

# PU.1 controls fibroblast polarization and tissue fibrosis

Thomas Wohlfahrt<sup>1</sup>, Simon Rauber<sup>1</sup>, Steffen Uebe<sup>2</sup>, Markus Luber<sup>1</sup>, Alina Soare<sup>1</sup>, Arif Ekici<sup>2</sup>, Stefanie Weber<sup>1</sup>, Alexandru-Emil Matei<sup>1</sup>, Chih-Wei Chen<sup>1</sup>, Christiane Maier<sup>1</sup>, Emmanuel Karouzakis<sup>3</sup>, Hans P. Kiener<sup>4</sup>, Elena Pachera<sup>3</sup>, Clara Dees<sup>1</sup>, Christian Beyer<sup>1</sup>, Christoph Daniel<sup>5</sup>, Kolja Gelse<sup>6</sup>, Andreas E. Kremer<sup>7</sup>, Elisabeth Naschberger<sup>8</sup>, Michael Stürzl<sup>8</sup>, Falk Butter<sup>9</sup>, Michael Sticherling<sup>10</sup>, Susetta Finotto<sup>11</sup>, Alexander Kreuter<sup>12</sup>, Mark H. Kaplan<sup>13</sup>, Astrid Jüngel<sup>3</sup>, Steffen Gay<sup>3</sup>, Stephen L. Nutt<sup>14,15</sup>, David W. Boykin<sup>16</sup>, Gregory M. K. Poon<sup>16</sup>, Oliver Distler<sup>3</sup>, Georg Schett<sup>1</sup>, Jörg H. W. Distler<sup>1</sup> & Andreas Ramming<sup>1\*</sup>

**Fibroblasts are polymorphic cells with pleiotropic roles in organ morphogenesis, tissue homeostasis and immune responses. In fibrotic diseases, fibroblasts synthesize abundant amounts of extracellular matrix, which induces scarring and organ failure. By contrast, a hallmark feature of fibroblasts in arthritis is degradation of the extracellular matrix because of the release of metalloproteinases and degrading enzymes, and subsequent tissue destruction. The mechanisms that drive these functionally opposing pro-fibrotic and pro-inflammatory phenotypes of fibroblasts remain unknown. Here we identify the transcription factor PU.1 as an essential regulator of the pro-fibrotic gene expression program. The interplay between transcriptional and post-transcriptional mechanisms that normally control the expression of PU.1 expression is perturbed in various fibrotic diseases, resulting in the upregulation of PU.1, induction of fibrosis-associated gene sets and a phenotypic switch in extracellular matrix-producing pro-fibrotic fibroblasts. By contrast, pharmacological and genetic inactivation of PU.1 disrupts the fibrotic network and enables reprogramming of fibrotic fibroblasts into resting fibroblasts, leading to regression of fibrosis in several organs.**

Fibroblasts play an important part in the maintenance of tissue integrity<sup>1,2</sup>. They are also critical during the response to tissue injury, which goes far beyond the deposition of extracellular matrix. In the context of inflammatory and neoplastic diseases<sup>2,3</sup>, fibroblasts can differentiate into an extracellular matrix-producing contractile phenotype that promotes progressive accumulation of extracellular matrix and the initiation of fibrotic disease<sup>4–6</sup>. By contrast, in chronic inflammatory diseases, such as rheumatoid arthritis, fibroblasts acquire an extracellular matrix-degrading catabolic phenotype<sup>7</sup>. Phenotypic differences between extracellular matrix-producing pro-fibrotic fibroblasts and catabolic pro-inflammatory fibroblasts also become evident with respect to their distribution in tissues: pro-inflammatory fibroblasts display an imprinted phenotype that forms hypertrophic lining layers in tissues such as the joints and grow in a similar manner to locally invasive tumours. By contrast, pro-fibrotic fibroblasts do not form such lining layers, and instead establish a diffuse arrangement within the connective tissues<sup>8–11</sup>.

Studying the transcriptional network that drives the polarization of fibroblasts into these two functionally opposing phenotypes revealed that PU.1, a transcription factor that belongs to the E26 transformation-specific (ETS) family, is highly expressed in extracellular matrix-producing fibrotic fibroblasts, but is silenced by epigenetic mechanisms in resting and extracellular matrix-degrading inflammatory

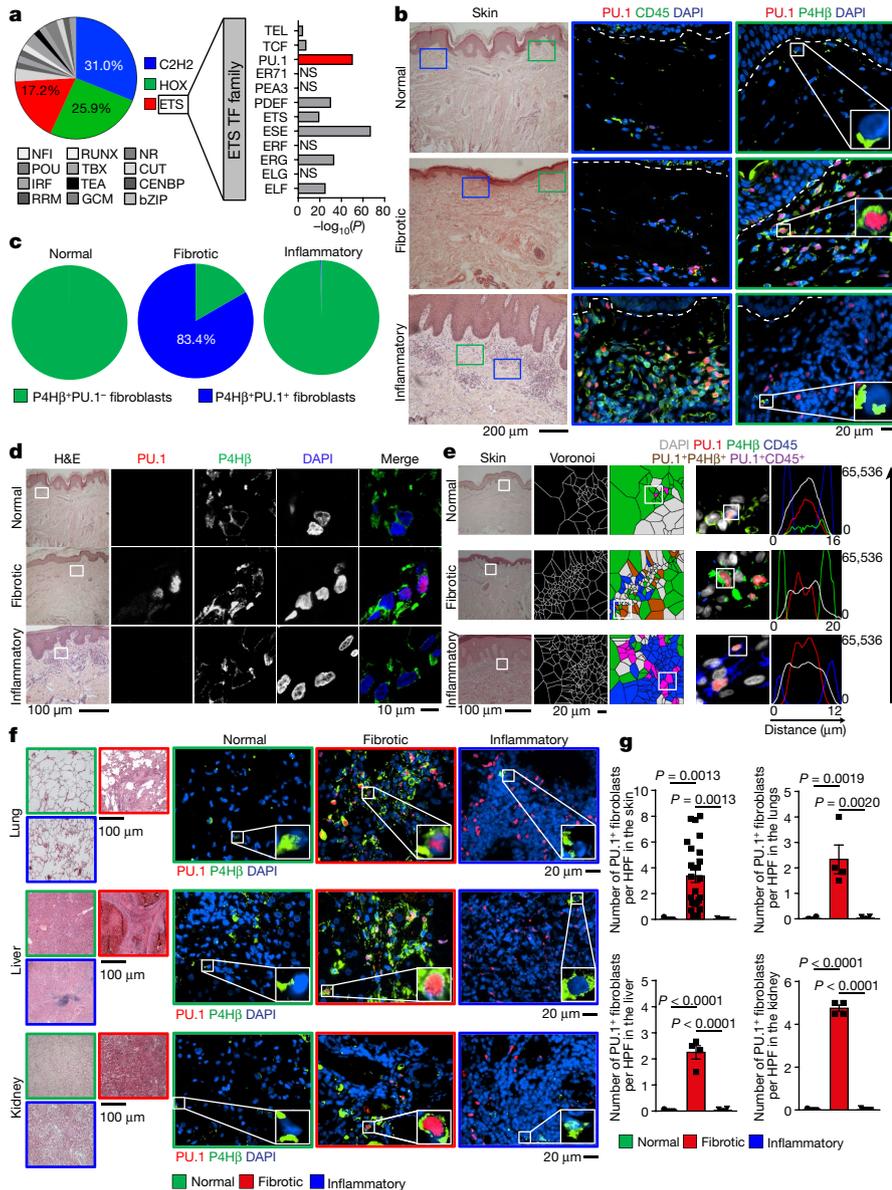
fibroblasts. PU.1 activity therefore acts as a genetic switch that promotes the extracellular matrix-producing fibrotic fibroblast fate.

## Fibrotic fibroblasts express PU.1

To understand the transcriptional network that induces the pro-fibrotic phenotype of fibroblasts, we extracted promoter sequences of differentially expressed genes from a published database of skin biopsy specimens from patients with systemic sclerosis compared to unaffected control subjects<sup>12</sup>. These promoter sequences were searched for the occurrence of motifs of 984 human transcription factors<sup>13</sup> using the HOMER software. For each of the selected motifs, the correlation between the occurrence of the motif in the gene promoter and the respective differential expression level in fibrotic tissues was examined in comparison to healthy samples. Analysis of variance (ANOVA) highlighted 58 transcription factors with an increased presence in the regulatory sequences of pro-fibrotic genes (Fig. 1a). A considerable proportion of those transcription factors belonged to the ETS family (17.2%). Of those ETS members, PU.1 showed the highest enrichment at promoters of pro-fibrotic genes.

PU.1 is a key factor in the differentiation of monocytes and B cells, and deregulation of PU.1 expression has been implicated as a central mechanism in the pathogenesis of leukaemia<sup>14,15</sup>. In fibrotic diseases, however, we observed prominent expression of PU.1 in prolyl

<sup>1</sup>Department of Internal Medicine 3 – Rheumatology and Immunology, Friedrich-Alexander-University (FAU) Erlangen-Nürnberg and Universitätsklinikum Erlangen, Erlangen, Germany. <sup>2</sup>Institute of Human Genetics, Friedrich-Alexander-University (FAU) Erlangen-Nürnberg, Erlangen, Germany. <sup>3</sup>Department of Rheumatology, University Hospital Zurich, Zurich, Switzerland. <sup>4</sup>Division of Rheumatology, Department of Medicine III, Medical University of Vienna, Vienna, Austria. <sup>5</sup>Department of Nephropathology, Friedrich-Alexander-University (FAU) Erlangen-Nürnberg, Erlangen, Germany. <sup>6</sup>Department of Trauma Surgery, Friedrich-Alexander-University (FAU) Erlangen-Nürnberg and Universitätsklinikum Erlangen, Erlangen, Germany. <sup>7</sup>Department of Internal Medicine 1, Friedrich-Alexander-University (FAU) Erlangen-Nürnberg and Universitätsklinikum Erlangen, Erlangen, Germany. <sup>8</sup>Division of Molecular and Experimental Surgery, Department of Surgery, Friedrich-Alexander-University (FAU) Erlangen-Nürnberg, Erlangen, Germany. <sup>9</sup>Quantitative Proteomics Group, Institute of Molecular Biology, Mainz, Germany. <sup>10</sup>Department of Dermatology, Friedrich-Alexander-University (FAU) Erlangen-Nürnberg and Universitätsklinikum Erlangen, Erlangen, Germany. <sup>11</sup>Department of Molecular Pneumology, Friedrich-Alexander-University (FAU) Erlangen-Nürnberg and Universitätsklinikum Erlangen, Erlangen, Germany. <sup>12</sup>Department of Dermatology, Venereology and Allergology, HELIOS St. Elisabeth Klinik Oberhausen, University Witten-Herdecke, Oberhausen, Germany. <sup>13</sup>Herman B. Wells Center for Pediatric Research, Indiana University School of Medicine, Indianapolis, IN, USA. <sup>14</sup>The Walter and Eliza Hall Institute of Medical Research, Molecular Immunology Division, Parkville, Victoria, Australia. <sup>15</sup>Department of Medical Biology, University of Melbourne, Parkville, Victoria, Australia. <sup>16</sup>Department of Chemistry, Georgia State University, Atlanta, GA, USA. \*e-mail: andreas.ramming@uk-erlangen.de



**Fig. 1 | PU.1 expression in leukocytes and fibroblasts from normal human tissues and tissues affected by inflammatory or fibrotic diseases.** **a**, Motif binding analysis of 984 transcription factors (TF) within promoter sequences of differentially expressed genes or gene families in skin biopsy specimens<sup>12</sup> from patients with systemic sclerosis ( $n = 61$ ) compared to unaffected controls ( $n = 36$ ) using HOMER findMotifs. The fold change in expression of differentially expressed genes ( $\log_2(\text{expression ratio})$ ) was calculated and a linear model with the formula  $\log_2(\text{expression ratio}) \propto \text{MotifOccurrences}$  was generated. Transcription factors are shown that had significantly increased motif occurrence ( $-\log_{10}(P)$ ) on pro-fibrotic genes as assessed by ANOVA. **b, d–f**, Representative immunofluorescence images obtained using wide-field (**b, e, f**) and confocal (**d**) microscopy of human skin, lung, liver and kidney biopsy specimens stained for PU.1 (red), CD45 or P4H $\beta$  (green) and with DAPI (blue); tissues were obtained from healthy individuals ( $n = 5$  per tissue type), patients with systemic sclerosis ( $n = 25$ ), plaque psoriasis ( $n = 7$ ), idiopathic pulmonary fibrosis ( $n = 4$ ), acute asthma ( $n = 5$ ), alcoholic liver cirrhosis ( $n = 4$ ), autoimmune hepatitis ( $n = 4$ ), cirrhotic kidney ( $n = 4$ ) and interstitial nephritis ( $n = 5$ ). Haematoxylin and eosin (H&E)-stained tissue specimens are included. **c**, Semi-quantification of PU.1<sup>+</sup> fibroblasts/total P4H $\beta$ <sup>+</sup> fibroblasts per high-power field (HPF). **e**, Voronoi mesh-based tessellated pictures amenable to computational simulation, immunofluorescence microscopy images and histograms of respective immunofluorescence signals. **g**, Semi-quantification of PU.1<sup>+</sup> fibroblasts per HPF. Six randomly chosen HPFs of each slide ( $n$  numbers are indicated in the legend for **b–f**) were used. Data are mean  $\pm$  s.e.m. of biologically independent samples.  $P$  values were determined by one-way ANOVA with Tukey's multiple comparison post hoc test. NS, not significant.

4-hydroxylase (P4H) $\beta$ <sup>+</sup> fibroblasts that lacked the haematopoietic fate markers CD45 and CD11b (Fig. 1b–g and Extended Data Fig. 1a, b). PU.1 was upregulated in fibroblasts of various fibrotic diseases. We also detected PU.1-expressing lymphocytes; however, staining with additional fibroblast markers<sup>16–19</sup> revealed that the majority of PU.1<sup>+</sup> cells in fibrotic tissues were indeed fibroblasts (Fig. 1b–g and Extended Data Fig. 1b–d). In contrast to the abundant expression of PU.1 in fibrotic tissues, PU.1<sup>+</sup> fibroblasts were not found in normal or inflamed tissues of the skin, lung, liver, kidney and joints (Fig. 1b–g and Extended Data Fig. 1a, c, e).

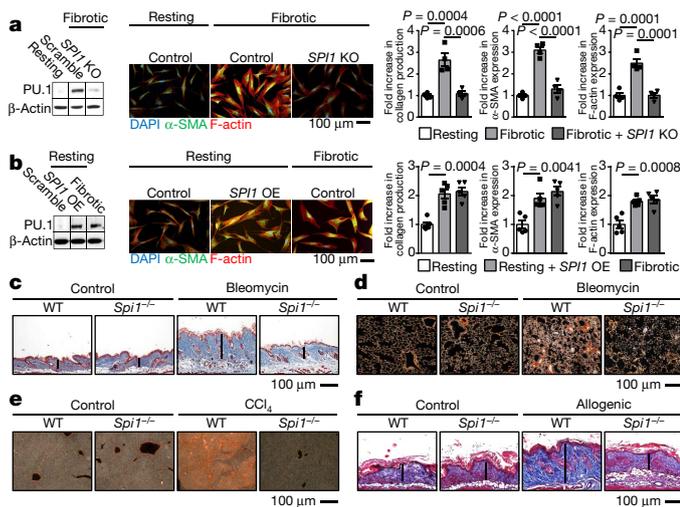
### PU.1 controls tissue fibrosis

To examine the physiological relevance of PU.1 expression in fibroblasts we used CRISPR–Cas9 technology to knock out the gene encoding PU.1 (in human, *SPI1*; in mouse, *Spi1*) in human fibroblasts isolated from fibrotic tissue. *SPI1* knockout fibrotic fibroblasts displayed reduced collagen release, and expression of  $\alpha$ -smooth muscle actin ( $\alpha$ -SMA) and F-actin was similar to the expression levels in resting fibroblasts without affecting cell viability (Fig. 2a and Extended Data Fig. 1f). Conversely, PU.1 overexpression in human resting fibroblasts induced the transition of resting fibroblasts from healthy donors to a highly activated, pro-fibrotic phenotype with upregulation of collagen

release,  $\alpha$ -SMA and F-actin (Fig. 2b and Extended Data Fig. 1g). Next, we addressed the functional effect of PU.1-expressing fibroblasts in several mouse models of fibrosis that resembled different fibrotic conditions across different organs<sup>20</sup>. Similar to humans, PU.1 was expressed in fibroblasts from mouse models of fibrosis, but not in the non-fibrotic controls (Extended Data Figs. 2a–g, 3a–j). Fibroblast-specific knockout of *Spi1* ameliorated fibrosis in these models (Fig. 2c–f, Extended Data Fig. 1h–k).

### Mechanisms that control PU.1 expression

Because PU.1 is required for tissue fibrosis, we next examined the potential mechanisms that could account for its differential expression between fibrotic and inflammatory fibroblasts. Inflammatory stimuli—such as tumour necrosis factor (TNF)—did not influence PU.1 expression in resting, fibrotic or inflammatory fibroblasts (Extended Data Fig. 4a). Similarly, short-term stimulation with pro-fibrotic mediators—such as transforming growth factor (TGF)- $\beta$ —did not convert resting or inflammatory fibroblasts into PU.1-expressing fibroblasts (Extended Data Fig. 4b). Persistent TGF- $\beta$  activity, as found in fibrotic diseases<sup>21–23</sup>, also failed to induce PU.1 in resting or inflammatory fibroblasts (Extended Data Fig. 4c). In fibrotic fibroblasts, however, the basal levels of PU.1 were further upregulated by TGF- $\beta$  in a



**Fig. 2 | Fibrogenic potential of PU.1-expressing fibroblasts.** **a**, CRISPR-Cas9-mediated knock-out (KO) of *SPI1* in human fibrotic fibroblasts.  $n = 4$  per group. **b**, *SPI1*-overexpressing (OE) human resting fibroblasts.  $n = 5$  per group. **a**, **b**, Knockout and overexpression of *SPI1* were measured by western blot analysis. Representative immunofluorescence images of fibroblasts stained for  $\alpha$ -SMA (green), F-actin (red) and with DAPI (blue) are included. Collagen production as well as  $\alpha$ -SMA and F-actin expression were quantified. **c–f**, Representative images of trichrome or Sirius red-stained tissue sections of fibrosis models of wild-type (WT) and *Spi1*<sup>-/-</sup> mice (*Spi1*<sup>fl/fl</sup>; *Col1a2*<sup>creER</sup> mice were used for skin models or *Spi1*<sup>fl/fl</sup>; *Col6*<sup>cre</sup> mice for lung and liver models). **c**, **d**, Bleomycin-induced skin (**c**;  $n = 6$  per group) and lung (**d**;  $n = 6$  per group) fibrosis model. NaCl-treated mice were used as controls. **e**, Carbon tetrachloride (CCl<sub>4</sub>)-induced liver fibrosis model ( $n = 5$  per group). Mice treated with oil were used as controls. **f**, Sclerodermatous chronic graft-versus-host disease model ( $n = 6$  per group). Lines indicate tissue thickness. Data are mean  $\pm$  s.e.m. of biologically independent samples. *P* values were determined by one-way ANOVA with Tukey's multiple comparison post hoc test.

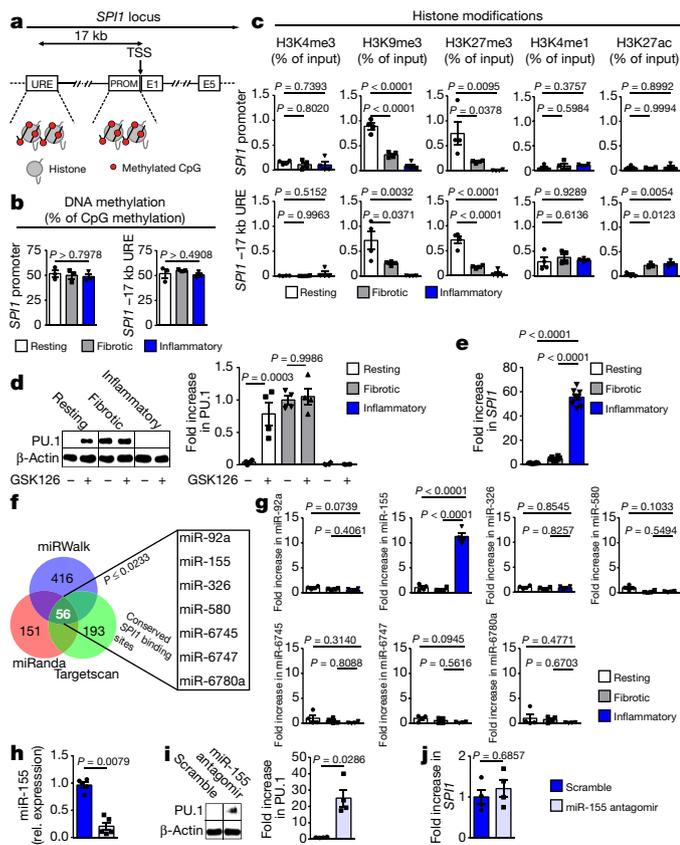
SMAD3-dependent manner (Extended Data Fig. 4d, e). Fibrotic fibroblasts maintained constant levels of PU.1 through several passages, in contrast with normal and inflammatory fibroblasts (Extended Data Fig. 4f).

Because PU.1 expression was maintained in cell culture over multiple passages, we considered whether epigenetic mechanisms had a major role in its regulation<sup>24,25</sup>. Differences in epigenetic programming have previously been related to the development of fibrotic diseases<sup>24,26</sup>. Therefore, we analysed the epigenetic signatures of the *SPI1* locus (Fig. 3a) in human resting, fibrotic and inflammatory fibroblasts. Although DNA methylation has been shown to have a central role in fibroblast activation<sup>27,28</sup>, we did not observe any major differences in DNA methylation at the *SPI1* promoter and enhancer regions among fibroblast phenotypes (Fig. 3b). However, the promoter and the -17 kb upstream regulatory element (URE) of the *SPI1* locus were dominated by the presence of repressive histone 3 lysine 9 trimethylation (H3K9me3) and H3K27me3 in resting fibroblasts. This finding is consistent with increased expression levels of the H3K27 trimethyltransferase enhancer of zeste homologue 2 (EZH2)<sup>29</sup> (Fig. 3c and Extended Data Fig. 4g). Resting fibroblasts showed a poised -17 kb URE (H3K4me1 and H3K27me3), which became active in fibrotic and inflammatory fibroblasts through co-localized H3K27 acetylation (Fig. 3c). Exposure to GSK126<sup>30</sup>, an inhibitor of EZH2 methyltransferase activity, induced expression of PU.1 in resting fibroblasts (Fig. 3d and Extended Data Fig. 4h, i). By contrast, and consistent with the absence of H3K27me3 marks, incubation with GSK126 did not further increase PU.1 expression in fibrotic fibroblasts. In inflammatory fibroblasts, however, the repressive marks H3K9me3 and H3K27me3 were absent, but this was not sufficient for detectable amounts of PU.1 protein (Fig. 3c, d). We observed transcriptional activity at the

*SPI1* locus and detectable *SPI1* mRNA levels (Fig. 3e), suggesting that post-transcriptional regulation might prevent the translation of PU.1 in inflammatory fibroblasts. microRNAs (miRNAs) regulate fibroblast growth and activation<sup>10</sup>. We identified seven potential miRNAs with conserved *SPI1*-binding sites (Fig. 3f) and found that miR-155 was significantly upregulated in inflammatory fibroblasts compared with resting and fibrotic fibroblasts, consistent with previous reports<sup>31,32</sup> (Fig. 3g). As in B cells<sup>33</sup>, inactivation of miR-155 by antagonists (Fig. 3h and Extended Data Fig. 4j, k) induced expression of PU.1 in inflammatory fibroblasts (Fig. 3i), suggesting that *SPI1* in inflammatory fibroblasts was post-transcriptionally regulated by miR-155. However, PU.1 might be degraded by another factor, and protein expression might have been restored because the expression of PU.1 was further increased by inhibition of miR-155. Therefore, we analysed *SPI1* mRNA levels during inhibition of miR-155. We detected stable expression levels of *SPI1* mRNA in miR-155-inhibited cells compared with cells transfected with scrambled antagonists, indicating that it was unlikely that PU.1 was regulated by an independent factor that was affected by miR-155 (Fig. 3j). Together, these findings suggest that two independent mechanisms regulate PU.1 expression in fibroblasts at the level of transcription and translation, determining the functional state of these cells.

### PU.1 induces polarization in fibrotic fibroblasts

Next, we analysed the molecular mechanisms of PU.1-induced fibroblast polarization. Chromatin immunoprecipitation (ChIP) analysis showed that PU.1 binds to the promoters of pro-fibrotic genes, such as *ACTA2* and *COL1A1* (Fig. 4a). Similar to other ETS proteins, PU.1 binds to DNA sites that contain a 5'-GGAA-3' core consensus sequence<sup>34</sup>. However, in contrast to other ETS transcription factors, PU.1 is strongly selective for binding sites in which the 5'-GGAA-3' core is flanked by upstream AT-rich flanking sequences<sup>34</sup>. DB1976<sup>35,36</sup> is a heterocyclic diamidine that competitively blocks PU.1 binding to DNA with minimal effects on other ETS transcription factors because of the high specificity of PU.1 for AT-rich flanking sequences of the 5'-GGAA-3' core (Fig. 4b). DB1976 decreased the transcription of *COL1A1*, reduced the expression of type I collagen and  $\alpha$ -SMA and inhibited the expression of F-actin in fibrotic fibroblasts at least to the levels of resting fibroblasts at non-toxic concentrations (Fig. 4c–e and Extended Data Fig. 4l, m). RNA-sequencing (RNA-seq) analysis and subsequent gene set enrichment analysis (GSEA) demonstrated that incubation with DB1976 inhibited the pro-fibrotic gene signature of fibrotic fibroblasts<sup>37–42</sup> without effects on apoptosis-related and inflammatory Gene Ontology (GO)-defined gene sets (Fig. 4f–i). DB1976 induced a gene expression pattern comparable to that of resting fibroblasts (Fig. 4j). Conversely, GSEA of resting fibroblasts co-transfected with *SPI1* revealed upregulation of the pro-fibrotic gene set and no effects on apoptosis-related, inflammatory or monocyte gene sets (Fig. 4k and Extended Data Fig. 5a). Additional treatment with DB1976 completely blocked the pro-fibrotic effects of PU.1 overexpression (Extended Data Fig. 5b). In three-dimensional full-thickness skin organoids, overexpression of PU.1 in resting fibroblasts increased the expression of collagen and  $\alpha$ -SMA, as well as increasing the thickening of the skin organoid (Fig. 4l). RNA-seq data were validated using an integrative analysis that compared RNA-seq and PU.1 ChIP followed by sequencing (ChIP-seq) data. In total 1,247 genes (8.1% of all expressed genes) were found to be significantly differentially expressed between untreated fibrotic fibroblasts and fibrotic fibroblasts treated with DB1976 ( $q < 0.05$ ). A significant majority of differentially expressed genes ( $n = 989$ ; 79.3%) was associated with a PU.1 ChIP-seq peak. In addition to promoter regions, we also identified a substantial number of PU.1-binding sites in distal regions that were more than 50 kb away from known or predicted transcription start sites, reflecting the ability of PU.1 to control transcription through distal enhancers (Extended Data Fig. 5c). To address the question of whether those PU.1-binding regions were associated with the regulation of the respective genes, several ENCODE datasets comprising DNase-seq and histone



**Fig. 3 | Epigenetic and post-transcriptional regulation of PU.1 in human fibroblasts.** **a**, Schematic diagram of the *SPI1* locus. E1–E5, exons 1–5; PROM, promoter; TSS, transcriptional start site; URE, upstream regulatory element. **b–h**, Ex vivo experiments with primary human resting, fibrotic and inflammatory fibroblasts. **b**, DNA methylation analysis of promoters and UREs ( $n = 3$ ) in fibroblasts. **c**, The occurrence of histone modifications at the *SPI1* promoter and UREs was assessed by ChIP and qPCR relative to input DNA.  $n = 4$  per group. **d**, Representative western blot and semi-quantitative analyses of PU.1 expression in resting, fibrotic and inflammatory fibroblasts in the presence or absence of GSK126 for 96 h.  $n = 4$ . **e**, Quantitative analysis of *SPI1* mRNA levels.  $n = 8$  per group. **f**, Prediction of miRNA binding sites within the *SPI1* mRNA by miRWalk ( $n = 416$  hits), TargetsScan ( $n = 193$  hits) and miRanda ( $n = 151$  hits). The overlap of possible miRNAs from all three tools were further restricted to  $P \leq 0.0233$  predicted by miRWalk<sup>49,50</sup>. **g**, Respective miRNA expression levels relative to expression levels in resting fibroblasts.  $n = 4$ . **h**, The expression of miR-155 was reduced in inflammatory fibroblasts co-transfected with miR-155-specific or scramble antagonomirs as control ( $n = 5$  per group). **i**, Representative western blot and semi-quantitative analyses of PU.1 expression in inflammatory fibroblasts co-transfected with appropriate miR-155-specific or scramble antagonomirs. PU.1 expression is shown relative to  $\beta$ -actin.  $n = 4$ . **j**, *SPI1* mRNA expression levels in inflammatory fibroblasts in the presence and absence of miR-155 antagonomirs.  $n = 4$ . Data are mean  $\pm$  s.e.m. of biologically independent samples.  $P$  values were determined either by one-way ANOVA with Tukey's multiple comparison post hoc test or two-tailed Mann–Whitney  $U$ -test if two groups were compared.

ChIP-seq data were used for the unbiased identification of active regulatory elements within the respective genes. PU.1 ChIP-seq peaks at various  $q$ -value thresholds showed a marked overlap with these predicted active regulatory elements (Extended Data Fig. 5d). These results underpin the regulatory function of PU.1 within fibrotic genes.

### PU.1 switches inflammatory to fibrotic fibroblasts

We next investigated whether forced expression of PU.1 could repolarize inflammatory into fibrotic fibroblasts. Indeed, ectopic PU.1 expression in inflammatory fibroblasts resulted in the upregulation

of fibrosis-associated genes instead of genes for extracellular matrix-degrading proteins and inflammatory mediators (Extended Data Fig. 5e). In three-dimensional micro-mass organoids that resembled the synovial membrane, inflammatory fibroblasts showed a reduced ability to form lining layers after PU.1 overexpression (Extended Data Fig. 5f). Instead, inflammatory fibroblasts that were forced to express PU.1 acquired an extracellular matrix-producing, pro-fibrotic phenotype that showed expression of  $\alpha$ -SMA, increased collagen deposition and thickening of the dermal compartment in full-thickness skin organoids (Extended Data Fig. 5g). Inhibition of miR-155 induced PU.1 expression in inflammatory fibroblasts and ingenuity pathway analysis revealed that the transcriptional network of miR-155 included several inflammatory targets such as NF- $\kappa$ B, in addition to fibrosis-related targets (data not shown). Accordingly, inhibition of miR-155 alone was not sufficient to induce a fibrotic phenotype in inflammatory fibroblasts (Extended Data Fig. 5h). Consistent with the broader effects of miR-155, simultaneous blockade of miR-155 and PU.1 inhibited transcription of pro-fibrotic genes and induced expression of inflammatory mediators and metalloproteinases (Extended Data Fig. 5h). These results highlight that fibrotic and inflammatory mediators are tightly balanced by a complex network of transcriptional and post-transcriptional factors. Within this network, PU.1 inhibition is sufficient to block transcription of fibrotic gene clusters.

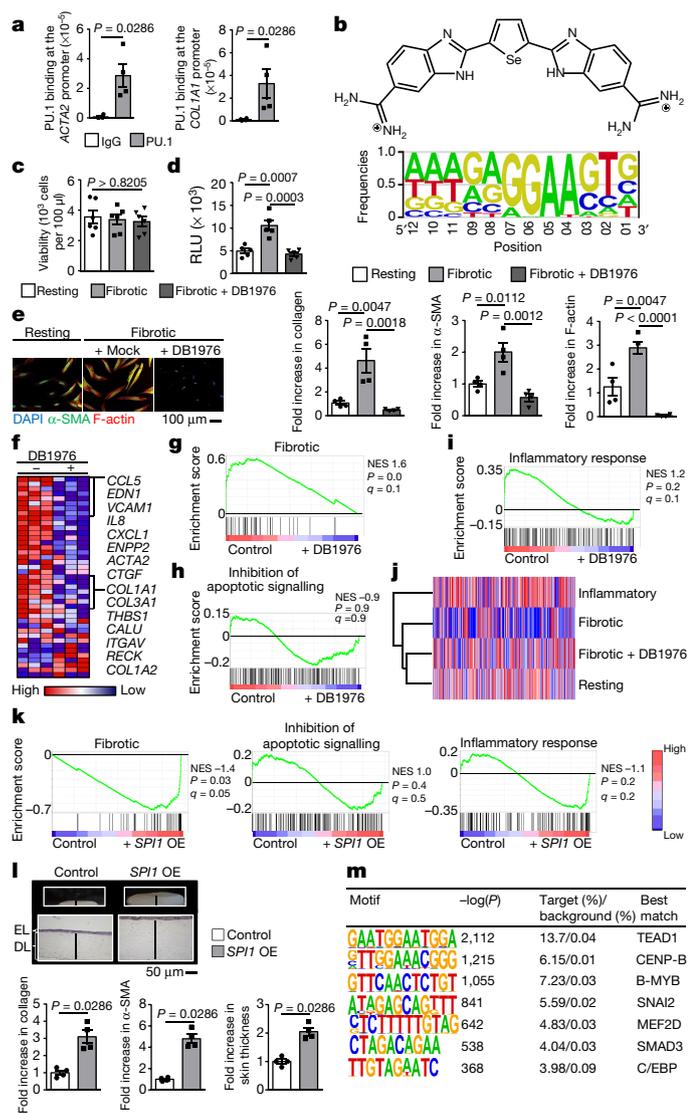
### PU.1 anchors differentiation towards fibrotic fibroblasts

Previously published reports show that PU.1 collaboratively interacts with other transcription factors at closely spaced binding sites to shape the phenotype of a cell<sup>43</sup>. To address whether PU.1 alone is sufficient to induce a fibrotic phenotype, we performed ChIP-seq and additionally investigated transcription factor binding in the vicinity of occupied PU.1 sites. We identified binding motifs for several pro-fibrotic factors, including TEAD1, CENP-B, B-MYB, SNAI2, MEF2D, SMAD3 and C/EBP, in the vicinity of PU.1 peaks. Notably, this represents a different set of factors from those known to collaboratively interact with PU.1 in monocytes and B cells<sup>44</sup> (Fig. 4m). TEAD1 was further validated and showed robust expression levels in resting, fibrotic and inflammatory fibroblasts (Extended Data Fig. 6a). Key pro-fibrotic genes were screened for PU.1 ChIP-seq peaks and potential flanking TEAD1-binding sites. ChIP analysis revealed binding of TEAD1 to predicted regions in key pro-fibrotic genes including *HTR2B*, *ACTA2*, *COL1A1*, *COL1A2*, *CTGF* (also known as *CCN2*), *ITGAV*, *LPAR3*, *PDGFC*, *THBS1* and *TGFB1* (Extended Data Fig. 6b).

To confirm the necessity of the transcriptional network that includes PU.1 to induce a fibrotic phenotype, PU.1-transfected inflammatory fibroblasts were cultured under neutral, fibrotic and inflammatory conditions. As expected, PU.1 induced pro-fibrotic mediators in inflammatory fibroblasts cultured under neutral conditions. However, the expression of key fibrotic factors was substantially facilitated by TGF- $\beta$ -related, fibrotic culture conditions. By contrast, TNF-rich, inflammatory conditions interfered with the fibrotic effects of PU.1 (Extended Data Fig. 6c). These results corroborate the coordinating role of PU.1 as a susceptibility anchor within the network of factors that drives the differentiation towards a fibrotic phenotype.

### Inhibition of PU.1 controls tissue fibrosis

Finally, we investigated pharmacological targeting of PU.1 as a potential strategy to prevent uncontrolled fibrotic tissue remodelling. DB1976 showed anti-fibrotic effects in vivo in various fibrosis models and across several organs. Treatment with DB1976 not only prevented bleomycin-mediated skin fibrosis, but also induced regression of pre-established fibrosis (Extended Data Fig. 7a–d). Treatment with DB1976 in anti-fibrotic concentrations did not affect body weight, pain and distress levels of mice (Extended Data Fig. 7e, f). At the cellular level, we did not detect disturbance of haematopoiesis, alterations in haematopoietic and mesenchymal stem cells, defects in B cell development in the bone marrow or T cell maturation within the thymus after DB1976 treatment (Extended Data Fig. 8a–j).



**Fig. 4 | Regulation of pro-fibrotic genes by PU.1.** **a**, ChIP of PU.1 binding at promoters of *ACTA2* and *COL1A1* analysed by qPCR relative to input DNA.  $n = 4$  each. **b**, Chemical structure of DB1976 and the genomic consensus. **c**, CCK-8 toxicity assay of human fibroblasts stimulated with DB1976 for 96 h.  $n = 5$ . **d**, Luciferase activity of human fibroblasts transfected with a reporter vector that contains the full promoter region of *COL1A1* 24 h after stimulation with or without DB1976.  $n = 5$  each. **e**, Resting and fibrotic fibroblasts treated with or without DB1976 (collagen production, and  $\alpha$ -SMA and F-actin expression per HPF relative to control). Representative immunofluorescence images are shown.  $n = 4$ . **f–i**, RNA-seq of human fibrotic fibroblasts treated with or without DB1976 for 96 h. **f**, Heat map of the pro-fibrotic gene signature. *IL8* is also known as *CXCL8*. **g–i**, GSEA of RNA-seq signals of GO-defined gene clusters. NES, normalized enrichment score. **j**, Hierarchical dendrogram of RNA-seq profiles of resting, inflammatory and fibrotic fibroblasts stimulated with or without DB1976. **k**, GSEA of resting fibroblasts co-transfected with *SPI1* (OE) or scrambled vector ( $n = 4$  each); GO-defined gene signatures were assessed. **l**, Three-dimensional full-thickness skin organoid model of human resting fibroblasts co-transfected with *SPI1* or scrambled vector ( $n = 4$  each); the epidermal layer (EL) and dermal layer (DL) are marked. Collagen content (hydroxyproline assay),  $\alpha$ -SMA expression and skin thickness are quantified per HPF. Data are normalized to control. **m**, Transcription-factor-binding motifs enriched in the 400 bp region surrounding the PU.1-binding sites that are specific for fibrotic fibroblasts are shown after correction for a random-GC genomic background.  $n = 3$  per group. Data are mean  $\pm$  s.e.m. of biologically independent samples.  $P$  values were determined according to previously published methods<sup>51</sup> (**g–i**, **k**) and a previously published study<sup>43</sup> (**m**), by one-way ANOVA with Tukey's multiple comparison post hoc test (**c–e**) or two-tailed Mann–Whitney  $U$ -test (**a**, **l**).

## Discussion

Our data demonstrate that expression of PU.1 is effectively silenced in fibroblasts during tissue homeostasis. When the epigenetic control of PU.1 is lost and PU.1 expression is induced, fibroblasts differentiate into a fibrotic phenotype that includes the transcription of numerous pro-fibrotic mediators. PU.1 has previously been mainly implicated in the regulation of haematopoiesis, for which PU.1 is indispensable during normal myeloid and lymphoid development<sup>14,15</sup> and determines the fate of respective progenitors<sup>45–48</sup>. We found that the majority of PU.1-expressing cells in fibrotic tissues, in multiple disease settings, are of a mesenchymal phenotype. PU.1 polarized resting fibroblasts and even repolarized extracellular matrix-degrading inflammatory fibroblasts to an extracellular matrix-producing fibrotic phenotype.

PU.1 is associated with a network of pro-fibrotic factors including members of the TEAD–HIPPO, canonical TGF- $\beta$ –SMAD and API signalling pathways. Other transcription factors with fibrotic abilities, such as SNAI2 and myocyte enhancer factor (MEF) 2, bind in close vicinity to PU.1-binding sites within the genome and may contribute to the recruitment of the transcription machinery that drives the switch towards the fibrotic phenotype. Motif enrichment does not establish whether or not respective factors are required for collaborative binding. However, we show that the simultaneous induction of TGF- $\beta$ -related mediators facilitates the pro-fibrotic properties of PU.1, whereas TNF-rich, inflammatory settings interfere with PU.1. In line with previous reports in monocytes<sup>43</sup>, these results suggest that the crosstalk between PU.1 and factors that are enriched in the vicinity of PU.1 sites drive fibroblast polarization. Also analogous to its lineage-defining function in monocytes, our results demonstrate that PU.1 has a major coordinating role within this complex network of transcription factors in fibroblasts, as the inactivation of PU.1 alone is sufficient to prevent fibrotic polarization.

These findings also suggest that PU.1 inhibition may represent an effective therapeutic approach to treat a wide range of fibrotic diseases. Inactivation of PU.1 effectively reverted the fibrotic phenotype of fibroblasts to a resting state and induced the regression of tissue fibrosis. Furthermore, the level of PU.1 inhibition necessary to revert the functional phenotype of fibroblasts and alleviate fibrosis appears to be substantially lower than the one necessary for inhibition of haematopoietic cell differentiation. Targeting of PU.1 may thus provide a therapeutic option to not only efficiently but also safely interfere with excessive matrix deposition and enable the restoration of tissue homeostasis in fibrotic diseases.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at <https://doi.org/10.1038/s41586-019-0896-x>.

Received: 6 October 2017; Accepted: 21 December 2018;

Published online: 30 January 2019

- Driskell, R. R. et al. Distinct fibroblast lineages determine dermal architecture in skin development and repair. *Nature* **504**, 277–281 (2013).
- Kalluri, R. & Zeisberg, M. Fibroblasts in cancer. *Nat. Rev. Cancer* **6**, 392–401 (2006).
- Flavell, S. J. et al. Fibroblasts as novel therapeutic targets in chronic inflammation. *Br. J. Pharmacol.* **153**, S241–S246 (2008).
- Palumbo-Zerr, K. et al. Orphan nuclear receptor NR4A1 regulates transforming growth factor- $\beta$  signaling and fibrosis. *Nat. Med.* **21**, 150–158 (2015).
- Ramming, A., Dees, C. & Distler, J. H. From pathogenesis to therapy — perspective on treatment strategies in fibrotic diseases. *Pharmacol. Res.* **100**, 93–100 (2015).
- Chakraborty, D. et al. Activation of STAT3 integrates common profibrotic pathways to promote fibroblast activation and tissue fibrosis. *Nat. Commun.* **8**, 1130 (2017).
- McInnes, I. B. & Schett, G. Cytokines in the pathogenesis of rheumatoid arthritis. *Nat. Rev. Immunol.* **7**, 429–442 (2007).
- Parsonage, G. et al. Global gene expression profiles in fibroblasts from synovial, skin and lymphoid tissue reveals distinct cytokine and chemokine expression patterns. *Thromb. Haemost.* **90**, 688–697 (2003).
- Kiener, H. P. et al. Synovial fibroblasts self-direct multicellular lining architecture and synthetic function in three-dimensional organ culture. *Arthritis Rheumatol.* **62**, 742–752 (2010).

10. Wynn, T. A. & Ramalingam, T. R. Mechanisms of fibrosis: therapeutic translation for fibrotic disease. *Nat. Med.* **18**, 1028–1040 (2012).
11. Frank-Bertonec, M. et al. Epigenetically-driven anatomical diversity of synovial fibroblasts guides joint-specific fibroblast functions. *Nat. Commun.* **8**, 14852 (2017).
12. Assassi, S. et al. Dissecting the heterogeneity of skin gene expression patterns in systemic sclerosis. *Arthritis Rheumatol.* **67**, 3016–3026 (2015).
13. Jolma, A. et al. DNA-binding specificities of human transcription factors. *Cell* **152**, 327–339 (2013).
14. Scott, E. W., Simon, M. C., Anastasi, J. & Singh, H. Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science* **265**, 1573–1577 (1994).
15. McKercher, S. R. et al. Targeted disruption of the PU.1 gene results in multiple hematopoietic abnormalities. *EMBO J.* **15**, 5647–5658 (1996).
16. Scanlan, M. J. et al. Molecular cloning of fibroblast activation protein  $\alpha$ , a member of the serine protease family selectively expressed in stromal fibroblasts of epithelial cancers. *Proc. Natl Acad. Sci. USA* **91**, 5657–5661 (1994).
17. Wu, M. et al. Identification of cadherin 11 as a mediator of dermal fibrosis and possible role in systemic sclerosis. *Arthritis Rheumatol.* **66**, 1010–1021 (2014).
18. Uccelli, A., Moretta, L. & Pistoia, V. Mesenchymal stem cells in health and disease. *Nat. Rev. Immunol.* **8**, 726–736 (2008).
19. Sheikh, H., Yarwood, H., Ashworth, A. & Isacke, C. M. Endo180, an endocytic recycling glycoprotein related to the macrophage mannose receptor is expressed on fibroblasts, endothelial cells and macrophages and functions as a lectin receptor. *J. Cell Sci.* **113**, 1021–1032 (2000).
20. Soare, A., Ramming, A., Avouac, J. & Distler, J. H. W. Updates on animal models of systemic sclerosis. *J. Scleroderma Relat. Disord.* **1**, 266–267 (2016).
21. Gabrielli, A., Avvedimento, E. V. & Krieg, T. Scleroderma. *N. Engl. J. Med.* **360**, 1989–2003 (2009).
22. Sonnylal, S. et al. Postnatal induction of transforming growth factor  $\beta$  signaling in fibroblasts of mice recapitulates clinical, histologic, and biochemical features of scleroderma. *Arthritis Rheumatol.* **56**, 334–344 (2007).
23. Györfi, A. H., Matei, A. E. & Distler, J. H. W. Targeting TGF- $\beta$  signaling for the treatment of fibrosis. *Matrix Biol.* **68–69**, 8–27 (2018).
24. Zeybel, M. et al. Multigenerational epigenetic adaptation of the hepatic wound-healing response. *Nat. Med.* **18**, 1369–1377 (2012).
25. Ramming, A., Druzd, D., Leipe, J., Schulze-Koops, H. & Skapenko, A. Maturation-related histone modifications in the PU.1 promoter regulate Th9-cell development. *Blood* **119**, 4665–4674 (2012).
26. Van De Wiel, M. A. et al. Bayesian analysis of RNA sequencing data by estimating multiple shrinkage priors. *Bioinformatics* **14**, 113–128 (2013).
27. Bechtel, W. et al. Methylation determines fibroblast activation and fibrogenesis in the kidney. *Nat. Med.* **16**, 544–550 (2010).
28. Noda, S. et al. Simultaneous downregulation of KLF5 and Fli1 is a key feature underlying systemic sclerosis. *Nat. Commun.* **5**, 5797 (2014).
29. Cao, R. et al. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298**, 1039–1043 (2002).
30. McCabe, M. T. et al. EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. *Nature* **492**, 108–112 (2012).
31. Christmann, R. B. et al. miR-155 in the progression of lung fibrosis in systemic sclerosis. *Arthritis Res. Ther.* **18**, 155 (2016).
32. Migita, K. et al. TNF- $\alpha$ -induced miR-155 regulates IL-6 signaling in rheumatoid synovial fibroblasts. *BMC Res. Notes* **10**, 403 (2017).
33. Vigorito, E. et al. microRNA-155 regulates the generation of immunoglobulin class-switched plasma cells. *Immunity* **27**, 847–859 (2007).
34. Pham, T. H. et al. Mechanisms of *in vivo* binding site selection of the hematopoietic master transcription factor PU.1. *Nucleic Acids Res.* **41**, 6391–6402 (2013).
35. Stephens, D. C. et al. Pharmacologic efficacy of PU.1 inhibition by heterocyclic dication: a mechanistic analysis. *Nucleic Acids Res.* **44**, 4005–4013 (2016).
36. Munde, M. et al. The unusual monomer recognition of guanine-containing mixed sequence DNA by a dithiophene heterocyclic diamidine. *Biochemistry* **53**, 1218–1227 (2014).
37. Farina, G., Lafyatis, D., Lemaire, R. & Lafyatis, R. A four-gene biomarker predicts skin disease in patients with diffuse cutaneous systemic sclerosis. *Arthritis Rheumatol.* **62**, 580–588 (2010).
38. Sanchez-Lopez, E. et al. Targeting colorectal cancer via its microenvironment by inhibiting IGF-1 receptor–insulin receptor substrate and STAT3 signaling. *Oncogene* **35**, 2634–2644 (2016).
39. Scholten, D. et al. Migration of fibrocytes in fibrogenic liver injury. *Am. J. Pathol.* **179**, 189–198 (2011).
40. Vollmann, E. H. et al. Identification of novel fibrosis modifiers by *in vivo* siRNA silencing. *Mol. Ther. Nucleic Acids* **7**, 314–323 (2017).
41. Henderson, N. C. et al. Targeting of  $\alpha_v$  integrin identifies a core molecular pathway that regulates fibrosis in several organs. *Nat. Med.* **19**, 1617–1624 (2013).
42. Sakai, N. et al. LPA $_1$ -induced cytoskeleton reorganization drives fibrosis through CTGF-dependent fibroblast proliferation. *FASEB J.* **27**, 1830–1846 (2013).
43. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
44. Gosselin, D. et al. Environment drives selection and function of enhancers controlling tissue-specific macrophage identities. *Cell* **159**, 1327–1340 (2014).
45. Bakri, Y. et al. Balance of MafB and PU.1 specifies alternative macrophage or dendritic cell fate. *Blood* **105**, 2707–2716 (2005).
46. Chang, H. C. et al. The transcription factor PU.1 is required for the development of IL-9-producing T cells and allergic inflammation. *Nat. Immunol.* **11**, 527–534 (2010).
47. Anderson, K. L. et al. Transcription factor PU.1 is necessary for development of thymic and myeloid progenitor-derived dendritic cells. *J. Immunol.* **164**, 1855–1861 (2000).
48. Rosenbauer, F. & Tenen, D. G. Transcription factors in myeloid development: balancing differentiation with transformation. *Nat. Rev. Immunol.* **7**, 105–117 (2007).
49. Dweep, H., Sticht, C., Pandey, P. & Gretz, N. miRWalk – database: prediction of possible miRNA binding sites by “walking” the genes of three genomes. *J. Biomed. Inform.* **44**, 839–847 (2011).
50. Dweep, H. & Gretz, N. miRWalk2.0: a comprehensive atlas of microRNA-target interactions. *Nat. Methods* **12**, 697 (2015).
51. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).

**Acknowledgements** We thank M. Pascual, K. Dreißigacker, K.-T. Yang, R. Kleinlein and M. Spiller-Becker for technical assistance; and R. Palmisano and P. Tripal of the Optical Imaging Center Erlangen (OICE) for helping with confocal images. This work was supported by grants RA 2506/3-1, RA 2506/4-1, DI 1537/9-1, DI 1537/11-1, DE 2414/2-1, AK 144/2-1, DFG FOR2438, SCHE 1583/7-1, SPP1468, CRC1181 (A01, C01) and SFB TR221 (B04) of the German Research Foundation, the Bundesministerium für Bildung und Forschung (BMBF; project Metharthros), the Marie Curie project OSTEOIMMUNE, the TEAM project of the European Union (305549) and the IMI funded project RTCure (115142), Else Kröner-Fresenius-Stiftung 2014\_A184, grants A64, D28 and J40 of the IZKF in Erlangen, grant 16-10-05-1-Ramming of the ELAN-Foundation Erlangen, grant 2017.129.1 of the Wilhelm Sander-Stiftung, grant SNF310030\_166259 (to O.D.), NIH (R01 A1057459 and AI095282 to M.H.K.; H1.129063 to G.M.K.P.).

**Reviewer information** Nature thanks T. Graf, M. Kolb, M. T. Longaker, C. Nerlov and the other anonymous reviewer(s) for their contribution to the peer review of this work.

**Author contributions** T.W., S.R., J.H.W.D. and A.R. designed the study; T.W., S.R., S.U., M.L., A.S., A.E., S.W., A.-E.M., C.-W.C., C.M., E.K., H.P.K., E.P., C. Dees, F.B. and A.R. acquired data; T.W., S.R., S.U., M.L., A.S., S.W., A.-E.M., C.-W.C., E.K., H.P.K., C.B., E.N., A.J., S.G., O.D., G.S., J.H.W.D. and A.R. interpreted data; C.B., C. Daniel, K.G., A.E.K., E.N., M. Stürzl, F.B., M. Sticherling, S.F., A.K., M.H.K., A.J., S.L.N., D.W.B., G.M.K.P., O.D., G.S., J.H.W.D. and A.R. provided materials; T.W., S.R., G.S., J.H.W.D. and A.R. prepared the manuscript.

**Competing interests** The authors declare no competing interests.

#### Additional information

**Extended data** is available for this paper at <https://doi.org/10.1038/s41586-019-0896-x>.

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-019-0896-x>.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

**Correspondence and requests for materials** should be addressed to A.R.

**Publisher's note**: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

## METHODS

**Ethical compliance and experimental approaches.** This project complied with all relevant ethical regulations regarding animal research and human studies. Experiments were done in a blinded fashion except when specifically indicated. There were no exclusion criteria for the human and animal experiments. Mice were stratified according to sex and then randomized into the different groups. Cells from human donors were also randomized.

**Patient characteristics.** Skin biopsies were obtained from 25 patients with systemic sclerosis according to the 2013 American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) criteria<sup>52</sup>, 7 patients with plaque psoriasis and 21 age- and sex-matched healthy volunteers. Lung tissue was obtained from 4 patients with idiopathic pulmonary fibrosis, 5 patients with asthma and 5 matched non-inflammatory/non-fibrotic controls. Liver samples were obtained from 4 patients with alcoholic liver cirrhosis<sup>53</sup>, 4 samples from patients with autoimmune hepatitis<sup>54</sup> and 5 matched non-inflammatory/non-fibrotic controls. To investigate fibrotic kidney tissue, we used cirrhotic kidneys from 4 patients with end-stage renal disease after renal transplantation or hydronephrosis. Kidney tissues from 5 patients with interstitial nephritis were used as controls. Normal kidney tissues were obtained from macroscopically normal portions of kidneys that had been surgically excised owing to the presence of a localized neoplasm ( $n = 5$ ). Synovial tissue specimens were obtained from 5 patients with rheumatoid arthritis who fulfilled the 2010 ACR classification criteria for rheumatoid arthritis<sup>55</sup> as well as 5 patients with osteoarthritis. Normal synovium was used as control tissue, which was obtained from surgery specimens of patients with no articular disease ( $n = 4$ ). Written informed consent was obtained from all subjects. The study was approved by the ethical committee of the University of Erlangen-Nürnberg.

**Mice.** Wild-type C57/BL6NRj mice were purchased from Janvier. *Spi1<sup>fl/fl</sup>* mice<sup>46</sup> were bred in-house. To selectively inactivate PU.1 in fibroblasts, *Spi1<sup>fl/fl</sup>* mice were crossed with either *Col1a2<sup>creER</sup>* mice<sup>56</sup> or *Col6<sup>cre</sup>* mice<sup>57</sup> to generate *Spi1<sup>fl/fl</sup>;Col1a2<sup>creER</sup>* or *Spi1<sup>fl/fl</sup>;Col6<sup>cre</sup>* mice, respectively. Cre-mediated recombination was induced by repeated intraperitoneal (i.p.) injections of tamoxifen over 5 days. Control groups were injected with corn oil. *Spi1<sup>GFP</sup>* reporter mice<sup>58</sup> were provided by S.L.N. All mice were bred under specific pathogen-free conditions, and all studies were approved by the animal ethical committee of the government of Unterfranken, Würzburg, Germany. The study has complied with all relevant ethical regulations.

**Cell culture.** Human dermal fibroblasts were isolated from ten patients with systemic sclerosis (fibrotic fibroblasts) and ten age- and sex-matched healthy volunteers (resting fibroblasts). After enzymatic digestion of the skin biopsies with collagenase type II (Merck) for 3 h at 37 °C, the digested tissues were filtered using a 100- $\mu$ m nylon filter and centrifuged at 1,400 r.p.m. for 5 min. The pellet was resuspended in DMEM/F12 medium containing 10% heat-inactivated fetal bovine serum (FBS), 25 mM HEPES, 100 U ml<sup>-1</sup> penicillin, 100  $\mu$ g ml<sup>-1</sup> streptomycin, 2 mM L-glutamine, 2.5  $\mu$ g ml<sup>-1</sup> amphotericin B (Thermo Fisher Scientific) and 0.2 mM ascorbic acid (Sigma-Aldrich). Synovial fibroblasts (inflammatory fibroblasts) were isolated from inflamed joints of 6 patients with rheumatoid arthritis. Fibroblasts were cultured for three passages and then quality checked for a pure fibroblast population before use in experiments. Fibroblasts were negative for CD31, CD45, CD326 (also known as EpCAM) and KRT14 and positive for collagen-1, PDGFR $\alpha$  and vimentin. Representative FACS plots are presented in Extended Data Fig. 9a–c. Resting, fibrotic and inflammatory fibroblasts from passages 3–8 showed homogeneous characteristics with regard to proliferation, migration and invasion capacity (Extended Data Fig. 9d, e) as assessed using a xCELLigence Real-Time Cell Analyzer (RTCA) instrument (ACEA Biosciences). As indicated, fibroblasts were transfected with 0.1  $\mu$ g of either pUNO1 empty vector (control) or pUNO1-hSPI1a (both from InvivoGen) plasmids using the 4D-Nucleofector (Lonza). Gene silencing was achieved using either 3  $\mu$ g control CRISPR–Cas9 plasmid or 3  $\mu$ g of PU.1 CRISPR–Cas9 KO plasmid (h2) (Santa Cruz Biotechnology). miR-155 silencing was performed using the anti-hsa-miR-155-5p miScript miRNA inhibitor or the miScript negative control inhibitor (1,500 ng) (both from QIAGEN). In selective experiments, cells were incubated with recombinant TGF- $\beta$  (5 ng ml<sup>-1</sup>) (PeproTech), recombinant human TNF (10 ng ml<sup>-1</sup>) (ImmunoTools) and/or a combination of one or both of the following: GSK126 (1  $\mu$ M) (Selleck Chemicals) or DB1976 (2.5  $\mu$ M; provided by D.W.B. and G.M.K.P.).

**Cell viability and cytotoxicity assays.** Cell viability of cultured cells was quantified using the Cell Counting Kit (CCK)-8 (Dojindo Molecular Technologies) and an MRX ELISA reader (Dynex Technologies).

**Preparation of micro-mass cultures.** Micro-mass organ culture experiments were performed as described elsewhere<sup>9</sup>. Synovial or dermal fibroblasts were released from the culture dish using TrypLE (Thermo Fisher Scientific) following transfection with plasmids as described above. Cells were resuspended in ice-cold Matrigel Matrix (BD Biosciences) at a density of  $5 \times 10^6$  cells per ml. Then, 40- $\mu$ l droplets of the cell suspension were placed onto non-adherent 12-well culture dishes (Thermo Fisher Scientific). Gelation was allowed for 45 min at 37 °C. Afterwards, the gel

was overlaid with basal culture medium (DMEM, supplemented with penicillin, streptomycin, L-glutamine, nonessential amino acid solution, insulin–transferrin–selenium (Thermo Fisher Scientific), 0.1 mmol l<sup>-1</sup> of ascorbic acid, 10 ng ml<sup>-1</sup> of TNF and 10% heat-inactivated FBS. The floating three-dimensional cultures were maintained for 3 weeks; the medium was changed twice a week.

**Three-dimensional full-thickness skin organoids.** Three-dimensional full-thickness skin organoids were generated by submerging transfected fibroblasts ( $1 \times 10^5$  cells per ml) in neutralization solution (232.5 ml DMEM/F12, 7.5 ml FBS, 7.5 ml 3 M HEPES, 2.5 ml chondroitin sulfate, 10 mg ml<sup>-1</sup> rat tail collagen type 1). This mixture (500  $\mu$ l) was filled in cell-culture inserts with porous membranes (8  $\mu$ m) and a 15-mm diameter providing a growth area of 1.13 cm<sup>2</sup> (Greiner Bio One). The dermal components were cultured for one day in DMEM supplemented with 10% FBS, 5 ng ml<sup>-1</sup> TGF- $\beta$  and 1% penicillin–streptomycin at 37 °C, 5% CO<sub>2</sub> and atmospheric O<sub>2</sub>. To build up full-thickness skin models, the epidermal component was generated by seeding  $5 \times 10^5$  normal human epidermal keratinocytes resuspended in Epilife medium with 1% human keratinocyte growth supplement (E1 medium; Gibco) and with extra 1.44 mM CaCl<sub>2</sub> (denoted as E2 medium) on the apical surface of the dermal components on the following day. After a submersed incubation of the models in E2 medium for 16 h at 37 °C, 5% CO<sub>2</sub> and atmospheric O<sub>2</sub>, the medium was aspirated and the airlift-interface culture was initiated. Full-thickness skin models were cultured in E2 medium supplemented with 0.125 mM L-ascorbic acid 2-phosphate and 10 ng ml<sup>-1</sup> keratinocyte growth factor (Sigma Aldrich) for an additional 5–10 days and the level of the culture medium was adjusted to the meniscus of the skin models<sup>59</sup>.

**Real-time monitoring of cell proliferation, migration and invasion.** The real-time proliferation assay was performed using the xCELLigence RTCA system (ACEA Biosciences) according to the manufacturer's instructions. Cells were seeded at a density of 25% in an E-plate in cell-culture medium (10% FBS) and measured every hour for 7 days. As a negative control, serum-starved (0.1% FBS) medium was used. For assessment of cell migration and invasion, CIM-plates 16 were used according to the manufacturer's instructions. In brief,  $2 \times 10^4$  cells were plated in serum-starved (0.1% FBS) medium in the upper chamber. The lower chambers were filled with cell-culture medium containing 10% FBS or with serum-starved medium as control. For invasion assays, the experimental setup of the migration assay was slightly modified as the upper chambers were loaded with 20  $\mu$ l of a 1:10 dilution of Matrigel to create a three-dimensional biomatrix film in each well before cell loading. Cell status was measured by electrical impedance and the relative change between impedance measured at any time ( $t$ ) and baseline; respective values are displayed as the dimensional parameter 'cell index'. The obtained data were analysed using the xCELLigence RTCA software. Results are presented as a curve over time.

**Reporter assays.** Human fibroblasts were transfected with a *COL1A1* luciferase reporter plasmid (Active Motif) using the 4D-Nucleofector. Luciferase assays were performed using the Renilla Luciferase Assay System according to the manufacturer's instructions (Progema). Relative light units were obtained with a Luminoskan Ascent instrument with automated well-wise injection (Thermo Fisher Scientific). Relative light units were normalized to the protein concentration, as determined by a Bradford protein assay according to the manufacturer's protocol (Bio-Rad).

**Quantification of collagen proteins.** The amount of soluble collagen in cell-culture supernatants was quantified using the SirCol collagen assay (Biorcolor). The total collagen content of tissue samples was determined by hydroxyproline assays<sup>60</sup>.

**Histological analysis.** Formalin-fixed, paraffin-embedded skin sections (2–5  $\mu$ m) were deparaffinized and stained with haematoxylin and eosin, Sirius red or trichrome. Dermal thickness was analysed at four different sites in each mouse in a blinded manner. For direct visualization of collagen fibres, Sirius red staining was performed (Sigma-Aldrich). For evaluation of lung tissue, the Ashcroft score was used as described elsewhere<sup>61</sup>. Liver cirrhosis evaluation (Scheuer score) was performed as described elsewhere<sup>62</sup>.

**Fluorescence imaging.** Epitopes were retrieved from deparaffinized sections using a heat-induced method. In brief, sections were alternately bathed in boiling sodium citrate buffer (10 mM sodium citrate, pH 6.0) or Tris-EDTA buffer (10 mM Tris base, 1 mM EDTA, 0.05% Tween 20, pH 9.0). Each bathing step was repeated five times for 2 min following a washing step in distilled water for 5 min.

For cryo-sections, tissues were placed in OCT compound (Tissue-Tek) (Newcomer Supply) and then snap-frozen in liquid nitrogen and cut to 7- $\mu$ m slices. Sections were washed in distilled water after thawing and fixed in 4% PBS-buffered formaldehyde for 10 min following another washing step in PBS.

Next, sections were blocked for 1 h in PBS supplemented with 5% BSA and 2% horse serum. Primary antibodies were incubated overnight at 4 °C, secondary antibodies and DAPI after an intense washing step for 2 h at ambient temperature. Consecutive staining was performed to minimize cross-reactivity. Cross-reactivity was blocked by pre-incubation with species-specific immunoglobulins. The following antibodies were used:  $\alpha$ -SMA (1:500, clone 1A4, Sigma-Aldrich), cadherin-11 (1:100, polyclonal, LS-B2308, LS-Bio), CD45R/B220 (1:500, clone

RA3-6B2, eBioscience), collagen I (1:200, clone COL1, Abcam), collagen I (1:500, polyclonal, ab21286, Abcam), CD11c (1:100, clone N418, Abcam), CD45 (1:500, polyclonal, ab10558, Abcam), F4/80 (1:200, polyclonal, ab100790, Abcam), F4/80 (1:100, clone CI:A3-1, Abcam), fibroblast activation protein (1:5,000, polyclonal, ab28244, Abcam), MRC2 (1:1,000, polyclonal, ab70132, Abcam), vimentin (1:500, clone VI-10, Abcam), CD3ε (1:100, clone 145-2C11), CD45-BV421 (1:500, clone 30-F11), CD49f-APC (1:1,000, clone GoH3), CD117-APC (1:1,000, clone 2B8), EpCAM-APC/Cy7 (1:1,000, clone G8.8), KRT14 (1:1,000, polyclonal, 905301) (all Biolegend), CD11b (1:100, clone M1/70), CD31 (1:20, polyclonal, AF3628), Ly6G/GR-1 (1:200, clone RB6-8C5) (all R&D Systems), PDGFRα-PE/Cy7 (1:100, clone APA5, Thermo Fisher Scientific), prollyl 4-hydroxylase subunit beta (1:50, clone 3-2B12, Acris), PU.1 (1:200, polyclonal, 2266, Cell Signaling) and vimentin-Alexa 647 (1:50, clone V9, Santa Cruz Biotechnology). As secondary antibodies for immunohistochemistry rabbit Alexa 594 (1:200, polyclonal, A-11037), rabbit Alexa 488 (1:200, polyclonal, A-11034), rabbit Alexa 647 (1:500, polyclonal, A-21443), rat Alexa 647 (1:500, polyclonal, A-21472), mouse Alexa 488 (1:200, polyclonal, A11001), mouse Alexa 647 (1:500, polyclonal, A-21236) and goat Alexa 647 (1:500, polyclonal, A-21447) were used (all Thermo Fisher Scientific). As IgG controls for immunohistochemistry goat IgG (sc-2028), rabbit IgG (sc-2027), rat IgG (sc-2026) and mouse IgG (sc-2025) were used (all Santa Cruz Biotechnology). The specificity of the respective antibodies was confirmed by corresponding IgG staining (Extended Data Fig. 10a, b). The F-actin cytoskeleton was visualized with rhodamine-conjugated phalloidin (1:250, R415, Thermo Fisher Scientific). In addition, cell nuclei were stained using DAPI (1:800, sc-3598, Santa Cruz Biotechnology). Six randomly chosen HPFs (0.125 mm<sup>2</sup>) at 200-fold magnification per patient or healthy volunteer were evaluated by two experienced researchers in a blinded manner. Stained cells were visualized either using a Nikon Eclipse Ni-U microscope (Nikon) or an inverted CLSM-1P Leica SP5 II microscope (Leica). Representative images were reconstructed using the ImageJ distribution Fiji<sup>63,64</sup>. Voronoi-tessellated pictures were generated as described elsewhere<sup>65</sup>. For quantification of F-actin, the same microscope settings were used for each HPF. The mean intensity of rhodamine-conjugated phalloidin was measured as raw integrated density (RawIntDen) divided by the area of the cell using ImageJ. Myofibroblasts were identified as single cells that were double-positive for α-SMA and collagen and not directly adjacent to CD31<sup>+</sup> endothelial cells (Extended Data Fig. 10c); α-SMA and collagen double-positive cells were counted in three randomly chosen HPFs of *n* specimens per mouse in a blinded manner at 200-fold magnification.

**ChIP assays.** ChIP assays were performed using the ChIP-IT Express Kit (Active Motif). In brief, 10 μg of sonicated chromatin extract was incubated with antibodies against H3K27me3, H3K27ac, H3K9me3, H3K4me1 and H3K4me3 (39155, 39297, 39161, 39135 and 39915; all from Active Motif), PU.1, SMAD3 (2266 and 9523; both from Cell Signaling Technology) and TEAD1 (610923 from BD Biosciences) or normal rabbit or mouse IgG antibody (sc-2027 X or sc-2025; both from Santa Cruz Biotechnology). Purification was performed using the Chromatin IP DNA Purification Kit (Active Motif) and bound sequences were determined by quantitative real-time PCR using primers listed in Supplementary Table 1.

**ChIP-seq.** Single-end reads were generated from PU.1 precipitated (2266, Cell Signaling), input and IgG control DNA on an Illumina HiSeq 2500 system. Alignment to the GRCh37 reference genome was performed with bwa mem version 0.7.14-r1136<sup>66</sup>. MACS version 2.1.1.20160309<sup>67</sup> was used to call peaks for each sample, using both input and IgG alignments as controls. In preparation for a motif enrichment search, two different region files were generated from the peak files as output by MACS. First, a file with the union of all three peak files was generated. Second, a file with only the flanking regions (200 bp in either direction) of each peak region in the union file (vicinity analysis) was created. Motif enrichment analysis was then performed using HOMER software version 4.9.1<sup>43</sup>. For region of interest (ROI)-based ChIP-seq peak and RNA-seq overlap analysis, ROIs were defined as 225 kb upstream and downstream of differentially expressed genes<sup>68</sup>. Within each ROI, the most significantly enriched peak from ChIP-seq analysis was determined. Differentially expressed genes from RNA-seq were mapped to the annotation of PU.1 ChIP-seq peaks (*q* < 0.05). For the identification of active regulatory elements, 11 ENCODE datasets containing DNase-sequencing and histone ChIP-seq data of human dermal and lung fibroblast (ENCFF128ARX, ENCFF148DHA, ENCFF195SIN, ENCFF328XNN, ENCFF350PQN, ENCFF392WNX, ENCFF524YEK, ENCFF811YTI, ENCFF965XKX, ENCFF073ILZ and wgEncodeBroadHistoneNhlh3k27acStdSig) were used. For each file, the 99th percentile of all enrichment values was set as the threshold. The regions beyond this threshold from all files were merged to determine the percentage of ChIP-seq peaks at various *q*-value thresholds overlapping these regions.

**RNA-seq.** RNA-seq was performed on an Illumina HiSeq 2500 system with a read length of 100 bp (forward only). After adaptor trimming and filtering using

cutadapt version 1.9.1<sup>69</sup>, reads were mapped to the Ensembl GRCh37 human reference using STAR version 2.5.2a<sup>70</sup>. Features were counted with subread featureCounts version 1.5.1<sup>71</sup> (count > 5 as threshold) on the Ensembl GRCh37 release 85 genome annotation. All further analysis was performed in R version 2.15.3 using the DESeq2 package<sup>72,73</sup>. GSEA was performed using GSEA version 3.0 software (Broad Institute)<sup>51,74</sup>. The statistical significance was assessed using 10,000 random permutations of the gene set with a signal-to-noise metric for ranking genes. A FDR-corrected value of *q* < 0.25 was considered to be significant. Gene sets were obtained from the Molecular Signatures Database (MSigDB) version 6.1 or created based on published signature genes (fibrotic cluster).

**Bisulfite pyrosequencing.** Genomic DNA was prepared from fibroblasts using the QIAamp DNA Blood Mini kit (QIAGEN). The DNA (1 μg) was bisulfite-modified using the EpiTect bisulfite kit (QIAGEN). PCR amplification of bisulfite-modified DNA (2 μl) was performed using AmpliTaq Gold polymerase (Thermo Fisher Scientific). The PCR program was 95 °C for 4 min; 5 × 95 °C for 30 s, 52 °C for 90 s, 72 °C for 2 min; 25 × 95 °C for 30 s, 52 °C for 90 s, 72 °C for 90 s; 72 °C for 4 min. The following primers were used: PU.1 promoter forward 5'-TAGTAAGTTAGGAGGGTAGTGGGTG-3'; biotin reverse 5'-CCCCATCCTAAAAACTCTACATTA-3'; Pyro-seq forward 5'-GTTGGGTTGGTGGAGGAGT-3'; PU.1 enhancer forward 5'-GGTTGTAGTTGTTTTTGT TTTTATAT-3'; biotin reverse 5'-CTAAACATCCCCCTAAAACCTAAC-3'; Pyro-seq forward 5'-AGTTATATAGGAAGTATGTG-3'. The PCR products were visualized with agarose gel electrophoresis. Afterwards, they were directly sequenced using the PyroMark Q48 Autoprep according to the manufacture instructions (QIAGEN).

**Cell isolation and flow cytometry.** Mice were euthanized by cervical dislocation under anaesthesia and dissected to generate single-cell suspensions from the lung, spleen, thymus and/or bone marrow. Fat was thoroughly removed from the dissected organs and their capsules were opened to ensure good drainage of the digestive solution, which consisted of RPMI 1640 medium supplemented with 1 mg ml<sup>-1</sup> collagenase D from *Clostridium histolyticum* and 0.2 mg ml<sup>-1</sup> DNase I, grade II from bovine pancreas (both from Roche Diagnostics). For the digestion of liver and lung samples, the digestive solution was enriched with 0.1 mg ml<sup>-1</sup> Dispase II (Roche Diagnostics). Lung, spleen and thymus were digested in 1 ml digestion medium at 37 °C for 1 h on a thermo shaker at 500 r.p.m. (Eppendorf). Pipetting every 20 min ensured good dissociation of the tissue. Tibias were cut off at both ends and bones were flushed with PBS to collect bone marrow. The resulting single-cell suspensions were filtered through 70-μm cell strainers and washed in a larger volume of RPMI 1640 supplemented with 10 mM EDTA and 10% FBS. Red blood cells were lysed after digestion by applying in-house-made ACK buffer for 1 min. Lysis was stopped by adding a sufficient amount of 10 × PBS to generate a 1 × solution. Cells were then washed in PBS supplemented with 5 mM EDTA and 2% FBS and filtered through 40-μm cell strainers. Skin and liver samples were centrifuged through a gradient to remove debris (Debris Removal Kit, Miltenyi Biotec). For flow cytometry analysis, 1 × 10<sup>6</sup> cells of the resulting single-cell suspensions were incubated in 100 μl of diluted antibody solution in V-shaped plates for 20 min on ice. The following antibodies were used: CD3ε-Pacific Blue or CD3ε-PE/Cy7 (1:500 or 1:100, respectively, clone 145-2C11), CD4-FITC (1:1,500, clone RM4-5), CD8α-APC (1:300, clone 53-6.7), CD11b-PE/Cy7 or CD11b-APC (each 1:1,000, clone M1/70), CD11c-PE/Cy7 or CD11c-APC/Cy7 (each 1:200, clone N418), CD25-PE or CD25-PE/Cy7 (each 1:500, clone PC61), CD29-PE (1:1,000, clone HMβ1-1), CD31-APC (1:1,000, clone WM59), CD31-PerCP/Cy5.5 (1:1,000, clone 390), CD34-PerCP/Cy5.5 (1:500, clone HM34), CD44-PE (1:2,000, clone IM7), CD45-BV421 (1:2,000, clone 30-F11), CD45-PerCP/Cy5.5 (1:1,000, clone H130), CD45R/B220-FITC or CD45R/B220-APC/Cy7 (each 1:500, clone RA3-6B2), CD49f-APC (1:1,000, clone GoH3), CD115-PE (1:100, clone AFS98), CD117-APC or CD117-BV480 (each 1:100, clone 2B8), CD117-PE (1:500, clone 104D2), CD127/IL7R-PE/Cy7 (1:100, clone A7R34), EpCAM-APC/Cy7 (1:500, clone G8.8), EpCAM-FITC (1:200, clone 9C4), PDGFRα-PE (1:100, clone 16A1), PU.1-PE (1:1,000, clone 7C2C34), TER119-APC/Cy7 (1:100, clone TER119) (all Biolegend), CD45-V500 (1:1,000, clone 30F11), CD105-BV421 (1:100, clone MJ7/18) (all BD Biosciences), COL1A1-FITC (1:200, clone 5D8-G9, Merck), KRT14-PE (1:1,000, clone LL002, Novus Biologicals), PDGFRα-PE/Cy7 (1:1,000, clone APA5, eBiosciences) and vimentin-Alexa 647 (1:2,000, clone V9, Santa Cruz Biotechnology). For viability staining in flow cytometry Zombie Violet (1:1,000, 423113, Biolegend), DAPI (0.1 μg ml<sup>-1</sup>, D9542, Sigma-Aldrich) and eFluor780 (1:4,000, 65-0865-14, eBiosciences) were used. Blocking of Fc receptors was performed before staining with fluorophore-labelled antibodies. Mouse blood (50 μl collected in EDTA) was incubated with the respective antibodies for 20 min at 4 °C. Afterwards, 450 μl of RBC Lysis/Fixation Buffer (1 ×) (Biolegend) was added for another 15 min at room temperature. After washing twice with PBS, cells were resuspended in PBS supplemented with 5 mM EDTA and 2% FBS and filtered through 40-μm cell strainers. All flow cytometry analyses were performed on a Gallios or Cytotex-S flow cytometer (both Beckman Coulter) equipped with

3 laser- (405 nm, 488 nm and 633 nm) and 10 fluorescence-detection channels and analysed using Kaluza version 1.5 (Beckmann) or CytExpert. Gating was performed as shown in Extended Data Fig. 8.

**Gene-expression analysis.** Total RNA was extracted from single-cell suspensions using either the Nucleo Spin RNA isolation kit (Macherey-Nagel) or the miRNeasy Mini Kit (QIAGEN). Subsequently, 1 µg of RNA was used to transcribe mRNA to cDNA following standard protocols. Real-time PCR was performed in triplicates using either the SYBR Select Master Mix (Thermo Fisher Scientific) or the miScript SYBR Green PCR Kit and miScript Primer Assay (both from QIAGEN) and a QuantStudio 6 Flex System (Thermo Fisher Scientific). Expression of target genes was calculated using the  $\Delta C_t$  comparative method for relative quantification after normalization. Samples without enzyme in the reverse-transcription reaction (non-RT controls) were used as negative controls. Unspecific signals caused by primer dimers were excluded by non-template controls and by dissociation curve analysis. *Actb*, *let-7b* or *mir-15a* were used to normalize for the amounts of cDNA within each sample. The following miScript primer assays were used: *let-7b* (MS00003122), *miR-15a* (MS00003178), *miR-92a-2* (MS00032137), *miR-155* (MS00031486), *miR-326* (MS00003948), *miR-580* (MS00010227), *miR-6745* (MSC0075916), *miR-6747* (MS00046515) and *miR-6780a* (MS00046872) (all from QIAGEN). Primer sequences are listed in Supplementary Table 1.

**In silico analysis of potential miRNA binding sites to SPI1.** For predictions of potential miRNA binding to *SPI1*, miRWalk<sup>49</sup>, miRanda<sup>75</sup> and Targetscan<sup>76</sup> were used. The overlap of possible miRNAs from all three tools were further restricted to  $P < 0.0233$  as well as *SPI1* conserved binding sites predicted by miRWalk<sup>49,50</sup>.

**Western blot analysis.** Protein concentrations were determined using the Bradford protein assay according to the manufacturer's protocol (Bio-Rad). Equal amounts of protein were loaded on a Tris-glycine-buffered gel. Proteins were separated by SDS-PAGE and transferred to PVDF membrane. The membrane was incubated with the appropriate primary antibody and HRP-conjugated secondary antibodies (DAKO). The following antibodies were used:  $\beta$ -actin (1:10,000, clone A5441, Sigma-Aldrich), collagen I (1:1,000, clone COL-1, Abcam), TEAD1 (1:500, clone 610923, BD Biosciences), EZH2 (1:2,000, polyclonal, 4905), total histone H3 (1:1,000, polyclonal, 9715), tri-methyl-histone H3(Lys27) (1:1,000, polyclonal, 9733), total SMAD3 (1:1,000, polyclonal, 9513) or phospho-SMAD3 (1:1,000, polyclonal, 9520) and PU.1 polyclonal (1:500, 2266) (all Cell Signaling). As secondary antibodies in western blot anti-mouse (1:1,500, polyclonal, P0447) or anti-rabbit (1:2,000, polyclonal, P0448) HRP-conjugated secondary antibodies (all DAKO) were used. Blots were visualized using enhanced chemiluminescence (ECL).  $\beta$ -Actin was used as a loading control. Western blots were quantified using ImageJ Software (version 1.46r).

**Animal studies of fibrosis.** The role of PU.1 in fibrosis was investigated in five different mouse models. Fibrosis was induced in 6–8-week-old littermates of the stated background. In the first model, bleomycin-induced skin fibrosis was induced by local injections of bleomycin at a concentration of 0.5 mg ml<sup>-1</sup> in defined areas of 1 cm<sup>2</sup> at the upper back every other day for 4 weeks (6 weeks of age, mixed genders)<sup>77</sup>. Mice treated with subcutaneous injections of 0.9% NaCl were used as controls. Second, in the *Tsk1* model (10 weeks of age, mixed genders), a dominant mutation in the gene that encodes fibrillin-1 results in activated TGF- $\beta$  signalling in *Tsk1* fibroblasts and progressive, generalized hypodermal thickening within 10 weeks after birth<sup>70</sup>. Third, bleomycin-induced pulmonary fibrosis was induced by a single intratracheal application of bleomycin (0.025 U, 8 weeks of age, males) using a high-pressure syringe (Penn-Century). Mice in which equal volumes of 0.9% NaCl were injected were used as a control<sup>78</sup>. Fourth, CCl<sub>4</sub>-induced hepatic fibrosis was induced by i.p. injections of CCl<sub>4</sub> diluted in sunflower oil (week 1: 1:31 dilution; week 2: 1:15 dilution; week 3: 1:7 dilution; week 4–6: 1:3 dilution) in mice (8 weeks of age, mixed genders)<sup>79</sup> three times per week. Sunflower oil was used in the control group. Fifth, an LP/J to C57BL/6 minor histocompatibility antigen-mismatched model, which reflects clinical and pathologic symptoms of human sclerodermatous chronic graft-versus-host disease, was used<sup>80</sup>. Recipient C57BL/6 mice underwent total body irradiation with a single dose of 9.5 Gy. Each recipient mouse received  $5 \times 10^7$  splenocytes dissolved in 100 µl PBS within 6 h after irradiation from either C57BL/6 in a syngeneic or LP/J in an allogeneic, multiple minor mismatched transplantation via eye vein injection. As indicated, mice were treated with DB1976<sup>81</sup>. DB1976 was dissolved in water and applied i.p. Controls received NaCl. In the preventive model of bleomycin-induced skin fibrosis, DB1976 was injected i.p. simultaneously with bleomycin or NaCl applications for 4 weeks. In the therapeutic model of bleomycin-induced skin fibrosis, mice were prechallenged with bleomycin for 3 weeks to induce robust skin fibrosis. After 3 weeks, treatment with DB1976 or NaCl as control was initiated, while injections with bleomycin were continued. After a total of 6 weeks of bleomycin and 3 weeks of treatment with DB1976 or NaCl, mice were euthanized and the extent of fibrosis was compared to control mice. In the bleomycin-induced lung fibrosis model, DB1976 was injected for 4 weeks. In the model of CCl<sub>4</sub>-induced liver fibrosis, mice were treated for 6 weeks. ADVIA 120 analyser (version 3.1.8.0-MS; Siemens

Healthcare Diagnostics) was used for the analysis of red blood cell counts, white blood cell count, thrombocytes and reticulocyte count. In all mouse experiments, body weight as well as pain and distress levels were monitored every second day. Pain and distress was evaluated as follows: 0, no signs of stress, mouse is active, in good condition, calm and has a normal appetite; 1, no/mild signs of stress, mouse is active but shows some signs of restlessness; 2, mild pain and distress, mouse is not well groomed, is slightly hunched and has a lower appetite; 3, moderate stress, mouse moves slowly and shows signs of depression; 4, severe pain, mouse loses substantial weight and is non-responsive to touch. If symptoms became worse, mice were excluded from the analysis and euthanized<sup>82</sup>.

**Statistical analysis.** Results were visualized and analysed with Prism version 7 (GraphPad Software) and are depicted as the mean  $\pm$  s.e.m. if not stated otherwise. For a two-group comparison, a Mann–Whitney *U*-test for nonparametric data was used. When two groups of samples were compared for iterating parameters or more than two groups of samples were compared, a one-way ANOVA was used. Tukey's range test was used as post hoc analysis of ANOVA. Significance levels are indicated as suggested by Prism Software: NS,  $P > 0.5$ , \* $P \leq 0.05$ , \*\* $P \leq 0.01$ , \*\*\* $P \leq 0.001$ .

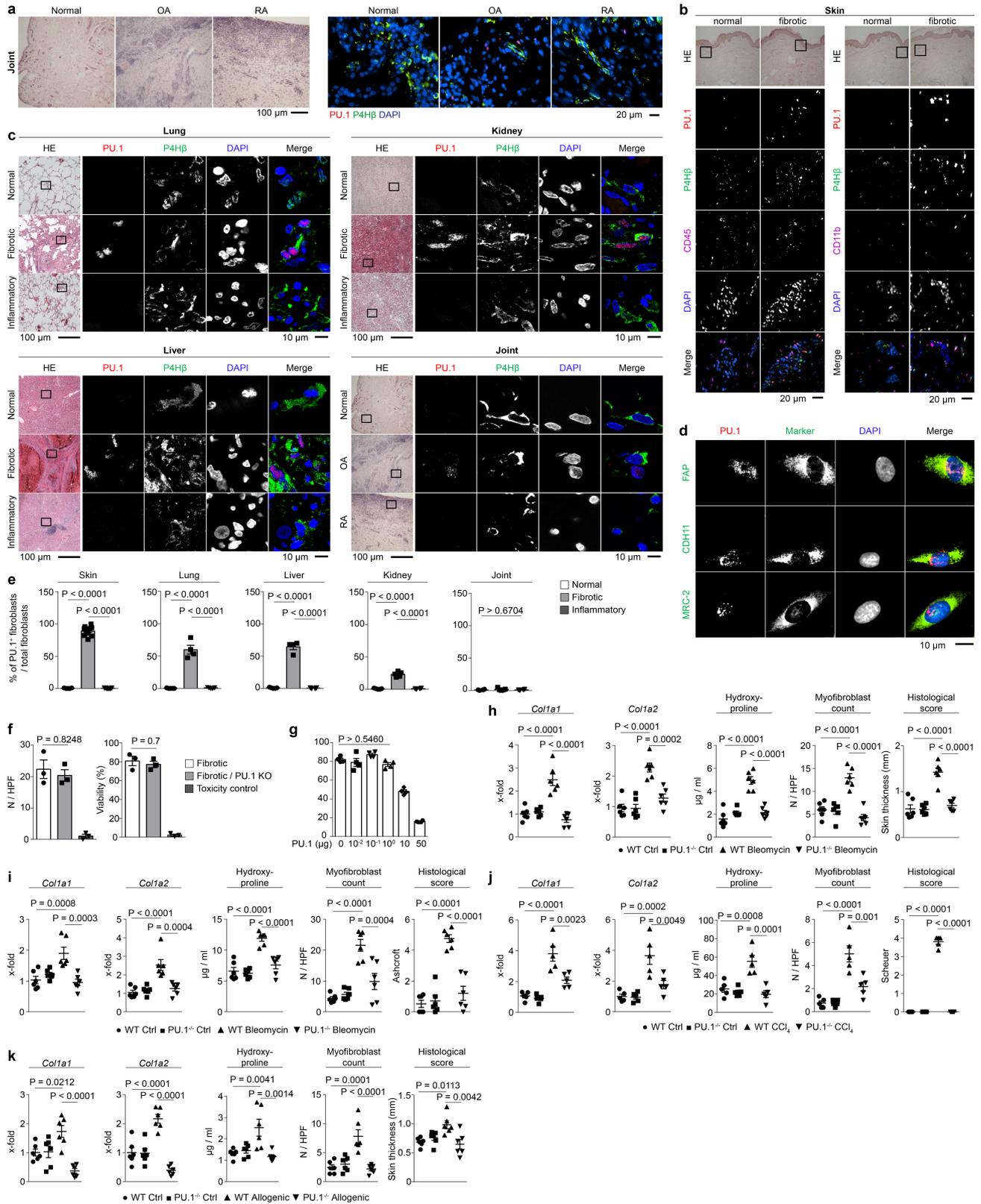
**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

The RNA-seq data used in this study have been deposited in the Gene Expression Omnibus (GEO) under the accession number GSE122334. The sequencing data from the ChIP-seq experiments have been submitted to the National Center for Biotechnology Information (NCBI) database under BioProject PRJNA480591; the project includes the following biosamples: SUB4300598, SUB4300595, SUB4300592, SUB4300591, SUB4300589, SUB4300587, SUB4300586, SUB4300583 and SUB4300579; the FASTQ data was uploaded to the NCBI Sequence Read Archive (SRA) under accession number SRP168244.

- van den Hoogen, F. et al. 2013 classification criteria for systemic sclerosis: an American college of rheumatology/European league against rheumatism collaborative initiative. *Ann. Rheum. Dis.* **72**, 1747–1755 (2013).
- Singal, A. K., Bataller, R., Ahn, J., Kamath, P. S. & Shah, V. H. ACG clinical guideline: alcoholic liver disease. *Am. J. Gastroenterol.* **113**, 175–194 (2018).
- EASL Clinical Practice Guidelines. Autoimmune hepatitis. *J. Hepatol.* **63**, 971–1004 (2015).
- Aletaha, D. et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann. Rheum. Dis.* **69**, 1580–1588 (2010).
- Zheng, B., Zhang, Z., Black, C. M., de Crombrughe, B. & Denton, C. P. Ligand-dependent genetic recombination in fibroblasts: a potentially powerful technique for investigating gene function in fibrosis. *Am. J. Pathol.* **160**, 1609–1617 (2002).
- Armaka, M. et al. Mesenchymal cell targeting by TNF as a common pathogenic principle in chronic inflammatory joint and intestinal diseases. *J. Exp. Med.* **205**, 331–337 (2008).
- Nutt, S. L., Metcalf, D., D'Amico, A., Polli, M. & Wu, L. Dynamic regulation of PU.1 expression in multipotent hematopoietic progenitors. *J. Exp. Med.* **201**, 221–231 (2005).
- Rossi, A., Appelt-Menzel, A., Kurdyn, S., Walles, H. & Groeber, F. Generation of a three-dimensional full thickness skin equivalent and automated wounding. *J. Vis. Exp.* **96**, e52576 (2015).
- Akhmetshina, A. et al. Dual inhibition of c-abl and PDGF receptor signaling by dasatinib and nilotinib for the treatment of dermal fibrosis. *FASEB J.* **22**, 2214–2222 (2008).
- Ashcroft, T., Simpson, J. M. & Timbrell, V. Simple method of estimating severity of pulmonary fibrosis on a numerical scale. *J. Clin. Pathol.* **41**, 467–470 (1988).
- Scheuer, P. J. Classification of chronic viral hepatitis: a need for reassessment. *J. Hepatol.* **13**, 372–374 (1991).
- Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
- Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
- Jarjour, M. et al. Fate mapping reveals origin and dynamics of lymph node follicular dendritic cells. *J. Exp. Med.* **211**, 1109–1122 (2014).
- Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at <https://arxiv.org/abs/1303.3997> (2013).
- Zhang, Y. et al. Model-based analysis of ChIP-seq (MACS). *Genome Biol.* **9**, R137 (2008).
- Shen, Y. et al. A map of the cis-regulatory sequences in the mouse genome. *Nature* **488**, 116–120 (2012).
- Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
- Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
- Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
- Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

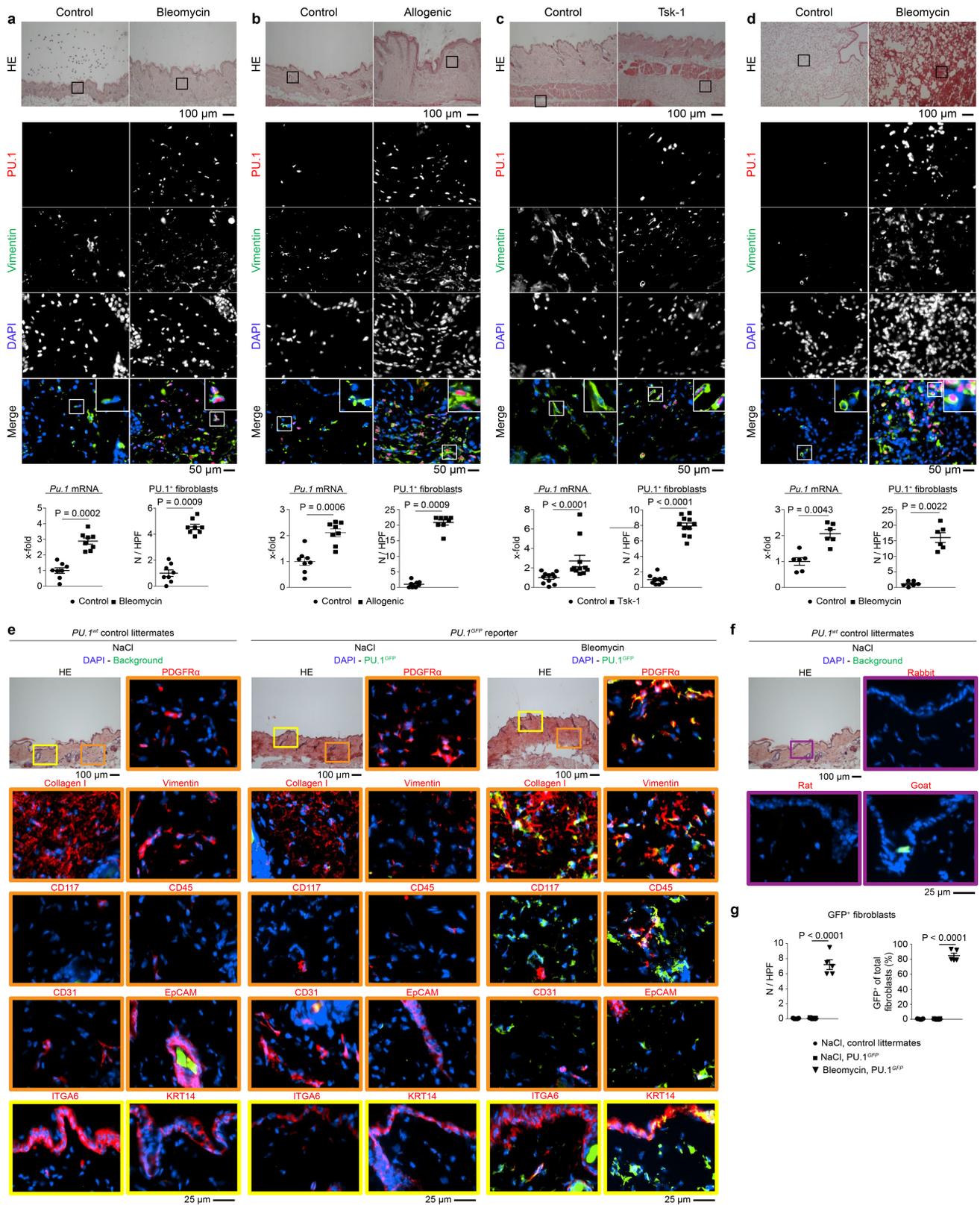
73. R Core Team. *R: A Language and Environment for Statistical Computing*. <http://www.R-project.org/> (R Foundation for Statistical Computing, Vienna, 2013).
74. Mootha, V. K. et al. PGC-1 $\alpha$ -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003).
75. Enright, A. J. et al. MicroRNA targets in *Drosophila*. *Genome Biol.* **5**, R1 (2003).
76. Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P. & Burge, C. B. Prediction of mammalian microRNA targets. *Cell* **115**, 787–798 (2003).
77. Dees, C. et al. Platelet-derived serotonin links vascular disease and tissue fibrosis. *J. Exp. Med.* **208**, 961–972 (2011).
78. Hecker, L. et al. NADPH oxidase-4 mediates myofibroblast activation and fibrogenic responses to lung injury. *Nat. Med.* **15**, 1077–1081 (2009).
79. Popov, Y. et al. Tissue transglutaminase does not affect fibrotic matrix stability or regression of liver fibrosis in mice. *Gastroenterology* **140**, 1642–1652 (2011).
80. DeClerck, Y., Draper, V. & Parkman, R. Clonal analysis of murine graft-vs-host disease. II. Leukokines that stimulate fibroblast proliferation and collagen synthesis in graft-vs. host disease. *J. Immunol.* **136**, 3549–3552 (1986).
81. Munde, M. et al. Structure-dependent inhibition of the ETS-family transcription factor PU.1 by novel heterocyclic diamidines. *Nucleic Acids Res.* **42**, 1379–1390 (2014).
82. Burkholder, T., Foltz, C., Karlsson, E., Linton, C. G. & Smith, J. M. Health evaluation of experimental laboratory mice. *Curr. Protoc. Mouse Biol.* **2**, 145–165 (2012).



Extended Data Fig. 1 | See next page for caption.

**Extended Data Fig. 1 | PU.1-expressing fibroblasts control tissue fibrosis.** **a–c**, Representative images of immunofluorescence (**a**, **b**) and confocal (**c**) microscopy of human skin, lung, liver, kidney and joint biopsy specimens stained for PU.1 (red), P4H3 (green), CD45 or CD11b (purple), and with DAPI (blue). Tissues were obtained from healthy individuals ( $n = 5$  each), idiopathic pulmonary fibrosis ( $n = 4$ ), acute asthma ( $n = 5$ ), alcoholic liver cirrhosis ( $n = 4$ ), autoimmune hepatitis ( $n = 4$ ), cirrhotic kidney ( $n = 4$ ), interstitial nephritis ( $n = 5$ ), osteoarthritis (OA;  $n = 5$ ) and rheumatoid arthritis (RA;  $n = 5$ ). Haematoxylin and eosin (HE)-stained tissue specimens are included. **d**, Representative immunofluorescence images ( $n = 4$ ) of explanted fibrotic fibroblasts stained for PU.1 (red) and one of the following markers (green): FAP, CDH11 or MRC2; nuclei were stained with DAPI (blue). **e**, Semi-quantification of PU.1<sup>+</sup> fibroblasts/total P4H3<sup>+</sup> fibroblasts per HPF. Tissues were obtained from healthy individuals ( $n = 5$  each), patients with systemic sclerosis ( $n = 10$ ), plaque psoriasis ( $n = 7$ ), idiopathic pulmonary fibrosis ( $n = 4$ ), acute asthma ( $n = 5$ ), alcoholic liver cirrhosis ( $n = 4$ ),

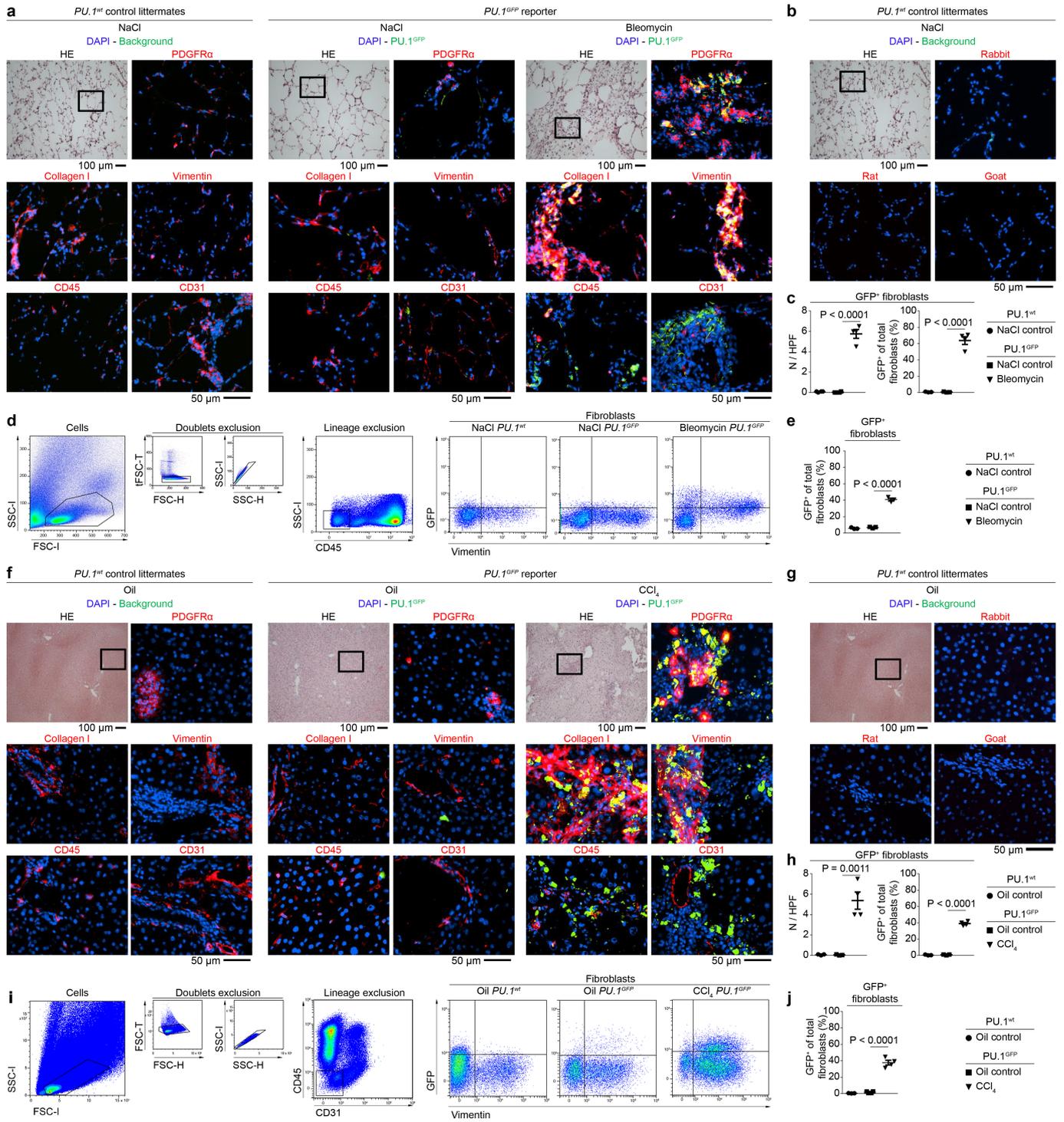
autoimmune hepatitis ( $n = 4$ ), cirrhotic kidney ( $n = 4$ ) and interstitial nephritis ( $n = 5$ ), osteoarthritis ( $n = 5$ ) and rheumatoid arthritis ( $n = 5$ ). **f**, Cell counts and viability of CRISPR–Cas9-mediated PU.1 knockout in human fibrotic fibroblasts compared to unaffected control fibroblasts and fibroblasts treated with 50% DMSO as toxic control ( $n = 3$  each). Cells were counted per HPF. **g**, Resting fibroblasts co-transfected with different amounts of a *SP11* expression plasmid as indicated ( $n = 4$  each). Cell viability of fibroblasts was determined by CCK-8 toxicity assay. **h–k**, Relative *Col1a1* and *Col1a2* mRNA levels, hydroxyproline concentration, myofibroblast counts per HPF and respective histological scores (skin thickness, Ashcroft, Scheuer) in bleomycin-induced skin fibrosis (**h**;  $n = 6$  per group), bleomycin-induced lung fibrosis (**i**;  $n = 6$  per group) and CCl<sub>4</sub>-induced liver fibrosis (**j**;  $n = 5$  per group) models, as well as a sclerodermatous chronic graft-versus-host disease model (**k**;  $n = 6$  per group). Data are mean  $\pm$  s.e.m. of the indicated number of independent experiments. *P* values were determined by one-way ANOVA with Tukey's multiple comparison post hoc test.



Extended Data Fig. 2 | See next page for caption.

**Extended Data Fig. 2 | PU.1-expressing fibroblasts in different mouse models of fibrosis.** **a**, Representative haematoxylin and eosin and immunofluorescence images of a bleomycin-induced skin fibrosis model ( $n = 8$  per group). Mice injected with NaCl were used as controls. **b**, Mouse model of sclerodermatous chronic graft-versus-host disease ( $n = 8$  per group). Syngeneic transplanted mice were used as controls. **c**, Fibrosis model of tight skin 1 (*Tsk1*) mice ( $n = 11$  per group). **d**, Model of bleomycin-induced pulmonary fibrosis ( $n = 6$  per group). Controls received intratracheal application of NaCl. Representative haematoxylin and eosin and immunofluorescence images of respective tissues stained for PU.1 (red), vimentin (green), and with DAPI (blue) are included. Total *Spi1* mRNA in the respective tissues was measured by qPCR. Absolute counts of PU.1-expressing fibroblasts were analysed per HPF. **e, f**, Mouse model of bleomycin-induced skin fibrosis ( $n = 5$  per group). Controls received NaCl. Representative haematoxylin and

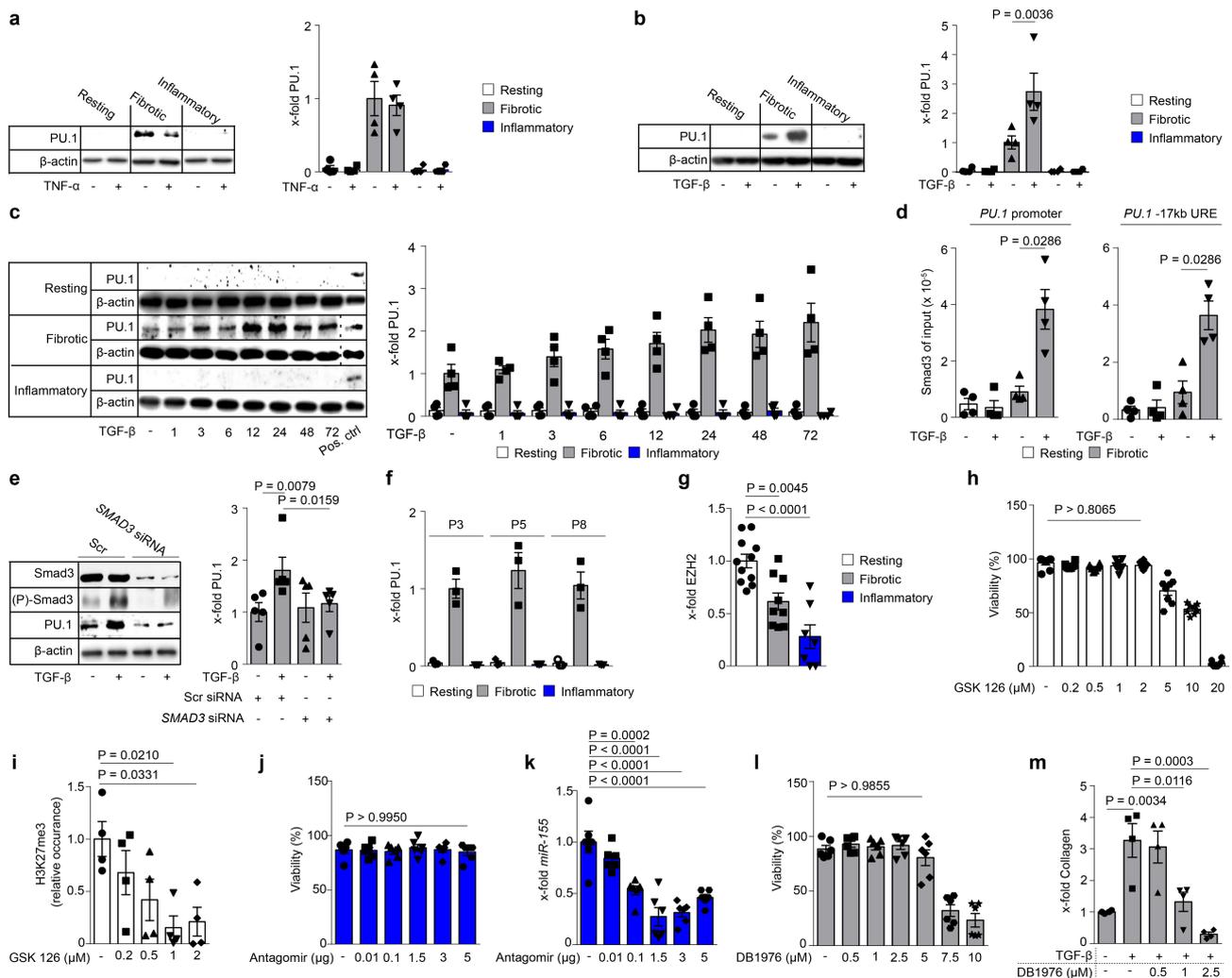
immunofluorescence images of frozen serial tissue sections; boxed areas in the haematoxylin and eosin-stained sections indicate the representative histological regions (yellow, orange and purple) of the correspondingly framed immunofluorescence panels. **e**, Tissues from control littermates or *Spi1<sup>GFP</sup>* reporter mice were stained with DAPI (blue) and the indicated antibodies (red). **f**, IgG control of NaCl-treated control littermates of *Spi1<sup>GFP</sup>* reporter mice ( $n = 3$  per group). **g**, Semi-quantitative analysis of PU.1 (GFP)-expressing fibroblasts. Absolute counts of PU.1-expressing fibroblasts were analysed per HPF (respective  $n$  is given in **e**). Control images of GFP<sup>+</sup> tissue sections are shown in Extended Data Fig. 10d. Data are mean  $\pm$  s.e.m. of the indicated number of independent experiments.  $P$  values were determined by either one-way ANOVA with Tukey's multiple comparison post hoc test or two-tailed Mann-Whitney  $U$ -test if two groups were compared.



Extended Data Fig. 3 | See next page for caption.

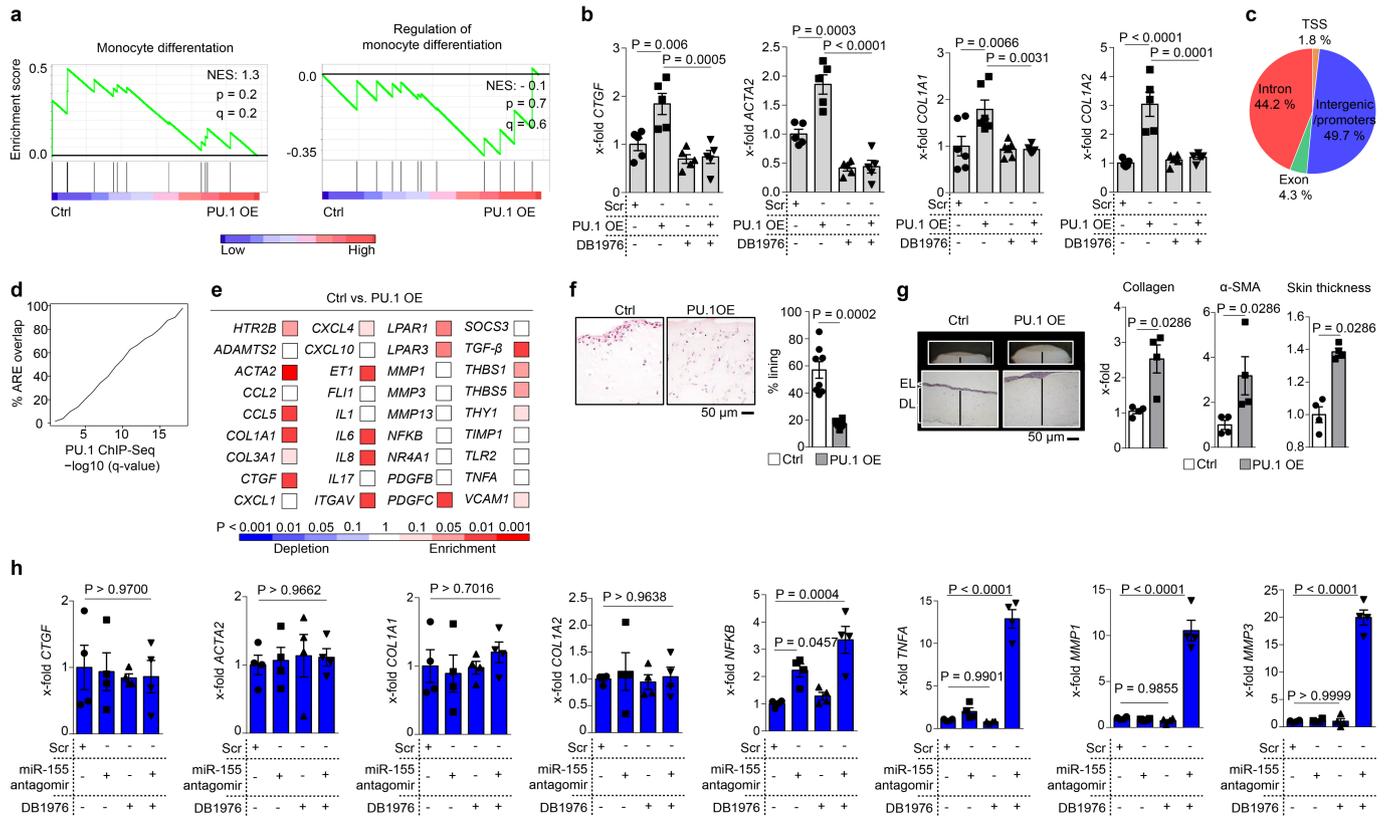
**Extended Data Fig. 3 | PU.1-expressing fibroblasts in bleomycin-induced lung and CCl<sub>4</sub>-induced liver fibrosis.** **a–e**, Mouse model of bleomycin-induced lung fibrosis ( $n = 4$  per group). Controls received NaCl. **a**, Representative haematoxylin and eosin and immunofluorescence images of frozen serial tissue sections of control littermates or *Spi1<sup>GFP</sup>* reporter mice stained with DAPI (blue) and the indicated antibodies (red). **b**, IgG control of NaCl-treated control littermates of *Spi1<sup>GFP</sup>* reporter mice ( $n = 3$  per group). **c**, Semi-quantitative analysis of PU.1 (GFP)-expressing fibroblasts ( $n = 4$  each). Absolute counts of PU.1-expressing fibroblasts were analysed per HPF. Control images of GFP<sup>+</sup> tissue sections are shown in Extended Data Fig. 10e. **d, e**, Flow cytometry analysis of digested lungs. **d**, Gating strategy to characterize GFP<sup>+</sup> cells. **e**, Quantitative analysis of PU.1 (GFP)-expressing fibroblasts ( $n = 3$  each). The percentage of CD45<sup>-</sup>vimentin<sup>+</sup> PU.1-expressing fibroblasts per lung sample is shown. **f–j**, Mouse model of CCl<sub>4</sub>-induced liver fibrosis ( $n = 4$ ). Controls received oil. **f**, Representative haematoxylin and eosin and immunofluorescence images of frozen serial tissue sections of control littermates or *Spi1<sup>GFP</sup>*

reporter mice stained with DAPI (blue) and the indicated antibodies (red). **g**, IgG control of sunflower oil-treated control littermates of *Spi1<sup>GFP</sup>* reporter mice ( $n = 4$  per group). **h**, Semi-quantitative analysis of PU.1 (GFP)-expressing fibroblasts ( $n = 3$  each). Absolute counts of PU.1-expressing fibroblasts were analysed per HPF. Control images of GFP<sup>+</sup> tissue sections are shown in Extended Data Fig. 10f. **i, j**, Flow cytometry analysis of digested livers. **i**, Gating strategy to characterize GFP<sup>+</sup> cells. **j**, Quantitative analysis of PU.1 (GFP)-expressing fibroblasts ( $n = 4$  each). The percentage of CD31<sup>-</sup>CD45<sup>-</sup>vimentin<sup>+</sup> PU.1-expressing fibroblasts per liver sample is shown. Data are mean  $\pm$  s.e.m. of the indicated number of biologically independent samples. *P* values were determined by one-way ANOVA with Tukey's multiple comparison post hoc test. **a, b, f, g**, Boxed areas in the haematoxylin and eosin-stained sections indicate the representative histological regions of the corresponding immunofluorescence panels. Experiments were repeated three times independently with similar results.



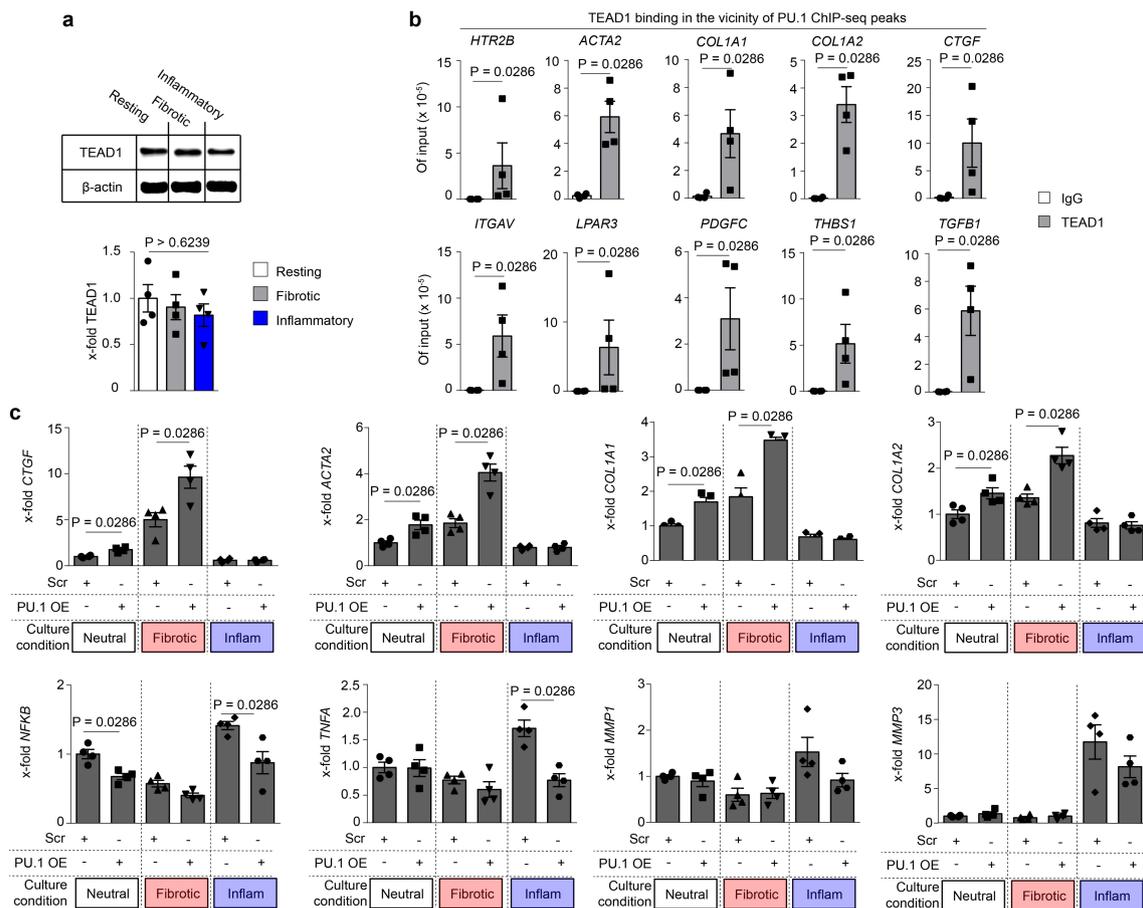
**Extended Data Fig. 4 | Regulation of PU.1 expression in fibroblasts.** **a–c**, PU.1 expression levels of primary human fibroblasts. Representative western blot and semi-quantitative analysis of PU.1 protein expression in resting (isolated from normal skin), fibrotic (isolated from fibrotic skin of patients with systemic sclerosis) and inflammatory (isolated from inflamed joints of patients with rheumatoid arthritis) fibroblasts stimulated with or without TNF for 24 h (**a**), with or without TGF- $\beta$  for 24 h (**b**) or for up to 72 h (**c**) ( $n = 4$  per group). Protein extracts of fibrotic fibroblasts were used as positive control in each lane. **d**, ChIP analysis ( $n = 4$  each) assessing the binding of SMAD3 to the *SPI1* promoter and its  $-17$  kb URE is shown. **e**, Short interfering RNA (siRNA)-mediated knockdown of *SMAD3* in fibrotic fibroblasts stimulated with or without TGF- $\beta$  for 24 h ( $n = 5$ ). Scrambled (scr) siRNA was used as control. **f**, Expression levels of PU.1 in primary human resting, fibrotic and inflammatory fibroblasts ( $n = 3$  each) cultured ex vivo for several passages. **g**, Expression levels of EZH2 in resting ( $n = 11$ ), fibrotic ( $n = 9$ ) and inflammatory ( $n = 7$ ) fibroblasts relative to  $\beta$ -actin as assessed by western blot analysis. Results

are presented relative to resting fibroblasts. **h, i**, Resting fibroblasts treated with different concentrations of GSK126 as indicated ( $n = 3$  each). **h**, Cell viability of fibroblasts was determined by CCK-8 toxicity assay. **i**, Expression levels of H3K27me3 relative to total H3 as assessed by western blot analysis. Results are presented relative to untreated control. **j, k**, Inflammatory fibroblasts treated with different concentrations of miR-155 antagonists as indicated ( $n = 3$  each) to investigate cell viability by CCK-8 toxicity assay (**j**) and *mir-155* expression levels relative to *let-7b* as assessed by qPCR (**k**). Results are presented relative to cells co-transfected with scrambled antagonists. **l, m**, Fibrotic fibroblasts treated with different concentrations of DB1976 to analyse cell viability by CCK-8 toxicity assay ( $n = 6$ ) (**l**) and DB1976 dose-dependent effects on TGF- $\beta$ -induced collagen production ( $n = 4$  each) (**m**). Results are presented relative to untreated control. Data are mean  $\pm$  s.e.m. of the indicated number of independent experiments.  $P$  values were determined by one-way ANOVA with Tukey's multiple comparison post hoc test.



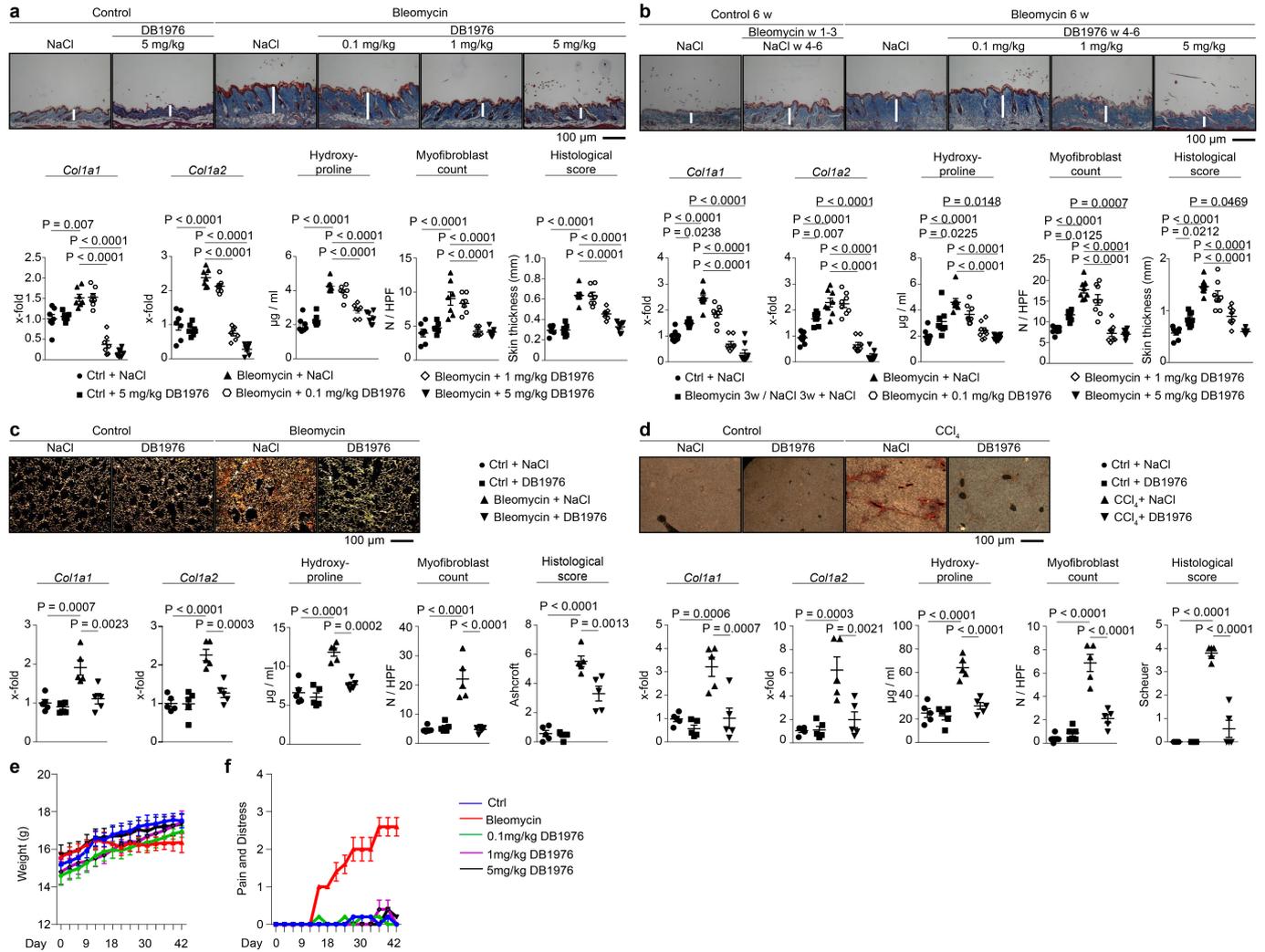
**Extended Data Fig. 5 | Pro-fibrotic potential of PU.1.** **a**, GSEA of quantitative RNA-seq signals of GO-defined monocyte-related gene clusters in human resting fibroblasts co-transfected with PU.1 ( $n = 4$ ). Resting fibroblasts co-transfected with a control plasmid were used as controls ( $n = 4$ ). **b**, mRNA expression levels of indicated transcripts in human resting fibroblasts treated with or without DB1976 and simultaneously co-transfected with or without *SPI1* expression plasmid (pUNO.1-hSPI1 (PU.1 OE),  $n = 5$  per group) as assessed by qPCR. Cells co-transfected with a scramble (scr) plasmid were used as control. Results are presented relative to cells co-transfected with scramble. **c**, Genomic annotation of PU.1-binding sites defined by ChIP-seq analysis in primary human fibrotic fibroblasts. **d**, Annotation of PU.1 ChIP-seq peaks ( $n = 3$  each) at various  $q$ -value thresholds to active regulatory elements (AREs). For unbiased identification of active regulatory elements, 11 ENCODE datasets from DNase-sequencing and histone ChIP-seq were used as described in the Methods;  $q$  values are those provided by MACS2 call-peak<sup>67</sup>. **e**, Differentially expressed genes from gene sets of inflammatory fibroblasts co-transfected with *SPI1* (PU.1 OE) or scramble vector as control (ctrl). Gene sets include fibrosis-associated, inflammatory and matrix-degrading pathways determined by qPCR ( $n = 4$  each). Colours

represent the significance levels of the observed changes in the respective expression levels in PU.1 overexpression compared to control. **f**, Micro-mass organoids of inflammatory fibroblasts co-transfected with *SPI1* or scramble vector in the presence of TNF for 21 days ( $n = 8$  per group). Sections of micro-mass organoids were stained with haematoxylin and eosin. Lining fibroblasts were quantified relative to total number of cells per HPF. **g**, Three-dimensional full-thickness skin organoid model of inflammatory fibroblasts co-transfected with *SPI1* or scramble vector. The collagen content was measured by hydroxyproline assay;  $\alpha$ -SMA expression and skin thickness were quantified per HPF ( $n = 4$  each). **h**, mRNA expression levels of indicated transcripts in primary human inflammatory fibroblasts treated with or without DB1976 and simultaneously co-transfected with or without miR-155 antagonists ( $n = 4$  each). Results are presented relative to cells co-transfected with scramble (scr) antagonists. Data are mean  $\pm$  s.e.m. of the indicated number of independent experiments.  $P$  values were determined either according to a previous study<sup>51</sup> (**a**), by one-way ANOVA with Tukey's multiple comparison post hoc test (**b**, **h**) or by two-tailed Mann-Whitney  $U$ -test (**e-g**).



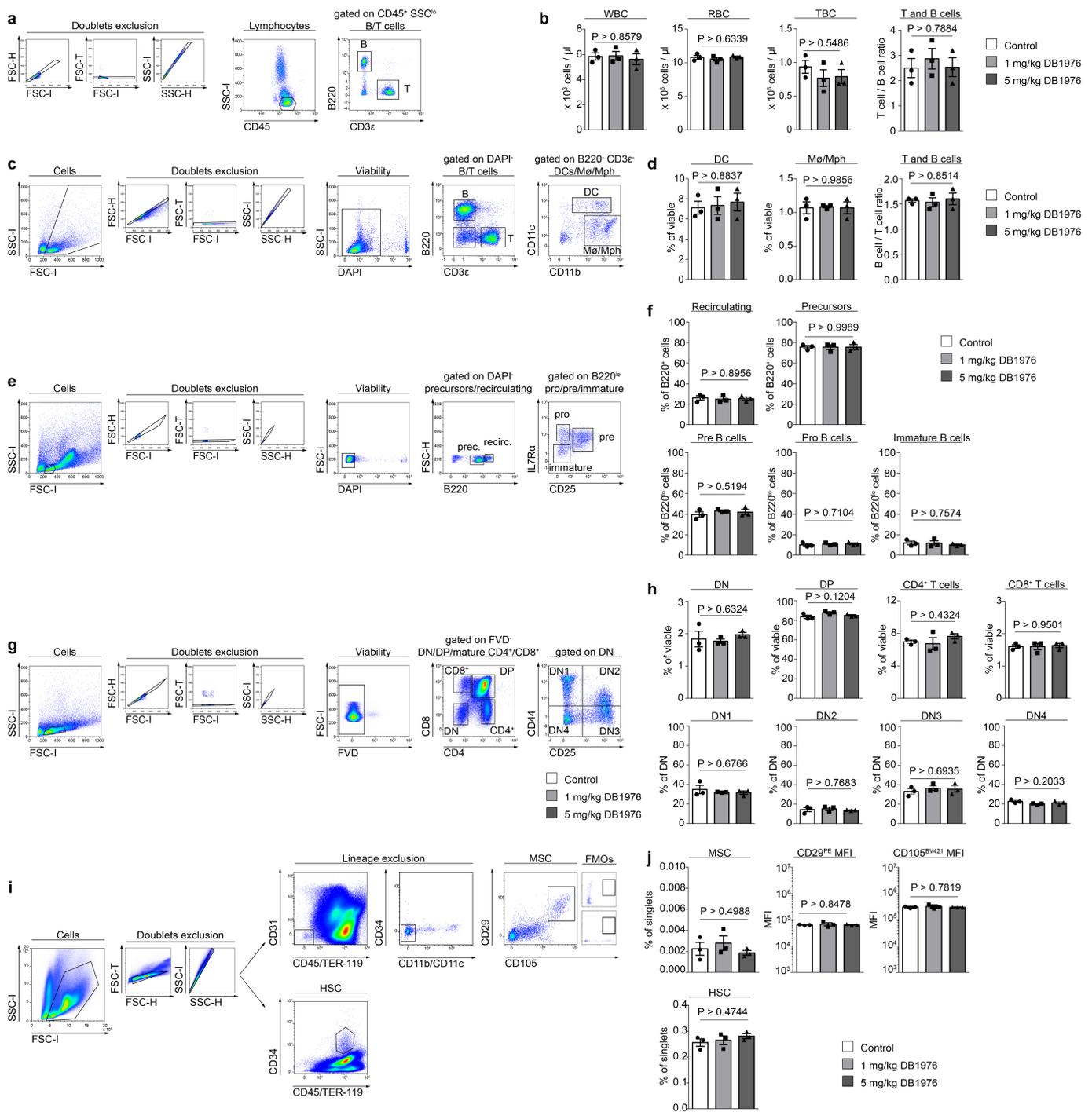
**Extended Data Fig. 6 | PU.1 anchors differentiation towards fibrotic fibroblasts in a network of flanking factors including TEAD1.** **a**, TEAD1 expression levels of primary human fibroblasts. Representative western blot and semi-quantitative analysis of TEAD1 expression in resting, fibrotic and inflammatory fibroblasts ( $n = 4$  each). **b**, ChIP analysis of TEAD1 binding at regions of genes with a fibrotic signature in the vicinity of PU.1-binding sites. DNA fragments of human fibrotic fibroblasts were immunoprecipitated with anti-TEAD1 and analysed by qPCR relative to input DNA ( $n = 4$  per group). Results are compared to IgG control. Signature pro-fibrotic genes were screened for PU.1 ChIP-seq peaks

and potential flanking TEAD1-binding sites. **c**, mRNA expression levels of indicated transcripts in primary human inflammatory fibroblasts co-transfected with *SP11* or scramble plasmid ( $n = 4$  each); cells were cultured under neutral conditions (serum-starved medium only) or in the presence of TGF- $\beta$  (fibrotic) or TNF (inflammatory (inflam)). Results are presented relative to scramble under neutral culture conditions. Data are mean  $\pm$  s.e.m. of the indicated number of biologically independent samples.  $P$  values were determined either by one-way ANOVA with Tukey's multiple comparison post hoc test or two-tailed Mann-Whitney  $U$ -test if two groups were compared.



**Extended Data Fig. 7 | PU.1 silencing in experimental fibrosis.** **a–f**, Experimental fibrosis models. Representative images of trichrome-stained (**a, b**) or Sirius red-stained (**c, d**) tissue sections, mRNA levels of *Col1a1* and *Col1a2*, hydroxyproline content, myfibroblast counts and respective histological scores (skin thickness, Ashcroft, Scheuer) in mice treated with or without DB1976. Mice treated with NaCl or oil were used as controls. **a, b**, Bleomycin-induced skin fibrosis model with preventive (**a**;  $n = 7$ ) or therapeutic (**b**;  $n = 8$ ) treatment; in the latter, regression of pre-established fibrosis was evaluated since mice were challenged with bleomycin for 3 weeks to induce robust skin fibrosis before treatment with

DB1976 was initiated, while injections with bleomycin were continued. As an additional control, mice were injected with bleomycin for 3 weeks followed by injections with NaCl for another 3 weeks. **c**, Bleomycin-induced lung fibrosis model ( $n = 5$ ). **d**, CCl<sub>4</sub>-induced liver fibrosis model ( $n = 5$ ). Body weights (**e**) and levels of pain and distress (**f**) of DB1976-treated mice were monitored every second day ( $n = 5$  each). Mice challenged with subcutaneous injections of bleomycin were used as positive controls. Data are mean  $\pm$  s.e.m. of the indicated number of biologically independent samples.  $P$  values were determined by one-way ANOVA with Tukey's multiple comparison post hoc test.

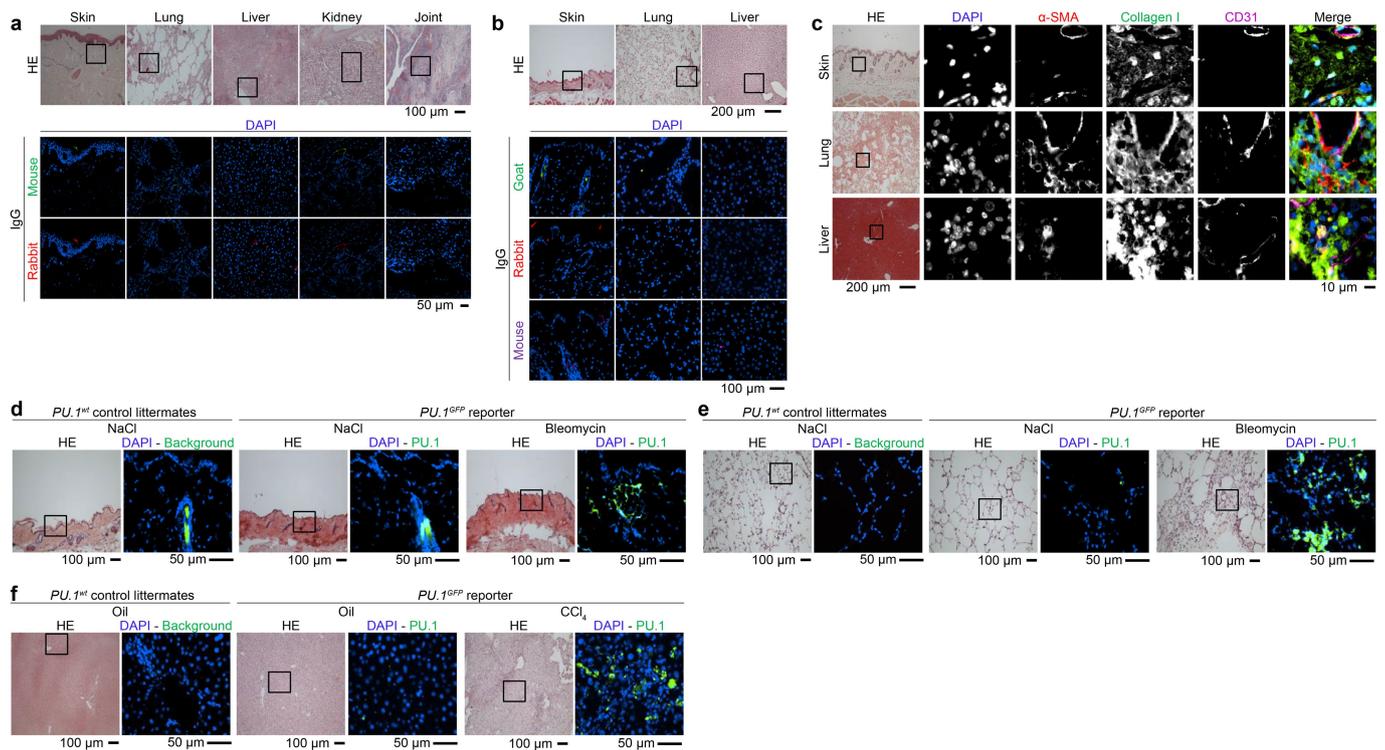


### Extended Data Fig. 8 | Effects of DB1976 in anti-fibrotic concentrations on haematopoietic cells and bone marrow-derived stem cells.

**a, c, e, g, i**, Flow cytometry gating strategy to identify different peripheral blood cell (**a**) and splenic cell (**c**) populations, B cell precursors and mature B cells in the bone marrow (**e**), T cell precursors and mature T cells in the thymus (**g**) or bone marrow-derived mesenchymal stem cells (MSCs) and haematopoietic stem cells (HSCs) (**i**) in mice treated with different concentrations of DB1976 or NaCl ( $n = 3$  each) for 6 weeks. FMO, fluorescence minus one controls. **b**, White blood cell (WBC), red blood cell (RBC) count, numbers of thrombocytes (TBC) and the T to B cell ratio in the peripheral blood. **d**, Quantification of splenic

monocytes (M $\phi$ ), macrophages (Mph), dendritic cells (DC) and the T to B cell ratio. **f**, Frequencies of indicated B cell populations. **h**, Frequency of indicated thymocyte subsets. DN, double-negative thymocytes (based on the expression of CD25 and CD44); DP, double-positive thymocytes. **j**, Percentage of Lin<sup>-</sup>CD29<sup>+</sup>CD105<sup>+</sup> MSCs in the bone marrow, mean fluorescence intensity (MFI) of CD29<sup>+</sup> MSCs and CD105<sup>+</sup> MSCs and percentage of CD45<sup>+</sup>CD34<sup>+</sup> HSCs in the bone marrow are shown. Data are mean  $\pm$  s.e.m. of the indicated number of biologically independent samples. *P* values were determined by one-way ANOVA with Tukey's multiple comparison post hoc test.





**Extended Data Fig. 10 | Control images of human and mouse tissues.**  
**a–c**, Representative haematoxylin and eosin and immunofluorescence images of paraffin-embedded (a–c) human skin, lung, liver, kidney and joint tissues (a) or mouse skin, lung and liver tissues (b) stained with DAPI and Ig controls as indicated ( $n = 5$  each). **c**, Representative images of haematoxylin and eosin and immunofluorescence images of mouse biopsy specimens ( $n = 5$  each) of fibrotic skin, lung and liver stained

with DAPI, and for  $\alpha$ -SMA, collagen I and CD31. **d–f**, Representative haematoxylin and eosin and immunofluorescence images of frozen tissue sections of control littermates or *Spi1<sup>GFP</sup>* reporter mice stained with DAPI (blue). **d**, Mouse model of bleomycin-induced skin fibrosis ( $n = 5$  per group). Controls received NaCl. **e**, Mouse model of bleomycin-induced lung fibrosis ( $n = 4$  per group). Controls received NaCl. **f**, Mouse model of CCl<sub>4</sub>-induced liver fibrosis ( $n = 4$  per group). Controls received oil.

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a | Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated
- Clearly defined error bars  
*State explicitly what error bars represent (e.g.  $SD$ ,  $SE$ ,  $CI$ )*

*Our web collection on [statistics for biologists](#) may be useful.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

Illumina HiSeq Control Software 2.2.68; Molecular Signatures Database (MSigDB) v6.1; Beckman Coulter Gallios; Beckman Coulter Cytoflex-S

Data analysis

bcl2fastq v2.17.1.14; bwa v0.7.14-r1136; bedtools v2.25.0; STAR v2.5.2a; cutadapt v1.9.1; fqtrim v0.9.5; subread 1.5.3; DESeq2 v1.14.1; R v3.4.3; MACS 2.1.1.2016030; Prism version 7 GraphPad Software; xCELLigence RTCA software; ImageJ 1.46r; R version 2.15.3; Kaluza version1.5; CytExpert; HOMER software version 4.9.1; GSEA v3.0 software; G\*Power software 3.1

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

ChIP-seq data are available through NCBI BioProject, Accession No. PRJNA480591; none of the figures have directly associated raw data; data are available on request until final publication.

## Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences  Behavioural & social sciences  Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/authors/policies/ReportingSummary-flat.pdf](https://www.nature.com/authors/policies/ReportingSummary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The sample size was determined from preliminary experiments. We performed a power analysis with G*Power software 3.1. To assess the effect size d, we calculated the respective mean values and standard deviations. Using a $\alpha$ -error of 0.05 and a power of 0.8, we determined the optimal sample size for our study as $n \geq 3$ .
Data exclusions	No data were excluded from the analysis.
Replication	Experimental findings were reliably reproduced in at least three independent experiments.
Randomization	Mice were stratified according to sex and then randomized into the different groups. Cells and tissues from human donors were also randomized.
Blinding	Experiments were done in a blinded fashion except when specifically indicated.

## Reporting for specific materials, systems and methods

### Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Unique biological materials
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants

### Methods

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Antibodies

Antibodies used

The following antibodies were used for immunohistochemistry (IHC):  $\alpha$ -Smooth muscle clone 1A4 (1/500, Sigma-Aldrich), Cadherin 11 polyclonal (LS-B2308, 1/100, LS-Bio), CD45R/B220 clone RA3-6B2 (1/500, eBioscience), Collagen I clone COL1 (1/200), Collagen I polyclonal (ab21286, 1/500), CD11c clone N418 (1/100), CD45 polyclonal (ab10558, 1/500), F4/80 polyclonal (ab100790, 1/200), F4/80 clone Cl:A3-1 (1/100), fibroblast activation protein polyclonal (ab28244, 1/5000), MRC2 polyclonal (ab70132, 1/1000), Vimentin clone VI-10 (1/500) (all abcam), CD3 $\epsilon$  clone 145-2C11 (1/100), CD45-BV421 clone 30-F11 (1/500), CD49f-APC clone GoH3 (1/1000), CD117-APC clone 2B8 (1/1000), EpCAM-APC/Cy7 clone G8.8 (1/1000), KRT14 polyclonal (905301, 1/1000) (all Biolegend), CD11b clone M1/70 (1/100), CD31 polyclonal (AF3628, 1/20), Ly6G/GR-1 clone RB6-8C5 (1/200) (all R&D Systems), PDGFR $\alpha$ -PE/Cy7 clone APA5 (1/100, Thermo Fisher Scientific), Prolyl 4-hydroxylase subunit beta clone 3-2B12 (1/50, Acris), PU.1 polyclonal (2266, 1/200, Cell Signaling) and Vimentin-Alexa 647 clone V9 (1/50, Santa Cruz Biotechnology). As secondary antibodies in IHC Rabbit-Alexa 594 polyclonal (A-11037, 1/200), Rabbit-Alexa 488 polyclonal

(A-11034, 1/200), Rabbit-Alexa 647 polyclonal (A-21443, 1/500), Rat-Alexa 647 polyclonal (A-21472, 1/500), Mouse-Alexa 488 polyclonal (A11001, 1/200), Mouse-Alexa 647 polyclonal (A-21236, 1/500) and Goat-Alexa 647 polyclonal (A-21447, 1/500) were used (all Thermo Fisher Scientific). For IgG controls in IHC Goat IgG (sc-2028), Rabbit IgG (sc-2027), Rat IgG (sc-2026) and Mouse IgG (sc-2025) were used (all Santa Cruz Biotechnology).

The following antibodies were used for flow cytometry: CD3e-Pacific Blue or -PE/Cy7 clone 145-2C11 (1/500 or 1/100), CD4-FITC clone RM4-5 (1/1,500), CD8a-APC clone 53-6.7 (1/300), CD11b-PE/Cy7 or -APC clone M1/70 (each 1/1000), CD11c-PE/Cy7 or -APC/Cy7 clone N418 (each 1/200), CD25-PE or PE/Cy7 clone PC61 (each 1/500), CD29-PE clone HMβ1-1 (1/1000), CD31-APC clone WM59 (1/1000), CD31-PerCP/Cy5.5 clone 390 (1/1000), CD34-PerCP/Cy5.5 clone HM34 (1/500), CD44-PE clone IM7 (1/2000), CD45-BV421 clone 30-F11 (1/2000), CD45-PerCP/Cy5.5 clone HI30 (1/1000), CD45R/B220-FITC or -APC/Cy7 clone RA3-6B2 (each 1/500), CD49f-APC clone GoH3 (1/1000), CD115-PE clone AFS98 (1/100), CD117-APC or -BV480 clone 2B8 (each 1/100), CD117-PE clone 104D2 (1/500), CD127/IL7R-PE/Cy7 clone A7R34 (1/100), EpCAM-APC/Cy7 clone G8.8 (1/500), EpCAM-FITC clone 9C4 (1/200), PDGFRα-PE clone 16A1 (1/100), PU.1-PE clone 7C2C34 (1/1000), TER119-APC/Cy7 clone TER119 (1/100) (all Biolegend), CD45-V500 clone 30F11 (1/1000), CD105-BV421 clone MJ7/18 (1/100) (all BD Biosciences), COL1A1-FITC clone 5D8-G9 (1/200, Merck), KRT14-PE clone LL002 (1/1000, Novus Biologicals), PDGFRα-PE/Cy7 clone APA5 (1/1000, eBiosciences) and Vimentin-Alexa 647 clone V9 (1/2000, Santa Cruz Biotechnology). For viability staining in flow cytometry Zombie Violet (423113, 1/1000, Biolegend), DAPI (D9542, 0.1µg/ml, Sigma-Aldrich) and eFluor780 (65-0865-14, 1/4000, eBiosciences) were used.

The following antibodies were used for western blot: β-Actin clone A5441 (1/10000, Sigma-Aldrich), Collagen I clone COL-1 (1/1000, abcam), TEAD1 clone 610923 (1/500, BD Biosciences), EZH2 polyclonal (4905, 1/2000), total histone H3 polyclonal (9715, 1/1000), Tri-Methyl-Histone H3(Lys27) polyclonal (9733, 1/1000), total polyclonal (9513, 1/1000) or phospho-Smad3 polyclonal (9520, 1/1000) and PU.1 polyclonal (2266, 1/500) (all Cell Signaling). As secondary antibodies in western blot anti-mouse polyclonal (P0447, 1/1500) or anti-rabbit polyclonal (P0448, 1/2000) HRP-conjugated secondary antibodies (all Dako) were used.

#### Validation

All antibodies are commercially available and validated by the manufacturer.

## Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

#### Laboratory animals

Male or female wildtype C57/BL6NRj, PU.1fl/fl X Col6Cre, PU.1fl/fl X Col1a2CreER, PU.1GFP reporter mice (mean age 8 weeks, range 6-10) were used in this study.

#### Wild animals

The study did not involve wild animals.

#### Field-collected samples

The study did not involve samples collected from the field.

## Human research participants

Policy information about [studies involving human research participants](#)

#### Population characteristics

Skin biopsies were obtained from 25 patients with systemic sclerosis according to the 2013 American College of Rheumatology (ACR) / European League Against Rheumatism (EULAR) criteria, seven patients with plaque psoriasis and 21 age- and sex-matched healthy volunteers. Lung tissue was obtained from four patients with idiopathic pulmonary fibrosis (IPF), five patients with asthma and five matched non-inflammatory/non-fibrotic controls. Liver samples were obtained from four patients with alcoholic liver cirrhosis, four samples from patients with autoimmune hepatitis (AIH) and five matched non-inflammatory/non-fibrotic controls. To investigate fibrotic kidney tissue we used cirrhotic kidneys from four patients with end-stage renal disease after renal transplantation or hydronephrosis. Kidney tissues from five patients with interstitial nephritis served as controls. Normal kidney tissues were obtained from macroscopically normal portions of kidneys surgically excised due to the presence of a localized neoplasm (n = 5). Synovial tissue specimens were obtained from five patients with rheumatoid arthritis (RA) who fulfilled the 2010 ACR classification criteria for RA as well as five patients with osteoarthritis (OA). Normal synovium was used as control tissue, which was obtained from surgery specimen of patients with no articular disease (n = 4).

#### Recruitment

Human samples were obtained from research volunteers of the University Hospital Erlangen. Written informed consent was obtained from all subjects. There was no self-selection bias involved.

## ChIP-seq

### Data deposition

Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).

Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

#### Data access links

*May remain private before publication.*

The sequencing data from the ChIP-seq experiments have been submitted to the National Center for Biotechnology Information (NCBI) database under BioProject PRJNA480591;

#### Files in database submission

The project includes following biosamples: SUB4300598, SUB4300595, SUB4300592, SUB4300591, SUB4300589, SUB4300587, SUB4300586, SUB4300583 and SUB4300579; the FASTQ data was uploaded to the NCBI Sequence Read

Archive (SRA).	
Genome browser session (e.g. <a href="#">UCSC</a> )	<a href="http://genome-euro.ucsc.edu/cgi-bin/hgTracks?hgS_doOtherUser=submit&amp;hgS_otherUserName=GPGraphics&amp;hgS_otherUserSessionName=ets1_ChIPseq">http://genome-euro.ucsc.edu/cgi-bin/hgTracks?hgS_doOtherUser=submit&amp;hgS_otherUserName=GPGraphics&amp;hgS_otherUserSessionName=ets1_ChIPseq</a>
<b>Methodology</b>	
Replicates	3 samples used; mean peak correlation 55% (total base pairs)
Sequencing depth	Single end reads (total, mapped): Ssc112_IgG (15801820, 2705858); Ssc112_Input (23656975, 22144346); Ssc112_PU1 (8516025, 3004831); Ssc118_IgG (11278708, 6553411); Ssc118_Input (26577761, 25384936); Ssc118_PU1 (5808879, 4718787); Ssc126_IgG (4674418, 2877667); Ssc126_Input (31033753, 28406995); Ssc126_PU1 (10714304, 4482625)
Antibodies	PU.1 (no. 2266, from Cell Signaling Technology, Danvers, USA)
Peak calling parameters	effective genome size = 2.70e+09; band width = 300; model fold = [5, 50]; qvalue cutoff = 5.00e-02; Regional lambda range [1000 bps, 10000 bps]
Data quality	FASTQ quality check with FastQC (all passed); Enrichment >= 5x and q <= 0.05: Ssc112 394; Ssc118 338; Ssc126 924
Software	Alignment with bwa mem v0.7.14-r1136; peak calling with MACS 2.1.1.20160309

## Flow Cytometry

### Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

### Methodology

Sample preparation	Mice were sacrificed by cervical dislocation under anesthesia and dissected to generate single cell suspensions from the lung, spleen, thymus and/or bone marrow. Fat was thoroughly removed from the dissected organs and their capsules were opened to ensure good drainage of the digestive solution consisting of RPMI 1640 medium supplemented with 1 mg/ml Collagenase D from Clostridium histolyticum and 0.2 mg/ml DNase I, grade II from bovine pancreas (both from Roche Diagnostics, Mannheim, Germany). For the digestion of liver and lung samples, the digestive solution was enriched of 0.1 mg/ml Dispase II (Roche Diagnostics). Lung, spleen and thymus were digested in 1 ml digestion medium at 37 ° C for 1 h on a thermo shaker at 500 rpm (Eppendorf, Hamburg, Germany). Pipetting after every 20 min ensured good dissociation of the tissue. Tibia was cut off at both ends and bones were flushed with PBS to collect bone marrow. The resulting single cell suspensions were filtered through 70 µm cell strainers and washed in a larger volume of RPMI 1640 supplemented with 10 mM EDTA and 10 %FBS. Red blood cells were lysed after digest applying self-made ACK buffer for 1 min. Lysis was stopped adding a qs of 10X PBS to generate a 1X solution. Cells were then washed in PBS supplemented with 5 mM EDTA and 2 % FBS and filtered through 40 µm cell strainers. Skin and liver samples were centrifuged through a gradient to remove debris (Debris Removal Kit, Miltenyi Biotec).
Instrument	All flow cytometric analysis was performed on a Gallios or Cytoflex-S flow cytometer (both Beckmann Coulter, Krefeld, Germany) equipped with 3 Laser (405nm, 488nm, 633nm) and 10 fluorescence detection channels
Software	Beckmann's proprietary software Kaluza version1.5 or CytExpert.
Cell population abundance	10 <sup>6</sup> cells per sample were obtained.
Gating strategy	described in Material and Methods and Extended Data Figures.

- Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.