



OPEN

Global gene expression analysis of systemic sclerosis myofibroblasts demonstrates a marked increase in the expression of multiple NBPF genes

Giuseppina Abignano^{1,2,5}, Heidi Hermes^{3,5}, Sonsoles Piera-Velazquez³, Sankar Addya⁴, Francesco Del Galdo^{1,2,5}✉ & Sergio A. Jimenez^{3,5}✉

Myofibroblasts are the key effector cells responsible for the exaggerated tissue fibrosis in Systemic Sclerosis (SSc). Despite their importance to SSc pathogenesis, the specific transcriptome of SSc myofibroblasts has not been described. The purpose of this study was to identify transcriptome differences between SSc myofibroblasts and non-myofibroblastic cells. Alpha smooth muscle actin (α -SMA) expressing myofibroblasts and α -SMA negative cells were isolated employing laser capture microdissection from dermal cell cultures from four patients with diffuse SSc of recent onset. Total mRNA was extracted from both cell populations, amplified and analyzed employing microarrays. Results for specific genes were validated by Western blots and by immunohistochemistry. Transcriptome analysis revealed 97 differentially expressed transcripts in SSc myofibroblasts compared with non-myofibroblasts. Annotation clustering of the SSc myofibroblast-specific transcripts failed to show a TGF- β signature. The most represented transcripts corresponded to several different genes from the Neuroblastoma Breakpoint Family (NBPF) of genes. NBPF genes are highly expanded in humans but are not present in murine or rat genomes. In vitro studies employing cultured SSc dermal fibroblasts and immunohistochemistry of affected SSc skin confirmed increased NBPF expression in SSc. These results indicate that SSc myofibroblasts represent a unique cell lineage expressing a specific transcriptome that includes very high levels of transcripts corresponding to numerous NBPF genes. Elevated expression of NBPF genes in SSc myofibroblasts suggests that NBPF gene products may play a role in SSc pathogenesis and may represent a novel therapeutic target.

Systemic Sclerosis (SSc) is a systemic autoimmune disease of unknown etiology characterized by progressive fibrosis of skin and multiple internal organs and a severe proliferative vasculopathy^{1–4}. The cells responsible for tissue fibrosis and fibroproliferative processes are myofibroblasts, a population of mesenchymal cells displaying unique biological functions including cell mobility and tissue contraction, and the high expression of α -smooth muscle actin (α -SMA) and numerous molecules associated with fibrotic processes^{5–8}. The accumulation of myofibroblasts in affected tissues and their expression of a persistent profibrotic phenotype are crucial in the development of the fibroproliferative process in SSc. There is also strong evidence to indicate that these cells regulate the extent and rate of progression of the fibrotic alterations in SSc, and, therefore, play an important role in determining the clinical course, response to therapy, prognosis, and overall mortality of the disease^{9–13}. Indeed, it has been shown that the number of myofibroblasts present in affected SSc skin correlates with the severity of SSc-associated tissue fibrosis¹².

¹Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Beckett Street, WTBB 6.14, Leeds LS97TF, UK. ²NIHR Leeds Biomedical Research Centre, Leeds Teaching Hospitals NHS Trust, Leeds, UK. ³Jefferson Institute of Molecular Medicine and Scleroderma Center, Thomas Jefferson University, 233 S. 10th Street, Room 509 BLSB, Philadelphia, PA 19107-5541, USA. ⁴Center for Genomics, Kimmel Cancer Center, Thomas Jefferson University, Philadelphia PA 19107-5541, USA. ⁵These authors contributed equally: Giuseppina Abignano, Heidi Hermes, Francesco Del Galdo and Sergio A. Jimenez. ✉email: f.delgaldo@leeds.ac.uk; sergio.jimenez@jefferson.edu

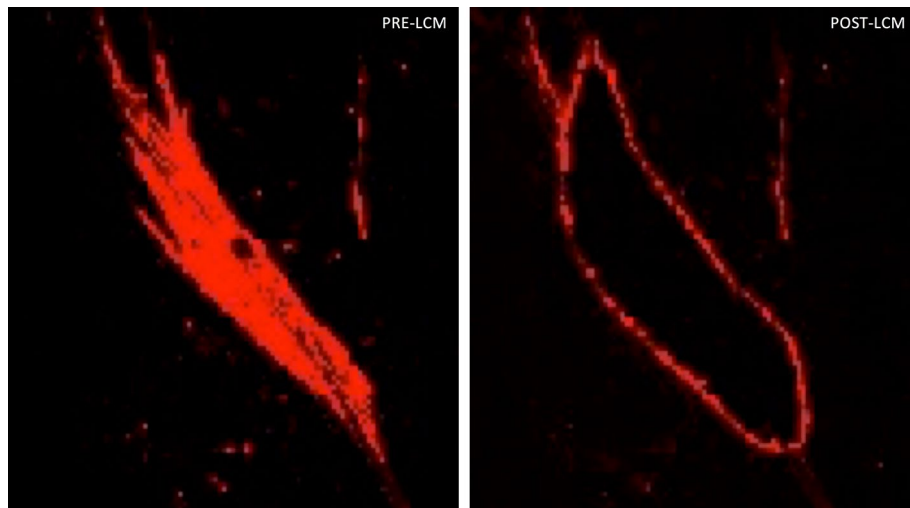


Figure 1. Laser capture microdissection of SSc myfibroblasts from monolayer cultures of dermal fibroblasts expanded from affected SSc skin in vitro. The image shows immunofluorescence staining (red) for α -SMA before (PRE-LCM) and the empty space following its removal (POST-LCM) employing laser capture microdissection. The image is representative of 250 cells per experiment and four different experiments from cultures obtained from four different dcSSc skin biopsies. Original magnification $400\times$.

The origins of the myfibroblasts responsible for the exaggerated and uncontrolled production of collagen and other extracellular matrix proteins in SSc have not been completely elucidated. Extensive studies have shown that these cells may originate from several sources, including expansion and activation of resident tissue fibroblasts¹³, migration and tissue accumulation of bone marrow-derived circulating fibrocytes^{14,15}, or from trans-differentiation of epithelial or endothelial cells to mesenchymal cells, complex processes known, respectively, as epithelial to mesenchymal transition^{16–19} and endothelial to mesenchymal transition^{20–23}. The relative contribution and importance of the different myfibroblast sources to the fibrotic alterations in SSc is unknown, and it is likely to depend on the stage of evolution of the disease as well as on the specific organ involved. Furthermore, it is likely that myfibroblasts originated from different cellular sources or present in different tissues may display unique features or play distinct roles in the pathogenesis or perpetuation of tissue fibrosis in SSc^{7–13,24}. For example, it has been shown that the global gene expression profile of SSc myfibroblasts present in lungs from patients with SSc-associated interstitial lung disease examined employing single cell RNA sequencing assays displays unique gene expression signatures compared to two different types of fibroblasts²⁵. However, the gene expression profile of SSc dermal myfibroblasts has not been examined and it is not known whether these cells may represent a specific cellular phenotype. Thus, the purpose of this study was to analyze the specific transcriptome of pure SSc myfibroblasts. To accomplish this goal myfibroblasts expressing α -SMA actin were isolated employing laser capture microdissection (LCM) from monolayer cultures of cells derived from skin biopsies from affected SSc skin.

Results

Laser capture microdissection of α -SMA-positive myfibroblasts and α -SMA-negative fibroblasts. To determine the global transcriptome of α -SMA positive SSc myfibroblasts we examined monolayer cell cultures that had been expanded from biopsies excised from the leading (proximal) edge of forearm skin lesions of four patients with dcSSc of recent onset. Passage 3 cells were dissociated and subcultured directly on Leica LCM slides and when the slides were essentially fully covered by growing cells, a quick fixation/staining protocol on ice was performed to preserve the mRNA integrity. Alpha-SMA positive myfibroblasts and α -SMA-negative fibroblasts were isolated employing LCM (Fig. 1) and the mRNA of the separate populations was amplified. Real Time PCR was used to confirm the enrichment of α -SMA in the myfibroblasts. The results showed that α -SMA expression was an average of 370% greater in the α -SMA-positive myfibroblasts compared to the α -SMA-negative fibroblasts as described elsewhere²⁶.

Differential gene expression of α -SMA-positive SSc myfibroblasts compared with α -SMA-negative SSc fibroblasts.

The global mRNA expression pattern of α -SMA-positive SSc myfibroblasts isolated employing LCM from monolayer cell cultures derived from affected SSc skin identified 97 genes differentially expressed (greater than 1.5 fold with a p value of < 0.5) compared to the transcriptome of LCM captured α -SMA-negative SSc fibroblasts. Twenty one transcripts were downregulated and 76 upregulated (Supplementary Data S1). A representative heat map of the corresponding genes as an average of the two sets of samples (4 α -SMA-negative samples vs. 4 α -SMA-positive samples) is shown in Fig. 2A. A heat map showing the 8 samples in one single figure, and 4 additional heat maps showing separate images of each individual pair comparison are included in Supplementary Figs. S1 and S2. The full microarray data are publicly available through NCBI GEO website (<http://www.ncbi.nlm.nih.gov/geo/>). Annotation Clustering analysis indicated that

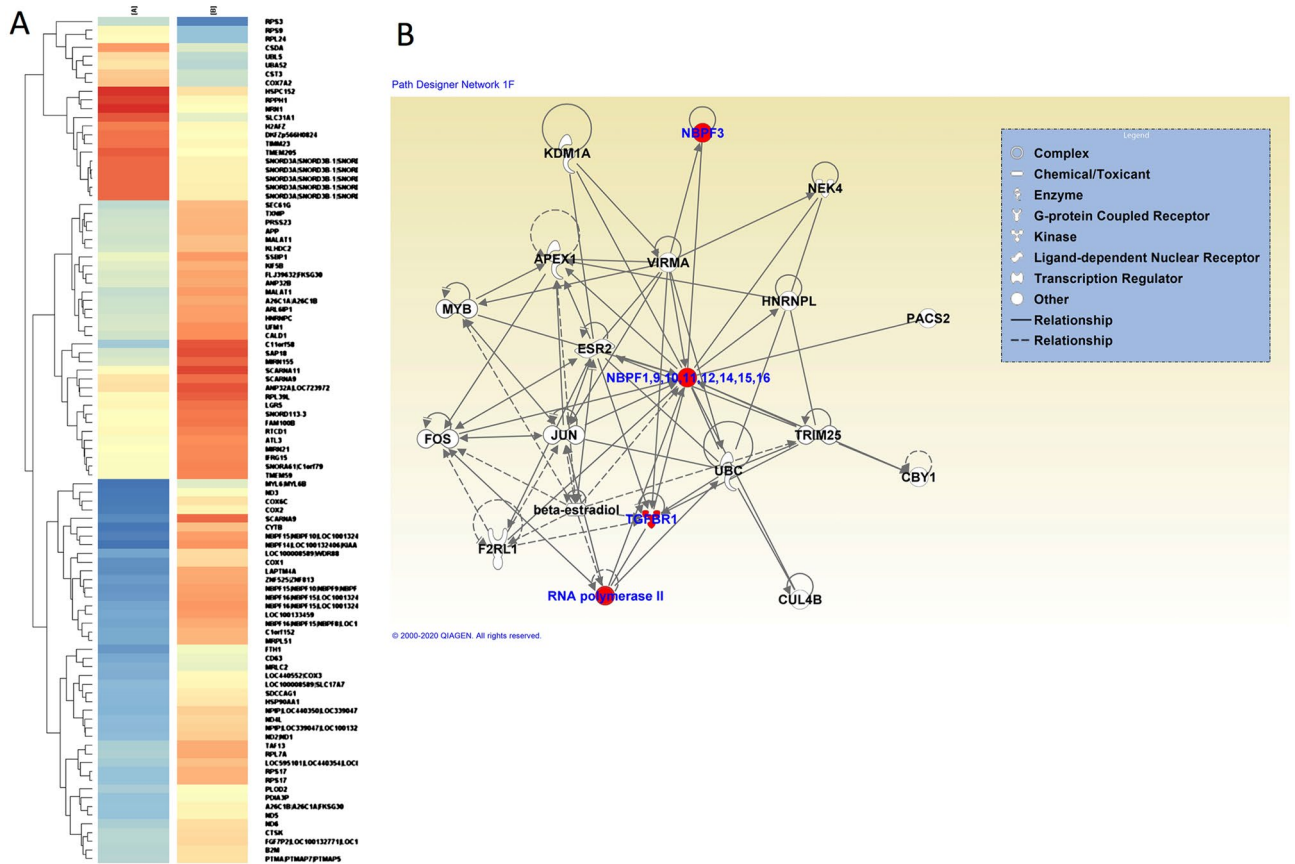


Figure 2. Hierarchic clustering of the gene expression levels in SSc myofibroblasts compared to non-activated fibroblasts. Comparison of gene expression levels of α -SMA positive SSc myofibroblasts and α -SMA negative SSc fibroblasts isolated from the same culture dish employing LCM. (A) The image shows the heat map containing the genes as an average of the two sets of samples (4 α -SMA-negative “A” samples vs. 4 α -SMA-positive “B” samples). Gene expression levels are depicted as color variation from red (high expression) to blue (low expression). The gene expression microarrays were performed employing Affymetrix gene chips (human gene 1.0 ST arrays), and the analysis was performed utilizing the Agilent Gene Spring Software 11.5 also from Affymetrix. (B) Schematic representation of the genes interacting with NBPf genes according to Ingenuity Pathway Analysis (IPA), performed to obtain biological and functional network analysis and utilizing ingenuity.com 8.0 software. The complete list of the genes and pathways identified by Ingenuity CanPath is shown in Supplementary Table S1.

the transcripts with the highest enrichment score corresponded to members of the NBPf cluster which displayed an enrichment score of 2.93. Multiple NBPf transcripts were found in the array, that corresponded to at least six different NBPf genes (Fig. 2A). Canonical pathway analysis according to Ingenuity software identified numerous relevant genes related to inflammatory and fibrotic pathways. A diagram of the genes interacting with NBPf genes is shown in Fig. 2B. A full list of CanPath analysis is summarized in Supplementary Table S1. Data of the quality and quantity of RNA samples used in the microarray experiments are shown in Supplementary Fig. S4 reporting the list of the RNA samples with their concentrations, the 260/280 ratio for demonstration of their high purity and the nanodrop analysis.

NBPf protein expression in systemic sclerosis affected skin. To determine whether the increased expression of NBPf observed in the microarray of the LCM SSc myofibroblasts was also present in vivo, we analysed by immunohistochemistry skin biopsies from three dcSSc patients and biopsies from three healthy subjects obtained from the same anatomical region (forearm). Immunohistochemistry studies indicated that epithelial cells within the epidermis showed a strong and homogeneous NBPf staining in both healthy control skin and in SSc skin biopsies. In contrast, in SSc skin biopsies virtually all the cellular elements within the papillary dermis, showed a strong positivity for NBPf whereas in the dermis of healthy control skin NBPf expression was confined to fewer cells (Fig. 3A). These observations were confirmed by Western Blot analysis of cell extracts from confluent cultures of dermal fibroblasts established from patients with dcSSc and from normal control skin. These studies showed a substantial increase in protein bands corresponding to numerous NBPf transcript levels in cultures of SSc dermal fibroblasts compared to normal dermal fibroblasts (Fig. 3B, C).

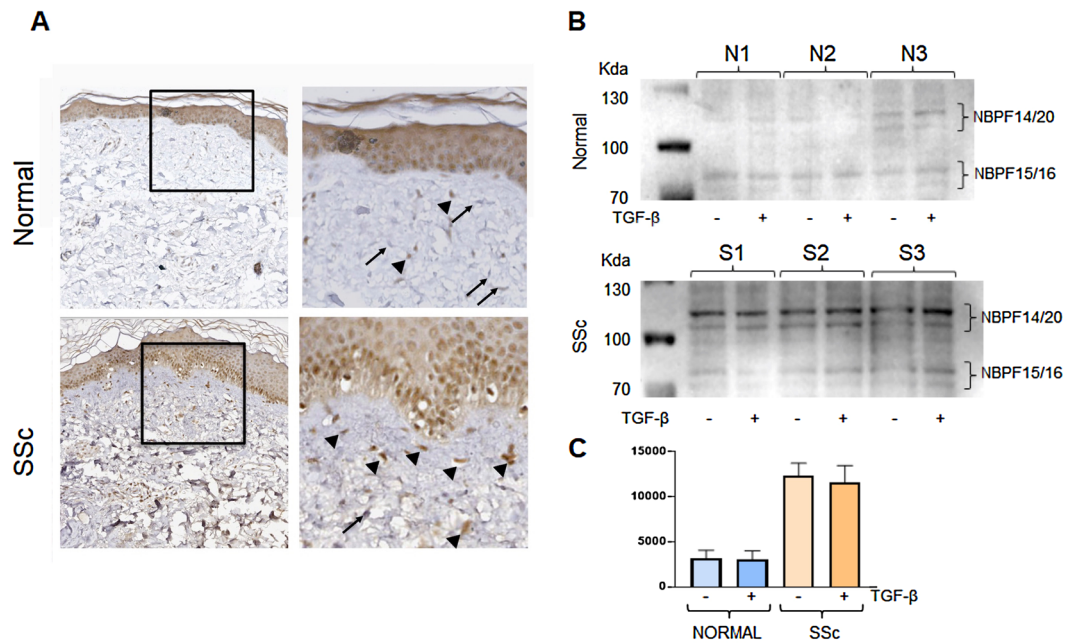


Figure 3. Increased expression of NBPF in affected SSc skin and in cultured SSc dermal fibroblasts. **(A)** Immunohistochemistry for NBPF on one representative skin biopsy from a healthy control (normal) or from a SSc patient. Note that fibroblasts within healthy control skin did not show any staining for NBPF (black arrows) whereas fibroblasts within SSc dermis were intensely stained in brown (arrow heads). Original magnification 20X. Panels are representative of five independent skin biopsies. Panels on the right are a tenfold magnification of the squared inset in the panels on the left. **(B)** Western blots of cell extracts from three different normal fibroblast cell lines and three different SSc dermal fibroblast cell lines cultured in monolayers with or without TGF- β 1 (10 ng/ml) for 24 h employing an NBPF antibody. The NBPF bands shown have been identified based on their expected molecular weights. **(C)** Bar graph displaying a quantitative assessment of the data shown in **(B)**. The gel blots used to prepare the images shown in **(B)** and **(C)** were cropped above the 130 KD and below the 70 KD molecular size marker bands because the migration of the relevant NBPF molecular species was expected to be between these molecular sizes. There were no other crops in the gel blots. Full blots of Western Blot included in this figure are shown in Supplementary Fig. S5.

Dermal fibroblast expression of NBPF in vitro is not stimulated by TGF- β . To examine whether the expression levels of NBPF gene products were modified by exposure to TGF- β , qRT-PCR and Western blot analysis of normal dermal fibroblasts cultured in vitro in the presence of TGF- β were examined. The results showed that incubation of the cells with 10 ng/ml of human recombinant TGF- β for either 24 or 48 h failed to induce any detectable increase in the expression of NBPF at either the mRNA or the protein level (data not shown). Similar results were obtained at the protein level following treatment of cultured normal and SSc dermal fibroblasts with TGF- β for 24 h as shown in Fig. 3B, C.

Discussion

Extensive experimental evidence indicates that myofibroblasts are the key cellular elements in normal tissue repair responses^{5–8} and have been shown to play a crucial role in the pathogenesis of tissue fibrosis in SSc and other fibrotic disorders^{9–12}. It has already been shown that these cells may originate from numerous sources^{7,13,24}, including TGF- β -induced activation of quiescent fibroblasts^{7,8,24} as well as from circulating fibrocytes^{14,15} or from epithelial or endothelial cells undergoing a phenotypic transition into mesenchymal cells^{16–23}. However, there is extreme heterogeneity in their proportion compared to non-myofibroblasts in the skin of SSc patients^{9–13}. The unequivocal evidence that within affected SSc skin there is remarkable heterogeneity and variability in the abundance of myofibroblasts compared to that of quiescent fibroblasts is of remarkable interest taking into account the premise that essentially all cells within a skin biopsy are exposed to the same microenvironment in vivo.

Our study aimed to identify the differences in gene expression between α -SMA-positive LCM isolated myofibroblasts and LCM isolated α -SMA-negative fibroblasts derived from monolayer cultures established from skin biopsies from patients with diffuse SSc of recent onset. Microarray analysis indicated that 97 transcripts were differentially expressed in the myofibroblasts compared to the non-myofibroblasts isolated employing LCM from the same monolayer cultures of SSc skin derived cells. Among these, some transcripts were clearly related to profibrotic activation, whereas others were ribosomal genes, mitochondrial genes involved in oxidative phosphorylation, or genes involved in cell to cell adhesion and in antigen processing and presentation. A remarkable observation was that the most enriched family of genes within the LCM SSc myofibroblast signature was the NBPF family, which was represented by transcripts corresponding to at least 6 different NBPF genes. This

observation is of substantial interest owing to the fact that the NBPF genes are highly enriched in humans as a result of a marked genetic expansion and human lineage-specific amplification and selection that occurred in the human genome following the evolutionary separation of *Homo sapiens* from chimpanzees^{27,28}. The NBPF encoded genes represent a large gene family comprising as many as 25 distinct genes that are localized almost exclusively within chromosome 1. This large gene family received the nomenclature of NBPF genes because one of its members (NBPF1) was found to be disrupted by a chromosomal translocation in a neuroblastoma patient. Although, at the present, the function(s) of NBPF gene products have not been fully elucidated it has become apparent that besides a possible role in the development of neuroblastoma, the proteins encoded by these genes have important functions in brain development and cognitive functions²⁹. However, the full functional range of these genes is not known. Indeed, it has been recently demonstrated that NBPF encoded proteins may participate in inflammatory/immunologic reactions owing to their role as putative transcription factors responsive to NFκB³⁰. While the lack of functional data on NBPF prevents to imply a specific role in SSc pathogenesis, the remarkable increase in the expression of transcripts corresponding to several NBPF genes in myofibroblasts obtained from SSc patients supports the possibility that these highly expanded human lineage genes could be involved in the pathogenesis of a disease such as SSc that displays human-specific expression^{31,32}. On the other hand, it is possible that NBPF is regulated by matrix stiffness since all the experiments were performed under regular culture conditions, which corresponds to a stiff matrix. Numerous recent studies have indicated a highly important role of matrix stiffness and of the biophysical properties of the extracellular matrix in the regulation of gene expression including a marked profibrotic effect of a stiff matrix on fibroblast/myofibroblast biosynthetic patterns^{33,34}. This aspect is of high relevance and will need to be examined in greater detail. However, it should be emphasized that since both types of laser captured cells studied here were obtained from the same culture matrix it is unlikely that matrix stiffness played a role in the markedly increased differential expression of NBPF gene transcripts we observed.

It is generally accepted that TGF-β plays a crucial role in the pathogenesis of SSc-associated fibrosis^{35–38}. Furthermore, it has been shown in numerous *in vitro* studies that TGF-β is capable of inducing a potent stimulation of the conversion of quiescent mesenchymal fibroblastic cells to activated myofibroblasts³⁹. However, a recent study that examined regulators involved in the phenotype of cultured primary SSc dermal fibroblasts employing RNA interference failed to demonstrate a role for TGF-β in the induction of SSc myofibroblasts⁴⁰. The data presented here are in agreement with these latter observations.

Limitations of our study included the small number of enrolled patients and controls and the lack of correction for multiple comparison, however the results of this study clearly indicate that SSc myofibroblasts may represent a specific lineage of mesenchymal cells with a specific transcriptional signature that could determine a functional phenotype playing a crucial role in the pathogenesis of the SSc-related fibrotic process. The identification of a specific myofibroblast cellular phenotype in SSc with a unique gene expression profile quite distinct from that induced by TGF-β could pave the way to a more precise understanding of the complex pathogenesis of tissue fibrosis in SSc and may be of great value for the development of novel therapies for this currently incurable disease.

Methods

Patients and skin biopsies. Full thickness skin biopsies were surgically obtained from the forearms of four adult patients with SSc of recent onset, which was defined as disease duration of less than 18 months from the appearance of clinically detectable skin induration. All participants provided written informed consent to participate in this study. Informed consent procedure and all experiments were approved by the Institutional Review Board of Thomas Jefferson University (IRB#: 06F.186). All experiments were performed in accordance with relevant guidelines and regulations. The patients satisfied the 2013 ACR/EULAR criteria for the classification of SSc⁴¹ and had the diffuse cutaneous clinical subset (dcSSc) as defined by LeRoy et al.⁴². The SSc skin biopsies were obtained from the proximal edge of expanding forearm lesions. One portion of the biopsies was fixed in formaldehyde and embedded in paraffin for diagnostic and other histopathological and immunohistochemical studies. The remaining tissues were used for isolation and establishment of monolayer cultures of dermal cell lines as described previously^{43,44}. All participants provided written informed consent to participate in this study. Informed consent procedure was approved by the Institutional Review Board of Thomas Jefferson University (IRB#: 06F.186) as described previously^{43,44}.

Cell cultures. Cell cultures were established from SSc skin biopsies obtained from the leading edge of forearm lesions from untreated patients with dcSSc of recent onset and from the dorsal side of forearms from normal volunteers as described previously^{43,44}. The cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) (Invitrogen, Carlsbad, CA) supplemented with 10% FBS, antibiotics and glutamine (complete medium) until confluent and were used before reaching passage 6.

Laser capture microdissection (LCM). For LCM, third passage monolayer cultures of dermal cells derived from the skin biopsies of four SSc patients were dissociated and plated directly onto polyethylene naphthalate membrane-coated slides (Leica Microsystems, Wetzlar, Germany), which had been pretreated with RNase Zap (Ambion, Life Technologies, NY), sterilized with 100% ethanol, and allowed to air dry under an ultraviolet light prior to being placed in a standard polystyrene tissue culture dish. Following 48 h culture, the cells were fixed with acetone and incubated for 7 min on dry ice with a mouse anti-α-SMA primary antibody (1:50, Thermo Fisher Scientific, Fremont, CA), followed by 3 min incubation with fluorescein rabbit anti-mouse Cy3 secondary antibody (1:200, Sigma-Aldrich, Saint Louis, MO). The antibody diluent consisted of 2% BSA in PBS with added RNase inhibitor (1 U/μl final concentration). Cold PBS was used for all interval rinses. The

slides were subsequently dehydrated in serial ethanol dilutions and placed on dry ice. LCM and pressure cell capturing were performed immediately following antibody staining according to a modification of previously published procedures⁴⁵. The laser capture microdissected cells were collected in RNA-ase free tubes to avoid RNA degradation, and were stored at -80°C until processed as described in detail previously²⁶.

Immunohistochemistry. For Immunohistochemistry, paraffin-embedded skin sections from 3 dcSSc patients and 3 sex- and age-matched healthy controls, were incubated with heat retrieval solution (Access Revelation, Menapath, A. Menarini Diagnostics, UK) and then transferred to the slide wash following heat mediated antigen retrieval according to the manufacturer's protocol. Slides were then incubated with 3% H_2O_2 for 20 min to inhibit endogenous peroxidase activity and for additional 20 min with 1% casein, for protein-blocking. Neuroblastoma Breakpoint Family (NBPF) was detected by incubation at a 1:50 dilution, for 2 h at room temperature with a polyclonal anti-NBPF antibody that detects endogenous levels of total NBPF including NBPFs 1/9/10/12/14/15/16 and 20 (NBPF1-71614; Novus Biologicals, Littleton, CO). Slides were then incubated with NovocastraTM Post Primary (Leica Biosystems, UK) for 30 min, washed in TBS and incubated with NovolinkTM Polymer for 30 min to detect the tissue-bound primary antibody. Sections were further incubated with the substrate/chromogen, 3,3'-diaminobenzidine (DAB), prepared from NovocastraTM DAB Chromogen and NovolinkTM DAB substrate buffer, according to the manufacturer's protocol. Sections were counterstained with hematoxylin.

RNA isolation, amplification and quantitative real time PCR. Two hundred and fifty α -SMA positive SSc myofibroblasts and two hundred and fifty α -SMA negative SSc fibroblasts were harvested by LCM as described above and then processed for RNA extraction (RNAeasy kit; Qiagen, Valencia, CA) as described previously^{26,46}. Normal and SSc human dermal fibroblasts were cultured in monolayers until reaching confluency and then they were treated with trypsin-EDTA, washed in PBS and then processed for RNA extraction (RNAeasy kit; Qiagen, Valencia, CA) according to the manufacturer's instructions. Amplified cDNA was prepared using nuGEN WT-Ovation RNA amplification system (nuGEN, San Carlos, CA) followed by the NuGEN WT-exon module. Total RNA (2 μg) was reverse-transcribed using Super-script-II reverse transcriptase (Invitrogen, Carlsbad, CA). Real-time PCR was performed using SYBR green Master mix chemistry employing a standard amplification protocol on a Bio-Rad MyiQ real-time PCR system (Bio-Rad, Hercules, CA). Reactions were conducted as described previously^{26,46}. All primers were designed using Primer Express (Applied Biosystems, Foster City, CA) and validated employing the National Center for Biotechnology Information (NCBI) database for specificity. The primers used to amplify NBPF sequences were: Forward-TAAGGGAGAAGTTGCGGAA; and Reverse-AGTGAGGAGGGCCTGGAGAT.

Global gene expression profiling of LCM-isolated SSc myofibroblasts. Total RNA was extracted from LCM isolated SSc myofibroblasts and non-myofibroblastic SSc cells using the Qiagen RNeasy Mini kit (Qiagen, Valencia, CA). A DNase I digestion step was included to eliminate DNA contamination. Four separate pairs of samples were studied. Assessment for each group was performed in triplicate. RNA was quantified on a NanoDrop spectrophotometer (Nanodrop- ThermoScientific, Wilmington, DE), followed by RNA quality assessment on an Agilent 2100 bioanalyzer (Agilent, Palo Alto, CA, USA). Amplification of cDNA was performed using the Ovation Pico WTA-system V2 RNA amplification system (NuGen Technologies, Inc.). Briefly, 50 ng total RNA was reverse transcribed using a chimeric cDNA/mRNA primer, and a second complementary cDNA strand was synthesized. Purified cDNA was then amplified with ribo-SPIA enzyme and SPIA DNA/RNA primers (NuGen Technologies, Inc.). Amplified DNA was purified with Qiagen MinElute (Qiagen) reaction cleanup kit. The concentration of purified cDNA was measured using the NanoDrop spectrophotometer and 2.5 μg cDNAs were fragmented and chemically labeled with biotin to generate biotinylated ST-cDNA using FL-Ovation cDNA biotin module V2 (NuGen Technologies, Inc.).

Global gene expression microarrays were performed employing Affymetrix gene chips (human gene 1.0 ST array; Affymetrix, Santa Clara, CA) as described previously^{26,46}. The gene chips were hybridized with fragmented and biotin-labeled target (2.5 μg) in 110 μl of hybridization cocktail. Target denaturation was performed at 99°C for 2 min and then 45°C for 5 min followed by hybridization for 18 h. Arrays were then washed and stained using Gene chip Fluidic Station 450, using Affimetrix Gene chip hybridization wash and stain kit. Chips were scanned on an Affymetrix Gene Chip Scanner 3000, using Command Console Software. Data analyses were performed using Gene Spring software 11.5 (Agilent Technologies, Inc., Santa Clara, CA). The probe set signals were calculated with the Iterative Plier 16 normalization algorithm; baseline to median of all samples was used as baseline option. The data were filtered by percentile and a lower cut off was set at 25. The criteria for differentially expressed genes were set at ≥ 1.5 -fold changes. Statistical analysis was performed to compare the results obtained from 4 samples of α -SMA-positive SSc myofibroblasts displaying the highest α -SMA transcript levels to the results obtained from 4 samples of α -SMA-negative SSc fibroblasts using T-Test unpaired with p -value equal or less than 0.5. Heat maps were generated from the differentially expressed gene list. The list of differentially expressed genes was loaded into Ingenuity Pathway Analysis (IPA) 8.0 software (<http://www.ingenuity.com>) to perform biological network and functional analyses.

Western blot analysis. Dermal fibroblasts isolated from SSc patients and healthy donors were cultured in monolayers with and without TGF- β 1 (10 ng/mL) for 24 h, then harvested and lysed in RIPA lysis buffer containing protease inhibitors. For Western blots, aliquots of the extracts containing 30 μg of protein were separated by electrophoresis in 8% TrisGlycine polyacrylamide gels and the separated proteins were electroblotted onto nitrocellulose membranes (Invitrogen). The membranes were blocked in PBS/0.5% dry milk/0.01% normal goat

serum for 1 h at room temperature. The transferred proteins were incubated overnight at 4 °C with an NBPF antibody that recognizes NBPF species 1, 9, 10, 12, 14, 15, 16 and 20 (Assay Biotech, Fremont CA). Rabbit secondary antibody coupled to peroxidase and the ECL detection system (Thermo Scientific Pierce) were employed for detection. Quantitative assessment of the protein in the blots was performed employing quantitative densitometry. The identity of the various NBPF protein bands was assigned based on their estimated molecular weight.

Data availability

The full microarray data are publicly available through NCBI GEO website (<http://www.ncbi.nlm.nih.gov/geo/>).

Received: 30 December 2020; Accepted: 8 September 2021

Published online: 14 October 2021

References

- Gabrielli, A., Avvedimento, E. V. & Krieg, T. Scleroderma. *N. Engl. J. Med.* **360**, 1989–2003 (2009).
- Yannick, A. *et al.* Systemic sclerosis. *Nat. Rev. Dis. Primers* **1**, 15002 (2015).
- Denton, C. P. & Khanna, D. Systemic sclerosis. *Lancet* **390**, 1685–1699 (2017).
- Varga, J. & Abraham, D. Systemic sclerosis: A prototypic multisystem fibrotic disorder. *J. Clin. Invest.* **117**, 557–567 (2007).
- Gabbiani, G. The myofibroblast in wound healing and fibrocontractive diseases. *J. Pathol.* **200**, 500–503 (2003).
- Tomasek, J. J., Gabbiani, G., Hinz, B., Chaponnier, C. & Brown, R. A. Myofibroblasts and mechano-regulation of connective tissue remodeling. *Nat. Rev. Mol. Cell. Biol.* **3**, 349–363 (2002).
- Hinz, B. *et al.* Recent developments in myofibroblast biology: Paradigms for connective tissue remodeling. *Am. J. Pathol.* **180**, 1340–1355 (2012).
- Darby, I. A., Zakuan, N., Billet, F. & Desmoulière, A. The myofibroblast, a key cell in normal and pathological tissue repair. *Cell. Mol. Life Sci.* **73**, 1145–1157 (2016).
- Abraham, D. J., Eckes, B., Rajkumar, V. & Krieg, T. New developments in fibroblast and myofibroblast biology: Implications for fibrosis and scleroderma. *Curr. Rheumatol. Rep.* **9**, 136–143 (2007).
- Gyftaki-Venieri, D. A., Abraham, D. J. & Ponticos, M. Insights into myofibroblasts and their activation in scleroderma: Opportunities for therapy?. *Curr. Opin. Rheumatol.* **30**, 581–587 (2018).
- van Caam, A., Vonk, M., van den Hoogen, F., van Lent, P. & van der Kraan, P. Unraveling SSc pathophysiology; the myofibroblast. *Front. Immunol.* **9**, 2452 (2018).
- Kissin, E. Y., Merkel, P. A. & Lafyatis, R. Myofibroblasts and hyalinized collagen as markers of skin disease in systemic sclerosis. *Arthritis Rheum.* **54**, 3655–3660 (2006).
- Korman, B. Evolving insights into the cellular and molecular pathogenesis of fibrosis in systemic sclerosis. *Transl. Res.* **209**, 77–89 (2019).
- Bucala, R., Spiegel, L. A., Chesney, J., Hogan, M. & Cerami, A. Circulating fibrocytes define a new leukocyte subpopulation that mediates tissue repair. *Mol. Med.* **1**, 71–81 (1994).
- Bellini, A. & Mattoli, S. The role of the fibrocytes, a bone marrow-derived mesenchymal progenitor, in reactive and reparative fibroses. *Lab. Invest.* **87**, 858–870 (2007).
- Thiery, J. P., Acloque, H., Huang, R. Y. & Nieto, M. A. Epithelial-mesenchymal transitions in development and disease. *Cell* **139**, 871–890 (2009).
- Guarino, M., Tosoni, A. & Nebuloni, M. Direct contribution of epithelium to organ fibrosis: Epithelial-mesenchymal transition. *Hum. Pathol.* **40**, 1365–1376 (2009).
- Stone, R. *et al.* Epithelial-mesenchymal transition in tissue repair and fibrosis. *Cell Tissue Res.* **365**, 495–506 (2016).
- Rosa, I., Romano, E., Fioretto, B. S. & Manetti, M. The contribution of mesenchymal transitions to the pathogenesis of systemic sclerosis. *Eur. J. Rheumatol.* **7**, S157–S164 (2020).
- Jimenez, S. A. Role of endothelial to mesenchymal transition in the pathogenesis of the vascular alterations in systemic sclerosis. *ISRN Rheumatol.* **2013**, 835948 (2013).
- Manetti, M. *et al.* Endothelial-to-mesenchymal transition contributed to endothelial dysfunction and dermal fibrosis in systemic sclerosis. *Ann. Rheum. Dis.* **76**, 924–934 (2017).
- Mendoza, F. A., Piera-Velazquez, S., Farber, J. L., Feghali-Bostwick, C. & Jimenez, S. A. Endothelial cells expressing endothelial and mesenchymal cell gene products in lung tissue from patients with systemic sclerosis-associated interstitial lung disease. *Arthritis Rheumatol.* **68**, 210–217 (2016).
- Piera-Velazquez, A. & Jimenez, S. A. Endothelial to mesenchymal transition: Role in physiology and in the pathogenesis of human disease. *Physiol. Rev.* **99**, 1281–1324 (2019).
- Kramann, R., DiRocco, D. P. & Humphreys, B. D. Understanding the origin, activation and regulation of matrix-producing myofibroblasts for treatment of fibrotic disease. *J. Pathol.* **231**, 273–289 (2013).
- Valenzi, E. *et al.* Single-cell analysis reveals fibroblast heterogeneity and myofibroblasts in systemic sclerosis-associated interstitial lung disease. *Ann. Rheum. Dis.* **78**, 1379–1387 (2019).
- Wasson, C. W. *et al.* Long non-coding RNA HOTAIR drives EZH2-dependent myofibroblast activation in systemic sclerosis through miRNA 34a-dependent activation of NOTCH. *Ann. Rheum. Dis.* **79**, 507–517 (2020).
- Vandepoele, K., Van Roy, N., Staes, K., Speleman, F. & van Roy, F. A novel gene family NBPF: Intricate structure generated by gene duplications during primate evolution. *Mol. Biol. Evol.* **22**, 2265–2274 (2005).
- Vandepoele, K. *et al.* A constitutional translocation t(1;17) (p36.2;q11.2) in a neuroblastoma patient disrupts the human NBPF1 and ACCN1 genes. *PLoS ONE* **3**, e2207 (2008).
- Dumas, L. J. *et al.* DUF1220-domain copy number implicated in human brain-size pathology and evolution. *Am. J. Hum. Genet.* **91**, 444–454 (2012).
- Zhou, F. *et al.* NBPF is a potential DNA-binding transcription factor that is directly regulated by NF- κ B. *Int. J. Biochem. Cell Biol.* **45**, 2479–2490 (2013).
- Cooper, D. N. & Kehrer-Sawatzki, H. Exploring the potential relevance of human-specific genes to complex disease. *Hum. Genom.* **5**, 99–107 (2011).
- Jimenez, S. A. & Piera-Velazquez, S. Potential role of human-specific genes, human-specific microRNAs and human-specific non-coding regulatory RNAs in the pathogenesis of systemic sclerosis and Sjögren's syndrome. *Autoimmun. Rev.* **12**, 1046–1051 (2013).
- Hinz, B., Mc Cullock, C. A. & Coehlo, N. M. Mechanical regulation of myofibroblast phenocopy and collagen contraction. *Exp. Cell Res.* **379**, 119–128 (2019).
- Schulz, J. N. *et al.* New developments on skin fibrosis—Essential signals emanating from the extracellular matrix for the control of myofibroblasts. *Matrix Biol.* **68–69**, 522–532 (2018).

35. Denton, C. P., Black, C. M. & Abraham, D. J. Mechanisms and consequences of fibrosis in systemic sclerosis. *Nat. Clin. Pract. Rheumatol.* **2**, 134–144 (2006).
36. Varga, J. & Whitfield, M. L. Transforming growth factor-beta in systemic sclerosis (scleroderma). *Front. Biosci. (Sch. Ed.)* **1**, 226–235 (2009).
37. Jimenez, S. A., Castro, S. V. & Pitera-Velazquez, S. Role of growth factors in the pathogenesis of tissue fibrosis in systemic sclerosis. *Curr. Rheumatol. Rev.* **6**, 283–294 (2010).
38. Lafyatis, R. Transforming growth factor β -at the centre of systemic sclerosis. *Nat. Rev. Rheumatol.* **10**, 706–719 (2014).
39. Carthy, J. M. TGF β signaling and the control of myofibroblast differentiation: Implications for chronic inflammatory disorders. *J. Cell. Physiol.* **233**, 98–106 (2018).
40. Chadli, L. *et al.* Identification of regulators of the myofibroblast phenotype of primary dermal fibroblasts from early diffuse systemic sclerosis patients. *Sci. Rep.* **9**, 4521 (2019).
41. van den Hoogen, F. *et al.* 2013 classification criteria for systemic sclerosis: An American college rheumatology/European league against rheumatism collaborative initiative. *Ann. Rheum. Dis.* **72**, 1747–1755 (2013).
42. LeRoy, E. C. *et al.* Scleroderma (systemic sclerosis): Classification, subsets and pathogenesis. *J. Rheumatol.* **15**, 202–205 (1988).
43. Jimenez, S. A., Feldman, G., Bashey, R. I., Bienkowski, R. & Rosenbloom, J. Co-ordinate increase in the expression of type I and type III collagen genes in progressive systemic sclerosis fibroblasts. *Biochem. J.* **237**, 837–843 (1986).
44. Del Galdo, F. *et al.* Decreased expression of caveolin 1 in patients with systemic sclerosis: crucial role in the pathogenesis of tissue fibrosis. *Arthritis Rheum.* **58**, 2854–2865 (2008).
45. Tadros, Y., Ruiz-Deya, G., Crawford, B. E., Thomas, R. & Abdel-Mageed, A. B. In vivo proteomic analysis of cytokine expression in laser capture-microdissected urothelial cells of obstructed ureteropelvic junction procured by laparoscopic dismembered pyeloplasty. *J. Endourol.* **17**, 333–336 (2003).
46. Del Galdo, F., Wermuth, P. J., Addya, S., Fortina, P. & Jimenez, S. A. NF κ B activation and stimulation of chemokine production in normal human macrophages by the gadolinium-based magnetic resonance contrast agent Omniscan: Possible role in the pathogenesis of nephrogenic systemic fibrosis. *Ann. Rheum. Dis.* **69**, 2024–2033 (2010).

Acknowledgements

FDG and GA were supported by National Institute for Health Research (NIHR) Leeds Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The expert assistance of Alana Pagano in the preparation of this manuscript is gratefully acknowledged.

Author contributions

F.D.G. and S.A.J. designed the study. G.A., H.H., S.P.V., S.A. performed study experiments. G.A. wrote the manuscript. F.D.G. and S.A.J. contributed to writing and critical revisions. All authors read and approved the final manuscript.

Funding

NIH/NIAMS Grant RO-119616 to SAJ; EULAR ODP Grant to FDG.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-99292-y>.

Correspondence and requests for materials should be addressed to F.G. or S.A.J.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021