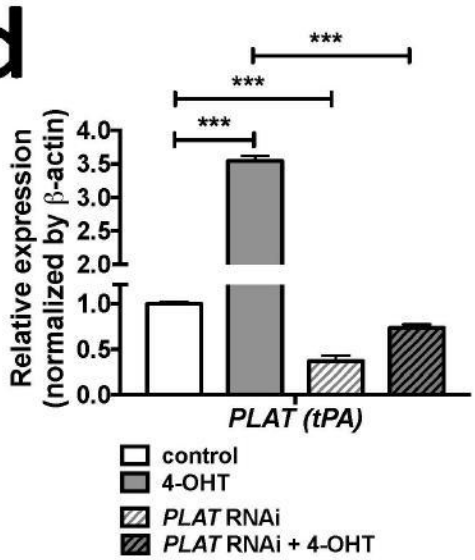
b



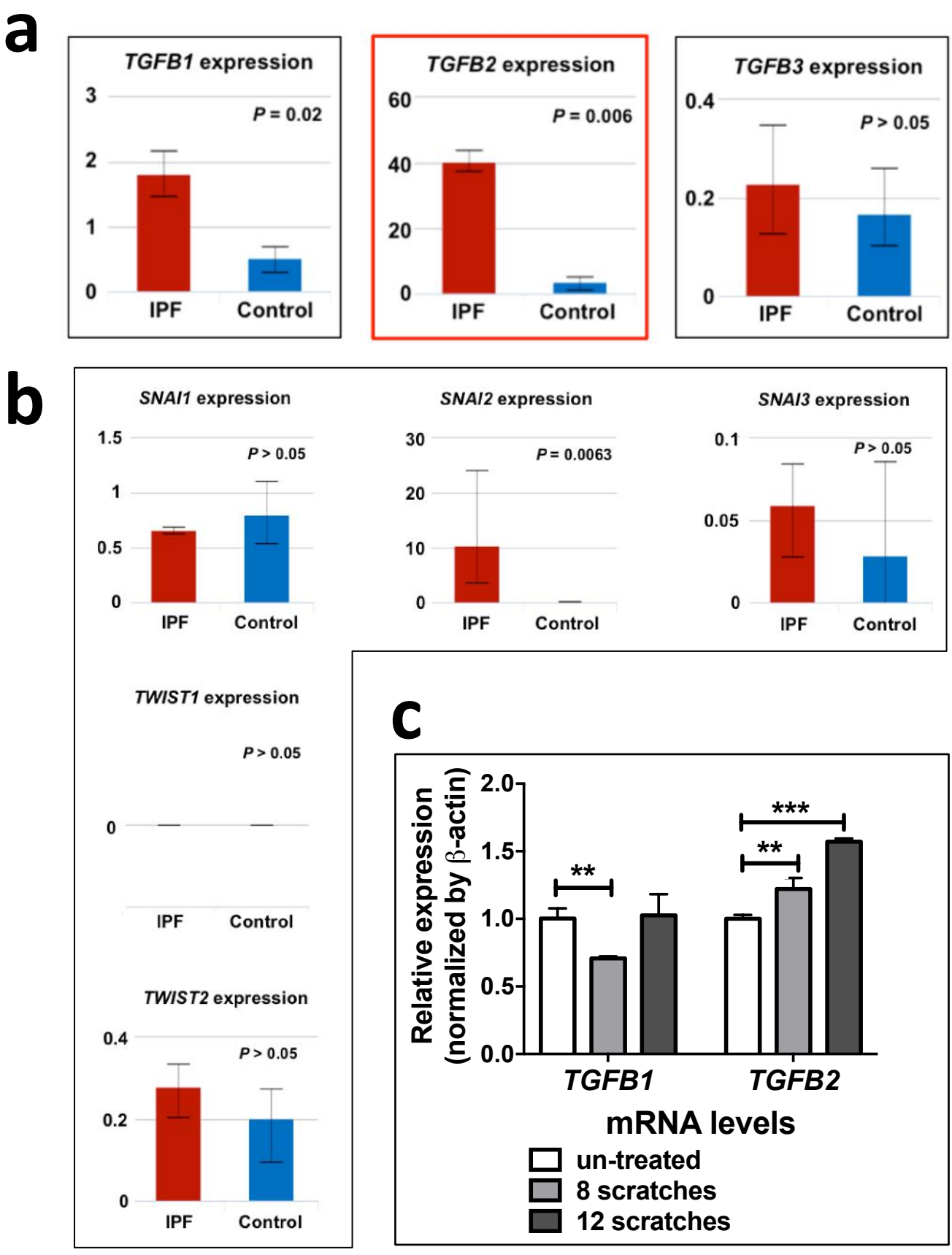
c



d



Supple Figure S7



Supplementary Table S1

Secretome analysis														LGEA analysis	
protein.Entry	protein.Accession	protein.Description	Control 1	Control 2	Control 3	Mean Control	4OHT_1	4OHT_2	4OHT_3	Mean 4OHT	Present_NumFiles	4OHT/Control	*t* test	*P* value	fold change
sp	P00750	TPA_HUMAN Tissue-type plasminogen activator OS=Homo sapiens GN=PLAT PE=1 SV=1	0.03805945	0.01625878	1.42580102	0.493373083	5.43590936	3.76467321	3.91557126	4.37205128	4	8.86155209	0.0054235	0.0029	62.186
sp	O43665	RGS10_HUMAN Regulator of G-protein signaling 10 OS=Homo sapiens GN=RGS10 PE=1 SV=2	0.03805945	0.01625878	0.07417637	0.042831532	0.99216271	0.780106	0.60354859	0.7919391	3	18.48962817	0.0027397	0.0173	31.993
sp	Q9UBP4	DKK3_HUMAN Dickkopf-related protein 3 OS=Homo sapiens GN=DKK3 PE=1 SV=2	5.93056295	5.83114799	4.56300733	5.441572754	10.0672589	9.12225748	6.76727499	8.65226378	6	1.590029974	0.0405159	0.0016	22.862
sp	P09382	LEG1_HUMAN Galectin-1 OS=Homo sapiens GN=LGALS1 PE=1 SV=2	26.24369	24.4859987	25.7932936	25.50766077	32.8702499	34.1803729	33.0724352	33.3743527	6	1.30840507	0.0002942	0.0078	21.273
sp	P49006	MRP_HUMAN MARCKS-related protein OS=Homo sapiens GN=MARCKSL1 PE=1 SV=2	0.03805945	0.01625878	0.07417637	0.042831532	1.22915429	0.81883881	1.03388924	1.02729411	3	23.98452876	0.0011912	0.0448	8.948
sp	Q9BUF5	TBB6_HUMAN Tubulin beta-6 chain OS=Homo sapiens GN=TUBB6 PE=1 SV=1	1.59108512	1.6288924	1.70757031	1.642515942	2.42474405	1.99892431	1.96350613	2.12905816	6	1.296217653	0.0329479	0.0085	7.791
sp	Q01813	PFKAP_HUMAN ATP-dependent 6-phosphofructokinase_ platelet type OS=Homo sapiens GN=PFKP PE=1 SV=2	1.41330745	1.34007068	0.99123851	1.248205545	2.99812274	2.50893872	2.45329649	2.65345265	6	2.12581386	0.0029099	0.0103	4.928
sp	O00154	BACH_HUMAN Cytosolic acyl coenzyme A thioester hydrolase OS=Homo sapiens GN=ACOT7 PE=1 SV=3	5.52472907	4.71961366	5.08920541	5.111182716	7.68256975	8.53283108	7.77277898	7.99605994	6	1.564424592	0.0012627	0.0157	4.724
sp	Q16643	DREB_HUMAN Drebrin OS=Homo sapiens GN=DBN1 PE=1 SV=4	1.28240299	1.27421348	0.07417637	0.876930945	2.36175646	2.03966836	1.9269384	2.10945441	5	2.405496598	0.0432139	0.0027	4.643
sp	P08727	K1C19_HUMAN Keratin_ type I cytoskeletal 19 OS=Homo sapiens GN=KRT19 PE=1 SV=4	16.5180996	12.9025304	16.2798462	15.23349207	23.4078944	23.8995049	26.3436831	24.5503608	6	1.611604266	0.003245	0.0206	4.214
sp	O95865	DDAH2_HUMAN N(G)_N(G)-dimethylarginine dimethylaminohydrolase 2 OS=Homo sapiens GN=DDAH2 PE=1 SV=1	0.03805945	0.59171003	0.07417637	0.234648615	1.22556913	3.90442274	4.6763973	3.26879639	4	13.93060167	0.0459206	0.0237	3.299
sp	P08729	K2C7_HUMAN Keratin_ type II cytoskeletal 7 OS=Homo sapiens GN=KRT7 PE=1 SV=5	4.32725875	4.56625134	3.79641725	4.229975778	7.48748748	8.05830553	8.50840587	8.01806629	6	1.8955348	0.0005285	0.0013	3.216
sp	Q9NVA2	SEP11_HUMAN Septin-11 OS=Homo sapiens GN=SEPT11 PE=1 SV=3	1.57415868	1.13117274	0.97297195	1.226101122	2.2307126	3.16447689	2.8112982	2.7354959	6	2.231052436	0.0098376	0.004	2.985
sp	Q16555	DPYL2_HUMAN Dihydropyrimidinase-related protein 2 OS=Homo sapiens GN=DPYSL2 PE=1 SV=1	2.68078712	2.34254329	2.10650798	2.376612799	3.84755891	3.69629067	4.17248673	3.90544543	6	1.643282169	0.0021756	0.009	2.707
sp	Q96HC4	PDLI5_HUMAN PDZ and LIM domain protein 5 OS=Homo sapiens GN=PDLIM5 PE=1 SV=5	0.03805945	0.01625878	0.07417637	0.042831532	1.42312393	1.51615911	1.08566335	1.34164879	3	31.32385725	0.000596	0.0108	2.637
sp	P14618	KPYM_HUMAN Pyruvate kinase PKM OS=Homo sapiens GN=PKM PE=1 SV=4	34.1364177	34.2956157	41.1210583	36.51769721	44.7813329	41.5010682	46.4058265	44.2294092	6	1.211177391	0.046935	0.018	2.134
sp	Q15942	ZYX_HUMAN Zyxin OS=Homo sapiens GN=ZYX PE=1 SV=1	4.62041663	5.84959897	4.55543831	5.008484634	7.11327073	6.31643212	7.42846302	6.95272195	6	1.388188736	0.0221407	0.0048	2.119
sp	P33316	DUT_HUMAN Deoxyuridine 5'-triphosphate nucleotidohydrolase_ mitochondrial OS=Homo sapiens GN=DUT PE=1 SV=4	0.66413732	0.01625878	1.08337868	0.587924926	3.09795096	1.77269097	2.29507769	2.38857321	5	4.062718048	0.0219907	0.0262	2.049
sp	Q9H299	SH3L3_HUMAN SH3 domain-binding glutamic acid-rich-like protein 3 OS=Homo sapiens GN=SH3BGRL3 PE=1 SV=1	10.6936023	11.6126062	10.6132761	10.97316155	13.5005461	13.5527846	12.8667001	13.306677	6	1.212656616	0.0038884	0.0329	1.971
sp	Q9HA64	KT3K_HUMAN Ketosamine-3-kinase OS=Homo sapiens GN=FN3KRP PE=1 SV=2	0.03805945	0.01625878	0.07417637	0.042831532	0.8801573	1.3146447	1.45076137	1.21518779	3	28.3713361	0.0024689	0.011	1.845
sp	P35237	SPB6_HUMAN Serpin B6 OS=Homo sapiens GN=SERPINB6 PE=1 SV=3	1.84898793	2.04193852	1.8771163	1.922680918	2.79704459	2.88861106	3.64011849	3.10859138	6	1.616800455	0.0123326	0.0399	1.781
sp	P67936	TPM4_HUMAN Tropomyosin alpha-4 chain OS=Homo sapiens GN=TPM4 PE=1 SV=3	6.7261055	6.07247215	7.76056393	6.853047193	9.1039629	8.09742761	8.41492282	8.53877111	6	1.245981659	0.0425782	0.0447	1.718
sp	P15311	EZRI_HUMAN Ezrin OS=Homo sapiens GN=EZR PE=1 SV=4	24.1979949	22.4183019	21.3273715	22.64788942	26.6794785	30.7285466	30.7930708	29.4003653	6	1.298150337	0.0133946	0.0204	1.679
sp	O14907	TX1B3_HUMAN Tax1-binding protein 3 OS=Homo sapiens GN=TAX1BP3 PE=1 SV=2	0.03805945	0.01625878	0.07417637	0.042831532	1.88357001	4.36947595	4.53555723	3.59620106	3	83.96153243	0.0143496	0.0023	1.585
sp	O43707	ACTN4_HUMAN Alpha-actinin-4 OS=Homo sapiens GN=ACTN4 PE=1 SV=2	24.6135439	25.0097507	25.5202034	25.04783267	32.0447354	30.9006059	34.5628056	32.5027156	6	1.297625869	0.0025856	0.0246	1.507