Skin Gene Expression Correlates of Severity of Interstitial Lung Disease in Systemic Sclerosis

Shervin Assassi, MD, MS[1], **Minghua Wu, MD, PhD**[1], **Filemon K. Tan, MD, PhD**[1], **Jeffrey Chang, PhD**[1], **Tiffany A. Graham, MD**[1], **Daniel E. Furst, MD**[2], **Dinesh Khanna, MD, MS**[2], **Julio Charles, MS**[1], **Emma C. Ferguson, MD**[1], **Carol Feghali-Bostwick, PhD**[3], and **Maureen D. Mayes, MD, MPH**[1]

[1]University of Texas Health Science Center at Houston

[2](current address: University of Michigan, Ann Arbor): University of California, Los Angeles

[3]University of Pittsburgh, Pittsburgh, Pennsylvania

Abstract

Objective—We undertook this hypothesis-generating study to identify skin transcripts correlating with severity of interstitial lung disease (ILD) in systemic sclerosis (SSc).

Methods—Skin biopsy samples from 59 patients enrolled in the Genetics versus Environment in Scleroderma Outcome Study (GENISOS) cohort or an open-label imatinib study (baseline visit) were examined by global gene expression analysis using Illumina HT-12 arrays. Skin transcripts correlating with concomitantly obtained forced vital capacity (FVC) values and the modified Rodnan skin thickness score (MRSS) were identified by quantitative trait analysis. Also, immunofluorescence staining for selected transcripts was performed in affected skin and lung tissue. Plasma levels of CCL2, soluble SELP, and soluble P-selectin glycoprotein ligand 1 (sPSGL-1) were examined in all patients enrolled in the GENISOS cohort (n = 266).

Results—Eighty-two skin transcripts correlated significantly with FVC. This gene list distinguished patients with more severe ILD (FVC <70% predicted) in unsupervised hierarchical clustering analysis ($P < 0.001$). These genes included SELP, CCL2, and matrix metalloproteinase 3, which are involved in extravasation and adhesion of inflammatory cells. Among the FVC correlates, 8 genes (CCL2, HAPLN3, GPR4, ADCYAP1, WARS, CDC25B, PLP1, and STXBP6) also correlated with the MRSS. Immunofluorescence staining revealed that SELP and CCL2 were also overexpressed in affected skin and lung tissue from SSc patients compared to those from controls. Plasma levels of CCL2 and sPSGL-1 correlated with concomitantly obtained FVC values ($r = -0.22$, $P = 0.001$ and $r = 0.17$, $P = 0.015$, respectively). This relationship was independent of

potential confounders (age, sex, ethnicity, smoking status, anti–topoisomerase I positivity, treatment with immunosuppressive agents, MRSS, disease type, and disease duration).

Conclusion—A limited number of skin transcripts including genes involved in extravasation and adhesion of inflammatory cells correlate with severity of ILD.

Systemic sclerosis (SSc) is characterized by the triad of vasculopathy, immune dysregulation, and fibrosis and is associated with high morbidity and mortality. Interstitial lung disease (ILD) is the primary cause of SSc-related mortality (1,2), and the available treatment options for this disease manifestation have limited efficacy (3,4). Furthermore, the course of lung involvement in SSc is highly variable. Even though deterioration of pulmonary function is slowly progressive in many SSc patients, ~15% have a rapidly progressive course (5). Clinicians are currently unable to predict reliably early in the course of disease which patients will develop significant ILD. Therefore, the treatment of SSc-related ILD is delayed until fibrosis has clearly occurred in the pulmonary tissue.

Gene expression profiling of affected end organs has provided a valuable resource for development of biomarkers in the field of oncology. Although the lung is a prominently affected end organ in SSc, its inaccessibility has precluded the widespread use of pulmonary tissue for research and clinical purposes. The global gene expression profile of SSc pulmonary tissue has only been studied in patients with end-stage disease undergoing lung transplantation (6).

Skin is another prominently affected and easily accessible organ in patients with SSc. Global gene expression studies of the skin tissue of patients with SSc have demonstrated a distinct gene expression profile compared to controls; an inflammatory activation pattern and a fibrotic signature were seen (7,8). In a larger study with 24 SSc patients, patients with diffuse cutaneous SSc (dcSSc) could be subdivided into 3 distinct groups and patients with limited cutaneous SSc (lcSSc) into 2 groups based on skin gene expression profiling. A subgroup of patients with lcSSc and dcSSc showed a pattern of inflammatory genes that included interferon-inducible genes (9) and interleukin-13 (IL-13)–inducible genes (10). Another subgroup of SSc patients with diffuse skin involvement showed a fibrotic gene expression profile containing transforming growth factor β (TGFβ)–responsive genes (11). In that study, none of the patients with dcSSc in the subgroup without activation of TGFβ response genes had ILD (defined as a dichotomous outcome based on high-resolution computed tomography [HRCT] results). There are no published reports of skin transcript correlates of severity of ILD in patients with SSc.

In the present study, we investigated the skin transcript correlates of ILD severity in a large group of patients with SSc. Considering that skin tissue can be obtained during routine clinical practice, the results of the present hypothesis-generating study can provide valuable information for identification of novel biomarkers and therapeutic targets.

PATIENTS AND METHODS

Patients were recruited from the Genetics versus Environment in Scleroderma Outcome Study (GENISOS) (5) or at the baseline visit of an investigator-initiated, open-label phase I/IIa imatinib study (12). All patients either fulfilled the American College of Rheumatology preliminary classification criteria for SSc (13) or had at least 3 of 5 CREST syndrome (calcinosis, Raynaud's phenomenon, esophageal dysmotility, sclerodactyly, telangiectasias) symptoms, with sclerodactyly being mandatory.

Pulmonary function testing (PFT) was performed within 1 month of the skin biopsy date. The predicted values were calculated according to the patient's age, height, weight, sex, and

ethnicity using consistent reference values (14). All PFT results were reviewed by a pulmonologist, and studies that did not fulfill the American Thoracic Society/European Respiratory Society criteria for PFT were excluded. Forced vital capacity (FVC; % predicted) was the primary outcome measure; FVC % predicted is a validated outcome measure for severity of ILD in randomized controlled studies of patients with SSc (15). In the current study, FVC was examined as a continuous outcome for the correlation analysis. In the hierarchical clustering analysis, the presence of moderate-to-severe ILD was defined as FVC <70% predicted and the presence of HRCT chest findings indicative of pulmonary fibrosis. As a secondary outcome measure, the diffusing capacity for carbon monoxide (DLCO; % predicted) was determined concomitantly.

The modified Rodnan skin thickness score (MRSS) (16) was used to assess the severity of skin involvement. In order to avoid issues arising from interobserver variability, the gene expression correlates of the MRSS were investigated only in 31 patients who had been examined by the same physician (MDM).

Skin global gene expression studies

Punch skin biopsy samples (3 mm) were obtained from the arms of study subjects and were put immediately in RNAlater solution (Qiagen) and stored at −80°C. RNA was prepared by mechanical disruption of the stored samples, followed by isolation of RNA using RNeasy Fibrous Tissue (Qiagen) according to the manufacturer's instructions. The RNA quality and yield were assessed with an Agilent 2100 Bioanalyzer and a NanoDrop Technologies ND-1000 Spectrophotometer. All microarray experiments were performed in 1 batch. Two hundred nanograms of total RNA was amplified and purified using a TotalPrep RNA Amplification Kit (Applied Biosystems/Ambion). The amplified complementary DNA was hybridized on Illumina HT-12 arrays, and the data were extracted with Illumina Genomestudio software. An initial clustering analysis based on the date of biopsy, RNA extraction, or hybridization did not show any batch effects resulting from technical artifacts. Microarray data from this study are available from NCBI GEO (http://www.ncbi.nlm.nih.gov/geo/) using accession no. GSE47162. The protocol for obtaining skin biopsy samples was approved by the local institutional review board, and all patients provided fully informed, voluntary consent.

Microarray data analysis

The raw data were exported into and analyzed with BRB-ArrayTools (National Cancer Institute, National Institutes of Health). The data were normalized according to the quantile method. Genes whose log intensity variance was in the bottom 75th percentile were filtered out. The correlation of transcript with FVC, DLCO, and MRSS was examined by Spearman's rank correlation. A transcript was considered correlated with FVC, DLCO, and MRSS (continuous outcome) if Spearman's rank correlation coefficient (r_s) was >0.3 or < −0.3, as well as if the *P* value was less than 0.01. We did not compensate in this method for multiple testing as this study was hypothesis generating.

A more stringent method was also used to explore the genes that correlate with FVC, DLCO, and MRSS. In this analysis, a multivariable permutation test was performed to detect genes that correlate with the outcome with a false discovery rate (FDR) of <5%.

Fisher's exact probability test was used to compare frequency of patients with ILD in gene expression groups identified by unsupervised clustering analysis. Multivariable analysis was performed with the R software environment (www.r-project.org).

The Significance Analysis of Microarrays (SAM) method with an FDR of <5% (17) was used to identify genes that were differentially expressed between disease type (lcSSc versus dcSSc), antibody subsets, and treatment subgroups of SSc patients. The sets of correlated or differentially expressed genes were also modeled in IPA.

Quantitative reverse transcription–polymerase chain reaction (qRT-PCR) was also performed in 2 genes (SELP and CCL2). The expression values were normalized to GAPDH levels. Relative quantification was performed using the C_t method, where $\Delta\Delta C_t$ values were calculated based on GAPDH and transcript levels in controls (18).

Immunofluorescence staining of skin and lung tissue

Immunofluorescence staining was performed on skin biopsy samples obtained from patients with SSc and unaffected controls. Additionally, lung tissue samples from patients with SSc and idiopathic pulmonary fibrosis (IPF) were investigated. These samples were obtained from patients who underwent lung transplantation at the University of Pittsburgh Medical Center under a protocol approved by the institutional review board. Also, normal lung tissue specimens were obtained from organ donors whose lungs were not used for transplantation. Both skin and lung tissue were embedded in paraffin.

Five-micrometer sections of skin and lung tissue were prepared and processed for immunofluorescence. The samples were deparaffinized, rehydrated, and immersed in Tris buffered saline and 0.1% Tween 20 buffer and treated with 10 m*M* citrate buffer at 95°C for 10 minutes. After blocking with goat serum for 1 hour, primary antibodies against CCL2 (monocyte chemotactic protein 1) (MAB2791; R&D Systems) and P-selectin (antigen CD62) (ab6632; Abcam) were used in separate experiments. In addition to the above-mentioned antibodies, primary antibodies against a-smooth muscle actin (a-SMA) (ab5694; Abcam), CD68 (sc-9139; Santa Cruz Bio-technology), and CD31 (ab32457; Abcam) were incubated overnight at 4°C. The secondary antibodies fluorescein isothiocyanate–conjugated anti-mouse IgG and phycoerythrin-conjugated anti-rabbit IgG (Jackson ImmunoResearch) were used, followed by DAPI nuclear staining and viewing under an Olympus BX60 microscope. Immunofluorescence intensity was quantified by Photoshop (Adobe) image analysis. The proportion of cells with double staining (protein of interest plus marker protein) was determined by calculating the ratio of double-positive cells to single-positive cells on 4 representative images.

Chemokine determination

Levels of CCL2, soluble SELP (sSELP), and soluble P-selectin glycoprotein ligand 1 (sPSGL-1) were determined in the baseline samples from 266 SSc patients and 97 age-, ethnicity-, and sex-matched unaffected controls. The patients included all subjects enrolled in the GENISOS as of 2009 (5). Plasma was collected in EDTA blood collection tubes, centrifuged, aliquoted, and stored at −80° C. The plasma samples had undergone only 2 freeze/thaw cycles before protein measurements. All samples were measured in duplicate. Levels of CCL2 were determined by enzyme-linked immunosorbent assay (ELISA) using electro-chemiluminescence multiplex assays (Meso Scale Discovery). Levels of sSELP and sPSGL-1 were determined using commercially available ELISA kits (eBioscience) according to the manufacturer's instructions. The chemokine levels were log-transformed for the analysis. Univariable and multivariable linear regression analysis was used to examine the correlation of cytokine levels with FVC values.

RESULTS

Skin biopsy samples obtained from 59 patients with SSc were examined by global gene expression analysis. As shown in Supplementary Table 1 (available on the *Arthritis & Rheumatism* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.38101/abstract), these patients had a mean disease duration of 7.6 years, and 35 of them (59.3%) had diffuse cutaneous involvement. Fourteen patients were treated with immunosuppressive agents at the time of biopsy. ILD was present in 23 patients (39%).

Skin gene expression correlating with severity of ILD

A total of 11,819 transcripts met the initial filtering criteria. Next, the correlation of transcripts with the concomitantly obtained FVC values (continuous variable) was examined. In this analysis, 82 skin transcripts correlated with FVC (see Supplementary Table 2, available on the *Arthritis & Rheumatism* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.38101/abstract). The overlap between these genes and previously published lists of genes induced by TGFβ (11) and IL-13 (10) was 11% (9 of 82) and 8.5% (7 of 82), respectively.

Next, we conducted an unsupervised hierarchical clustering of SSc samples using this gene list (82 transcripts). This analysis showed a significant separation between subjects with ILD and those without ($P < 0.001$); 19 of 23 patients (83%) with an FVC <70% predicted clustered together, while the majority of patients (24 of 36 [67%]) without ILD clustered in a separate group (Figure 1).

We also performed similar unsupervised clustering analysis using the above-mentioned lists of TGFβ-inducible genes (n = 562) (11) and IL-13–inducible genes (n = 294) (10). The use of the list of TGFβ-inducible genes revealed a trend toward separation between patients with and those without ILD ($P = 0.086$). As shown in Supplementary Figure 1 (available on the *Arthritis & Rheumatism* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.38101/abstract), although 78% of patients (18 of 23) with ILD (FVC <70% predicted) clustered together, the majority of patients without ILD also clustered in that group (19 of 36 [53%]). The use of the list of IL-13–inducible genes resulted in a significant separation between patients with and those without ILD ($P = 0.029$). The majority of ILD patients (17 of 23 [74%]) were grouped together, while 58% of patients without ILD clustered separately (see Supplementary Figure 2, http://onlinelibrary.wiley.com/doi/10.1002/art.38101/abstract).

Forty-four transcripts showed a negative correlation with FVC, indicating that their levels are increased in patients with more severe ILD (lower FVC). Next, we modeled these 44 transcripts in the Ingenuity Pathways Knowledge Base. The top 3 canonical pathways represented were atherosclerosis signaling (CCL2, matrix metalloproteinase 3 [MMP-3], and SELP; $P = 0.0038$), role of IL-17F in allergic inflammatory airway disease (CCL2 and insulin-like growth factor 1 [IGF-1]; $P = 0.0053$), and granulocyte adhesion and diapedesis (CCL2, MMP-3, and SELP; $P = 0.009$). The most prominently represented molecular functions involved cell-to-cell signaling and cellular movement. Genes involved in diapedesis/extravasation, such as SELP ($r_s = -0.52$), CCL2 ($r_s = -0.37$), membrane metalloendopeptidase ($r_s = -0.44$), and MMP-3 ($r_s = -0.38$), were among the most significantly overrepresented genes. These genes are involved in adhesion of immune cells to the endothelium. Of note, IGF-1, which has been implicated in SSc pathogenesis (6,19), also correlated with severity of ILD ($r_s = -0.35$). Each of these genes was significantly correlated with FVC independent of sex, age, disease duration, diffuse cutaneous involvement, and treatment with immunosuppressive agents, as determined in a multivariable linear model (data not shown). When we used the more stringent method (FDR <5%), only SELP correlated significantly with the concomitantly obtained FVC.

As shown in Figure 2, the correlation of selected transcripts (SELP and CCL2) with FVC was confirmed using qRT-PCR. Similar to the microarray results, the correlation of FVC with SELP (Spearman's $\rho = -0.49$, $P < 0.001$) was stronger than its correlation with CCL2 (Spearman's $\rho = -0.34$, $P = 0.009$). In the multivariable analysis, the association of FVC (as a continuous variable) with SELP ($P = 0.007$) and CCL2 ($P = 0.041$) was independent of sex, age, disease duration, diffuse cutaneous involvement, treatment with immunosuppressive agents, and MRSS.

We also examined the genes that correlated with the concomitantly obtained DLCO. A total of 397 genes correlated with DLCO. As shown in Supplementary Table 3 (http://onlinelibrary.wiley.com/doi/10.1002/art.38101/abstract), 326 genes had a negative correlation. Forty-two genes (including SELP, CCL2, MMP-3, and IGF-1) were shared between the lists of genes that correlated with FVC and DLCO, with all these genes showing concordant direction of association. When we used the more stringent method (FDR of <5% in the multivariable analysis), only 7 genes (GGT5, HAPLN3, STAB1, CETP, PKN1, SLC2A3, and PPAPDC3) correlated with DLCO.

Skin transcripts correlating with the severity of skin disease

In the hypothesis-generating analysis, 1,169 transcripts correlated with the MRSS. In the more stringent analysis, 276 transcripts correlated with the MRSS, of which 184 showed a positive correlation (see Supplementary Table 4, http://onlinelibrary.wiley.com/doi/10.1002/art.38101/abstract). The top 3 canonical pathways represented among these 276 transcripts were leukocyte extravasation signaling ($P = 5.2 \times 10^{-7}$), granulocyte adhesion and diapedesis ($P = 4.6 \times 10^{-4}$), and hepatic fibrosis ($P = 6.9 \times 10^{-4}$). The most important molecular functions were cellular development and proliferation, as well as cellular movement.

Eight transcripts (CCL2, HAPLN3, GPR4, ADCYAP1, WARS, CDC25B, PLP1, and STXBP6) were shared between the MRSS and FVC correlates, of which only CCL2 was also included in the more restricted list of MRSS correlates (FDR <5%). As expected, all 8 of these genes showed a discordant direction of association between MRSS and FVC (i.e., transcripts showing a positive correlation with MRSS had a negative correlation with FVC and vice versa).

Genes that were differentially expressed between the clinical subgroups

SAM analysis revealed no differentially expressed genes when patients were subgrouped based on the following characteristics: disease type (diffuse versus limited), treatment with immunosuppressive agents, anti–topoisomerase I positivity, and anticentromere antibody positivity. Comparison of anti–RNA polymerase III–positive patients to the remainder of patients revealed 24 down-regulated genes (see Supplementary Table 5, http://onlinelibrary.wiley.com/doi/10.1002/art.38101/abstract).

Increased SELP and CCL2 immunofluorescence staining in skin and lung tissue from patients with SSc

SELP was selected for the immunofluorescence experiments because it showed the highest negative correlation with FVC ($r_s = -0.52$). Furthermore, immunostaining for CCL2 was performed because of its relevance to the pathogenesis of SSc (10,20) and its moderate negative correlation with FVC ($r_s = -0.37$) in the current study. An increased level of CCL2 immunofluorescence staining has been previously reported in the lesional skin of patients with early dcSSc. The CCL2 staining colocalized with vascular structures and perivascular infiltrates but also appeared throughout the dermis, although with weaker intensity (20). In a similar experiment, we performed immunofluorescence staining with anti–P-selectin

antibodies on skin biopsy samples from normal controls (n = 5) and from lesional skin of patients with SSc (n = 4). Skin biopsy samples from SSc patients showed higher SELP expression than did samples from unaffected healthy controls (Figure 3). Similar to CCL2, the SELP staining colocalized to endothelial cells, macrophages, and a-SMA–positive cells. The SELP/a-SMA–positive cells were located in perivascular areas, as well as throughout the dermis. Of note, the lowest SELP expression among patients was seen in a patient with normal FVC (110% predicted).

Immunofluorescence staining with anti–P-selectin and anti-CCL2 antibodies was also performed on lung tissue from patients with advanced SSc (n = 4), patients with IPF (n = 4), and unaffected controls (n = 5). Lung tissue from healthy controls had low levels of P-selectin protein (Figure 3) and CCL2 protein (Figure 4). In contrast, lung tissue from patients with SSc or IPF had markedly higher SELP and CCL2 staining, localizing to endothelial cells, macrophages, and a-SMA–positive cells (Figures 3 and 4). The a-SMA–positive cells with staining were not confined to the perivascular areas. These findings demonstrate that SELP and CCL2 show high expression levels in skin as well as in lung tissue of patients with SSc, mainly in endothelial cells, vascular and perivascular inflammatory infiltrates, and myofibroblasts. The proportion of cells with double staining in the investigated skin and lung slides is presented in Table 1.

Plasma CCL2 levels correlate independently with severity of ILD in patients with SSc

CCL2, sSELP, and sPSGL-1 were selected for our confirmatory studies of baseline plasma samples from 266 patients (disease duration 5 years) enrolled in the GENISOS cohort and from 97 matched, unaffected controls. (For further characteristics of the 266 patients, see Supplementary Table 6, available on the *Arthritis & Rheumatism* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.38101/abstract.) The plasma CCL2 levels were higher in patients than in controls ($P < 0.001$) (see Supplementary Figure 3, http://onlinelibrary.wiley.com/doi/10.1002/art.38101/abstract). Moreover, they correlated inversely with the concomitantly obtained FVC values in the univariable analysis (r = −0.22, $P = 0.001$) (Figure 5). This relationship remained significant ($P = 0.015$) after adjustment for the potential confounding effects of sex, age, ethnicity, smoking status, disease duration, disease type, anti–topoisomerase I positivity, and treatment with immunosuppressive agents. The association of FVC with CCL2 remained significant even after the model was extended by including the MRSS ($P = 0.021$). This indicates that higher CCL2 levels correlate with more severe ILD independent of other potential confounding factors in patients with SSc. Subgrouping the patients based on an FVC threshold value of 70% predicted revealed that patients with moderate-to-severe ILD had higher CCL2 levels (b = 0.26 [95% confidence interval 0.07, 0.44], $P = 0.007$). Plasma CCL2 levels also correlated with the concomitantly obtained DLCO values (r = −0.19, $P = 0.005$).

The plasma sSELP levels were also higher in patients than in controls ($P < 0.001$) (see Supplementary Figure 3, http://onlinelibrary.wiley.com/doi/10.1002/art.38101/abstract). Although the plasma sSELP level did not show a significant correlation with the concomitantly obtained FVC value, the direction of relationship paralleled the skin transcript findings (r = −0.11, $P = 0.117$) (Figure 5). The relationship also did not become significant after adjustment for the above potential confounders. However, plasma sSELP levels correlated significantly with DLCO values (r = −0.21, $P = 0.002$).

Plasma sPSGL-1 showed a trend toward lower levels in patients than in controls ($P = 0.055$) (Supplementary Figure 3). Furthermore, there was a positive correlation between sPSGL-1 and FVC (r = 0.17, $P = 0.015$) (Figure 5). This association was independent of sex, age, ethnicity, smoking status, disease duration, disease type, anti–topoisomerase I positivity, MRSS, and treatment with immunosuppressive agents ($P = 0.046$). There was also a trend

toward lower plasma sPSGL-1 levels in patients with moderate-to-severe ILD (b = −0.09 [95% confidence interval −0.18, 0.01], *P* = 0.082).

DISCUSSION

To our knowledge, the present study is the first to examine skin transcript correlates of ILD severity in SSc. In this large skin gene expression study (n = 59), a limited number of transcripts correlated with FVC. Transcripts involved in extravasation and attracting of inflammatory cells to endothelium showed the strongest correlation with ILD severity. We also observed an overexpression of these genes at the protein level in the skin and lung tissue of patients with SSc. Furthermore, a correlation of plasma CCL2 and sPSGL-1 levels with ILD severity was confirmed in a large SSc cohort after adjustment for potential confounders.

Several transcripts involved in extravasation, such as SELP and CCL2, have been previously studied in SSc. In an earlier study, plasma SELP levels were found to be higher in SSc patients than in unaffected controls. Within SSc patients, SELP showed a trend toward correlation with perceived shortness of breath. A correlation analysis with FVC was not performed (21). In the present study, sSELP levels did not show a significant correlation with FVC, although SELP transcript levels in the skin had the strongest correlation with FVC. Levels of sSELP correlated significantly with DLCO. PSGL-1 binds with high affinity to SELP and is found on most leukocytes (22). In addition to a membrane-bound form, there is a soluble form of PSGL-1 that competes with the membrane-bound form in binding to SELP. Therefore, sPSGL-1 can act as an inhibitor of adhesion/extravasation and can directly inhibit SELP function. In the present study, we observed a positive correlation between sPSGL-1 and FVC, indicating that patients with more severe ILD have lower levels of sPSGL-1. This relationship was independent of potential confounders, including the concomitantly obtained MRSS.

In previous studies, serum CCL2 levels were also higher in SSc patients than in unaffected controls. Furthermore, serum CCL2 levels correlated with pulmonary vital capacity at the cross-sectional and longitudinal levels in 31 SSc patients. However, an adjustment for potential confounders was not performed because the sample size was small (23). In another study, skin CCL2 transcript levels correlated with severity of skin involvement, measured by the MRSS (10). In vitro blockage of CCL2 reduced the expression of *α*-SMA, indicating that CCL2 may be important in initiating or maintaining myofibroblast proliferation leading to fibrosis (20). Furthermore, blockage of CCL2 in 2 murine SSc models (the bleomycin model [24] and sclerodermatous graft-versus-host disease [10]) attenuated fibrosis. CCL2 was also one of the top 20 differentially regulated genes in the pulmonary tissue of patients with advanced SSc-related ILD (6). Our results support the importance of CCL2 as a potential biomarker and therapeutic target in SSc.

It is noteworthy that we did not observe a correlation of ILD severity with primary fibrotic genes such as collagen, TGF*β*, or connective tissue growth factor. This might be because skin fibrosis peaks early during the course of SSc and improves later (25), while fibrosis in pulmonary tissue continues to progress even at later stages of disease (5). Therefore, primary fibrotic genes in the skin might not reflect changes in lung fibrosis. As expected, fibrotic genes (including collagen) correlated with the MRSS. Furthermore, a greater number of genes showed a correlation with the MRSS than with FVC, and the magnitude of this correlation was higher (r_s ranging up to 0.81). A 4-gene panel consisting of cartilage oligomeric matrix protein (COMP), thrombospondin 1, IFI44, and sialic acid–binding Ig-like lectin 1 has been proposed as a biomarker that correlates closely with the concomitantly

obtained MRSS (26). In the present study, COMP correlated with the MRSS ($r_s = 0.57$), while we did not observe significant correlations with the other 3 genes.

Genes involved in extravasation are chemokines that regulate the migration and recruitment of specific leukocytes to endothelium and regions of inflammation. Therefore, they might represent fibrotic as well as vasculopathic components of SSc pathogenesis. Furthermore, several genes involved in extravasation, including SELP and CCL2, are regulated by IL-13, which has been implicated in pathogenesis of SSc (10). This might also explain why the IL-13 gene signature distinguished the patients with ILD better than did the TGFβ-inducible genes in the current study.

The immunofluorescence staining in the present study confirmed the overexpression of CCL2 and SELP at the protein level in the pulmonary tissue of patients with SSc and IPF. To our knowledge, this represents the first report of increased immunofluorescence staining of these 2 important transcripts in tissue from patients with SSc-related ILD. The colocalization experiments indicated that these molecules are expressed in monocyte/macrophages, endothelial cells, and α-SMA–positive cells. The fact that α-SMA costaining was not confined to the perivascular areas indicates that these 2 molecules are expressed beyond pericytes in myofibroblasts.

We also examined the correlation of plasma CCL2 levels with severity of ILD in a large multiethnic cohort of patients with early SSc. Confirming the findings of a smaller study (23), CCL2 levels showed a modest correlation with FVC. Similar results were observed with sPSGL-1, which strongly binds SELP. These correlations were independent of potential confounding factors such as age, disease type, smoking, treatment with immunosuppressive agents, and the MRSS. However, the skin CCL2 and SELP transcript levels showed stronger correlations with ILD severity than did the plasma CCL2 and sPSGL-1 levels, suggesting that skin transcripts of these molecules might be more beneficial for biomarker development. In general, the transcripts correlating with FVC in the present study are candidates for future longitudinal studies examining the predictive significance of their skin or plasma/serum protein levels for progression of ILD or response to treatment. A recently reported study examined the skin gene expression profiles of 7 patients with baseline MRSS 11 who were subsequently treated with mycophenolate mofetil (27). Four patients were classified as improvers and 3 were classified as nonimprovers in severity of skin involvement upon followup. All patients with MRSS improvement had an inflammatory skin gene expression signature, while patients without clinical improvement had a fibrotic or normal-like transcript profile. Progression of SSc-related ILD was not investigated in that study. It is conceivable that similar studies can develop predictive skin transcript biomarkers for SSc-related ILD if they are sufficiently powered and conducted in appropriately selected patient populations.

The present study has several strengths. A large number of patients were examined by global gene expression analysis. With 59 unique SSc patients, the present study represents the largest skin global gene expression study in SSc published to date. Furthermore, the concomitant measurement of FVC at the time of skin biopsy allowed us to investigate the correlates of ILD severity using a validated outcome measure (15) for this important disease complication. In addition, availability of detailed clinical data in the GENISOS cohort enabled us to adjust for potential confounders such as smoking, the MRSS, or disease duration.

The present study also has some limitations. We did not correct for multiple comparisons as this was primarily a hypothesis-generating investigation, but results of more stringent analysis have also been provided. Furthermore, our study was cross-sectional, which

precluded our investigating the predictive significance of skin transcripts. Furthermore, HRCT images were not available for the majority of patients because some patients with no PFT abnormalities had not participated in this diagnostic study, and the majority of patients with low FVC values had undergone HRCT in an outside facility where only the radiology report was available to us. Therefore, we could not investigate the extent of fibrosis (fibrosis score) on HRCT as a potential alternative outcome measure.

In summary, a limited number of skin transcripts, including genes involved in extravasation such as SELP and CCL2, correlated with severity of ILD in SSc. Given the accessibility of skin biopsy samples during routine clinical practice, skin transcript levels of these genes are attractive candidates for biomarker development in future longitudinal studies. Furthermore, our findings provide support for the potential utility of these molecules as therapeutic targets for SSc-related ILD.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Steen VD, Medsger TA. Changes in causes of death in systemic sclerosis, 1972–2002. Ann Rheum Dis. 2007; 66:940–4. [PubMed: 17329309]

2. Tyndall AJ, Bannert B, Vonk M, Airo P, Cozzi F, Carreira PE, et al. Causes and risk factors for death in systemic sclerosis: a study from the EULAR Scleroderma Trials and Research (EUSTAR) database. Ann Rheum Dis. 2010; 69:1809–15. [PubMed: 20551155]

3. Tashkin DP, Elashoff R, Clements PJ, Goldin J, Roth MD, Furst DE, et al. Cyclophosphamide versus placebo in scleroderma lung disease. N Engl J Med. 2006; 354:2655–66. [PubMed: 16790698]

4. Tashkin DP, Elashoff R, Clements PJ, Roth MD, Furst DE, Silver RM, et al. Effects of 1-year treatment with cyclophosphamide on outcomes at 2 years in scleroderma lung disease. Am J Respir Crit Care Med. 2007; 176:1026–34. [PubMed: 17717203]

5. Assassi S, Sharif R, Lasky RE, McNearney TA, Estrada YM, Draeger HT, et al. Predictors of interstitial lung disease in early systemic sclerosis: a prospective longitudinal study of the GENISOS cohort. Arthritis Res Ther. 2010; 12:R166. [PubMed: 20813056]

6. Hsu E, Shi H, Jordan RM, Lyons-Weiler J, Pilewski JM, Feghali-Bostwick CA. Lung tissues in patients with systemic sclerosis have gene expression patterns unique to pulmonary fibrosis and pulmonary hypertension. Arthritis Rheum. 2011; 63:783–94. [PubMed: 21360508]

7. Gardner H, Shearstone JR, Bandaru R, Crowell T, Lynes M, Trojanowska M, et al. Gene profiling of scleroderma skin reveals robust signatures of disease that are imperfectly reflected in the transcript profiles of explanted fibroblasts. Arthritis Rheum. 2006; 54:1961–73. [PubMed: 16736506]

8. Whitfield ML, Finlay DR, Murray JI, Troyanskaya OG, Chi JT, Pergamenschikov A, et al. Systemic and cell type-specific gene expression patterns in scleroderma skin. Proc Natl Acad Sci U S A. 2003; 100:12319–24. [PubMed: 14530402]

9. Milano A, Pendergrass SA, Sargent JL, George LK, McCalmont TH, Connolly MK, et al. Molecular subsets in the gene expression signatures of scleroderma skin. PLoS One. 2008; 3:e2696. [PubMed: 18648520]

10. Greenblatt MB, Sargent JL, Farina G, Tsang K, Lafyatis R, Glimcher LH, et al. Interspecies comparison of human and murine scleroderma reveals IL-13 and CCL2 as disease subset-specific targets. Am J Pathol. 2012; 180:1080–94. [PubMed: 22245215]

11. Sargent JL, Milano A, Bhattacharyya S, Varga J, Connolly MK, Chang HY, et al. A TGFβ-responsive gene signature is associated with a subset of diffuse scleroderma with increased disease severity. J Invest Dermatol. 2010; 130:694–705. [PubMed: 19812599]

12. Khanna D, Saggar R, Mayes MD, Abtin F, Clements PJ, Maranian P, et al. A one-year, phase I/IIa, open-label pilot trial of imatinib mesylate in the treatment of systemic sclerosis–associated active interstitial lung disease. Arthritis Rheum. 2011; 63:3540–6. [PubMed: 21769849]

13. Subcommittee for Scleroderma Criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Arthritis Rheum. 1980; 23:581–90. [PubMed: 7378088]

14. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S population. Am J Respir Crit Care Med. 1999; 159:179–87. [PubMed: 9872837]

15. Furst D, Khanna D, Matucci-Cerinic M, Clements P, Steen V, Pope J, et al. Systemic sclerosis—continuing progress in developing clinical measures of response. J Rheumatol. 2007; 34:1194–200. [PubMed: 17477486]

16. Clements P, Lachenbruch P, Seibold J, White B, Weiner S, Martin R, et al. Inter and intraobserver variability of total skin thickness score (modified Rodnan TSS) in systemic sclerosis. J Rheumatol. 1995; 22:1281–5. [PubMed: 7562759]

17. Tusher VG, Tibshirani R, Chu G. Significance analysis of micro-arrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A. 2001; 98:5116–21. [PubMed: 11309499]

18. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. Methods. 2001; 25:402–8. [PubMed: 11846609]

19. Hamaguchi Y, Fujimoto M, Matsushita T, Hasegawa M, Takehara K, Sato S. Elevated serum insulin-like growth factor (IGF-1) and IGF binding protein-3 levels in patients with systemic sclerosis: possible role in development of fibrosis. J Rheumatol. 2008; 35:2363–71. [PubMed: 19004037]

20. Carulli MT, Ong VH, Ponticos M, Shiwen X, Abraham DJ, Black CM, et al. Chemokine receptor CCR2 expression by systemic sclerosis fibroblasts: evidence for autocrine regulation of myofibroblast differentiation. Arthritis Rheum. 2005; 52:3772–82. [PubMed: 16320328]

21. Blann AD, Constans J, Carpentier P, Renard M, Satger B, Guerin V, et al. Soluble P selectin in systemic sclerosis: relationship with von Willebrand factor, autoantibodies and diffuse or localised/limited disease. Thromb Res. 2003; 109:203–6. [PubMed: 12757775]

22. Patel KD, Moore KL, Nollert MU, McEver RP. Neutrophils use both shared and distinct mechanisms to adhere to selectins under static and flow conditions. J Clin Invest. 1995; 96:1887–96. [PubMed: 7560080]

23. Hasegawa M, Fujimoto M, Matsushita T, Hamaguchi Y, Takehara K, Sato S. Serum chemokine and cytokine levels as indicators of disease activity in patients with systemic sclerosis. Clin Rheumatol. 2011; 30:231–7. [PubMed: 21049277]

24. Inoshima I, Kuwano K, Hamada N, Hagimoto N, Yoshimi M, Maeyama T, et al. Anti-monocyte chemoattractant protein-1 gene therapy attenuates pulmonary fibrosis in mice. Am J Physiol Lung Cell Mol Physiol. 2004; 286:L1038–44. [PubMed: 15064241]

25. Amjadi S, Maranian P, Furst DE, Clements PJ, Wong WK, Postlethwaite AE, et al. for the Investigators of the D-Penicillamine. Human Recombinant Relaxin. and Oral Bovine Type I Collagen Clinical Trials. Course of the modified Rodnan skin thickness score in systemic sclerosis clinical trials: analysis of three large multicenter, double-blind, randomized controlled trials. Arthritis Rheum. 2009; 60:2490–8. [PubMed: 19644851]

26. Farina G, Lafyatis D, Lemaire R, Lafyatis R. A four-gene biomarker predicts skin disease in patients with diffuse cutaneous systemic sclerosis. Arthritis Rheum. 2010; 62:580–8. [PubMed: 20112379]

27. Hinchcliff M, Huang CC, Wood TA, Matthew MJ, Martyanov V, Bhattacharyya S, et al. Molecular signatures in skin associated with clinical improvement during mycophenolate treatment in systemic sclerosis. J Invest Dermatol. 2013; 133:1979–89. [PubMed: 23677167]
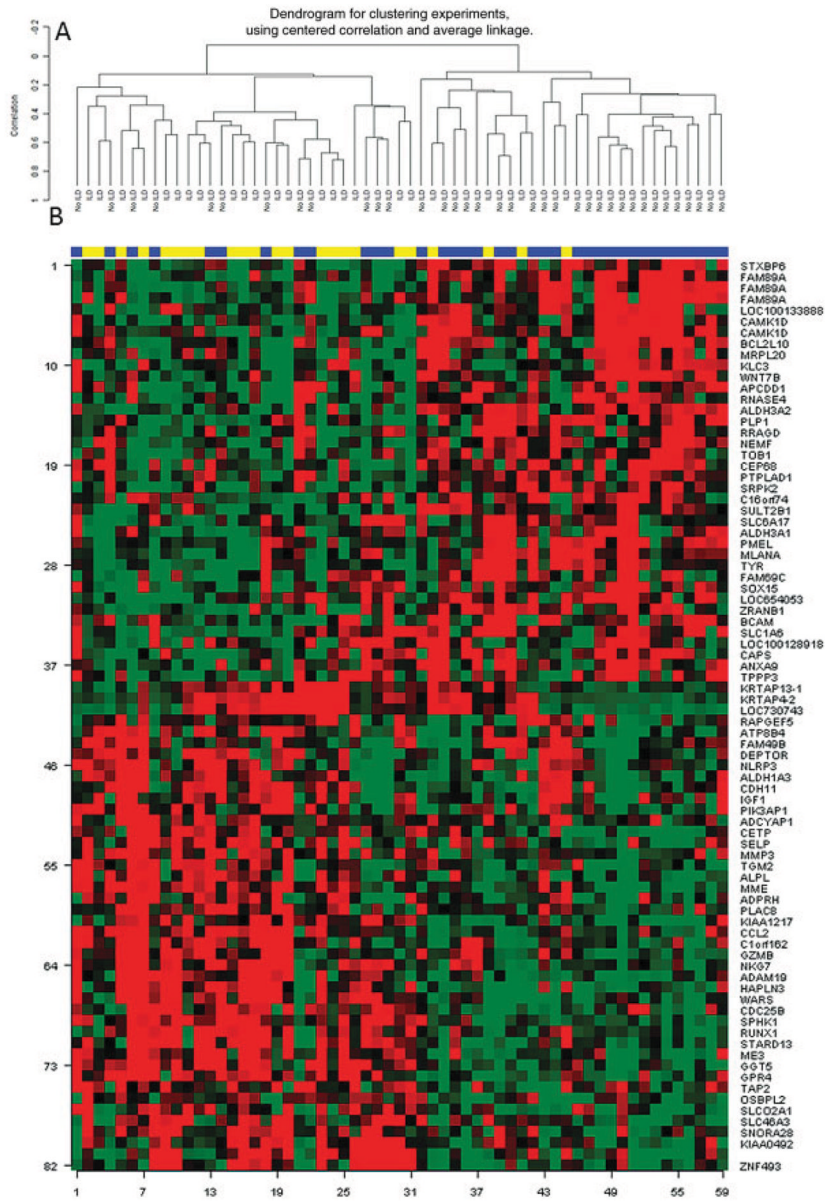
Figure 1.
Unsupervised hierarchical clustering of 82 skin transcripts that correlated with the concomitantly obtained forced vital capacity % predicted values. **A,** Only samples. **B,** Samples/genes. Samples are labeled according to study groups: those with interstitial lung disease (ILD) (yellow) and those with no ILD (purple).
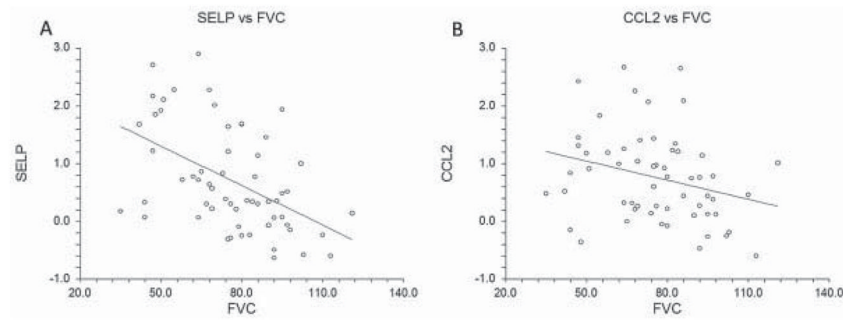
Figure 2.
Correlation of skin transcript levels of SELP (**A**) and CCL2 (**B**) with the concomitantly obtained forced vital capacity (FVC) % predicted value. The transcript levels have been log-transformed.
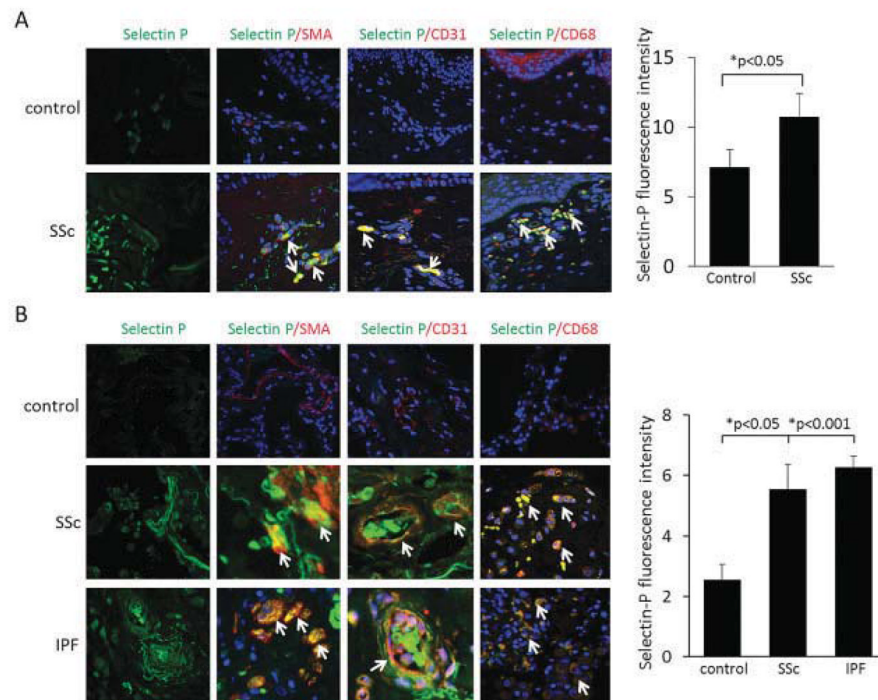
Figure 3.
Left, Immunofluorescence staining with anti–P-selectin antibodies in skin (**A**) and lung tissue (**B**) from normal controls and patients with systemic sclerosis (SSc) or idiopathic pulmonary fibrosis (IPF). **Arrows** indicate double-positive cells. Original magnification × 400. SMA = α-smooth muscle actin. Right, Fluorescence intensity of P-selectin in the 3 groups. Bars show the mean ± SD.
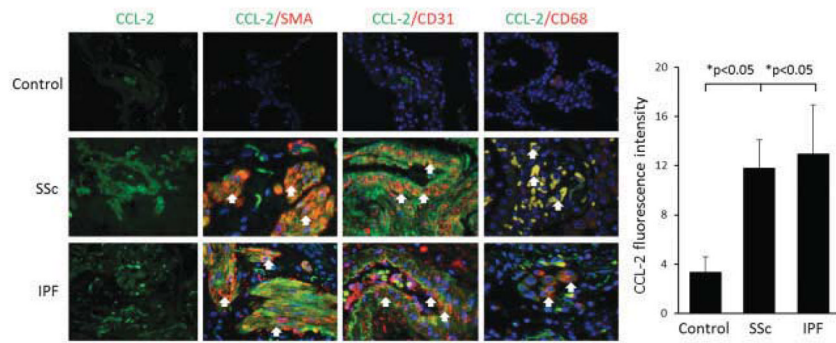
Figure 4.
Left, Immunofluorescence staining with anti-CCL2 antibodies in lung tissue from normal controls and patients with SSc or IPF. **Arrows** indicate double-positive cells. Original magnification × 400. Right, Fluorescence intensity of CCL2 in the 3 groups. Bars show the mean ± SD. See Figure 3 for definitions.
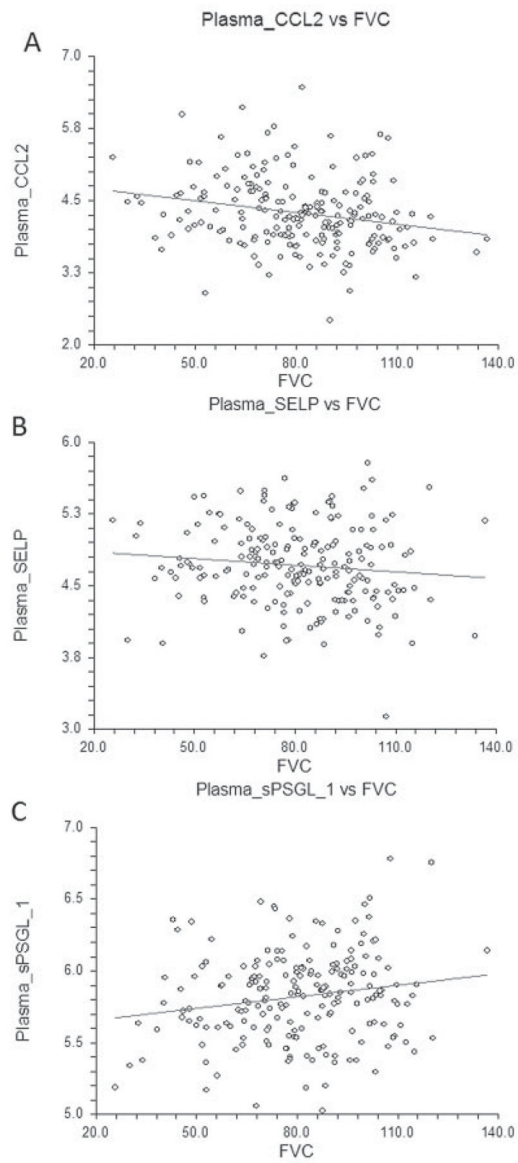
Figure 5.
Correlations of plasma levels of CCL2 (**A**), soluble SELP (**B**), and soluble P-selectin glycoprotein ligand 1 (sPSGL-1) (**C**) with concomitantly obtained forced vital capacity (FVC) % predicted values.

Table 1

Proportion of cells with double staining in the skin and lung tissue[*]

Tissue, sample	Protein of interest	CD31	CD68	a-SMA
Skin				
Control	P-selectin	9.9 ± 3.2	4.5 ± 0.1	9.7 ± 4.8
SSc	P-selectin	$49.2 \pm 10.8^{\dagger}$	$51.2 \pm 8.8^{\dagger}$	$27.6 \pm 3.8^{\dagger}$
Lung				
Control	P-selectin	20.8 ± 15.0	26.5 ± 2.0	5.4 ± 4.0
SSc	P-selectin	42.0 ± 8.3	$58.1 \pm 7.2^{\dagger}$	$63.9 \pm 6.6^{\dagger}$
IPF	P-selectin	$56.4 \pm 7.4^{\dagger}$	$69.2 \pm 3.2^{\dagger}$	$38.4 \pm 6.4^{\dagger}$
Lung				
Control	CCL2	22.9 ± 3.6	20.6 ± 5.3	13.6 ± 4.9
SSc	CCL2	$58.5 \pm 6.5^{\dagger}$	$48.5 \pm 6.8^{\dagger}$	$50.6 \pm 9.7^{\dagger}$
IPF	CCL2	$65.7 \pm 3.8^{\dagger}$	$59.9 \pm 7.2^{\dagger}$	$60.9 \pm 4.4^{\dagger}$

[*] Values are the mean ± SD proportion of cells with double staining (protein of interest plus marker protein). Proportions were determined by calculating the ratio of double-positive cells (i.e., P-selectin or CCL2 plus CD31 or CD68 or a-smooth muscle actin [a-SMA]) to single-positive cells (i.e., CD31 or CD68 or a-SMA only) on 4 representative images. SSc = systemic sclerosis; IPF = idiopathic pulmonary fibrosis.

[†] $P < 0.05$ versus control subjects.