

# The Circulating Cell-free microRNA Profile in Systemic Sclerosis Is Distinct from Both Healthy Controls and Systemic Lupus Erythematosus

Samantha O. Steen, Line V. Iversen, Anting Liu Carlsen, Mark Burton, Christoffer T. Nielsen, Søren Jacobsen, and Niels H.H. Heegaard

**ABSTRACT. Objective.** To evaluate the expression profile of cell-free circulating microRNA (miRNA) in systemic sclerosis (SSc), healthy controls (HC), and systemic lupus erythematosus (SLE).

**Methods.** Total RNA was purified from plasma and 45 different, mature miRNA were measured using quantitative PCR assays after reverse transcription. Samples (n = 189) were from patients with SSc (n = 120), SLE (n = 29), and from HC (n = 40). Expression data were clustered by principal components analysis, and diagnostically specific miRNA profiles were developed by leave-one-out cross-validation. Diagnostic probability scores were derived from stepwise logistic regression.

**Results.** Thirty-seven miRNA specificities were consistently detected and 26 of these were unaffected by SSc sample age and present in more than two-thirds of SSc samples. SSc cases showed a distinct expression profile with 14/26 miRNA significantly decreased (false discovery rate < 0.05) and 5/26 increased compared with HC. A 21-miRNA classifier gave optimum accuracy (80%) for discriminating SSc from both HC and SLE. The discrimination between HC and SSc (95% accuracy) was strongly driven by miRNA of the 17-92 cluster and by miR-16, -223, and -638, while SLE and SSc differed mainly in the expression of miR-142-3p, -150, -223, and -638. Except for a weak correlation between anti-Scl-70 and miR-638 ( $p = 0.048$ ), there were no correlations with other patient variables.

**Conclusion.** Circulating miRNA profiles are characteristic for SSc compared with both HC and SLE cases. Some of the predicted targets of the differentially regulated miRNA are of relevance for transforming growth factor- $\beta$  signaling and fibrosis, but need to be validated in independent studies. (First Release Nov 15 2014; J Rheumatol 2015;42:214-21; doi:10.3899/jrheum.140502)

*Key Indexing Terms:*

SYSTEMIC SCLEROSIS

microRNA

CLASSIFICATION

SYSTEMIC LUPUS ERYTHEMATOSUS

DIAGNOSIS

Systemic sclerosis (SSc) is a rare autoimmune disease characterized by chronic inflammation of connective tissues associated with cutaneous and organ fibrosis, occlusive vasculopathy, specific types of circulating autoantibodies, and a highly variable clinical course. The etiology of SSc is unknown, but perivascular inflammation is an early and central pathogenetic event and leads to aberrant activation

or increased synthesis of various collagen cross-linking enzymes and subsequent increased synthesis/cross-linking and decreased degradation of collagen<sup>1,2,3,4</sup>. The class of noncoding small RNA called microRNA (miRNA) regulates the expression of most protein-coding genes at the post-transcriptional level by interfering with translation of target mRNA. Typically, miRNA have multiple mRNA

*From the Department of Clinical Biochemistry, Immunology and Genetics, Statens Serum Institut; Department of Dermatology, Bispebjerg Hospital, and the Department of Rheumatology, Rigshospitalet, University of Copenhagen, Copenhagen; Department of Clinical Genetics, and the Department of Clinical Biochemistry and Pharmacology, Odense University Hospital, Odense; Research Unit of Human Genetics, and Clinical Biochemistry, Institute of Clinical Research, University of Southern Denmark, Odense, Denmark.*

*Supported by the Foundation for the Advancement of Medical Research, Bang's Foundation, and the Danish Rheumatism Association (R99-A1937).*

*S.O. Steen, MSc; A.L. Carlsen, PhD, Department of Clinical Biochemistry, Immunology and Genetics, Statens Serum Institut; L.V. Iversen, MD, PhD, Department of Dermatology, Bispebjerg Hospital, University of*

*Copenhagen; M. Burton, PhD, Department of Clinical Genetics, Odense University Hospital, and the Institute of Clinical Research, Research Unit of Human Genetics, University of Southern Denmark; C.T. Nielsen, MD, PhD; S. Jacobsen, MD, DMedSc, Department of Rheumatology, Rigshospitalet, University of Copenhagen; N.H.H. Heegaard, MD, DMedSc, DNatSc, Department of Clinical Biochemistry, Immunology and Genetics, Statens Serum Institut, and the Department of Clinical Biochemistry and Pharmacology, Odense University Hospital, and the Institute of Clinical Research, Clinical Biochemistry, University of Southern Denmark.*

*Address correspondence to Dr. N. Heegaard, Department of Autoimmunology and Biomarkers, Statens Serum Institut, Artillerivej 5, DK-2300 Copenhagen S, Denmark. E-mail: nhe@ssi.dk  
Accepted for publication October 1, 2014.*

targets in inflammatory and fibrosis pathways (including growth factors, enzymes, and cytokines), are themselves regulated by inflammatory cytokine/cytokine receptor networks<sup>5</sup>, and are dysregulated in cells of the innate and adaptive immune system from patients with autoimmune disorders<sup>6,7,8</sup>. Thus, specific miRNA have been shown to be differentially expressed in fibrotic diseases such as SSc (miRNA-29)<sup>9</sup> and idiopathic pulmonary fibrosis<sup>10</sup> (the miRNA-17~92 cluster). The miRNA-17~92 cluster targets, for example, fibrotic and antiangiogenic genes and DNA methyltransferase (DNMT)-1 expression and is downregulated in fibrotic lung tissue. miRNA-29, which also targets DNMT-1 expression, is depressed in SSc dermal fibroblasts and regulates type I and III collagen expression<sup>9</sup>. Differential expression of miRNA was also noted in a study of sclerodermic skin biopsies<sup>11</sup>. Expression profiles of miRNA may thus be associated with diagnosis and classification of SSc and constitute potential therapeutic targets. From a diagnostic viewpoint, the presence of a population of stable, circulating miRNA in the blood<sup>12,13</sup> that may include specific contributions from pathological cells and tissues in cancer and autoimmunity is valuable, and studies have demonstrated specific profiles associated with systemic autoimmune disease such as systemic lupus erythematosus (SLE)<sup>14</sup>. Because of the unique dominance of fibrosis in SSc, we hypothesized that the circulating miRNA profile in this condition might be characteristic and different from those in other systemic autoimmune diseases and possibly related to subtypes of SSc. To address this hypothesis, we investigated levels of a panel of 45 specific miRNA in plasma samples collected under controlled conditions from a cohort of 120 patients with SSc, 29 patients with SLE, and 40 healthy controls (HC).

## MATERIALS AND METHODS

**Patients and methods.** Our cross-sectional study included 120 patients with SSc with a median age of 57 years (range 22–79 yrs) fulfilling the American College of Rheumatology criteria<sup>15</sup> for SSc (Table 1). Patients were included consecutively over 2 years. Skin involvement was graded according to the modified Rodnan skin score<sup>16</sup>. Seven patients with SSc were receiving disease-modifying antirheumatic drugs (DMARD; azathioprine, methotrexate, or anti-malarials) and 6 were taking prednisolone (together with DMARD in 4 cases) at the time of inclusion. None of the patients had ever received therapy with biological drugs or prostanoid infusions within 6 months of inclusion. Concomitant treatment included calcium channel blockers, proton pump inhibitors, angiotensin-converting enzyme inhibitors, diuretics, statins, and others. None of the patients had known cancer or systemic infections, and none were pregnant at the time of inclusion. With the exception of 1 patient, all were white. As disease controls, 29 white patients with SLE (4 of these were men) were included from another study<sup>14</sup>. The SLE samples were collected within a few months and were subsequently stored as citrated platelet-poor plasma at –80°C for about 3.5 years before being used in our present study. Finally, 40 unmedicated, healthy individuals (33 women and 7 men) with a median age of 46 years (range 24–71 yrs) were included as HC and sampled within a few weeks in 2 rounds (36 HC in the first round) separated by a year. All participants were included after giving written informed consent. The study was approved by the local ethics committee (approval no.

H-B-2008-131) and carried out in accordance with the principles of the Declaration of Helsinki.

**Clinical, biochemical, and serological assessment.** The clinical characteristics of the patient population are summarized in Table 1. Data on disease history including clinical, biochemical, and pharmacological data were obtained by review of medical records, patient examination, and interview. Subsequently, the data were registered in a database. Based on the extent of cutaneous involvement, patients were classified as having limited cutaneous SSc (lcSSc; n = 79) or diffuse cutaneous SSc (dcSSc; n = 41)<sup>17</sup>. Serological assessment of antinuclear antibodies, including anticomplement antibodies (ACA), was performed by indirect immunofluorescence on human epithelial type 2 cell line cells, and anti-Scl-70 antibodies were demonstrated by ELISA.

**Blood sampling and isolation of platelet-poor plasma.** Venipuncture was performed using a 21-gauge needle, and after release of the tourniquet, the first tube obtained was always reserved for autoantibody analyses. Next, for microRNA analysis, the blood was collected into 3 × 9 ml citrate tubes (Vacuette sodium citrate 3.8%; Greiner Bio-One) and gently mixed 5 times. Immediately after collection, the blood cells were removed by centrifugation. All samples underwent a 2-step centrifugation procedure: 1800 g, 10 min, 21°C, followed by centrifugation of the supernatant 3000 g, 10 min, 21°C to obtain platelet poor plasma (PPP). The PPP was aliquoted, snap-frozen in liquid nitrogen, and stored at –80°C until analysis.

**RNA isolation and miRNA profiling.** Total RNA was purified from 100 µl-citrated PPP samples using Norgen Total RNA Purification Kit (Norgen Biotek Corp.), and a panel of miRNA was analyzed after reverse transcription using stem-loop primers, preamplification, and quantitative PCR (qPCR) using specific assays from Applied Biosystems. For the qPCR, a BioMark MX microfluidic platform (Fluidigm Corp.) allowing duplicate assays for 48 miRNA in 96 samples on 96.96 dynamic arrays in 1 operation was used, as detailed elsewhere<sup>18</sup>. The 48 miRNA assays included assays for 45 human and 3 *Caenorhabditis elegans* miRNA (cel-miR-39, -54, and -238; Supplementary Table 1 available online at jrheum.org). The mixture of 3 synthetic cel-miRNA was added to the lysis buffer for use as spike-in controls for technical normalization, as previously reported<sup>18</sup>.

**Data handling.** Only assays yielding Cq values below 30 were included, and assays for 3 miRNA (miR-129-5p, -210, and -31-5p) consistently failed, leaving a total of 42 miRNA for data analysis. Technical (using the cel-miRNA average) and row normalization using the average of miRNA (n = 20) detected in all samples was performed. The resulting  $\Delta\Delta Cq$  values increase with increasing miRNA levels. Individual miRNA levels were expressed in a linear form using the  $2^{-\Delta\Delta Cq}$  transformation, and SLE, controls, and SSc cases were compared using 1-way ANOVA with Tukey posttest, using p = 0.05 as the significance level. For multiple comparison, significance levels were corrected using false discovery rates (FDR)<sup>19</sup>. Principal components analysis (PCA) was based on row-normalized  $\Delta\Delta Cq$  values using Genesis (v.1.7.6.; genome.tugraz.at/genesisclient/genesisclient\_description.shtml).

Risk scores for SSc were derived from optimized gene profiles, i.e., the set of miRNA that resulted in maximal balanced accuracy (mean of specificity and sensitivity) for the SSc diagnosis. Three miRNA profiles were developed, distinguishing SSc from SLE, SSc from HC, and SSc from the rest (SLE and HC). The profiles were obtained using logistic regression classification from which optimal profiles were selected by leave-one-out cross-validation (LOOCV). Briefly, in this procedure, a single sample serves as a test sample and the remaining samples as a training set. This is repeated until all samples have been left out once and the accuracy of the classifier is determined by the correctly classified samples. In the training set, feature selection is necessary to avoid a small sample-per-feature ratio and provide better classification. The feature selection procedure consisted of 3 steps: (1) testing the miRNA in the training set for significance using univariate logistic regression; (2) ranking the miRNA features by their significance; and (3) finding the optimal number of features by sub-

Table 1. Demographic and clinical characteristics of the 120 patients with SSc. Arterial hypertension: blood pressure above 140/90 mmHg at study inclusion.

Characteristics	All	lcSSc	dcSSc
<b>Patients</b>			
Sex, n, males/females	19/101	3/76	16/25
Age, yrs, mean ± SD (range)	57 ± 12 (22–79)	59 ± 11	55 ± 12
Disease duration, yrs, mean ± SD (range)	12 ± 9 (0–53)	12 ± 9	11 ± 9
<b>Cutaneous involvement</b>			
mRSS, mean ± SD	11 ± 9	7 ± 3	20 ± 9
Cutaneous ulcers, n	22	10	12
Calcinosis, n (N)*	43 (107)	30	13
<b>Pulmonary and cardiac involvement</b>			
DLCO, mean ± SD % predicted (N)*	65.1 ± 19.2 (118)	64.2 ± 18.5 (77)	66.2 ± 20.5 (41)
FVC, mean ± SD % predicted (N)*	95.6 ± 21.2 (118)	99.3 ± 20.1 (77)	88.8 ± 21.8 (41)
Radiographic confirmed lung fibrosis, n	21	14	7
Elevated SPAP by echocardiography, n (N)*	10 (115)	6	4
<b>Serology, n</b>			
Antinuclear antibodies	115	74	41
Anti-Scl-70 antibodies	15	3	12
Anticentromere antibodies	48	41	7
DMARD	7	4	3
Current or previous smokers	82	57	25
Arterial hypertension	30	18	12
<b>Blood samples</b>			
Leukocytes, × 10 <sup>9</sup> /l, mean ± SD	7.2 ± 2.3	7.4 ± 2.5	7.7 ± 2.4
Neutrophils, × 10 <sup>9</sup> /l, mean ± SD	4.6 ± 1.9	4.7 ± 2.0	4.9 ± 2.2
Platelets, × 10 <sup>9</sup> /l, mean ± SD	272 ± 85	265 ± 86	269 ± 61
Lactate dehydrogenase, U/l, mean ± SD	186 ± 37	191 ± 40	185 ± 37

\* Reduced number of patients because of lack of information. N: total number of subjects in the group; SSc: systemic sclerosis; lcSSc: limited cutaneous SSc; dcSSc: diffuse cutaneous SSc; mRSS: modified Rodnan Skin score; FVC: forced vital capacity; SPAP: systolic pulmonary artery pressure; DMARD: disease-modifying antirheumatic drug.

quently adding 1 feature at a time in a “top-down” forward wrapper approach starting with the top 2 features of the ranked list; at each increment, the classification accuracy of the training samples was assessed using LOOCV. All calculations were performed using the open source R-environment (cran.r-project.org). The function “generalized linear model” was used for logistic regression classification, and regression coefficients were extracted from the summary (model) function.

## RESULTS

**Expression data and PCA.** Five of the 42 analyzable hsa-miRNA (let-7a-5p, miR-10a, -101, -125a-3p, -196a) were inconsistently detected (i.e., present only in less than a third of all samples) and therefore disregarded in the further analysis. The expression values of the remaining 37 miRNA (sorted according to abundance in HC in Supplementary Table 2, available online at jrheum.org) were used as input for a PCA. In the PCA score plot (Figure 1), HC (x) clustered together with a subset of the SSc samples (white dots) at high PC2 scores (i.e., PC2 > 0.05), whereas the remaining SSc samples were below this level and distinct from HC samples. PC2 is the second principal component, i.e., the weighted linear combination of all variables that explain the second-most of the variance of the multivariable data. No apparent separation of SSc and SLE samples was observed in this plot along either the PC1 or PC2 axis. Overall, the PCA suggested that the data contained infor-

mation that discriminated some SSc (and SLE) samples from HC, but that a distinct subset of SSc samples grouped with HC. Both HC and SLE samples were collected over short time intervals (the HC over 2 short periods) as compared to a considerably longer sampling interval for the SSc samples. Therefore, we next correlated PC2 scores with sampling date (Figure 2). From this analysis, it was apparent that SSc samples with a PC2 > 0.05, were the oldest and that there was a significant ( $p < 0.0001$ ,  $R^2 = 0.64$ ) linear correlation between sample age and PC2. Interestingly, both HC sample groups were located at PC2 > 0.05 while the SLE samples that were as old as the oldest SSc samples predominantly grouped below PC2 = 0.05. Thus, the combination of variables (miRNA expression values) that constitutes PC2 did not correlate with the age of the SLE samples.

Because of the noted influence of sample age on the SSc data variance, we next analyzed data for each miRNA independently to ascertain which miRNA contributed most to the sample age-dependent variability. Results are summarized in Table 2 and Supplementary Figure 1 (available online at jrheum.org). These data show that 14 miRNA were significantly decreased (FDR < 0.05; marked with downward arrows in Table 2) in SSc samples irrespective of sample age in comparison with HC, while 5 miRNA were significantly increased (upward arrows). The remaining

Table 2. Overview of 35 miRNA in SSc and HC samples, ranked according to fold-change (SSc/HC). miR-18a-5p and miR-181a are not included because of too few data. Rank is numbered from most (1) to least significant (35).

miRNA	Fold-change	Rank	FDR	Note
hsa-miR-29a	9.190	17	0.0002	*
hsa-miR-203	3.053	19	0.0002	*
hsa-miR-409-3p	2.585	10	0.0002	*
hsa-miR-181b	1.625	5	0.0002	↑
hsa-miR-638	1.613	4	0.0002	↑
hsa-miR-29b-3p	1.516	21	0.0009	↑
hsa-miR-590-5p	1.385	23	0.0041	↑
hsa-miR-150	1.206	28	0.0391	↑
hsa-miR-155	1.021	34	0.8232	NS
hsa-miR-383	0.986	35	0.9532	NS
hsa-miR-34a	0.853	33	0.4121	NS
hsa-miR-15a	0.807	30	0.2028	NS
hsa-miR-125b	0.753	31	0.2205	NS
hsa-let-7c	0.737	29	0.0831	NS
hsa-miR-142-3p	0.660	22	0.0028	↓
hsa-miR-208	0.637	32	0.3588	NS
hsa-miR-21	0.578	24	0.0041	*
hsa-miR-423-5p	0.525	27	0.0392	*
hsa-miR-27a-3p	0.493	20	0.0005	*
hsa-miR-342-3p	0.490	13	0.0002	↓
hsa-miR-146a	0.483	3	0.0002	↓
hsa-miR-20a	0.473	18	0.0002	↓
hsa-miR-375	0.463	25	0.0159	*
hsa-miR-29c-3p	0.448	15	0.0002	*
hsa-miR-92a	0.429	16	0.0002	↓
hsa-miR-145-5p	0.412	26	0.0159	↓
hsa-miR-16	0.285	14	0.0002	↓
hsa-miR-223	0.274	2	0.0002	↓
hsa-miR-192-5p	0.199	7	0.0002	↓
hsa-miR-221	0.075	11	0.0002	↓
hsa-miR-24	0.060	1	0.0002	↓
hsa-miR-146b	0.059	8	0.0002	↓
hsa-miR-17	0.052	12	0.0002	↓
hsa-miR-106a	0.041	6	0.0002	↓
hsa-miR-184	0.001	9	0.0002	*

\* Data values distribute according to sample age and are not included in the diagnostic modeling. miRNA: microRNA; SSc: systemic sclerosis; HC: healthy controls; NS: non-significant [FDR: (false discovery rate) > 0.05].

miRNA either showed clear SSc sample age-related correlations (miR-184, -21, -27a-3p, -29a, -29c-3p, -203, -375, -409-3p, 423-5p) affecting significance levels or were non-significantly differently expressed in the sample sets. In 2 cases (miR-18a-5p, -181a) there were too few (1/3 or more missing) data values in the SSc group to warrant inclusion in Table 2.

*A circulating miRNA profile for SSc.* Of the 35 miRNA included in Table 2, we excluded the 9 miRNA with expression levels that were clearly influenced by sample age. The remaining dataset of 26 miRNA was systematically tested for the best combination of miRNA, enabling discrimination between the SSc, the SLE, and the HC groups using logistic regression combined with optimizing

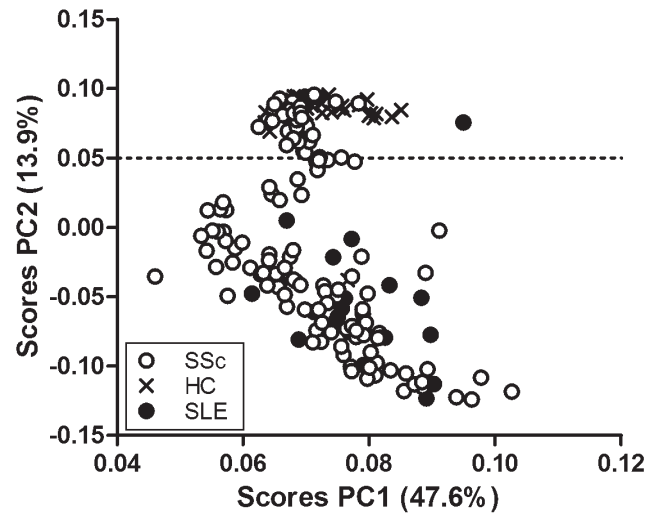


Figure 1. Principal components analysis score plot to visualize the degree of sample expression data variation that correlates with sample grouping (HC, SSc, and SLE). PC 1 and 2, together accounting for more than 60% of the variation of the data, are shown. A line is indicated at PC2 = 0.05, above which all normal samples, except 1, are located. HC: healthy controls; SSc: systemic sclerosis; SLE: systemic lupus erythematosus; PC: principal component.

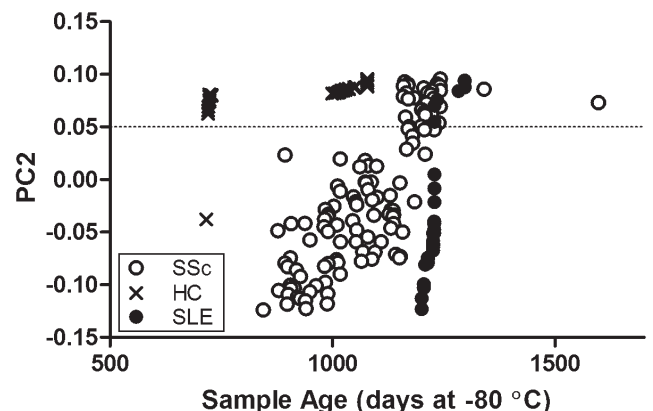
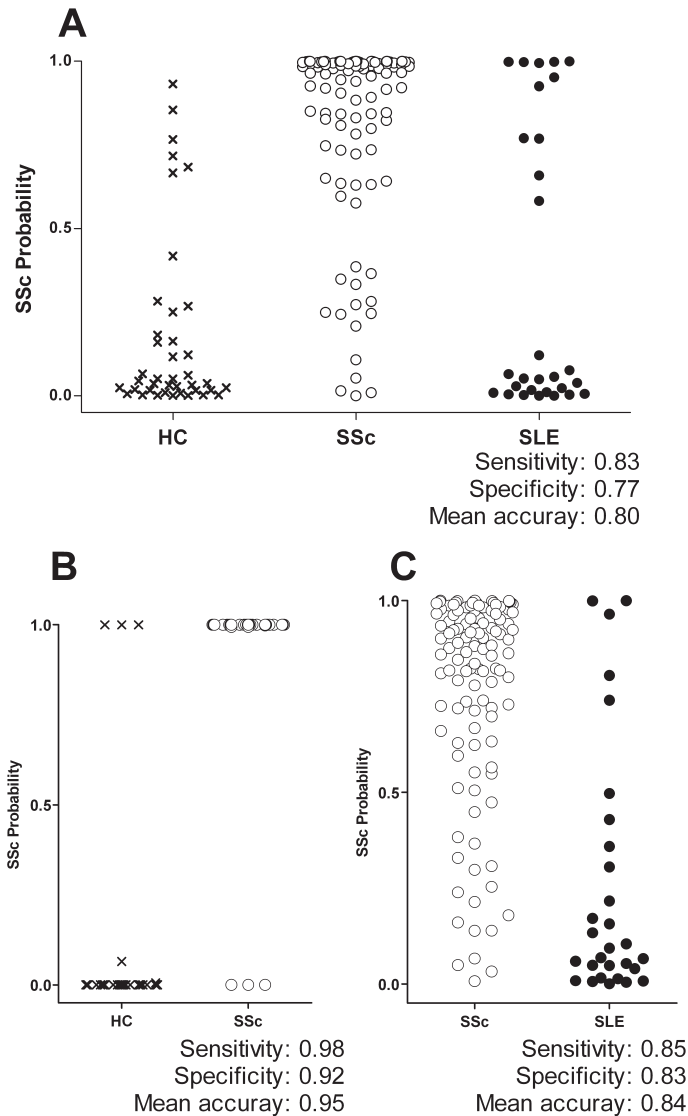


Figure 2. The contribution of sample age to the PC2 scores. SSc samples display a systematic linear correlation ( $R^2 = 0.64$ ) between sample age and PC2 scores, indicating a dependency of some miRNA levels with sample age. PC: principal component; SSc: systemic sclerosis; HC: healthy controls; SLE: systemic lupus erythematosus; miRNA: microRNA.

the number of features (miRNA) by adding 1 at a time from the top of the significance list. Each step was validated for the highest diagnostic accuracy (best mean of sensitivity and specificity) by an LOOCV approach. Briefly, in this procedure, a single sample serves as a test sample and the remaining samples as a training set. This is repeated until all samples have been left out once and the accuracy of the classifier is then determined by the correctly classified samples. Risk scores (disease probability) were developed for distinguishing SSc from SLE, SSc from HC, and SSc from the rest (SLE and HC). Probability plots based on the optimized classifiers are shown in Figure 3, and lists of

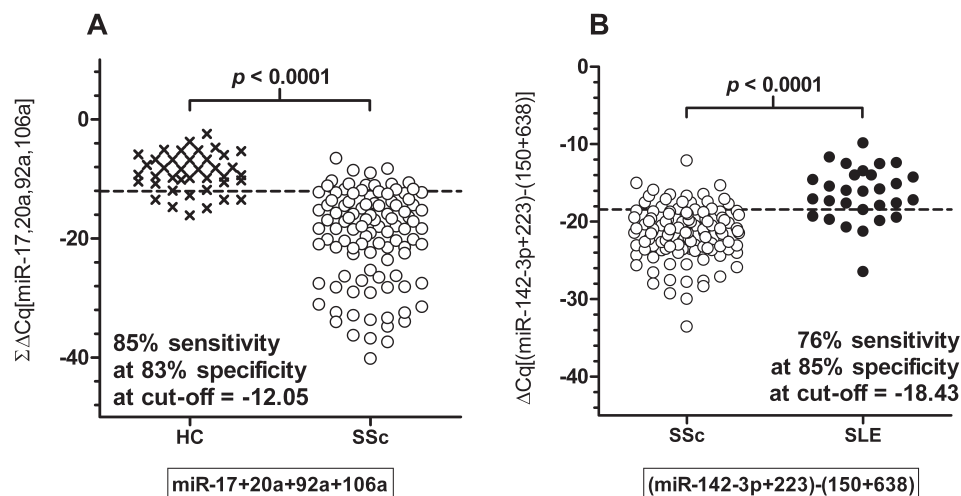


**Figure 3.** Probability plots showing the optimized diagnostic classification based on miRNA combinations derived by logistic regression to give the highest balanced diagnostic accuracy. Risk score equations are listed in Supplementary Table 3 (available online at jrheum.org). A. SSc versus all other samples (HC and SLE) based on 21 miRNA. B. SSc versus HC based on 22 miRNA. C. SSc versus SLE based on 14 miRNA. Specificity, sensitivity, and mean accuracy are given for all 3 models. miRNA: microRNA; SSc: systemic sclerosis; HC: healthy controls; SLE: systemic lupus erythematosus.

miRNA profiles (listed in order of decreasing univariate significance) and equations for risk scores are included in Supplementary Table 3 (available online at jrheum.org). The results showed that the optimal combination of miRNA for discriminating SSc cases from both SLE cases and HC consists of a 21 miRNA classifier yielding an analytical accuracy of 80% (Figure 3A). Looking at SSc versus HC and SSc versus SLE separately (Figure 3B and Figure 3C), the corresponding miRNA panels consisted of 22 and 14 miRNA, respectively. Using the respective scores in each

case gave receiver-operation characteristic (ROC) curves with an area under the curve (AUC) of > 0.95 (data not shown), and diagnostic accuracies of 0.95 for SSc versus HC and 0.84 for SSc versus SLE.

The main contributors to the diagnostic classifications were the miRNA-17~92 cluster members (miR-17, -20a, -92a, and -106a, down in SSc) that strongly separate SSc and HC (Figure 4A) with an AUC of the ROC curve of 0.92. The opposite directions of the changes in miR-142-3p and miR-223 (down in SSc, up in SLE) and miR-150 and



**Figure 4.** Simplified risk score plots based on 4 miRNA for (A) HC versus SSc and (B) SLE versus SSc. Mean accuracies are 0.84 and 0.81, respectively. The corresponding values for the optimized models (Figure 3) are 0.95 and 0.84, respectively. The stated p values are based on Mann-Whitney U tests. Optimal cutoff levels based on ROC curves (not shown) are indicated by dashed lines, and give the specificity and sensitivity performance indicated in the figure. miRNA: microRNA; HC: healthy controls; SSc: systemic sclerosis; SLE: systemic lupus erythematosus; ROC: receiver-operation characteristic.

miR-638 [up in SSc, down (miR-150) or unaltered (miR-638) in SLE] made these 4 miRNA strong classifiers for SSc compared with SLE (Figure 4B) with an AUC of the ROC curve of 0.87. Compared to the optimized profiles, these simplified 4 miRNA models lost comparatively little diagnostic power (accuracies of 0.84 vs 0.95 and 0.81 vs 0.84 for SSc-HC and SSc-SLE, respectively).

*Correlation of circulating miRNA profiles with clinical and paraclinical variables.* We found no correlations between miRNA levels and patients with SSc with different auto-antibody profiles (ACA-positive patients, n = 48/120; anti-Scl-70-positive patients, n = 15/120), except for a weakly significant difference (p = 0.048, Mann-Whitney U test) between a decreased level of miR-638 in the anti-Scl-70-positive group compared with the anti-Scl-70-negative group. The lcSSc and the dcSSc groups could not be discriminated on the basis of their miRNA-expression profiles. The cohort includes a total of 9 patients in treatment with DMARD and/or prednisolone. We did not observe any specific miRNA expression differences in this small subset of cases compared with the rest of the cohort (data not shown).

## DISCUSSION

To the best of our knowledge, our present study is the first investigation of circulating miRNA profiles in plasma samples from a cohort of more than 100 (n = 120) well-characterized patients with SSc compared with both HC and SLE cases. More than 2500 human miRNA are now known (www.mirbase.org), but less than 25% of these have been found in the circulation<sup>20</sup>, and the costs and amount of

RNA required to analyze each sample on individual printed arrays are prohibitive. Instead, we used a dynamic array for qPCR that enabled the analysis of many samples for a subset of miRNA. We chose a panel of 45 different miRNA based on previous studies of cellular and tissue miRNA in SSc and fibrosis and circulating miRNA in SLE and inflammation<sup>5,9,10,11,14,21,22</sup>, and found 35 of these to be present at measurable levels across samples. Despite a clear correlation of some of these miRNA with sample age-related factors in the SSc cohort, we were able, after eliminating these miRNA, to delineate circulating miRNA profiles that were highly characteristic of our SSc population. In the absence of an independent validation cohort, we applied a data analysis strategy with systematic internal cross validation (leave-one-out) to optimize models. In this way, highly diagnostically accurate models based on miRNA combinations could be extracted from data. In simplified terms, a combination of the miRNA-17~92 cluster that is significantly decreased in SSc and SLE, and miR-142-3p and -223 that in our study is decreased in SSc but increased in SLE, clearly discriminates between HC and the disease groups and between SSc and SLE.

Our finding of a significantly decreased level of miR-142-3p in SSc is in disagreement with a study of Japanese patients with SSc where the level of this miRNA was found to be increased even in comparison with SLE cases<sup>21</sup>. A very important difference, however, is that the study used sera while we here used plasma samples obtained under stringent and consistent conditions. While the overall ranking of abundance may correlate reasonably well for a number of miRNA in plasma and serum<sup>12</sup>, it has been

shown in several studies that the levels of individual miRNA in serum and plasma samples from the same patients do not correlate well<sup>18,23</sup> and indeed, that cellular miRNA content overwhelmingly influences the levels of supposedly cell-free miRNA (including the exosomal miRNA) in samples experiencing hemolysis or cell contamination or platelet activation, as will occur in coagulating blood<sup>20,24,25</sup>.

Interestingly, the miR-17~92 cluster has been linked with fibrosis in models of idiopathic pulmonary fibrosis, where it was shown that enhanced expression of the cluster attenuated fibrosis<sup>10</sup>. This cluster of miRNA is also known to be important for oncogenesis<sup>26</sup> and suppresses proliferation and angiogenesis in some types of cancers, but is carcinogenic in others<sup>27,28</sup>. A role in attenuating fibrosis and angiogenesis is consistent with our finding of decreased levels of 4 members of this cluster in the circulation of patients with SSc. Evidently, this cluster regulates the expression of numerous other genes in health and disease<sup>29</sup>, and its expression is also decreased in other autoimmune diseases such as SLE. Nevertheless, 2 very interesting targets of miR-17~92 were found in a study of the effects of miR-17-92 transfection on lung epithelial cell lines<sup>30</sup>. The most downregulated gene (with an mRNA predicted target sequence for miR-17-5p) was procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2, the gene encoding collagen telopeptide lysyl hydroxylase 2 (LH2). Also, as an additional target, the oxygen-dependent transcription factor HIF-1 $\alpha$  was established. LH2 has been established as a key profibrotic enzyme responsible for hydroxylation of collagen telopeptides participating in the increased pyridinoline cross-links in collagen in fibrosis<sup>2</sup>, and HIF-1 $\alpha$  is responsible for inducing the expression of a number of collagen prolyl and lysyl hydroxylases in fibroblasts under hypoxic conditions<sup>31</sup>. Decreased inhibition of the expression of these genes as a consequence of decreased miR-17~92 would thus be associated with increased fibrosis, including an increased fibrotic response to tissue hypoxia. While consistent hypotheses may thus be generated by the findings of our study, we do not, however, prove that the observed systemic decrease in the levels of the miR-17~92 cluster reflects the pathophysiological events in the skin and other sclerosing tissues and organs in SSc.

For miR-29a, which was previously reported to be downregulated in fibroblasts from patients with SSc<sup>9</sup>, we found an overall pattern of upregulation in the circulation of patients with SSc and SLE, even though to a markedly lesser extent in old SSc samples; this miRNA is not included in the diagnostic panel in our present study. The miR-29 family (29a, 29b, 29c) is involved in the regulation of DNA methylation<sup>32</sup> and it may be hypothesized that its systemic increase in patients with SSc is a response aimed at silencing gene expression important for SSc tissue pathology. Support of this notion requires further experimental work, and the

decrease of miR-29 observed in fibroblasts may simply be a secondary effect of suppressive regulation of miR-29 by transforming growth factor (TGF)- $\beta$ <sup>9</sup>.

The differential response of miR-142-3p (up in SLE, down in SSc compared with HC) is interesting, especially because of its known involvement in the regulation of expression of proteins important for TGF- $\beta$  signaling. Together with other growth factors and cytokines, TGF- $\beta$  is a skin fibroblast activator and a major inducer of fibrosis in SSc<sup>33</sup>. The expression of TGF- $\beta$  receptor I and a number of other genes related to TGF- $\beta$  signaling is predicted and experimentally verified to be directly regulated by miR-142-3p<sup>34,35</sup>. An increased TGF- $\beta$  signaling activity in SSc fibroblasts<sup>36,37</sup> and a decreased production of TGF- $\beta$  by lymphocytes from patients with SLE<sup>38</sup> may be consistent with the notion of a dysregulation of miR-142-3p in opposite directions in SSc and SLE, but this also requires further studies to be verified.

In the setting of early or undifferentiated connective tissue disease, the miRNA-based biomarker profile reported here may serve to identify patients likely to progress to SSc. However, the potential practical utility of the findings is limited because intra- and interindividual variability, reference ranges, longitudinal stability, preanalytical variables, and analytical variability are still largely unknown for circulating miRNA. Also, a simple method for direct quantitative real-time PCR analysis of miRNA in plasma does not yet exist. In addition, there is a need for validation in independent studies.

A specific profile of circulating miRNA is found in SSc plasma. Common denominators for the targets of these miRNA are TGF- $\beta$  signaling and enzymes participating in collagen cross-linking. Involvement of these miRNA in the pathogenesis of SSc should be experimentally verified, including studies of expression in affected versus unaffected tissues, and their validity as markers of SSc requires investigation in independent cohorts of patients with SSc and controls.

## ONLINE SUPPLEMENT

Supplementary data for this article are available online at [jrheum.org](http://jrheum.org).

## REFERENCES

1. Hachulla E, Launay D. Diagnosis and classification of systemic sclerosis. *Clin Rev Allergy Immunol* 2011;40:78-83.
2. van der Slot AJ, Zuurmond AM, Bardeol AF, Wijmenga C, Pruijs HE, Silence DO, et al. Identification of PLOD2 as telopeptide lysyl hydroxylase, an important enzyme in fibrosis. *J Biol Chem* 2003;278:40967-72.
3. van der Slot AJ, Zuurmond AM, van den Bogaardt AJ, Ulrich MM, Middelkoop E, Boers W, et al. Increased formation of pyridinoline cross-links due to higher telopeptide lysyl hydroxylase levels is a general fibrotic phenomenon. *Matrix Biol* 2004;23:251-7.
4. van der Slot AJ, van Dura EA, de Wit EC, De Groot J, Huizinga TW, Bank RA, et al. Elevated formation of pyridinoline cross-links by profibrotic cytokines is associated with enhanced lysyl hydroxylase 2b levels. *Biochim Biophys Acta* 2005;1741:95-102.

5. Nakashima T, Jinnin M, Yamane K, Honda N, Kajihara I, Makino T, et al. Impaired IL-17 signaling pathway contributes to the increased collagen expression in scleroderma fibroblasts. *J Immunol* 2012;188:3573-83.
6. O'Connell RM, Rao DS, Chaudhuri AA, Baltimore D. Physiological and pathological roles for microRNAs in the immune system. *Nat Rev Immunol* 2010;10:111-22.
7. Xiao C, Rajewsky K. MicroRNA control in the immune system: basic principles. *Cell* 2009;136:26-36.
8. Sonkoly E, Pivarsci A. Advances in microRNAs: implications for immunity and inflammatory diseases. *J Cell Mol Med* 2009;13:24-38.
9. Maurer B, Stanczyk J, Jungel A, Akhmetshina A, Trenkmann M, Brock M, et al. MicroRNA-29, a key regulator of collagen expression in systemic sclerosis. *Arthritis Rheum* 2010;62:1733-43.
10. Dakhllallah D, Batte K, Wang Y, Cantemir-Stone CZ, Yan P, Nuovo G, et al. Epigenetic regulation of miR-17~92 contributes to the pathogenesis of pulmonary fibrosis. *Am J Respir Crit Care Med* 2013;187:397-405.
11. Li H, Yang R, Fan X, Gu T, Zhao Z, Chang D, et al. MicroRNA array analysis of microRNAs related to systemic sclerosis. *Rheumatol Int* 2012;32:307-13.
12. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, et al. Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci U S A* 2008;105:10513-8.
13. Kroh EM, Parkin RK, Mitchell PS, Tewari M. Analysis of circulating microRNA biomarkers in plasma and serum using quantitative reverse transcription-PCR (qRT-PCR). *Methods* 2010;50:298-301.
14. Carlsen AL, Schetter AJ, Nielsen CT, Lood C, Knudsen S, Voss A, et al. Circulating microRNA expression profiles associated with systemic lupus erythematosus. *Arthritis Rheum* 2013;65:1324-34.
15. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. *Arthritis Rheum* 1980;23:581-90.
16. Clements P, Lachenbruch P, Siebold J, White B, Weiner S, Martin R, et al. Inter and intraobserver variability of total skin thickness score (modified Rodnan TSS) in systemic sclerosis. *J Rheumatol* 1995;22:1281-5.
17. LeRoy EC, Black C, Fleischmajer R, Jablonska S, Krieg T, Medsger TA Jr, et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. *J Rheumatol* 1988;15:202-5.
18. Heegaard NH, Schetter AJ, Welsh JA, Yoneda M, Bowman ED, Harris CC. Circulating micro-RNA expression profiles in early stage nonsmall cell lung cancer. *Int J Cancer* 2012;130:1378-86.
19. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc, Ser B* 1995;57:289-300.
20. Duttgupta R, Jiang R, Gollub J, Getts RC, Jones KW. Impact of cellular miRNAs on circulating miRNA biomarker signatures. *PLoS One* 2011;6:e20769.
21. Makino K, Jinnin M, Kajihara I, Honda N, Sakai K, Masuguchi S, et al. Circulating miR-142-3p levels in patients with systemic sclerosis. *Clin Exp Dermatol* 2012;37:34-9.
22. Sonkoly E, Pivarsci A. microRNAs in inflammation. *Int Rev Immunol* 2009;28:535-61.
23. McDonald JS, Milosevic D, Reddi HV, Grebe SK, Algeciras-Schimmich A. Analysis of circulating microRNA: preanalytical and analytical challenges. *Clin Chem* 2011;57:833-40.
24. Pritchard CC, Kroh E, Wood B, Arroyo JD, Dougherty KJ, Miyaji MM, et al. Blood cell origin of circulating microRNAs: a cautionary note for cancer biomarker studies. *Cancer Prev Res* 2012;5:492-7.
25. Landry P, Plante I, Ouellet DL, Perron MP, Rousseau G, Provost P. Existence of a microRNA pathway in anucleate platelets. *Nat Struct Mol Biol* 2009;16:961-6.
26. Olive V, Jiang I, He L. mir-17-92, a cluster of miRNAs in the midst of the cancer network. *Int J Biochem Cell Biol* 2010;42:1348-54.
27. Shan SW, Fang L, Shatseva T, Rutnam ZJ, Yang X, Du W, et al. Mature miR-17-5p and passenger miR-17-3p induce hepatocellular carcinoma by targeting PTEN, GalNT7 and vimentin in different signal pathways. *J Cell Sci* 2013;126:1517-30.
28. Shan SW, Lee DY, Deng Z, Shatseva T, Jeyapalan Z, Du WW, et al. MicroRNA MiR-17 retards tissue growth and represses fibronectin expression. *Nat Cell Biol* 2009;11:1031-8.
29. Mendell JT. miRiad roles for the miR-17-92 cluster in development and disease. *Cell* 2008;133:217-22.
30. Taguchi A, Yanagisawa K, Tanaka M, Cao K, Matsuyama Y, Goto H, et al. Identification of hypoxia-inducible factor-1 alpha as a novel target for miR-17-92 microRNA cluster. *Cancer Res* 2008;68:5540-5.
31. Gilkes DM, Bajpai S, Chaturvedi P, Wirtz D, Semenza GL. Hypoxia-inducible factor 1 (HIF-1) promotes extracellular matrix remodeling under hypoxic conditions by inducing P4HA1, P4HA2, and PLOD2 expression in fibroblasts. *J Biol Chem* 2013;288:10819-29.
32. Fabbri M, Garzon R, Cimmino A, Liu Z, Zanesi N, Callegari E, et al. MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B. *Proc Natl Acad Sci U S A* 2007;104:15805-10.
33. Ihn H, Yamane K, Kubo M, Tamaki K. Blockade of endogenous transforming growth factor beta signaling prevents up-regulated collagen synthesis in scleroderma fibroblasts: association with increased expression of transforming growth factor beta receptors. *Arthritis Rheum* 2001;44:474-80.
34. Danger R, Pallier A, Giral M, Martinez-Llordella M, Lozano JJ, Degauque N, et al. Upregulation of miR-142-3p in peripheral blood mononuclear cells of operationally tolerant patients with a renal transplant. *J Am Soc Nephrol* 2012;23:597-606.
35. Wang Y, Gu J, Roth JA, Hildebrandt MA, Lippman SM, Ye Y, et al. Pathway-based serum microRNA profiling and survival in patients with advanced stage non-small cell lung cancer. *Cancer Res* 2013;73:4801-9.
36. Leask A, Abraham DJ, Finlay DR, Holmes A, Pennington D, Shi-wen X, et al. Dysregulation of transforming growth factor beta signaling in scleroderma: overexpression of endoglin in cutaneous scleroderma fibroblasts. *Arthritis Rheum* 2002;46:1857-65.
37. Kawakami T, Ihn H, Xu W, Smith E, Leroy C, Trojanowska M. Increased expression of TGF-beta receptors by scleroderma fibroblasts: evidence for contribution of autocrine TGF-beta signaling to scleroderma phenotype. *J Invest Dermatol* 1998;110:47-51.
38. Ohtsuka K, Gray JD, Stimmler MM, Toro B, Horwitz DA. Decreased production of TGF-beta by lymphocytes from patients with systemic lupus erythematosus. *J Immunol* 1998;160:2539-45.