

# UCSF

## UC San Francisco Previously Published Works

### Title

An interleukin 6 responsive plasma cell signature is associated with disease progression in systemic sclerosis interstitial lung disease.

### Permalink

<https://escholarship.org/uc/item/7rw7r8xd>

### Journal

iScience, 26(11)

### Authors

Jia, Guiquan

Ramalingam, Thirumalai

Heiden, Jason

et al.

### Publication Date

2023-11-17

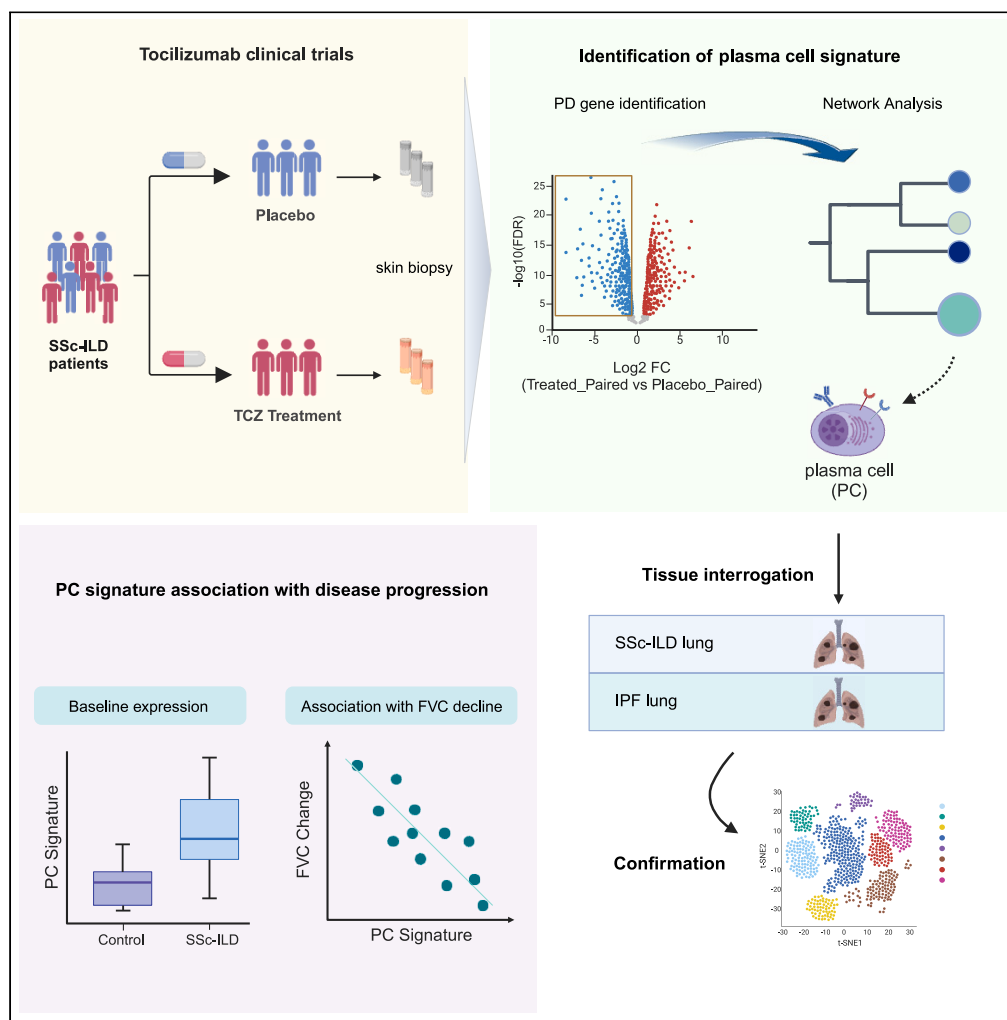
### DOI

10.1016/j.isci.2023.108133

Peer reviewed

Article

# An interleukin 6 responsive plasma cell signature is associated with disease progression in systemic sclerosis interstitial lung disease



Guiquan Jia,  
Thirumalai R.  
Ramalingam,  
Jason Vander  
Heiden, ..., Celia  
Lin, Dinesh  
Khanna, Joseph R.  
Arron

[jjia@gene.com](mailto:jjia@gene.com)

**Highlights**

Tocilizumab treatment benefits SSc-ILD patients with stabilization of lung function

Skin plasma cell (PC) gene signature is identified responsive to the treatment

The enriched PC signature is associated with lung function of SSc-ILD patients

The enriched PC signature is interrogated and verified in the lung of ILD patients



## Article

# An interleukin 6 responsive plasma cell signature is associated with disease progression in systemic sclerosis interstitial lung disease

Guiquan Jia,<sup>1,4,\*</sup> Thirumalai R. Ramalingam,<sup>1</sup> Jason Vander Heiden,<sup>1</sup> Xia Gao,<sup>1</sup> Daryle DePianto,<sup>1</sup> Katrina B. Morshead,<sup>1</sup> Zora Modrusan,<sup>1</sup> Nandhini Ramamoorthi,<sup>1</sup> Paul Wolters,<sup>2</sup> Celia Lin,<sup>1</sup> Dinesh Khanna,<sup>3</sup> and Joseph R. Arron<sup>1</sup>

**SUMMARY**

**Systemic sclerosis (SSc) interstitial lung disease (ILD) is among the leading causes of SSc-related morbidity and mortality. Tocilizumab (TCZ, anti-IL6RA) has demonstrated a reduced rate of pulmonary function decline in two phase 2/3 trials (faSScinate and focuSSced) in SSc-ILD patients. We performed transcriptome analysis of skin biopsy samples collected in the studies to decipher gene networks that were potentially associated with clinical responses to TCZ treatment. One module correlated with disease progression showed pharmacodynamic changes with TCZ treatment, and was characterized by plasma cell (PC) genes. PC signature gene expression levels were also significantly increased in both fibrotic SSc and IPF lungs compared to controls. scRNAseq analyses confirmed that PC signature genes were co-expressed in CD38 and CD138 expressing PC subsets in SSc lungs. These data provide insights into the potential role of PC in disease progression and mechanisms of action of TCZ in fibrotic interstitial lung diseases.**

**INTRODUCTION**

Systemic sclerosis (SSc) is an autoimmune disease characterized by fibrosis, vascular dysfunction, and immune dysregulation. SSc often initially presents as Raynaud's phenomenon, a manifestation of microvasculopathy, with patients testing positive for one or more of a well-defined suite of autoantibodies.<sup>1</sup> While the exact pathogenesis of SSc remains poorly understood, aberrant activation of immune cells abetted by autoantibodies and immune complexes can trigger repeated bouts of tissue injury and activate wound healing pathways that manifest over time as organ fibrosis. While a variable degree of skin involvement is almost universally present in SSc patients, significant morbidity arises from involvement of internal organs such as lungs, kidneys, and gastrointestinal tract. Indeed, interstitial lung disease (ILD) stands among the leading causes of death in SSc patients<sup>2,3</sup> and affects nearly 80% of SSc patients over their lifetimes. Though diffuse skin involvement and certain autoantibody repertoires are associated with greater risk of disease progression, there are no validated biomarkers shown to be clinically useful in prognosticating disease trajectory, or predicting treatment response.

Interleukin 6 (IL-6) has been postulated to be involved in various aspects of SSc pathogenesis, including both activation and recruitment of myeloid and B cells to the lesions.<sup>4-6</sup> Tocilizumab (TCZ), a monoclonal antibody against IL-6RA that inhibits IL-6 activity,<sup>7</sup> was recently tested in SSc patients in two large phase 2/3 trials.<sup>8,9</sup> Although TCZ treatment did not significantly improve skin thickening measured by modified Rodnan skin score (MRSS),<sup>10</sup> lung function, measured as change in forced vital capacity (FVC) over time, was significantly stabilized in patients treated with TCZ. A key question is whether the lack of significant benefit on the clinical MRSS outcome is reflected in pharmacodynamic effects in skin tissue. An objective assessment of changes in disease pathway activity via molecular analysis may be more relevant and robust in quantifying the mechanistic contribution of the targeted biology to SSc pathogenesis.

SSc lends itself to molecular characterization as lesional skin punch biopsies can be readily collected before and after treatment. Assessment of disease activity at the molecular level, using transcriptomic and/or proteomic approaches, can enable us to explore prognostic relationships between activity levels of various pathways to key clinical manifestations of the disease and their progression over time, pharmacodynamics of targeted pathways in response to treatment, and potentially explore predictive aspects of certain molecular signatures that enrich for clinical benefit with a given intervention. Such molecular approaches are instrumental to realizing the promise of stratified medicine such that the right patient receives the right drug at the right time in the disease course. Multiple studies have reported on the heterogeneity of skin gene signatures in patients with SSc,<sup>11-13</sup> and some have correlated these signatures with cross-sectional clinical metrics, but very few

<sup>1</sup>Genentech Inc, South San Francisco, CA 94080, USA

<sup>2</sup>UCSF, San Francisco, CA 94143, USA

<sup>3</sup>University of Michigan, Ann Arbor, MI 48104, USA

<sup>4</sup>Lead contact

\*Correspondence: [jjia@gene.com](mailto:jjia@gene.com)

<https://doi.org/10.1016/j.isci.2023.108133>



**Table 1. Characterization of demographics of subjects in comparison between RNAseq subset and the total cohort of FocuSSced**

Comparison of population characteristics between subjects in the total cohort and RNAseq subset

	RNAseq subset	Total cohort	p value
N	37	208	
Age (mean (SD))	51.27 (11.25)	48.09 (12.43)	n.s.
FVC%pred. (median [IQR])	81.35 [70.19, 93.44]	82.48 [71.12, 93.30]	n.s.
FVC change (median [IQR])	−0.05 [−0.16, 0.06]	−0.05 [−0.12, 0.06]	n.s.
DLCO%pred. (median [IQR])	74.29 [57.28, 89.30]	73.63 [63.21, 88.28]	n.s.
DLCO change (median [IQR])	−0.09 [−0.23, 0.01]	−0.05 [−0.17, 0.09]	n.s.
MRSS (median [IQR])	22.00 [17.00, 30.00]	19.00 [15.00, 25.00]	*
MRSS change (median [IQR])	−0.11 [−0.20, −0.01]	−0.12 [−0.20, −0.02]	n.s.

Comparison of population characteristics within PBO arm

	RNAseq subset	Total cohort	p value
n	19	104	
Age (mean (SD))	54.05 (10.74)	49.19 (12.66)	n.s.
FVC%pred. (median [IQR])	92.13 [81.48, 98.03]	86.84 [72.56, 95.96]	n.s.
FVC change (median [IQR])	−0.13 [−0.28, −0.03]	−0.08 [−0.16, 0.01]	n.s.
DLCO%pred. (median [IQR])	76.55 [69.66, 98.06]	76.15 [65.70, 85.90]	n.s.
DLCO change (median [IQR])	−0.14 [−0.31, −0.03]	−0.06 [−0.18, 0.07]	n.s.
MRSS (median [IQR])	20.00 [15.50, 29.00]	19.00 [15.00, 26.00]	n.s.
MRSS change (median [IQR])	−0.05 [−0.19, 0.02]	−0.11 [−0.20, 0.01]	n.s.

Comparison of population characteristics within TCZ arm

	RNAseq subset	Total cohort	p value
n	18	104	
Age (mean (SD))	48.33 (11.31)	46.98 (12.15)	n.s.
FVC%pred. (median [IQR])	72.89 [66.23, 77.70]	80.00 [69.69, 90.04]	n.s.
FVC change (median [IQR])	0.01 [−0.08, 0.18]	−0.02 [−0.10, 0.08]	n.s.
DLCO%pred. (median [IQR])	60.85 [51.15, 82.33]	71.46 [59.42, 89.24]	n.s.
DLCO change (median [IQR])	−0.04 [−0.12, 0.11]	−0.05 [−0.17, 0.11]	n.s.
MRSS (median [IQR])	25.50 [19.00, 30.75]	19.00 [15.00, 24.25]	**
MRSS change (median [IQR])	−0.18 [−0.23, −0.06]	−0.14 [−0.21, −0.06]	n.s.

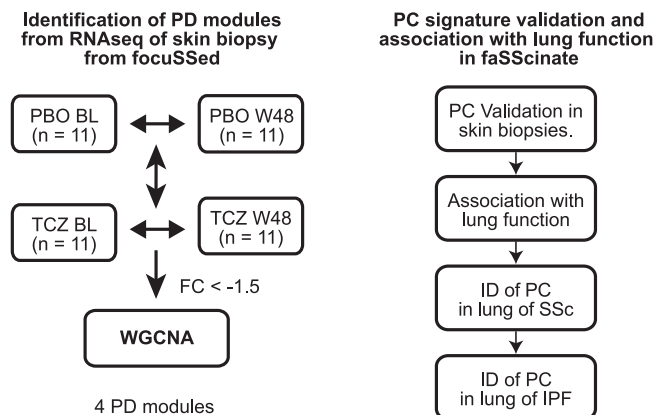
n.s.: no significance statistically; \*\* p value <0.05.

have done so in a large, well curated cohort of patients followed up systematically for an extended period of time to allow for the assessment of disease progression. In order to understand IL-6 modulated cellular and molecular signals dysregulated in SSc patient skin, we performed unbiased RNAseq analyses of skin punch biopsies collected as part of a large phase 3 study of TCZ in SSc, and identified a gene signature attributable to plasma cells (PCs) that is elevated in SSc skin and correlated with disease progression. This signature was blunted by TCZ treatment, thereby implicating a role for IL-6 in plasma cell (PC) mediated pathobiological processes in SSc.

## RESULTS

### Demographic profile of the subjects in the study

A total 212 subjects with diffuse cutaneous SSc were enrolled in the focuSSced study and randomly assigned to receive placebo or tocilizumab (TCZ) as detailed elsewhere.<sup>8</sup> 208 subjects with non-missing baseline and follow up clinical data were used in the post-hoc study. Among them, skin RNA samples were available from 37 subjects that consented for an optional, exploratory biopsy, collected once immediately prior to treatment (baseline, BL), and followed by a second biopsy at 48 weeks of treatment—including 18 from the active treatment arm and 18 from the placebo arm (of which 11 subjects in each arm had one biopsy at each time point) after the removal of outlier samples from RNAseq (see [STAR Methods](#) section for detail). Baseline demographics and disease characteristics for the subjects in the entire study and the subset with skin biopsy samples used in this analysis are presented in [Table 1](#).



**Figure 1. Flow chart of the data analysis**

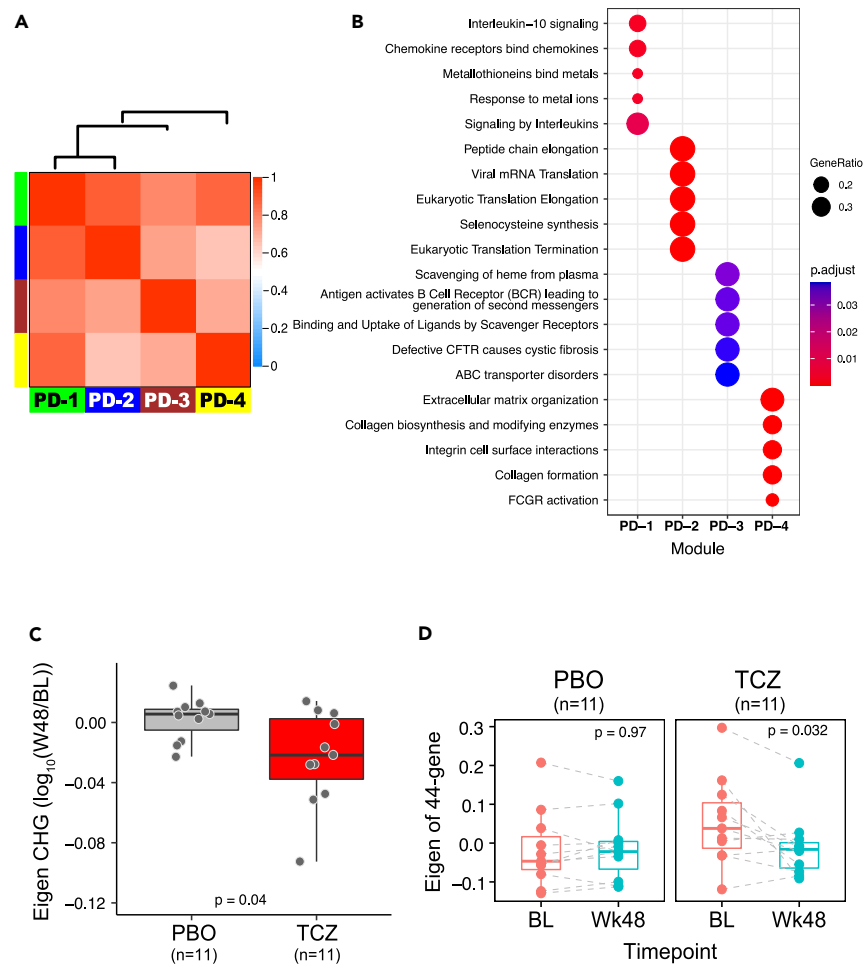
The figure summarizes the main procedures and steps in the entire data analysis; the left panel illustrates the analytical steps for the identification (ID) of pharmacodynamic (PD) modules, while the right panel shows the validation of plasma cell (PC) as PD module associated with the disease. BL: baseline; W48: week 48; FaSS: FaSScinat; PC: plasma cells; PD: pharmacodynamic; PBO: placebo group; TCZ: the group of treatment with tocilizumab; WGCNA: weighted gene co-expression network analysis.

The overall demographic profile was generally comparable between the biopsy subset and the total cohort, including age at baseline, gender distribution, baseline lung function as measured by percent predicted FVC (median FVC%pred.: 81.35 in the subset vs. 82.48 in the total cohort) and percent predicted diffusing capacity of lung for carbon monoxide (median DLCO%pred.: 74.29 in the subset, vs. 73.63 in the total). However, degree of skin involvement at baseline, as measured by the MRSS was higher in the biopsy subset than the total cohort (median MRSS: 22.0 vs. 19.0,  $p$  value <0.05). The key clinical endpoints, including change from baseline to week 48 in MRSS and in lung function, as measured by linear regression slope using all longitudinal measurements, were not meaningfully different between the subset and the entire cohort.

### Identification of a distinct plasma cell (PC) gene module from pharmacodynamic effects of TCZ treatment on skin transcriptional profiles

We performed RNAseq on bulk RNA extracted from available skin biopsies and queried whether any genes/pathways pharmacodynamically (PD) responded to TCZ treatment. We performed differential gene expression analysis in a design of time course with nested individuals for group-specific effects, where we compared baseline to week 48 expression changes intra-individually for individuals with both pre- and post-treatment biopsies prior to aggregating those changes for each group and then compared the aggregate changes between TCZ treatment and the placebo (PBO) groups. Overall, 534 genes from 764 DEG (defined as differentially expressed genes (DEG) with  $p$  value <0.05 & |fold change| > 1.5, as detailed in the Methods section) were significantly down-regulated in the TCZ treatment group compared to PBO over time. These TCZ down-regulated genes were advanced as a candidate PD gene set. To identify molecular pathways implicated by these genes, we used weighted gene co-expression network analysis (WGCNA) across all PD genes to construct co-expressed networks and module partitions (see [STAR Methods](#) section and [Figure 1](#) for details). These analyses identified a total of four modules (designated as PD-1 to PD-4) as shown in [Figure 2A](#) (detailed gene list in each module in [supplemental information](#)). We performed pathway enrichment analysis to identify biological pathways that were implicated by the genes in the modules and top 5 enriched pathways are listed in the dotplot ([Figure 2B](#)). The module PD-1 (including 33 genes) was enriched for pathways involved in the regulation of immune responses; module PD-2 (211 genes) was relatively proximal to PD-1 in correlation and comprised many genes involved in the pathways and cellular functions responsible for gene translation elongation and termination. The genes in the module PD-4 (44 genes) were involved in pathways related to wound healing processes, such as extracellular matrix organization, collagen biosynthesis, and macrophage activation, e.g., Fc-gamma receptor (FCGR) activation ([Figure 2B](#)). Some genes in PD-4 have shown significant reduction in expression as a macrophage gene set in response to TCZ treatment in the FaSScinat cohort,<sup>14</sup> including FCGR1A, FCGR3A, CD163L, C1QB, NNMT, and VSIG4.

Interestingly, the PD-3 module comprising 44 genes (gene set list in [supplemental information](#)) was highly enriched for transcripts associated with PCs (PC, also termed as antibody secreting cells), most of which were immunoglobulin variable region genes (39 out of 44) and others implicated in PC function and/or known to be selectively expressed in PC (MZB1, POU2AF1, DERL3, FAM30A, and CD79A). For instance, the gene MZB1, encoding marginal zone B and B1 cell specific protein, is required for T cell-independent immune responses and differentiation of PCs<sup>15</sup> and POU2AF1 is a B cell-specific transcriptional coactivator.<sup>16,17</sup> Therefore, we designated this gene module the PC module. PC module expression, represented by module eigengene (ME), was significantly reduced following TCZ treatment ([Figure 2C](#)). The heterogeneity in the degree of PD response of the PC module may in turn be related to the overall heterogeneity of the underlying disease among SSc subjects. Subjects showing the greatest PD effect of TCZ treatment generally tended to be those with higher baseline PC module expression within the group, whereas there was no significant change over time in this expression module in the placebo group ([Figure 2D](#); [Figure S1](#)).



**Figure 2. Identified PD modules from RNAseq of skin biopsy**

(A) The Module eigengene adjacency dendrogram (top) and heatmap with red and blue indicating highly related and unrelated modules, respectively, depicting the relationship of the modules of PD genes identified by WGCNA. The color row below the heatmap indicates PD module assignment: PD-1 (n = 33 genes), PD-2 (n = 211 genes), PD-3 (n = 44 genes), and PD-4 (n = 44 genes).

(B) Reactome pathway enrichment of each module gene set. Pathways were determined for each module gene using the compareCluster function with Reactome database. The top 5 of the most over-represented the classification of pathways are illustrated as dot plots, with the gene ratio denoted by size and the significance denoted by color. The p values were adjusted by the Benjamini-Hochberg method.

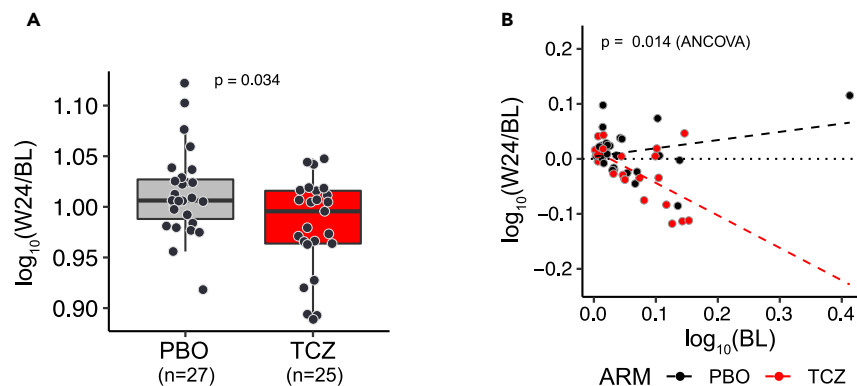
(C) PC module eigengene (ME) was down-regulated in TCZ treatment group compared with placebo group in the FocuSSced cohort. The eigengene change of PC module 48 weeks post treatment from pretreatment is expressed as log10 on y axis; the p value (Wilcoxon test) is indicated at the bottom of graph.

(D) Paired comparison of module eigengene at baseline and 48 weeks following-up. Intra-patient comparison showed a significant reduction of ME in patients with TCZ treatment but not with placebo (PBO).

Immunoglobulin (Ig) variable region genes are uniquely expressed in B-lineage cells and at very high levels in Ig-secreting cells such as plasmablasts and PCs, hence their presence in the PD-3 module specifically implicates Ig-secreting cells as a major component of this module. However, Ig variable region genes are not always well annotated or individually informative in gene expression datasets, due to their intrinsic complicated sequence profiling based on antigen dependency across individuals. Therefore, we used the first four non-immunoglobulin genes (MZB1, POU2AF1, DERL3, and FAM30A) in the module that we also found highly correlated in lungs later as an abbreviated PC gene signature in further analyses unless otherwise noted. We tested whether this four-gene signature could faithfully recapitulate overall PC ME and confirmed that it indeed highly and significantly correlated with the full PC module ( $R = 0.96$ ,  $p < 0.001$ ; [Figure S2A](#)) and showed the same PD effect of TCZ treatment ([Figure S2B](#)).

### Validation of PD effect on PC gene signature in an independent cohort

To validate our observations in an independent cohort, we used skin biopsy transcriptomic data available from the faSScinate study,<sup>14</sup> which recruited a similar profile of SSc patients and used a clinical trial design analogous to focuSSced, with a few notable differences.



**Figure 3. PC signature is down-regulated by TCZ treatment in FaSScinat cohort**

(A) PC signature is down-regulated in TCZ treatment arm compared with placebo arm. The eigengene change of 3-gene PC signature 24 weeks post treatment from pretreatment is expressed on y axis and p value labeled on the top of graph is from Wilcoxon test.

(B) Additional test with ANCOVA model confirmed the significant down-regulation of PC signature by TCZ treatment and the greater PD effects on the subjects with higher PC expression level. Scatterplot shows the function of change (follow-up/baseline measurements) against baseline measurements, both in log<sub>10</sub>. The dotted line represents perfect agreement (no change) and the dashed lines are fitted regression lines for TCZ treatment and placebo groups. P value was from ANCOVA test.

The follow-up biopsy was conducted 24 weeks after treatment, compared to 48 weeks in focuSSced, and the transcriptomic data were captured using a whole transcriptome microarray technology, as opposed to the RNA sequencing approach used for focuSSced. Given platform differences, only three of the four genes from the abbreviated PC module (MZB1, POU2AF1, and DERL3) were represented in the microarray dataset, and used as the PC gene set in this case. Consistent with the observations for the four-gene PC module in focuSSced, this three-gene set in FaSScinat showed significant down-regulation following TCZ treatment as compared to placebo control (Figure 3A) and highly correlated each other (Figure S3). Some of the samples in the placebo group that had higher PC eigengene values at baseline showed a spontaneous reduction at 24 weeks (data not shown), suggesting a possible “regression to the mean” (RTM) phenomenon commonly observed with repeated measures in clinical trials,<sup>18</sup> or a natural change in the underlying disease activity. In order to assess and adjust for the effect of RTM, we performed an analysis of covariance (ANCOVA),<sup>19</sup> factoring the baseline PC eigengene levels as a covariate (Figure 3B). This analysis showed that TCZ treatment had a statistically significant PD effect compared to the PBO arm. In addition, the divergence between the regression lines by increasing baseline PC gene expression levels confirmed the observation in focuSSced that higher PC expression levels at baseline corresponded to a greater potential to observe a PD effect with TCZ treatment.

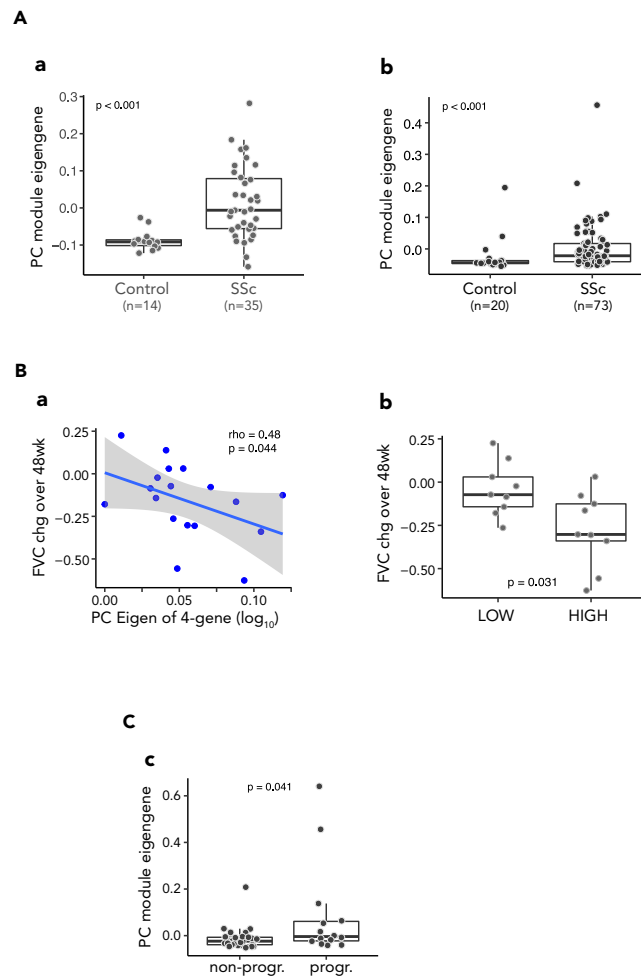
### Skin PC module expression is elevated and prognostic of FVC decline in SSc-ILD patients

The PC gene module showed elevated expression levels in SSc skin samples compared to age and gender matched healthy controls (Figure 4A) in both focuSSced (Figure 4A[a]) and faSScinat (Figure 4A[b]) clinical studies. This is consistent with the observed greater PD effect on the PC module in patients with higher baseline activity of the pathway, as they had a detectable window to show decreases in pathway activity.

As TCZ treatment was associated with a reduced rate of lung function decline in SSc patients in both focuSSced and faSScinat studies, we queried whether skin PC biology and gene expression levels would be informative in the context of SSc lung disease. In the placebo arm of the focuSSced study, where lung function progressively decreased over time, we found that patients with higher level of skin PC module gene expression at baseline showed greater lung disease progression, as measured by FVC change over time (using all the available FVC data from baseline to wk 48) (Figure 4B[a]). Specifically, subjects with higher than median levels of PC signature at baseline showed significantly larger FVC decline over time (Figure 4B[b]).

Although the skin PC eigengene at the baseline did not show significant correlation with continuous change in the FVC over 48 weeks in faSScinat (data not shown) possibly due to the eigengene value calculated from the gene expression derived from the microarray and/or unexplained baseline differences between the subjects in these two clinical trials,<sup>8,14,20</sup> using a combination of categorical definition of disease progression (FVC decrease >10% or DLCO decrease >15% over 48 weeks) which had greater prognostic significance than continuous change in the FVC,<sup>21</sup> subjects with “progressive ILD” showed higher PC eigengene values at baseline than non-progressors in the faSScinat study (Figure 4C).

Interestingly, baseline PC eigengene level did not show a significant correlation with longitudinal changes in skin thickness (MRSS), an endpoint that TCZ treatment failed to significantly impact in both the clinical trials, although there was a nonsignificant trend for a baseline correlation between PC score and MRSS. Nor was a significant correlation evident between PC eigengene and diffusing capacity of the lungs for carbon monoxide (DLCO), either at baseline level or the change over time (Figure S4).



**Figure 4. Elevated PC signature was associated with FVC decline in SSc-ILD**

(A) Baseline PC module is highly elevated in SSc skin compared with healthy controls. (a) PC module eigengene in FocuSSc; (b) 3-gene eigengene in FaSScinatate.

(B) Baseline PC module expression associated with FVC decline over 48 weeks.

(a) PC module eigengene (ME) in skin biopsies was correlated inversely with FVC change (percent predicted) in placebo group of the FocuSSc study. The Spearman correlation coefficient and p value are listed on the top of the plot. (b) Subjects dichotomized according to median baseline ME showed significant differences in FVC change. P value was from Wilcoxon test. LOW: the group with baseline ME less than the median; HIGH: the group with baseline ME greater than the median.

(C) The progressors in lung function decline showed higher PC gene expression in the placebo arm of FaSScinatate cohort. Progressors (progr.) were defined as subjects with FVC decline over 48 weeks  $\geq 10\%$  or DLCO decline  $>15\%$ . P value was from Wilcoxon test.

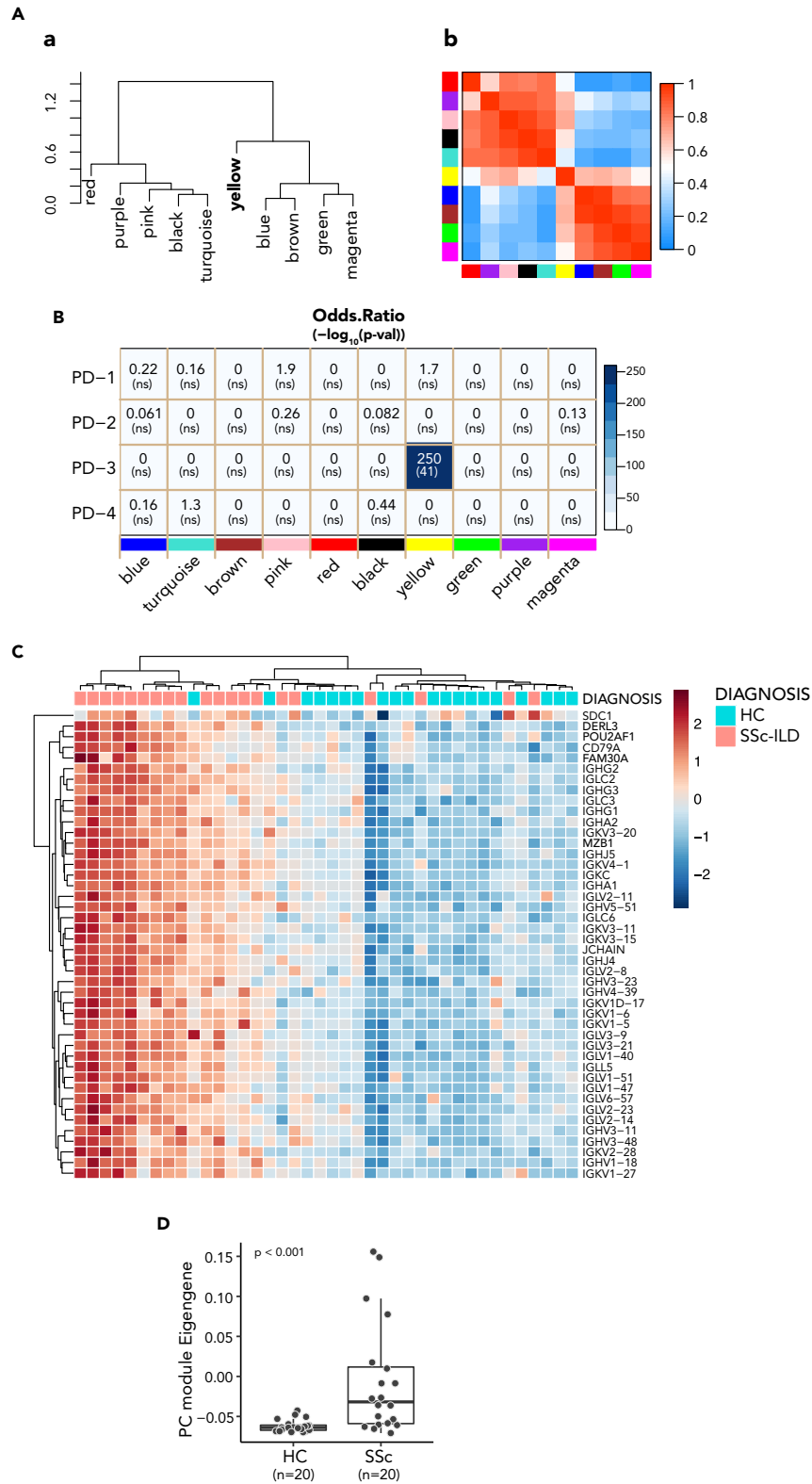
### Skin PC signature is enriched in lungs of SSc-ILD patients

MZB1 protein and MZB1-positive PC have been reported to be enriched in both human lung and skin fibrosis,<sup>22</sup> suggesting that PC accumulation as a common underlying pathological phenomenon in both skin and lung of SSc patients.

To build upon these observations, we generated RNAseq data from explanted SSc-ILD and control lung tissue and determined the DEG between the two groups. We subjected this subset of genes to WGCNA analysis to identify gene co-expression modules, coded by color names (Figure 5A). The modules aggregated into two main groups with network adjacency matrix, while one module (labeled yellow) was less strongly correlated with either of the two main groups.

We performed gene similarity analysis of pathways<sup>23</sup> to compare the lung co-expression modules to the skin PD modules to assess whether any were related to skin gene sets of interest. The overall results from the overlap analysis are shown in an odds ratio heatmap (Figure 5B). Intriguingly, the yellow module was identified in lung as most overlapping with the skin PC module (PD-3), whereas none of the other lung modules significantly overlapped with any of the four skin PD modules. The yellow module comprised 195 genes, among which 133 genes encode immunoglobulin genes, in addition to many other genes relevant to B cell and PC biology.





### Figure 5. Skin PC signature enriched in SSc-ILD lung

(A) Relationship between WGCNA derived module eigengenes from RNAseq of SSc-ILD lung.

(a) The eigengene dendrogram of DEG genes on hierarchical clustering of adjacency-based dissimilarity, and (b) relationship of modules on eigengene adjacency heatmap. The modules assigned in WGCNA were either text labeled in (a) or color labeled on both axes of the heatmap in (b), respectively. Colors in heatmap are from low adjacency (blue) to high adjacency (red).

(B) Heatmap showing the overall results from overlap analysis between the module genes from SSc-ILD lung and PD modules from SSc skin. Blue color gradient represents the odds ratio score. The number on the top in each square box of heatmap represents the odds ratio value from the analysis, while the number in the parentheses at the bottom represents the p value in  $-\log_{10}$ . The color labeled modules derived from the RNAseq of SSc-ILD lung are on the x axis and 4 PD modules from skin RNAseq are on y axis. ns represents no statistical significance with a cutoff of  $p < 0.05$ .

(C) Heatmap of elevated PC gene expression in RNAseq of SSc-ILD lungs compared with control lungs. PC module genes ( $n = 44$ ) in rows and subjects in columns were hierarchically clustered and the gene expression was transformed and standardized to Zscore. The conventional PC marker CD138 (encoded by SDC1) was included as a reference displayed on the top of row in the heatmap. The four PC signature genes are highlighted in bold in the label.

(D) Elevated PC ME in lung of SSc-ILD. PC ME of SSc-ILD lung was compared with controls. p value shown on the top of plot was derived from Wilcoxon test.

We performed a hierarchical clustering of the gene expression levels of the skin PC module ( $n = 44$ ) in lungs from subjects with SSc and controls (Figure 5C). These genes were coordinately expressed at higher levels in a majority of SSc lungs analyzed, relative to controls. Consistent with the expression of PC module in skin, the skin PC ME was significantly higher in SSc lungs compared to controls (Figure 5D).

It was infeasible given the limitations of the study design to sample lung tissue longitudinally in the context of TCZ treatment to determine whether the PD effect on the PC signature observed in skin is also evident in lung. However, given the evidence that the PC signature is elevated in both fibrotic lung and skin tissue and modulated by TCZ treatment in skin, and that TCZ treatment significantly stabilized lung function in SSc-ILD, it is plausible that increased tissue PC activity might actively contribute to lung disease progression in ILD.

Given the increased PC signature in skin and lung tissues, we explored whether the PC related transcripts were elevated in peripheral blood of SSc patients as well. Analysis of data from RNAseq of whole blood samples collected from the FaSScinate cohort at baseline showed that tissue PC signature genes were not co-expressed in a similar manner in blood as in tissues. Nevertheless, some PC genes (such as DERL3 and POU2AF1) were down-regulated in SSc-ILD blood compared the controls (Figure S5). The tissue PC module we discovered in response to IL6 inhibition may therefore be distinct from that of circulating leukocytes, which may reflect differential recruitment, retention, and/or redistribution of these cells in affected tissues.

### Skin PC signature from SSc-ILD is also enriched in IPF ILD lungs

Next we evaluated whether the increased expression of PC module genes was unique to SSc lungs, or if the feature is readily apparent in other fibrotic interstitial lung diseases as well. For this, we compared RNAseq of lung tissue explants from idiopathic pulmonary fibrosis (IPF) patients and that of controls, and analyzed the DEG in IPF patients compared to controls using WGCNA (Figure 6A). Interestingly, akin to SSc-ILD lungs, gene modules in IPF lungs aggregated into two major groups plus one module (yellow) that was less related to either of the major groups. The IPF yellow module was highly similar to the corresponding module in SSc lungs, displaying a large odds ratio and low p value (Figure 6B). This suggests that PC accumulation may not be a unique feature of lung fibrosis only in the autoimmune setting. Indeed, IPF lungs showed significantly higher expression of the PC signature (Figure 6C), suggesting that PCs might play a role in pathogenic processes in IPF as well.

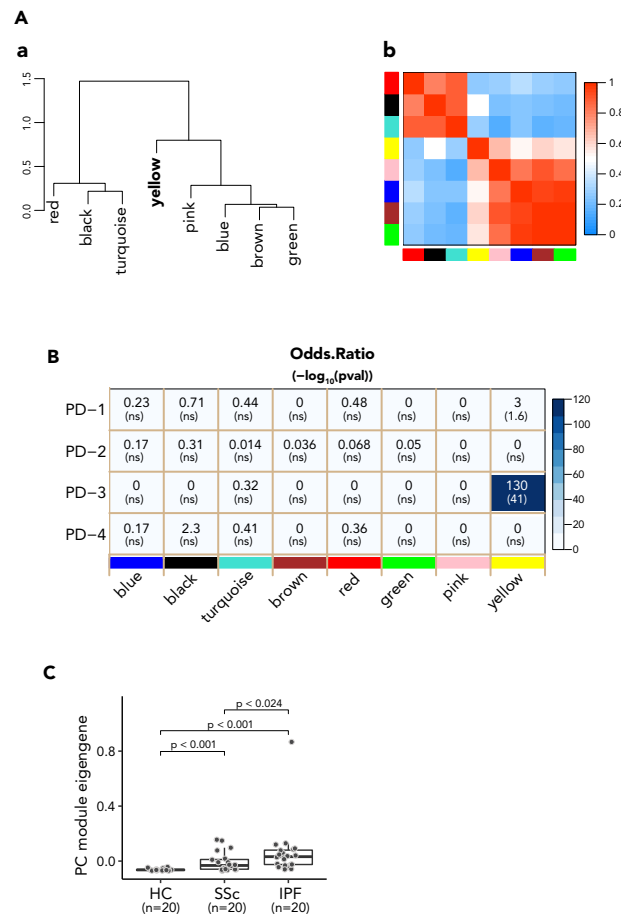
### Confirmation of plasma cells in ILD lungs by scRNAseq

To better delineate the cell population with PC signature and their changes in cellular phenotypes associated with fibrotic diseases, we examined scRNAseq of explanted lungs from ILD patients and controls. scRNAseq analysis showed that small distinct subsets (PC1 and PC2) expressed PC signature genes visualized in single-cell atlas with UMAP plot (Figure 7A) and stacked violin plot (Figure 7B). Their expression levels were higher in SSc-ILD lungs than in control lung (Figure 7B; Figure S7A, Data S3), especially in the PC2 subset which was only detected in SSc-ILD lung (Figure S7B). The clusters of cells also expressed CD38 and CD138 (gene name: SDC1) but were low or negative for CD19 (not shown) and CD20 (gene name: MS4A1) (Figure 7B), which is distinct from conventional CD19 and CD20 positive B lymphocyte cells (Figure 7A). It confirmed that a PC population was present in lung tissue that may contribute to the pathophysiology of ILD, in agreement with a previous report by Schiller et al.<sup>22</sup>

## DISCUSSION

Although understanding of the biological processes underpinning the pathogenesis of SSc has improved in recent years due to the vast and continuing progress being made in the domains of genomics, transcriptomics, and proteomics, a gap in their translation into safe, effective, and targeted therapies still exists. A way to narrow that gap would be to evaluate molecular changes in tissue samples and patient data collected in the course of rigorously designed clinical trials for further exploratory research. This would accelerate our understanding about the key molecular pathways operating in affected organs, and how they relate to disease severity, progression, and response (or lack of thereof) to any experimental treatments being studied.

Here, we sought to understand which molecular pathways changed in skin tissue in response to therapeutic blockade of IL-6 signaling in the context of two clinical trials, and how that information could be extrapolated to potentially address the unmet need in SSc. We used available skin punch biopsy samples from individuals enrolled and appropriately consented in the focuSSc cohort as a discovery set, and



**Figure 6. Skin PC signature enriched in IPF lung**

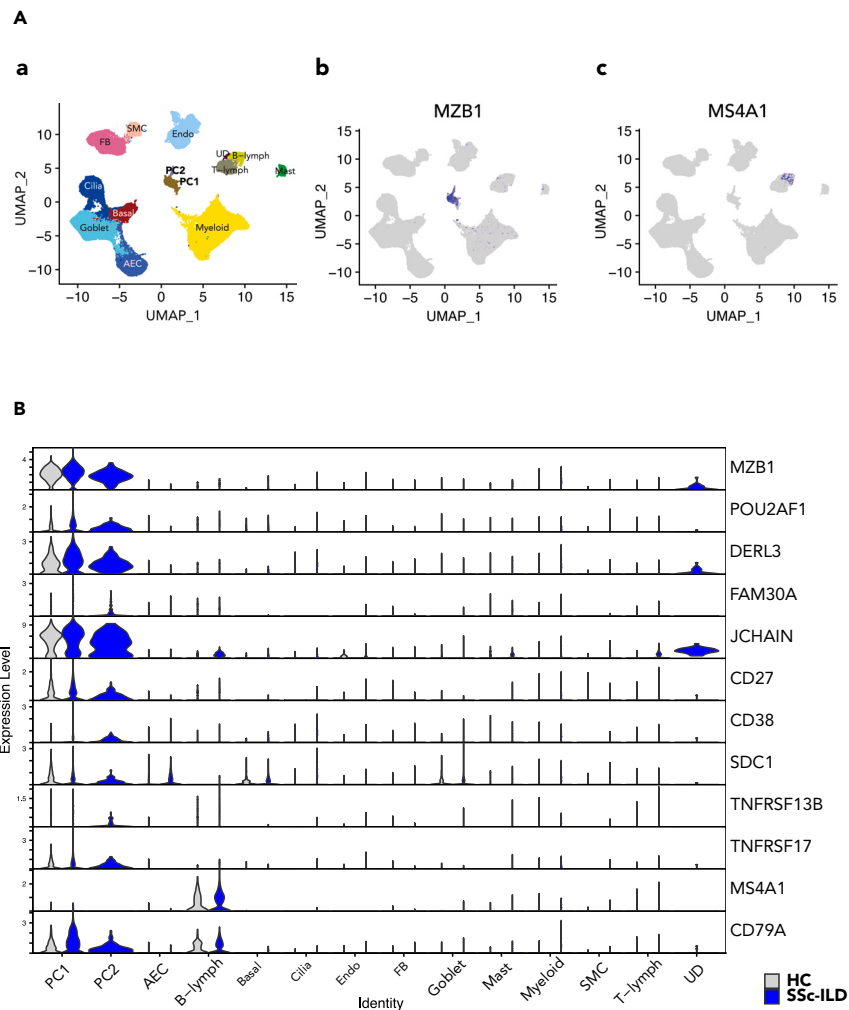
(A) Relationship between WGCNA derived module eigengenes from RNAseq of IPF lung. (a) The eigengene dendrogram of DEG genes on hierarchical clustering of adjacency-based dissimilarity, and (b) relationship of modules on eigengene adjacency heatmap depicted tight clusters of correlated eigengenes of modules. The modules assigned in WGCNA, were either text labeled in (a) or color labeled on both axes of the heatmap in (b), respectively. Colors in heatmap are from low adjacency (blue) to high adjacency (red).

(B) Heatmap showing the overall results from overlap analysis between the module genes from IPF lung and PD modules from SSc skin. Blue color gradient represents the score of odds ratios. The number on the top in each square box of heatmap represents the odds ratio of overlapping from the analysis, while the number in parentheses at the bottom represents the p value in  $-\log_{10}$ . The color labeled modules derived from the RNAseq of IPF lung are present on x axis and 4 PD modules from skin RNAseq are on y axis. ns represents no statistical significance with a cutoff of  $p < 0.05$ .

(C) PC ME in control and ILD lungs. PC MEs of SSc and IPF lungs were compared to each other and to controls ( $n = 20$  in each category as shown). P value shown on the top of plot was derived from Wilcoxon test.

performed an unbiased whole transcriptome analysis by clustering correlated genes together as modules using the WGCNA approach. This analysis for pharmacodynamic changes showed a PC enriched module that was significantly attenuated in response to TCZ treatment, which we confirmed in samples from the faSScinatate trial.

PCs are long-lived tissue-resident antibody secreting cells that generally reside in bone marrow, but are also reported to persist within chronically inflamed tissue such as in kidneys of patients with systemic lupus erythematosus (SLE),<sup>24</sup> central nervous system of patients with multiple sclerosis,<sup>25</sup> spleens of patients with auto-immune thrombocytopenia,<sup>26</sup> inflamed synovial tissue of rheumatoid arthritis patients<sup>27,28</sup> or thymus tissue of myasthenia gravis patients.<sup>29</sup> Long-lived autoreactive PCs implicated in persistent autoimmune inflammation have been reviewed by F. Hiepe et al.<sup>30</sup> Transcriptomic analyses have not only confirmed the existence of PCs in inflamed tissue but also provided some quantitative relationships with clinical metrics beyond the scope of what is achievable by histopathologic characterization alone. For instance, a transcriptomically derived PC signature was shown to be elevated in the peripheral blood and skin tissue of SSc patients.<sup>31</sup> While some genes used for the published PC signature are common with our tissue derived PC signature (e.g., JCHAIN), there are differences as well, attributable to the differences in how they were derived. Of note, we did not observe a PC signature in blood that correlated with tissue PC signatures (Figure S5).



**Figure 7. Confirmation of plasma cells in ILD lungs by scRNAseq**

(A) UMAP atlas of scRNAseq from explanted lung cells depicted an MZB1 expressing subset. The annotated subsets of SSc-ILD lung cells are shown on the left (a) as the reference for cell type position on the UMAP atlas and MZB1 positive (b) and MS4A1 (encoding CD20) (c) cell subsets are depicted with a color spectrum. AEC: alveolar epithelial cells; FB: fibroblasts; SMC: smooth muscle cells; Endo: endothelium; UD: undetermined cells; PC1/PC2, plasma cells.

(B) Skin PC signature genes are distinctly expressed in plasma cells in SSc-ILD lung tissue. The gene expression of interest and relevant to B cell development stages are presented in a stacked violin plot in comparison between the control and SSc, where genes related to specific B cell sub-types are listed on the right of y axis and the annotated clusters on x axis. The colors represent the sources of RNAseq, blue from SSc-ILD lung (SSc-ILD), whereas gray from the healthy control (HC). TNFRSF17: encoding BCMA; TNFRSF13B: encoding TAC1.

Production of pathogenic autoantibodies by self-reactive PCs is a hallmark of autoimmune diseases. We do not have direct data to support that the PCs in the skin tissue are pathogenic or produce autoantibodies. However, although none of the commonly tested systemic autoantibody titers were individually significantly associated with the skin PC signature level, a composite index derived by the sum of the individual autoantibody positivity for each subject, was associated with skin PC level, such that subjects scoring a composite index of two manifested a skin PC signature that was statistically significantly higher than those with a score of one or zero. This suggests that tissue PC score may reflect the totality of the activities of multiple different autoreactive PC (Figure S6). Because the entire spectrum of autoAb specificities in SSc is unknown and not conventionally measurable, a molecular index of total PC activity in tissue as determined by gene expression may provide complementary information to measuring a particular subset of serum autoAb titers. Furthermore, while PCs produce antibodies, they also can produce inflammatory and pro-fibrotic cytokines<sup>32</sup> which could contribute to tissue pathophysiology independently of antibody-mediated effects. However, this does not explain the disconnect between clinical effects of TCZ treatment on skin and lung outcomes. The lack of available pre- and post-treatment tissue samples from lung precludes a definitive assessment of whether the PD effect on PCs observed in skin is also true in lung tissue. The antibody-independent contributions of PCs to SSc tissue pathology and their roles in ILD represent areas for future investigation.

The putative autoimmune etiology of SSc-ILD is consistent with our observation that a tissue PC signature is elevated and associated with lung function progression. IPF, on the other hand, is associated with compromised regenerative capacity of alveolar epithelium and is not considered an autoimmune disease (though certain auto-antibodies have been reported in some patients<sup>33–36</sup>), and affects a demographically different population, with a more progressive trajectory. IPF and SSc-ILD, while displaying certain degrees of overlap, are generally differentiated both histopathologically and radiologically (UIP pattern vs. NSIP pattern, respectively). Therefore, our data confirming prior observations that PCs are enriched across multiple fibrotic ILDs despite different pathological patterns imply that dysregulated PC biology may represent a convergent pathological outcome of distinct inciting events. Tissue PCs may contribute to the progression of these apparently disparate ILDs by perpetuating a chronic inflammatory reaction that contributes to persistent tissue injury, fomenting fibrogenesis.

Interleukin-6 (IL-6) is a key regulator of B cell differentiation into PCs and exerts a significant impact on the pathogenesis of PC disorders, including plasmacytoma and myeloma.<sup>37</sup> Perturbed IL-6 production has been implicated in the development of polyclonal PC abnormalities and PC neoplasias.<sup>38,39</sup> Furthermore, IL-6 has been shown to play a crucial role in supporting the generation of long-lived PCs.<sup>40,41</sup>

Our identified distinct gene signature is enriched for B cell-related genes expressed in various B cell subsets, respectively. While we have referred to this signature as a “PC module,” we acknowledge that it may also reflect B cell activation and the infiltration of activated B cells into SSc tissues. To our knowledge, our analyses demonstrate for the first time that the tissue PC are not only associated with SSc disease at baseline but may also carry prognostic information to help predict clinical progression, and mechanistically link tissue PC levels to endogenous IL-6 activity—a relationship long suspected based on *in vitro* and preclinical observations but not previously demonstrated in human disease via a rigorously controlled, double blind randomized clinical trial setting with a selective IL-6 inhibitor.

Despite the limitations of small sample size from skin biopsies due to the patient consents in focuSSced and the post hoc nature of our study, a plasma cell-associated gene signature measured on a different technical platform in faSScinatc yielded similar pharmacodynamic effects. Furthermore, the biological plausibility of our findings is supported by preexisting literature.<sup>26,30</sup> Collectively, these features underscore the robustness of the finding, which may provide insight into the contribution of PCs and their regulation by IL-6 in fibrotic diseases. Based on our findings, it is possible that targeting PC numbers and/or function via different mechanisms could represent a viable interventional strategy to ameliorate disease progression in SSc and related fibrotic diseases.

### Limitations of the study

This study has some potential limitations and caveats. The PD effect by Tocilizumab primarily observed in the study is based on the skin biopsies from the clinical trials. The size of samples is limited by the consent of the patients and the original trial design. It is therefore subject to the statistical power and biases in analysis that may have influenced the effect estimates, although the effect has been demonstrated in two independent clinical cohorts in the study. The verification of the plasma gene signature and their change with additional approaches such as IHC imaging which would better support the findings is compromised by the tissue availability since the skin biopsies collected from the clinical trials conducted many years ago have been either lysed to retrieve RNA or destroyed according to the trial protocol by the time we performed the analyses.

Given current techniques to isolate intact PCs from tissues made it challenging to confirm specific tissue B-lineage subsets with the signature, emerging technique such as spatial transcriptomics and proteomics may enable it to be addressed in future studies.

### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- METHOD DETAILS
  - RNAseq dataset
  - Identification of DEGs in SSc skin and ILD lung in compared with the controls
  - Identification of DEGs by the effect of TCZ treatment in SSc skin
  - Construction of the gene co-expression network affected by TCZ treatment
  - Construction of the gene co-expression network in ILD lungs
  - Pathway enrichment analysis
  - Gene overlap analysis for the conserved PC module
  - Heatmap of genes in clustering
  - scRNAseq analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS
- ADDITIONAL RESOURCES

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2023.108133>.

## ACKNOWLEDGMENTS

We acknowledge all the patients and families who participated in the focuSSced and faSScinate clinical trials, and the clinical team who supported and implemented these trials. In addition, we thank Dr. M. Townsend for thoughtful discussion and help with editing, and D. Nicetto for editing. All studies were supported by Genentech.

## AUTHOR CONTRIBUTIONS

Conceptualization, G.J., T.R.R., and J.R.A.; methodology, G.J.; investigation, G.J.; writing – original draft, G.J. and T.R.R.; writing – review & editing, G.J., T.R.R., J.V.H., P.W., D.K., and J.R.A.; data curation: J.V.H.; resources, X.G., D.D., K.B.M., N.R., and Z.M.; supervision, C.L. and D.K.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 7, 2023

Revised: August 14, 2023

Accepted: September 30, 2023

Published: October 5, 2023

## REFERENCES

- Allanore, Y., Simms, R., Distler, O., Trojanowska, M., Pope, J., Denton, C.P., and Varga, J. (2015). Systemic sclerosis. *Nat. Rev. Dis. Primers* 1, 15002. <https://doi.org/10.1038/nrdp.2015.2>.
- Khanna, D., Lescoat, A., Roofeh, D., Bernstein, E.J., Kazerooni, E.A., Roth, M.D., Martinez, F., Flaherty, K.R., and Denton, C.P. (2022). Systemic Sclerosis–Associated Interstitial Lung Disease: How to Incorporate Two Food and Drug Administration–Approved Therapies in Clinical Practice. *Arthritis Rheumatol.* 74, 13–27. <https://doi.org/10.1002/art.41933>.
- Nikpour, M., and Baron, M. (2014). Mortality in systemic sclerosis. *Curr. Opin. Rheumatol.* 26, 131–137. <https://doi.org/10.1097/bor.000000000000027>.
- Giuggioli, D., Lumetti, F., Colaci, M., Fallahi, P., Antonelli, A., and Ferri, C. (2015). Rituximab in the treatment of patients with systemic sclerosis. Our experience and review of the literature. *Autoimmun. Rev.* 14, 1072–1078. <https://doi.org/10.1016/j.autrev.2015.07.008>.
- Bosello, S.L., Basile, U., Lorenzis, E.D., Canestrari, G., Parisi, F., Rucco, M., Birra, D., Gulli, F., Napodano, C., Pocino, K., et al. (2016). FRI0238 Increased Serum Free Light Chains of Immunoglobulins in Systemic Sclerosis Patients: Correlation with Lung Involvement and Inflammatory Milieu. *Ann. Rheum. Dis.* 75, 520. <https://doi.org/10.1136/annrheumdis-2016-eular.5735>.
- Kawaguchi, Y. (2017). Contribution of Interleukin-6 to the Pathogenesis of Systemic Sclerosis. *J. Scleroderma Relat. Disord.* 2, S6–S12. <https://doi.org/10.5301/jsrd.5000258>.
- Uchiyama, Y., Yorozu, K., Hashizume, M., Moriya, Y., and Mihara, M. (2008). Tocilizumab, a humanized anti-interleukin-6 receptor antibody, ameliorates joint swelling in established monkey collagen-induced arthritis. *Biol. Pharm. Bull.* 37, 1159–1163. <https://doi.org/10.1248/bpb.31.1159>.
- Khanna, D., Lin, C.J.F., Furst, D.E., Goldin, J., Kim, G., Kuwana, M., Allanore, Y., Matucci-Cerinic, M., Distler, O., Shima, Y., et al. (2020). Tocilizumab in systemic sclerosis: a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Respir. Med.* 8, 963–974. [https://doi.org/10.1016/s2213-2600\(20\)30318-0](https://doi.org/10.1016/s2213-2600(20)30318-0).
- Khanna, D., Jahreis, A., and Furst, D.E. (2017). Tocilizumab Treatment of Patients with Systemic Sclerosis: Clinical Data. *J. Scleroderma Relat. Disord.* 2, S29–S35. <https://doi.org/10.5301/jsrd.5000267>.
- Khanna, D., Furst, D.E., Clements, P.J., Allanore, Y., Baron, M., Czirkaj, L., Distler, O., Foeldvari, I., Kuwana, M., Matucci-Cerinic, M., et al. (2017). Standardization of the modified Rodnan skin score for use in clinical trials of systemic sclerosis. *J. Scleroderma Relat. Disord.* 2, 11–18. <https://doi.org/10.5301/jsrd.5000231>.
- Hinchcliff, M., Huang, C.-C., Wood, T.A., Matthew Mahoney, J., Martyanov, V., Bhattacharyya, S., Tamaki, Z., Lee, J., Carns, M., Podlasky, S., et al. (2013). Molecular Signatures in Skin Associated with Clinical Improvement During Mycophenolate Treatment in Systemic Sclerosis. *J. Invest. Dermatol.* 133, 1979–1989. <https://doi.org/10.1038/jid.2013.130>.
- Milano, A., Pendergrass, S.A., Sargent, J.L., George, L.K., McCalmont, T.H., Connolly, M.K., and Whitfield, M.L. (2008). Molecular Subsets in the Gene Expression Signatures of Scleroderma Skin. *PLoS One* 3, e2696. <https://doi.org/10.1371/journal.pone.0002696>.
- Pendergrass, S.A., Lemaire, R., Francis, I.P., Mahoney, J.M., Lafyatis, R., and Whitfield, M.L. (2012). Intrinsic Gene Expression Subsets of Diffuse Cutaneous Systemic Sclerosis Are Stable in Serial Skin Biopsies. *J. Invest. Dermatol.* 132, 1363–1373. <https://doi.org/10.1038/jid.2011.472>.
- Khanna, D., Denton, C.P., Jahreis, A., van Laar, J.M., Frech, T.M., Anderson, M.E., Baron, M., Chung, L., Fierlbeck, G., Lakshminarayanan, S., et al. (2016). Safety and efficacy of subcutaneous tocilizumab in adults with systemic sclerosis (faSScinate): a phase 2, randomised, controlled trial. *Lancet Lond Engl* 387, 2630–2640. [https://doi.org/10.1016/s0140-6736\(16\)00232-4](https://doi.org/10.1016/s0140-6736(16)00232-4).
- Andreani, V., Ramamoorthy, S., Pandey, A., Lupar, E., Nutt, S.L., Lämmermann, T., and Grosschedl, R. (2018). Cochaperone Mzb1 is a key effector of Blimp1 in plasma cell differentiation and  $\beta$ 1-integrin function. *P Natl Acad Sci Usa* 115, E9630–E9639. <https://doi.org/10.1073/pnas.1809739115>.
- Auer, R.L., Starczynski, J., McElwaine, S., Berton, F., Newland, A.C., Fegan, C.D., and Cotter, F.E. (2005). Identification of a potential role for POU2AF1 and BTG4 in the deletion of 11q23 in chronic lymphocytic leukemia. *Genes Chromosomes Cancer* 43, 1–10. <https://doi.org/10.1002/gcc.20159>.
- Chyra, Z., Samur, M.K., Aktas-Samur, A., Yao, Y., Derebail, S., Perini, T., Xu, Y., Morelli, E., Adamia, S., Park, W.D., et al. (2021). B Cell Transcriptional Coactivator POU2AF1 (BOB-1) Is an Early Transcription Factor Modulating the Protein Synthesis and Ribosomal Biogenesis in Multiple Myeloma: With Therapeutic Implication. *Blood* 138, 2670. <https://doi.org/10.1182/blood-2021-152121>.
- Morton, V., and Torgerson, D.J. (2005). Regression to the mean: treatment effect without the intervention. *J. Eval. Clin. Pract.* 11, 59–65. <https://doi.org/10.1111/j.1365-2753.2004.00505.x>.
- Barnett, A.G., van der Pols, J.C., and Dobson, A.J. (2005). Regression to the mean: what it is and how to deal with it. *Int. J. Epidemiol.* 34, 215–220. <https://doi.org/10.1093/ije/dyh299>.
- Denton, C.P., Yee, P., Ong, V.H., Yee, P., and Ong, V.H. (2020). News and failures from recent treatment trials in systemic sclerosis. *Eur. J. Rheumatol.* 7, 242–S248. <https://doi.org/10.5152/eurjrh.2020.19187>.
- Goh, N.S., Hoyle, R.K., Denton, C.P., Hansell, D.M., Renzoni, E.A., Maher, T.M., Nicholson, A.G., and Wells, A.U. (2017). Short-Term

- Pulmonary Function Trends Are Predictive of Mortality in Interstitial Lung Disease Associated With Systemic Sclerosis. *Arthritis Rheumatol.* 69, 1670–1678. <https://doi.org/10.1002/art.40130>.
22. Schiller, H.B., Mayr, C.H., Leuschner, G., Strunz, M., Staab-Weijnitz, C., Preisendörfer, S., Eckes, B., Moinzadeh, P., Krieg, T., Schwartz, D.A., et al. (2017). Deep Proteome Profiling Reveals Common Prevalence of MZB1-Positive Plasma B Cells in Human Lung and Skin Fibrosis. *Am J Resp Crit Care* 196, 1298–1310. <https://doi.org/10.1164/rccm.201611-2263oc>.
23. Taroni, J.N., Greene, C.S., Martyanov, V., Wood, T.A., Christmann, R.B., Farber, H.W., Lafyatis, R.A., Denton, C.P., Hinchcliff, M.E., Pioli, P.A., et al. (2017). A novel multi-network approach reveals tissue-specific cellular modulators of fibrosis in systemic sclerosis. *Genome Med.* 9, 27. <https://doi.org/10.1186/s13073-017-0417-1>.
24. Starke, C., Frey, S., Wellmann, U., Urbonaviciute, V., Schett, G., Winkler, T., and Voll, R.E. (2011). Autoantibody secreting cells and long-lived plasma cells are enriched within inflamed kidneys of NZB/NZW F1 lupus mice and sensitive to bortezomib treatment. *Ann. Rheum. Dis.* 70, A80. <https://doi.org/10.1136/ard.2010.149013.29>.
25. Racke, M.K. (2008). The role of B cells in multiple sclerosis: rationale for B-cell-targeted therapies. *Curr. Opin. Neurol.* 21, S9–S18. <https://doi.org/10.1097/01.wco.0000313359.61176.15>.
26. Mahévas, M., Michel, M., Weill, J.-C., and Reynaud, C.-A. (2013). Long-Lived Plasma Cells in Autoimmunity: Lessons from B-Cell Depleting Therapy. *Front. Immunol.* 4, 494. <https://doi.org/10.3389/fimmu.2013.00494>.
27. Cole, S., Walsh, A., Yin, X., Wechalekar, M.D., Smith, M.D., Proudman, S.M., Veale, D.J., Fearon, U., Pitzalis, C., Humby, F., et al. (2018). Integrative analysis reveals CD38 as a therapeutic target for plasma cell-rich pre-disease and established rheumatoid arthritis and systemic lupus erythematosus. *Arthritis Res. Ther.* 20, 85. <https://doi.org/10.1186/s13075-018-1578-z>.
28. Dong, W., Li, X., Liu, H., and Zhu, P. (2009). Infiltrations of plasma cells in synovium are highly associated with synovial fluid levels of APRIL in inflamed peripheral joints of rheumatoid arthritis. *Rheumatol. Int.* 29, 801–806. <https://doi.org/10.1007/s00296-008-0773-7>.
29. Hill, M.E., Shiono, H., Newsom-Davis, J., and Willcox, N. (2008). The myasthenia gravis thymus: A rare source of human autoantibody-secreting plasma cells for testing potential therapeutics. *J. Neuroimmunol.* 201–202, 50–56. <https://doi.org/10.1016/j.jneuroim.2008.06.027>.
30. Hiepe, F., Dörner, T., Hauser, A.E., Hoyer, B.F., Mei, H., and Radbruch, A. (2011). Long-lived autoreactive plasma cells drive persistent autoimmune inflammation. *Nat. Rev. Rheumatol.* 7, 170–178. <https://doi.org/10.1038/nrrheum.2011.1>.
31. Streicher, K., Morehouse, C.A., Groves, C.J., Rajan, B., Pilataxi, F., Lehmann, K.P., Brohawn, P.Z., Higgs, B.W., McKeever, K., Greenberg, S.A., et al. (2014). The Plasma Cell Signature in Autoimmune Disease: Gene Expression Signature to Monitor Plasma Cells. *Arthritis Rheumatol.* 66, 173–184. <https://doi.org/10.1002/art.38194>.
32. Thoreau, B., Chaigne, B., and Mouthon, L. (2022). Role of B-Cell in the Pathogenesis of Systemic Sclerosis. *Front. Immunol.* 13, 933468. <https://doi.org/10.3389/fimmu.2022.933468>.
33. Morita, H., Shimizu, Y., Nakamura, Y., Okutomi, H., Watanabe, T., Yokoyama, T., Soda, S., Ikeda, N., Shiobara, T., Miyoshi, M., et al. (2020). Auto-antibody evaluation in idiopathic interstitial pneumonia and worse survival of patients with Ro52/TRIM21 auto-antibody. *J. Clin. Biochem. Nutr.* 67, 199–205. <https://doi.org/10.3164/jcbn.20-5>.
34. Mignot, S., Taille, C., Besnard, V., Crestani, B., and Chollet-Martin, S. (2014). Detection of anti-periplakin auto-antibodies during idiopathic pulmonary fibrosis. *Clin. Chim. Acta* 433, 242. <https://doi.org/10.1016/j.cca.2014.03.027>.
35. Roh, J., Song, J., and Kim, D. (2009). Prevalence and Clinical Significance of Auto-Antibodies in Patients with Idiopathic Pulmonary Fibrosis. C45 Ild Connect Tissue Disord Ii, A4505. 10.1164/ajrccm-conference.2009.179.1\_ meetingabstracts.a4505
36. Adler, B.L., Boin, F., Wolters, P.J., Bingham, C.O., Shah, A.A., Greider, C., Casciola-Rosen, L., and Rosen, A. (2021). Autoantibodies targeting telomere-associated proteins in systemic sclerosis. *Ann. Rheum. Dis.* 80, 912–919. <https://doi.org/10.1136/annrheumdis-2020-218918>.
37. Hong, D.S., Angelo, L.S., and Kurzrock, R. (2007). Interleukin-6 and its receptor in cancer: Implications for Translational Therapeutics. *Cancer* 110, 1911–1928. <https://doi.org/10.1002/cncr.22999>.
38. Hirano, T., and Kishimoto, T. (1989). Interleukin 6 and plasma cell neoplasias. *Prog. Growth Factor Res.* 1, 133–142. [https://doi.org/10.1016/0955-2235\(89\)90007-0](https://doi.org/10.1016/0955-2235(89)90007-0).
39. Burger, R., Wandler, J., Antoni, K., Helm, G., Kalden, J.R., and Gramatzki, M. (1994). Interleukin-6 production in B-cell neoplasias and Castleman's disease: Evidence for an additional paracrine loop. *Ann. Hematol.* 69, 25–31. <https://doi.org/10.1007/bf01757344>.
40. van Laar, J.M., Breedveld, F.C., Melchers, M., Teng, Y.K.O., van der Zouwen, B., Mohammadi, R., Fischer, R., Margolis, L., Fitzgerald, W., et al. (2007). Sustained Secretion of Immunoglobulin by Long-Lived Human Tonsil Plasma Cells. *Am. J. Pathol.* 171, 917–927. <https://doi.org/10.2353/ajpath.2007.070005>.
41. Jourdan, M., Cren, M., Robert, N., Bolloré, K., Fest, T., Duperray, C., Guilloton, F., Høse, D., Tarte, K., and Klein, B. (2014). IL-6 supports the generation of human long-lived plasma cells in combination with either APRIL or stromal cell-soluble factors. *Leukemia* 28, 1647–1656. <https://doi.org/10.1038/leu.2014.61>.
42. Stifano, G., Sornasse, T., Rice, L.M., Na, L., Chen-Harris, H., Khanna, D., Jahreis, A., Zhang, Y., Siegel, J., and Lafyatis, R. (2018). Skin Gene Expression Is Prognostic for the Trajectory of Skin Disease in Patients With Diffuse Cutaneous Systemic Sclerosis. *Arthritis Rheumatol.* 70, 912–919. <https://doi.org/10.1002/art.40455>.
43. Gao, X., Jia, G., Guttman, A., DePianto, D.J., Morshead, K.B., Sun, K.-H., Ramamoorthi, N., Vander Heiden, J.A., Modrusan, Z., Wolters, P.J., et al. (2020). Osteopontin Links Myeloid Activation and Disease Progression in Systemic Sclerosis. *Cell Rep. Med.* 1, 100140. <https://doi.org/10.1016/j.xcrm.2020.100140>.
44. Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf.* 9, 559. <https://doi.org/10.1186/1471-2105-9-559>.
45. Botía, J.A., Vandrovцова, J., Forabosco, P., Guelfi, S., D'Sa, K., United Kingdom Brain Expression Consortium, Hardy, J., Lewis, C.M., Ryten, M., and Weale, M.E. (2017). An additional k-means clustering step improves the biological features of WGCNA gene co-expression networks. *BMC Syst. Biol.* 11, 47. <https://doi.org/10.1186/s12918-017-0420-6>.
46. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. <https://doi.org/10.1186/s13059-014-0550-8>.
47. Langfelder, P., and Horvath, S. (2012). Fast R Functions for Robust Correlations and Hierarchical Clustering. *J. Stat. Softw.* 46, i11.
48. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43, e47. <https://doi.org/10.1093/nar/gkv007>.
49. Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *Omics J Integr Biology* 16, 284–287. <https://doi.org/10.1089/omi.2011.0118>.
50. Yu, G., and He, Q.-Y. (2016). ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Mol. Biosyst.* 12, 477–479. <https://doi.org/10.1039/c5mb00663e>.
51. Mahoney, J.M., Taroni, J., Martyanov, V., Wood, T.A., Greene, C.S., Pioli, P.A., Hinchcliff, M.E., and Whitfield, M.L. (2015). Systems Level Analysis of Systemic Sclerosis Shows a Network of Immune and Proinflammatory Pathways Connected with Genetic Polymorphisms. *PLoS Comput. Biol.* 11, e1004005. <https://doi.org/10.1371/journal.pcbi.1004005>.
52. Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* 36, 411–420. <https://doi.org/10.1038/nbt.4096>.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Deposited data</b>		
bulk RNAseq of skin biopsy from focuSSced	This paper	GSE231694
Microarray of skin biopsy from faSScinate	Stifano et al. <sup>42</sup>	GSE106358
bulk RNAseq of whole blood PAXgene from faSScinate	This paper	GSE231694
bulk RNAseq of explanted lung tissue from ILD patients	This paper	GSE231694
scRNAseq of fibrotic lungs	Gao et al. <sup>43</sup>	GSE159354
<b>Software and algorithms</b>		
R v 4.0.2		<a href="http://www.R-project.org">http://www.R-project.org</a>
RStudio		<a href="http://www.rstudio.com/">http://www.rstudio.com/</a>
DESeq2 package	Bioconductor	<a href="https://bioconductor.org/packages/release/bioc/html/DESeq2.html">https://bioconductor.org/packages/release/bioc/html/DESeq2.html</a>
WGCNA package	CRAN	<a href="https://cran.r-project.org/web/packages/WGCNA/index.html">https://cran.r-project.org/web/packages/WGCNA/index.html</a>
clusterProfiler package	Bioconductor	<a href="https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html">https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html</a>
GeneOverlap package	Bioconductor	<a href="https://bioconductor.org/packages/release/bioc/html/GeneOverlap.html">https://bioconductor.org/packages/release/bioc/html/GeneOverlap.html</a>
Seurat package	CRAN	<a href="https://cran.r-project.org/web/packages/Seurat/index.html">https://cran.r-project.org/web/packages/Seurat/index.html</a>

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Guiquan Jia ([jjia@gene.com](mailto:jjia@gene.com)).

### Materials availability

This study did not generate new unique reagents.

### Data and code availability

- All RNA-seq data have been deposited at GEO and are publicly available as of the date of publication. Accession numbers are listed in the [key resources table](#).
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available as [supplemental information](#) or from the [lead contact](#) upon request.

## METHOD DETAILS

### RNAseq dataset

The RNA for bulk RNAseq was prepared from the skin biopsies collected from focuSSced cohort (NCT02453256). Briefly, the skin samples in RNeasy Lysis Buffer (Qiagen, Austin, TX, USA) were transferred to TRIzol (ThermoFisher Scientific, Waltham, MA, USA)-filled tubes and homogenized using a Miltenyi gentleMACS™ Dissociator. Afterwards, samples were centrifuged for 3 minutes at 8000g and the supernatants were collected. The total RNA was isolated from the supernatant using the RNeasy Fibrous Tissue mini kit. Resultant total RNA after assessed quantitatively (NanoDrop; ThermoFisher Scientific) and qualitatively (Bioanalyzer; Agilent, Santa Clara, CA, USA) were used for RNAseq.

Bulk RNAseq dataset was prepared by sequencing the RNA on the Illumina HiSeq 4000 platform and followed by aligned and annotated to human genomic database GRCh38, which comprised 58302 features and 76 samples, 62 from 37 subjects, the subset of the patients that consented for pre-post treatment skin biopsies in focuSSced study and 14 from age and gender matched healthy donors. The statistical diagnostic for flagging potential outlying samples was analyzed with the standard procedure of WGCNA<sup>44,45</sup> using all genes available. Three samples as the outliers were removed from the subset of SSc patients in this dataset. Among 36 subjects with SSc after outlier sample removal, 22 subjects had 2 biopsies, each from pre- (baseline) and post-treatment (week 48), where 11 subjects were in placebo (PBO) arm and 11 subjects in the TCZ treatment arm.

Microarray dataset of skin biopsy from faSScinate cohort (NCT01532869) was described in previous publication.<sup>14,42</sup>



Bulk RNAseq dataset of whole blood PAXgene generated in Genentech included 20 samples from patients in faSScinata (NCT01532869) and 10 age matched healthy donors and was aligned and annotated to the genome database GRCh38.

Bulk RNAseq dataset of explanted lung tissue from ILD patients was generated in Genentech using lung samples collected from SSc ILD and IPF patients during transplantation. It included 20 SSc ILD, 20 IPF and 20 healthy controls. It was aligned and annotated to the genome database GRCh38.

scRNAseq dataset was described in previous publications,<sup>43</sup> which included cells of lung tissues from 3 patients with systemic sclerosis and 3 normal controls (GSE159354).

### Identification of DEGs in SSc skin and ILD lung in compared with the controls

DESeq2 package<sup>46</sup> in R was applied to identify DEGs at baseline by comparing expression counts of baseline samples between disease and control. The adjusted p-value < 0.05, |fold change (FC)| ≥ 2, and mean of counts (baseMean, a parameter used in DESeq2) > 25 (about one quartile of the range of mean of counts for genes in RNAseq) were used as cut-off criteria for DEGs selection for further analysis indicated (detail of DEG information available in [supplemental information](#)).

### Identification of DEGs by the effect of TCZ treatment in SSc skin

DEG analysis with DESeq2 package was performed with likelihood ratio test (LRT) in a matrix design of time course analysis with fully-paired samples intra-individually to reach the differential expression (DE) between pre- (baseline, used as the reference) and post-treatment (week 48) within each arm and the gene expression impacted by the true effect of TCZ treatment was obtained by contrasting between two arms with placebo group as the reference. The PD genes were selected based on the criteria as of adjusted p-value < 0.05, FC ≤ -1.5, and mean of counts (baseMean) > 25 (about one quartile of range of mean of counts for genes in RNAseq). The information of PD genes is available in [supplemental information](#).

### Construction of the gene co-expression network affected by TCZ treatment

All PD genes selected on the effect of TCZ treatment and the subset of RNAseq from SSc skin samples in the treatment arm were used in the construction of weighted co-expression networks (modules) using the WGCNA<sup>44,47</sup> procedure with WGCNA R package available on the Comprehensive R Archive Network (<http://cran.r-project.org>). Briefly, the gene expression level in counts were transformed with the function of voom in limma<sup>48</sup> R package first. And then the network was rendered based on a pair-wise correlation using a  $\beta$  of 8 as a weight and 20% genes in the least connection from softConnectivity analysis within the package were excluded from the clustering. And finally co-expressed genes were clustered according to a topological overlap metric with the method of "ward.D2", and the modules assigned in default by colors with the dynamic tree cut algorithm at a minimum module size of 30 for gene number. The genes without annotated gene symbol were excluded (95 out of 427 genes). Module eigengene (ME) was calculated with a function in WGCNA package based on all gene expression within the module/signature and transformed to positive value with formula (ME – min (ME) +1) when required in the process of log transform.

### Construction of the gene co-expression network in ILD lungs

All DEGs from DEG analysis described in relevant section were applied to construct weighted co-expression modules for each case. The subset of RNAseq from the defined disease samples were used in the network construction using the same parameter setups in the procedure of WGCNA described above.

### Pathway enrichment analysis

In order to investigate gene module at the molecular and functional level, we performed Reactome pathways enrichment analysis (and other pathway databases if available) by using clusterProfiler<sup>49</sup> (<http://bioconductor.org/packages/release/bioc/html/clusterProfiler.html>), which calculates and compares enriched functional categories of each gene module.<sup>50</sup> Using gene ratios (the proportions of genes enriched in each category) and adjusted p-values, a dot plot was constructed to visualize the difference in enriched functional categories between the modules.

### Gene overlap analysis for the conserved PC module

The principle to detect module overlap network was elicited previously by Mohoney et al.<sup>51</sup> and Taroni et al.<sup>23</sup> As an alternative practice, R Package: GeneOverlap from BioConductor, was used. Given two sets of gene lists, this package calculated the overlaps between all pairs of lists from the two intersecting sets. Fisher's exact test was used based on the hypergeometric distribution to determine the p-value and odds ratio in comparison to a genomic background. The null hypothesis is that the odds ratio is no larger than 1. In our cases, the genomic background was scoped as the union size of all DEG input for the sets in comparison presented in the result section, respectively. The matrix table with odd ratio and p value were created to represent all pairwise overlap comparisons between the gene lists, where p value was transformed with  $-\log_{10}(x)$ . The p value less than 0.05 was labelled as no significance (n.s.) statistically.

### Heatmap of genes in clustering

A common method of visualizing gene expression as a heatmap was generated with pheatmap function from R package of pheatmap available from CRAN. All gene expression in counts from RNAseq was voom-transformed with limma<sup>48</sup> package for linear modeling and scaled in the heatmap. The genes and subjects in row and column of heatmap, respectively, were hierarchically clustered with the method of ward.D2.

### scRNAseq analysis

The scRNAseq data was processed and analyzed following the standard analytical flow on the website of Seurat<sup>52</sup> R package version 4.0.1. In brief, after counts normalization and IntegrationAnchors identification for each sample dataset, respectively, all sample datasets were integrated together with the functions in Seurat package accordingly. Then we performed graph-based unsupervised clustering of this integrated data using FindNeighbors (reduction = "pca", dims = 1:30) and FindClusters (resolution = 1) function of Seurat and visualized the clusters on uniform manifold approximation and projection (UMAP) plot using RunUMAP function of Seurat based on networks of the shared-nearest neighbor (SNN) graph. Cell types for each cluster were annotated via canonical cell type markers and then merged to the major cell types presented in airway and lung. The expression comparison of the gene of interest was plotted in stacked violin plot with the clusters of interest using the functions included in Seurat package.

Boxplots representing the nonzero percent makeup distributions of individual varieties of immune cell as a proportion of all immune cells per subject within each disease group.

### QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analyses and graphing were performed using R with a variety of relevant packages available either from CRAN or BioConductor. Statistically correlation coefficient and significant differences were assessed by the tests indicated in the graphs. For all analyses, a two-sided p-value less than 0.05 was considered statistically significant.

### ADDITIONAL RESOURCES

focuSSced cohort: NCT02453256. Link: <https://clinicaltrials.gov/study/NCT02453256>

faSScinate cohort: NCT01532869. Link: <https://clinicaltrials.gov/study/NCT01532869>.