Whole blood gene expression profiling distinguishes systemic sclerosis-overlap syndromes from other subsets

P Moinzadeh[1], P Frommolt[2], M Franitza[3], M R Toliat[3], K Becker[3], P Nürnberg[2,3], S I Nihtyanova[4], M Ahrazoglu[1], D Belz[1], N Hunzelmann[1], D Abraham[4], V H Ong[4], L Mouthon[5], R Hesselstrand[6], C P Denton[4], T Krieg[1]

[1]Department of Dermatology and Venerology, University of Cologne, Cologne, Germany, [2]CECAD University of Cologne, Germany, [3]Cologne Center for Genomics, University of Cologne, Cologne, Germany, [4]Centre for Rheumatology and Connective Tissue Diseases, UCL Medical School, Royal Free Hospital, London, United Kingdom, [5]Université Paris Descartes, National Reference Centre for Scleroderma and Systemic Vasculitis (LM, LG), Cochin Hospital, Paris, France, [6]Department of Rheumatology, Lund University, Lund, Sweden

Corresponding author:

Pia Moinzadeh, MD

Email: pia.moinzadeh@uk-koeln.de

Telephone: +49-221-47882286; Fax: +49-221-4784538

Department of Dermatology and Venereology, University of Cologne, Kerpener Str. 62, 50937 Cologne, Germany

Letter to the editor:

Dear Editor,

Amongst the autoimmune rheumatic diseases, systemic sclerosis (SSc) is especially challenging due to its clinical diversity, reflected by the extent of skin involvement and internal organ involvement[1]. Beside the two major subsets (lcSSc and dcSSc), a substantial number of SSc cases exhibit overlap features of another distinct rheumatic disease, summarized as SSc-overlap syndromes. A few of these are already well characterized by their specific clinical course and circulating antibodies. All these patients meet the criteria of SSc showing a limited expression of skin sclerosis. They, however, should be regarded as a separate SSc subset, distinct from lcSSc and dcSSc patients[2].

We have here used gene expression profiling in whole blood cells to differentiate SSc patients from healthy controls (HC) and to identify a specific gene expression signature and predictive gene sets for SSc-overlap syndromes. Whole blood of 150 SSc patients (86 patients classified as dcSSc/lcSSc and 64 as SSc-overlap syndromes) and 40 HC, collected at four different centres, the University of Cologne, the Royal Free Hospital in London, the University of Paris and the University of Lund, have been used. Subsequently, a molecular predictor was created by pairwise comparison of the SSc patients versus SSc-overlap syndromes. To demonstrate the predictive potential of our gene set, we used the 150 patients as a training set for Support Vector Machines (SVMs). Their performance was evaluated by a leave-one-out cross validation: For each patient, an SVM was trained on the remaining 149 patients and the SVM-based prediction for the patient was compared with the true clinical subtype. This set of genes was used to distinguish SSc and SSc-overlap syndromes. We used Ingenuity pathway analysis (IPA) to collect information about signaling pathways, networks and gene functions[3].

A comparison of all SSc patients with HC (FDR<0.05, FC>1.5) revealed that in total, 41 genes were differentially expressed. Network analysis revealed that most of these genes were functionally associated with antimicrobial and inflammatory responses as well as cell cycle, cell-to-cell signaling and interaction, cell death and survival (figure 1). Canonical pathway analysis identified interferon-signaling as the most significant pathway, involving seven of our 41 differentially expressed genes (IFIT1, IFIT3, OAS1, MX1, IFI35, IFI6, ISG15; p-value=1.21E01). 59% of genes have been previously described for

connective tissue diseases, inflammatory/infectious diseases and other inflammatory dermatological diseases e.g. LE, DM, Psoriasis vulgaris.

In addition, specific, non-overlapping genes clearly differentiated SSc-overlap patients from other patients with SSc. These mainly represent genes coding for proteins involved in the control of cellular metabolisms, apoptosis, cell movement, immune cell trafficking, cell proliferation, cell-to-cell signaling and interaction, and protein synthesis, but also the humoral and adaptive immune response. The two most interesting and significant pathways were iNOS Signaling and the Toll-like Receptor Signaling, which were significantly higher expressed in the SSc overlap patients (table 1).

The SVM-based predictions enabled us to distinguish patients with SSc and SSc-overlap syndromes with a success rate of 72%. Again, most of these genes were functionally associated with cellular growth, proliferation, cell-to-cell signaling and interaction, cellular growth and proliferation, immune cell trafficking as well as cell death and survival. Of all 89 genes, 57 were directly or indirectly connected with IFN regulation (figure 1).

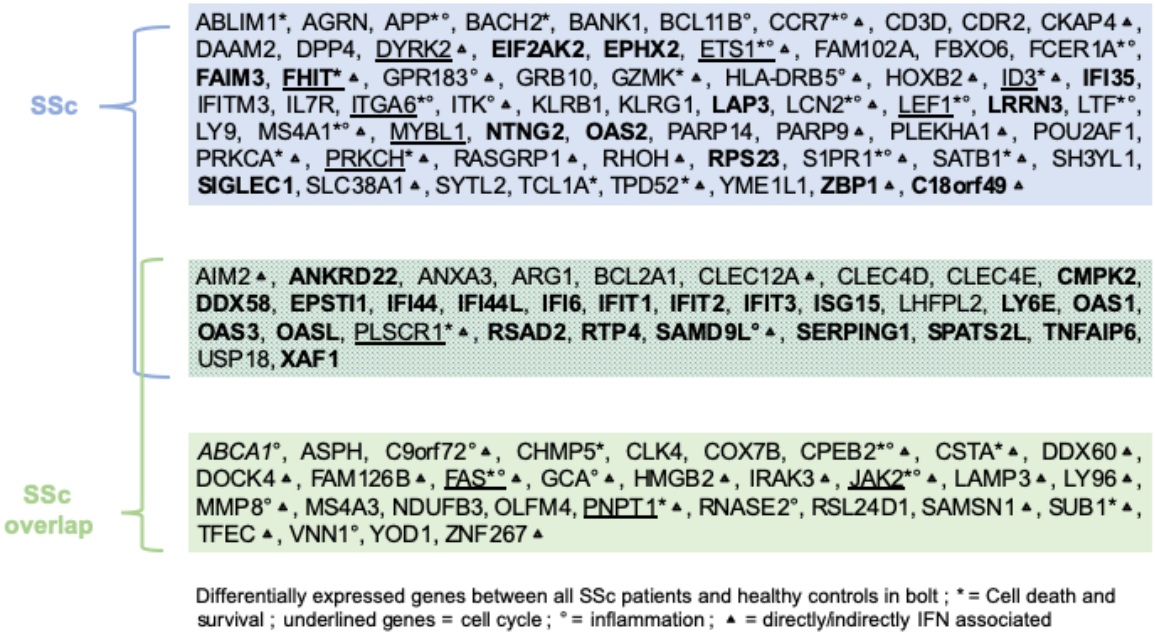
The presented differentially expressed genes of SSc patients compared to HC have shown that the interferon signature is a key player in SSc, similar to other autoimmune diseases[4].

In addition, these data clearly support the clinical view that SSc (lcSSc/dcSSc) and SSc-overlap syndromes represent separate subsets within the broader spectrum of SSc with specific genes for these two subgroups.

A key finding in our study is to define a set of predictive genes allowing the classification of patients into SSc and SSc-overlap patients based on a blood sample. Although additional validation and long term follow up studies are required this may enable us to identify patients at an early stage of the disease and to predict better the course of the disease in individual patients. This study indicates that gene expression profiling in whole blood cells has the potential to diagnose SSc patients early and to distinguish between SSc-overlap patients from the other subsets. It can therefore have a potential benefit for clinical practice and trial design to help with early case stratification and treatment choice.

Investigations of independent cohorts of patients and longitudinal studies are required to find out whether identification of certain genes regulated in the peripheral cells in the different subsets are stable or might reflect disease activity.

Figure 1: Specific differentially expressed genes for SSc (blue box; FC>1.5, FDR<0.05) and SSc-overlap subsets (green box; FC>2.0, FDR<0.05) together with their overlapping genes



Differentially expressed genes between all SSc patients and healthy controls in bolt ; * = Cell death and survival ; underlined genes = cell cycle ; ° = inflammation ; ▲ = directly/indirectly IFN associated

Table 1: Top 5 canonical pathways (ingenuity pathway analysis) for all SSc patients versus HCs, for overlapping genes in SSc and SSc-overlap syndromes and for specific genes in SSc and SSc-overlap patients.

Canonical pathway (TOP 5) SSc versus HC	Molecules	p-value	overlap
Interferone signaling	IFIT3,IFIT1,OAS1,MX1,IFI35,IFI6,ISG15	7.63E-13	19.4%, 7/36
Activation of IRF by Cytosolic Pattern Recognition Receptors	DDX58,ZBP1,IFIT2,ISG15	7.14E-06	3.8% 4/60
Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses	OAS1,OAS2,DDX58,EIF2AK2,OAS3	7.32E-06	3.8%, 5/131
Salvage Pathways of Pyrimidine Ribonucleotides	CMPK2,EIF2AK2	1.63E-02	2.1%, 2/94
Role of Lipids/Lipid Rafts in the Pathogenesis of Influenza	RSAD2	3.87E-02	5.3% 1/19
Canonical pathway (TOP 5) Overlapping in SSc and SSc-overlap vs HC	Molecules	p-value	overlap
Interferone signaling	OAS1,IFIT1,IFIT3,ISG15,IFI35,MX1,IFI6,IFITM3	2.31E-11	22.2%, 8/36
Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses	EIF2AK2,OAS1,PRKCA,OAS3,OAS2,DDX58,PRKCH	1.02E-05	5.3%, 7/131
Calcium-induced T Lymphocyte Apoptosis	PRKCA,HLA-DRB5,CD3D,PRKCH	3.84E-04	6.7%, 4/60
Activation of IRF by Cytosolic Pattern Recognition Receptors	ISG15,DDX58,ZBP1,IFIT2	3.84E-04	6.7%, 4/60
Leukocyte Extravasation Signaling	RHOH,PRKCA,ITGA6,ITK,PRKCH,RASGRP1	1.15E-03	2.9%, 6/205
Canonical pathway (TOP 5) specific SSc-overlap vs HC	Molecules	p-value	overlap
iNOS Signaling	LY96,JAK2,IRAK3	6.76E-05	7.0%, 3/43
Toll-like Receptor Signaling	LY96,IRAK3	7.43E-03	2.9%, 2/70
Oxidative Phosphorylation	COX7B,NDUFB3	1.13E-02	2.3%, 2/87
Airway Pathology in Chronic Obstructive Pulmonary Disease	MMP8,	1.29E-02	14.3%, 1/7
Type I Diabetes Mellitus Signaling	JAK2,FAS	1.39E-02	2.1%, 2/97
Canonical pathway (TOP 5) specific SSc vs HC	Molecules	p-value	overlap
Calcium-induced T Lymphocyte Apoptosis	PRKCA,HLA-DRB5,CD3D,PRKCH	5.24E-05	7.7%, 4/52
Leukocyte Extravasation Signaling	RHOH,PRKCA,ITK,ITGA6,PRKCH,RASGRP1	1.33E-04	3.0%, 6/199
NF-B Activation by Viruses	EIF2AK2,PRKCA,ITGA6,PRKCH	3.88E-04	4.6%, 4/87
Tec Kinase Signaling	RHOH,PRKCA,ITK,PRKCH,FCER1A	4.24E-04	3.1%, 5/160
iCOS-iCOSL Signaling in T Helper Cells	HLA-DRB5,CD3D,ITK,PLEKHA1	9.41E-04	3.6%, 4/110

References

1. Gabrielli A, Avvedimento EV, Krieg T. Scleroderma. *The New England journal of medicine* 2009;**360**:1989-2003.

2. Moinzadeh P, Aberer E, Ahmadi-Simab K, et al. Disease progression in systemic sclerosis-overlap syndrome is significantly different from limited and diffuse cutaneous systemic sclerosis. *Ann Rheum Dis* 2014.

3. Ganter B, Zidek N, Hewitt PR, Muller D, Vladimirova A. Pathway analysis tools and toxicogenomics reference databases for risk assessment. *Pharmacogenomics* 2008;**9**:35-54.

4. Delaney TA, Morehouse C, Brohawn PZ, et al. Type I IFNs Regulate Inflammation, Vasculopathy, and Fibrosis in Chronic Cutaneous Graft-versus-Host Disease. *J Immunol* 2016;**197**:42-50.