

1-21-2015

# Experimentally-Derived Fibroblast Gene Signatures Identify Molecular Pathways Associated with Distinct Subsets of Systemic Sclerosis Patients in Three Independent Cohorts

Michael E. Johnson  
*Dartmouth College*

J. Matthew Mahoney  
*Dartmouth College*

Jaclyn Taroni  
*Dartmouth College*

Jennifer L. Sargent  
*Dartmouth College*

Eleni Marmarelis  
*Dartmouth College*

*See next page for additional authors*

Follow this and additional works at: <https://digitalcommons.dartmouth.edu/facoa>

 Part of the [Medicine and Health Sciences Commons](#)

---

## Recommended Citation

Johnson, Michael E.; Mahoney, J. Matthew; Taroni, Jaclyn; Sargent, Jennifer L.; Marmarelis, Eleni; Wu, Ming-ru; Varga, John; Hinchcliff, Monique E.; and Whitfield, Michael L., "Experimentally-Derived Fibroblast Gene Signatures Identify Molecular Pathways Associated with Distinct Subsets of Systemic Sclerosis Patients in Three Independent Cohorts" (2015). *Open Dartmouth: Faculty Open Access Articles*. 2720.  
<https://digitalcommons.dartmouth.edu/facoa/2720>

---

**Authors**

Michael E. Johnson, J. Matthew Mahoney, Jaclyn Taroni, Jennifer L. Sargent, Eleni Marmarelis, Ming-ru Wu, John Varga, Monique E. Hinchcliff, and Michael L. Whitfield

RESEARCH ARTICLE

# Experimentally-Derived Fibroblast Gene Signatures Identify Molecular Pathways Associated with Distinct Subsets of Systemic Sclerosis Patients in Three Independent Cohorts

Michael E. Johnson<sup>1</sup>, J. Matthew Mahoney<sup>1</sup>, Jaclyn Taroni<sup>1</sup>, Jennifer L. Sargent<sup>1</sup>, Eleni Marmarelis<sup>1</sup>, Ming-Ru Wu<sup>1</sup>, John Varga<sup>2</sup>, Monique E. Hinchcliff<sup>2</sup>, Michael L. Whitfield<sup>1\*</sup>

**1** Department of Genetics, Geisel School of Medicine at Dartmouth, Hanover, NH, United States of America, **2** Northwestern University Feinberg School of Medicine, Department of Internal Medicine, Division of Rheumatology, Chicago, IL, United States of America

\* [Michael.L.Whitfield@Dartmouth.edu](mailto:Michael.L.Whitfield@Dartmouth.edu)



OPEN ACCESS

**Citation:** Johnson ME, Mahoney JM, Taroni J, Sargent JL, Marmarelis E, Wu M-R, et al. (2015) Experimentally-Derived Fibroblast Gene Signatures Identify Molecular Pathways Associated with Distinct Subsets of Systemic Sclerosis Patients in Three Independent Cohorts. PLoS ONE 10(1): e0114017. doi:10.1371/journal.pone.0114017

**Academic Editor:** Stamatis-Nick Liossis, University of Patras Medical School, GREECE

**Received:** June 18, 2014

**Accepted:** November 4, 2014

**Published:** January 21, 2015

**Copyright:** © 2015 Johnson et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Data from time courses generated here (IL-4, IL-13, S1P, and PDGF) have been submitted to GEO under accession numbers GSE56308 and GSE59785. The data used in this study from published sources are available from NCBI GEO at the following accessions: GSE9285 (Milano et al. 2008), GSE32413 (Pendergrass et al. 2012), and GSE45485 (Hinchcliff et al. 2013), GSE12493 (TGFβ from Sargent et al. 2009), GSE11130 (Chung et al. 2009), and GSE24125 (Rubins, et al. 2011).

## Abstract

Genome-wide expression profiling in systemic sclerosis (SSc) has identified four ‘intrinsic’ subsets of disease (fibroproliferative, inflammatory, limited, and normal-like), each of which shows deregulation of distinct signaling pathways; however, the full set of pathways contributing to this differential gene expression has not been fully elucidated. Here we examine experimentally derived gene expression signatures in dermal fibroblasts for thirteen different signaling pathways implicated in SSc pathogenesis. These data show distinct and overlapping sets of genes induced by each pathway, allowing for a better understanding of the molecular relationship between profibrotic and immune signaling networks. Pathway-specific gene signatures were analyzed across a compendium of microarray datasets consisting of skin biopsies from three independent cohorts representing 80 SSc patients, 4 morphea, and 26 controls. IFNα signaling showed a strong association with early disease, while TGFβ signaling spanned the fibroproliferative and inflammatory subsets, was associated with worse MRSS, and was higher in lesional than non-lesional skin. The fibroproliferative subset was most strongly associated with PDGF signaling, while the inflammatory subset demonstrated strong activation of innate immune pathways including TLR signaling upstream of NF-κB. The limited and normal-like subsets did not show associations with fibrotic and inflammatory mediators such as TGFβ and TNFα. The normal-like subset showed high expression of genes associated with lipid signaling, which was absent in the inflammatory and limited subsets. Together, these data suggest a model by which IFNα is involved in early disease pathology, and disease severity is associated with active TGFβ signaling.

**Funding:** This work was supported by the Scleroderma Research Foundation (SRF) (MLW, MH), P50AR060780 (MLW), P30AR061271 (MLW) and U01AR055063 (MLW and R. Lafyatis). JMM was supported by Award Number R25 CA134286-01 from the National Cancer Institute (NCI). JT was supported by Award Number T32GM008704 from the National Institute of General Medical Sciences (NIGMS). MEJ was also supported by a SYNERGY Grant from the Geisel School of Medicine at Dartmouth (to MLW). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIGMS, NIAMS, NCI or the NIH. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** MLW and MH have filed patents for gene expression biomarkers in SSc. Molecular signatures for diagnosing scleroderma (PCT/US2009/004089) and Gene expression signature in skin predicts response to mycophenolate mofetil (PCT/US2012/027020). MLW is the Scientific Founder of Celdara Medical, LLC. There are no further patents or products to declare from the work contained herein. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

## Introduction

Systemic sclerosis (SSc) is a progressive fibrotic disease of unknown etiology characterized by fibrosis of the skin and internal organs, vascular abnormalities, immune activation, and excessive extracellular matrix deposition. Heterogeneity of disease symptoms and outcomes remains a significant obstacle, though emerging data are beginning to provide insight. Clinical classifications of SSc are based primarily on the extent of skin and internal organ involvement, and SSc autoantibody profiles [1]. Multiple high-throughput gene expression analyses of patient skin biopsies have identified four SSc intrinsic subsets that span the two clinically identified subsets of limited (lSSc) and diffuse (dSSc) disease. Distinct molecular signaling pathways appear to underlie each subset, providing insights into the clinically observed heterogeneity between SSc patients that has confounded clinical trials. Analysis of serial biopsies over 6–12 months has shown the intrinsic subsets to be stable over this short time frame, but does not rule out the possibility of patients changing subsets over much longer time scales [2].

Previously, we have described associations between both a TGF $\beta$ -responsive gene signature and increased disease severity in the fibroproliferative subset of dSSc patients [3], and an IL-13/CCL2 gene signature and the inflammatory subset [4]. While these associations were suggestive, the studies were limited by the small number of samples available, and the absence of a validation cohort. In addition, these pathways accounted for only a fraction of the overall gene expression present within each of the intrinsic gene expression subset of SSc. Here, we have expanded our analyses to include ten additional inflammatory and fibrotic signaling pathways (three experimentally derived here for the first time; seven taken from the literature), and expanded on two others, to determine the genes induced, the unique and overlapping genes among the pathways, and how each contributes to the gene expression changes in SSc skin. Along with our prior analyses of TGF $\beta$ , these pathway gene signatures were compared against three independent SSc patient cohorts, which were merged into a single dataset, and stratified into intrinsic gene expression subsets. This allows us to assess the relative contribution of each signaling pathway to the gene expression changes seen in SSc skin.

The list of pathways analyzed here includes both pathway analyses previously performed within our own group, along with pathways strongly implicated by the primary literature, but without knowledge of how they stratify across a sample of the SSc patient population. Pathways suggested by the literature include platelet-derived growth factor (PDGF), sphingosine-1-phosphate (S1P), peroxisome proliferator-activated receptor gamma (PPAR $\gamma$ ), tumor necrosis factor alpha (TNF $\alpha$ ), interferon alpha (IFN $\alpha$ ), nuclear factor kappa-B (NF- $\kappa$ B), and innate immune signaling. The *in vivo* gene response to imatinib mesylate was also included in these analyses due to the overlapping functions of this drug, and its use as an experimental treatment for SSc [5].

IFN $\alpha$  signaling was strongly associated with early disease, while TGF $\beta$  signaling spanned both the inflammatory and fibroproliferative subsets, and was associated with more severe skin involvement. We find the fibroproliferative intrinsic subset to be more strongly associated with the PDGF gene signature, while the inflammatory subset is associated with a wide range of NF- $\kappa$ B activating pathways.

## Materials and Methods

### Skin biopsy data

Microarray data for scleroderma lesional and nonlesional skin biopsies and healthy controls used in this analysis have been described previously [2,6,7]. These data are publically available in the NCBI GEO database under accession numbers GSE9285, GSE32413, and GSE45485,

respectively. Additional skin biopsy microarrays not previously described elsewhere are also included in this dataset, and are available from the NCBI GEO database under accession number GSE59785. The analysis of human samples in this study was approved by the Committee for the Protection of Human Subjects at Dartmouth College (CPHS# 16631) and by the institutional review boards (IRB) of Northwestern University's Feinberg School of Medicine (STU00004428). All subjects in the study provided written consent, which was approved by the IRB review panels of Dartmouth College and Northwestern University Feinberg School of Medicine.

Batch effects evident between the three datasets were adjusted using ComBat [8] run as a GenePattern module using parametric and non-parametric settings. The statistical significance of batch bias before and after adjustment was assessed using guided principal component analysis (gPCA) and the first two unguided principal components were inspected (S1 Fig.). Genes were then selected using an intrinsic gene identifier algorithm [9] using a false discovery rate (FDR) sufficient to produce reproducible clusters (generally between 2000 and 3000 probes), clustered using Cluster 3.0 [10], and visualized with Java TreeView [11]. The distribution of intrinsic subset assignments in the original published datasets were compared to those determined after ComBat adjustment using a Chi-squared test.

## Experimental treatment and RNA preparation

Primary adult NHDFs were obtained from Cambrex Bioscience Inc. (East Rutherford, NJ); SSC fibroblasts were isolated from explanted lesional biopsies cultured for three passages in DMEM supplemented with 10% fetal bovine serum (FBS) (v/v) and 100 IU/mL penicillin-streptomycin. To measure pathway treatment responses,  $4 \times 10^5$  fibroblasts were seeded in 100 mm dishes, and cultured in DMEM supplemented with 10% FBS for 48 h; cells were then brought to quiescence in DMEM plus 0.1% FBS for 24 h. Cellular agonists (PDGF, R&D Systems, Minneapolis, MN; rosiglitazone (RZN), Cayman Chemical Company, Ann Arbor, MI; S1P, Sigma-Aldrich, St. Louis, MO; IL-4 and IL-13, Peprotech, Rocky Hill, NJ) were added to low serum media, and cells incubated for 0, 2, 4, 8, 12, and 24 h; baseline, zero hour time points were performed in triplicate. Following treatment, cells were lysed in RLT buffer supplemented with 0.1%  $\beta$ -mercaptoethanol, and total RNA isolated using RNeasy mini kits (Qiagen, Valencia, CA), according to the manufacturer's instructions. Pathway gene signatures were defined as all probes exhibiting a  $\geq 2$ -fold mean change in expression relative to controls at 12 and 24 h across all replicates. Data were filtered to include only probes showing an average correlation  $> 0.8$  relative to an idealized induction pattern (full induction at 2–24 h time points).

## Quantitative real-time PCR

Reverse transcription of total RNA (200 ng) was performed using SuperScript II reverse transcriptase (Invitrogen, San Diego, CA) to generate single-stranded complementary DNA; 1.0 mg cDNA was used for each qRT-PCR reaction. Taqman gene expression probes for CD36, THBD, and 18S were obtained from Life Technologies (Foster City, CA), and analyzed using the 7500 Fast Real-Time PCR system. Fold changes were calculated relative to 18S controls using the comparative  $C_t$  formula  $2^{-\Delta\Delta C_t}$  [12]. All experiments were performed in triplicate.

## Microarray procedures

Microarray hybridizations were performed as described previously [7]. Briefly, RNA quality was assessed using the Agilent 2100 Bioanalyzer, and quantified using a Thermo Scientific NanoDrop 2000 spectrophotometer. Total RNA (200 ng) was amplified and labeled using Agilent QuickAmp Labeling kits, as described previously [6]. Cy3-labeled sample and

Cy5-labeled Universal Human Reference RNA (Stratagene, La Jolla, CA) we co-hybridized onto Agilent SurePrint Human Genome 4 × 44k (G4112F) and 8 × 60k (G4851A) microarrays. Data were uploaded to the UNC microarray database, normalized, and filtered for spot quality and signal intensity. Microarray data from this paper have been deposited in the NCBI GEO database under accession numbers GSE56038 and GSE59785.

## Data analysis

Data analyses were performed for each of the 13 agonists: PDGF, S1P, RZN, TGFβ, IL-13, IL-4, IFNα, TNFα, Polyinosinic:polycytidylic acid (poly(I-C)), ionomycin-phorbol 12-myristate 13-acetate (ionomycin-PMA), dexamethasone (DEX), lipopolysaccharide (LPS), and imatinib mesylate. PDGF, S1P, and RZN time courses were performed as part of this analysis. TGFβ time courses were originally described by Sargent, *et al.* and are available from the NCBI GEO database under accession number GSE12493. Two additional IL-13 and IL-4 time courses each were performed adding to the data published in Greenblatt, *et al.* [4] and are available under accession number GSE56308. *In vitro* fibroblast treatment arrays for agonists IFNα, TNFα, poly(I-C), ionomycin-PMA, DEX, and LPS were originally described by Rubins, *et al.* [13], and are available from the NCBI GEO database under accession number GSE24125. *In vivo* imatinib mesylate treatment response microarrays were performed by Chung, *et al.* [5] using skin biopsies collected before and after treatment; these data are available from the NCBI GEO database under accession number GSE11130. A summary of all treatment-associated microarray data used in this study is presented in Table 1.

Pathway gene signatures for each of the treatments listed in Table 2 were defined as all genes up or downregulated  $\geq 2$ -fold across all of their corresponding 12 and 24 h time points, relative to untreated controls. The imatinib signature was determined based upon a *p*-value cutoff, as defined by Chung, *et al.* [5]. A centroid was created for each of the TGFβ, PDGF, S1P, RZN, IL-13, and IL-4 time courses by averaging the 12 and 24 h time points, and the centroids aligned to the MPH dataset using Agilent probe IDs. Data for IFNα, TNFα, poly(I-C), LPS, ionomycin-PMA, DEX, and imatinib were aligned by Entrez ID due to differences in microarray annotation; genes represented by multiple probes were averaged across all probes for both the treatment and MPH datasets. Pearson's correlation coefficients were calculated between each pathway gene signature centroid and the MPH dataset for each individual microarray; average correlations were then calculated for each gene signature for each of the four intrinsic subsets. *P* values quantifying the enrichment of pathway signatures within individual subsets were calculated based upon the average Pearson's correlation *r* coefficient using the standard method for Pearson's correlation *P* value calculations, with *n* defined as the number of genes within each pathway.

Pearson's correlation coefficients were used to quantify the contribution of a specific pathway to the gene expression within a given patient. These correlation 'scores' were then compared against clinically relevant factors including age, sex, modified Rodnan skin score (MRSS), biopsy site, and disease duration to identify the predictive value of each pathways for these clinical outcomes. Clinical comparisons were limited to dSSc patients only, using a single array per patient per time point collected; in cases where both lesional and non-lesional biopsies were collected only the lesional biopsy was considered. Comparisons of biopsy site were limited to clinically dSSc patients that provided paired lesional and non-lesional biopsies at a given time point; *n* denotes the number of patients included in each analysis. Continuous variables were compared using Pearson's correlation; categorical variables were analyzed by ANOVA. All statistical analyses were performed using IBM SPSS version 19. *P* values  $\leq 0.05$  were considered statistically significant.

Table 1. Overview of agonists used in this study.

Agonist	Primary Target	Source	Experimental Overview	Accession
PDGF	PDGFR	This study	<i>In vitro</i> human fibroblast time course	GSE56308
S1P	S1PR1–5	This study	<i>In vitro</i> human fibroblast time course	GSE56308
RZN	PPAR $\gamma$	This study	<i>In vitro</i> human fibroblast time course	GSE56308
IL-13	IL-13R $\alpha$ 1/IL-4R $\alpha$	This study	<i>In vitro</i> human fibroblast time course	GSE56308
IL-4	IL-4R $\alpha$	This study	<i>In vitro</i> human fibroblast time course	GSE56308
TGF $\beta$	TGF $\beta$	Sargent, <i>et al.</i> , 2009	<i>In vitro</i> human fibroblast time course	GSE12493
TNF $\alpha$	TNFR1/2	Rubins, <i>et al.</i> , 2011	24 h <i>in vitro</i> human fibroblast treatment	GSE24125
IFN $\alpha$	IFNAR	Rubins, <i>et al.</i> , 2011	24 h <i>in vitro</i> human fibroblast treatment	GSE24125
Poly(I-C)	TLR3	Rubins, <i>et al.</i> , 2011	24 h <i>in vitro</i> human fibroblast treatment	GSE24125
LPS	TLR4	Rubins, <i>et al.</i> , 2011	24 h <i>in vitro</i> human fibroblast treatment	GSE24125
DEX	Glucocorticoid receptors	Rubins, <i>et al.</i> , 2011	24 h <i>in vitro</i> human fibroblast treatment	GSE24125
Iono-PMA	T cells/Ca <sup>2+</sup> /PKC	Rubins, <i>et al.</i> , 2011	24 h <i>in vitro</i> human fibroblast treatment	GSE24125
Imatinib	Ableson kinase/PDGFR	Chung, <i>et al.</i> , 2009	Skin biopsies from <i>in vivo</i> clinical trial collected before and after treatment	GSE11130

doi:10.1371/journal.pone.0114017.t001

Table 2. Experimental pathway gene signatures and overlap with the MPH dataset.

Treatment	Number of probes passing filter <sup>a, b</sup>	Number of genes found in MPH dataset (% overlap) <sup>d</sup>
PDGF	1198	728 (60.8)
TGF $\beta$	946	842 (89.0)
S1P	848	825 (97.3)
IL-13	850	759 (89.3)
IL-4	1549	1415 (91.3)
RZN	222	128 (57.7)
LPS	1472	1185 (80.5)
PolyIC	4599	3749 (81.5)
TNF $\alpha$	1487	1184 (79.6)
IFN $\alpha$	262	223 (85.1)
Iono-PMA	3694	3040 (82.3)
Dex	1495	1151 (77.0)
Imatinib	1050 <sup>c</sup>	843 (80.3)

<sup>a</sup> Pathway gene signatures were defined as all genes up or downregulated  $\geq 2$ -fold across all 12 and 24 h time points, relative to untreated controls.

<sup>b</sup> IDs for PDGF, TGF $\beta$ , S1P, IL-13, IL-4, and RZN denote unique Agilent probe IDs. Entrez gene IDs were used for LPS, PolyIC, TNF $\alpha$ , IFN $\alpha$ , Iono-PMA, Dex, and imatinib; all genes represented by two or more probes were averaged in both the MPH dataset and individual gene signatures.

<sup>c</sup> The gene expression signature used for imatinib was determined based upon a p value cutoff, as defined by Chung, *et al.* [5].

<sup>d</sup> MPH overlap signifies the number of genes IDs from a given pathway also appearing in the MPH dataset; the low overlap percentages seen in both PDGF and PPAR $\gamma$  pathways is a result of platform differences, as both PDGF and PPAR $\gamma$  pathways were reanalyzed on Agilent 8  $\times$  60k DNA microarrays, while the MPH dataset includes only probes present in both 44k and 60k arrays.

doi:10.1371/journal.pone.0114017.t002

## Results

### Integrative analysis of the intrinsic subsets

*In vitro*, experimentally derived pathway signatures putatively deregulated in SSc provide an interpretive framework for previously generated skin biopsy data. Three distinct skin biopsy datasets consisting of 75 [6], 89 [2], and 165 ([7] and unpublished data) microarrays were merged using ComBat [8] to create a single microarray dataset (referred to as the MPH (Milano-Pendergrass-Hinchcliff) dataset). Together, these combined data include 329 microarray hybridizations from 287 unique biopsies representing 111 patients: 70 dSSc, 10 lSSc, 26 healthy controls, 4 morphea, and 1 eosinophilic fasciitis; one patient's diagnosis changed from lSSc to dSSc during the period of study. This combined dataset was used as a reference against which the relative contributions of different signaling pathways could be compared in a genome-wide meta-analysis.

### Functional gene expression groups

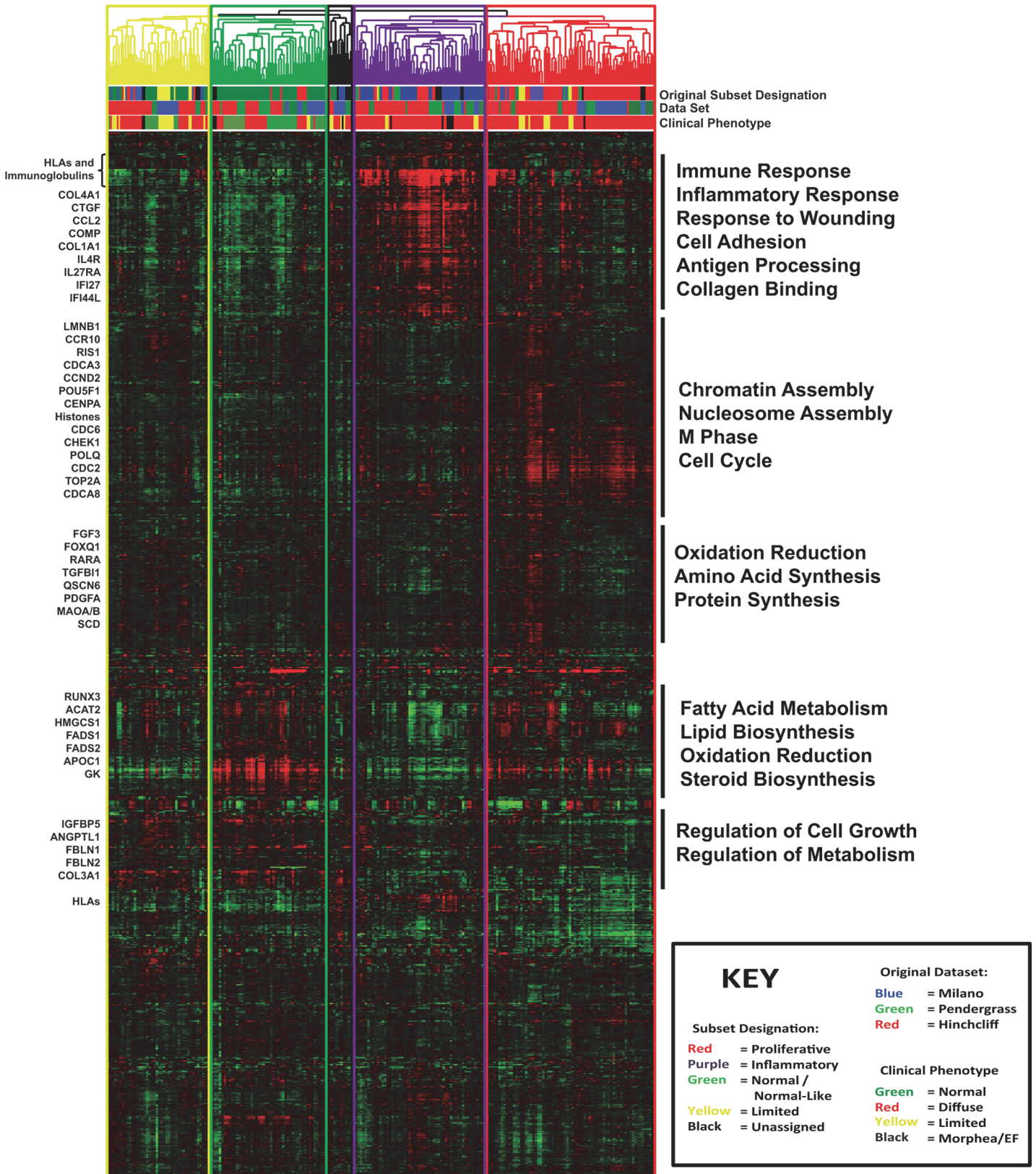
Clustering of the MPH dataset was performed as described previously [2,6,7], using the genes that showed the most intrinsic expression (e.g. the most consistent expression across all samples from a single patient but with the highest variance between different patients [9,14]). We selected 2316 probes covering 2189 unique genes at an estimated false discovery rate (FDR) of 0.65%. Average linkage hierarchical clustering was performed for both genes and arrays, recapitulating the four previously described 'intrinsic' subsets (fibroproliferative, inflammatory, limited, and normal-like; Fig. 1). A similar analysis performed using only a single array per patient revealed broadly similar results, indicating that the inclusion of multiple time points and technical replicates for some patients did not significantly affect the size of each subset (S2 Fig).

As the MPH dataset is composed of previously described biopsy samples, the intrinsic subsets assignments identified in this analysis are similar to those previously described. The methods required to merge these three datasets into one group does cause some samples at the edges of groups to be misclassified [2,6,7]. Subset assignments were largely consistent between the original and MPH datasets ( $p \leq 0.001$ , Chi-squared test; S1 Table). Strong reproducibility was observed in the inflammatory (47/75; 63%), and fibroproliferative (69/99; 70%) subsets where samples original classified as such were given the same classification here. The most common misclassification of fibroproliferative was to inflammatory (17 biopsies) and vice versa (16 inflammatory biopsies misclassified as fibroproliferative). Patients originally classified as the limited subset were typically classified as such here (12/19; 63%) and the most common misclassification was to fibroproliferative (6 biopsies). The normal-like subset showed the greatest variability (58/104 56%) with the majority of the misclassified samples (28 biopsies) being added to the limited dendrogram branch.

Gene clusters associated with each intrinsic subset were analyzed using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) to identify functional enrichment. Gene ontology (GO) biological process annotations recapitulated those previously described. The inflammatory subset include *inflammatory response*, *immune response*, *cell adhesion*, *angiogenesis*, and *antigen processing* ( $p < 0.001$ ; Fig. 1) and include multiple HLA and immunoglobulin genes, CTGF, CCL2, IL10RA, IL27RA, VEGFC, and genes associated with fibrosis (COMP, COL1A1, COL4A1, COL4A2, COL5A2, COL6A1, COL6A3, COL14A1, and COL15A1).

The fibroproliferative subset is significantly enriched for GO biological processes associated with the cell cycle including *chromatin assembly*, *nucleosome assembly*, *M phase*, and *cell cycle*





**Figure 1. Hierarchical clustering recreates intrinsic subsets.** Hierarchical clustering of the ComBat-merged MPH dataset recreates clear normal-like, fibroproliferative, inflammatory, and limited subsets. Clustering was performed on 2316 probes covering 2189 genes at an FDR of 0.65%, chosen based

upon their consistent expression within an individual patient, along with high variance between patients. The array tree is color coded to indicate new intrinsic subset designations (yellow = limited, green = normal-like, purple = inflammatory, red = fibroproliferative, and black = unassigned). Below the array tree, hash marks are used to indicate the original subset designation (TOP: green = normal-like, red = fibroproliferative, purple = inflammatory, yellow = limited, black = unassigned), the dataset of origin (MIDDLE: blue = Milano, green = Pendergrass, red = Hinchcliff), and the clinical diagnosis (BOTTOM: green = normal, red = diffuse scleroderma, yellow = limited scleroderma, black = morphea or eosinophilic fasciitis). Black bars indicate genes that clustered together hierarchically, with the most highly represented GO terms listed alongside each cluster.

doi:10.1371/journal.pone.0114017.g001

( $p < 0.001$ ; Fig. 1), and includes genes for cell cycle regulators CCNE1, CDCA5, CDKN2A, and CCNB2, as well as multiple histone genes.

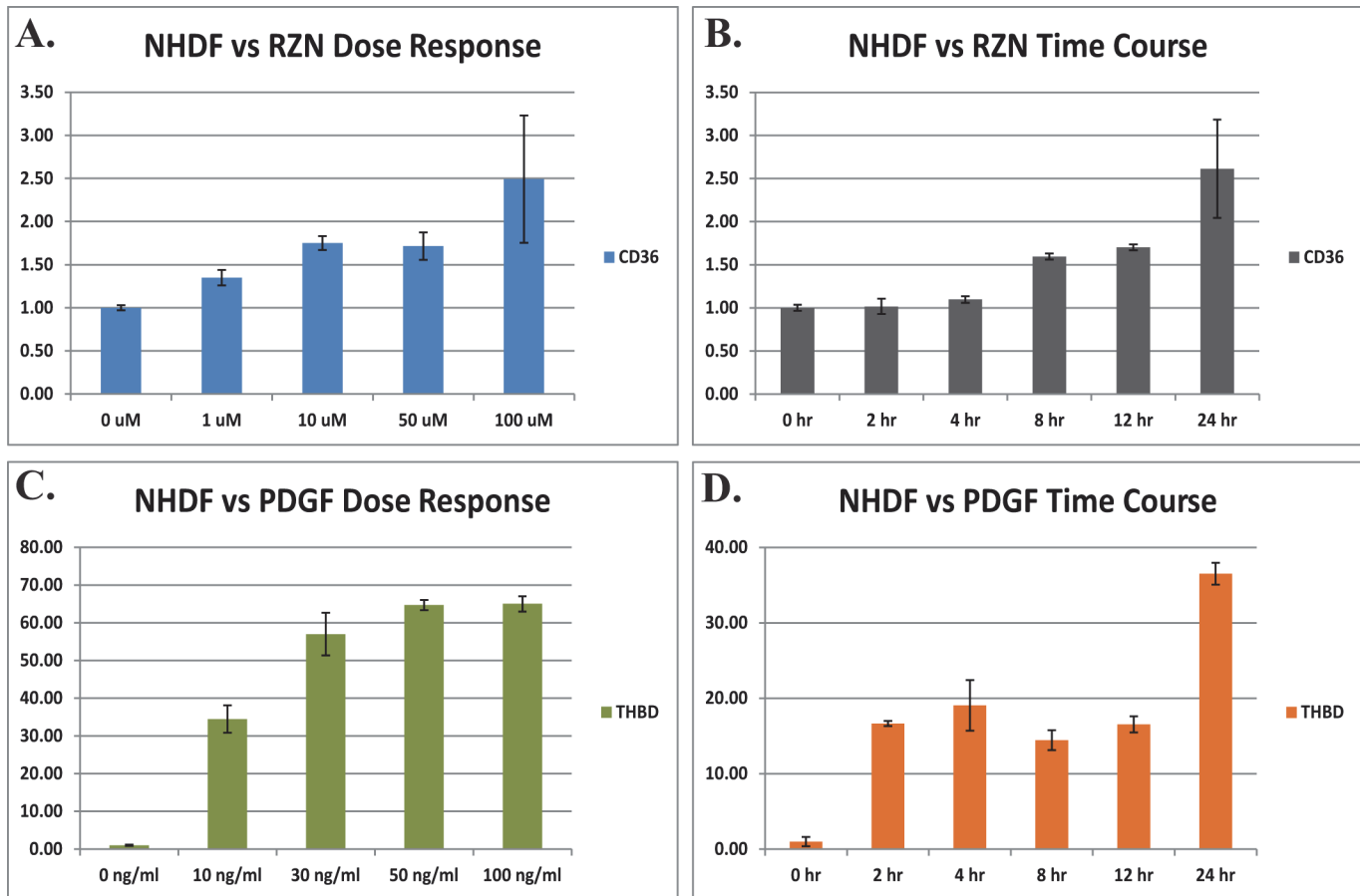
The normal-like and limited groups are defined primarily based upon the absence of immune or proliferation related gene expression, with the primary division between these groups driven by a strong lipid and fatty acid metabolism signature in the normal-like group which is significantly decreased in the limited subset. This lipid signature is characterized by GO biological processes of *fatty acid metabolism*, *lipid biosynthesis*, *oxidation reduction*, and *steroid biosynthesis* ( $p < 0.001$ ; Fig. 1). Genes principally involved in these processes include HMGCS1, fatty acid desaturases (FADS1 and FADS2), and acyl-CoA synthesis genes (ACADVL, ACAT2, ACOX2, and ACSL3).

## Generation of fibrotic pathway gene signatures in dermal fibroblasts

Using targets suggested by the literature, we performed treatment time courses for PDGF, S1P, and rosiglitazone (RZN), an agonist of PPAR $\gamma$ , in SSc and normal dermal fibroblasts to assess the role of each signaling pathway in SSc pathogenesis; we performed two additional time courses each for IL-4 and IL-13 to expand upon the work of Greenblatt *et al.* [4]. No significant differences were observed between the genes induced by the different treatments in SSc lesional and healthy control fibroblasts in culture, consistent with previous findings [3,15,16]. Optimal dosing for PDGF and RZN were determined experimentally (Fig. 2), with cellular responses measured by quantitative real-time PCR; dosing for S1P was chosen based upon published results [17,18]. A 10  $\mu$ M concentration of RZN resulted in a 1.7-fold induction of CD36 ( $p < 0.001$  vs. control), with only modest increases at higher concentrations (Fig. 2A). The gene expression response increased over the course of 24 h with 10  $\mu$ M (Fig. 2B). Accordingly, we chose 10  $\mu$ M for all RZN treatment time courses. Treatment with 30 ng/mL PDGF resulted in a 57-fold induction of thrombomodulin (THBD), with dosage above 50 ng/mL saturating ( $p < 0.001$  vs. control; Fig. 2C). Based upon these results a concentration of 30 ng/mL was used for all PDGF time course experiments. THBD expression increased sharply upon treatment with PDGF, with maximal induction seen at 24 h (Fig. 2D). Table 1 provides a summary of the time courses generated in this study. Each time course was analyzed independently (see methods) and then all pathways concatenated into a single data file (Fig. 3).

We first examined the genes and biological processes affected by each pathway treatment independently. Since IL-4, IL-13, and TGF $\beta$  have been described previously [3,4], transcriptional responses to these agonists are not discussed in detail.

The link between PDGF signaling and SSc pathogenesis is well established [19–22], with many PDGF-mediated effects, including cell proliferation, abnormal vascular development, and immune activation [23] having important implications for SSc. PDGF treatment of dermal fibroblasts resulted in the induction or suppression of 1198 probes covering 1099 unique genes (Table 2; S3 Fig.). Genes with increased expression were highly enriched for GO biological processes related to kinase activity, phosphorylation, wound healing, cytokine signaling, and smooth muscle cell proliferation ( $p < 0.001$ ). Specific genes include IL8R ligands (CXCL1, CXCL2, CXCL3), TNF receptor superfamily members (TNFRSF12A, TNFRSF21, TNFRSF6B, TNFRSF8), metallothioneins (MT1A, MT1B, MT1E, MT1L, MT2A), BCL2A1, CCL2, IFI44,



**Figure 2. Dosage response and induction of reporter genes following stimulation with PDGF and RZN.** RZN and PDGF concentrations were optimized for use in microarray treatment experiments by qRT-PCR using reporter genes CD36 and thrombospondin (THBD), respectively. NHDFs were treated with A) 0, 1, 10, 50, and 100 μM RZN or C) 0, 10, 30, 50, and 100 μg/mL PDGF for 24 h. Levels of CD36 and THBD were analyzed by qRT-PCR, and normalized to 18S rRNA. NHDFs were treated with B) 10 μM RZN and D) 30 ng/mL PDGF for 0, 0, 2, 4, 8, 12, and 24 h. Error bars indicate the standard deviation across three or more replicates; all time points were statistically significant relative to controls ( $p < 0.05$ ).

doi:10.1371/journal.pone.0114017.g002

and VEGF. Downregulated genes were enriched for GO biological processes associated with cell motility and migration, MAP kinase signaling, and Wnt receptor signaling. Genes downregulated by PDGF include CTGF, MAP3K8, and GATA6.

The lipid and fatty acid metabolism signature identified within the normal-like subset are indicative of increased PPAR $\gamma$  signaling, as suggested by Varga and coworkers [24–26]. PPAR $\gamma$  signaling exerts a potent anti-fibrotic response [27], and is antagonistic to TGF $\beta$  [25], suggesting a potential therapeutic role for this pathway in SSc. Activation of PPAR $\gamma$  signaling by RZN had only modest effects on fibroblasts in the absence of other signals. A total of 222 probes covering 219 unique genes were affected in this analysis, of which only 37 probes were upregulated including ADRP, ANGPTL4, and PDK4 (Table 2; S4 Fig.). Lowering of the 2-fold cutoff to 1.5-fold increased the overall number of probes to 985. This more permissive cutoff revealed enrichment for expected GO processes including regulation of lipid metabolism, lipid storage, and long-chain fatty acid synthesis ( $p < 0.05$ ). GO biological processes for downregulated genes are almost exclusively associated with cell cycle regulation, including the terms M phase, cell cycle, mitosis, nuclear division, spindle organization, and others ( $p < 0.001$ ); this result was seen with both 2 and 1.5-fold cutoffs.



**Figure 3. Clustering of pathway-regulated genes signatures reveals co-regulated and pathway specific modules.** A total of 2136 probes covering 2081 genes were identified which show  $\geq 2$ -fold average change in gene expression at 12–24 h in one or more of the six different pathways examined (IL-4, IL-13, S1P, TGF $\beta$ , PDGF, and RZN). Gene expression data from each of the eight time points (0, 0, 0, 2, 4, 8, 12, and 24 h) from each time course are shown. Black bars indicate genes that clustered together hierarchically, with the most highly represented GO terms listed alongside each cluster.

doi:10.1371/journal.pone.0114017.g003

S1P signaling has also been shown to play an important role in immune activation and regulation [28], with potent pro-fibrotic effects seen in both normal and SSc fibroblasts [29]. As S1P levels are regulated in part through TGF $\beta$  [18], this suggests both unique and overlapping functions associated with this pathway. S1P treatment induced the most diverse responses of any of the agonists tested, with  $\geq 2$ -fold induction or suppression seen in 848 probes covering 749 unique genes (Table 2; S5 Fig.). Upregulated GO biological processes included immune activation, inflammatory and wounding responses, regulation of cell death, and proliferation ( $p < 0.001$ ). Prominently induced pathways include IL8R, TGF $\beta$ , Toll-like receptor, PPAR, and VEGF signaling, along with substantial activation of interferon-inducible proteins, such as IFI44. Downregulated GO biological processes include metabolism of sugars (glucose, hexose, and monosaccharide), antigen processing and presentation, immune response, fatty acid synthesis, and cell adhesion ( $p < 0.001$ ).

### Identification of specific and overlapping functions for each pathway

Significant overlap exists between pathway gene signatures, particularly for fibrotic genes, making it difficult to identify pathway-specific effects. To better delineate the genes induced by multiple pathways (common) and those induced by a single pathway (specific), all probes showing  $\geq 2$ -fold change in expression across all 12 and 24 h time points were concatenated from each of our treatment pathways, and hierarchically clustered to identify functional gene clusters. Pathways included in this analysis were PDGF, RZN, and S1P, along with our expanded IL-4 and IL-13 time courses, and our previous data examining TGF $\beta$ -induced gene expression [3]. A total of 2136 probes covering 2081 genes were identified in one or more of the six pathways considered (Fig. 3); probes not present on both the 4 $\times$ 44k and 8 $\times$ 60k microarray platforms were excluded from this analysis.

The clustered data revealed several areas of divergence that may be important in the pathogenesis of SSc (Fig. 3). Cluster 1 is highly enriched for virtually all cell cycle associated genes present in this dataset and showed induction by PDGF at 12 and 24 h time points, while substantial downregulation was seen in all other pathways. Clusters 3 and 5 were most strongly associated with TGF $\beta$  signaling, exhibiting a strong decrease in lipid and steroid biosynthesis (Fig. 3, cluster 3), with increased expression of genes associated with cell differentiation, migration, and wound healing including CTGF and COL3A1; these genes were largely unaffected in the five other pathways tested.

Clusters 2 and 6 were selectively upregulated in S1P, exhibiting strong induction of multiple TLRs and interferon-inducible proteins, indicating a clear role for this pathway in innate immunity. Surprisingly, S1P showed a strong induction of the interferon-inducible proteins commonly observed in SSc and Lupus PBMC samples [30,31]. IL-8-related signaling (e.g. IL-8, CXCL1-3) was induced by both S1P and PDGF (Fig. 3, cluster 6), although PDGF lacked many of the other genes associated with innate immunity induced by S1P, including IL-6, NFKBIA, NFKBIE, TLR1, TLR2, and TLR4 (Fig. 3).

Cluster 7 was most strongly associated with IL-4/IL-13 signaling. GO terms associated with this cluster include Jak/STAT signaling, amino acid synthesis and transport, and extracellular matrix organization. CCL2 was among the genes highly upregulated in this cluster, consistent with previous findings [4]; however, increased CCL2 expression was also observed in S1P and

PDGF treatments, illustrating that activation of multiple signaling pathways can induce CCL2 expression.

In addition to pathway-specific effects, substantial convergence of pathways was also observed. Gene expression patterns are highly similar in both IL-4 and IL-13 signaling pathways due to their convergence on the shared IL4RA receptor (Fig. 3). Pathway-specific variations exist, though modest to strong downregulation is seen throughout cluster 4 for IL-4, IL-13, S1P, TGF $\beta$ , and PDGF, while the same pathways show consistent upregulation in clusters 8 and 10. Cluster 8 is most strongly activated in TGF $\beta$ , and includes many of the biological responses associated with fibrogenesis, including robust induction of epithelial to mesenchymal transition, cell motility, and Wnt signaling; however, this cluster is also upregulated to varying degrees in IL-4, IL-13, S1P, and PDGF, suggesting widespread convergence on these genes typically associated with fibrosis. Cluster 10, is consistently upregulated by all six pathways and is characterized by induction of multiple cellular biological processes including protein complex synthesis and mRNA regulation.

Together these analyses identify important pathway-specific effects of each agonist, including wound healing, cell proliferation, and immune activation. Furthermore, these analyses provide important data regarding many of the genes associated with fibrosis, and shows their regulation by multiple pathways in dermal fibroblasts. A pdf containing the full data from Fig. 3 is available among the supplemental materials (S6 Fig.).

### Curation of NF- $\kappa$ B-related signaling pathways and the imatinib response signature

Next, additional microarray data probing the response of dermal fibroblasts to a wide range of immunological perturbations were downloaded from the NCBI GEO database (Table 1). These pathways are particularly relevant to SSc due to the inflammatory gene expression observed in our skin biopsy dataset. *In vitro* fibroblast treatment data were obtained for TNF $\alpha$ , IFN $\alpha$ , lipopolysaccharide (LPS), polyinosinic-polycytidylic acid (poly(I-C)), ionomycin plus phorbol-12-myristate-13-acetate (ionomycin-PMA), and dexamethasone (DEX), [13]. TNF $\alpha$  and IFN $\alpha$  are among the first cytokines expressed during an innate immune response, and are important for the generation of adaptive T cell responses [32]. TNF $\alpha$  plays a major role in both acute and chronic inflammation [33], while IFN $\alpha$  acts as an important mediator of antiviral activity [34]. Both LPS and poly(I-C) initiate innate immune responses through Toll-like receptors, activating TLR4 and TLR3, respectively. Ionomycin-PMA raises intracellular Ca<sup>+</sup> levels, and induces protein kinase C (PKC) activation [35,36]. DEX is a synthetic glucocorticoid steroid which functions as a potent anti-inflammatory. Due to differences in platforms, gene annotation, and experimental design, microarray data from each of these treatments were processed independently; genes represented by multiple probes were averaged across all probes for both the treatment and MPH datasets. Each set of genes constitutes a 'signature' for that pathway.

The final set of data included in this study was taken from a case report study performed by Chung, *et al.* [5] examining the effect of imatinib mesylate on two dSSc patients. Imatinib is a selective tyrosine kinase inhibitor which blocks phosphorylation of Abelson kinase (Abl), a target of both TGF $\beta$  and PDGF, as well as the PDGF receptor (PDGFR). Microarray analyses were performed using skin biopsies collected before and after treatment, with the imatinib response signature determined based upon a *p*-value cutoff. We used the list of 1050 imatinib response genes as published by Chung *et al.* [5] (Table 1).

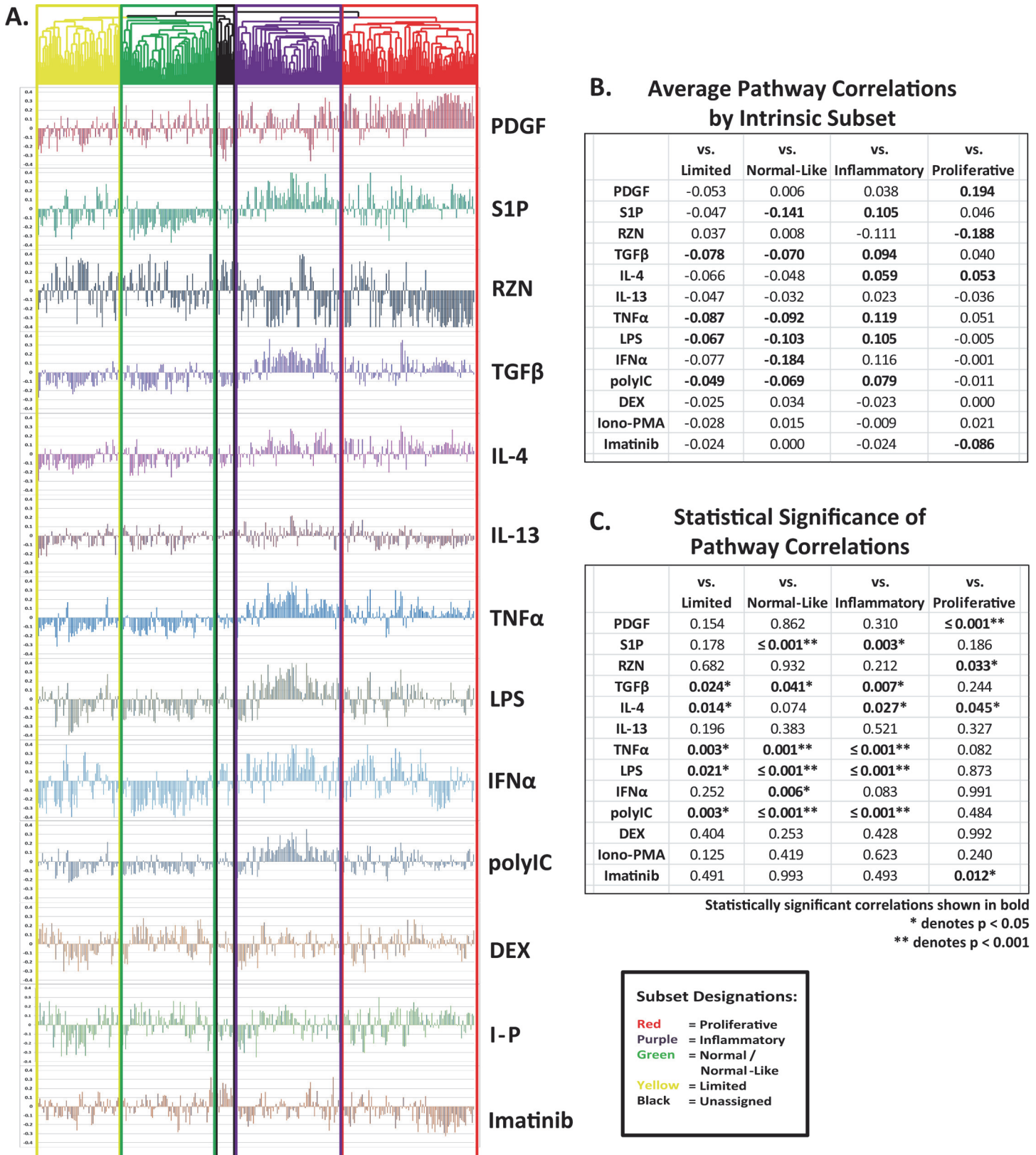
## Contributions of individual pathways within each intrinsic subset of disease

To identify the contribution of each pathway to the overall gene expression profile observed in patient biopsies, Pearson's correlations were performed comparing each of the thirteen gene expression signatures against the corresponding probes extracted from the MPH skin biopsy dataset. Due to differences in DNA microarray platforms, not every probe or Entrez gene ID induced by a pathway was present in the MPH dataset. The number of probes and Entrez gene IDs for each pathway, and the corresponding number present in the MPH dataset are shown in [Table 2](#). Correlations were then averaged for each intrinsic gene expression subset, as indicated by the colored coded dendrogram ([Fig. 4A](#)).

The relative contributions of each of the thirteen pathways to the differential gene expression that defined the four intrinsic subsets are represented by their average Pearson's correlation scores ([Fig. 4B](#)) and associated  $p$  values ([Fig. 4C](#)). The strongest correlation between any intrinsic subset and pathway signature was seen between the fibroproliferative subset and PDGF (average correlation = 0.194;  $p < 0.001$ ; [Fig. 4B and C](#)); the only other pathway showing a significant positive correlation to the fibroproliferative subset was IL-4, which exhibited modest, but consistent activation across the inflammatory subset, and variable activation in the fibroproliferative subset (average correlation = 0.053;  $p = 0.045$ ; [Fig. 4B and C](#)). Surprisingly, the average TGF $\beta$  correlation for the entire fibroproliferative subset was only 0.040 ( $p = 0.244$ ; [Fig. 4B and C](#)), indicating that sustained TGF $\beta$  expression is variable across patients in this subset. This observation was surprising however, given previous results obtained using the same TGF $\beta$  time courses and skin biopsy microarray data solely from Milano, *et al.* [3]. The fibroproliferative subset of arrays originally described in Milano, *et al.* does continue to show a positive correlation with TGF $\beta$ , however this average correlation is significantly lower than that seen with PDGF (TGF $\beta$  average correlation = 0.088;  $p = 0.011$  vs. PDGF average correlation = 0.258;  $p < 0.001$ ). RZN exhibited a significant negative correlation to this subset (average correlation = -0.188;  $p = 0.033$ ; [Fig. 4B and C](#)), due primarily to the strong downregulation of cell cycle genes seen in this pathway.

The data of Chung *et al.* [5] suggested that dSSc patients in the fibroproliferative subset were most likely to benefit from tyrosine kinase inhibitor therapy. A direct comparison of the imatinib response signature to the MPH dataset supports this view, with a significant negative correlation evident for the fibroproliferative subset (average correlation = -0.086;  $p = 0.012$ ; [Fig. 4B and C](#)). Imatinib also shows a strong negative correlation to the PDGF gene signature (= -0.824); this correlation is significantly stronger than with TGF $\beta$  (= -0.220), suggesting that the PDGF gene expression signature may also be a good predictor of response to imatinib.

The inflammatory subset is positively associated with the widest array of pathway signatures, likely due to the convergence of many of these pathways at NF- $\kappa$ B. The LPS, TNF $\alpha$ , S1P, poly(I-C), and TGF $\beta$  gene signatures are all significantly enriched within this subset ( $p < 0.001$  for all; [Fig. 4B and C](#)). The IFN $\alpha$  gene signature is also enriched in this subset, though this correlation fails to reach statistical significance due to the relatively small number of genes in this pathway (average correlation = 0.116;  $p = 0.083$ ). Enrichment of TGF $\beta$  signaling within this subset was surprising based on our prior findings; from these data it suggests that TGF $\beta$  signaling spans both the inflammatory and fibroproliferative subsets. TGF $\beta$  has been shown to induce both pro- and anti-inflammatory responses depending upon the presence of other cytokines [37], and can activate NF- $\kappa$ B by means of TGF $\beta$ -associated kinase 1 (TAK1) [38,39]. The strong correlation between S1P and the inflammatory subset, but not the fibroproliferative subset, was also surprising given the well documented roles for S1P in fibrosis, cell proliferation, and immune activation. Evidence from these analyses indicates a much stronger role for S1P in



**Figure 4. Correlations between pathway-specific gene signatures and patient gene expression profiles.** Pearson correlations were performed between each of the thirteen pathway-specific gene signatures and the corresponding probes in the MPH dataset. A. Pathway gene signatures are defined



as all probes exhibiting  $\geq 2$ -fold average change in gene expression across all 12 and 24 h time points for a given treatment (Table 2). Correlations were repeated across each of the 329 arrays and aligned using the array dendrogram from Fig. 1. Boxes representing each of the four intrinsic subsets (normal-like = green, fibroproliferative = red, inflammatory = purple, limited = yellow) are shown; arrays not clustering with any defined subset are indicated in black. B. Average Pearson's correlations for each pathway across each of the intrinsic subsets are provided. C. *P* values quantifying the enrichment of pathway signatures within individual subsets were calculated based upon the average Pearson's correlation, with statistically significant correlations highlighted in bold.

doi:10.1371/journal.pone.0114017.g004

immune activation in SSc. Combined, these correlations suggest a role for innate immune signaling through NF- $\kappa$ B as an important mediator of pathology within the inflammatory subset.

Consistent with our prior studies, both the IL-4 and IL-13 gene signatures are associated with the inflammatory subset. The IL-4 pathway is significantly enriched suggesting a role for T<sub>H</sub>2-like immune responses (average correlation = 0.059;  $p = 0.027$ ; Fig. 4B and C) in this subset. Despite its strong correlation with IL-4 (= 0.942;  $p < 0.001$ ), the IL-13 signature initially showed only weak correlation to this subset ( $p = 0.521$ ); however this difference was largely an artifact of 2-fold cutoff, as the IL-4 signature is almost twice the size of the IL-13 signature (1415 genes for IL-4 vs. 759 genes for IL-13). An equivalently sized 1415 gene signature from the IL-13 treatment showed enrichment in the inflammatory subset, although this correlation failed to reach statistical significance ( $p = 0.101$ ).

The limited and normal-like subsets show very similar gene expression, exhibiting negative correlations to almost all of the pathways tested. These negative correlations were particularly strong among the pathways activated in the inflammatory subset; S1P, TGF $\beta$ , TNF $\alpha$ , LPS, and poly(I-C) ( $p < 0.05$  for all; Fig. 4B and C), indicative of a more immunologically quiescent form of disease. The primary distinction between the two subsets was the high level of gene expression associated with lipid signaling in the normal-like subset. Surprisingly, the RZN gene signature exhibited no enrichment within this subset despite being an agonist for many of the upregulated genes. This absence of correlation is likely due to the low number of genes positively affected by RZN in the fibroblast, indicating that that fibroblasts are not the primary source of lipid signaling seen in these patients.

### TGF $\beta$ is associated with increased disease severity while IFN $\alpha$ is associated with early disease

Pearson's correlations for each of the thirteen pathways were compared against clinically relevant factors including age, sex, skin score (MRSS), biopsy site, and disease duration to identify specific associations between individual pathways and disease outcomes. Clinical variables including lung disease, gastrointestinal involvement, renal disease, Raynaud's severity, race, and autoantibody profile were not considered due to insufficient data across the multiple skin biopsy cohorts analyzed. Clinically limited SSc, morphea, and eosinophilic fasciitis patients were excluded from this analysis due to underlying differences in MRSS, age, and disease duration between clinical subsets, which limited to the analysis solely to dSSc patients. We limited the analysis to a single microarray per patient per time point collected; in cases where both lesional and non-lesional biopsies were collected only the lesional biopsy was considered.

Multiple signaling pathways exhibited strong correlations with MRSS (Table 3). Of the six agonists with significant correlation to MRSS (S1P, RZN, IL-4, TGF $\beta$ , TNF $\alpha$ , and imatinib), TGF $\beta$  was by far the strongest overall predictor of severity of skin involvement, with a correlation score nearly double that of the next highest pathway (average Pearson's correlation = 0.385 for TGF $\beta$  vs. 0.210 for IL-4).

In addition to MRSS, the TGF $\beta$  gene signature was also strongly associated with biopsy site, showing a significant increase in TGF $\beta$  activation in lesional skin (average correlation = 0.058 vs. 0.002 in forearm and back, respectively;  $p = 0.001$ ; Table 3). Alternatively, PDGF signaling

Table 3. Associations with clinical outcomes.

		IL-4	IL-13	PDGF	RZN	S1P	TGFβ	DEX	IFNα	Iono-PMA	LPS	PolyIC	TNFα	Imatinib
mRSS	Pearson	0.210	-0.108	0.134	-0.210	0.183	0.385	0.095	-0.057	0.073	0.117	0.098	0.179	-0.182
<i>n</i> = 116	<i>P</i> value	<b>0.023*</b>	0.136	0.064	<b>0.004*</b>	<b>0.011*</b>	<b>≤ 0.001*</b>	0.189	0.431	0.312	0.106	0.178	<b>0.013*</b>	<b>0.011*</b>
Age	Pearson	0.003	0.040	-0.106	0.069	-0.017	-0.038	-0.054	0.034	-0.033	0.033	0.025	-0.023	0.203
<i>n</i> = 116	<i>P</i> value	0.979	0.584	0.144	0.344	0.812	0.601	0.461	0.641	0.647	0.654	0.730	0.749	<b>0.005*</b>
Disease Duration	ANOVA	2.386	0.372	0.868	0.453	2.141	0.084	2.998	5.098	2.240	0.075	0.145	1.313	3.067
<i>n</i> = 115	<i>P</i> value	0.125	0.543	0.354	0.502	0.146	0.773	0.086	<b>0.026*</b>	0.137	0.785	0.704	0.254	0.083
Sex	ANOVA	0.002	1.451	0.150	0.013	0.278	0.480	0.201	0.428	0.749	0.590	0.768	0.142	0.060
<i>n</i> = 116	<i>P</i> value	0.963	0.230	0.699	0.910	0.599	0.489	0.654	0.514	0.388	0.443	0.382	0.707	0.806
Biopsy Site	ANOVA	0.012	0.335	3.962	2.325	0.613	12.170	0.986	3.311	2.273	0.719	1.476	0.551	7.097
<i>n</i> = 152	<i>P</i> value	0.914	0.564	<b>0.048*</b>	0.129	0.435	<b>0.001*</b>	0.322	0.071	0.134	0.398	0.226	0.459	<b>0.009*</b>

Pearson’s correlations comparing each of the arrays and pathways tested were used to quantify the overall contribution of a given pathway within an individual patient. These scores were then compared against clinically relevant factors including age, sex, modified Rodnan skin score (MRSS), biopsy site, and disease duration to assess the predictive value of each pathway for disease outcomes. Early disease was defined as disease ≤ 2 years after first non-Raynaud’s phenomenon symptoms. Comparisons were performed with clinically diffuse patients only, using a single array per patient for each time point collected. Comparisons of biopsy site were limited to clinically diffuse patients which provided paired lesional and non-lesional biopsies at a given time point; *n* denotes the number of patients included in each analysis. Continuous variables were compared using Pearson’s correlation; categorical variables were analyzed by ANOVA.

\* denotes *p* < 0.05.

doi:10.1371/journal.pone.0114017.t003

appears elevated in non-lesional back skin (average correlation = 0.126 vs. 0.078 in back and forearm, respectively; *p* = 0.048; Table 3). These observations suggest subtle, but reproducible differences between lesional and unaffected skin [2,6,15], and may reflect differences between TGFβ and PDGF-driven disease.

Disease duration showed a significant negative correlation to IFNα pathway activation (Table 3), indicating a spike in IFNα signaling early in disease pathogenesis (defined as ≤ 2 years after first non-Raynaud’s symptoms), followed by downregulation of this pathway as disease progresses. Other inflammatory signals, including S1P and IL-4 were also higher in early disease though these signals did not reach statistical significance (*p* > 0.05; Table 3).

Finally, comparisons between the inflammatory and fibroproliferative subsets are suggestive of a weak association between disease stage and subset, with the fibroproliferative subset containing 8 of 11 patients with disease lasting ≥ 8 years, though this enrichment was not statistically significant (*p* = 0.104 by Chi-squared test). Both age and sex were comparable between subsets.

Taken together, these data suggest that IFNα signaling and other immune activation pathways may play a role in early disease pathogenesis, while TGFβ signaling is most strongly associated with disease severity. The observation that TGFβ spans the inflammatory and fibroproliferative subsets suggests a mechanistic connection may exist between these groups, driven in part by TGFβ signaling.

## Discussion

Scleroderma is a clinically heterogeneous disease that is likely to be caused by a network of pathways with distinct and overlapping effects. One way of determining the degree to which each pathway contributes to disease pathogenesis is to have a list of genes induced by each pathway in the primary cell type of interest, dermal fibroblasts. The data presented here

provide a framework by which we can query and dissect the molecular signaling pathways that contribute to each of the intrinsic subsets.

The inflammatory subset shows strong positive correlations with a wide array of signaling pathways, with significant overlap in the induced genes. The most obvious point of convergence is NF- $\kappa$ B, a signaling intermediary shared between LPS, poly(I-C), IFN $\alpha$ , TNF $\alpha$ , S1P, and TGF $\beta$ . Indeed, many of these pathways appear to be directly linked in SSc; TLR signaling was found to induce strong upregulation of both type I IFNs (IFN $\alpha$  and IFN $\beta$ ) and TGF $\beta$  in SSc skin and fibroblasts [40], providing a mechanism through which these signals may be linked. This convergence on TLRs and NF- $\kappa$ B is consistent with reports implicating innate immune activation in SSc pathogenesis [41].

In addition to NF- $\kappa$ B-mediated signaling, activation of other pathways within the inflammatory subset suggests distinct cell populations that may contribute to SSc pathology, providing hypotheses that can be tested experimentally. Strong IL-4-related gene expression in the inflammatory subset is consistent with T<sub>H</sub>2-like immune responses in these patients. Combined with the clear co-occurrence of TGF $\beta$  and innate immune signals, these data suggest a central role for alternatively activated (M2) macrophages in the inflammatory subset of SSc. M2 macrophages are known to be induced by a combination of T<sub>H</sub>2 cytokines, such as IL-4 and IL-13, in combination with TGF $\beta$  [42], and likely play key roles in SSc pathogenesis. Evidence for M2 macrophages has been observed in SSc lesional skin [43], lung [44–46], and serum [47], showing that these cells are likely to be involved in the initiation of fibrosis.

In addition to T<sub>H</sub>2-like immune responses, growing evidence suggests a role for T<sub>H</sub>17 cells in the pathogenesis of SSc with clear differences between diffuse and limited disease [48–52]. T<sub>H</sub>17-like immune responses have been implicated in a wide range of autoimmune conditions, including multiple sclerosis, systemic lupus erythematosus, psoriasis, neuromyelitis optica, Crohn's disease, inflammatory bowel disease, and rheumatoid arthritis, suggesting a common mechanism of pathology associated with autoimmunity [53–56]. Parallels drawn between SSc and other autoimmune diseases may help to explain some of the contradictory signals seen in SSc, including activation of type I IFNs within the inflammatory subset. Under normal conditions type I IFNs are potent inhibitors of T<sub>H</sub>17 activity [51]; however, in many autoimmune diseases these signals actually enhance T<sub>H</sub>17 responses, exacerbating disease [53]. While the TGF $\beta$  and TNF $\alpha$  gene expression signatures seen in some patients in the inflammatory subset, in conjunction with pervasive inflammatory infiltrates, are consistent with a T<sub>H</sub>17-like immune response [37], additional pathway analyses examining other cytokines, such as IL-6 and IL-17, will be necessary to determine the relative contribution of T<sub>H</sub>17-like responses in each of the intrinsic subsets, as well as the presence of these signals over time.

Analysis of clinical covariates revealed a clear association between the TGF $\beta$  gene signature and increased MRSS severity, consistent with previous findings [3]. The strong association between the TGF $\beta$  gene signature and clinically affected forearm skin likely reflects the increased fibrosis at these sites.

The gene expression signature most strongly associated with the fibroproliferative subset was PDGF, which was not measured in our prior work [3]. The association is driven primarily by the strong upregulation of cell cycle and other proliferation-related genes, in contrast to TGF $\beta$ , which is traditionally regarded as an inhibitor of cell proliferation. Emerging evidence suggests that TGF $\beta$  signaling may span the inflammatory and fibroproliferative subsets, providing a potential mechanistic link between these two groups (Mahoney *et al.*, Submitted). If we were to consider an interpretation of the intrinsic subsets as mechanistic stops in disease progression rather than independent groups, expression of TGF $\beta$  during the initial inflammatory phase may play a role in the initiation of fibrosis, while sustained expression of TGF $\beta$  may induce greater expression of PDGF [57], leading to a more proliferative form of disease.

In addition to TGF $\beta$ , the timing of IFN $\alpha$  signaling may play a role in regulating the transition from the inflammatory to fibroproliferative subset. Under certain conditions, type I interferons are capable of inhibiting both PDGF activation and PDGF-mediated collagen expression [58]. Downregulation of IFN signaling would remove these inhibitory signals, hastening the transition to a more PDGF-driven, proliferative form of disease. Such a process may explain some of the negative treatment outcomes associated with anti-IFN $\alpha$  therapy in SSc, including a worsening of disease symptoms following therapy [59]. Such an outcome highlights the need for a better understanding of the interrelationship of SSc associated pathways, how they may change during disease progression, and if combination therapies could more effectively stop disease progression.

Beyond the actions of TGF $\beta$  alone, the maintenance and progression of fibrotic phenotypes has been shown to be driven in part by the mechanical environment [60]. Specific evidence regarding this phenomenon has recently been extended to SSc, with changes in the cell-matrix sufficient to perpetuate pro-fibrotic responses, even in the absence of other stimuli [61]. As heightened matrix stiffness has been shown to increase signaling through PDGFR [62], this suggests a mechanism by which physical changes in affected tissues can perpetuate disease after the initial inflammation has been resolved. Clearance of inflammation alone may therefore be insufficient for resolving disease phenotypes.

Patients clustering to the limited and normal-like subsets exhibited near-zero to negative correlations against all thirteen agonists tested, indicative of a non-proliferative, immunologically quiescent state of disease. Further longitudinal studies will be necessary to determine how these patients progress from a clinical standpoint, and whether they transition into another more active subset of disease over time.

One possible model suggested by our analysis of patient biopsy data is that of a cascade of signaling pathways generating the progressive disease we know as SSc. A progressive model of pathogenesis, in which each intrinsic subset represents a distinct phase of disease progression, provides the simplest interpretation of the data. A weakness of this model is that we have not been able to capture patients changing subsets when analyzing patients longitudinally over 6 to 12 months. However, this could simply mean that patients move between intrinsic subsets very slowly over time or in a way that is difficult to capture experimentally with longitudinal biopsies.

Direct validation of this progressive model of disease pathogenesis has not been performed due to the absence of appropriate model systems, and the duration of time necessary to observe these changes in patients; however, all of the agonists and cell types implicated in this model have been well documented in SSc. Agonists such as TGF $\beta$  [3,63,64], PDGF [20,65], IL-4 [66], IL-13 [4,66], IFN $\alpha$  [67], S1P [68], and TNF $\alpha$  [69] are present in the skin, sera, and bronchoalveolar fluid of SSc patients, while cell types such as M2 macrophages [43,45] and T<sub>H</sub>2 cells [70] have also been described. While considerable effort will be necessary to validate such a model, it provides a framework from which to link seemingly divergent observations into a single, comprehensive model of disease pathogenesis. Longitudinal studies examining gene expression and cytokine profiles, along with direct confirmation of the cell types involved in each step, will be necessary to clearly define the processes underlying each stage of disease progression.

## Supporting Information

**S1 Fig. Principal Component Analysis of Merged Datasets.** The statistical significance of batch bias before and after adjustment was assessed using guided principal component analysis (gPCA) and the first two unguided principal components were inspected. The proportion of the variance associated with each unguided principal component is labeled on the axes. *P*

values  $\leq 0.05$  are indicative of significant batch bias.  
(EPS)

**S2 Fig. Hierarchical clustering recreates intrinsic subsets.** Hierarchical clustering of the ComBat-merged MPH dataset recreates clear normal-like, fibroproliferative, inflammatory, and limited subsets. Clustering was performed on 2316 probes covering 2189 genes at an FDR of 0.65%, chosen based upon their consistent expression within an individual patient, along with high variance between patients. The array tree is color coded to indicate new intrinsic subset designations (yellow = limited, green = normal-like, purple = inflammatory, red = fibroproliferative, and black = unassigned). Below the array tree, hash marks are used to indicate the original subset designation (TOP: green = normal-like, red = fibroproliferative, purple = inflammatory, yellow = limited, black = unassigned), the dataset of origin (MIDDLE: blue = Milano, green = Pendergrass, red = Hinchcliff), and the clinical diagnosis (BOTTOM: green = normal, red = diffuse scleroderma, yellow = limited scleroderma, black = morphea or eosinophilic fasciitis). Black bars indicate genes that clustered together hierarchically, with the most highly represented GO terms listed alongside each cluster.  
(EPS)

**S3 Fig. Hierarchical clustering of PDGF time courses.** Normal human dermal fibroblasts and SSc-derived dermal fibroblasts were treated with 30 ng/mL PDGF, with samples harvested at 0, 2, 4, 8, 12, and 24 h. Data shown include all probes exhibiting  $\geq 2$ -fold change in expression relative to untreated controls across all 12 and 24 h time points. Genes were clustered using Cluster 3.0, and visualized with Java TreeView.  
(EPS)

**S4 Fig. Hierarchical clustering of RZN time courses.** Normal human dermal fibroblasts were treated with 10  $\mu$ M RZN, with samples harvested at 0, 2, 4, 8, 12, and 24 h. Data shown include all probes exhibiting  $\geq 2$ -fold change in expression relative to untreated controls across all 12 and 24 h time points. Genes were clustered using Cluster 3.0, and visualized with Java TreeView.  
(EPS)

**S5 Fig. Hierarchical clustering of S1P time courses.** Normal human dermal fibroblasts and were treated with S1P, with samples harvested at 0, 2, 4, 8, 12, and 24 h. Data shown include all probes exhibiting  $\geq 2$ -fold change in expression relative to untreated controls across all 12 and 24 h time points. Genes were clustered using Cluster 3.0, and visualized with Java TreeView.  
(EPS)

**S6 Fig. Searchable version of Fig. 3.** A searchable version of Fig. 3 including gene names for all probes exhibiting a  $\geq 2$ -fold average change in gene expression at 12–24 h in one or more of the six different pathways examined.  
(EPS)

**S1 Table. Patients included in this study.** Full clinical data and associated pathway correlation scores for all patients and biopsies included in this study.  
(XLS)

**S2 Table. 2-fold IDs for all pathways included in this study.** Agilent probe IDs and Entrez gene IDs of all genes up- or downregulated  $\geq 2$ -fold across all 12 and 24 h time points for each pathway tested.  
(XLSX)

## Acknowledgments

MEJ would like to thank members of the Whitfield lab for helpful discussions.

## Author Contributions

Conceived and designed the experiments: MEJ MLW. Performed the experiments: MEJ JLS EM MRW. Analyzed the data: MEJ JMM JT MLW. Contributed reagents/materials/analysis tools: MEJ JMM JT JV MEH MLW. Wrote the paper: MEJ JMM JT JV MEH MLW.

## References

1. LeRoy EC, Black C, Fleischmajer R, Jablonska S, Krieg T, et al. (1988) Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. *J Rheumatol* 15: 202. PMID: [3361530](#)
2. Pendergrass SA, Lemaire R, Francis IP, Mahoney JM, Lafyatis R, et al. (2012) Intrinsic Gene Expression Subsets of Diffuse Cutaneous Systemic Sclerosis Are Stable in Serial Skin Biopsies. *J Invest Dermatol* 132: 1363–1373. doi: [10.1038/jid.2011.472](#) PMID: [22318389](#)
3. Sargent JL, Milano A, Bhattacharyya S, Varga J, Connolly MK, et al. (2009) A TGF $\beta$ -responsive gene signature is associated with a subset of diffuse scleroderma with increased disease severity. *J Invest Dermatol* 130: 694–705. doi: [10.1038/jid.2009.318](#) PMID: [19812599](#)
4. Greenblatt MB, Sargent JL, Farina G, Tsang K, Lafyatis R, et al. (2012) Interspecies Comparison of Human and Murine Scleroderma Reveals IL-13 and CCL2 as Disease Subset-Specific Targets. *Am J Pathol* 180: 1080–1094. doi: [10.1016/j.ajpath.2011.11.024](#) PMID: [22245215](#)
5. Chung L, Fiorentino D, BenBarak M, Adler A, Mariano M, et al. (2009) Molecular Framework for Response to Imatinib Mesylate in Systemic Sclerosis. *Arth Rheum* 60: 584–591. doi: [10.1002/art.24221](#) PMID: [19180499](#)
6. Milano A, Pendergrass SA, Sargent JL, George LK, McCalmont TH, et al. (2008) Molecular Subsets in the Gene Expression Signatures of Scleroderma Skin. *PLoS ONE* 3: e2696. PMID: [18648520](#)
7. Hinchcliff M, Huang C-C, Wood TA, Mahoney JM, Martyanov V, et al. (2013) Molecular Signatures in Skin Associated with Clinical Improvement during Mycophenolate Treatment in Systemic Sclerosis. *J Invest Dermatol*.
8. Johnson WE, Li C, Rabinovic A (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8: 118–127. PMID: [16632515](#)
9. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, et al. (2000) Molecular portraits of human breast tumours. *Nature* 406: 747–752. PMID: [10963602](#)
10. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. *PNAS* 95: 14863–14868. PMID: [9843981](#)
11. Saldanha AJ (2004) Java Treeview—extensible visualization of microarray data. *Bioinformatics* 20: 3246–3248. PMID: [15180930](#)
12. Livak KJ, Schmittgen TD (2001) Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the  $2^{-\Delta\Delta CT}$  Method. *Methods* 25: 402–408. PMID: [11846609](#)
13. Rubins KH, Hensley LE, Relman DA, Brown PO (2011) Stunned silence: gene expression programs in human cells infected with monkeypox or vaccinia virus. *PLoS ONE* 6: e15615. PMID: [21267444](#)
14. Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, et al. (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *PNAS* 98: 10869–10874.
15. Whitfield ML, Finlay DR, Murray JI, Troyanskaya OG, Chi JT, et al. (2003) Systemic and cell type-specific gene expression patterns in scleroderma skin. *PNAS* 100: 12319–12324. PMID: [14530402](#)
16. Gardner H, Shearstone JR, Bandaru R, Crowell T, Lynes M, et al. (2006) Gene profiling of scleroderma skin reveals robust signatures of disease that are imperfectly reflected in the transcript profiles of explanted fibroblasts. *Arth Rheum* 54: 1961–1973. PMID: [16736506](#)
17. Huang LS, Berdyshev E, Mathew B, Fu P, Gorshkova IA, et al. (2013) Targeting sphingosine kinase 1 attenuates bleomycin-induced pulmonary fibrosis. *FASEB J*. doi: [10.1096/fj.12-219634](#) PMID: [23315259](#)
18. Donati C, Marseglia G, Magi A, Serrati S, Cencetti F, et al. (2011) Sphingosine 1-Phosphate Induces Differentiation of Mesoangioblasts towards Smooth Muscle. A Role for GATA6. *PLoS ONE* 6: e20389. doi: [10.1371/journal.pone.0020389](#) PMID: [21629665](#)

19. Pandolfi A, Florita M, Altomare G, Pigatto P, Donati MB, et al. Increased plasma levels of platelet-derived growth factor activity in patients with progressive systemic sclerosis; 1989. Royal Society of Medicine. pp. 1–4.
20. Ludwicka A, Ohba T, Trojanowska M, Yamakage A, Strange C, et al. (1995) Elevated levels of platelet derived growth factor and transforming growth factor-beta 1 in bronchoalveolar lavage fluid from patients with scleroderma. *J Rheumatol* 22: 1876–1883. PMID: [8991985](#)
21. Iwayama T, Olson LE (2013) Involvement of PDGF in Fibrosis and Scleroderma: Recent Insights from Animal Models and Potential Therapeutic Opportunities. *Curr Rheumatol Reports* 15: 1–6. doi: [10.1007/s11926-012-0304-0](#) PMID: [23307576](#)
22. Olson LE, Soriano P (2009) Increased PDGFR $\alpha$  Activation Disrupts Connective Tissue Development and Drives Systemic Fibrosis. *Dev Cell* 16: 303–313.
23. Olson LE, Soriano P (2011) PDGFR $\beta$  signaling regulates mural cell plasticity and inhibits fat development. *Dev Cell* 20: 815–826.
24. Wei J, Bhattacharyya S, Varga J (2010) Peroxisome proliferator-activated receptor gamma: innate protection from excessive fibrogenesis and potential therapeutic target in systemic sclerosis. *Curr Opin Rheumatol* 22: 671–676. doi: [10.1097/BOR.0b013e32833de1a7](#) PMID: [20693905](#)
25. Wei J, Ghosh AK, Sargent JL, Komura K, Wu M, et al. (2010) PPAR $\gamma$  downregulation by TGF $\beta$  in fibroblast and impaired expression and function in systemic sclerosis: a novel mechanism for progressive fibrogenesis. *PLoS ONE* 5: e13778.
26. Ghosh AK, Bhattacharyya S, Lakos G, Chen S-J, Mori Y, et al. (2004) Disruption of transforming growth factor  $\beta$  signaling and profibrotic responses in normal skin fibroblasts by peroxisome proliferator-activated receptor  $\gamma$ . *Arth Rheum* 50: 1305–1318.
27. Ghosh AK, Bhattacharyya S, Wei J, Kim S, Barak Y, et al. (2009) Peroxisome proliferator-activated receptor- $\gamma$  abrogates Smad-dependent collagen stimulation by targeting the p300 transcriptional coactivator. *FASEB J* 23: 2968–2977.
28. Bode C, Gräler MH (2012) Immune regulation by sphingosine 1-phosphate and its receptors. *Arch Immunol Ther Exper* 60: 3–12. PMID: [22159476](#)
29. Bu S, Asano Y, Bujor A, Highland K, Hant F, et al. (2010) Dihydrosphingosine 1-phosphate has a potent antifibrotic effect in scleroderma fibroblasts via normalization of phosphatase and tensin homolog levels. *Arth Rheum* 62: 2117–2126. doi: [10.1002/art.27463](#) PMID: [20309867](#)
30. Assassi S, Mayes MD, Arnett FC, Gourh P, Agarwal SK, et al. (2010) Systemic sclerosis and lupus: Points in an interferon-mediated continuum. *Arth Rheum* 62: 589–598. doi: [10.1002/art.27224](#) PMID: [20112391](#)
31. Baechler EC, Battiwalla FM, Karypis G, Gaffney PM, Ortmann WA, et al. (2003) Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. *PNAS* 100: 2610–2615. PMID: [12604793](#)
32. Tompkins WA (1999) Immunomodulation and therapeutic effects of the oral use of interferon-alpha: mechanism of action. *J Interferon Cytokine Res* 19: 817–828. PMID: [10476925](#)
33. Chu W-M (2013) Tumor necrosis factor. *Cancer Letters* 328: 222–225.
34. Choubey D, Moudgil KD (2011) Interferons in autoimmune and inflammatory diseases: regulation and roles. *J Interferon Cytokine Res* 31: 857–865. PMID: [22149411](#)
35. Boldrick JC, Alizadeh AA, Diehn M, Dudoit S, Liu CL, et al. (2002) Stereotyped and specific gene expression programs in human innate immune responses to bacteria. *PNAS* 99: 972–977. PMID: [11805339](#)
36. Diehn M, Alizadeh AA, Rando OJ, Liu CL, Stankunas K, et al. (2002) Genomic expression programs and the integration of the CD28 costimulatory signal in T cell activation. *Sci Signaling* 99: 11796. PMID: [12195013](#)
37. Steinman L (2007) A brief history of TH17, the first major revision in the TH1/TH2 hypothesis of T cell-mediated tissue damage. *Nature Med* 13: 139–145. PMID: [17290272](#)
38. Hamidi A, von Bulow V, Hamidi R, Winssinger N, Barluenga S, et al. (2012) Polyubiquitination of Transforming Growth Factor  $\beta$  (TGF $\beta$ )-associated Kinase 1 Mediates Nuclear Factor- $\kappa$ B Activation in Response to Different Inflammatory Stimuli. *J Biol Chem* 287: 123–133.
39. Gingery A, Bradley EW, Pederson L, Ruan M, Horwood NJ, et al. (2008) TGF-beta coordinately activates TAK1/MEK/AKT/NF $\kappa$ B and SMAD pathways to promote osteoclast survival. *Exper Cell Res* 314: 2725–2738. doi: [10.1016/j.yexcr.2008.06.006](#) PMID: [18586026](#)
40. Farina GA, York MR, Di Marzio M, Collins CA, Meller S, et al. (2010) Poly (I: C) drives type I IFN-and TGF $\beta$ -mediated inflammation and dermal fibrosis simulating altered gene expression in systemic sclerosis. *J Invest Dermatol* 130: 2583–2593. doi: [10.1038/jid.2010.200](#) PMID: [20613770](#)

41. O'Reilly S (2014) Innate immunity in systemic sclerosis pathogenesis. *Clin Sci* 126: 329–337.
42. Atamas SP, White B (2003) Cytokine regulation of pulmonary fibrosis in scleroderma. *Cytokine Growth Factor Rev* 14: 537–550. PMID: [14563355](#)
43. Higashi-Kuwata N, Makino T, Inoue Y, Takeya M, Ihn H (2009) Alternatively activated macrophages (M2 macrophages) in the skin of patient with localized scleroderma. *Experimental Dermatol* 18: 727–729. doi: [10.1111/j.1600-0625.2008.00828.x](#) PMID: [19320738](#)
44. Homer RJ, Herzog EL (2010) Recent advances in pulmonary fibrosis: implications for scleroderma. *Curr Opin Rheumatol* 22: 683–689. doi: [10.1097/BOR.0b013e32833ddcc9](#) PMID: [20693906](#)
45. Pechkovsky DV, Prasse A, Kollert F, Engel KM, Dentler J, et al. (2010) Alternatively activated alveolar macrophages in pulmonary fibrosis—mediator production and intracellular signal transduction. *Clin Immunol* 137: 89–101. doi: [10.1016/j.clim.2010.06.017](#) PMID: [20674506](#)
46. Mathai SK, Gulati M, Peng X, Russell TR, Shaw AC, et al. (2010) Circulating monocytes from systemic sclerosis patients with interstitial lung disease show an enhanced profibrotic phenotype. *Laboratory Invest* 90: 812–823. doi: [10.1038/labinvest.2010.73](#) PMID: [20404807](#)
47. Nakayama W, Jinnin M, Makino K, Kajihara I, Makino T, et al. (2012) Serum levels of soluble CD163 in patients with systemic sclerosis. *Rheumatol International* 32: 403–407. PMID: [21120485](#)
48. Mathian A, Parizot C, Dorgham K, Trad S, Arnaud L, et al. (2012) Activated and resting regulatory T cell exhaustion concurs with high levels of interleukin-22 expression in systemic sclerosis lesions. *Ann Rheum Dis* 71: 1227–1234. doi: [10.1136/annrheumdis-2011-200709](#) PMID: [22696687](#)
49. Murata M, Fujimoto M, Matsushita T, Hamaguchi Y, Hasegawa M, et al. (2008) Clinical association of serum interleukin-17 levels in systemic sclerosis: Is systemic sclerosis a Th17 disease? *J Dermatol Sci* 50: 240–242. PMID: [18329249](#)
50. Radstake TR, van Bon L, Broen J, Hussiani A, Hesselstrand R, et al. (2009) The pronounced Th17 profile in systemic sclerosis (SSc) together with intracellular expression of TGF $\beta$  and IFN $\gamma$  distinguishes SSc phenotypes. *PLoS ONE* 4: e5903. doi: [10.1371/journal.pone.0005903](#) PMID: [19536281](#)
51. Rodríguez-Reyna TS, Furuzawa-Carballeda J, Cabiedes J, Fajardo-Hermosillo LD, Martínez-Reyes C, et al. (2012) Th17 peripheral cells are increased in diffuse cutaneous systemic sclerosis compared with limited illness: a cross-sectional study. *Rheumatology international* 32: 2653–2660.
52. Deleuran B, Abraham DJ (2007) Possible implication of the effector CD4 $^{+}$  T-cell subpopulation TH17 in the pathogenesis of systemic scleroderma. *Nat Clin Pract Rheum* 3: 682–683.
53. Axtell RC, Raman C (2012) Janus-like effects of type I interferon in autoimmune diseases. *Immunol Rev* 248: 23–35. doi: [10.1111/j.1600-065X.2012.01131.x](#) PMID: [22725952](#)
54. Brand S (2009) Crohn's disease: Th1, Th17 or both? The change of a paradigm: new immunological and genetic insights implicate Th17 cells in the pathogenesis of Crohn's disease. *Gut* 58: 1152–1167.
55. Camporeale A (2012) IL-6, IL-17, and STAT3: a holy trinity in auto-immunity? *Front Biosci* 17: 2306–2326. PMID: [22652781](#)
56. Kimura A, Kishimoto T (2011) Th17 cells in inflammation. *Int Immunopharmacol* 11: 319–322.
57. Trojanowska M (2008) Role of PDGF in fibrotic diseases and systemic sclerosis. *Rheumatology* 47: v2–v4. PMID: [18784131](#)
58. Mallat A, Preaux AM, Blazejewski S, Rosenbaum J, Dhumeaux D, et al. (1995) Interferon alfa and gamma inhibit proliferation and collagen synthesis of human Ito cells in culture. *Hepato* 21: 1003–1010. PMID: [7705772](#)
59. Black CM, Silman AJ, Herrick AI, Denton CP, Wilson H, et al. (1999) Interferon- $\alpha$  does not improve outcome at one year in patients with diffuse cutaneous scleroderma: Results of a randomized, double-blind, placebo-controlled trial. *Arth Rheum* 42: 299–305.
60. Hinz B (2009) Tissue stiffness, latent TGF- $\beta$ 1 activation, and mechanical signal transduction: implications for the pathogenesis and treatment of fibrosis. *Curr Rheumatol Reports* 11: 120–126.
61. Gerber EE, Gallo EM, Fontana SC, Davis EC, Wigley FM, et al. (2013) Integrin-modulating therapy prevents fibrosis and autoimmunity in mouse models of scleroderma. *Nature* 503: 126–130. doi: [10.1038/nature12614](#) PMID: [24107997](#)
62. Brown XQ, Bartolak-Suki E, Williams C, Walker ML, Weaver VM, et al. (2010) Effect of substrate stiffness and PDGF on the behavior of vascular smooth muscle cells: implications for atherosclerosis. *J Cell Physiol* 225: 115–122. doi: [10.1002/jcp.22202](#) PMID: [20648629](#)
63. Falanga V, Gerhardt CO, Dasch JR, Takehara K, Ksander GA (1992) Skin distribution and differential expression of transforming growth factor  $\beta$ 1 and  $\beta$ 2. *J Dermatol Sci* 3: 131–136. PMID: [1498091](#)
64. Sfikakis PP, McCune BK, Tsokos M, Aroni K, Vayiopoulos G, et al. (1993) Immunohistological demonstration of transforming growth factor- $\beta$  isoforms in the skin of patients with systemic sclerosis. *Clin Immunol Immunopathol* 69: 199–204. PMID: [8403557](#)



65. Gay S, Jones RE, Huang G-q, Gay RE (1989) Immunohistologic demonstration of platelet-derived growth factor (PDGF) and sis-oncogene expression in scleroderma. *J Invest Dermatol* 92: 301–303. PMID: [2918236](#)
66. Hasegawa M, Fujimoto M, Kikuchi K, Takehara K (1997) Elevated serum levels of interleukin 4 (IL-4), IL-10, and IL-13 in patients with systemic sclerosis. *J Rheumatol* 24: 328–332. PMID: [9034992](#)
67. Wu M, Assassi S (2013) The role of type 1 interferon in systemic sclerosis. *Front Immunol* 4. PMID: [24046769](#)
68. Tokumura A, Carbone LD, Yoshioka Y, Morishige J, Kikuchi M, et al. (2009) Elevated serum levels of arachidonoyl-lysophosphatidic acid and sphingosine 1-phosphate in systemic sclerosis. *Int J Med Sci* 6: 168. PMID: [19521548](#)
69. Bolster MB, Ludwicka A, Sutherland SE, Strange C, Silver RM (1997) Cytokine concentrations in bronchoalveolar lavage fluid of patients with systemic sclerosis. *Arth Rheum* 40: 743–751. PMID: [9125259](#)
70. Truchetet M-E, Brembilla NC, Montanari E, Allanore Y, Chizzolini C (2011) Increased frequency of circulating Th22 in addition to Th17 and Th2 lymphocytes in systemic sclerosis: association with interstitial lung disease. *Arth Res Ther* 13: R166. doi: [10.1186/ar3486](#) PMID: [21996293](#)