A TGFβ-Responsive Gene Signature Is Associated with a Subset of Diffuse Scleroderma with Increased Disease Severity

Jennifer L. Sargent1, Ausra Milano1, Swati Bhattacharyya2, John Varga2, M. Kari Connolly3, Howard Y. Chang4 and Michael L. Whitfield1

Systemic sclerosis is a complex disease with widespread skin fibrosis and variable visceral organ involvement. Since transforming growth factor-β (TGFβ) has been implicated in driving fibrosis in systemic sclerosis, a mechanism-derived gene expression signature was used to assay TGFβ-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 TGFβ, 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 TGFβ responsive. Expression of the TGFβ-responsive signature was examined in skin biopsies from 17 dSSc, seven limited SSc (lSSc), three morphea patients, and six healthy controls. The TGFβ-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 TGFβ-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 TGFβ-responsive signature is found in only a subset of dSSc patients who could be targeted by specific therapies.

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; Mariq 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-β (TGFβ) in vitro are associated with the “diffuse-proliferation” subset of dSSc patients.

Skin fibrosis, generally considered to be driven by TGFβ-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 TGFβ have been observed in SSC skin biopsies (Gabrielli et al., 1993; Sfikakis et al., 1993) and TGFβ-activated fibroblasts produce collagen I, III, and V in addition to other matrix proteins such as glycoaminoglycans (Wynn, 2008). Numerous studies have identified TGFβ-regulated genes that show increased expression in SSC fibroblasts, including fibronectin (Xu et al., 1991), collagens α1(1) and α1(III) (Kuroda and Shinkai, 1997), collagen-α2(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 TGFβRI and TGFβRII mRNA levels and surface protein expression when compared with healthy dermal fibroblasts (Kubo et al., 2002), suggesting activation of the TGFβ 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-TGFβ 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 TGFβ-response is found in the gene expression subsets of SSC. We defined a TGFβ-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 TGFβ-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 TGFβ-responsive signature is found in the “diffuse-proliferation” subset of dSSc skin biopsies. This indicates that TGFβ deregulation, or a related pathway, contributes to pathogenesis in this subset.

RESULTS

We identified the genes responsive to TGFβ 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 biopsies, 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 TGFβ 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 TGFβ 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 pmol TGFβ and the cellular response measured by quantitative real-time PCR (qRT-PCR). We found 50 pmol TGFβ resulted in an 11-fold induction of plasminogen activator inhibitor 1 (PAI1) mRNA, whereas treatment with 100 pmol TGFβ resulted in 16-fold induction (Figure 1). Concentrations of 200–300 pmol resulted in decreased relative PAI1 levels, indicating that a concentration higher than 100 pmol was saturating. We chose to use 50 pmol for all further experiments.

Quiescent cells were exposed to 50 pmol TGFβ and total RNA was collected at 0, 2, 4, 8, 12, and 24 hours after

### Table 1. Primary fibroblast cell cultures

<table>
<thead>
<tr>
<th>Cell line</th>
<th>Sex</th>
<th>Age</th>
<th>Disease duration</th>
<th>Biopsy site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control 1</td>
<td>Male</td>
<td>47</td>
<td>—</td>
<td>ND¹</td>
</tr>
<tr>
<td>Control 2</td>
<td>Female</td>
<td>58</td>
<td>—</td>
<td>Forearm</td>
</tr>
<tr>
<td>Scleroderma 1</td>
<td>Female</td>
<td>62</td>
<td>1 year</td>
<td>Forearm</td>
</tr>
<tr>
<td>Scleroderma 2</td>
<td>Male</td>
<td>41</td>
<td>1 year</td>
<td>Forearm</td>
</tr>
</tbody>
</table>

Clinical characteristics of individuals from whose biopsies dermal fibroblasts were obtained.  
¹Adult dermal fibroblast primary cell line obtained from a commercial source. Site not determined (ND).

![Figure 1. Dosage and time dependence of plasminogen activator inhibitor 1 (PAI1) mRNA expression after TGFβ treatment.](image-url)
treatment. A mock time course without TGFβ, 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 TGFβ. 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 TGFβ in adult dermal fibroblasts**

To characterize the genome-wide response to TGFβ 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 TGFβ-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 TGFβ targets in dermal fibroblasts is shown (Figure 2b). Examination of these probes in the mock, no TGFβ, control shows that the observed changes in gene expression are specific to a TGFβ response (Figure 2a).

To identify differential TGFβ responses between normal and scleroderma fibroblasts, the data from TGFβ-treated normal fibroblast was compared to that of TGFβ-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 TGFβ responses in cultured lung fibroblasts from SSc patients and normal controls (Chambers et al., 2003).

**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 TGFβresponsive (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.
To summarize the biological programs represented by genes in the TGFβ-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 TGFβ 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 TGFβ treatment included proteinaceous extracellular matrix, extracellular matrix, and extracellular space, associated with collagen production, extracellular matrix deposition, and remodeling (Figure 2c). TGFβ 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 TGFβ 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 TGFβ downregulates the proliferation signature (Whitfield et al., 2006) under low-serum conditions.

**TGFβ-responsive gene expression underlies the “diffuse-proliferation” intrinsic subset**

To determine whether the TGFβ-responsive signature was expressed in the intrinsic subsets of scleroderma, we analyzed the expression of the 894 TGFβ-responsive signature probes (674 genes) in the data set of Milano et al. (2008), which included dSSc, ISSc, 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

---

**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. $^*P<0.001$. 

---
compared to the sample groupings reported using an intrinsic gene set (Milano et al., 2008). The group expressing the TGFβ-responsive signature is similar to the diffuse-proliferation subset (Supplementary Table S1), suggesting that deregulated TGFβ 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), endothe-
lin-1, which is a downstream mediator of TGFβ 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 TGFβ 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 TGFβ-responsive signature include only dSSc skin biopsies, and this enrichment is statistically significant (P < 1 x 10^{-4}, χ²-test). One patient classified as diffuse-proliferation and one patient classified as inflammatory did not show consistent expression of the TGFβ-responsive signature (Supplementary Table S1). Therefore, there is not a precise overlap of the TGFβ-responsive group and the proliferative group identified by Milano et al. The group without the TGFβ-responsive signature contains a subset of the dSSc biopsies, all lSSc biopsies, morphea samples, and healthy controls. Therefore, the TGFβ-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.

**TGFβ-responsive gene expression in dSSc patients**

To determine whether the TGFβ-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 TGFβ-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 TGFβ-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 TGFβ-activated group to be a single statistically significant cluster. The TGFβ-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 TGFβ 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 TGFβ-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) TGFβ-activated because of high positive correlation coefficients associated with these samples, and the group on the right (black dendrogram) TGFβ-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 TGFβ treatment. Genes found concordantly upregulated in TGFβ-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 (PDGFRα and IL6R) and phospholipids binding (PLD1 and PIK3C2B).

Genes induced in the TGFβ-treated fibroblasts, but with low expression in the TGFβ-activated skin biopsies, were enriched for developmental processes and signaling (RUNX2, BMP2, and BMP6). Genes showing decreased expression in TGFβ-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 TGFβ signature genes in patient samples**

We validated the expression of three genes from the TGFβ-responsive signature, which were also the most highly differentially expressed between the groups with and without the TGFβ-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 TGFβ superfamily, and actin-α2. 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 TGFβ-activated versus TGFβ-not-activated groups.

The TGFβ-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 TGFβ-activated group showed phenotypic differences from those in the TGFβ-not-activated group. Since patients dSSc2 and dSSc8 could not be accurately assigned, they were excluded from the analysis, leaving 10 patients in the TGFβ-activated group and five in the TGFβ-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 χ²-test used to assess differences between the groups (Table 2). When individual biopsies were considered, the TGFβ-activated group showed significantly higher skin scores (mean = 26.9 ± 2.04) than the TGFβ-not-activated group (mean = 17.8 ± 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 TGFβ-activated samples (7 of 16) as compared with TGFβ-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 TGFβ-activated group had ILD, whereas none of the five patients in the TGFβ-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 TGFβ 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 TGFβ-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 TGFβ 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 TGFβ-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 TGFβ-responsive signature is not found in patients with lSSc, morphea, or in healthy controls. Given the established role for TGFβ 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 TGFβ 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 is caused by 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,
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.

Table 2. Clinical parameters associated with the TGFβ-responsive signature

<table>
<thead>
<tr>
<th>Clinical parameter</th>
<th>Activated (16 biopsies)</th>
<th>Not activated (10 biopsies)</th>
<th>P-value</th>
<th>Activated (10 patients)</th>
<th>Not activated (5 patients)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient age (years)</td>
<td>46.3 ± 2.55</td>
<td>50.6 ± 2.33</td>
<td>0.26</td>
<td>48.0 ± 3.43</td>
<td>50.6 ± 3.50</td>
<td>0.64</td>
</tr>
<tr>
<td>Disease duration (years)</td>
<td>7.63 ± 1.41</td>
<td>4.00 ± 1.19</td>
<td>0.086</td>
<td>9.00 ± 2.04</td>
<td>4.40 ± 4.07</td>
<td>0.14</td>
</tr>
<tr>
<td>MRSS</td>
<td>26.9 ± 2.04</td>
<td>17.8 ± 1.95</td>
<td>0.0061</td>
<td>25.2 ± 2.63</td>
<td>17.8 ± 2.92</td>
<td>0.11</td>
</tr>
<tr>
<td>Lung disease</td>
<td>7/16</td>
<td>0/10</td>
<td>0.014²</td>
<td>5/10</td>
<td>0/5</td>
<td>0.053</td>
</tr>
<tr>
<td>GI involvement</td>
<td>14/16</td>
<td>6/10</td>
<td>0.11</td>
<td>9/10</td>
<td>3/5</td>
<td>0.17</td>
</tr>
<tr>
<td>Renal disease</td>
<td>2/16</td>
<td>0/10</td>
<td>0.24</td>
<td>2/10</td>
<td>0/5</td>
<td>0.28</td>
</tr>
<tr>
<td>Raynaud severity</td>
<td>5.93 ± 0.57</td>
<td>7.00 ± 0.99</td>
<td>0.33</td>
<td>5.33 ± 0.78</td>
<td>7.00 ± 1.48</td>
<td>0.29</td>
</tr>
<tr>
<td>Digital ulcers</td>
<td>0.50 ± 0.18</td>
<td>0.80 ± 0.38</td>
<td>0.44</td>
<td>0.50 ± 0.22</td>
<td>0.80 ± 0.58</td>
<td>0.57</td>
</tr>
<tr>
<td>Immunosuppressive therapy</td>
<td>6/16</td>
<td>2/10</td>
<td>0.069</td>
<td>5/10</td>
<td>1/5</td>
<td>0.26</td>
</tr>
<tr>
<td>Anti-topoisomerase-1³</td>
<td>6/12</td>
<td>2/6</td>
<td>0.502</td>
<td>3/6</td>
<td>1/3</td>
<td>0.64</td>
</tr>
</tbody>
</table>

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 χ²-test (²P < 0.01, ³P < 0.02). Calculations were performed by biopsy (columns 2–4) and by patient (columns 6–7).

³Data 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.
only a weak association was found between the TGFβ-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 TGFβ-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 TGFβ-responsive genes are also responsive to other profibrotic cytokines and signaling molecules such as IL13, IL4, spingosine-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 TGFβ-activated group.

The 674 genes of the TGFβ-responsive signature defined here were compared to two previously published studies of TGFβ-responsive genes in human fibroblasts. Chambers et al. (2003) analyzed human fetal lung fibroblasts and identified 122 TGFβ-responsive genes, of which 58 (48%) were identified as TGFβ-responsive in our study. Renzoni et al. (2004) analyzed adult lung fibroblasts and identified 128 TGFβ-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 TGFβ-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 TGFβ-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

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 TGFβ-activated and TGFβ-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.
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 motil, 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, and penicillin-streptomycin (100 IU/ml) at 37°C in a humidified atmosphere with 5% CO2. 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, counterstained with 4’-6-diamidino-2-phenylindole, and visualized with an Olympus BX51 microscope.

RNA preparation
For analysis of the TGFβ response, 4 × 10⁵ 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. TGFβ 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 TGFβ, cells were lysed in RLT buffer with β-mercaptoethanol and total RNA was isolated with RNeasy mini kits (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 log₂ of the lowest-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 log₂ ratios to Cy3/Cy5 ratios and then T0-transformed using the average of triplicate 0-hour samples. TGFβ-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 TGFβ treatment time courses. The optimal fold-change threshold was determined by comparing genes induced or repressed by the addition of TGFβ over a range of threshold values to a list of 15 known TGFβ 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 TGFβ targets in the microarray time-course data. Analyzing this same list of genes in two published TGFβ 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 TGFβ-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 mg) was used for each qRT-PCR reaction. Primer probes sets for PAI1 (NM_001001557.1), actin-β (NM_001141945.1), differentiation factor-6 (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 (Ci) 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 Ci formula 2−ΔΔCt (Livak and Schmittgen, 2001), where ΔCt is the difference between the target gene (PAI1) and the 18S rRNA control, and ΔΔCt is the difference between the ΔCt value of the target gene and the average of the ΔCt 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).
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/TGFbeta/

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.

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).

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

REFERENCES


