Biomarker discovery in systemic sclerosis: state of the art

Francesco Bonella1
Giuseppe Patuzzo2
Claudio Lunardi2

1Interstitial and Rare Lung Disease Unit, Ruhrlandklinik University Hospital, University of Duisburg-Essen, Essen, Germany; 2Department of Medicine, University of Verona, Verona, Italy

Abstract: Systemic sclerosis (SSc) is an autoimmune disease characterized by immune dysfunction and by abnormalities of the microvasculature with vascular obliteration, eventually leading to fibrosis of the skin, gastrointestinal tract, lungs, heart, and kidney. The etiology and pathogenesis of SSc remain unclear, despite recent significant progress in the field. Immune activation and microangiopathy are followed by widespread organ fibrosis, leading to organ failure and increased mortality. The production of inflammatory cytokines and growth factors after tissue injury, as well as the presence of circulating autoantibodies, provide a source of biomarkers with potential diagnostic and prognostic applications in the clinical routine. Two principal approaches exist to discover and characterize biomarkers. The proof-of-concept approach verifies the ability of known proteins, generally involved in the pathogenesis, to correlate with disease phenotype and outcome. A proteomic approach does not need prior knowledge of the proteins or of their function, but it requires high-performance and time-consuming techniques. In this review, we highlight the most recent findings in biomarkers used to characterize SSc organ involvement, to stratify the patients, and to assess the response to treatment.

Keywords: systemic sclerosis, biomarkers, proteomics, gene expression profiling

Introduction

Systemic sclerosis (SSc) is a rare and complex autoimmune disease that involves the connective tissue of skin and internal organs, with remarkable heterogeneity in the clinical features and course of the disease, resulting in high morbidity and mortality. The disease is characterized by severe and diffuse endothelial cell damage and by an overproduction and accumulation of collagen and other extracellular matrix (ECM) proteins, resulting in the thickening of the skin and fibrosis of the affected organs. The pathogenetic mechanisms involve three interactive components represented by vascular dysfunction and injury, immune system activation, and increased secretion of collagen by fibroblasts. The prevalence of SSc is between 50 and 300 cases per one million inhabitants, and an incidence ranging from 2.3 to 22.8 cases per million inhabitants per year in the general population.

The complexity of the underlying pathogenetic mechanisms suggests that disease prediction may depend on a combinatorial analysis of many mediators rather than single molecules. As a sign of the times, the new classification criteria for SSc by the American College of Rheumatology/European League Against Rheumatism incorporate autoantibodies (Abs), specifically anti-Scl70, anti-RNA polymerase (RNA-pol) III, and anticentromere (ACA), as supportive tools for the classification of SSc. This is a milestone in the evolution of thinking regarding the role of serology.
in the clinical management of SSc. Multicohort studies have reported variable antibody frequency and distribution based on race, but the clinical associations remain very similar.\textsuperscript{5-10} It is worthwhile to point out that in the last 2 years, two new Abs have been identified through mass spectrometry (MS) and included in the serology panel of SSc.\textsuperscript{11,12}

In a near future perspective, the application of new technologies like gene expression profiling and proteomic wide analysis will offer the opportunity to improve discrimination of SSc phenotypes and personalize treatment. Particularly promising are the proteomic studies, because proteome results from the interaction between genetic background and environmental factors, and it may be considered the signature of a disease, avoiding the limitations of RNA transcriptional and post-transcriptional changes.\textsuperscript{13}

Several attempts have been made in the last years to identify biomarkers able to predict disease outcome and response to treatment, but the majority of these studies originate from single centers and need further validation in large cohorts.

### Pathogenesis of SSc

The majority of biomarkers investigated in SSc are not only epiphenomena, but molecules involved in the pathogenesis of the disease. Therefore, an introduction to the pathogenetic mechanisms underlying SSc is necessary.

As for other autoimmune diseases, genetic and environmental factors play a pivotal role in the pathogenesis of the disease. Among environmental factors, infectious agents may be involved in the pathogenesis of the disease with different mechanisms. The molecular mimicry mechanism (ie, the similarity between autoantigens and proteins of viral origin) has been the most widely studied.\textsuperscript{14} It is worth mentioning that there is a certain structural similarity between retroviral proteins and topoisomerase I, the antigen recognized by anti-Sc1-70 Abs. Moreover, human cytomegalovirus may contribute to the onset of SSc through its ability to infect endothelial cells and to induce cell activation and apoptosis. Patients affected by SSc present with antibodies directed against the human cytomegalovirus-derived protein UL94; such antibodies are able to induce apoptosis of human endothelial cells and to stimulate fibroblast proliferation – two typical aspects of SSc.\textsuperscript{14-17}

Endothelial cell damage is considered the primary event in the pathogenesis of SSc, leading to the increased expression of adhesion molecules; recruitment and activation of leucocytes; thickening of the basal lamina; proliferation of the pericytes, fibroblasts, and smooth muscle cells; platelet adhesion and activation; and formation of the perivascular infiltrate.\textsuperscript{18,19} Perivascular infiltrate is initially characterized by myeloid dendritic cells, by CD4+ T-lymphocytes, and by mastocytes, and subsequently by polymorphonucleates, B-lymphocytes, and T-lymphocytes.\textsuperscript{18,20-23} Growth factors, oxidative stress, and cytokines stimulate transdifferentiation of the fibroblasts into myofibroblasts within the vessel wall, with consequent thickening of the vessel wall itself and further reduction of the regional blood flow by narrowing the vessel lumen. This is exacerbated by activation of platelets and of the coagulation system, with the formation of microthrombi. Chronic perivascular inflammation, in advanced sclerodermic lesions, is replaced by fibrosis, which contributes to the permanent alteration of the vessel architecture, with thickening of the vessel wall, and progressive narrowing of the lumen. The composition of the ECM varies according to the stage of the disease: proteoglycans and elastic fibers such as fibrillin are typical of the first phases of the fibrotic process; while type I collagen is found in the advanced phases.

Irrespective of the progressive loss of capillaries and of the high plasma levels of vascular endothelial growth factor (VEGF) in response to tissue hypoxia, a deficient angiogenesis is observed.\textsuperscript{19} Many consequences of endothelial damage are related to an imbalance in vasoactive factors, with hyperproduction of endothelin (ET)-1 and hypoproduction of nitric oxide (NO) and prostacyclin. An imbalance of these mediators has an effect not only on vascular tone, but also on inflammation, platelet activation, myofibroblastic induction, and fibrosis.\textsuperscript{24,25}

### Tools and techniques used in biomarker discovery

In Table 1, the principal tools and techniques used in biomarker discovery are shown. For the sake of brevity, we will describe only some of the tools in detail.

### Gene expression and candidate gene studies

Gene expression profiling has become an increasingly applied tool, not only to reveal pathogenetic aspects of disease, but also to better characterize SSc subsets.\textsuperscript{26} In one of the first studies in this sense, Milano et al\textsuperscript{27} analyzed the genome-wide patterns of gene expression with DNA microarrays in skin biopsies from distinct scleroderma subsets (17 patients with diffuse SSc [dSSc], seven patients with limited cutaneous SSc [iSSc], three patients with morphea, and six healthy controls). Analysis by hierarchical clustering demonstrates nearly identical patterns of gene expression in 17 out of 22 of the forearm and back skin pairs of SSc patients. Distinct
patterns of gene expression separate patients with dSSc from those with lSSc without overlapping with healthy controls. Moreover, a 177-gene signature was associated with the severity of skin disease in dSSc. In a further study, the same group reported a transforming growth factor (TGF)-β-responsive gene signature isolated from primary dermal fibroblasts, which was associated with increased disease severity in dSSc patients. Comparative studies with classical serum biomarkers are needed to confirm the value of gene signatures in SSc.

With regard to SSc-interstitial lung disease (ILD), the gene expression profiles of SSc fibrotic lung seem to differ significantly from idiopathic pulmonary fibrosis (IPF). Seibold et al found a strong correlation of polymorphism in the mucin 5B (MUC5B) promoter region with pulmonary fibrosis, and Putman et al found a correlation of the minor allele with fibrotic abnormalities in the general population. By contrast, no association was found for MUC5B with SSc-ILD in three different cohorts, suggesting different pathophysiology.

In the skin biopsies of SSc patients from the Genetics versus Environment in Scleroderma Outcome Study (GENISOS), several transcripts including genes involved in extravasation and adhesion of inflammatory cells (the most representative, ELP, CCL2, and MMP3) correlated with the severity of ILD. In the GENISOS cohort, genome-wide association study (GWAS) was performed in samples from more than 5,000 individuals with SSc. A new susceptibility

---

<table>
<thead>
<tr>
<th>Table 1 Principal tools and techniques used in biomarker discovery</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Technologies</strong></td>
</tr>
</tbody>
</table>
| Genomic | • Genome-wide methods  
- Microarrays: whole-genome expression array  
- Serial analysis of gene expression (SAGE)  
- Expressed sequence tags (ESTs) analysis  
- Gene expression profiling based on alternative RNA splicing  
- Tangerine expression profiling  
• Individual sequences  
- Real-time (RT)-polymerase chain reaction (PCR)  
- Competitive RT-PCR  
- RNase protection assay  
- T-cell receptor expression analysis  
• Analysis of single-cell gene expression  
• RNA amplification  
• Monitoring in vivo gene expression by magnetic resonance imaging (MRI) or positron emission tomography (PET) | Gene expression in healthy and diseased tissues, identification of single nucleotide polymorphisms (SNPs), correlation genotype-phenotype |
| Epigenomic | • DNA methylation profiling  
- Methylation at cytosine based sequences (CpG)  
- Quantitative methyl-SNP analysis  
- Histone endonuclease-linked detection of methylation sites of DNA (HELMET)  
• Histone modification profiling acetylation, methylation, phosphorylation, SUMOylation  
• Noncoding (nc)RNA profiling: long ncRNA and micro (mi)RNA – PCR and nanostring | Changes in the regulation of gene activity and expression independent from the DNA sequence |
| Proteomic | • Two-dimensional (2D) gel electrophoresis (GE) and difference gel electrophoresis (DIGE)  
• Isotope-coded affinity tag (ICAT)  
• Mass spectrometry (MS)  
- Matrix-assisted laser desorption/ionization (MALDI)  
- Time-of-flight MALDI (MALDI-TOF)  
- Quantitative tandem MS  
- Imaging MS (IMS)  
- Liquid chromatography–MS  
- Surface-enhanced laser desorption/ionization (SELDI) and SELDI-TOF | Identification of proteins in fluids and tissue, detecting structural and functional variations of proteins (post-translational modifications, turnover, transport, and excretion) |
| Metabolomic | • MS-based kits  
• Urinary profiling by capillary electrophoresis  
• Lipomic technologies | Technology that complements proteomics and transcriptomics, focusing on metabolites |

locus for SSc was demonstrated at CD247 and MHC, and IRF5 and STAT4 were confirmed as genetic risk factors for SSc. In a second study from the GENISOS cohort and two additional validation cohorts, single nucleotide polymorphisms (SNPs) from the three susceptibility loci obtained through GWAS were investigated. A SNP in the IRF5 promoter region (rs4728142), associated with lower IRF5 transcript levels, was predictive of longer survival and milder ILD in patients with SSc. Moreover, the IRF5 rs4728142 minor allele correlated with higher forced vital capacity (FVC)% predicted at enrolment ($P=0.019$).

It is known that ethnicity may significantly impact disease traits. A recent GWAS study in a Han Chinese cohort of 387 SSc patients did not confirm the association between the rs2056626 of CD247 and SSc or any SSc subtypes. Moreover, another study demonstrated an association of gene polymorphisms in RHOB and FAM167A-BLK, which are proteins involved in intracellular protein trafficking, with susceptibility to SSc in a Chinese Han population. Interestingly, the loci of the SNPs in the RHOB region were reported to be different between Asian and Caucasian populations.

The results of all these genetic association studies suggest, once again, the complex nature of genetics of SSc.

New data confirm that the gene expression profiles of SSc-ILD and SSc-pulmonary hypertension (PH) lung tissue are different. Hsu et al. characterized the molecular fingerprint of lung tissue of 33 SSc patients who underwent lung transplantation. Fibrotic lungs were enriched not only for genes for collagen types I and III, IGFBPs, tissue inhibitors of metalloproteases 1 (TIMPs-1), and interferon gamma receptor (IFNGR), which are classically involved in fibrosis, but also for insulin growth factor (IGF). Conversely, SSc-PH lungs showed specific genes including chemokines (CCL-2, CXCL10) and HLA (−DR and −A) genes.

Gene expression profiling of peripheral blood mononuclear cells (PBMCs) is a powerful and less invasive tool for the identification of biomarkers involved in disease processes. Comparing the transcript profiles of PBMCs from 30 patients with idiopathic PH, 19 SSc without PH, and 42 with SSc-PH, Cheadle et al. found that a specific signature for erythrocyte maturation was enriched in patients with PH, suggesting a reaction to the increasing hypoxemia in this condition. A similar study on PBMC gene expression in 36 SSc patients reported the increased expression of nine genes ($ICAM1$, $IFNGR1$, $IL1B$, $IL13Rα1$, $JAK2$, $AIF1$, $CCR1$, $ALAS2$, and $TIMP2$) in ISSc-PH patients.

Markers of vascular injury (VCAM-1, VEGF, and von Willebrand factor [vWF]) were also found to be elevated in the serum of SSc-PH patients.

**Proteomic approach**

Multiplex protein technology has the potential to identify biomarkers for the differentiation, classification, and improved understanding of the pathogenesis of complex diseases. In comparison to the proof-of-concept approach, protein technology is time consuming and more expensive. This is probably one of the reasons why proteome-wide studies in SSc are limited.

Proteome-wide analysis through protein chip array (generally, surface-enhanced laser desorption/ionization [SELDI-TOF]) followed by MS for the characterization of the peaks allows for the screening of fluids, tissues, and cells for a large number of putative transcripts at the same time. In the study cited earlier on plasma S100A8/A9 as a biomarker for SSc, the authors efficaciously used SELDI-TOF-MS to identify proteins differentially expressed in SSc patients and healthy controls. Similarly, the discovery and validation of serum CXCL4 as a biomarker for SSc was reached through a complex proteome-wide analysis.

**microRNA expression profiling**

Epigenetic mechanisms, including alterations in DNA methylation, histone code modifications, and changes in microRNA (miRNA) expression levels, can be defined as the “missing bridge” in understanding the pathogenesis of SSc. miRNAs are highly stable small noncoding RNAs, which function in RNA silencing and post-transcriptional regulation of gene expression. Since they can be detected in biologic fluids and tissue, and given that they are resistant to variations in sample handling and temperature, miRNAs represent potential biomarkers. Recent reports have described the differential expression of numerous regulatory miRNAs in SSc mainly in skin fibroblasts, several of which are important for TGF-β pathways and for downstream signaling cascades (Table 2).

Among the investigated miRNAs, some of them are likely to correlate with disease subsets and severity. For example, SSc patients with lower serum miR-150 levels have more severe clinical manifestations, like digital ulcers (DUs) and PH, but not ILD. Similarly, patients with lower serum miR-196a levels show a significantly higher modified Rodnan Skin Score (mRSS) and a higher prevalence of pitting scars than patients with higher miR-196a levels.
Table 2 miRNAs investigated in SSc, their potential function, and molecular targets

<table>
<thead>
<tr>
<th>miRNA</th>
<th>Regulation</th>
<th>Tissue</th>
<th>Function</th>
<th>Target</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-21</td>
<td>Up</td>
<td>Skin FBs, murine dermal FBs</td>
<td>Profibrotic</td>
<td>SMAD7</td>
<td>52</td>
</tr>
<tr>
<td>miR-29</td>
<td>Down</td>
<td>FBs, murine dermal FBs</td>
<td>Antifibrotic</td>
<td>COL1A1, COL1A2, COL3A1, FBN1</td>
<td>52,56</td>
</tr>
<tr>
<td>miR-31</td>
<td>Up</td>
<td>Skin, FBs</td>
<td>Unclear</td>
<td>COL1A1, COL1A2</td>
<td>52</td>
</tr>
<tr>
<td>miR-129-5p</td>
<td>Down</td>
<td>FBs</td>
<td>Unclear</td>
<td>COL1A1, COL1A2</td>
<td>232</td>
</tr>
<tr>
<td>miR-145</td>
<td>Down</td>
<td>Skin, FBs</td>
<td>Unclear</td>
<td>SMAD3</td>
<td>52</td>
</tr>
<tr>
<td>miR-146</td>
<td>Up</td>
<td>Skin, FBs</td>
<td>Unclear</td>
<td>SMAD4</td>
<td>52</td>
</tr>
<tr>
<td>miR-503</td>
<td>Up</td>
<td>Skin, FBs</td>
<td>Unclear</td>
<td>SMAD7</td>
<td>52</td>
</tr>
</tbody>
</table>

Circulating miRNAs

| miR-7 | Up in dcSSc, down in lcSSc | FBs, skin, serum | Antifibrotic | COL1A1, COL1A2 | 233,234 |
| miR-let-7a | Down | FBs, serum | Unclear | COL1A1, COL1A2 | 235 |
| miR-29a | Down | Hair shaft, serum | Antifibrotic | COL1A1, COL1A2 | 236 |
| miR-92-a | Up | FBs, serum | Unclear | MMP-1 | 237 |
| miR-142 | Up | Serum | Unclear | Integrin αV | 238 |
| miR-150 | Down | FBs, serum | Profibrotic | Integrin β3 | 53 |
| miR-196a | Down | FBs, serum, and hair shafts | Antifibrotic | COL1A1, COL1A2 | 239,240 |

Note: Human cells in origin, unless otherwise specified.

Abbreviations: miRNA, microRNA; SSc, systemic sclerosis; FB, fibroblast; SMAD, intracellular proteins that transduce extracellular signals from transforming growth factor-β ligands; COL1, type 1 collagen; dcSSc, diffuse systemic sclerosis; lcSSc, limited cutaneous systemic sclerosis; MMP-1, matrix metalloproteinase 1.

MiR-29 has been extensively studied, and its antifibrotic effects in several major fibrotic disorders, including SSc, are well known.55 The downregulation of miR-29 is correlated with the upregulation of COL1A1 messenger (m) RNA.55 Maurer et al56 identified miR-29 as a key regulator of collagen expression in SSc. Interestingly, SSc patients presenting with lower miR-29a levels had significantly higher right ventricular systolic pressure, as estimated by Doppler echocardiography, than those with normal miR-29a levels, suggesting a role in the pathogenesis of PH.57 The identification of specific miRNAs as markers of disease subsets and severity could lead, in the future, to the development of disease- and stage-specific targeted therapies.

Biomarkers for diagnosis, prognosis, and treatment response

Clinical manifestations and diagnosis of SSc

Raynaud’s phenomenon is usually the first manifestation of SSc and it may precede the development of other clinical features by months or years, depending on the subset of the disease (ie, the dSSc or lcSSc form, respectively). From a clinical point of view, Raynaud’s phenomenon occurs when the delicate balance between vasodilation and vasoconstriction is disturbed, in favor of vasoconstriction.58 Raynaud’s phenomenon is an exaggerated physiological response of the extremities to exposure to cold or emotional stress. Its classical presentation is characterized by the fingers turning white (ischemia), then blue (cyanosis) and, finally, red (reperfusion). Antinuclear antibodies are usually present at the time of Raynaud’s phenomenon onset. Indeed, the finding of a positive antinuclear antibody in a patient with Raynaud’s phenomenon suggests the need for further screening in order to identify SSc at an early stage.

Chronic ischemia leads to a reduction of the finger pad substance with consequent loss of small areas of tissue. DUs and digital gangrene are caused by a severe degree of ischemia. Ulcers that spontaneously occur on the fingertips are due exclusively to ischemia, whereas those over the extensor surface of the proximal interphalangeal, metacarpophalangeal, ulnar styloid, and elbow joints are due to a combination of poor perfusion in areas of stretched skin or in areas of repeated minor trauma.59

The fibrosing process results in thickened skin; gastrointestinal dysmotility; heart involvement of the pericardium, myocardium, and cardiac conduction system; renal failure; and lung involvement with ILD and pulmonary arterial hypertension (PAH).60 Moreover, arthralgia and peripheral neuropathy due to nerve entrapment are frequently present.

The thickening of the skin is the hallmark feature of SSc. At the cutaneous level, fibrosis begins at the deep derma and extends to involve the entire derma and the more superficial layers. The earliest findings may be puffiness, swelling, and decreased flexibility to the joint and tendons.61 Skin thickening is frequently accompanied by hyperpigmentation, giving a salt-and-pepper appearance. As SSc advances to the fibrotic stage, the skin becomes more thickened until atrophy occurs.
The obliterative small-vessel vasculopathy is also responsible for PAH and scleroderma renal crisis (SRC). Usually, in SSc patients, PAH develops late in the course of the disease. It is more frequent in patients with lSSc than in patients with dSSc. Typically, the development of PAH is related to the presence of ACA antibodies and anti-Th/To antibody, which are associated with lSSc, and with anti-U3RNP antibodies. PAH may be initially silent. Early symptoms are nonspecific (a sense of generalized weakness on exertion) with dyspnea appearing later. Echocardiography plays a pivotal role in the follow-up of SSc patients in order to identify patients with possible PAH, and right-heart catheterization is the gold standard to confirm the diagnosis.

SRC was the most common cause of death in SSc patients prior to the introduction of angiotensin-converting enzyme inhibitors. It is characterized by an abrupt rise in blood pressure over days to weeks, as well as by rapidly progressive renal failure, if untreated. SRC occurs in the diffuse form of the disease. Risk factors include the use of corticosteroids and cyclosporine A, and the presence of anti-RNA-pol antibodies. Clinical signs and symptoms are those of severe hypertension and can include headaches, stroke, and heart failure. The creatinine and plasma levels of renin are elevated; urinalysis shows proteinuria and microscopic hematuria.

ILD is a frequent complication of SSc. In SSc patients, ILD most commonly occurs within 10 years after diagnosis and it may be seen in patients with both the limited and diffuse subsets of the disease. The presence of antitopoisoomerase I antibodies (ATA) and anti-U11/U12 antibody is associated with an increased risk of ILD. Early lung disease is frequently asymptomatic; however, dry cough and dyspnea in a patient with SSc should raise the possibility of ILD, and a work up for ILD should be performed.

With regard to the diagnostic criteria, the new combined American College of Rheumatology (ACR)/European League Against Rheumatism clinical classification criteria, have been designed to improve the deficiencies of the earlier 1980 ACR clinical classification criteria. The advances in diagnostic techniques for Abs and nailfold videocapilloscopy (NVC). Indeed, the new criteria also incorporate Abs – specifically, anti-Scl70 (or ATA), RNAP, and ACA – which provide support for the classification and subtyping of SSc.

**Autoantibody subsets and organ involvement in SSc**

In Table 3, the Abs occurring with higher frequencies (>5%) in SSc patients and their association with organ involvement are reported. For frequently occurring antibodies, such as ACA and ATA, reliable detection systems based on enzyme-linked immunosorbent assay (ELISA) or other binding tests have been developed. For other antibodies like RNAPs, there are no standardized testing procedures, and the choice of the detection method is a critical point, impacting the sensitivity/specificity and reliability of the results.

ATA have been consistently found to be associated with dSSc, pulmonary fibrosis, and severe renal crisis, while ACA are present in up to 90% of patients with lSSc and are linked to PH and, rarely, to ILD. ACA-positive patients show a better prognosis than ATA-positive patients.

Serum Abs to small nuclear ribonucleoproteins (RNPs) have been found in patients with SSc and other connective tissue diseases (CTDs). Among the anti-RNP Abs, anti-U1 and anti-U3 (also called antifibrillarin) are the most frequent in SSc patients, while anti-U5 and anti-U4/U6 are rare. Lung fibrosis has been reported in 79% of the anti-U11/U12 RNP AB-positive SSc patients, while PH has been described in anti-U3-positive patients; moreover, anti-U11/U12 RNP-positive patients showed a higher mortality from pulmonary fibrosis than antibody-negative patients.

Anti-RNA-pol I, II, and III Abs are associated with severe clinical manifestations of SSc, especially anti-RNA-pol III. Although a reduced incidence of lung fibrosis has been found in anti-RNA-pol III-positive patients, these Abs are strongly associated with the development of SRC and the progression of skin disease. Additionally, anti-RNA-pol Abs have been found to be associated with a twofold hazard ratio (HR) for the development of cancer within 36 months of the clinical onset of SSc.

Antiendothelial cell Abs (AECA) are reported in 22%–86% of patients with SSc. These Abs are involved in the pathogenesis of autoimmune diseases with vascular involvement, but some AECA may bind to fibroblasts in vivo, triggering fibroblast activation. AECA positivity has also been noted in patients with other CTDs (rheumatoid arthritis [RA], systemic lupus erythematosus [SLE], and Wegener granulomatosis). Their presence has been associated with a higher frequency of pulmonary fibrosis and impairment of lung function tests, particularly the diffusing capacity of the lung for carbon monoxide (DLCO). Recently, a new set of Abs, called anti-RuvBL1/2, two highly conserved eukaryotic proteins implicated in transcription, DNA repair, and small nucleolar RNP assembly, has been associated with SSc in overlap with myositis; they occur with a frequency of 2% in Asian and North American populations.
The pathogenetic mechanisms underlying dermal fibrosis include inflammation, fibroblast recruitment, differentiation, and the deposition of ECM, and they are paradigmatic for other organs involved in SSc.

The gold standard for assessing SSc skin fibrosis is the mRSS, which is based on the determination of skin thickness at 17 distinct areas of the body (grading from 0 to 3 per area). The mRSS is widely used during evaluation in clinical routine, and as a surrogate endpoint in clinical trials. Despite the good intraobserver variability of the mRSS, its interobserver reproducibility is low. Therefore, biomarker investigations for skin and lung fibrosis in SSc are arising and include a broad spectrum of applications, ranging from the measurement of mediators in the peripheral blood and imaging (NVC), to gene expression profiling and proteomics. Several analytes have been identified as potential serum and plasma SSc biomarkers, mostly derived from matrix remodeling and fibrosis molecular pathways. Serum osteopontin (OPN), matrix metalloproteinase (MMP), and Krebs von den Lungen 6 (KL-6) have been found to correlate directly with the extent of skin fibrosis; moreover, adiponectin concentrations are lower in SSc patients with diffuse cutaneous form (dcSSc). None of these analytes are currently used during clinical routine.

OPN is a matrix protein with proinflammatory and profibrotic properties, binding to cell surface integrins and CD44; it modulates signaling in a wide variety of cell types, like fibroblasts and myofibroblasts. Wu et al demonstrated high plasma OPN levels in SSc patients, both with limited and diffuse disease, as compared to healthy controls. Recently, an OPN SNP, previously reported to be associated with

### Table 3 Specific nuclear antigen and nonantinuclear autoantibodies and organ involvement in SSc

<table>
<thead>
<tr>
<th>Specific nuclear antigen antibodies</th>
<th>Lung</th>
<th>Skin</th>
<th>DU</th>
<th>Joint/muscle</th>
<th>GI, renal, cardiac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Histones (AHA)*</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
</tr>
<tr>
<td>U1-RNP**</td>
<td>+ (PH)</td>
<td>+ (PH)</td>
<td>+ (PH)</td>
<td>+ (PH)</td>
<td>+ (PH)</td>
</tr>
<tr>
<td>U3-RNP*</td>
<td>+ (ILD, PH)</td>
<td>+ (limited)</td>
<td>+ (limited)</td>
<td>+ (limited)</td>
<td>+ (limited)</td>
</tr>
<tr>
<td>U11-RNP**</td>
<td>+ (severe ILD)</td>
<td>+ (severe ILD)</td>
<td>+ (severe ILD)</td>
<td>+ (severe ILD)</td>
<td>+ (severe ILD)</td>
</tr>
<tr>
<td>U12-RNP**</td>
<td>+ (severe ILD)</td>
<td>+ (severe ILD)</td>
<td>+ (severe ILD)</td>
<td>+ (severe ILD)</td>
<td>+ (severe ILD)</td>
</tr>
<tr>
<td>RNP*</td>
<td>+ (PH)</td>
<td>+ (PH)</td>
<td>+ (PH)</td>
<td>+ (PH)</td>
<td>+ (PH)</td>
</tr>
<tr>
<td>SSA (Ro60)**</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
</tr>
<tr>
<td>Ro52 (TRIM21)**</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
</tr>
<tr>
<td>Ku*</td>
<td>+ (myositis)</td>
<td>+ (myositis)</td>
<td>+ (myositis)</td>
<td>+ (myositis)</td>
<td>+ (myositis)</td>
</tr>
<tr>
<td>ScI-70***</td>
<td>+ (ILD)</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
</tr>
<tr>
<td>ACA/CENP (A-E)</td>
<td>+ (PH, ILD)</td>
<td>+ (limited)</td>
<td>+ (limited)</td>
<td>+ (limited)</td>
<td>+ (limited)</td>
</tr>
<tr>
<td>Th/To*</td>
<td>+ (ILD, PH)</td>
<td>+ (ILD, PH)</td>
<td>+ (ILD, PH)</td>
<td>+ (ILD, PH)</td>
<td>+ (ILD, PH)</td>
</tr>
<tr>
<td>RNA-pol-I*</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
</tr>
<tr>
<td>RNA-pol-II*</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
</tr>
<tr>
<td>RNA-pol-III*</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
</tr>
<tr>
<td>Anti-RuvBL1/2</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
</tr>
<tr>
<td>Anti-tRNA synthetases PL-7***</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
</tr>
<tr>
<td>PMSc2***</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
</tr>
<tr>
<td>Non-ANA</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
</tr>
<tr>
<td>AECA</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
</tr>
<tr>
<td>EIF-2</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
<td>+ (ILD)</td>
</tr>
<tr>
<td>AFA</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
<td>+ (diffuse)</td>
</tr>
<tr>
<td>Anti-M3R</td>
<td>+ (GAVE)</td>
<td>+ (GAVE)</td>
<td>+ (GAVE)</td>
<td>+ (GAVE)</td>
<td>+ (GAVE)</td>
</tr>
<tr>
<td>AMA M2</td>
<td>+ (PBC)</td>
<td>+ (PBC)</td>
<td>+ (PBC)</td>
<td>+ (PBC)</td>
<td>+ (PBC)</td>
</tr>
</tbody>
</table>

**Abbreviations:** ACA, anti-centromere antibody; AECA, anti-endothelial cells antibodies; AFA, antifibrillarin antibodies; AHA, anti-histone antibodies; AMA, antimitochondrial antibodies; ANA, antinuclear autoantibodies; CENP, centromere associated protein; DU, digital ulcer; EIF-2, eukaryotic initiation factor 2; GAVE, gastric antral vascular ectasia; GI, gastrointestinal; ILD, interstitial lung disease; M3R, muscarinic receptor 3; PBC, primary biliary cirrhosis; PH, pulmonary hypertension; PM, polymyositis; RNA-pol, RNA polymerase; RNP, ribonucleoprotein; SSA, anti-Sjogren's syndrome-related antigen A; SSc, systemic sclerosis.

**Notes:** *Nucleolar pattern at Hep-2 immunofluorescence (IF); **speckled pattern at IF; *cytoplasmic pattern at IF. Antibodies with rare incidence (<5%) have been excluded. + indicates established association (several studies); +/- indicates insufficient body of evidence or controversial data. Blank spaces in the table indicate non reported. © Georg Thieme Verlag KG. Adapted with permission from Bonella F, Costabel U. Biomarkers in connective tissue disease-associated interstitial lung disease. Semin Respir Crit Care Med. 2014;35(2):181–200.
SLE, has been found to be more frequent in SSc patients than in healthy controls, suggesting a role of OPN in SSc susceptibility.

MMPs belong to the family of proteases, and they are enzymes capable of degrading all kinds of ECM proteins and processing a number of bioactive molecules. Serum levels of MMP-9 and MMP-12 have been found to be elevated in SSc patients with the dcSSc form, and they also correlate with mRSS, ILD, and the presence of DUs. In particular, MMP-12 has been found to be strongly expressed in keratinocytes, dermal endothelial cells, myofibroblasts, and inflammatory cells in the skin of patients with SSc, and it correlates with peripheral vascular damage.

KL-6 is a glycoprotein belonging to MUC class I. KL-6 is mainly produced by regenerating pneumocytes type II and, in Japan, it is routinely used to assess the severity of lung fibrosis and other conditions. KL-6 is not only an epiphenomenon; it shows profibrotic and antiapoptotic effects on lung fibroblasts. The precise epitope structure recognized by the anti-KL-6 monoclonal (m)Ab is still unclear, but candidate carbohydrate epitopes for KL-6 mAb have been identified in O-linked glycans of MUC1 containing 6′sulfo-Gal/GalNAc. In a small cohort, Bonella et al found that KL-6, but not surfactant proteins (SP)-D, strongly correlated to skin involvement (mRSS) and a disease activity index. This finding led the authors to speculate that circulating levels of KL-6, similar to other pleonastic molecules, may directly contribute to chronic, persistent systemic fibrosis in a TGF-β-additive way in SSc.

Adiponectin is a downstream target of PPAR-gamma, a nuclear receptor involved in obesity and metabolism. Serum adiponectin concentrations are measured in diabetic patients to assess response to treatment with thioglitazones. Recent studies found an inverse correlation between the extent of skin disease and serum adiponectin levels in SSc patients. Adiponectin was decreased in dcSSc patients when compared to healthy controls and other rheumatic diseases.

Biomarkers associated with DUs
DUs in SSc are the main clinical manifestation of peripheral vascular disease and a severe complication requiring early medical intervention. High levels of serum placental growth factor, low counts of circulating endothelial progenitor cells (EPCs), and high levels of interleukin (IL)-6 have been found to be independent predictors of the occurrence of new DUs. Elevated serum IL-6 in early dcSSc has also been shown to identify a subgroup of patients with more severe skin disease, ILD, and reduced survival.

The Capillaroscopy Skin Ulcer Index was proposed in 2009 by Sebastiani et al, and it was recently validated in 229 SSc patients as a tool to identify patients with a high risk for developing DUs at 3 months from the initial evaluation, independently from the video capillaroscopy devices. A simpler score based on mean capillary loss at each finger has been also developed to predict digital trophic lesions in SSc patients with a follow-up of 6–12 months.

The number of studies investigating possible correlations between NVC severity patterns and SSc disease activity are increasing. In a recent investigation of two independent cohorts, NVC patterns were found to predict novel future severe organ involvement with a stronger odds ratio according to the more severe NVC patterns (early, active, and late).

Biomarkers of ILD
ILD occurs more often in SSc than in any other CTD and it is radiologically evident in up to two-thirds of patients with SSc. The most prevalent histologic pattern identified in SSc is fibrotic Nonspecific interstitial pneumonia followed by the usual interstitial pneumonia lung injury pattern.

The Scleroderma Lung Study (SLS) and GENISOS collected the most significant cohorts aimed to identify biomarkers for predicting SSc-ILD and its outcomes. A number of Abs (Table 3) and serum proteins (Table 4) correlating with the presence or severity of ILD in the context of SSc have been investigated. Even if the majority of lung-derived biomarkers are shared between idiopathic ILD and SSc-ILD, some differences exist, reflecting the peculiar pathogenetic mechanisms of the two conditions.

Abs associated with the presence of lung fibrosis
While there is enough evidence that ATA positivity is associated with a greater risk of lung fibrosis, it remains unclear whether it is associated with more progressive ILD and further investigations are needed.

Serum Abs to small nuclear RNPs have been found not only in SSc patients, but also in other CTDs. Some Abs recognize individual RNPs, while others are directed against a complex of RNPs. Among them, anti-U1 and anti-U3 are the most frequent in SSc patients, while anti-U5 and anti-U4/U6 are rare. Lung fibrosis has been reported in 79% of the anti-U11/U12 RNP Ab-positive SSc patients; moreover, anti-U11/U12 RNP positivity seems to be linked to higher mortality in SSc-ILD patients. Despite more severe skin involvement, patients with anti-RNA-pol antibody positivity have significantly lower frequencies of telangiectasias, inflammatory myopathy, restrictive lung...
disease, and serious cardiac abnormalities than those with ATA positivity.\textsuperscript{72}

AECA are a heterogeneous class of antibodies whose role in the pathogenesis of autoimmune diseases with vascular involvement has been extensively studied.\textsuperscript{75} Some AECA may directly activate fibroblasts and cause fibrosis.\textsuperscript{76} AECA positivity, reported in 22\%–86\% of patients with SSc, has also been seen in other CTDs.\textsuperscript{75} AECA have been

\begin{table}[h]
\centering
\begin{tabular}{|l|l|l|l|}
\hline
\textbf{Organ and cellular source} & \textbf{Protein} & \textbf{Evidence} & \textbf{Reference} \\
\hline
\textbf{Lung} & & & \\
Damaged alveolar epithelium & SP-A & + & 129 \\
 & SP-D\textsuperscript{†} & +++ & 93,108,127,130,145,148 \\
 & CXCL9 & +/- & 152 \\
 & CXCL10 & +/- & 151,152,244,245 \\
 & CXCL12 & + & 153 \\
\hline
\textbf{Extracellular matrix (ECM)} & MMP-7 & + & 137 \\
 & MMP-9 & + & 88 \\
 & MMP-12 & + & 89 \\
 & ADAM-12 (reduced) & + & 139 \\
\hline
\textbf{Activated macrophages} & CCL8 & + & 144–146,148 \\
 & CCL2 & +++ & 102,150–152,154,245 \\
 & IL-6 & +++ & 102,152,154,246,247 \\
 & IL-10 & + & 102,152 \\
 & TNF-alpha & + & 152,164 \\
 & CTGF & + & 248 \\
 & YKL-40 & + & 133 \\
\hline
\textbf{Activated lymphocytes T} & IL-4 & +/- & 152 \\
 & IL-13 & + & 247 \\
 & IL-17A & +/- & 247 \\
 & IL-22 & + & 162 \\
 & IL-23 & + & 161\textsuperscript{*} \\
\hline
\textbf{Activated lymphocytes B} & sFLC\textsuperscript{*} & + & 168 \\
\hline
\textbf{Activated platelets} & CXCL4 & +++ & 46,155 \\
\hline
\textbf{Damaged endothelium} & ET-1 & + & 171–173,175 \\
 & VCAM-1 & + & 174,175 \\
 & E-Selectin & + & 175 \\
 & VEGF & + & 175 \\
\hline
\textbf{Acute-Phase proteins} & CRP\textsuperscript{‡} & +++ & 249 \\
 & PTX3 & + & 206 \\
\hline
\textbf{Skin} & & & \\
 ECM & OPN & + & 85 \\
 Epithelia (airway/intestine) & MMP 9 and 12 & + & 88,89 \\
 Adipocytes & KL-6 & +/- & 93 \\
 & APN (reduced) & + & 95,96 \\
\hline
\textbf{Vascular involvement} & & & \\
 Peripheral (DU) & PIGF & + & 100 \\
 Pulmonary hypertension & EPC (reduced) & + & 100 \\
 & IL-6 & + & 101 \\
 & BNP & +++ & 197–200,250,251 \\
 & ET-1 & + & 173,175 \\
 & GDF-15 & + & 202,203 \\
 & PTX3 & + & 205,206 \\
 & FGF (reduced) & + & 206 \\
\hline
\end{tabular}
\caption{Circulating biomarkers associated with organ involvement in SSc and their cellular source}
\end{table}

Notes: Otherwise indicated, serum levels of the biomarkers are increased. +/- indicates controversial (non confirmatory studies); + indicates one single center study; ++ indicates more single center studies; +++ indicates large cohort studies; \textsuperscript{†}GENISOS cohort study; \textsuperscript{*}free light chains of immunoglobulins; \textsuperscript{‡}CSRG (Canadian Scleroderma Research Group); confirmed strong association with restrictive functional impairment but not with ILD. The evidence of clinical utility is scored in the table. \textcopyright Georg Thieme Verlag KG. Adapted with permission from Bonella F, Costabel U. Biomarkers in connective tissue disease-associated interstitial lung disease. \textit{Semin Respir Crit Care Med}. 2014;35(2):181–200.\textsuperscript{241}

Abbreviations: ADAM, A disintegrin and metalloproteinase; APN, adiponectin; BNP, brain natriuretic peptide; CCL, chemokine (C-C motif) ligand; CRP, C reactive protein; CTGF, connective tissue growth factor; CXCL, chemokine (C-X-C motif) ligand; DU, digital ulcers; ET-1, endothelin 1; EPC, circulating endothelial progenitor cells; FGF, fibroblast growth factor; GDF-15, growth differentiation factor 15; IL, interleukin; ILD, interstitial lung disease; KL-6, Krebs von den Lungen-6; MMP, matrix metalloproteinase; OPN, osteopontin; sFLC, free light chains of immunoglobulins; PIGF, placental growth factor; PTX3, pentraxin 3; SP, surfactant protein; SSC, systemic sclerosis; TNF-alpha, tumor necrosis factor alpha; YKL-40, chitinase-3-like protein 1; VCAM-1, vascular cell adhesion protein 1; VEGF, vascular endothelial growth factor.
associated with a higher frequency of ILD and lower levels of DLCO. 77

Human leukocyte antigens haplotypes

Human leukocyte antigens (HLA), PTPN22, CTGF, TGF-β, and PDGF are established molecules involved in the pathophysiology of SSc112,113 but the number of studies investigating these molecules as biomarkers for SSc-ILD is limited. 113 Odani et al113 reported that the prevalence of HLA-DRB5 gene carriers was increased in Japanese SSc patients with ILD relative to those without ILD or to healthy controls in both cohorts. The HLA-DRB5*01:05 allele was significantly more frequent in SSc patients with ILD than in SSc patients without ILD or in healthy controls. 113 These associations were also confirmed in a validation cohort. The authors concluded that the HLA-DRB5*01:05 allele could be linked to ILD susceptibility in SSc patients.113

Lung epithelium derived proteins

Although the studies on lung epithelial derived proteins provided the largest evidence for these markers as predictors of SSc-ILD, there are no specific biomarkers to distinguish idiopathic ILD from SSc-ILD. Most of these proteins, especially KL-6, SP-D, and CCL-18, correlate with the extent of lung fibrosis and can predict short-term decline of lung function.

KL-6 (MUC1) (the general characteristics of KL-6 have been presented earlier) is probably the most established biomarker for the assessment of ILD, at least in the Japanese population. 90 Serum KL-6 is elevated not only in idiopathic interstitial pneumonia, but also in ILD of known origin, including hypersensitivity pneumonitis, drug-induced pneumonitis, sarcoidosis, and CTD-ILD. 90 KL-6 has been investigated in several studies in SSc-ILD patients, 93,108,114–119 and its serum levels seem to correlate inversely with DLCO. 93,108 The presence of the MUC-1 568 adenosine to guanine polymorphism (rs4072037) has been found to influence serum KL-6 levels in Caucasian and Japanese ethnicities. 120,121 In the most recent study on KL-6 in SSc, Kumánovics et al119 assessed serum KL-6 levels over time from 173 SSc patients. The negative correlation of serum KL-6 with lung function parameters, independently of the time of investigation, as well with the probability of death were confirmed.119 However, no significant differences were seen in the deterioration and improvement rates between groups with normal and elevated KL-6 levels at baseline.119 The authors concluded that even if KL-6 seems to be useful as a marker of lung fibrosis in SSc, its role as disease activity marker for SSc remains questionable.119

Cube cell 16-kDa (CC16) is a low-molecular-weight protein produced by bronchiolar Cube cells122 and released in consequence to lung injury. 122–124 Increased CC16 levels have also been found in serum and bronchoalveolar lavage fluid (BALF) of patients with sarcoidosis. 122,125,126 Recently, Hasegawa et al127 compared the value of CC16, KL-6, and SP-D serum levels for the diagnosis and monitoring of pulmonary fibrosis SSc, SLE, and healthy controls. A CC16 level of 46.0 ng/mL or higher was diagnostic of ILD (sensitivity [Se]: 52%; specificity [Sp]: 89%), slightly inferior to KL-6 (Se: 85%; Sp: 85%), but comparable with that of SP-D (Se: 71%; Sp: 77%).127 Serum levels of CC16 correlated inversely with pulmonary function tests (FVC and DLCO) and disease activity in patients with SSc-ILD.127

SP-D and SP-A are secreted by type II alveolar epithelial cells and airway Cube cells; they contribute to stabilize the alveolar surface tension and play a pivotal role in the homeostasis of the host–defense system. 128 Serum levels of SP-A and SP-D are significantly higher in SSc patients with ILD than in those without. 93,129,130 Serum SP-D serum levels are negatively correlated with FVC and DLCO in SSc-ILD patients. 129,130 In one study of 42 SSc patients, SP-D levels were more sensitive than SP-A in detecting ILD, as defined by high-resolution computed tomography (HRCT), but they are less specific (Se is 77% and Sp is 83% for SP-D versus 33% and 100% for SP-A). 128 In 66 SSc patients from the SLS cohort, SP-D levels >90 ng/mL demonstrated good Se (89%) and Sp (80%) in the assessment of “alveolitis”, as defined by bronchoalveolar lavage (BAL) neutrophilia and/or HRCT ground glass density. 108 A positive correlation between SP-D serum levels and ILD worsening has been described in a small prospective study of 35 patients with SSc-ILD followed up to 10 years. 130 Moreover, Yanaba et al116 showed retrospectively that longitudinal SP-D levels were useful to assess the response to immunosuppressive therapy, decreasing by clinical improvement.

Chitinase-like protein YKL-40 is produced by activated macrophages and it acts as a growth factor for connective tissue cells. It is elevated in the sera of patients with diseases characterized by inflammation, tissue remodeling, or fibrosis. 131 Serum and BALF YKL-40 levels are predictors of survival in IPF. 132 YKL-40 is increased in the sera of SSc patients and correlates with airway obstruction, reduced DLCO, and poor prognosis. 133 The presence of single nucleotide polymorphisms (SNPs) in the gene encoding YKL-40
(CHI13-329 and CHI13-321) is associated with the interindividual variability of the levels of this marker in serum.\textsuperscript{132}

Calcium-binding proteins S100A8 and S100A9 (also named calgranulin A, myeloid-related protein [MRP]8, and calgranulin B, MRP14) are low-molecular-weight members of the S100 family of calcium-binding proteins involved in the initiation and amplification of inflammation\textsuperscript{134} They are located in the cytoplasm and nucleus of a wider range of cells. Xu et al\textsuperscript{134} found that plasma concentrations of S100A8 and S100A9 were significantly higher in dcSSc patients than in normal controls and lcSSc patients, but above all in those with dcSSc with lung and kidney involvement. Recently, van Bon et al\textsuperscript{13} confirmed in a double-cohort study that S100A8/A9 levels were elevated in the plasma of SSC patients, although more so in lcSSc than in dcSSc. This discrepancy between the findings of Xu et al\textsuperscript{134} and van Bon et al\textsuperscript{13} could be ascribed to ethnicity or to differences in the assay used. Since S100A8/A9 are considered to be ligands of the Toll-like receptor (TLR)-4, and since they are able to induce fibroblast proliferation and pulmonary fibrosis, van Bon et al\textsuperscript{13} conclude that the correlation between lung fibrosis and S100A8/A9 levels may not only be an epiphenomenon.

**Biomarkers derived from ECM**

Characteristics of MMPs and their inhibitors (TIMP) have been described. Among SSC patients, higher serum levels of MMP-7, an established prognostic biomarker for IPF,\textsuperscript{135} and MMP-9 (this one was also elevated in BALF) were found in patients with ILD when compared to those without, and they were associated with lower DLCO. The association of these MMPs with ILD progression was not evaluated in these studies.\textsuperscript{88,136,137}

Circulating levels of MMP-12 have been reported to correlate with FVC ($r=-0.82$), and the expression of MMP-12 is significantly increased in SSc-ILD lungs when compared to normal controls.\textsuperscript{39} TIMP-1 levels were found to be associated with the presence of ILD in 62 SSC patients and to correlate inversely with DLCO.\textsuperscript{138}

A disintegrin and metalloproteinase (ADAM)12 is one of the MMP-type ADAMs and exerts extracellular protease and cell-binding functions. ADAM12 is expressed in two alternative forms, which are a membrane-anchored form (ADAM12-L) and a short secreted form (ADAM12-S).\textsuperscript{139} Taniguchi et al\textsuperscript{139} measured serum ADAM12-S levels in 61 SSc patients and found a negative correlation with HRCT fibrosis score and a positive one with the ground glass score, indicating interstitial inflammation. ADAM12-S levels showed a positive correlation with FVC% predicted, but not with DLCO. On the basis of these results, the authors concluded that serum ADAM12-S seems to be more linked to inflammation than fibrosis\textsuperscript{139}

In a recent study of Oka et al\textsuperscript{140} MMP-1, MMP-8, MMP-9, and TIMP-3 were upregulated in the sera of 25 CTD patients (RA, SSc, and polymyositis/dermatomyositis) with acute-onset ILD. In contrast, MMP-3 and TIMP-2 were downregulated. The ratio of MMP-3 to MMP-1 serum levels, reduced in patients who died, has been proposed by the authors as a predictor of poor outcomes in these patients.\textsuperscript{140} These results need validation in a larger cohort of CTD-ILD patients.

Lysyl oxidase-like (LOXL) enzymes are proteins necessary for stability of the ECM, cross-linking type I collagen fibrils. LOXL-2 is highly expressed in the fibrotic lungs of IPF patients and is implicated in maintaining the fibrotic microenvironment.\textsuperscript{141,142} A Phase II clinical trial with simtuzumab (GS-6624), a mAb targeting soluble LOXL-2, is currently ongoing in IPF (NLM identifier: NCT01769196).

Serum levels of LOXL-2 have also been found to be significantly elevated in SSc, particularly in patients with dcSSc disease.\textsuperscript{143} Interestingly, serum LOXL-2 concentration correlated with the mRSS in patients without lung fibrosis.\textsuperscript{143} Whether targeting anti-LOXL-2 could be a potential therapeutic strategy for SSc-ILD should be addressed by a specific clinical trial.

**Chemokines and ligands**

Among the studied chemokines, CCL18 and CXCL4 are those with the largest evidence as biomarkers for detecting lung fibrosis and for assessing lung disease severity. While CCL18 shows some predictive value for ILD progression, but is not specific for SSc, CXCL4 strongly correlates with the presence of fibrosis in SSc patients.

CCL18, previously known as pulmonary and activation-regulated chemokine, is constitutively expressed at high levels in the lungs, mainly by alveolar macrophages, and is chemotactic for a variety of mononuclear cells.\textsuperscript{144} CCL18 is profibrotic and has been found increased in the serum and BALF of IPF and SSc-ILD patients.\textsuperscript{144} Three longitudinal studies in a total of 116 patients with SSc-ILD investigated the correlation of serum CCL18 with ILD activity, as determined by HRCT, lung function, and BAL analysis, with promising results.\textsuperscript{144,146} Tiev et al\textsuperscript{146} analyzed serum CCL18 levels in 83 SSc patients with ILD over a 4-year follow-up period. At a cutoff of 187 ng/mL, serum CCL18 was independently predictive of ILD worsening (HR: 5.4; $P=0001$). These results were similar to those by Prasse et al\textsuperscript{147} in IPF...
patients. However, Elhaj et al.\textsuperscript{149} compared the predictive significance of CCL18 and SP-D for the disease outcomes of SSc-ILD in the GENISOS cohort, in 266 patients.\textsuperscript{109,148} For 83 patients, it was possible to collect a complete follow-up for biomarkers over 4 years. In contrast to the previous studies,\textsuperscript{144,146} Elhaj et al.\textsuperscript{148} did not show a significant predictive significance of baseline serum CCL18 for long-term changes in FVC. Moreover, the correlation between baseline serum CCL18 and concomitant or 1-year follow-up FVC or DLCO was not confirmed. A subgroup analysis based on disease type (limited/diffuse) or ATA positivity did not yield significant results for short- or long-term predictive values of SP-D or CCL18.\textsuperscript{148}

CCL2, CXCL10, and CXCL12 were found to be upregulated in the serum of SSc patients and to correlate with the presence of ILD.\textsuperscript{149–153} CCL2 concentrations were also found to be elevated in BALF and they were associated with the presence of ILD, correlating with lung function parameters and computed tomography scores.\textsuperscript{154}

CXCL4 is an angiogenic and profibrotic factor that inhibits the expression of the antifibrotic cytokine interferon-γ and upregulates profibrotic cytokines IL-4 and IL-13.\textsuperscript{155} A large multicohort study of 779 SSc patients (462 with lcSSc and 317 with dcSSc) demonstrated that CXCL4 is the predominant protein secreted by plasmacytoid dendritic cells in SSc, both in circulation and in the skin.\textsuperscript{46} CXCL4 levels correlated with skin and lung fibrosis and with PAH. Moreover, only CXCL4 among the studied biomarkers (CCL2, CXCL10, CCL5, vWF, and CCL18) predicted the risk and progression of SSc.

**ILs and their ligands**

ILs are cytokines with pleiotropic function, and they cannot be considered specific of organ involvement. Among ILs investigated in SSc patients with lung fibrosis, IL-6 seems to be the most promising for application in clinical routine to assess disease severity and to predict outcomes.

Elevated serum and BALF levels of IL-6, a cytokine linked to the Th2 lymphocyte subset, have been described in several lung diseases.\textsuperscript{156–158} Recently, the prognostic value of serum IL-6 was assessed in a large cohort of 212 patients with SSc-ILD.\textsuperscript{102} In the exploratory analysis, only serum IL-6 among IL-8, IL-10, CCL2, CXCL10, VEGF, fibroblast growth factor (FGF)-2, and CX3CL1 was an independent predictor of DLCO decline in both IPF and SSc-ILD. At a cutoff of 7.7 pg/mL, serum IL-6 was predictive of a decline in FVC (HR: 2.58) and in DLCO (HR: 3.2) within the 1st year, and predictive of death within the first 30 months (HR: 2.69).\textsuperscript{102}

Wutte et al.\textsuperscript{159} found that serum levels of another pleiotropic cytokine, IL-15, a survival and growth factor for T- and B-lymphocytes, were strongly correlated with impaired lung function and ILD in SSc.

Levels of IL-17A and IL-23, connected to the Th17 lymphocyte subset, are elevated in SSc and associated with the presence of ILD.\textsuperscript{160,161} Circulating IL-22- and IL-17-producing T-cells are increased in patients with SSc-ILD when compared to those without.\textsuperscript{162} These findings support the hypothesis that Th22 may also be involved in pathological processes leading to SSc; the increased frequency of Th22 cells appears to be of interest as a biomarker in SSc.\textsuperscript{162}

Tumor necrosis factor (TNF)-related apoptosis-inducing ligand (TRAIL), a member of the TNF family of cytokines linked with apoptosis, is involved in the pathophysiology of some autoimmune diseases and cancer.\textsuperscript{163} Mean serum TRAIL levels were found to be significantly higher in SSc patients than in RA patients and healthy controls ($P<0.001$). Moreover, serum TRAIL levels correlated with the HRCT score.\textsuperscript{164}

**Biomarkers from activated B-cells**

It is increasingly thought that humoral immunity and B-cells play an important role in the pathophysiology of SSc, producing specific Abs and profibrotic cytokines.\textsuperscript{165,166} SSc patients have abnormal blood B-cell compartments with an expansion of naïve B-cells and activated memory B-cells.\textsuperscript{167} In a recent study by Lanteri et al.,\textsuperscript{168} levels of free light chains of immunoglobulins (sFLC) were assessed in the serum from 134 SSc patients and 401 healthy controls. In comparison to other markers of B-cell activation (immunoglobulin [Ig]G, IgA, beta2-microglobulin, and B-cell activating factor), sFLC serum levels correlated with mRSS and were independently associated with the presence of ILD and its severity.\textsuperscript{168} Since sFLC can bind antigens or activate some cells, such as mast cells, involved in the pathogenesis of lung fibrosis,\textsuperscript{169} these findings suggest that sFLC could have a direct role in the pathogenesis of SSc-ILD.\textsuperscript{168}

**Biomarkers derived from endothelial cells**

Endothelial cell injury is involved in the pathogenesis of fibrosis.\textsuperscript{170} ETs are potent peptides regulating both cell proliferation and turnover of the ECM. Increased ET-1 levels have been found in BALF and in lung tissue from patients with SSc.\textsuperscript{171,172} Increased serum levels of soluble vascular cell adhesion molecule-1, soluble E-selectin, VEGF, and ET-1 have been described in patients with SSc and lung fibrosis.\textsuperscript{173–175} AECA have been discussed earlier.
Circulating cells as biomarkers

Circulating endothelial cells (CECs) and EPCs have been studied in different conditions characterized by vasculopathy. High levels of CECs and EPCs have been reported in patients with SSc and seem to correlate with disease activity, with DU, and with PAH. Moreover EPCs are higher in the early stages of the disease, suggesting an attempt to repair endothelial cell damage and to favor revascularization.176,177

Fibrocytes are circulating precursors for fibroblasts. The fibrocytes, and their CD14-positive mononuclear precursors, are recruited from the blood to organs through multiple chemokine signaling pathways (particularly, the CCL2/CCR2 and CXCL12/CXCR4 pathways).178,179 Fibrocytes have been found increased in the blood of patients with IPF, ranging from 2.7% of the nucleated cells in stable IPF to 14.5% in the acute exacerbation of IPF (1% fibrocytes are present in healthy subjects).180 Recently, Borie et al181 detected fibrocytes in the alveolar space in 55% of SSc patients through BAL. Since fibrocyte detection was not associated with the capacity to grow fibroblasts from BALF in cell culture, the physiological role and their interaction with circulating fibrocytes remain unclear.182

BAL cell profiles

BAL differential cell counts in SSc are of questionable diagnostic and prognostic value.182 Nevertheless, a study on 73 SSc patients showed an association with BAL cellularity and the ILD pattern.183 A positive correlation between BAL neutrophils, the extent of disease on HRCT, and poor survival has also been described in SSC-ILD.184–186 In the study by Goh et al,187 BAL cellular profiles in 141 patients with SSc-ILD did not show a prognostic value for disease progression, since disease severity had been quantified by pulmonary function tests and HRCT. BAL neutrophilia, independent of disease severity, was associated with early mortality (within 2 years), but not with long-term prognosis.

In the SLS cohort, BAL cell counts were performed in 201 patients as part of a randomized placebo-controlled trial of cyclophosphamide versus placebo.188 Abnormal cellularity was present in 101 patients, and a subgroup was identified with more severe lung function, more extensive ground glass opacity, and more extensive fibrosis in the right middle lobe. The cell differential was not an independent predictor of disease progression or response to cyclophosphamide at 1 year.189

Nailfold capillary patterns and ILD

NVC is a noninvasive diagnostic tool that permits the detection of the main local microvascular alterations in SSc, characterized by enlargement of the capillary loops, a loss of capillaries, disruption of the orderly appearance of the capillary bed, and distortion of the capillaries.189 Two studies found a positive correlation between the vascular deletion score in NVC and skin thickness, the presence of anti-Scl-70 antibodies, signs of peripheral ischemia, esophageal dysfunction, and pulmonary disease.180 Recently, Corrado et al192 highlighted the differences in the capillary pattern between SSc-ILD and IPF. They pointed out that capillaroscopic changes in patients with lung fibrosis could indicate the first signs of a systemic disease involving the lungs, thus requiring particular attention in detecting the appearance of immunological/clinical signs of CTD.192

Biomarkers of PH associated with SSc

Pulmonary (arterial) hypertension in SSc (SSc-PH) can occur as a primary lung manifestation in association with ILD and/or as a consequence of left-sided heart disease. Patients with combined SSc-ILD and PH have a worse prognosis in comparison to those with ILD alone, with a 3-year survival of 47%.193 A severe discrepancy in the reduction in DLCO compared to lung volumes and the development of hypoxia can be early signs of PH.194,195 Since echocardiography has low Sp for the diagnosis of PH,196 and given that right-sided cardiac catheterization is invasive and cannot be used routinely for monitoring purposes, several serum biomarkers have been investigated. Among the studied biomarkers, brain natriuretic peptide (BNP) is widely applied in the clinical routine for detecting right ventricular dysfunction and indirectly monitoring PH.

Natriuretic peptides include BNP and N-terminal pro-BNP (NT-pro-BNP), and they are released in response to ventricular stretch by cardiomyocytes. Natriuretic peptides have been established as markers of right ventricular dysfunction in PH.197,198 BNP was also found to be associated with poor prognosis in patients with chronic lung diseases.199 In a study of 90 ILD patients – including 18 CTD-ILD – Corte et al200 found that serum BNP correlated with pulmonary pressures estimated by echocardiography. Furthermore, a serum BNP ≥ 20 pmol/L was associated with a 14-fold increased mortality compared to patients with BNP < 4 pmol/L.200 However, since natriuretic peptide levels are released in response to right ventricle strain, markers of earlier pulmonary vasculopathy are still needed.

ET-1 has been proposed as a biomarker for both vascular involvement and lung fibrosis development in patients with SSc.173,175 However, Camargo et al showed that serum levels
of ET-1 and vWF (a carrier protein for coagulation factor VIII), which can predict the development of pulmonary hypertension in SSc patients, were not correlated with SSc disease stage.201

Growth differentiation factor (GDF)-15 is a secreted member of the TGF-β cytokine superfamily and its levels are elevated in the serum of patients with cardiovascular diseases. Meadows et al202 demonstrated that patients with SSc-PH had significantly higher mean circulating levels of GDF-15 in plasma compared with patients with SSc without PH (422±369.5 pg/mL versus 108±193 pg/mL, respectively; \( P=0.004 \)). GDF-15 levels correlated positively with estimated right ventricular systolic pressure on echocardiogram and plasma levels of the amino terminal propeptide form of BNP.202 GDF-15 levels >125 pg/mL were associated with reduced survival. In a further study of 119 SSc patients, increased GDF-15 serum levels correlated with disease activity and the extent of lung fibrosis, including lung function, at follow-up.203

Pentraxin 3 (PTX3) is a multifunctional pattern recognition protein that can suppress FGF-2, has nonredundant roles in antimicrobial innate immunity, and is a component of the ECM, contributing to fibrosis.204 In a recent study from a Japanese cohort of 171 SSc patients, Shirai et al205 found PTX3 elevated in the serum of SSc patients who had DUs or PH, while the FGF-2 concentration was reduced. Multivariate analysis selected PTX3 as an independent parameter for the presence of DUs and PAH.205 Since, in a previous study, a correlation between elevated PTX3 and pulmonary fibrosis, but not PH,206 was seen, the study by Shirai et al205 is the first to address serum PTX3 as a biomarker for PH in SSc.

Biomarkers for liver involvement
Primary biliary cirrhosis (PBC) is the most common liver disease usually associated with scleroderma.207 It is a chronic liver disease characterized by cholestasis and immune-mediated nonsuppurative cholangitis, mainly affecting interlobular and septal bile ducts.208 PBC usually presents with pruritus followed by progressive jaundice. Antimitochondrial antibodies (AMA) are diagnostic.208 The reported prevalence of clinically evident PBC among patients with SSc is 2%–2.5%.209,210 The prevalence of AMA positivity among SSc patients is estimated to be 5%,210 and ACA are present in 25% of patients with PBC. However, a high prevalence of AMA-negative PBC in patients with SSc has also been described.211 Dual isotype ELISA, which detects multiple mitochondrial and nuclear Abs specific for PBC, has been developed to enhance the detection of PBC in patients that were found to be AMA negative by immunoﬂuorescence.212 Prince et al213 demonstrated that AMA can precede clinical symptoms of PBC with a median time of 5 years; therefore, patients require long-term monitoring of liver function tests and abdominal ultrasound.213

**Multiple biomarker assessment**
SSc is a biologically complex disease; its multiple pathogenetic patterns suggest that the best predictive tool should come from the combination of many mediators.3 The multibiomarker disease activity tool for RA is based on 12 biomarkers and was found to be consistently associated with clinical disease activity at baseline and over time, and to correlate with damage progression, as measured by radiography.214,215

Several attempts have been made to compare or combine different biomarkers at the same time for assessing disease activity and organ involvement in SSc, but mostly with a limited number of molecules.95,118,129

Beirne et al126 investigated a panel of 30 circulating biomarkers in order to differentiate between the presence or absence of lung fibrosis in 20 SSc patients, as compared with healthy controls and sarcoidosis patients. Interestingly, only a 17-analyte biomarker model correctly classified 90% of study individuals to their respective disease group, and another 14-biomarker panel correctly identified the presence of lung fibrosis in SSc patients.216 Although promising, these results need validation in a large cohort.

Biomarkers for prognosis and treatment response
The management of a patient with SSc depends greatly on disease subset, duration, and organ involvement. Full clinical assessment, including serological profile, capillaroscopy, gastroenteric assessment, chest X-ray and high-resolution tomography, pulmonary function testing, echocardiography, and electrocardiography are mandatory in every new SSc patient.

The therapeutic goals in SSc include minimization of damage from early inflammation and autoimmunity, restoration of vascular homeostasis, promotion of repair of structural connective tissue, and resolution of scarring. Treatment depends on the active problems, and immunosuppressive agents are used when there is active progressive skin or organ disease. A pivotal role is played by the local and general treatment of DUs and skin ulcers, which greatly limit the quality of life of the patients.217 Critical digital ischemia or recurrent DUs are now treated with phosphodiesterase type 5 inhibitors, prostacyclin analogs, and ET receptor antagonists.
The currently used immunosuppressive drugs for SSc act on two important aspects of the pathogenesis of the disease – inflammation and fibrosis – although, nonspecific drugs able to block or revert fibrosis have been identified so far.\textsuperscript{218} Cyclophosphamide is the only immunosuppressive agent that has been shown to be of benefit for SSc-associated lung fibrosis and skin disease. Retrospective cohort analyses and prospective open-label studies have suggested that mycophenolate mofetil may improve skin disease and lung fibrosis. In end-stage disease, lung transplantation represents a viable therapeutic option to consider for patients who have limited extrapulmonary manifestations, and for those with severe PAH when ET-1 receptor antagonists (bosentan, macitentan, ambrisentan), phosphodiesterase-5 inhibitors (sildenafil and tadalafil), and prostanooids (epoprostenol, treprostinil, and iloprost) have failed.\textsuperscript{219}

Among biologics, rituximab – an anti-CD20 mAb – has emerged as a potential treatment for SSc, improving lung function and skin scores in pilot and prospective open-label studies.

Autologous hematopoietic stem cell transplantation was investigated in a small series of patients with dcSSc who obtained a sustained improvement of skin fibrosis and stabilization of major organ function. Its utility still needs to be proven in larger clinical trials.

Although numerous exploratory studies have been published on serum and genetic markers to predict disease outcomes,\textsuperscript{68} markers to assess the response to treatment, or to drive clinical decisions, have been poorly investigated.

In a recent study of 173 SSc patients from a single center, Kumánovics et al\textsuperscript{120} investigated the predictive value of KL-6 serum levels for the outcome of ILD, including the response to treatment with cyclophosphamide. Serum levels of KL-6 significantly decreased in patients receiving cyclophosphamide treatment in spite of the fact that the spirometry results (FVC and DLCO) did not show a significant change.\textsuperscript{120}

In a recent small prospective study, plasma ET-1 and NO were assessed at baseline and after 24 weeks of treatment with bosentan in ten SSc patients with PH.\textsuperscript{220} Plasma ET-1 levels were significantly decreased after 24 weeks of bosentan therapy ($P<0.0001$), and ET-1 levels of SSc patients with PH decreased to a level comparable to that in patients without PH. Although a decrement in plasma ET-1 was induced by bosentan, plasma ET-1 levels remained high when compared with those of healthy controls.\textsuperscript{220} These findings reflect the necessity to interpret biomarkers, not on the basis of absolute cutoffs, but considering the percentage changes from baseline in single patients.

Since repeating blood sampling over time can be problematic in advanced SSc patients, noninvasive biomarkers, like exhaled compounds, are alluring. NO, a key cellular mediator, is both a powerful endogenous vasodilator and a highly reactive biomarker of inflammation. NO is produced in high quantities during alveolar inflammation.\textsuperscript{221} It has been shown that exhaled NO is related to the presence and severity of SSc-ILD.\textsuperscript{222,223} A recent study by Tiev et al\textsuperscript{224} assessed the efficacy of immunosuppressive treatment in 19 nonsmoking SSc-ILD patients, by measuring changes of exhaled NO over time after six cycles of cyclophosphamide. Patients with high levels of exhaled NO (>8.5 ppb), were the most responsive to cyclophosphamide therapy, reflecting that alveolar inflammation identifies SSc patients with a greater chance of benefitting from cyclophosphamide treatment.

Gene expression profiling has been also introduced as a tool to assess response to treatment in SSc. Chung et al\textsuperscript{225} compared gene expression patterns in two SSc patients before and after 1 month of therapy with imatinib mesylate, a tyrosine kinase inhibitor. A set of 1,050 genes modified by imatinib in both patients was demonstrated, involving multiple functional pathways (eg, cell proliferation, matrix and vascular remodeling, and growth factor signaling). Another recent, prospective study by Hinchcliff et al\textsuperscript{226} analyzed the molecular profile of skin biopsies in SSc patients treated with mycophenolate mofetil (MMF). Of the seven patients treated with MMF, four were improvers and three nonimprovers. A specific, baseline 321-gene expression signature was identified in the skin and was found to correlate with mRSS improvement during MMF treatment, but only in responders.\textsuperscript{226} DNA microarray seems to be promising in identifying potential responders to treatment, but it still needs further investigation in a larger cohort.

**True perspectives: personalized medicine in SSc**

The clinical management of patients with SSc remains a challenge and it currently involves known and newly proposed therapeutic approaches. Indeed, deeper knowledge of the cellular and molecular regulators of vasculopathy, immune response, and fibrosis has led to the identification of novel targets that possibly interrupt the vascular damage and the fibrotic process in SSc.\textsuperscript{227-229}

We suggest that the development, validation and application of some of the biomarkers we have described, singly or more likely in combination, will have a deep impact on SSc, and will play a pivotal role in early diagnosis, classification,
and management, as well as in the design of clinical trials with novel therapeutic agents.

It is worthwhile to mention that a novel treatment directed against integrins has been shown to control fibrosis in an animal model of SSC. However, combination therapy may be necessary to control the complex biological network active in SSC, leading to a control of disease progression. Finally, the knowledge of the genetic background and of mechanisms that play a major role in each patient may lead to an individually tailored treatment.

Disclosure

The authors report no conflicts of interest in this work.

References


