HHS Public Access

Author manuscript

Arthritis Rheumatol. Author manuscript; available in PMC 2024 January 01.

Published in final edited form as:

Arthritis Rheumatol. 2023 January ; 75(1): 108–119. doi:10.1002/art.42281.

RNA-Seq analysis identifies alterations of the primary cilia gene *SPAG17* and *SOX9* locus non-coding RNAs in systemic sclerosis

Elisha D.O. Roberson, PhD[1,2,*], **Mary Carns, MS**[3], **Li Cao, MD**[1], **Kathleen Aren, MPH**[3], **Isaac A. Goldberg, BA**[3], **David J. Morales-Heil, PhD**[1], **Benjamin D. Korman, MD**[3], **John P. Atkinson, MD**[1], **John Varga, MD**[3,4,†,*]

[1]Department of Medicine, Division of Rheumatology, Washington University, St. Louis, MO, USA.

[2]Department of Genetics, Washington University, St. Louis, MO, USA.

[3]Feinberg School of Medicine, Scleroderma Program, Northwestern University, Chicago, IL, USA.

[4]Department of Internal Medicine, Division of Rheumatology, University of Michigan, Ann Arbor, MI, USA.

Abstract

Objective.—Systemic sclerosis (**SSc**) is characterized by immune activation, vasculopathy, and unresolving fibrosis in the skin, lungs, and other organs. We performed RNA-Seq analysis on skin biopsies and peripheral blood mononuclear cells (**PBMCs**) from SSc patients and controls to better understand SSc pathogenesis.

Methods.—We analyzed these data to 1) test for case-control differences, and 2) identify genes whose expression levels correlate with SSc severity as measured by local skin score, modified Rodnan skin score (**MRSS**), forced vital capacity (**FVC**), or diffusion capacity for carbon monoxide (**DLCO**).

Results.—We found that PBMCs from SSc patients showed a strong type 1 interferon signature. This signal replicated in the skin, with additional signals for increased extracellular matrix (**ECM**) genes, classical complement pathway activation, and the presence of B cells. Notably, we observed a marked decrease in the expression of *SPAG17*, a cilia component, in SSc skin. We identified genes that correlated with MRSS, DLCO, and FVC in SSc PBMCs and skin using weighted gene co-expression analysis (**WGCNA**). These genes were largely distinct from the case/control differentially expressed genes. In PBMCs, type 1 interferon signatures negatively correlated with DLCO. In SSc skin, ECM gene expression positively correlated with MRSS. Network analysis of SSc skin genes correlated with clinical features identified the non-coding RNAs *SOX9-AS1* and *ROCR*, both near the *SOX9* locus, as highly connected, “hub-like” genes in the network.

[*]Corresponding authors John Varga, M.D., University of Michigan Medical School, Division of Rheumatology, Ann Arbor, MI 48104, vargaj@med.umich.edu; Elisha D.O. Roberson, Ph.D., Washington University in St. Louis, 660 South Euclid Avenue, MSC 8045-0020-10, St. Louis, MO 63110, eroberson@wustl.edu.
[†]Current address University of Michigan Medical School, Division of Rheumatology, Ann Arbor, MI 48104

CONFLICT OF INTEREST

None declared.

Conclusion.—These results identify non-coding RNAs and *SPAG17* as novel factors potentially implicated in SSc pathogenesis.

INTRODUCTION

Systemic sclerosis (**SSc**), is a complex orphan disease characterized by autoantibodies, vasculopathy of small vessels, and synchronous / unresolving fibrosis in multiple organs (1, 2). There is substantial patient-to-patient heterogeneity in clinical features, disease severity, and the rates of progression. Currently, there are few effective treatments for SSc. Moreover, there is a lack of molecular biomarkers that reliably predict clinical course, reflect disease activity, or identify rational therapeutic targets (3).

One approach to improve our understanding of the evolution and progression of the disease is through transcriptomics. Previous primary and secondary analyses of transcriptome data in SSc used microarray, bulk RNA-Seq, and single-cell RNA-Seq approaches. These studies revealed molecular heterogeneity among individual transcriptomes, increased type 1 interferon signaling, potential molecular subtypes, and altered cell populations in the skin (4-13). We sought to further clarify molecular disruptions in SSc, to find correlations with clinical measures of disease activity, and to determine if expression-trait correlation gene sets overlap with the case/control differential expression gene sets. We used prospective collection of skin and PBMC samples from control subjects and patients with SSc followed by bulk RNA-Seq. At each visit, disease severity was assessed by the local skin score, modified Rodnan skin score (**MRSS**), and pulmonary function testing. For RNA-Seq, we used a ribosomal depletion method to permit the detection of both nascent and mature mRNA along with non-coding RNAs lacking a poly(A) tail. This method may be more sensitive for genes with low expression levels or short half-life than poly(A)-based RNA-Seq methods, enabling us to identify potentially overlooked contributors to SSc. These methodologies allowed us to examine categorical differences between SSc patients and unaffected controls, as well as to identify genes whose expression is correlated with alterations in established clinical parameters of disease progression.

RESULTS

Study cohort and demographics

Systemic sclerosis cases (n=21) were recruited from the Northwestern Scleroderma Clinic and fulfilled classification criteria for SSc (1). These patients were further classified into limited cutaneous SSc (**lcSSc**; n=5), diffuse cutaneous SSc (**dcSSc**; n=14), SSc sine scleroderma (**SSS**; n=1), and very early diagnosis of systemic sclerosis (**VEDOSS**; n=1). Controls were volunteers with no history of an autoimmune or inflammatory disease (n=14). At each study visit, we obtained whole-blood and two 3 mm skin punch biopsies. We also obtained pulmonary function tests, and the same observer assessed the MRSS and local skin score. Seven dcSSc, two lcSSc, and one SSS individual volunteered for a second sampling and assessment at a 6-month follow-up visit.

For group-wise demographic summaries, we included controls and subjects with either lcSSc or dcSSc (n=19; Table 1). Control subjects were younger than SSc patients. Within

the SSc cohort, the lcSSc and dcSSc subsets were balanced for age and disease duration (Table 2). There were more women in the dcSSc than the lcSSc subset, while self-declared ethnicity was similar between the two. Medications at sample collection are listed in ST29.

SSc PBMCs demonstrate evidence of increased type 1 interferon signaling

We sought to characterize gene expression changes of SSc PBMCs since it is a minimally invasive tissue source. For cases, we only included baseline lcSSc and dcSSc samples to avoid bias toward individuals sampled more than once. There were 147 genes with decreased expression in SSc PBMCs (113 at least −1.5-fold) and 100 genes with increased expression (61 at least 1.5-fold; Fig. 1A; Table ST1). The most significantly decreased genes included *GALNTL6* (polypeptide N-aceltylgalactosaminyltransferase-like 6; −4.09 fold-change [**FC**]), *GPM6A* (glycoprotein M6A; −3.98 FC), *SLC4A10* (solute carrier family 4 member 10; −3.33 FC), and *COL4A3* (collagen type IV alpha 3 chain; −4.36 FC. The most consistent pathway enrichment amongst genes with decreased expression in SSc was for collagen pathways due to decreases in *COL4A3* and *COL4A4* (Fig. 1B; Table ST2).

Genes with significantly increased expression in SSc PBMCs compared to controls included *FAM13A* (family with sequence similarity 13, member A; 2.49), *E2F2* (E2F transcription factor; 2.28), *MYOF* (myoferlin; 1.83), and *TMEM178B* (transmembrane protein 178B; 1.36). SNPs in *FAM13A* are associated with an increased risk of pulmonary fibrosis (14) and liver cirrhosis (15). We tested the genes with increased expression in SSc PBMCs for pathway enrichment (Fig. 1C; Table ST3). The most significant pathways were related to type 1 interferon activity, including interferon-alpha/beta and type 1 interferon signaling categories. Target genes of IRF5 and IRF9 transcription factors were significantly enriched in the genes with increased expression. *IRF5* is highly expressed in M1 polarized human macrophages (16). A high prevalence of M1-like macrophage signatures has been detected in SSc skin (10).

Previous studies have observed similar increased expression of type 1 interferon-stimulated genes in SSc PBMCs as well as in skin (17, 18). It is also interesting to consider whether some of these signatures, including the increase in IRF5 and its known high expression in M1-polarized macrophages, indicate a specific role for this macrophage subpopulation in systemic sclerosis.

SSc skin shows evidence of substantial immune activation, increased complement component expression, and loss of ciliary protein *SPAG17*

PBMC procurement is minimally invasive, but may not reflect the molecular biology of affected tissues. The skin may therefore provide more insight into the molecular defects in SSc. There were 526 genes significantly decreased (226 at least −1.5-fold) and 1,200 genes increased (816 at least 1.5-fold; Fig. 2A; Table ST4) in SSc compared to control skin biopsies. These results help highlight the advantage of studying affected tissue.

The most significantly altered gene in the entire study was a decrease of *SPAG17* in SSc skin samples (sperm-associated antigen 17; −4.67 FC; adjusted p-value 3.22×10^{-12}). To our knowledge, this gene had not been previously associated with differential

expression in systemic sclerosis skin. Its expression level is relatively low, and studies using hybridization microarrays may not have been sensitive enough to detect its expression versus the background fluorescence. The low expression might also cause it to be filtered out of some RNA-Seq studies. SPAG17 protein is required for the function of primary cilia and male fertility (19). Mice deficient in *Spag17* have bone abnormalities such as decreased femur length and disrupted femur morphology (20). Its role in skin and immune cells is not particularly clear, though as part of the primary cilia it could be involved in signaling. The decreased expression of *SPAG17* does not require appreciable fibrosis, as SSc sine scleroderma skin has decreased *SPAG17* expression levels comparable to lcSSc and dcSSc (Fig. S2). The similarity of the skin transcriptomes from patients with SSS, lcSSc, and dcSSc is also observed by principal components analysis of the top case/control differentially expressed genes (supplemental material and Fig. S4).

Another significantly decreased gene was *LGR5* (leucine-rich repeat-containing G-protein coupled receptor 5; −2.70 FC). *LGR5* is a member of the G protein-coupled receptor family that is an important target and modulator of Wnt/β-catenin signaling. *LGR5* expression is a marker for intestinal villi tip telocytes in mice that maintain the correct differentiation gradient on the villus axis via non-canonical Wnt signaling (21). Dermal telocytes are reduced in the fibrotic skin and internal organs of individuals with SSc (22, 23). If these dermal telocytes also help to coordinate differentiation and/or signaling in the skin, their loss (perhaps detected by the reduction of *LGR5*) may directly contribute to the development of SSc fibrosis.

Only two known pathways were enriched among decreased genes (Fig. 2B; Table ST5): monocarboxylic acid biosynthesis and axonemal central apparatus.

Genes increased in SSc included *ARHGAP45* (rho GTPase activating protein 45; 2.22 FC), *COTL1* (coactosin-like protein 1; 2.18 FC), *IRF8* (interferon regulatory factor 8; 2.57 FC), and *COMP* (cartilage oligomeric matrix protein; 9.63 FC). The extensive list of genes increased in SSc skin would allow us to posit interesting hypotheses for almost any one of them. We, therefore, chose to check these genes for the enrichment of known molecular pathways as well. We found a total of 366 significant enrichments (Fig. 2C; Table ST6). This included enrichment for targets of the transcription factors IRF2, IRF4, IRF5, IRF7, IRF8, IRF9, and ISGF3. The RNA expression of transcription factors *IRF1* (2.65 FC), *IRF5* (1.48 FC), *IRF7* (1.99 FC), and *IRF8* (2.57 FC) were all significantly increased in the skin as well.

Some of the enrichments were variations on similar themes, such as immune cell adhesion, migration, and differentiation, B cell activation and proliferation, extracellular matrix deposition, and classical antibody-mediated complement activation. The complement theme was mainly due to increased expression of complement genes (including *C1QB, CR1, C5AR1*, and *C7*) and immunoglobulins. Complement activation usually leads to the assembly of the membrane attack complex (**MAC**) that can insert into membranes and lyse cells. The MAC is composed of complement proteins C5b, C6, C7, C8, and C9. MAC fragments are deposited in the dermal vasculature of systemic sclerosis patients, supporting a role for antibody-induced complement activation in SSc vasculopathy (24).

These data suggest that aberrant activation of complement may partially mediate cutaneous tissue damage in SSc.

Gene expression in SSc PBMCs and skin correlate with skin fibrosis and lung functions but are mostly missed by differential expression analysis

Group-wise gene expression analysis categorizes samples as test (SSc) or reference (control). This is a reasonable way to determine general features of a trait but ignores the heterogeneity of phenotypes within the trait. SSc in particular has substantial clinical heterogeneity. We used weighted gene co-expression network analysis (**WGCNA**) to correlate gene expression with skin severity (MRSS and forearm local skin score) and lung function (FVC and DLCO). This is a somewhat different approach from other studies that have correlated MRSS to gene expression arrays or lung disease to lung biopsy expression (25-27). We chose to test both skin and lung phenotypes, and had paired samples from PBMCs and skin for each individual. We included all SSc samples for which there was a matching MRSS, FVC, or DLCO measurement for the visit, i.e. follow-ups were included as separate measurements. The tally of positive and negative correlations by tissue is listed in Table ST28.

For PBMCs, there were significant correlations between gene expression and DLCO (Table ST8), forearm skin score (Table ST11), and MRSS (Table ST13). The genes negatively correlated with DLCO were enriched for type 1 interferon signaling and proteasomal antigen processing/presentation (Table ST9). The PBMC genes that positively correlated with DLCO were only enriched for targets of microRNA hsa-miR-6082 (*DTWD2, FXN*, and *ZFP30*; Table ST10).

Forearm-specific skin score correlations (Table ST11) were more difficult to interpret. The negative correlations did not show any pathway enrichments. Positive correlations were only enriched for the phosphoribosylformylglycinamidine synthase pathway, but it was due to a single gene (*PFAS*; Table ST12).

There were also correlations between PBMC gene expression and MRSS (Table ST13). Genes negatively correlated with MRSS were enriched for pathways associated with protein folding, unfolded proteins, and ER stress (Table ST14). There were a few pathways associated with positively correlated genes, including some related to Wnt signaling regulation and mitochondrial functions (Table ST15).

One might assume that the differentially expressed genes in case/control gene expression studies may also be the key genes and pathways that drive progression and therefore may correlate with disease severity. This does not appear to be the case for SSc PBMCs, as most of the genes with significant correlations to severity were not detected in the case/control analysis (Fig. S6).

The SSc skin transcriptomes had significant correlations with DLCO (Table ST16), FVC (Table ST19), and MRSS (ST21). For DLCO, many negative correlations were with ribosomal proteins, leading to enrichment of pathways related to ribosome function (Table ST17). SSc skin genes that positively correlated with DLCO did not fall into a consistent

theme (Table ST18). Negative correlations with FVC were enriched for sterol/cholesterol biosynthesis and alpha-linolenic/linolenic acid metabolism pathways (Table ST20). The positive correlations with FVC did not have any significant pathway enrichment.

Skin genes negatively correlated with MRSS were enriched in cell fate and synaptic categories (Table ST22). Genes positively correlated with MRSS were enriched for genes associated with the extracellular matrix (Table ST23). This is a nice confirmation, as increasing MRSS should indicate increasing fibrosis, secondary to increased matrix deposition.

Similar to what we observed in PBMCs, there is little overlap between the genes differentially expressed in SSc skin and skin genes that correlated with lung function or skin fibrosis parameters (Fig. S7). Overall, 90% of the differentially expressed genes did not correlate with any clinical parameters, and 93% of the clinical trait-associated genes were not differentially expressed. This highlights the disparity between the two methods and suggests that novel targets for clinical treatment and biomarkers may be identified using severity correlation rather than case/control differential expression.

Some genes were correlated with the same trait in both skin and PBMCs (Table ST24). For each gene, we considered the tissues concordant if the direction of effect was the same in each tissue, and discordant if the direction of effect was opposite. There were only overlaps between tissues for DLCO and MRSS. DLCO had 12 concordant genes and 7 discordant genes between tissues. MRSS had 11 concordant and 2 discordant genes. These two gene sets are intriguing to consider for further study as molecular biomarkers of disease activity, regardless of whether the effect is concordant or discordant, as a blood draw might be as informative as a skin punch biopsy.

We were also interested in whether there are some genes where PBMC expression could be as informative as a skin biopsy. A total of 91 genes were differentially expressed in skin and PBMCs (n=24), and/or correlated with a clinical trait in both tissues (n=69). We used Pearson correlation with all samples to see if the expression level of these genes correlated between tissues. Out of these candidates, 9 genes were significantly correlated between tissues (ST32). The PBMC expression of several genes related to interferon activity (*IFI44, IF44L, OAS1, OAS3*, and *RSAD2*) correlated with the expression in skin, and perhaps could be assayed without skin biopsies.

Non-coding RNAs *SOX9-AS1* and *ROCR* are central, highly connected nodes in the SSc skin gene-gene correlation network

Given that we had a list of genes that correlated with different traits and their normalized expression, the next thing we looked at was the gene-gene correlation network. We focused only on genes that significantly correlated with at least one clinical trait. Examining the network characteristics can help identify genes that act as signaling hubs or otherwise have co-expression with other members of the network.

After ranking each gene by degree and page rank (measures of network connectedness and centrality), the top-ranked PBMC gene was 2'-5'-oligoadenylate synthetase 2 (*OAS2*)

with a degree of 29 and page rank of 0.017 (Table ST25; Fig. 3). The oligoadenylate synthetases are interferon-response genes, which is consistent with the increased type 1 interferon signaling signatures in SSc PBMCs. The 4th highest ranked gene, *EPSTI1*, is linked to macrophage function. It's thought to have a key role in classical M1 polarization of macrophages, as a murine knockout of *Epsti1* has few M1 polarized macrophages along with a significant expansion of M2 polarized macrophages (28). The top-ranked genes in SSc PBMCs (all genes with a degree of at least 10) are inflammatory genes that correlated with DLCO.

Conversely, the top 41 genes in the SSc skin network (ranked by degree and page rank) all correlated with MRSS (Table ST26; Fig. 4). The top-ranked SSc skin gene was an antisense transcript, *SOX9-AS1* (degree = 49; page rank = 0.005). This suggests an important advantage of using stranded RNA library kits: with an unstranded kit it is impossible to tell sense from antisense transcripts if they overlap at the same locus. The protein-coding *SOX9* transcript was correlated with MRSS as well but had a lower degree of 9. The long non-coding RNA *ROCR* is also a highly connected gene in the SSc skin network (degree = 44; page rank 0.004). It is located in the same genomic locus as *SOX9* and *SOX9-AS1*. It is worth noting that *SPAG17* expression in SSc skin was negatively correlated with MRSS (−0.51 correlation). Since *SPAG17* is a low expression transcript, we were unable to generate a *SPAG17* network to identify co-expressed genes. Therefore further study is required to understand the genes co-expressed with *SPAG17* in relevant skin cell types, such as fibroblasts.

DISCUSSION

Systemic sclerosis is a complex and progressive inflammatory/fibrotic disease. In the current study, we show that the sensitivity of RNA-Seq can lead to new discoveries and that applying complementary approaches to the same data can reveal distinct trends. There are advantages to using strand-specific, ribosomal depletion RNA-Seq library kits: increased ability to detect non-polyadenylated transcripts, increased sensitivity for nascent and short half-life transcripts, and the ability to distinguish overlapping antisense transcripts. However, a disadvantage is a reduced power for transcript-level discoveries. For our data, focusing only on spliced transcripts would have used only a fraction of the available sequencing data, since much of the aligned sequence mapped to introns. There is also distinct value to using different tissues. Blood samples are informative in their own right, but comparing skin punches between SSc patients and controls reveals different gene sets more directly related to the ongoing molecular pathology.

The predominant signal in PBMCs was for type 1 interferon signaling, with enrichment for targets of IRF5 and IRF9. M1 polarized macrophages are known to have high *IRF5* expression (16). We observed that *CD86*, another marker of M1 macrophages, was increased in SSc PBMCs. Previous work has shown that a mix of monocytes with signs of M1 and M2 polarization in the blood is associated with interstitial lung disease in systemic sclerosis (29). The lack of an increase of M2 markers doesn't necessarily exclude their presence. Bulk RNA-Seq is not ideal for identifying cell populations and the question of macrophage

fraction (and activation state) in SSc PBMCs would be better answered with single-cell RNA-Seq, flow cytometry, or mass cytometry.

The skin data provide an even more informative perspective. The most significantly different gene in the entire analysis was *SPAG17*, which has not traditionally been a top dysregulated gene in SSc transcriptome studies. Genes are often named for the tissue in which they are first discovered, which may or may not be the only tissue they're expressed in or even the tissue with the highest expression. SPAG17 is a central pair protein that is critical in the formation of primary cilia, and its knockout is neonatal lethal in mice (30). As suggested by the name, alterations in *SPAG17* can lead to infertility in mouse models and for humans with certain rare missense variants (19, 31). But it is important to bear in mind that this gene has critical functions beyond sperm motility since it is part of the primary cilia. Missense changes in *SPAG17* can cause abnormal bone length (20, 32). Common variants in *SPAG17* are associated with body length in early life and height in adulthood (33, 34). Novel mutations and rare variants in components of the primary cilia can lead to primary ciliary dyskinesia (**PCD**). The most common effects are in the ears (chronic ear infections, hearing loss), sinuses (chronic sinus congestion), and lungs (recurrent pneumonia, chronic cough). *SPAG17* rare variants can lead to a PCD-like phenotype in mice and have been shown to cause human PCD as well (35, 36). The PCD phenotypes are largely driven by the altered ability of the cilia to beat. In systemic sclerosis skin, a defect of beating cilia doesn't make the most sense. Non-motile primary cilia are involved in signaling. One potential hypothesis is that primary cilia have an anti-fibrotic signaling role. The reduced expression could then lead to increased profibrotic signaling. Yet to be elucidated is whether mouse *Spag17* hypomorphs have increased fibrosis susceptibility, which cells in the skin are affected by reduced *SPAG17* expression, and why the *SPAG17* expression is decreased in the first place, i.e. whether it is a primary or secondary effect. One possibility is that *SPAG17* is lost in the transition from fibroblast to myofibroblast. Further research will be required to dissect these possibilities, particularly since the mouse germline knockout is neonatal lethal.

We detected a decreased expression of *LGR5* in SSc skin. The recent finding that *LGR5* is a marker for mouse intestinal villi tip telocytes begs the question of whether it is a marker of skin telocytes. The intestinal tip telocytes had increased expression of *Wnt5a* compared to the intestinal crypt telocytes, suggesting that non-canonical Wnt signaling is important in those cells (21). Skin telocytes are known to form multiple contacts to ECM and other cells, such as adipose cells and fibroblasts (37). One possible function of these interconnections is to provide support for the other cells within the skin matrix. It is, however, also possible that these cells are critical for the transduction of signals within the skin, and therefore may play a direct role in the evolution of fibrosis in SSc.

SSc skin had enrichment for pathways related to classical (antibody-mediated) complement activation. There is a growing body of evidence that complement activation and subsequent damage play a role in systemic sclerosis endothelial damage. The terminal effector of complement damage, the membrane attack complex, has been observed in the small vessels of affected systemic sclerosis skin and the muscle endothelium of patients with systemic sclerosis-associated myositis (24, 38). This indicates local tissue damage, regardless of how it is triggered, may be complement-mediated. This possibility is perhaps even more

evident in scleroderma renal crisis (**SRC**), which has a sudden onset of severe hypertension and acute renal failure. The kidneys of some individuals with SRC show deposition of complement C3b in renal arterioles (39, 40). Complement deposition without substantial inflammation and the presence of thrombotic microangiopathy is also a hallmark of atypical hemolytic uremic syndrome (**aHUS**). Familial aHUS is often caused by aberrant regulation of complement activation, particularly via genetic variants in complement factor H (41). The first-line therapy for aHUS is eculizumab, a monoclonal antibody to C5 (42). There is some evidence that eculizumab might also be effective for SRC (43, 44). However, this still leaves unanswered whether endothelial complement activation is a major driver of SSc skin vasculopathy.

The network analysis of PBMC and skin transcriptomes with skin fibrosis and lung function parameters demonstrated that both tissues are informative for different traits, opening the possibility of the development of minimally invasive, quantitative biomarkers of disease activity. This would be a boon to clinical trials, particularly if observable parameters, such as MRSS, don't tell the whole story about internal disease progression. Given the relatively small number of samples, further study is warranted to validate these findings. A particularly interesting facet of the clinical trait correlation in our study is the suggestion that SOX9 is a critical player in fibrosis. Expression of the *ROCR* and *SOX9-AS1* non-coding RNAs, as well as *SOX9* itself, was significantly correlated with fibrosis in SSc skin. Importantly, *SOX9-AS1* and *ROCR* were two of the most highly interconnected genes in the SSc skin gene-gene co-expression network, suggesting a key role in mediating the progression of fibrosis. Both are thought to help increase *SOX9* levels, perhaps via non-canonical TGF-β signaling through Wnt/β-catenin (45, 46). The exact mechanism is not well defined. We previously demonstrated that blocking Wnt/β-catenin signaling with C-82 restored subdermal adipogenesis in patients with SSc (47). However, since these are non-coding RNAs, it is currently unclear whether their effect is primarily through a role in regulating *SOX9* or through an alternative mechanism unrelated to *SOX9*. In particular, non-coding RNAs also act as linkers between DNA and protein. One such example is the *HOTAIR* non-coding RNA that mediates repression of some *HOX* loci by recruiting complexes to repress chromatin in those regions (48, 49). Further dissection of the role and mechanism of action of *SPAG17* in profibrotic signaling and a better understanding of the roles of *ROCR* and *SOX9-AS1* in the skin are intriguing areas for future research.

MATERIALS & METHODS

Clinical assessment

We completed standardized evaluations to establish SSc diagnosis, as well as presence/severity of organ involvement, as previously described (50, 51). We determined modified Rodnan Skin Score and local (forearm) skin scores at each visit. The high-resolution CT of the chest, echocardiography, and pulmonary function testing was performed as standard of care (52). At each research visit, we collected PBMCs and two skin punch biopsies (stabilized in RNAlater).

Genomics methods and analysis – abbreviated

Due to article word limits, detailed methods are in the supplement. Briefly, we used RNA extracted from the blood and PBMC samples (miRNeasy minikit #217004) to create created stranded, ribosomal depletion libraries (Takara/Clontech #634876). We trimmed raw sequencing data with cutadapt, aligned to the human genome with RNASTAR, and counted read-pairs per gene with featureCounts (53-55). We then calculated differential expression with DESeq2 and pathway enrichment with gProfileR, and performed network analyses with WGCNA (56-58).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This study was supported by the Rheumatic Disease Core Center (NIAMS P30-AR048335; EDOR, LC, JPA) and the Rheumatic Diseases Research Resource-based Center (NIAMS P30-AR073752; EDOR) at Washington University. JV was supported by the Northwestern University Clinical and Translational Sciences Institute (UL1-TR000150). JPA and EDOR were supported by the Washington University in St. Louis Institute of Clinical and Translational Sciences [ICTS] (UL1-TR000448). DJM-H was supported by an NIH training grant (T32-AR007279-36). Some of the analysis was performed with support from the Washington University Center for High-Performance Computing (S10-OD018091). Sequencing data were generated at the Genome Technology Access Center (GTAC@MGI) at Washington University School. GTAC was partially supported by the Siteman Cancer Center (P30-CA91842) and the ICTS (UL1-TR000448). The research presented herein represents the views of the authors and does not necessarily reflect the views of the NIH.

DATA AVAILABILITY

Data used for analysis, such as gene counts used for DESeq2 and demographic information per sample, are available from FigShare. https://figshare.com/projects/2021_Roberson_lab_systemic_sclerosis_transcriptome_data/118698

Code used for data analysis is available as a repository on GitHub. https://github.com/RobersonLab/2021_ssc_rnaseq

These were prospectively collected samples for controlled data access. FASTQ files will be available for general research use in dbGAP accession phs002902.v1.p1.

References

1. van den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League against Rheumatism collaborative initiative. Arthritis and rheumatism. 2013;65(11):2737–47. [PubMed: 24122180]
2. Allanore Y, Simms R, Distler O, Trojanowska M, Pope J, Denton CP, et al. Systemic sclerosis. Nature reviews Disease primers. 2015;1:15002.
3. Varga J, Roberson ED. Genomic advances in systemic sclerosis: It's time for precision. Arthritis & rheumatology. 2015;67(11):2801–5. [PubMed: 26239971]
4. Derrett-Smith EC, Martyanov V, Chighizola CB, Moinzadeh P, Campochiaro C, Khan K, et al. Limited cutaneous systemic sclerosis skin demonstrates distinct molecular subsets separated by a cardiovascular development gene expression signature. Arthritis Res Ther. 2017;19(1):156. [PubMed: 28676069]

5. Assassi S, Wang X, Chen G, Goldmuntz E, Keyes-Elstein L, Ying J, et al. Myeloablation followed by autologous stem cell transplantation normalises systemic sclerosis molecular signatures. Ann Rheum Dis. 2019;78(10):1371–8. [PubMed: 31391177]
6. Beretta L, Barturen G, Vigone B, Bellocchi C, Hunzelmann N, De Langhe E, et al. Genome-wide whole blood transcriptome profiling in a large European cohort of systemic sclerosis patients. Annals of the rheumatic diseases. 2020;79(9):1218–26. [PubMed: 32561607]
7. Brkic Z, van Bon L, Cossu M, van Helden-Meeuwsen CG, Vonk MC, Knaapen H, et al. The interferon type I signature is present in systemic sclerosis before overt fibrosis and might contribute to its pathogenesis through high BAFF gene expression and high collagen synthesis. Ann Rheum Dis. 2016;75(8):1567–73. [PubMed: 26371289]
8. Higgs BW, Liu Z, White B, Zhu W, White WI, Morehouse C, et al. Patients with systemic lupus erythematosus, myositis, rheumatoid arthritis and scleroderma share activation of a common type I interferon pathway. Ann Rheum Dis. 2011;70(11):2029–36. [PubMed: 21803750]
9. Pendergrass SA, Lemaire R, Francis IP, Mahoney JM, Lafyatis R, Whitfield ML. Intrinsic gene expression subsets of diffuse cutaneous systemic sclerosis are stable in serial skin biopsies. The Journal of investigative dermatology. 2012;132(5):1363–73. [PubMed: 22318389]
10. Skaug B, Khanna D, Swindell WR, Hinchcliff ME, Frech TM, Steen VD, et al. Global skin gene expression analysis of early diffuse cutaneous systemic sclerosis shows a prominent innate and adaptive inflammatory profile. Ann Rheum Dis. 2020;79(3):379–86. [PubMed: 31767698]
11. Assassi S, Swindell WR, Wu M, Tan FD, Khanna D, Furst DE, et al. Dissecting the heterogeneity of skin gene expression patterns in systemic sclerosis. Arthritis & rheumatology. 2015;67(11):3016–26. [PubMed: 26238292]
12. Apostolidis SA, Stifano G, Tabib T, Rice LM, Morse CM, Kahaleh B, et al. Single Cell RNA Sequencing Identifies HSPG2 and APLNR as Markers of Endothelial Cell Injury in Systemic Sclerosis Skin. Front Immunol. 2018;9:2191. [PubMed: 30327649]
13. Karimizadeh E, Sharifi-Zarchi A, Nikaein H, Salehi S, Salamatian B, Elmi N, et al. Analysis of gene expression profiles and protein-protein interaction networks in multiple tissues of systemic sclerosis. BMC Med Genomics. 2019;12(1):199. [PubMed: 31881890]
14. Fingerlin TE, Murphy E, Zhang W, Peljto AL, Brown KK, Steele MP, et al. Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. Nat Genet. 2013;45(6):613–20. [PubMed: 23583980]
15. Zhang Y, Wang S, Wang C, Xiao J, Zhang S, Zhou H. High expression of FAM13A was associated with increasing the liver cirrhosis risk. Mol Genet Genomic Med. 2019;7(3):e543. [PubMed: 30604588]
16. Krausgruber T, Blazek K, Smallie T, Alzabin S, Lockstone H, Sahgal N, et al. IRF5 promotes inflammatory macrophage polarization and TH1-TH17 responses. Nat Immunol. 2011;12(3):231–8. [PubMed: 21240265]
17. Duan H, Fleming J, Pritchard DK, Amon LM, Xue J, Arnett HA, et al. Combined analysis of monocyte and lymphocyte messenger RNA expression with serum protein profiles in patients with scleroderma. Arthritis and rheumatism. 2008;58(5):1465–74. [PubMed: 18438864]
18. Assassi S, Mayes MD, Arnett FC, Gourh P, Agarwal SK, McNearney TA, et al. Systemic sclerosis and lupus: points in an interferon-mediated continuum. Arthritis and rheumatism. 2010;62(2):589–98. [PubMed: 20112391]
19. Kazarian E, Son H, Sapao P, Li W, Zhang Z, Strauss JF, et al. *SPAG17* Is Required for Male Germ Cell Differentiation and Fertility. Int J Mol Sci. 2018;19(4):1252. [PubMed: 29690537]
20. Teves ME, Sundaresan G, Cohen DJ, Hyzy SL, Kajan I, Maczis M, et al. *Spag17* deficiency results in skeletal malformations and bone abnormalities. PloS one. 2015;10(5):e0125936. [PubMed: 26017218]
21. Bahar Halpern K, Massalha H, Zwick RK, Moor AE, Castillo-Azofeifa D, Rozenberg M, et al. Lgr5+ telocytes are a signaling source at the intestinal villus tip. Nat Commun. 2020;11(1):1936. [PubMed: 32321913]
22. Manetti M, Rosa I, Messerini L, Guiducci S, Matucci-Cerinic M, Ibba-Manneschi L. A loss of telocytes accompanies fibrosis of multiple organs in systemic sclerosis. J Cell Mol Med. 2014;18(2):253–62. [PubMed: 24467430]

23. Manetti M, Guiducci S, Ruffo M, Rosa I, Faussone-Pellegrini MS, Matucci-Cerinic M, et al. Evidence for progressive reduction and loss of telocytes in the dermal cellular network of systemic sclerosis. J Cell Mol Med. 2013;17(4):482–96. [PubMed: 23444845]
24. Scambi C, Ugolini S, Jokiranta TS, De Franceschi L, Bortolami O, La Verde V, et al. The local complement activation on vascular bed of patients with systemic sclerosis: a hypothesis-generating study. PloS one. 2015;10(2):e0114856. [PubMed: 25658605]
25. Farina G, Lafyatis D, Lemaire R, Lafyatis R. A four-gene biomarker predicts skin disease in patients with diffuse cutaneous systemic sclerosis. Arthritis and rheumatism. 2010;62(2):580–8. [PubMed: 20112379]
26. Christmann RB, Sampaio-Barros P, Stifano G, Borges CL, de Carvalho CR, Kairalla R, et al. Association of Interferon- and transforming growth factor beta-regulated genes and macrophage activation with systemic sclerosis-related progressive lung fibrosis. Arthritis & rheumatology. 2014;66(3):714–25. [PubMed: 24574232]
27. Rice LM, Ziemek J, Stratton EA, McLaughlin SR, Padilla CM, Mathes AL, et al. A longitudinal biomarker for the extent of skin disease in patients with diffuse cutaneous systemic sclerosis. Arthritis & rheumatology. 2015;67(11):3004–15. [PubMed: 26240058]
28. Kim YH, Lee JR, Hahn MJ. Regulation of inflammatory gene expression in macrophages by epithelial-stromal interaction 1 (Epsti1). Biochem Biophys Res Commun. 2018;496(2):778–83. [PubMed: 29217193]
29. Trombetta AC, Soldano S, Contini P, Tomatis V, Ruaro B, Paolino S, et al. A circulating cell population showing both M1 and M2 monocyte/macrophage surface markers characterizes systemic sclerosis patients with lung involvement. Respir Res. 2018;19(1):186. [PubMed: 30249259]
30. Teves ME, Zhang Z, Costanzo RM, Henderson SC, Corwin FD, Zweit J, et al. Sperm-associated antigen-17 gene is essential for motile cilia function and neonatal survival. Am J Respir Cell Mol Biol. 2013;48(6):765–72. [PubMed: 23418344]
31. Xu X, Sha YW, Mei LB, Ji ZY, Qiu PP, Ji H, et al. A familial study of twins with severe asthenozoospermia identified a homozygous SPAG17 mutation by whole-exome sequencing. Clin Genet. 2018;93(2):345–9. [PubMed: 28548327]
32. Cordova-Fletes C, Becerra-Solano LE, Rangel-Sosa MM, Rivas-Estilla AM, Alberto Galan-Huerta K, Ortiz-Lopez R, et al. Uncommon runs of homozygosity disclose homozygous missense mutations in two ciliopathy-related genes (SPAG17 and WDR35) in a patient with multiple brain and skeletal anomalies. Eur J Med Genet. 2018;61(3):161–7. [PubMed: 29174089]
33. van der Valk RJ, Kreiner-Moller E, Kooijman MN, Guxens M, Stergiakouli E, Saaf A, et al. A novel common variant in DCST2 is associated with length in early life and height in adulthood. Hum Mol Genet. 2015;24(4):1155–68. [PubMed: 25281659]
34. Kim JJ, Lee HI, Park T, Kim K, Lee JE, Cho NH, et al. Identification of 15 loci influencing height in a Korean population. J Hum Genet. 2010;55(1):27–31. [PubMed: 19893584]
35. Abdelhamed Z, Lukacs M, Cindric S, Omran H, Stottmann RW. A novel hypomorphic allele of Spag17 causes primary ciliary dyskinesia phenotypes in mice. Dis Model Mech. 2020;13(10):dmm045344. [PubMed: 32988999]
36. Andjelkovic M, Minic P, Vreca M, Stojiljkovic M, Skakic A, Sovtic A, et al. Genomic profiling supports the diagnosis of primary ciliary dyskinesia and reveals novel candidate genes and genetic variants. PloS one. 2018;13(10):e0205422. [PubMed: 30300419]
37. Rusu MC, Mirancea N, Manoiu VS, Valcu M, Nicolescu MI, Paduraru D. Skin telocytes. Annals of anatomy = Anatomischer Anzeiger : official organ of the Anatomische Gesellschaft. 2012;194(4):359–67. [PubMed: 22226149]
38. Corallo C, Cutolo M, Volpi N, Franci D, Agliano M, Montella A, et al. Histopathological findings in systemic sclerosis-related myopathy: fibrosis and microangiopathy with lack of cellular inflammation. Ther Adv Musculoskelet Dis. 2017;9(1):3–10. [PubMed: 28101144]
39. Okroj M, Johansson M, Saxne T, Blom AM, Hesselstrand R. Analysis of complement biomarkers in systemic sclerosis indicates a distinct pattern in scleroderma renal crisis. Arthritis Res Ther. 2016;18(1):267. [PubMed: 27863511]

40. Perez NA, Morales MLA, Sanchez RS, Salas RMO, Puebla RAF, Hernandez ME. Endothelial lesion and complement activation in patients with Scleroderma Renal Crisis. J Bras Nefrol. 2019;41(4):580–4. [PubMed: 30806445]
41. Noris M, Caprioli J, Bresin E, Mossali C, Pianetti G, Gamba S, et al. Relative role of genetic complement abnormalities in sporadic and familial aHUS and their impact on clinical phenotype. Clin J Am Soc Nephrol. 2010;5(10):1844–59. [PubMed: 20595690]
42. Cofiell R, Kukreja A, Bedard K, Yan Y, Mickle AP, Ogawa M, et al. Eculizumab reduces complement activation, inflammation, endothelial damage, thrombosis, and renal injury markers in aHUS. Blood. 2015;125(21):3253–62. [PubMed: 25833956]
43. Devresse A, Aydin S, Le Quintrec M, Demoulin N, Stordeur P, Lambert C, et al. Complement activation and effect of eculizumab in scleroderma renal crisis. Medicine (Baltimore). 2016;95(30):e4459. [PubMed: 27472742]
44. Uriarte MH, Larrarte C, Rey LB. Scleroderma Renal Crisis Debute with Thrombotic Microangiopathy: A Successful Case Treated with Eculizumab. Case Rep Nephrol. 2018;2018:6051083. [PubMed: 30425869]
45. Zhang W, Wu Y, Hou B, Wang Y, Deng D, Fu Z, et al. A SOX9-AS1/miR-5590-3p/SOX9 positive feedback loop drives tumor growth and metastasis in hepatocellular carcinoma through the Wnt/beta-catenin pathway. Mol Oncol. 2019;13(10):2194–210. [PubMed: 31402556]
46. Barter MJ, Gomez R, Hyatt S, Cheung K, Skelton AJ, Xu Y, et al. The long non-coding RNA ROCR contributes to SOX9 expression and chondrogenic differentiation of human mesenchymal stem cells. Development. 2017;144(24):4510–21. [PubMed: 29084806]
47. Lafyatis R, Mantero JC, Gordon J, Kishore N, Carns M, Dittrich H, et al. Inhibition of beta-Catenin Signaling in the Skin Rescues Cutaneous Adipogenesis in Systemic Sclerosis: A Randomized, Double-Blind, Placebo-Controlled Trial of C-82. The Journal of investigative dermatology. 2017;137(12):2473–83. [PubMed: 28807667]
48. Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, et al. Long noncoding RNA as modular scaffold of histone modification complexes. Science. 2010;329(5992):689–93. [PubMed: 20616235]
49. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell. 2007;129(7):1311–23. [PubMed: 17604720]
50. Hinchcliff M, Huang CC, Wood TA, Matthew Mahoney J, Martyanov V, Bhattacharyya S, et al. Molecular signatures in skin associated with clinical improvement during mycophenolate treatment in systemic sclerosis. The Journal of investigative dermatology. 2013;133(8):1979–89. [PubMed: 23677167]
51. Johnson ME, Mahoney JM, Taroni J, Sargent JL, Marmarelis E, Wu MR, et al. Experimentally-derived fibroblast gene signatures identify molecular pathways associated with distinct subsets of systemic sclerosis patients in three independent cohorts. PloS one. 2015;10(1):e0114017. [PubMed: 25607805]
52. Richardson C, Agrawal R, Lee J, Almagor O, Nelson R, Varga J, et al. Esophageal dilatation and interstitial lung disease in systemic sclerosis: A cross-sectional study. Semin Arthritis Rheum. 2016;46(1):109–14. [PubMed: 27033049]
53. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet journal. 2011;17(1):10–2.
54. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15–21. [PubMed: 23104886]
55. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923–30. [PubMed: 24227677]
56. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559. [PubMed: 19114008]
57. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550. [PubMed: 25516281]

58. Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Research. 2019;47(W1):W191–W8. [PubMed: 31066453]

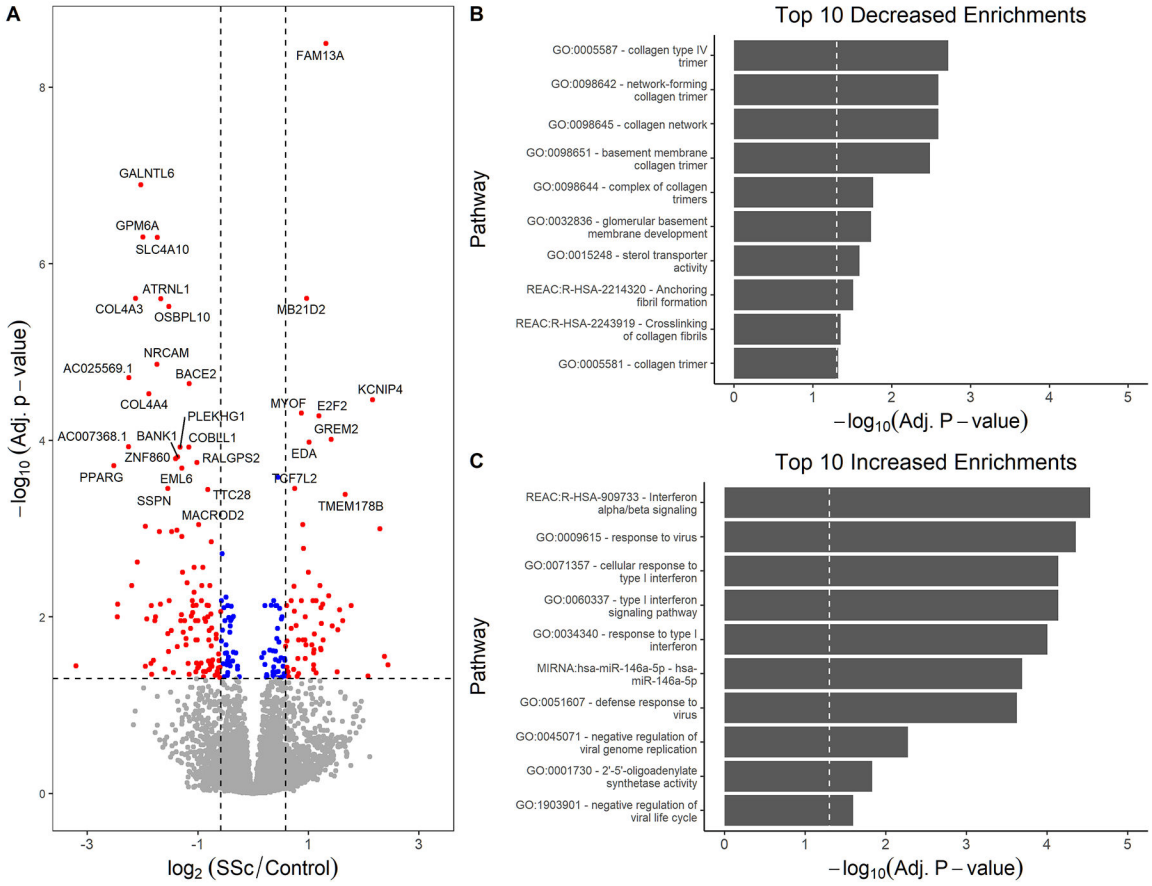

Fig. 1 –. Systemic sclerosis PBMCs have strong enrichments for increased type 1 interferon signaling

A. Standard volcano plot. The x-axis shows the effect size on a log2 scale of systemic sclerosis over control. The y-axis is the significance of the difference (−log10 scale). The vertical dashed lines show the 1.5-fold cutoff. The horizontal line is the 0.05 adjusted significance threshold. Genes meeting a 1.5-fold cutoff and transcriptome-wide significance are in red. Significant points that don't meet the fold-change cutoff are blue. Genes that weren't significantly different are in grey. **B.** Shows the enriched pathways for genes decreased in SSc PBMCs. The x-axis shows the −log10 of the adjusted p-values. The dashed vertical line shows the 0.05 adjusted p-value cutoff. The y-axis shows the pathway. The most significant enrichments include collagen and sterol transporter. **C.** The axes are the same as in **B**, but the pathways tested were for genes at least 1.5-fold increased in SSc PBMCs. The predominant signal is increased type 1 interferon signaling.

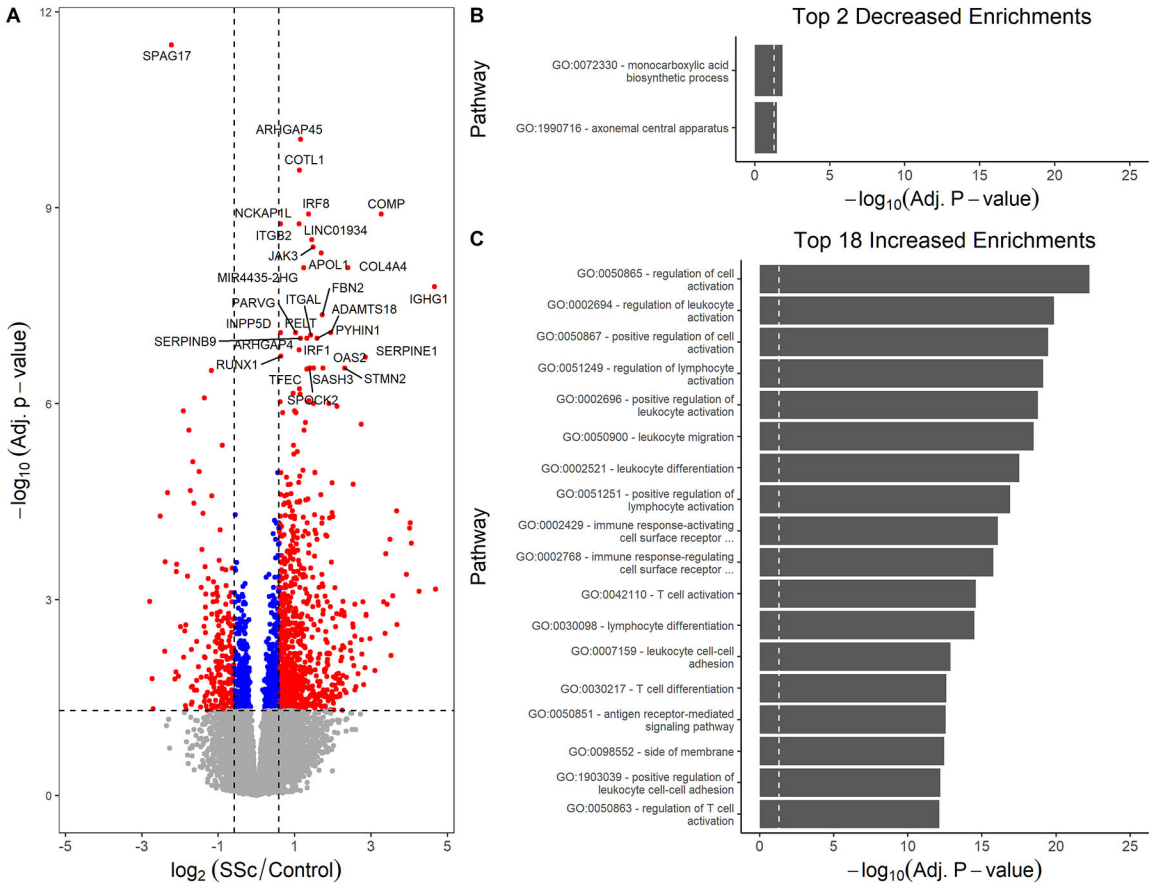

Fig. 2 –. SSc skin has decreased primary cilia protein *SPAG17* and increased immune activation

A. Volcano plot showing the significance and effect-sizes of gene expression in systemic sclerosis versus control skin. The log2 fold-change is on the x-axis, and the −log10 adjusted significance is on the y-axis. The vertical lines indicated 1.5-fold up and down cutoffs. The horizontal line shows the 0.05 adjusted significance threshold. The most significant difference was a decrease in *SPAG17* in SSc skin. More genes were significantly increased in SSc than significantly decreased. Both **B.** and **C.** show the gProfilerR pathway enrichments. The x-axis shows the −log10 of the significance for each enrichment. The y-axis shows the name of the enriched pathway. **B.** There were few significant enrichments for decreased genes, including the ciliary axoneme. **C.** The most significant enrichments for genes increased in SSc were related to immune cell activation, indicating the migration of immune cells into the skin.

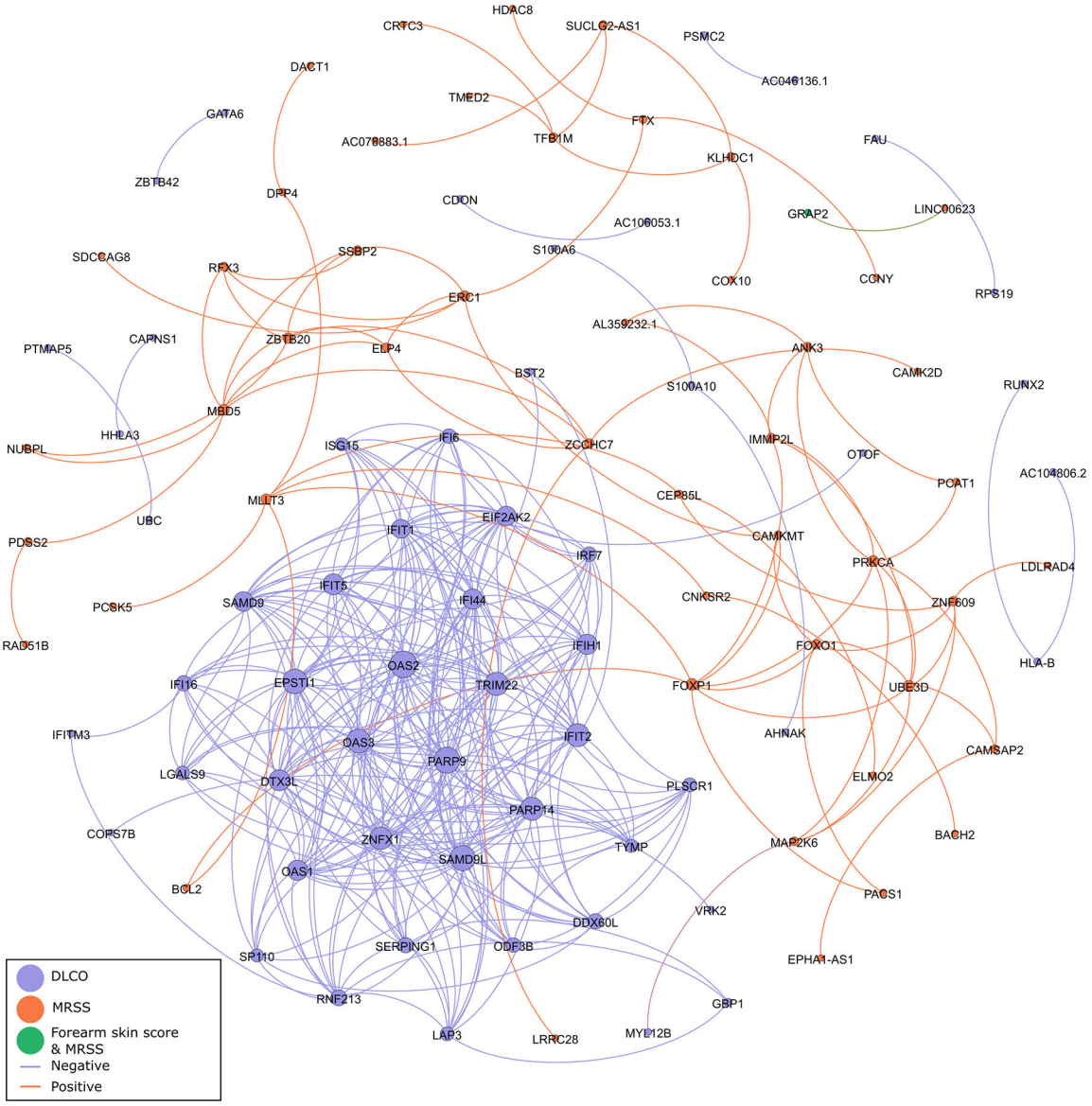

Fig. 3 –. Interferon-responsive genes are the most highly connected trait-correlated genes in SSc PBMCs

Shown is a network diagram for genes correlated with at least one clinical trait. Each node is an individual gene, sized by weighted degree and filled by the trait association. Each edge is colored for whether the gene-gene correlation is positive or negative, with a minimum cutoff of 0.80. The interferon-responsive genes, such as *OAS1, OAS2*, and *OAS3*, and *IFIT1/2*, were the most highly interconnected for PBMCs.

Abbreviations: PBMC, peripheral blood mononuclear cells; DLCO, diffusion capacity of the lungs for carbon monoxide; MRSS, modified Rodnan skin score.

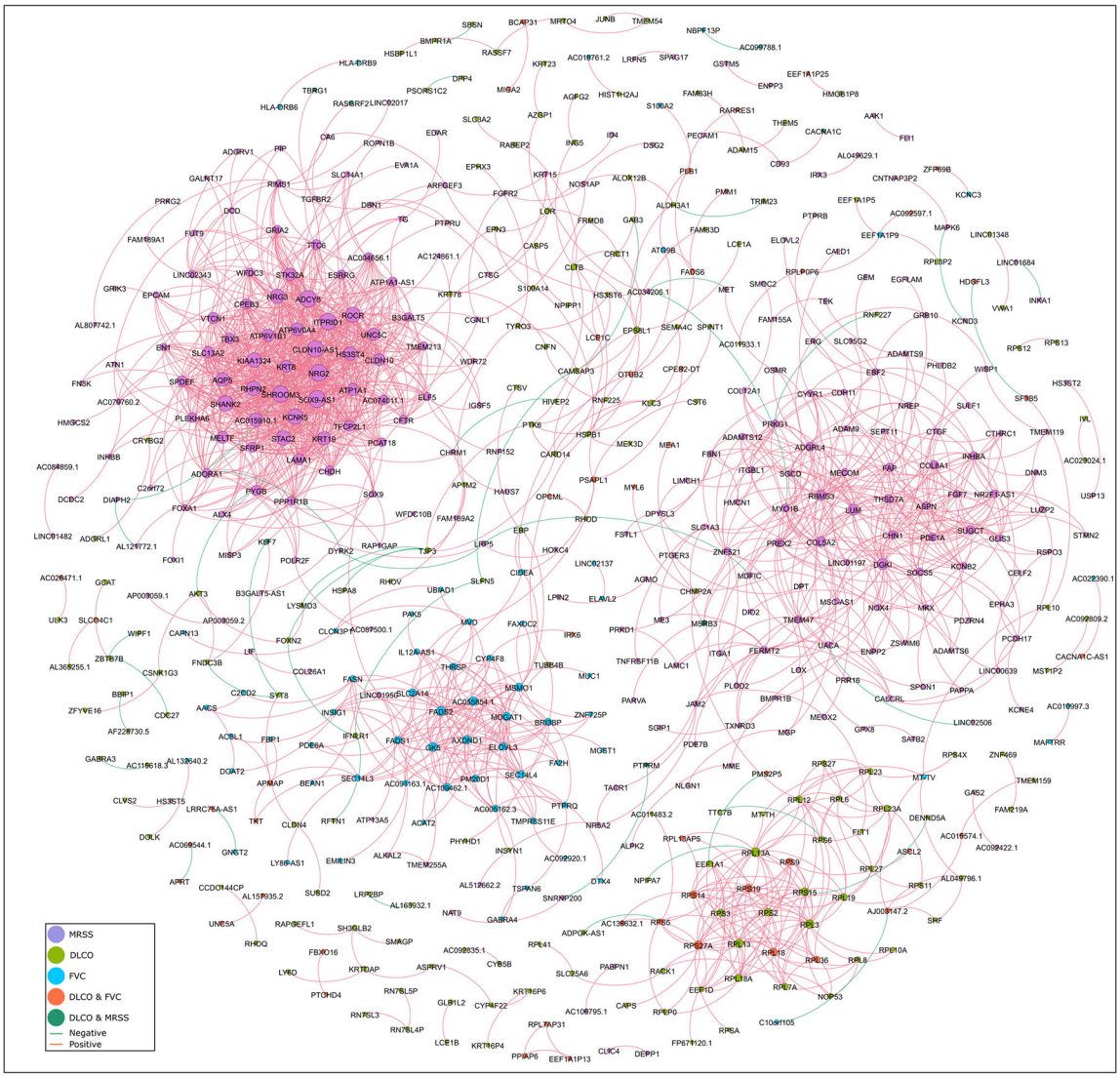

Fig. 4 –. *SOX9* locus genes are among the most highly connected trait-associated genes in SSc skin

Shown is a network diagram for genes correlated with at least one clinical trait in the skin. Each node is an individual gene, sized by weighted degree and colored by the trait association. The edges between nodes are colored for whether the gene-gene correlation is positive or negative, with a minimum cutoff of 0.80. Some genes correlated with fibrosis formed a relatively separate sub-network from those associated with lung function or both DLCO and fibrosis. The most connected genes in this sub-network included *SOX9-AS1* and *ROCR*, which are both non-coding and located at the *SOX9* locus. The overlap of correlations between fibrosis and lung function were primarily ribosomal proteins. A separate sub-network of MRSS-correlated genes was enriched for matrix proteins, such as *COL5A2* and *COL8A1*.

Abbreviations: DLCO, diffusion capacity of the lungs for carbon monoxide; FVC, forced vital capacity; MRSS, modified Rodnan skin score.

Table 1 –

Case/control demographics

Basic demographic information for our study cohort. All values are “n (%)” unless otherwise indicated. The controls were individuals without a self-reported history of autoimmune disease. SSc met ACR criteria for a definitive diagnosis. The controls were significantly younger than the SSc cases (two-sided student t-test) but balanced for sex (Fisher’s exact test). The control and SSc cohorts were predominantly white, and the population background was significantly different between the two (Fisher’s exact test).

	Control (n=14)	SSc (n=19)
Age mean (SD)	32.6 (10.8)	49.7 (11.4)
Female	9 (64.3)	15 (78.9)
Ethnicity		
Asian	0 (0.0)	2 (10.5)
Black	3 (21.4)	2 (10.5)
Hispanic	5 (35.7)	0 (0.0)
White	6 (42.9)	15 (78.9)

Abbreviations: SSc, systemic sclerosis; SD, standard deviation.

Table 2 –

Systemic sclerosis subtype demographics

Shown are the basic demographic and antibody staining pattern information for the two main systemic sclerosis sub-groups: diffuse disease and limited disease. Values in the table are represented as “mean (standard deviation)” unless otherwise indicated as “n (%)”. The MRSS, body mass index, forced vital capacity, diffusion capacity for CO, and total lung capacity were determined at repeat visits. For each individual, we used the “worst” observed value as the representative value.

	dcSSc (n=14)	lcSSc (n=5)
Age mean	49.7 (11.7)	49.8 (11.9)
Female (%)	13 (92.9)	2 (40.0)
Mean duration, months	42.7 (33.2)	36.2 (22.8)
MRSS	24.1 (10.2)	8.8 (5.0)
Forearm MRSS	1.9 (0.9)	0.8 (0.8)
BMI	26.7 (5.5)	29.5 (6.3)
FVC *	75.1 (15.6)	82.6 (12.5)
DLCO corrected *	66.2 (20.8)	80.0 (17.4)
TLC *	85.6 (17.2)	85.8 (11.7)
Ethnicity (%)		
Asian	1 (7.1)	1 (20.0)
Black	2 (14.3)	0 (0.0)
Hispanic	0 (0.0)	0 (0.0)
White	11 (78.6)	4 (80.0)
Immunofluorescence † **(%)**		
Centromere	1 (8.3)	2 (40.0)
Homogeneous	3 (25.0)	0 (0.0)
Nucleolar	4 (33.3)	1 (20.0)
Speckled	7 (58.3)	3 (60.0)

Abbreviations: MRSS, Modified Rodnan Skin Score; BMI, body mass index; FVC, forced vital capacity; DLCO, diffusion capacity for carbon monoxide; TLC, total lung capacity.

*These lung function parameters were calculated as percent estimated maximum for age and sex.

†Immunofluorescence data was only available for 12/14 dcSSc individuals. Percents were calculated based on this availability.