



Published in final edited form as:

*J Invest Dermatol.* 2010 March ; 130(3): . doi:10.1038/jid.2009.318.

## A TGF $\beta$ -Responsive Gene Signature Is Associated with a Subset of Diffuse Scleroderma with Increased Disease Severity

Jennifer L. Sargent<sup>1</sup>, Ausra Milano<sup>1</sup>, Swati Bhattacharyya<sup>2</sup>, John Varga<sup>2</sup>, M. Kari Connolly<sup>3</sup>, Howard Y. Chang<sup>4</sup>, and Michael L. Whitfield<sup>1</sup>

<sup>1</sup>Department of Genetics, Dartmouth Medical School, Hanover, New Hampshire, USA

<sup>2</sup>Division of Rheumatology, Department of Medicine, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, USA

<sup>3</sup>Department of Dermatology and Medicine (Rheumatology), University of California, San Francisco, California, USA

<sup>4</sup>Program in Epithelial Biology and Cancer Biology Program, Stanford University School of Medicine, Stanford, California, USA

### Abstract

Systemic sclerosis is a complex disease with widespread skin fibrosis and variable visceral organ involvement. Since transforming growth factor- $\beta$  (TGF $\beta$ ) has been implicated in driving fibrosis in systemic sclerosis, a mechanism-derived gene expression signature was used to assay TGF $\beta$ -responsive gene expression in the skin of patients with systemic sclerosis (SSc). Primary dermal fibroblasts from patients with diffuse SSc (dSSc) and healthy controls were treated with TGF $\beta$ , and the genome-wide gene expression was measured on DNA microarrays over a time course of 24 hours. Eight hundred and ninety-four probes representing 674 uniquely annotated genes were identified as TGF $\beta$  responsive. Expression of the TGF $\beta$ -responsive signature was examined in skin biopsies from 17 dSSc, seven limited SSc (lSSc), three morphea patients, and six healthy controls. The TGF $\beta$ -responsive signature was expressed in 10 out of 17 dSSc skin biopsies, but was not found in lSSc, morphea, or healthy control biopsies. Expression of dSSc the TGF $\beta$ -responsive signature stratifies patients into two major groups, one of which corresponds to the “diffuse-proliferation” intrinsic subset that showed higher modified Rodnan skin score and a higher likelihood of scleroderma lung disease. The TGF $\beta$ -responsive signature is found in only a subset of dSSc patients who could be targeted by specific therapies.

---

© 2009 The Society for Investigative Dermatology

Correspondence: Dr Michael L. Whitfield, Department of Genetics, Dartmouth Medical School, 7400 Remsen, Hanover, New Hampshire 03755, USA. michael.whitfield@dartmouth.edu.

#### Data access

Microarray data have been deposited at GEO (<http://www.ncbi.nlm.nih.gov/geo/>; accession no. GSE12493). The complete data set and searchable versions of the figures are available at <http://whitfieldlab.dartmouth.edu/TGFB/>.

#### SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at <http://www.nature.com/jid>

#### CONFLICT OF INTEREST

ML Whitfield has filed a patent application for gene expression biomarkers in scleroderma. This intellectual property has been licensed to a startup company, Celdara Medical, LLC, that is aiming to translate the discovery into clinical use. Dr Whitfield holds an interest in this company. The remaining authors state no conflict of interest.

## INTRODUCTION

Systemic sclerosis (SSc) is a systemic autoimmune disease characterized by skin fibrosis, internal organ dysfunction, vascular abnormalities, and immune activation. Clinical presentation is heterogeneous and patients with clinically indistinguishable early-stage disease progress to different clinical endpoints. There is a need for objective and quantitative measures of disease severity that can identify relative likelihoods of disease progression.

The most widely recognized classification of SSc groups patients into systemic sclerosis with diffuse scleroderma (dSSc) and systemic sclerosis with limited scleroderma (lSSc), largely by the degree of skin involvement (LeRoy *et al.*, 1988). Efforts have been made to further distinguish the heterogeneity in SSc patients (Barnett *et al.*, 1988a, b; Ferri *et al.*, 2002; Scussel-Lonzetti *et al.*, 2002; Maricq and Valter, 2004), but no further disease classification has been widely accepted.

Genome-wide measurement of gene expression with DNA microarrays can easily distinguish skin biopsies from patients with dSSc from healthy controls (Whitfield *et al.*, 2003; Gardner *et al.*, 2006), and identification of gene expression-based subsets has been reported (Milano *et al.*, 2008). Multiple distinct gene expression signatures were identified among patients with scleroderma (Milano *et al.*, 2008), with some of these groups delineating existing divisions such as those between dSSc and lSSc, and others reflecting more subtle changes in the cellular infiltrates or in the expression of different cellular processes. The signaling pathways underlying these subsets have not been mapped. Here we extend these findings by showing genes responsive to transforming growth factor- $\beta$  (TGF $\beta$ ) *in vitro* are associated with the “diffuse-proliferation” subset of dSSc patients.

Skin fibrosis, generally considered to be driven by TGF $\beta$ -activated fibroblasts, is a hallmark of dSSc (LeRoy *et al.*, 1989; Smith and LeRoy, 1990; Cotton *et al.*, 1998; Leask and Abraham, 2004; Varga, 2004; Leask, 2006; Verrecchia *et al.*, 2006; Ihn, 2008). Elevated levels of TGF $\beta$  have been observed in SSc skin biopsies (Gabielli *et al.*, 1993; Sfrikakis *et al.*, 1993) and TGF $\beta$ -activated fibroblasts produce collagen I, III, and V in addition to other matrix proteins such as glycoaminoglycans (Wynn, 2008). Numerous studies have identified TGF $\beta$ -regulated genes that show increased expression in SSc fibroblasts, including fibronectin (Xu *et al.*, 1991), collagens  $\alpha$ 1(I) and  $\alpha$ 1(III) (Kuroda and Shinkai, 1997), collagen- $\alpha$ 2(1) (Jinnin *et al.*, 2006), endoglin (Leask *et al.*, 2002), connective tissue growth factor (Leask, 2004), and cartilage oligomeric matrix protein (Farina *et al.*, 2006). SSc fibroblasts also show an increase in TGF $\beta$ R1 and TGF $\beta$ R2 mRNA levels and surface protein expression when compared with healthy dermal fibroblasts (Kubo *et al.*, 2002), suggesting activation of the TGF $\beta$  pathway in these cells. Activation of the platelet-derived growth factor (PDGF) pathway by stimulatory autoantibodies to the PDGF receptor has also been shown and may target many of the same genes (Baroni *et al.*, 2006). A clinical trial using anti-TGF $\beta$  therapy in dSSc patients has been reported, but the results were inconclusive (Denton *et al.*, 2007).

We report the use of a mechanism-derived gene expression signature to determine whether a TGF $\beta$ -response is found in the gene expression subsets of SSc. We defined a TGF $\beta$ -responsive signature over time courses of 24 hours in four independent cultures of dermal fibroblasts (two healthy, two dSSc) and identified 894 probes representing 674 TGF $\beta$ -responsive genes. The signature was analyzed *in vivo* in a microarray data set analyzing skin biopsies from patients with scleroderma and healthy controls (Milano *et al.*, 2008). The TGF $\beta$ -responsive signature is found in the “diffuse-proliferation” subset of dSSc skin biopsies. This indicates that TGF $\beta$  deregulation, or a related pathway, contributes to pathogenesis in this subset.

## RESULTS

We identified the genes responsive to TGF $\beta$  on a genome-wide scale with DNA microarrays in adult dermal fibroblasts from patients with dSSc and healthy controls. Four independent primary fibroblast cultures were analyzed, two from dSSc lesional forearm skin biopsy, one from a healthy control forearm skin biopsy, and one commercially available adult dermal fibroblast culture (Table 1). To compare our results to prior studies of TGF $\beta$  responses in fibroblasts, we used conditions similar to those previously reported (Chambers *et al.*, 2003; Renzoni *et al.*, 2004). Fibroblasts were cultured for 7–9 passages post explant in 10% serum and then brought to quiescence in 0.1% serum for 24 hours before TGF $\beta$  treatment. Fewer than 1% of the cells were in S-phase after 24 hours in 0.1% serum, determined by BrdU incorporation (data not shown).

To determine optimal experimental conditions, a dosage curve was generated with 50–300 pM TGF $\beta$  and the cellular response measured by quantitative real-time PCR (qRT-PCR). We found 50 pM TGF $\beta$  resulted in an 11-fold induction of plasminogen activator inhibitor 1 (PAI1) mRNA, whereas treatment with 100 pM TGF $\beta$  resulted in 16-fold induction (Figure 1). Concentrations of 200–300 pM resulted in decreased relative PAI1 levels, indicating that a concentration higher than 100 pM was saturating. We chose to use 50 pM for all further experiments.

Quiescent cells were exposed to 50 pM TGF $\beta$  and total RNA was collected at 0, 2, 4, 8, 12, and 24 hours after treatment. A mock time course without TGF $\beta$ , but in otherwise identical conditions, was performed in a single culture of normal dermal fibroblasts. The mock time course identifies genes that may respond to treatment effects that are independent of TGF $\beta$ . Total RNA from each sample was amplified, labeled, and hybridized to whole-genome DNA microarrays in a common reference design. Eight microarrays were hybridized for each time course resulting in a total of 40 microarray hybridizations.

### Genome-wide response to TGF $\beta$ in adult dermal fibroblasts

To characterize the genome-wide response to TGF $\beta$  across the four independent fibroblasts cultures, probes were selected that changed at least 1.74-fold in at least eight of the 32 arrays. Eight hundred and ninety-four TGF $\beta$ -responsive probes were selected representing 674 uniquely annotated genes (Figure 2a and Supplementary Data File 1). The gene expression data for eight probes corresponding to seven known TGF $\beta$  targets in dermal fibroblasts is shown (Figure 2b). Examination of these probes in the mock, no TGF $\beta$ , control shows that the observed changes in gene expression are specific to a TGF $\beta$  response (Figure 2a).

To identify differential TGF $\beta$  responses between normal and scleroderma fibroblasts, the data from TGF $\beta$ -treated normal fibroblast was compared to that of TGF $\beta$ -treated SSc fibroblasts with significance analysis of microarrays (SAM) using slope and area functions, in a two-class unpaired time-course analysis. Only a single gene, Early Growth Response-1 (EGR1), was significantly differentially induced at an false discovery rate (FDR) <5%. Since one would expect this number or more to be called as falsely significantly by chance alone, this is likely a false positive. Another recent study also failed to find significant differences in the TGF $\beta$  responses in cultured lung fibroblasts from SSc patients and normal controls (Chambers *et al.*, 2003).

### TGF $\beta$ module map

To summarize the biological programs represented by genes in the TGF $\beta$ -responsive signature, we created a module map that identifies groups of Gene Ontology (GO) terms

(biological process, cellular component, or molecular function) coordinately regulated by TGF $\beta$  in cultured adult dermal fibroblasts. Each row shows a unique module that is a set of coordinately regulated genes involved in similar processes.

Modules with significantly enriched genes ( $P < 0.05$ , hypergeometric distribution) in at least 16 of the 20 arrays were selected. A subset of modules is shown (Figure 2c); the full figure with all modules is available as Supplementary Figure S1. Those induced after TGF $\beta$  treatment included proteinaceous extracellular matrix, extracellular matrix, and extracellular space, associated with collagen production, extracellular matrix deposition, and remodeling (Figure 2c). TGF $\beta$  is a mediator of fibrosis in different organs including skin, kidneys, lungs, and liver, and is necessary for tissue remodeling and wound healing (Varga and Whitfield, 2009; Wynn, 2008).

Modules associated with development programs included angiogenesis, skeletal development, muscular development, vasculature development, multicellular organismal development, organ morphogenesis, and anatomical structure morphogenesis. The role of TGF $\beta$  and its family members in embryogenesis, tissue specification, and organ development through to tissue maintenance and repair have been extensively characterized (Wall and Hogan, 1994; Padgett *et al.*, 1997; Mummery, 2001).

Among the enriched, downregulated GO terms were mitotic cell cycle, cell division, and chromosomal segregation, indicating that TGF $\beta$  downregulates the proliferation signature (Whitfield *et al.*, 2006) under low-serum conditions.

### **TGF $\beta$ -responsive gene expression underlies the “diffuse-proliferation” intrinsic subset**

To determine whether the TGF $\beta$ -responsive signature was expressed in the intrinsic subsets of scleroderma, we analyzed the expression of the 894 TGF $\beta$ -responsive signature probes (674 genes) in the data set of Milano *et al.* (2008), which included dSSc, lSSc, morphea, and healthy controls. Expression data for each of the 894 probes was extracted from the set of 75 microarrays and organized by hierarchical clustering (Figure 3). Organization of the samples was compared to the sample groupings reported using an intrinsic gene set (Milano *et al.*, 2008). The group expressing the TGF $\beta$ -responsive signature is similar to the diffuse-proliferation subset (Supplementary Table S1), suggesting that deregulated TGF $\beta$  signaling, or a related pathway, may underlie this subset. Several related pathways and modulators have been implicated in SSc, including endoglin, which has increased expression on endothelial cells (Dharmapatni *et al.*, 2001), and fibroblasts (Leask *et al.*, 2002), endothelin-1, which is a downstream mediator of TGF $\beta$  responses, (Shi-Wen *et al.*, 2007), and the related PDGF pathway, implicated by the presence of stimulatory autoantibodies to the PDGF receptor (Baroni *et al.*, 2006). Activation of the PDGF pathway in SSc has proven controversial (Classen *et al.*, 2009). Consistent with deregulation of the TGF $\beta$  and PDGF pathways, several studies have demonstrated that imatinib mesylate prevents fibrosis in murine models of fibrosis (Daniels *et al.*, 2004; Distler *et al.*, 2007; Akhmet-shina *et al.*, 2008). Additionally, a case report showed a gene expression response associated with disease improvement in two patients after treatment with imatinib mesylate (Chung *et al.*, 2009).

Samples with high expression of the TGF $\beta$ -responsive signature include only dSSc skin biopsies, and this enrichment is statistically significant ( $P < 1 \times 10^{-4}$ ,  $\chi^2$ -test). One patient classified as diffuse-proliferation and one patient classified as inflammatory did not show consistent expression of the TGF $\beta$ -responsive signature (Supplementary Table S1). Therefore, there is not a precise overlap of the TGF $\beta$ -responsive group and the proliferative group identified by Milano *et al.* The group without the TGF $\beta$ -responsive signature contains a subset of the dSSc biopsies, all lSSc biopsies, morphea samples, and healthy controls.

Therefore, the TGF $\beta$ -responsive signature, a set of genes independent from the 995 intrinsic genes, is expressed in a subset of dSSc patients, but is not expressed in lSSc, morphea, or healthy controls.

### TGF $\beta$ -responsive gene expression in dSSc patients

To determine whether the TGF $\beta$ -responsive gene signature was consistently found in both lesional and non-lesional skin, or associated with specific clinical phenotypes, we analyzed the signature in the diffuse patients alone. The expression values for each of the 894 genes in the TGF $\beta$ -responsive signature were extracted from the skin biopsy data set and the 53 microarrays representing only patients with dSSc and normal controls. Samples were ordered by hierarchical clustering; the resulting sample dendrogram shows clear bifurcation of the skin biopsies (Figure 4a). The left branch, highlighted in red, is comprised solely of samples from dSSc patients, while the right branch includes the remaining dSSc samples and all healthy controls. Samples from two patients, dSSc2 and dSSc8, split between the two groups and, therefore, could not be assigned to either group. Of the patients who were consistently classified, each showed deregulation of the TGF $\beta$ -responsive gene signature in both lesion forearm and non-lesion back biopsies. Therefore, the systemic nature of the disease extends to specific pathways.

The significance of the classification was determined with Statistical Significance of Clustering (SigClust) (Liu *et al.*, 2008), which tests the robustness of the sample bifurcation and found that the clustering is highly significant ( $P < 0.001$ ). We find the TGF $\beta$ -activated group to be a single statistically significant cluster. The TGF $\beta$ -not-activated group forms two statistically significant clusters. We did not further investigate this additional sub-grouping given the relatively small size of the groups.

Alignment of the skin biopsy gene expression data with that from the *in vitro* TGF $\beta$  time courses reveals that expression of the signature is heterogeneous (Figure 4b). We identified the subset of the 894 probes driving the sample bifurcation. A two-class unpaired SAM analysis identified 474 probes to be significantly differentially expressed between the two groups (FDR < 1.15%; Figure 5). The centroid for the 474 differentially expressed probes is shown (Figure 5 and Supplementary Data File 2). The activation of the TGF $\beta$ -responsive signature in each patient sample was determined by calculating the Pearson correlation coefficients between the centroid and the gene expression for each skin biopsy (Figure 5). For simplicity, we have termed the group on the left (red dendrogram) TGF $\beta$ -activated because of high positive correlation coefficients associated with these samples, and the group on the right (black dendrogram) TGF $\beta$ -not activated owing to the predominantly negative correlation coefficients.

The genes in Figure 5 can be classified into two groups based on their expression patterns in the fibroblast time courses and skin biopsies. Those concordantly and those discordantly expressed between fibroblasts and skin, and further characterized as up- or downregulated in fibroblasts upon TGF $\beta$  treatment. Genes found concordantly upregulated in TGF $\beta$ -stimulated fibroblasts and skin included canonical targets ID3, PDGFA, SPHK1 (Supplementary Table S2), genes involved in organ and tissue development (HOXB2 and HOXC8), and genes associated with apoptosis (ENC1 and CARD4). Genes concordantly downregulated included those involved in cell signaling (PDGFRA and IL6R) and phospholipids binding (PLD1 and PIK3C2B).

Genes induced in the TGF $\beta$ -treated fibroblasts, but with low expression in the TGF $\beta$ -activated skin biopsies, were enriched for developmental processes and signaling (RUNX2, BMPR2, and BMP6). Genes showing decreased expression in TGF $\beta$ -treated fibroblasts, but increased expression in the skin biopsies, were those associated with cell cycle (the



proliferation signature; Whitfield *et al.*, 2006; CDCA2, CDCA8, CCNB2, and CENPF). High expression of these genes was observed previously in the diffuse-proliferation subset and we found increased KI67 staining in the epidermis of the skin biopsies (Milano *et al.*, 2008).

### Validation of TGF $\beta$ signature genes in patient samples

We validated the expression of three genes from the TGF $\beta$ -responsive signature, which were also the most highly differentially expressed between the groups with and without the TGF $\beta$ -responsive gene signature (Figure 6). We analyzed the expression of E2F7, which is also a cell-cycle-regulated gene (Whitfield *et al.*, 2002), growth differentiation factor-6, which is a member of the TGF $\beta$  superfamily, and actin- $\alpha$ 2. Each gene was measured in representative patient samples by qRT-PCR (Figure 6a–c) and compared to the expression levels from the DNA microarray (Figure 6d–f), in all cases the gene expression follows the same trends between the TGF $\beta$ -activated *versus* TGF $\beta$ -not-activated groups.

### The TGF $\beta$ -responsive signature is associated with more severe skin and lung disease

We examined the severity and incidence of available clinical parameters, calculated both by patient and by biopsy, to determine whether the dSSc patients in the TGF $\beta$ -activated group showed phenotypic differences from those in the TGF $\beta$ -not-activated group. Since patients dSSc2 and dSSc8 could not be accurately assigned, they were excluded from the analysis, leaving 10 patients in the TGF $\beta$ -activated group and five in the TGF $\beta$ -not-activated group. Each group was analyzed for differences in clinical covariates (Table 2). Student's *t*-tests were used to analyze differences in patient age, disease duration as defined by onset of first non-Raynaud symptoms, mean modified Rodnan skin score (MRSS) (0–51), Raynaud phenomenon (0–10), and digital ulcers (0–3). All other parameters for which we had clinical data, including incidence of interstitial lung disease (ILD), impaired renal function, and gastrointestinal involvement were scored as either present or absent and a  $\chi^2$ -test used to assess differences between the groups (Table 2). When individual biopsies were considered, the TGF $\beta$ -activated group showed significantly higher skin scores (mean =  $26.9 \pm 2.04$ ) than the TGF $\beta$ -not-activated group (mean =  $17.8 \pm 1.95$ ;  $P = 0.0061$ ). When performed on a per patient basis, the difference in the means is weakly significant ( $P = 0.11$ ). Examination of the dot plots for MRSS shows consistent difference in the means, with no significant outliers (Figure 7). We found that ILD was significantly more prevalent among the TGF $\beta$ -activated samples (7 of 16) as compared with TGF $\beta$ -not-activated group (0 of 10;  $P = 0.014$ ; Table 2), with an odds ratio of  $\approx 16.58$ . When patients rather than biopsies are considered, then five out of 10 patients in the TGF $\beta$ -activated group had ILD, whereas none of the five patients in the TGF $\beta$ -not-activated group had ILD ( $P = 0.053$ ). Since this type of exploratory analysis has not been performed previously, we are more concerned about type-II errors (false negatives) than type-I errors (false positives). As such, we consider the  $P$ -value of 0.053 to be strongly suggestive, although it does not meet a strict 0.05 significance level, and should be considered a hypothesis that should be tested in a larger cohort of patients. The analysis by biopsy is important because it shows the TGF $\beta$  signature is reproducible from different skin samples (lesion forearm and non-lesion back) from the same patient. We found only a weak association with disease duration (biopsies,  $P = 0.086$ ; patients,  $P = 0.14$ ). Examination of the dot plots of disease duration shows that the TGF $\beta$ -activated group is skewed by two outliers from patients with disease durations of 20 years (Figure 7). We found no significant associations with any of the other clinical variables, including potential immune modulating therapies or with anti-topoisomerase-I status (Table 2).

## DISCUSSION

Experimental determination of a mechanism-derived TGF $\beta$  signature in fibroblasts has allowed us to analyze the genes responsive to this pathway in skin biopsies from patients with scleroderma and healthy controls. This has shown that the TGF $\beta$ -responsive gene signature is expressed heterogeneously across the spectrum of scleroderma, but shows high expression in a subset of dSSc patients with higher MRSS and increased incidence of lung involvement. The TGF $\beta$ -responsive signature is not found in patients with lSSc, morphea, or in healthy controls. Given the established role for TGF $\beta$  in mediating fibrosis, it is not surprising that such a signature shows increased expression in dSSc skin, but the heterogeneity of expression among the SSc patients was not expected. This provides further evidence for subsets of patients that can be distinguished by gene expression and demonstrates that this heterogeneity extends to the pathway level. Most importantly, it suggests that TGF $\beta$  activation, or activation of a related pathway such as PDGF, may underlie the “diffuse-proliferation” subset of scleroderma.

The finding that a gene signature expressed in skin is associated with the occurrence of lung disease is surprising and to our knowledge is previously unreported. ILD is the leading cause of death among patients with dSSc and approximately 70% of patients show evidence of significant lung pathology at autopsy (Ostojic *et al.*, 2007; Steen and Medsger, 2007). Recent work has developed tools and methods for diagnosis, staging, and characterization of ILD in dSSc patients (Goh *et al.*, 2008); however, biomarkers that reliably predict who will develop lung complications before they become symptomatic would be beneficial. Surprisingly, only a weak association was found between the TGF $\beta$ -responsive gene signature and disease duration, raising the possibility that the signature is stable over the disease course in this subset of patients.

Concordance between the TGF $\beta$ -responsive gene signature defined *in vitro* and that found *in vivo* is variable. Gene expression observed in skin is likely influenced by multiple soluble factors and signals from the extracellular matrix. Thus, the heterogeneity of the signature in skin likely results from differences between single-cell-type cultures stimulated with a single cytokine *versus* the complexity of whole skin that contains multiple cell types in distinct tissue-specific niches. In addition, many TGF $\beta$ -responsive genes are also responsive to other profibrotic cytokines and signaling molecules such as IL13, IL4, sphingosine-1 phosphate (JL Sargent and ML Whitfield, unpublished data), endoglin (Dharmapatni *et al.*, 2001; Leask *et al.*, 2002; Fujimoto *et al.*, 2006; Wipff *et al.*, 2007), endothelin-1 (Shi-Wen *et al.*, 2007), but also PDGF and fibroblast growth factor (Gu and Iyer, 2006). Both endothelin-1 (Shi-Wen *et al.*, 2007) and PDGF have been implicated in SSc pathogenesis (Baroni *et al.*, 2006). Therefore, multiple profibrotic signaling pathways converge on this set of genes, which almost certainly contributes to the signature variability in the TGF $\beta$ -activated group.

The 674 genes of the TGF $\beta$ -responsive signature defined here were compared to two previously published studies of TGF $\beta$ -responsive genes in human fibroblasts. Chambers *et al.* (2003) analyzed human fetal lung fibroblast and identified 122 TGF $\beta$ -responsive genes, of which 58 (48%) were identified as TGF $\beta$ -responsive in our study. Renzoni *et al.* (2004) analyzed adult lung fibroblasts and identified 128 TGF $\beta$ -responsive genes, of which 33 (26%) were identified in our study. Only 11 genes were identified by both Renzoni and Chambers, all of which are found to be TGF $\beta$ -responsive here. The limited overlap found among the different signatures is not surprising given the differences in fibroblasts from different anatomical sites (Chang *et al.*, 2002; Rinn *et al.*, 2006), the use of different microarray platforms, and different analysis techniques.

We have previously described subsets of scleroderma patients distinguished by unique gene expression signatures found in both lesional (forearm) and non-lesional (lower back) skin (Milano *et al.*, 2008). The association of the TGF $\beta$ -responsive signature with both the lesional and non-lesional skin is an important validation of the systemic gene expression in skin.

## MATERIALS AND METHODS

The gene expression data from lesional forearm and non-lesional back skin of scleroderma, and of healthy skin biopsies, has been previously described (Milano *et al.*, 2008). All work with human subjects or materials was performed with adherence to the Declaration of Helsinki Principles. All subjects signed consent forms approved by the Committee on Human Research at the University of California, San Francisco. Patients met the ACR criteria for systemic sclerosis and were further defined as the diffuse subset. The Committee on the Protection of Human Subjects approved all protocols at Dartmouth Medical School.

All clinical measurements were taken at time of biopsy and have been described in detail previously (Milano *et al.*, 2008). Briefly, as part of the patients' routine scleroderma standard of care, patients were assessed for MRSS on a 51-point scale, disease duration since first onset of non-Raynaud's symptoms, a self-reported Raynaud severity score on a 10-point scale, and presence or absence of digital ulcers on a three-point scale. Also recorded were presence (+) or absence (-) of gastrointestinal involvement, scleroderma lung disease, which included patients defined as having pulmonary fibrosis or patients with ILD as determined by ground glass appearance, fibrosis or honeycombing on high-resolution computerized tomography, renal disease, and anti-topoisomerase antibodies. Data from patients on immunosuppressive therapy (mycophenolate mofetil, methotrexate, azathioprine, or prednisone >10 mg daily) were also recorded.

### Cells and cell culture

Primary adult human dermal fibroblasts were purchased from Cambrex Bioscience Inc. (East Rutherford, NJ). Additional adult dermal fibroblasts were isolated from explanted healthy or dSSc lesional forearm skin biopsies (Table 1) cultured for at least three passages in DMEM, 10% (v/v) fetal bovine serum, penicillin-streptomycin (100 IU/ml) at 37 °C in a humidified atmosphere with 5% CO<sub>2</sub>. Cells were passaged every 7 days for 7–10 passages prior to treatment.

### BrdU staining

Cell proliferation was assessed with BrdU Labeling and Detection kit I (Roche Applied Sciences, Penzberg, Germany). BrdU incorporation was detected according to the manufacturer's instructions, counter-stained with 4'-6-diamidino-2-phenylindole, and visualized with an Olympus BX51 microscope.

### RNA preparation

For analysis of the TGF $\beta$  response,  $4 \times 10^5$  cells were grown in 100-mm dishes and cultured in DMEM with 10% fetal bovine serum for 48 hours. Cells were brought to quiescence by culturing in low-serum media, DMEM with 0.1% fetal bovine serum, for 24 hours. TGF $\beta$  derived from human platelets (R&D Systems, Minneapolis, MN) was diluted into low-serum media and used to treat cells for 0, 2, 4, 8, 12, and 24 hours. Following treatment with TGF $\beta$ , cells were lysed in RLT buffer with  $\beta$ -mercaptoethanol and total RNA was isolated with RNeasy minikits (Qiagen, Valencia, CA).



## Microarray procedures

Microarray protocols have been described in detail (Milano *et al.*, 2008). Total RNA (300–500 ng) was amplified and labeled according to Agilent Low RNA Input Fluorescent Linear Amplification protocols. Each experimental RNA sample was labeled with Cy3-CTP and competitively hybridized against Universal Human Reference RNA (Stratagene, La Jolla, CA) labeled with Cy5-CTP on Agilent 44,000-element Human microarrays.

Microarrays were scanned with a GenePix 4000B scanner and acquired images quantified with the GenePix Pro 5.1 software (Axon Instruments, Foster City, CA). Technical artifacts and poor-quality spots were flagged and excluded from further analysis. Data were loaded to the UNC Microarray Database.

## Data analysis

Data were analyzed as  $\log_2$  of the lowess-normalized Cy5/Cy3 ratio. A total of 18,696 probes with fluorescent signal at least 1.5 greater than local background and of good quality in at least 80% of arrays were selected for analysis. Each data table was multiplied by negative one to convert the  $\log_2$  ratios to Cy3/Cy5 ratios and then T0-transformed using the average of triplicate 0-hour samples. TGF $\beta$ -responsive probes were selected by fold-change cutoff and genes were hierarchically clustered using the Cluster 3.0 software (Eisen *et al.*, 1998). Centroid values for each gene were calculated by averaging the T0-transformed data for the 12- and 24-hour time points across all TGF $\beta$  treatment time courses. The optimal fold-change threshold was determined by comparing genes induced or repressed by the addition of TGF $\beta$  over a range of threshold values to a list of 15 known TGF $\beta$  targets in human fibroblasts compiled from the literature (Supplementary Table S2). At a 1.74 fold-change threshold we identified seven of the 15 (47%) known TGF $\beta$  targets in the microarray time-course data. Analyzing this same list of genes in two published TGF $\beta$  signatures from pulmonary fibroblasts identified three (20%) (Renzoni *et al.*, 2004) and six (40%) (Chambers *et al.*, 2003) of the fibroblasts targets. A less stringent fold-change threshold did not significantly increase the fraction of known targets identified. Therefore, we elected to use the 1.74 fold-change threshold to define TGF $\beta$ -responsive probes.

Module maps were created with Genomica (Segal *et al.*, 2003). Expression data for probes that mapped to the same gene were averaged. Probes lacking entrez gene identifiers were excluded, yielding 10,664 unique genes for analysis. Gene sets of GO biological processes for the module map analysis were obtained from <http://genomica.weizmann.ac.il>.

Microarray data analyzing scleroderma skin are available from GEO (accession no. GSE9285) (Milano *et al.*, 2008). The patient data set consisting of biopsies from 17 dSSc patients and six healthy controls (53 arrays in total) was downloaded from UNC Microarray Database. A total of 40,818 probes with fluorescent signal at least 1.2-fold greater than local background and of good quality in at least 80% of arrays were selected for analysis.

Statistical analysis of clustering of the patient samples was performed with SigClust (Liu *et al.*, 2008). Heat maps were generated and visualized using TreeView version 1.0.13 (Eisen *et al.*, 1998).

## Quantitative real-time PCR

Total RNA (100–200 ng) was reverse-transcribed into single-stranded complementary DNA using SuperScript II reverse transcriptase (Invitrogen, San Diego, CA). Complementary DNA (1.25  $\mu$ g) was used for each qRT-PCR reaction. Primer probes sets for PAI1 (NM\_00602.2), 18S (X03205.1), E2F7 (NM\_203394.2), growth differentiation factor-6 (NM\_001001557.1), actin- $\alpha$ 2 (NM\_001141945.1), and glyceraldehyde-3-phosphate

dehydrogenase (NM\_002046.3) were obtained from Applied Biosystems (Foster City, CA) and analyzed using either the 7300 Real-Time PCR System or the 7,500 Fast Real-Time PCR system. The number of cycles required to generate a detectable fluorescence above background ( $C_t$ ) was measured for each sample. For dose–response and time-course analyses, fold changes were calculated relative to the average of triplicate untreated samples by the comparative  $C_t$  formula  $2^{-\Delta\Delta C_t}$  (Livak and Schmittgen, 2001), where  $\Delta C_t$  is the difference between the target gene (PAI1) and the 18S rRNA control, and  $\Delta\Delta C_t$  is the difference between the  $\Delta C_t$  value of the target gene and the average of the  $\Delta C_t$  values of untreated triplicate samples.

### Statistical methods

Statistical calculations were performed using Prism for Macintosh V4.0a (GraphPad Software Inc., La Jolla, CA) or Microsoft Excel with the add-in Winstat. Where indicated, SAM was used to identify significantly differentially expressed genes (Tusher *et al.*, 2001).

### Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

### Acknowledgments

This work was supported by research grants from the Scleroderma Research Foundation to MKC, HYC, and MLW. Additional support for MLW came from HHMI Biomedical Research Support award no. 76200-560801 to Dartmouth College. HYC received additional support from the Congressionally Directed Medical Research Program. AM also received support from the NIH Autoimmunity and Connective Tissue Biology Training grant (AR007576) from the National Institute of Arthritis, Musculoskeletal and Skin diseases (NIAMS).

### References

- Akhmet-shina A, Dees C, Pilecky M, Maurer B, Axmann R, Jungel A, et al. Dual inhibition of c-abl and PDGF receptor signaling by dasatinib and nilotinib for the treatment of dermal fibrosis. *FASEB J*. 2008; 22:2214–22. [PubMed: 18326784]
- Barnett AJ, Miller M, Littlejohn GO. The diagnosis and classification of scleroderma (systemic sclerosis). *Postgrad Med J*. 1988a; 64:121–5. [PubMed: 3050937]
- Barnett AJ, Miller MH, Littlejohn GO. A survival study of patients with scleroderma diagnosed over 30 years (1953–1983): the value of a simple cutaneous classification in the early stages of the disease. *J Rheumatol*. 1988b; 15:276–83. [PubMed: 3361537]
- Baroni SS, Santillo M, Bevilacqua F, Luchetti M, Spadoni T, Mancini M, et al. Stimulatory autoantibodies to the PDGF receptor in systemic sclerosis. *N Engl J Med*. 2006; 354:2667–76. [PubMed: 16790699]
- Chambers RC, Leoni P, Kaminski N, Laurent GJ, Heller RA. Global expression profiling of fibroblast responses to transforming growth factor-beta1 reveals the induction of inhibitor of differentiation-1 and provides evidence of smooth muscle cell phenotypic switching. *Am J Pathol*. 2003; 162:533–46. [PubMed: 12547711]
- Chang HY, Chi JT, Dudoit S, Bondre C, van de RM, Botstein D, et al. Diversity, topographic differentiation, and positional memory in human fibroblasts. *Proc Natl Acad Sci USA*. 2002; 99:12877–82. [PubMed: 12297622]
- Chung L, Fiorentino DF, Benbarak MJ, Adler AS, Mariano MM, Paniagua RT, et al. Molecular framework for response to imatinib mesylate in systemic sclerosis. *Arthritis Rheum*. 2009; 60:584–91. [PubMed: 19180499]
- Classen JF, Henrohn D, Rorsman F, Lennartsson J, Lauwerys BR, Wikstrom G, et al. Lack of evidence of stimulatory autoantibodies to platelet-derived growth factor receptor in patients with systemic sclerosis. *Arthritis Rheum*. 2009; 60:1137–44. [PubMed: 19333949]

- Cotton SA, Herrick AL, Jayson MI, Freemont AJ. TGF beta – a role in systemic sclerosis? *J Pathol.* 1998; 184:4–6. [PubMed: 9582519]
- Daniels CE, Wilkes MC, Edens M, Kottom TJ, Murphy SJ, Limper AH, et al. Imatinib mesylate inhibits the profibrogenic activity of TGF-beta and prevents bleomycin-mediated lung fibrosis. *J Clin Invest.* 2004; 114:1308–16. [PubMed: 15520863]
- Denton CP, Merkel PA, Furst DE, Khanna D, Emery P, Hsu VM, et al. Recombinant human anti-transforming growth factor beta1 antibody therapy in systemic sclerosis: a multicenter, randomized, placebo-controlled phase I/II trial of CAT-192. *Arthritis Rheum.* 2007; 56:323–33. [PubMed: 17195236]
- Dharmapatni AA, Smith MD, Ahern MJ, Simpson A, Li C, Kumar S, et al. The TGF beta receptor endoglin in systemic sclerosis. *Asian Pac J Allergy Immunol.* 2001; 19:275–82. [PubMed: 12009077]
- Distler JH, Jungel A, Huber LC, Schulze-Horsel U, Zwerina J, Gay RE, et al. Imatinib mesylate reduces production of extracellular matrix and prevents development of experimental dermal fibrosis. *Arthritis Rheum.* 2007; 56:311–22. [PubMed: 17195235]
- Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA.* 1998; 95:14863–8. [PubMed: 9843981]
- Farina G, Lemaire R, Korn JH, Widom RL. Cartilage oligomeric matrix protein is overexpressed by scleroderma dermal fibroblasts. *Matrix Biol.* 2006; 25:213–22. [PubMed: 16520029]
- Ferri C, Valentini G, Cozzi F, Sebastiani M, Michelassi C, La Montagna G, et al. Systemic sclerosis: demographic, clinical, and serologic features and survival in 1,012 Italian patients. *Medicine (Baltimore).* 2002; 81:139–53. [PubMed: 11889413]
- Fujimoto M, Hasegawa M, Hamaguchi Y, Komura K, Matsushita T, Yanaba K, et al. A clue for telangiectasis in systemic sclerosis: elevated serum soluble endoglin levels in patients with the limited cutaneous form of the disease. *Dermatology.* 2006; 213:88–92. [PubMed: 16902284]
- Gabrielli A, Di Loreto C, Taborro R, Candela M, Sambo P, Nitti C, et al. Immunohistochemical localization of intracellular and extracellular associated TGF beta in the skin of patients with systemic sclerosis (scleroderma) and primary Raynaud's phenomenon. *Clin Immunol Immunopathol.* 1993; 68:340–9. [PubMed: 8370185]
- 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]
- Goh NS, Desai SR, Veeraraghavan S, Hansell DM, Copley SJ, Maher TM, et al. Interstitial lung disease in systemic sclerosis: a simple staging system. *Am J Respir Crit Care Med.* 2008; 177:1248–54. [PubMed: 18369202]
- Gu J, Iyer VR. PI3K signaling and miRNA expression during the response of quiescent human fibroblasts to distinct proliferative stimuli. *Genome Biol.* 2006; 7:R42. [PubMed: 16737555]
- Ihn H. Autocrine TGF-beta signaling in the pathogenesis of systemic sclerosis. *J Dermatol Sci.* 2008; 49:103–13. [PubMed: 17628443]
- Jinnin M, Ihn H, Mimura Y, Asano Y, Tamaki K. Potential regulatory elements of the constitutive upregulated alpha2(I) collagen gene in scleroderma dermal fibroblasts. *Biochem Biophys Res Commun.* 2006; 343:904–9. [PubMed: 16564026]
- Kubo M, Ihn H, Yamane K, Tamaki K. Upregulated expression of transforming growth factor-beta receptors in dermal fibroblasts of skin sections from patients with systemic sclerosis. *J Rheumatol.* 2002; 29:2558–64. [PubMed: 12465152]
- Kuroda K, Shinkai H. Gene expression of types I and III collagen, decorin, matrix metalloproteinases and tissue inhibitors of metalloproteinases in skin fibroblasts from patients with systemic sclerosis. *Arch Dermatol Res.* 1997; 289:567–72. [PubMed: 9373715]
- Leask A. Transcriptional profiling of the scleroderma fibroblast reveals a potential role for connective tissue growth factor (CTGF) in pathological fibrosis. *Keio J Med.* 2004; 53:74–7. [PubMed: 15247510]
- Leask A. Scar wars: is TGFbeta the phantom menace in scleroderma? *Arthritis Res Ther.* 2006; 8:213. [PubMed: 16774692]

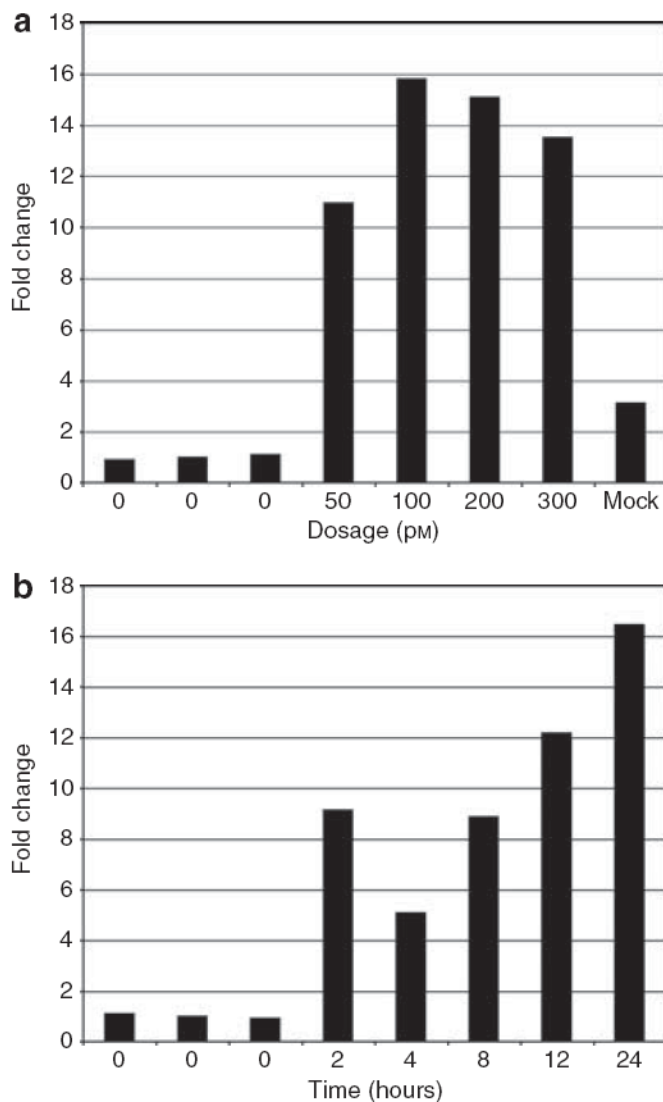
- Leask A, Abraham DJ. TGF-beta signaling and the fibrotic response. *FASEB J*. 2004; 18:816–27. [PubMed: 15117886]
- Leask A, Abraham DJ, Finlay DR, Holmes A, Pennington D, Shi-Wen X, et al. Dysregulation of transforming growth factor beta signaling in scleroderma: overexpression of endoglin in cutaneous scleroderma fibroblasts. *Arthritis Rheum*. 2002; 46:1857–65. [PubMed: 12124870]
- LeRoy EC, Black C, Fleischmajer R, Jablonska S, Krieg T, Medsger TA Jr, et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. *J Rheumatol*. 1988; 15:202–5. [PubMed: 3361530]
- Leroy EC, Smith EA, Kahaleh MB, Trojanowska M, Silver RM. A strategy for determining the pathogenesis of systemic sclerosis. Is transforming growth factor beta the answer? *Arthritis Rheum*. 1989; 32:817–25. [PubMed: 2665755]
- Liu Y, Hayes DN, Nobel A, Marron JS. Statistical significance of clustering for high-dimension, low-sample size data. *J Am Stat Assoc*. 2008; 103:1281–93.
- 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]
- Maricq HR, Valter I. A working classification of scleroderma spectrum disorders: a proposal and the results of testing on a sample of patients. *Clin Exp Rheumatol*. 2004; 22:S5–13. [PubMed: 15344590]
- 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]
- Mummary CL. Transforming growth factor beta and mouse development. *Microsc Res Tech*. 2001; 52:374–86. [PubMed: 11170296]
- Ostojic P, Cerinic MM, Silver R, Highland K, Damjanov N. Interstitial lung disease in systemic sclerosis. *Lung*. 2007; 185:211–20. [PubMed: 17717851]
- Padgett RW, Savage C, Das P. Genetic and biochemical analysis of TGF beta signal transduction. *Cytokine Growth Factor Rev*. 1997; 8:1–9. [PubMed: 9174659]
- Renzoni EA, Abraham DJ, Howat S, Shi-Wen X, Sestini P, Bou-Gharios G, et al. Gene expression profiling reveals novel TGFbeta targets in adult lung fibroblasts. *Respir Res*. 2004; 5:24. [PubMed: 15571627]
- Rinn JL, Bondre C, Gladstone HB, Brown PO, Chang HY. Anatomic demarcation by positional variation in fibroblast gene expression programs. *PLoS Genet*. 2006; 2:e119. [PubMed: 16895450]
- Scussel-Lonzetti L, Joyal F, Raynauld JP, Roussin A, Rich E, Goulet JR, et al. Predicting mortality in systemic sclerosis: analysis of a cohort of 309 French Canadian patients with emphasis on features at diagnosis as predictive factors for survival. *Medicine (Baltimore)*. 2002; 81:154–67. [PubMed: 11889414]
- Segal E, Shapira M, Regev A, Pe'er D, Botstein D, Koller D, et al. Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. *Nat Genet*. 2003; 34:166–76. [PubMed: 12740579]
- Sfikakis PP, McCune BK, Tsokos M, Aroni K, Vayiopoulos G, Tsokos GC. Immunohistological demonstration of transforming growth factor-beta isoforms in the skin of patients with systemic sclerosis. *Clin Immunol Immunopathol*. 1993; 69:199–204. [PubMed: 8403557]
- Shi-Wen X, Renzoni EA, Kennedy L, Howat S, Chen Y, Pearson JD, et al. Endogenous endothelin-1 signaling contributes to type I collagen and CCN2 overexpression in fibrotic fibroblasts. *Matrix Biol*. 2007; 26:625–32. [PubMed: 17681742]
- Smith EA, LeRoy EC. A possible role for transforming growth factor-beta in systemic sclerosis. *J Invest Dermatol*. 1990; 95:125S–7S. [PubMed: 2258629]
- Steen VD, Medsger TA. Changes in causes of death in systemic sclerosis, 1972–2002. *Ann Rheum Dis*. 2007; 66:940–4. [PubMed: 17329309]
- Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA*. 2001; 98:5116–21. [PubMed: 11309499]
- Varga J. Antifibrotic therapy in scleroderma: extracellular or intracellular targeting of activated fibroblasts? *Curr Rheumatol Rep*. 2004; 6:164–70. [PubMed: 15016348]

- Varga J, Whitfield ML. Transforming growth factor-beta in systemic sclerosis (scleroderma). *Front Biosci.* 2009; 1:226–35.
- Verrecchia F, Mauviel A, Farge D. Transforming growth factor-beta signaling through the Smad proteins: role in systemic sclerosis. *Autoimmun Rev.* 2006; 5:563–9. [PubMed: 17027893]
- Wall NA, Hogan BL. TGF-beta related genes in development. *Curr Opin Genet Dev.* 1994; 4:517–22. [PubMed: 7950318]
- 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 USA.* 2003; 100:12319–24. [PubMed: 14530402]
- Whitfield ML, George LK, Grant GD, Perou CM. Common markers of proliferation. *Nat Rev Cancer.* 2006; 6:99–106. [PubMed: 16491069]
- Whitfield ML, Sherlock G, Saldanha AJ, Murray JI, Ball CA, Alexander KE, et al. Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol Biol Cell.* 2002; 13:1977–2000. [PubMed: 12058064]
- Wipff J, Kahan A, Hachulla E, Sibia J, Cabane J, Meyer O, et al. Association between an endoglin gene polymorphism and systemic sclerosis-related pulmonary arterial hypertension. *Rheumatology (Oxford).* 2007; 46:622–5. [PubMed: 17166870]
- Wynn TA. Cellular and molecular mechanisms of fibrosis. *J Pathol.* 2008; 214:199–210. [PubMed: 18161745]
- Xu WD, Leroy EC, Smith EA. Fibronectin release by systemic sclerosis and normal dermal fibroblasts in response to TGF-beta. *J Rheumatol.* 1991; 18:241–6. [PubMed: 2023218]

## Abbreviations

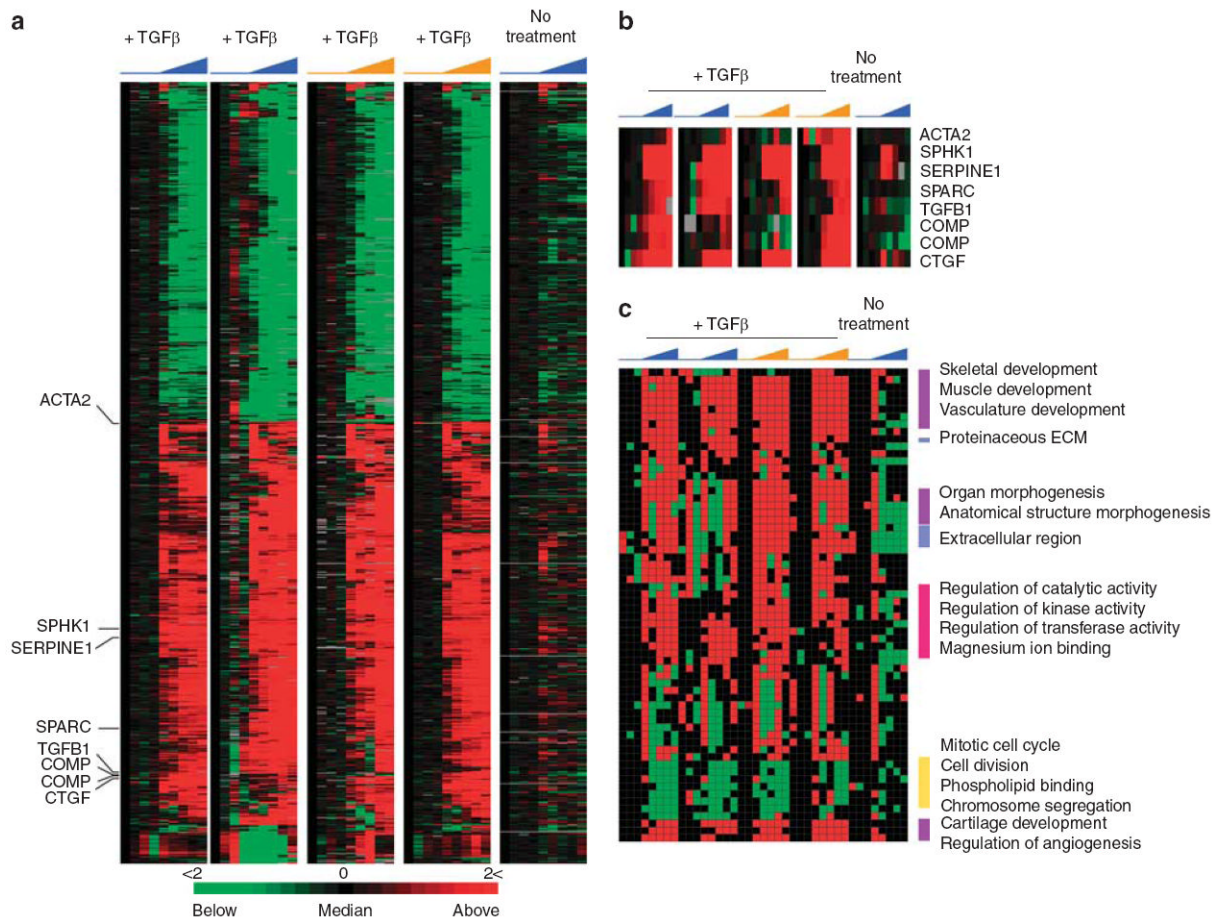
<b>dSSc</b>	diffuse systemic sclerosis
<b>COMP</b>	cartilage oligomeriz matrix protein
<b>GO</b>	Gene Ontology
<b>ILD</b>	interstitial lung disease
<b>ISSc</b>	limited systemic sclerosis
<b>MRSS</b>	modified Rodnan skin score
<b>PDGF</b>	platelet-derived growth factor
<b>qRT-PCR</b>	quantitative real-time PCR
<b>SAM</b>	significance analysis of microarrays
<b>SSc</b>	systemic sclerosis
<b>TGFβ</b>	transforming growth factor-β





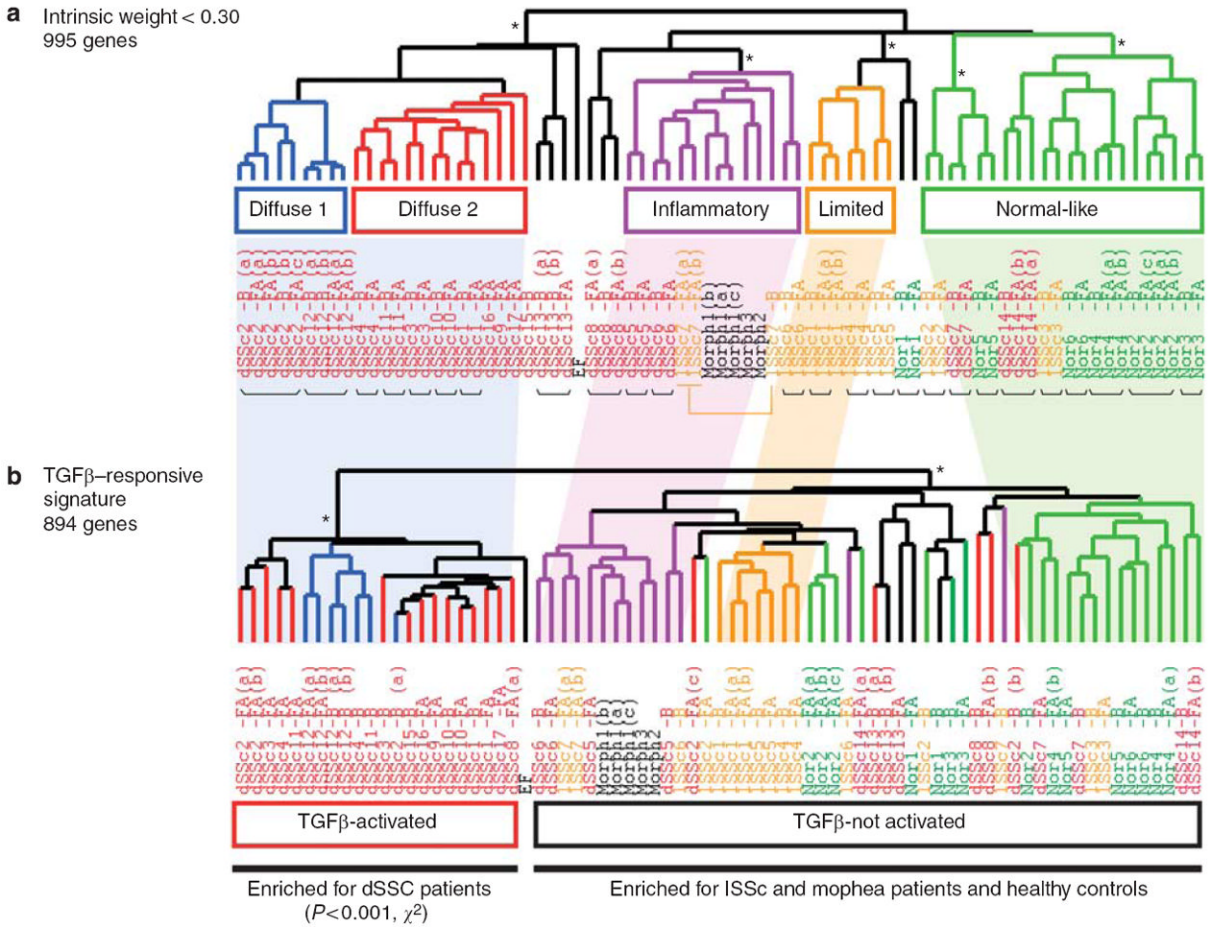
**Figure 1. Dosage and time dependence of plasminogen activator inhibitor 1 (PAI1) mRNA expression after TGF $\beta$  treatment**

To optimize the concentration of TGF $\beta$  and the length of time course to analyze, cells were treated with varying doses of TGF $\beta$ , and time points were collected from 2 to 24 hours. (a) Normal dermal fibroblasts (NDFs) in 0.1% serum were treated with 50, 100, 200, or 300 pM TGF $\beta$  for 2 hours. Levels of PAI1 mRNA were measured in triplicate by Taqman qRT-PCR, normalized to 18S rRNA; fold change is relative to the average of three independent, untreated samples. (b) NDFs were treated with 50 pM TGF $\beta$  and PAI1 mRNA levels were measured in triplicate at 0, 2, 4, 8, 12, and 24 hours post-treatment. Fold change is relative to the average of the zero time points.



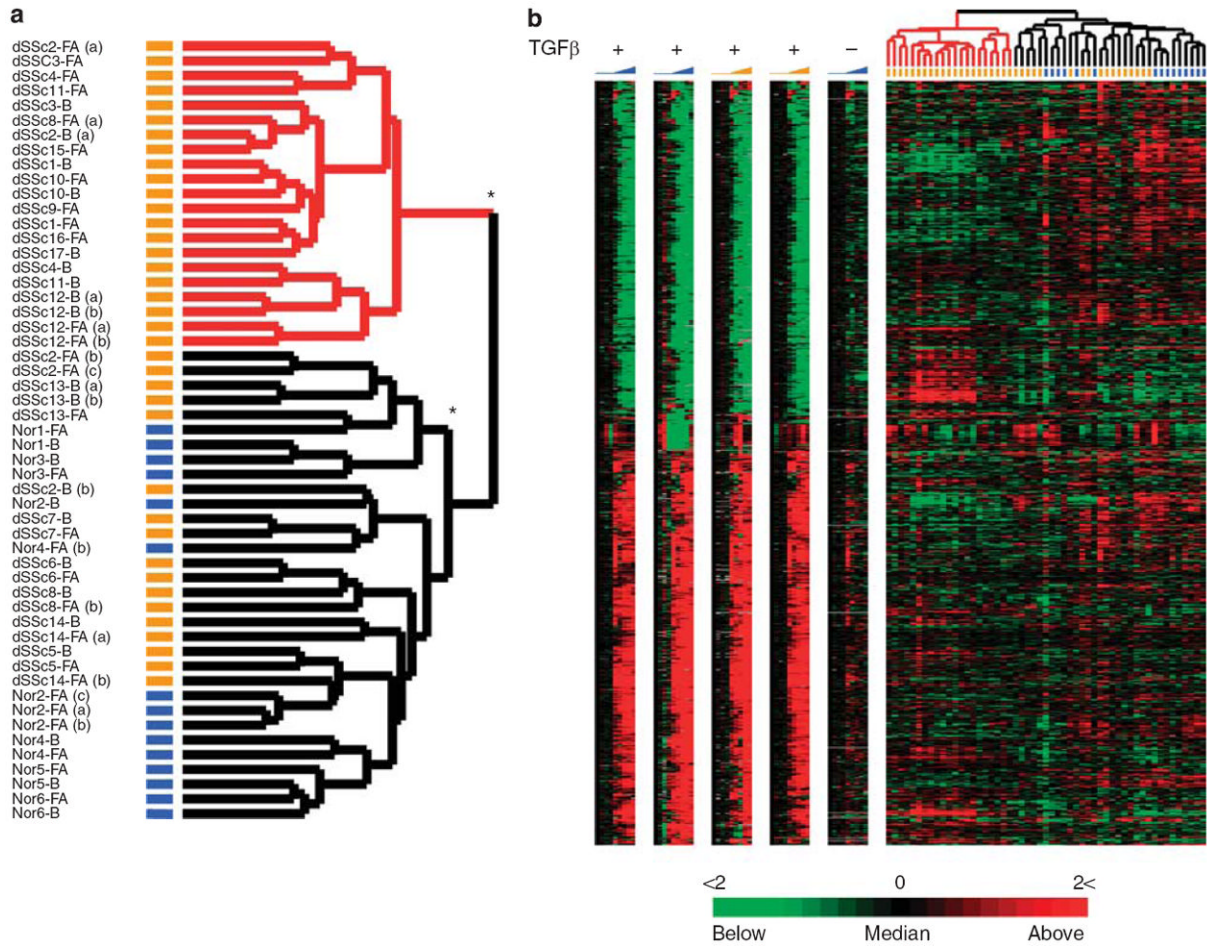
**Figure 2. Genome-wide response to TGFβ in adult dermal fibroblasts**

(a) Shown are the 894 probes, representing 674 annotated genes, with a 1.74-fold or greater change in expression over 24 hours following treatment with 50 pM TGFβ. Four independent primary cell cultures were treated with TGFβ, two from healthy control subjects (blue), and two from SSc patients (orange). A mock time course was performed using identical conditions with the omission of TGFβ. Time of treatment, from 0 to 24 hours, is indicated. Each row represents a probe, and each column represents a time point. (b) Expression data for the genes previously reported as being TGFβ responsive (Supplementary Table S2) and found among the 894 probes. (c) Module map of GO terms in the genome-wide response to TGFβ. Each column represents a microarray, and each row represents an enriched GO term. Only modules that were significantly enriched ( $P < 0.05$ ) in at least 16 microarrays analyzed are shown. Select modules are indicated.



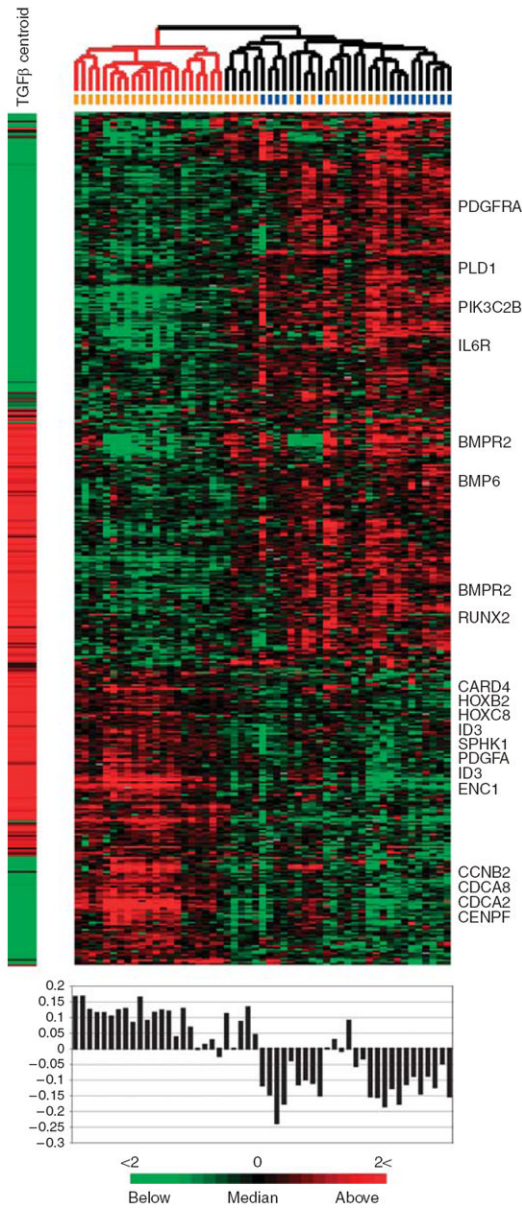
**Figure 3. TGFβ-responsive genes are deregulated in the “diffuse-proliferation” subset of scleroderma**  
**(a)** Intrinsic subsets of scleroderma as described by Milano *et al.* (2008). **(b)** Organization of 75 arrays from Milano *et al.* by hierarchical clustering of 894 TGFβ-responsive signature genes. Shading is indicative of positioning of subsets identified by intrinsic genes relative to that identified by the TGFβ-responsive signature. \* $P < 0.001$ .





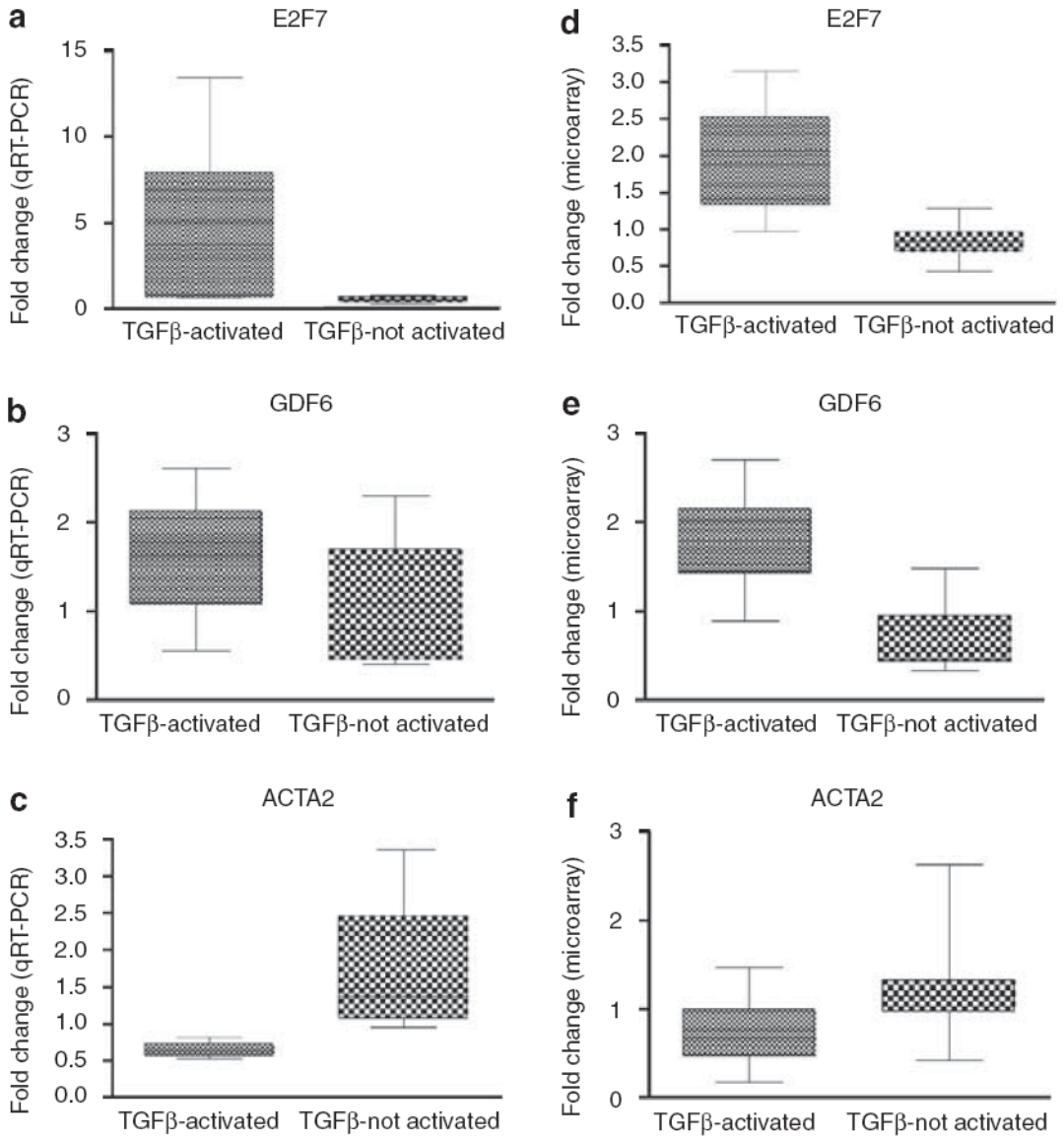
**Figure 4. The TGFβ-responsive signature distinguishes a subset of dSSc patients**

(a) Patient sample dendrogram resulting from hierarchical clustering of 53 arrays probing gene expression in skin biopsies of patients with dSSc (orange bars) and healthy control (blue bars). The samples were clustered using the 894 TGFβ-responsive probes that comprise the signature. Two major groups of samples are evident: TGFβ-activated (red) and TGFβ-not-activated (black). Technical replicates are designated by a letter (a, b or c) following patient and biopsy site identification. Statistically significant clusters as determined by SigClust are marked by \* ( $P < 0.001$ ). (b) Individual TGFβ time courses are aligned with the gene expression data from dSSc and healthy control biopsies, and illustrate the heterogeneity of the *in vitro*-derived TGFβ-responsive signature in skin biopsies.

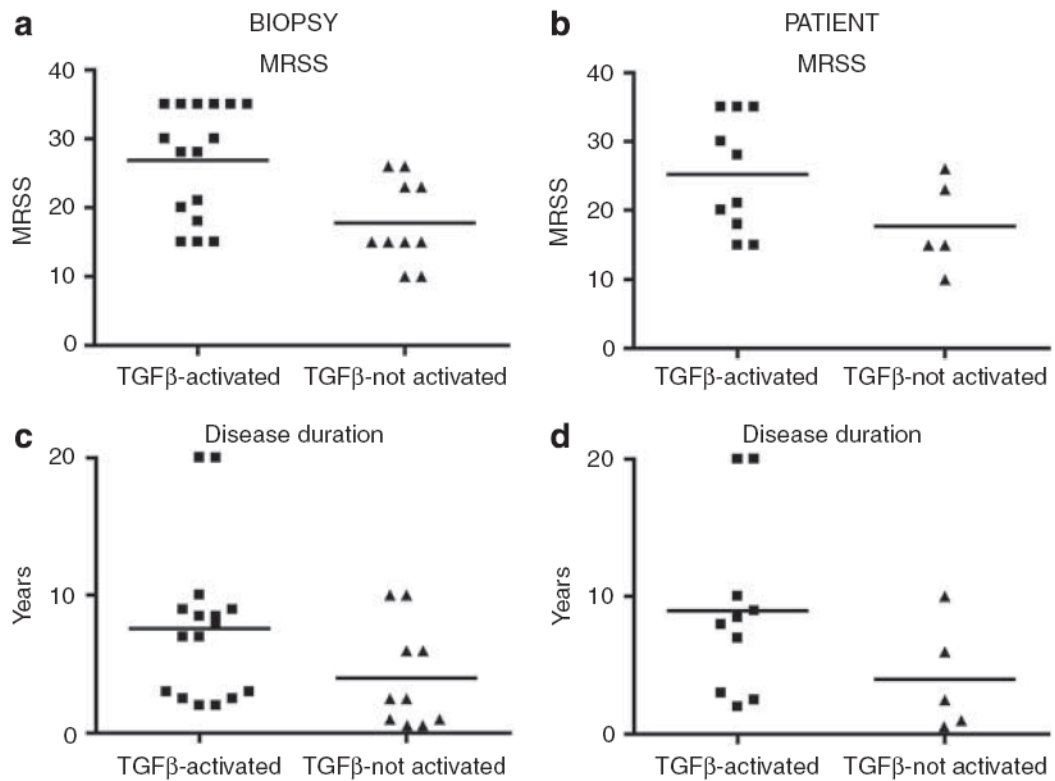


**Figure 5. A subset of the TGFβ-responsive signature ideally differentiates the two groups** SAM was used to identify 474 probes from the initial set of 894 probes that showed consistent, statistically significant differential expression between the TGFβ-activated and TGFβ-not-activated groups. The centroid representing the average of the TGFβ-responsive gene signature at maximal induction (12 and 24 hours) is shown to the left of the heat map. Pearson correlations between the centroid and each array were calculated and are plotted directly beneath each array.





**Figure 6. Validation of TGFβ-responsive genes in dSSc skin**  
Relative mRNA levels of three genes, E2F7 (**a, d**), growth differentiation factor-6 (**b, e**), and actin-α2 (**c, f**) were determined using Fast Taqman qRT-PCR on select patient samples from the TGFβ-activated and TGFβ-not-activated patient groups. Trends of mRNA levels determined by qRT-PCR were reflective of those measured on DNA microarrays (**d-f**). All data were normalized to the mean relative expression ratio.



**Figure 7. Distributions of MRSS and disease duration of dSSc patients**

Distributions of MRSS at the time of biopsy were plotted by biopsy (**a**) and by patient (**b**) for the TGF $\beta$ -activated and TGF $\beta$ -not-activated groups. Distributions of disease duration were also plotted by biopsy (**c**) and by patient (**d**) for the two groups. *P*-values for all comparisons are given in Table 2.

**Table 1**

## Primary fibroblast cell cultures

Cell line	Sex	Age	Disease duration	Biopsy site
Control 1	Male	47	—	ND <sup>1</sup>
Control 2	Female	58	—	Forearm
Scleroderma 1	Female	62	1 year	Forearm
Scleroderma 2	Male	41	1 year	Forearm

Clinical characteristics of individuals from whose biopsies dermal fibroblasts were obtained.

<sup>1</sup> Adult dermal fibroblast primary cell line obtained from a commercial source. Site not determined (ND).

Table 2

Clinical parameters associated with the TGFβ-responsive signature

Clinical parameter	Activated (16 biopsies)	Not activated (10 biopsies)	P-value	Activated (10 patients)	Not activated (5 patients)	P-value
Patient age (years)	46.3±2.55	50.6±2.33	0.26	48.0±3.43	50.6±3.50	0.64
Disease duration (years)	7.63±1.41	4.00±1.19	0.086	9.00±2.04	4.40±4.07	0.14
MRSS	26.9±2.04	17.8±1.95	<b>0.0061<sup>1</sup></b>	25.2±2.63	17.8±2.92	0.11
Lung disease	7/16	0/10	<b>0.0142</b>	5/10	0/5	<b>0.053</b>
GI involvement	14/16	6/10	0.11	9/10	3/5	0.17
Renal disease	2/16	0/10	0.24	2/10	0/5	0.28
Raynaud severity	5.93±0.57	7.00±0.99	0.33	5.33±0.78	7.00±1.48	0.29
Digital ulcers	0.50±0.18	0.80±0.38	0.44	0.50±0.22	0.80±0.58	0.57
Immunosuppressive therapy	9/16	2/10	0.069	5/10	1/5	0.26
Anti-topoisomerase-I <sup>3</sup>	6/12	2/6	0.502	3/6	1/3	0.64

MRSS, modified Rodnan skin score; TGFβ, transforming growth factor-β.

Statistical associations of clinical parameters with the TGFβ-activated and TGFβ-not-activated groups of patients. Associations with MRSS, disease duration, patient age, Raynaud severity, and digital ulcers were calculated using Student's *t*-tests (mean ± SEM). Associations with ILD, GI involvement, renal disease, and presence of anti-topoisomerase antibodies were determined by  $\chi^2$ -test

<sup>1</sup> ( $P<0.01$ ),<sup>2</sup> ( $P<0.02$ ),

Calculations were performed by biopsy (columns 2–4) and by patient (columns 6–7).

<sup>3</sup> Data on the anti-topoisomerase-I status were only available for six patients (12 biopsies) in the activated group and three patients (six biopsies) patients in the not activated group.