

AWARD NUMBER: W81XWH-13-1-0452

TITLE: Predicting Disease Progression in Scleroderma with Skin and Blood Biomarkers

PRINCIPAL INVESTIGATOR: Maureen D. Mayes, M.D., M.P.H.

CONTRACTING ORGANIZATION: The University of Texas Health Science Center
Houston, TX 77030-5401

REPORT DATE: December 2016

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE December 2016		2. REPORT TYPE Final		3. DATES COVERED 23 Sep 2013 - 22 Sep 2016	
4. TITLE AND SUBTITLE Predicting Disease Progression in Scleroderma with Skin and Blood Biomarkers				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER W81XWH-13-1-0452	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Maureen D. Mayes, M.D., M.P.H. E-Mail: Maureen.D.Mayes@uth.tmc.edu				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) The University of Texas Health Science Center at Houston 7000 Fannin St., UCT 1006 Houston, TX 77030-5401				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT Scleroderma (Systemic Sclerosis, SSc) is a chronic, incurable autoimmune disease associated with high morbidity and mortality primarily due to lung disease. There is a large variability in individual patients' courses and current predictors of disease progression are inadequate. The overall objective of the proposed research is to develop reliable predictors for clinical outcomes in scleroderma, utilizing the biospecimens and longitudinal clinical data in the GENISOS cohort combining data from multiple areas to develop robust prediction models for ILD progression. The GENISOS cohort is a unique and valuable resource for biomarker development; no other early SSc cohort exists that has serial biological samples linked to longitudinal clinical data and genetic markers. GENISOS is an inception cohort that avoids survival bias inherent in studies of prevalent cases (mean disease duration at GENISOS entry = 2.5 years, eligibility criterion mandates disease duration ≤ 5 years). The cohort was originally established in 1998 and this funding support permitted us to enroll additional patients, continue follow-up on previously enrolled subjects, collect bio-specimens (DNA, RNA in PAXgene tubes, serum, plasma, skin biopsies), perform laboratory studies and analysis as detailed in the Overall Project Summary section of the final report.					
15. SUBJECT TERMS Scleroderma, Systemic Sclerosis, GENISOS (Genes versus Environment in Scleroderma Outcome Study), Interstitial Lung Disease, Predictors of Outcome, Cytokines, DNA, RNA					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	124	19b. TELEPHONE NUMBER (include area code)

Predicting Disease Progression in Scleroderma with Skin and Blood Biomarkers

Proposal Log Number PR120687 (HRPO Log Number A-17770.1)

Award Number W81XWH-13-1-0452



PI: Maureen D. Mayes, M.D., M.P.H.

Org: University of Texas Health Science Center at Houston Award Amount: \$1,127,763.00

Study/Project Aim(s)

- Aim 1. To study genetic susceptibility variants (DNA) (identified from our previous GWAS, Immunochip and HLA studies) as predictors of progressive disease, in the GENISOS cohort
- Aim 2. To identify blood and skin gene expression profiles (RNA) predictive of progressive disease
- Aim 3. To identify the cytokines/analytes (protein) predictive of disease course utilizing multiplex assays
- 4. To build multivariable models with identified clinical and molecular predictors, utilizing advanced variable reduction and longitudinal analysis strategies

Approach - Funding for this project started 23 Sep 2013. The GENISOS (early scleroderma disease) cohort provides the subjects from which sample collection (DNA, RNA [PAXgene], skin biopsies, monocytes, and serum), clinical characterization and autoantibodies are obtained. Data collection is captured in an electronic database for analysis.



SCLERODERMA RESEARCH PROGRAM

In the 3 years of this award, we have enrolled **91 new** cases and conducted **405 follow-up** visits. Sample collection has met our proposed goal as has clinical characterization, data collection and entry as well as chest CAT-scan fibrosis scoring.

Timeline and Cost

Activities	CY	13	14	15	16
Text (Major aim/study/milestone)					
Text (Major aim/study/milestone)					
Text (Major aim/study/milestone)					
Text (Major aim/study/milestone)					
Estimated Budget (\$K)		\$85.3	\$85.3	\$85.3	

Updated: 19 DECEMBER 2016

Goals/Milestones (relevant to this period):

Task 1: IRB & DOD HRPO approvals – **Completed initial approvals, ongoing annual reviews up to date**

Task 2a: Collection of DNA samples: **86 samples**

Task 3a: Collection of Skin biopsy samples: **238 biopsies**

Task 4a: Collection of monocyte samples: **355**

Task 7: Calculation of fibrosis score of HRCTs (months 0-36) **Dr. Ferguson has reviewed and scored a total of 98 HRCTs.**

Task 8: Expansion of the GENISOS cohort: Cumulative **91 new enrollees, 405 follow-up visits.**

Task 9: Maintenance and expansion of the GENISOS data base (months 0-36) **tracking of all skin biopsies and blood samples has been added to the database, all visit data has been checked for quality (through our data quality measures) and entered.**

Budget Expenditure To-Date

Projected Expenditure: **\$1,127,763.00** (Years 01-03)

Actual Expenditure: **\$1,122,056.58** (23 Sept 2013 thru 22 Sept 2016)

Table of Contents

	<u>Page</u>
1. Introduction.....	1
2. Keywords.....	1
3. Overall Project Summary.....	1
4. Key Research Accomplishments.....	8
5. Conclusion.....	9
6. Publications, Abstracts, and Presentations.....	10
7. Inventions, Patents and Licenses.....	14
8. Reportable Outcomes.....	14
9. Other Achievements.....	14
10. References.....	14
11. Appendices.....	15

**FINAL PROGRESS REPORT for W81XWH-13-1-0452
for period September 23, 2013 – September 22, 2016
“Predicting Disease Progression in Scleroderma with Skin and Blood Biomarkers”
Principal Investigator: Maureen D. Mayes, M.D., M.P.H.**

1. INTRODUCTION:

Scleroderma (Systemic Sclerosis, SSc) is a chronic, incurable autoimmune disease associated with high morbidity and mortality primarily due to lung involvement (1). There is a large variability in individual patients' courses and current predictors of disease progression are inadequate (2). The overall objective of the proposed research was to develop more reliable predictors for clinical outcomes in scleroderma, utilizing the bio-specimens and longitudinal clinical data in the GENISOS cohort combining data from multiple areas to develop prediction models for ILD progression. The GENISOS cohort is a unique and valuable resource for biomarker development; no other early SSc cohort exists that has serial biological samples linked to longitudinal clinical data and genetic markers (3). GENISOS is an inception cohort that avoids survival bias inherent in studies of prevalent cases (mean disease duration at GENISOS entry = 2.5 years, eligibility criterion mandates disease duration \leq 5 years). The cohort was originally established in 1998 and this funding support permitted us to enroll additional patients, continue follow-up on previously enrolled subjects, collect bio-specimens (DNA, RNA in PAXgene tubes, serum, plasma, skin biopsies), perform laboratory studies and analysis as detailed below in OVERALL PROJECT SUMMARY.

2. KEYWORDS:

Scleroderma
Systemic Sclerosis
GENISOS (Genes versus Environment in Scleroderma Outcome Study)
Interstitial Lung Disease
Predictors of Outcome
Cytokines
DNA
RNA

3. OVERALL PROJECT SUMMARY:

This report summary is divided into 4 sections – the first (**3.a.**) is a tally of cumulative subjects enrolled and followed up; the second section (**3.b.**) summarizes sample collection which supports the experimental sections; section **3.c.** summarizes the meeting schedules necessary to coordinate this multi-step research project; and the fourth and longest section (**3.d.**) provides a summary of the data resulting from this work as they pertain to each Task separately.

3.a. Recruitment and follow-up of subjects in the GENISOS cohort (Task 8)

Recruitment of new subjects (Task 8) has continued throughout all 36 months of this project. The total number of new enrollees since the initiation of this project is **91 new subjects**. In addition, **405 follow-up visits** have been conducted. The proposed recruitment estimate was 120

new subjects. Although we have fallen somewhat short of this goal, we have collected more samples (see Table below) than originally proposed which provides us with additional statistical power to analyze changes over time and predictors of these changes.

3.b. Sample Collection (Tasks 2 through 4).

Cumulative sample collection to support the experiments in Tasks 2 through 4 is summarized in the Table below. Note that the number of skin biopsies (both initial biopsy at baseline entry into the cohort as well as follow-up biopsies) exceeded the proposed/estimated number. Along with this, the collection of monocytes, DNA, serum and PAXgene (for RNA gene expression) samples also exceeded our initial proposed/estimated number.

TABLE: Cumulative Sample Collection for GENISOS patients from 23 SEP 2013 through 23 SEPT 2016

CUMULATIVE		1st through 12th Quarter			Proposed		Sample Collection – GENISOS Patients		
Quarters	Time/Months	Skin bx Baseline	Skin bx	Monocyte Baseline	Monocyte F/U	DNA Baseline	Serum/PAXgene Baseline	Serum/PAXgene F/U	
1 thru	0-36	50	25	50	25	50	50	88	
CUMULATIVE		1st through 12th Quarter			ACTUAL		Sample Collection – GENISOS Patients		
1 thru 12	0-36	75	163	72	283	86	74 serum/ 76 PAXgene	352 serum/ 344 PAXgene	

F/U = Follow-Up; Bx = biopsy.

3.c. Coordination of projects and personnel

Weekly laboratory meetings and monthly research data meetings have been held to discuss the results of this project and to review current and future plans. The weekly laboratory meetings are attended by the PI (Mayes), the Co-Investigator (Assassi), the lab manager (Charles), the research assistant (Hau), the project coordinator (Theodore) and the Rheumatology Division manager (Gurtovnik) to review administrative issues regarding laboratory supplies and other issues. The monthly research data meetings are attended by investigators Drs. Mayes, Assassi, Pedroza and Ferguson as well as Mr. Theodore and Mr. Charles.

3.d. SUMMARY of PROGRESS per SOW TASKS:

Task 1: Institutional Review Board (IRB) (in months 1-2) and DOD Human Research Protection Office (HRPO). The initial IRB approvals were obtained early in Year 1 and annual reviews and renewals were obtained. Documentation has been provided to the DOD HRPO. There have been no substantive modifications to the protocol and no unanticipated problems involving risk to the subjects were encountered.

“Task 2: Collection of DNA samples, genotyping and analysis of genetic data (Specific Aim 1)”

“2.a. Collection of DNA samples (months 1-36): Forty new patients will be enrolled annually by Drs. Mayes and Assassi into the GENISOS cohort at UTHealth. DNA will be extracted from baseline blood samples by the Research Assistant in the laboratories of Division of Rheumatology at UTHealth.”

A total of 91 new subjects have been enrolled which, as noted in section 3.A. above, is less than initially estimated. DNA has been extracted from all blood samples and is stored in the Rheumatology Research Laboratory at UT Houston. These samples, as well as previously collected DNA samples, serve as the basis for the studies performed below.

“2.b. Genotyping by Taqman Assay: Genotyping data are available on the majority of patients in the GENISOS cohort through (previous) genome wide association study and Immuchip efforts. Genotyping by Taqman assays will be performed for selected susceptibility loci in the newly enrolled patients (120 by the end of funding period).”

“2.c. Analysis of genetic data.”

“2.d. Manuscript on genetic predictors of disease progression.”

We genotyped thirteen idiopathic interstitial pneumonia (IIP)-related single nucleotide polymorphisms (SNPs) by Taqman assay and analyzed them for their association with SSc overall or for the presence and severity of SSc-related lung disease (ILD). Surprisingly, we did NOT find an association with any of these SNPs. Our results add new evidence that SSc and SSc-related ILD are genetically distinct from IIP even though they share phenotypic similarities. (PMID 26792595)

In a related study, we continued analysis of SNPs associated with SSc and reported, along with our international colleagues and additional SSc cases and controls, that a polymorphism in *CCR6* was associated with Scl-70 (antitopoisomerase or ATA) positive SSc. The *CCR6* gene encodes for a chemokine receptor which plays a critical role in IL-17 driven autoimmunity. Scl-70 is the autoantibody that is associated with severe and progressive interstitial lung disease in SSc. Thus SSc-cases with this polymorphism may benefit from agents that block the IL-17 pathway. (PMID26314374)

“Task 3: Collection of skin samples, RNA extraction, gene expression analysis (Specific Aim 2)”

“3.a. Collection of skin biopsy samples. Drs. Mayes and Assassi will perform 60 skin biopsies per year at UTHealth.”

“3.b. RNA extraction and global gene expression study in skin samples. Purified RNA will be extracted from stored skin biopsy samples by the Research Assistant using commercially available kits in the laboratories of Division of Rheumatology at UTHealth. The RNA quantity and quality will be assessed by Nanodrop and Bio-analyzer in the CTSA Microarray Core Laboratories at UTHealth. Global gene expression profiling will also be performed in the Microarray Core Laboratories.”

“3.c. Analysis of skin global gene expression data. The analysis of skin gene expression data will be completed by Drs. Assassi and Gorlova in consultation with Dr. Michael Whitfield at Dartmouth Medical School (Dartmouth, NH).”

“3.d. Manuscript on skin gene expression predictors of disease progression.”

The cumulative number of skin biopsies is 238 (75 on newly enrolled subjects and 163 repeat/follow-up biopsies as noted in the Table above) which exceeded our initial target (note: this number of biopsies is approved by IRB) and which strengthens our ability to determine differences between normal and SSc and among the subsets of SSc.

This project has been divided into two sections – the first involved 61 GENISOS subjects and 36 controls (reported in a manuscript discussed in the following paragraph) and the second part (skin biopsies on 127 unique GENISOS subjects some of whom have serial/follow-up biopsies) will expand and confirm our previous findings and permit longitudinal analysis with prognostic implications. This second phase is in process. This is the largest and most comprehensive study of this type in SSc and indeed in any autoimmune connective tissue disease. Results of gene expression studies on the initial 61 GENISOS patients identified 2,754 differentially expressed transcripts when compared to controls. Clustering analysis revealed 2 prominent transcriptomes in SSc patients: the keratin and fibro-inflammatory signatures. Higher keratin transcript scores were associated with shorter disease duration and interstitial lung disease, while higher fibro-inflammatory scores were associated with diffuse cutaneous involvement and a higher modified Rodnan skin thickness score. There was considerable heterogeneity across patients. These findings may be useful in stratifying patients for targeted therapies or for predicting the response to immunosuppression (PMID 26238292).

The second phase of our gene expression studies will examine changes over time within individual patients and across patients to identify predictors for better/worse clinical course.

“Task 4: Collection of monocyte samples, RNA extraction, gene expression analysis (Specific Aim 2)”

“4.a. Collection of monocyte samples; Blood samples will be collected in CPT tubes from newly enrolled patients and patients seen at year 1 visit. 80 monocyte samples have been already collected. Drs. Mayes and Assassi will conduct the GENISOS visit and collect the clinical data. The monocyte samples will be purified and the RNA will be extracted by the Research Assistant in the laboratories of Division of Rheumatology at UTHealth. The purified RNA samples will be stored in -80 freezer.”

“4.b. Global gene expression study in monocyte samples”

“4.c. Analysis of monocyte global gene expression data”

Total collection of CPT tubes for monocyte isolation is 355 (72 at baseline and 352 in follow-up). Purified RNA from these samples are stored in -80 freezers.

Regarding analysis approaches, Drs. Mayes and Assassi have communicated with our consultant, Dr. Michael Whitfield, Associate Professor in the Department of Genetics at Dartmouth Medical School and Director of the Translation Genomics Laboratory at the Dartmouth Institute for Quantitative Biomedical Sciences through multiple emails and phone

calls. Several face-to-face meetings have occurred between Drs. Mayes, Assassi and Whitfield to discuss approaches to analysis of microarray gene expression data. These meetings have been scheduled to coincide with the annual American College of Rheumatology Conference, at the bi-annual Scleroderma Research Workshop meeting (August 2015) and at several other meetings so as to save costs since travel to these meetings was already planned. In addition, multiple discussions have been conducted through electronic communication.

Global gene expression studies of the monocyte samples will await the gene expression data from our skin biopsy study (Task 3 above) as the results of the skin biopsies gene expression will inform the monocyte studies. These further studies will be funded from non-DOD sources.

“Task 5: Cytokine level determination and data analysis”

“5.a. Multiplex assays in rapid/slow progressor groups. Human DiscoveryMAP v 1.0 multiplex assays for the analyte determination in already collected 62 samples (rapid and slow progressor samples-Stage I of Specific Aim 3) has been performed by Myriad Ruled Based Medicine (Austin, TX).”

“5.c. Design of custom-made multiplex assay. Custom-made multiplex assays will be designed in collaboration with Myriad Ruled Based Medicine (Austin, TX) based on the identified cytokines in the Subtask 5.b.”

“5.d. Determination of cytokine levels by custom-made assays. The levels of selected cytokines will be determined in all collected baseline samples (currently available 331 + newly enrolled 80= 411) utilizing custom-made assays. These assays will be performed by Myriad Ruled Based Medicine (Austin, TX).”

The levels of 72 plasma cytokine/analytes were measured in 91 patient samples and 40 age-, gender- and ethnicity matched controls. All patients had SSc-ILD based on imaging. The main focus of this analysis was to identify plasma cytokine/analytes which predict the course of SSc-ILD. Plasma samples were investigated by highly sensitive multiplex assays including Simoa in Myriad-Rule Based Medicine (Austin, TX). In this important analysis, four plasma cytokine/analytes were predictive of ILD progression. Higher levels in the below cytokines were associated with slower ILD progression as measured by annualized rate of change in FVC% predicted:

- B Lymphocyte Chemoattractant (BLC) → p=0.009
- Beta-2-Microglobulin (B2M) → p=0.015
- Macrophage Migration Inhibitory Factor (MIF) → p= 0.033
- Macrophage Inflammatory Protein-1 beta (MIP-1 beta) → p= 0.044

BLC, MIF, and MIP-1 are important inflammatory cytokines. BLC is selectively chemotactic for B-cells while MIF and MIP-1 are prominent interferon inducible chemokines. These cytokines cumulatively represent an inflammatory subset of SSc that are more amendable to immunosuppression, thus are associated with better ILD outcome. This finding is also supported by our observation at the whole blood gene expression level. In those studies, patients with a prominent interferon signature were more likely to respond to immunosuppression.

We also have measured levels of the aforementioned cytokines in the baseline serum samples from SCOT trial. Longitudinal FVC% data from this trial will be shared with us next month (January 2017). We plan to publish the combined results of the GENISISOS cohort (discovery cohort) and SCOT (validation) in a manuscript.

In addition, and as reported in our last progress report we investigated the role of pneumoprotein predictors of early ILD progression in the subjects in the GENISISOS cohort who had ILD verified by imaging. Specifically, we studied CCL18 (also known as pulmonary activation-regulated chemokine or PARC, as we had previously observed that PARC had a significant correlation with short-term decline in functional vital capacity (FVC) as a surrogate of ILD progression. However, in this cohort of patients with verified ILD there was no statistical significant relationship between CCL18/PARC levels and FVC progression ($p=0.45$). We also measured KL-6 by a commercially available, validated ELISA kit. Baseline KL-6 levels were higher in patients than in controls ($p<0.0001$). Baseline higher KL-6 levels (as a continuous variable) were predictive of a faster rate of FVC% decline at the one year follow-up ($b=-0.03$, $p=0.04$). Upon categorizing KL-6 using a cut-off of 1273 u/mL based on the optimal cut-off previously determined by a Japanese study (4), the predictive significance of KL-6 remained in both the univariate ($p=0.01$) and multivariable analyses after accounting for race, disease duration, age, and treatment with immunosuppressive agents. Although CCL-18 was higher in patients than controls (<0.0001), its levels did not predict rate of FVC decline ($p=0.4$). These data were presented at the 2016 American College of Rheumatology meeting and the manuscript is currently being prepared.

In this very select cohort we also analyzed whether autoantibodies and other various clinical and demographic variables could predict FVC decline and observed that Scl-70 performed by immunodiffusion was the only variable that predicted faster FVC decline in patients with SSc related ILD. Interestingly, the same antibody performed by a chemoluminescent or a line-blot immunoassay were not predictive of FVC decline. The discrepancy observed between different methods of Scl-70 determination is important as it might have relevant implications for enrichment strategies in clinical trials of SSc-ILD. These data were presented in an abstract form at the 2016 Scleroderma World Congress (Lisbon, Portugal February 2016) and the manuscript is under preparation for submission.

**“Task 6: Multivariable models with identified clinical and molecular predictors
(Specific Aim 4)**

6.a. Random forest and longitudinal analyses (months 32-34)

**6.b. Manuscript on multivariable models predictive of disease progression
(months 35-36)”**

For this aim, we have conducted a novel multi-level analysis combining whole blood gene expression data with plasma proteins. This analytic method is called Similarity Network Fusion analysis (Wang et al. Nature Methods 2014). As shown in Figure 1, networks of patient and control samples are generated at each

molecular level (e.g. whole blood gene expression and proteins in the circulation) and then combined into a single network. We are currently analyzing the skin gene expression data from the GENISOS cohort which will provide a third level of molecular data. After addition of the skin gene expression data, we will create clusters of patients and link those patient groups to pulmonary outcomes.

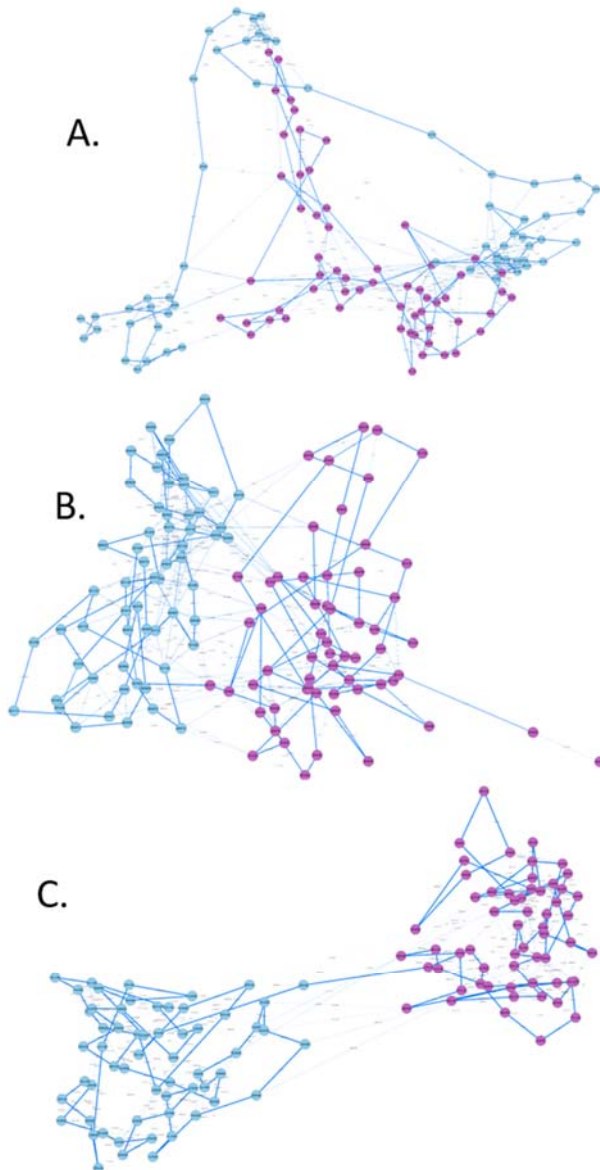


Figure 1: Similarity network fusion analysis of multilevel molecular data. A. Whole blood gene expression data; B. Protein data in the circulation; C. Combined molecular data
Red: SSc and blue: controls

“Task 7: Calculation of fibrosis score of high resolution chest CTs (months 0-36)”
“Dr. Ferguson will determine the fibrosis score on the high resolution chest CTs obtained in the GENISOS cohort.”

Cumulatively, our radiology co-investigator, Dr. Ferguson, has reviewed and scored 98 high resolution chest computed tomography scans (chest HRCTs). The degree of fibrosis correlates with the pulmonary function testing in terms of stabilization or decline. These scores are used in our evaluation of SSc-related interstitial lung disease (ILD) in subjects in the GENISOS cohort and are used, in addition to pulmonary function test results, to confirm the presence or absence of ILD in the cohort.

“Task 8: Expansion of the GENISOS cohort (months 0-36)”

“The baseline and follow-up GENISOS visits will be conducted by Drs. Mayes and Assassi at UTHealth. This is necessary for capturing the longitudinal clinical data and collecting the bio-specimens for the proposed research.”

The cumulative total since the initiation of this project is **91 newly recruited subjects** and **405 follow-up visits**. All longitudinal clinical data have been obtained and bio-specimens have been collected (See Table above)

“Task 9: Maintenance and expansion of GENISOS data base (months 0-36)”

“The Data Base Manager at UTHealth in close collaboration with Mr. Tony Mattar (Computer Task Force, Inc., Troy, MI) will maintain the GENISOS data base. The data base will also be expanded to accommodate the genetic and cytokine data. They also will establish an interface for connecting the gene expression data with the clinical data base.”

In the twelfth quarter of this project, Ms. Deepthi Nair (database manager) and Mr. Samuel Theodore (Study Coordinator) have continued to work together, as they have through all 12 quarters, to ensure that the clinical data entry is accurate and complete. Quality control (QC) checks are in place. It is only after the QC process has been completed on data for each visit that the visit data are uploaded into the “permanent” database and available for use in analyses. In collaboration with Mr. Tony Mattar (database consultant) they have reviewed, verified and uploaded the clinical information on cohort subjects both new enrollees and those with follow-up visits. The link between clinical and laboratory data and the skin biopsy specimen tracking (as established in the first quarter) is working well. The clinical and laboratory data have been queried repeatedly to provide the data field values (for example, pulmonary function test data over time) to correlate with the genetic (Task 2), gene expression (Tasks 3 and 4) and cytokine data (Task 5).

4. KEY RESEARCH ACCOMPLISHMENTS:

- Skin gene expression patterns: Prominent fibro-inflammatory and keratin gene expression signatures are present in SSc skin which show significant heterogeneity within this patient population; this finding will be useful for stratifying

- patients for targeted therapies or predicting response to immunosuppression. (PMID 26238292)
- Anti-Ku antibodies characterize a group of SSc patients who are more likely to have SSc-related lung disease and muscle inflammation but whose survival is no different from the overall SSc population. (PMID 267583908)
 - Meta-analysis of genetic risk factors identified a novel shared locus, *IRF4*, for the risk of SSc and RA highlighted the usefulness of a cross-disease GWAS meta-analysis strategy in the identification of common risk loci for autoimmune disease. (PMID 27111665)
 - Genetic susceptibility loci of idiopathic interstitial pneumonia (IIP) do **NOT** represent risk for SSc-interstitial lung disease (SSc-ILD) – this is an important finding with implications for therapy in that agents which have been shown to be effective in IIP may not be effective in SSc-ILD. (PMID 26792595)
 - CCR6 genetic polymorphisms are associated with Scl-70 positive SSc which implicates the IL-17 immune pathway as being important in at least some cases of SSc and, thus, these subjects may benefit from IL-17 blocking strategies. (PMID26314374)
 - In our cytokine studies we have found that KL6 and not CCL-18 is a predictor of early progression in systemic sclerosis related interstitial lung disease.

5. CONCLUSION:

Scleroderma (systemic sclerosis, SSc) is a markedly heterogeneous disease ranging from mild symptoms that do not affect survival to severe and rapid disease with significant morbidity and the highest mortality of all of the autoimmune diseases. The importance of our work lies in the potential of precision medicine, that is, the accurate characterization of SSc subsets at the onset of disease in order to target treatment to the individual's particular pathway responsible for disease activity and progression.

By expanding and following the GENISOS cohort, we have collected detailed clinical data in this inception cohort that characterizes and documents clinical course.

We have identified – through our skin gene expression studies described above – that there are different gene expression signatures in SSc skin which will be helpful in stratifying patients for targeted therapies. Our future and expanded gene expression studies (underway) will further refine and develop this approach.

We have identified – through our cytokine studies – that the secreted protein KL6 and not CCL-18 (as reported by others) is a predictor of early progression in SSc-related ILD. Our full cytokine profile is pending.

FUTURE DIRECTIONS

It has become clear from our current and ongoing genotyping results, gene expression studies and cytokine analysis that the complexity of SSc is even greater than we at first anticipated. New and sophisticated statistical methods that have recently been developed will help in analysis and interpretation of these data in a combined fashion.

Specifically, future directions include the following: 1) completion of our skin biopsy gene expression studies which will be the largest and most detailed such study in SSc; 2) completion of cytokine analysis including changes over time in serial serum samples; 3) analysis of these combined data sets to develop (and then publish) predictive models.

6. PUBLICATIONS, ABSTRACTS AND PRESENTATIONS:

a. All Manuscripts:

(1) PUBLICATIONS – Lay Press

Nothing to report.

(2) PUBLICATIONS – Peer-Reviewed Scientific Journals (n=10)

1. Hoa S, Hudson M, Troyanov Y, Proudman S, Walker J, Stevens W, Nikpour M, Assassi S, **Mayes MD**, Wang M, Baron M, Fritzler MJ; Canadian Scleroderma Research Group (CSRG); Australian Scleroderma Interest Group (ASIG); Genetics versus Environment in Scleroderma Outcome Study (**GENISOS**). Single-specificity anti-Ku antibodies in an international cohort of 2140 systemic sclerosis subjects: clinical associations. *Medicine (Baltimore)*. 2016 Aug;95(35):e4713. doi: 10.1097/MD.0000000000004713. PMID: 27583908
2. López-Isac E, Martín JE, Assassi S, Simeón CP, Carreira P, Ortego-Centeno N, Freire M, Beltrán E, Narváez J, Alegre-Sancho JJ; Spanish Scleroderma Group, Fernández-Gutiérrez B, Balsa A, Ortiz AM, González-Gay MA, Beretta L, Santaniello A, Bellocchi C, Lunardi C, Moroncini G, Gabrielli A, Witte T, Hunzelmann N, Distler JH, Riekemasten G, van der Helm-van Mil AH, de Vries-Bouwstra J, Magro-Checa C, Voskuyl AE, Vonk MC, Molberg Ø, Merriman T, Hesselstrand R, Nordin A, Padyukov L, Herrick A, Eyre S, Koeleman BP, Denton CP, Fonseca C, Radstake TR, Worthington J, **Mayes MD**, Martín J. Cross-disease Meta-analysis of Genome-wide Association Studies for Systemic Sclerosis and Rheumatoid Arthritis Reveals IRF4 as a New Common Susceptibility Locus. *Arthritis Rheumatol*. 2016 Apr 25. doi: 10.1002/art.39730. [Epub ahead of print] PMID: 27111665
3. Prins BP, Abbasi A, Wong A, Vaez A, Nolte I, Franceschini N, Stuart PE, Guterriez Achury J, Mistry V, Bradfield JP, Valdes AM, Bras J, Shatunov A; PAGE Consortium; International Stroke Genetics Consortium; Systemic Sclerosis consortium; Treat OA consortium; DIAGRAM Consortium; CARDIoGRAMplusC4D Consortium; ALS consortium; International Parkinson's Disease Genomics Consortium; Autism Spectrum Disorder Working Group of the Psychiatric Genomics Consortium; CKDGen consortium; GERAD1 Consortium; International Consortium for Blood Pressure; Schizophrenia Working Group of the Psychiatric Genomics Consortium; Inflammation Working Group of the CHARGE Consortium, Lu C, Han B, Raychaudhuri S, Bevan S, **Mayes MD**, Tsoi LC, Evangelou E, Nair RP, Grant SF, Polychronakos C, Radstake TR, van Heel DA, Dunstan ML, Wood NW, Al-Chalabi A, Dehghan A, Hakonarson H, Markus HS, Elder JT, Knight J, Arking DE, Spector TD, Koeleman BP, van Duijn CM, Martin J, Morris AP, Weersma RK, Wijmenga C, Munroe PB, Perry JR, Pouget JG, Jamshidi Y, Snieder H, Alizadeh BZ. Investigating the Causal Relationship of C-Reactive Protein with 32 Complex Somatic and Psychiatric Outcomes: A Large-Scale Cross-Consortium Mendelian Randomization Study. *PLoS Med*. 2016 Jun 21;13(6):e1001976. doi: 10.1371/journal.pmed.1001976. eCollection 2016 Jun. PMID:

27327646

4. Wu M, Assassi S, Salazar GA, Pedroza C, Gorlova OY, Chen WV, Charles J, Taing ML, Liao K, Wigley FM, Hummers LK, Shah AA, Hinchcliff M, Khanna D, Schioppa E, Phillips K, Furst DE, Steen V, Baron M, Hudson M, Zhou X, Pope J, Jones N, Docherty P, Khalidi NA, Robinson D, Simms RW, Silver RM, Frech TM, Fessler BJ, Fritzler MJ, Molitor JA, Segal BM, Movahedian M, Martín J, Varga J, **Mayes MD**. Genetic susceptibility loci of idiopathic interstitial pneumonia do not represent risk for systemic sclerosis: a case control study in Caucasian patients. *Arthritis Res Ther*. 2016 Jan 20;18(1):20. doi: 10.1186/s13075-016-0923-3. PMID: 26792595
5. Assassi S, Swindell WR, Wu M, Tan FK, Khanna D, Furst DE, Tashkin DP, Jahan-Tigh RR, **Mayes MD**, Gudjonsson JE, Chang JT. Dissecting the heterogeneity of skin gene expression patterns in systemic sclerosis. *Arthritis Rheumatol*. 2015 Nov; 67(11):3016-26. PMID: 26238292
6. Ochoa E, Martin JE, Assassi S, Beretta L, Carreira P, Simeón CP, Koumakis E, Dieude P, Allanore Y, García-Hernández FJ, Espinosa G, Castellvi I, Trapiella JL, Rodriguez L, González-Gay MA, Egurbide MV, Sáez L, Callejas-Rubio JL, Vargas JA, Hunzelmann N, Riemekasten G, Witte T, Distler JHW, Kreuter A, Lunardi C, Santaniello A, Tan FK, Arnett FC, Shiels PG, Herrick A, Worthington J, Vonk MC, Koeleman BP, Radstake TRDJ, **Mayes MD**, Martin J. Confirmation of CCR6 as a risk factor for antitopoisomerase I antibodies in systemic sclerosis. *Clin Exp Rheumatol*. 2015 Jul-Aug;33(4 Suppl 91):S31-55. PMID: 26314374
7. Salazar GA, Assassi S, Wigley F, Hummers L, Varga J, Hinchcliff M, Khanna D, Schioppa E, Phillips K, Furst DE, Steen V, Baron M, Hudson M, Taillefer SS, Pope J, Jones N, Docherty P, Khalidi NA, Robinson D, Simms RW, Silver RM, Frech TM, Fessler BJ, Molitor JA, Fritzler MJ, Segal BM, Al-Kassab F, Perry M, Yang J, Zamanian S, Reveille JD, Arnett FC, Pedroza C, **Mayes MD**. Antinuclear antibody-negative systemic sclerosis. *Semin Arthritis Rheum*. 2015 Jun;44(6):680-6. doi: 10.1016/j.semarthrit.2014.11.006. PMID: 25578738
8. Peterson LK, Jaskowski TD, **Mayes MD**, Tebo AE. Detection of anti-U3-RNP/fibrillarin IgG antibodies by line immunoblot assay has comparable clinical significance to immunoprecipitation testing in systemic sclerosis. *Immunol Res* 2015 Oct 14 (Epub ahead of print). PMID 26467972.
9. Wodkowski M, Hudson M, Proudman S, Walker J, Stevens W, Nikpour M, Assassi S, **Mayes MD**, Tatibouet S, Wang M; Canadian Scleroderma Research Group (CSRG); Australian Scleroderma Interest Group (ASIG); Genetics versus Environment in Scleroderma Outcome Study (**GENISOS**), Baron M, Fritzler MJ. Clinical correlates of monospecific anti-PM75 and anti-PM100 antibodies in a tri-nation cohort of 1574 systemic sclerosis subjects. *Autoimmunity*. 2015 Sep 3:1-10. [Epub ahead of print] PMID: 26334795
10. Wodkowski M, Hudson M, Proudman S, Walker J, Stevens W, Nikpour M, Assassi S, **Mayes M**, Wang M, Baron M, Fritzler M. Monospecific anti-Ro52/TRIM21 antibodies in a tri-nation cohort of 1574 systemic sclerosis subjects – Evidence of an association with interstitial lung disease and worse survival. *Clin Exp Rheumatol*. 2015 Jul-Aug;33 (4 Suppl 91):S131-5. PMID 26315678

(3) PUBLICATIONS: INVITED ARTICLES (N=2)

1. Wu M, Mayes MD. Insights into the genetic basis of systemic sclerosis: Immunity in human disease and SSc mouse models. *Adv Genomics Genet*. 2014, 4:143-151.

2. Salazar G, Mayes MD Genetics, Epigenetics, and Genomics of Systemic Sclerosis. *Rheum Dis Clin North Am.* 2015 Aug;41(3):345-66. doi: 10.1016/j.rdc.2015.04.001. Epub 2015 May 20. PMID: 26210123

(4) ABSTRACTS and PRESENTATIONS:

1. Salazar GA, Kuwana M, Wu M, **Mayes MD**, Ying J, Charles J, Assassi S. KL6 and not CCL-18 is a predictor of early progression in systemic sclerosis related interstitial lung disease. Presented at the annual meeting American College of Rheumatology, November 2016. Manuscript in preparation.
2. Pauling JD, Lu H, Betteridge ZE, Salazar G, Assassi S, **Mayes M**, McHugh NJ. Clinical Associations of Anti-Eukaryotic Initiation Factor 2B (anti-eIF2B) Antibodies in a Large Cohort of Patients with Anti-Nuclear Antibody Negative Systemic Sclerosis. Presented at the annual meeting American College of Rheumatology, November 2016. * Manuscript prepared and ready for submission.
3. Wodkowsk M, Hudson M, Proudman S, Walker J, Stevens W, Nikpour M, Assassi S, **Mayes MD**, Wang M, Canadian Scleroderma Research Group (CSRG), Australian Scleroderma Interest Group (ASIG), Genetics versus environment in Scleroderma Outcome Study (GENISOS), Baron M, Fritzler MJ. Anti-topoisomerase antibodies are negatively associated with smoking and positively associated with HLADPB1*1301 in systemic sclerosis (SSc): preliminary evidence for a possible gene-environment interaction. Presented at the annual meeting American College of Rheumatology, November 2016.
4. Salazar G, Hagan J, Wu M, Guo X, Zhou X., Charles J, Mayes MD, Assassi S. The Global MicroRNAa Profile of Systemic Sclerosis Whole Skin/Dermal Fibroblasts and the Role of the Xq26.3 miRNA Cluster As a TGF-b Pathway Positive Feedback Mechanism. Presented at the annual meeting American College of Rheumatology,. November 2015.
5. Martin J, Ochoa E, Martin JE, Assassi S, Beretta L, Carreira P, Simeon CP, Koumakis E, Dieude P, Allanore Y, García-Hernández FJ, Espinosa G, Barranco IV, Trapiella L, Rodriguez L, González-Gay MA, Egurbide MV, Saez L, Callejas JL, Vargas-Hitos JA, Hunzelmann N, Riemekasten G, Witte T, Distler JHW, Kreuter A, Lunardi C, Santaniello A, Tan FK, Arnett FC, Shields P, Herrick AL, Worthington J, Vonk MC, Koeleman BPC, Radstake TRDJ, Mayes M. Systemic Sclerosis Patients with Antitopoisomerase Antibodies Showed Significant Association with CCR6 Polymorphisms. Presented at the 2014 annual meeting American College of Rheumatology, Washington, DC. November 2014. *Arthritis Rheumatol* 2014 Oct;66(S10):S328. * Manuscript PMID 26314374
6. Salazar G, Fritzler M, Hudson M, Mejia Otero C, Estrada-Y-Martin R, Nair D, Pedroza C, Mayes M, Assassi S. Scl-70 autoantibody determined by immunodiffusion is a clinical predictor of faster interstitial lung disease progression in systemic sclerosis. Presented at the 4th Systemic Sclerosis World Congress, Lisbon, Portugal Feb 2016. *Manuscript in preparation.
7. Salazar G, Hagan J, Wu M, Guo X, Zhuo X, Charles J, Distler O, Pachera E, Mayes M, Assassi S. The global microRNA profile of systemic sclerosis whole skin/dermal fibroblasts and the role of the XQ26.3 miRNA cluster as a TGF-beta pathway positive feedback mechanism. Presented at the 4th Systemic Sclerosis World Congress, Lisbon, Portugal Feb 2016.
8. Volkmann E, Tashkin D, Elashoff R, Tseng CH, Clements P, Roth M, Furst D, Khanna D, Charles J, Mayes M, Assassi S. CXCL4: A potential biomarker for systemic

- sclerosis-related interstitial lung disease. Presented at the 4th Systemic Sclerosis World Congress, Lisbon, Portugal Feb 2016. * Manuscript submitted.
9. Elena Lopez-Isac, Lara Bossini-Castillo, Sandra G Guerra, Shervin Assassi, Carmen P. Simeón, Patricia E. Carreira, Norberto Ortego-Centeno, Paloma García de la Peña, Lorenzo Beretta, Alessandro Santaniello, Chiara Bellocchi¹ Claudio Lunardi, Gianluca Moroncini, Armando Gabrielli, Gabriela Riemekasten, Torsten Witte, Nicolas Hunzelmann, Alexander Kreuter, Jorg HW. Distler, Alexandre E. Voskuyl, J.K. de Vries-Bouwstra, Ariane L. Herrick, Jane Worthington, Christopher P. Denton, Carmen Fonseca, T.R.D.J. Radstake, Maureen D Mayes and Javier Martín. TYK2 and Systemic Sclerosis Susceptibility: a New Associated Locus in the IL-12 Pathway. Presented at American College of Rheumatology Annual Meeting, 11/2015
 10. Lara Bossini-Castillo, Enrique Raya, John D. Reveille, Shervin Assassi, Olga Y Gorlova, Dinesh Khanna, Daniel E. Furst, Suzanne Kafaja, Robert W. Simms, Robert Lafyatis, Tracy M. Frech, Virginia D. Steen, Barri J. Fessler, Jerry A. Molitor, Marta Alarcón-Riquelme, Javier Martín and Maureen D Mayes. A Genome-Wide Association Study of a Hispanic Systemic Sclerosis Cohort. Presented at American College of Rheumatology Annual Meeting, 11/2015
 11. Gloria Salazar, John Hagan, Minghua Wu, Xinjian Guo, Xiaodong Zhou, Julio Charles, Maureen D Mayes and Shervin Assassi, The Global MicroRNA Profile of Systemic Sclerosis Whole Skin/Dermal Fibroblasts and the Role of the Xq26.3 miRNA Cluster As a TGF- β Pathway Positive Feedback Mechanism. Presented at American College of Rheumatology Annual Meeting, 11/2015
 12. Elizabeth R. Volkman, Donald P. Tashkin, Michael Roth, Chi-hong Tseng, Holly LeClair, Philip J. Clements, Daniel E. Furst, Maureen D Mayes, Julio Charles, Dinesh Khanna, Robert Elashoff, Shervin Assassi and Scleroderma Lung Study II Group. CXCL4 Does Not Predict Extent or Progression of Interstitial Lung Disease in Systemic Sclerosis. Presented at American College of Rheumatology Annual Meeting, 11/2015
 13. Shervin Assassi, Maureen D Mayes, Claudia Pedroza, Jeffrey T. Chang, Daniel E. Furst, Leslie J. Crofford, Richard Nash, Peter McSweeney, Mary Ellen Csuka, Ellen Goldmuntz, Lynette Keyes-Elstein, Paul Wallace and Keith Sullivan. Immunoablation Followed By Autologous Stem Cell Transplantation in Systemic Sclerosis Patients Decreases Significantly the Interferon Signature. Presented at American College of Rheumatology Annual Meeting, 11/2014.
 14. Salazar GA, Hagan J, Wu M, Mayes MD, Reveille HD and Assassi S. The Global miRNA Whole Blood Profile in Systemic Sclerosis and Its Correlation with Serum Cytokine Levels. Presented at American College of Rheumatology Annual Meeting, 11/2014.
 15. Assassi, S, Swindell WR, Wu M, Tan FK, Khanna D, Furst DEF, Tashkin DP, Jahan-Tigh RR, Mayes MD, Gudjonsson JE, Chang JT. Dissecting the Heterogeneity of Skin Gene Expression Patterns in Systemic Sclerosis. Presented at American College of Rheumatology Annual Meeting, 11/2014.
 16. Wu M, Assassi S, Salazar GA, Gorlova OY, Chen W, Charles J, Wigley F, Hummers L, Varga J, Hinchcliff M, Khanna D, Schiopu E, Phillips K, Furst DE, Steen V, Baron M, Hudson M, Taillefer SS, Pope J, Jones N, Markland J, Docherty P, Khalidi NA, Robinson D, Simms R, Silver R, Frech TM, Fessler B, Molitor J, Fritzler M, Segal B, Al-Kassab F, Yang J, Mayes MD. Genetic Susceptibility Loci of Idiopathic Interstitial Pneumonitis do not Represent Risk for systemic sclerosis. Presented at American College of Rheumatology Annual Meeting, 11/2014.

6.b – PRESENTATIONS – see Abstracts 1 through 8 above; all abstracts were presented at national or international meetings.

7. INVENTIONS, PATENTS AND LICENSES:

Nothing to report.

8. REPORTABLE OUTCOMES:

- a. CCR6 genetic polymorphisms are associated with Scl-70 positive SSc which implicates the IL-17 immune pathway as being important SSc
- b. Identification of a SNP in IRF4 that is a novel risk factor for SSc.
- c. KL6 and not CCL-18 is a predictor of early progression in systemic sclerosis related interstitial lung disease.
- d. Identification of the keratin signature as being important in SSc-related skin gene expression.
- e. Publication of the largest skin gene-expression in SSc with implications for stratifying subjects by disease pathway and potentially as a means to determine response to therapy.
- f. Publication of the first microRNA whole blood profiling in SSc.
- g. The discovery of the keratin signature in SSc skin described in Abstract # 2 above.
- h. The discovery that the susceptibility genes of idiopathic interstitial lung disease are quite distinct from those of SSc-related ILD.

9. OTHER ACHIEVEMENTS:

A repository of DNA, RNA, serum and plasma has resulted from the sample collection for this study. This repository is linked to longitudinal detailed clinical data for access and analysis of future predictors of outcome.

10. REFERENCES:

1. Elhai M, Meune C, Avouac J, Kahan A, Allanore Y: Trends in mortality in patients with systemic sclerosis over 40 years: a systematic review and meta-analysis of cohort studies. *Rheumatology (Oxford)* 2012 June;51(6):1017-26. PMID: 21900368
2. Assassi S, Sharif R, Lasky RE, McNearney TA, Estrada-Y-Martin RM, Draeger H, Nair DK, Fritzler MJ, Reveille JD, Arnett FC, Mayes MD; GENISOS Study Group. Predictors of interstitial lung disease in early systemic sclerosis: a prospective longitudinal study of the GENISOS cohort. *Arthritis Res Ther.* 2010;12(5):R166. doi: 10.1186/ar3125. PMID: 20813056.
3. Reveille JD, Fischbach M, McNearney T, Friedman AW, Aguilar MB, Lisse J, Fritzler MJ, Ahn C, Arnett FC; GENISOS Study Group. Systemic sclerosis in 3 US ethnic groups: a comparison of clinical, sociodemographic, serologic, and immunogenetic determinants. *Semin Arthritis Rheum.* 2001 Apr;30(5):332-46. PMID: 11303306.
4. Kuwana M, Shirai Y, Takeuchi T. Elevated Serum Krebs von den Lungen-6 in Early Disease Predicts Subsequent Deterioration of Pulmonary Function in Patients with Systemic Sclerosis and Interstitial Lung Disease. *J Rheumatol.* 2016 Oct;43(10):1825-1831. PMID: 27481907.

5. Wang B, Mezlini AM, Demir F, Fiume M, Tu Z, Brudno M, Haibe-Kains B, Goldenberg A. Similarity network fusion for aggregating data types on a genomic scale. Nat Methods. 2014 Mar;11(3):333-337. PMID 24464287.

11. APPENDICES:

(1) List of personnel (not salaries) receiving pay from the research effort:

Personnel	Role	Percent Effort
Dr. Maureen Mayes	Principal Investigator	20%
Dr. Shervin Assassi	Co-Investigator	0% (K-award)
Dr. Claudia Pedroza	Co-Investigator	10%
Dr. Emma Ferguson	Co-Investigator	2.5%
Samuel Theodore	Research Coordinator	100%
Hau Pham (who replaced Kelley Liao as of 5/23/2016)	Research Assistant	50%
Kelley Liao (who replaced Miranda Taing as of 5/26/2015)	Research Assistant	50%
Miranda Taing	Research Assistant	50%

(2) Copies of Journal Articles

Single-specificity anti-Ku antibodies in an international cohort of 2140 systemic sclerosis subjects: clinical associations

S. Hoa, MD^{a,b}, M. Hudson, MD, MPH^{a,b,c,*}, Y. Troyanov, MD^{d,e}, S. Proudman, MBBS^{f,g}, J. Walker, MBBS, PhD^h, W. Stevens, MBBS, PhDⁱ, M. Nikpour, MBBS, PhD^{i,j}, S. Assassi, MD^k, M.D. Mayes, MD, MPH^k, M. Wang, MSc^b, M. Baron, MD^{a,b,c}, M.J. Fritzler, MD, PhD^l, Canadian Scleroderma Research Group (CSRG); Australian Scleroderma Interest Group (ASIG); Genetics versus Environment in Scleroderma Outcome Study (GENISOS)

Abstract

Autoantibodies directed against the Ku autoantigen are present in systemic sclerosis (SSc) and have been associated with myositis overlap and interstitial lung disease (ILD). However, there is a paucity of data on the clinical correlates of anti-Ku antibodies in the absence of other SSc-specific antibodies. The aim of this study was to assess the clinical correlates of single-specificity anti-Ku in SSc.

An international (Canada, Australia, USA, Mexico) cohort of 2140 SSc subjects was formed, demographic and clinical variables were harmonized, and sera were tested for anti-Ku using a line immunoassay. Associations between single-specificity anti-Ku antibodies (i.e., in isolation of other SSc-specific antibodies) and outcomes of interest, including myositis, ILD, and survival, were investigated.

Twenty-four (1.1%) subjects had antibodies against Ku, and 13 (0.6%) had single-specificity anti-Ku antibodies. Subjects with single-specificity anti-Ku antibodies were more likely to have ILD (58% vs 34%), and to have increased creatine kinase levels (>3× normal) at baseline (11% vs 1%) and during follow-up (10% vs 2%). No difference in survival was noted in subjects with and without single-specificity anti-Ku antibodies.

This is the largest cohort to date focusing on the prevalence and disease characteristics of single-specificity anti-Ku antibodies in subjects with SSc. These results need to be interpreted with caution in light of the small sample. International collaboration is key to understanding the clinical correlates of uncommon serological profiles in SSc.

Abbreviations: ACA = anticentromere antibodies, ACR = American College of Rheumatology, ANA = antinuclear antibodies, ARNAP = anti-RNA polymerase III antibodies, ASIG = Australian Scleroderma Interest Group, ATA = antitopoisomerase I

Editor: Francesco Carubbi.

Financial disclosures: The CSRG received funds and/or gifts in kind from the Canadian Institutes of Health Research (CIHR) (grant n. FRN 83518), the Scleroderma Society of Canada and its provincial Chapters, Scleroderma Society of Ontario, Sclérodémie Québec, Cure Scleroderma Foundation, Scleroderma Society of Saskatchewan, INOVA Diagnostics Inc. (San Diego, CA), Euroimmun (Lübeck, Germany), Fonds de la recherche en santé du Québec (FRSQ), the Canadian Arthritis Network (CAN), the Arthritis Society Research Chair (University of Calgary) and the Lady Davis Institute for Medical Research of the Jewish General Hospital, Montreal, QC.

The CSRG has also received educational grants from Pfizer and Actelion pharmaceuticals. The ASIG receives unrestricted grants from Actelion, Pfizer and GSK, and is also supported by Scleroderma Australia. The GENISOS cohort receives funding from DoDW81XWH-13-1-0452.

Conflicts of interest: SH is funded as a clinical fellow by the Université de Montréal Rheumatology Program Abbvie educational grant. MH is funded by the Fonds de la recherche en Santé du Québec. NN holds an NHMRC research fellowship (APP1071735). SA is funded by the NIH/NIAMS K23AR061436. MDM is funded by the NIH/NIAMS AR055258.

The funding sources had no role in the design of the study, analysis of the data, preparation of the manuscript and decision to submit for publication.

The authors report no conflicts of interest.

Supplemental Digital Content is available for this article.

^a Department of Medicine, McGill University, Montreal, Quebec, Canada, ^b Lady Davis Institute, Jewish General Hospital, Montreal, Quebec, Canada, ^c Division of Rheumatology, Jewish General Hospital, Montreal, Quebec, Canada, ^d Division of Rheumatology, Hôpital du Sacre-Coeur de Montreal, Montreal, Quebec, Canada, ^e Department of Medicine, Université de Montréal, Montreal, Quebec, Canada, ^f Rheumatology Unit, Royal Adelaide Hospital, Adelaide, Australia, ^g Discipline of Medicine, University of Adelaide, Bedford Park, Australia, ^h Department of Allergy and Immunology, Flinders Medical Centre, Bedford Park, Australia, ⁱ Department of Rheumatology, St. Vincent's Hospital Melbourne, Fitzroy, Victoria, Australia, ^j Department of Medicine, The University of Melbourne at St. Vincent's Hospital, Melbourne, Victoria, Australia, ^k Division of Rheumatology and Immunogenetics, University of Texas Health Science Centre at Houston, Houston, TX, ^l Faculty of Medicine, University of Calgary, Calgary, Alberta, Canada.

* Correspondence: Dr Marie Hudson, Jewish General Hospital, Room A-725, 3755 Côte Sainte-Catherine Road, Montreal, Quebec H3T 1E2, Canada (e-mail: marie.hudson@mcgill.ca).

Copyright © 2016 the Author(s). Published by Wolters Kluwer Health, Inc. All rights reserved.

This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Medicine (2016) 95:35(e4713)

Received: 29 April 2016 / Received in final form: 3 August 2016 / Accepted: 4 August 2016

<http://dx.doi.org/10.1097/MD.0000000000004713>

(or anti-Scl70) antibodies, CI = confidence interval, CK = creatine kinase, CSRG = Canadian Scleroderma Research Group, DNA-PKcs = DNA-dependent protein kinase catalytic subunit, EULAR = European League Against Rheumatism, GENISOS = Genetics versus Environment in Scleroderma Outcome Study, HLA = human leukocyte antigen, HRCT = high-resolution computed tomography, IIF = indirect immunofluorescence, ILD = interstitial lung disease, LIA = line immunoassay, mPAP = mean pulmonary artery pressure, mRSS = modified Rodnan Skin Score, OR = odds ratio, PCWP = pulmonary capillary wedge pressure, SD = standard deviation, SLE = systemic lupus erythematosus, sPAP = systolic pulmonary artery pressure, SSc = systemic sclerosis.

Keywords: anti-Ku antibodies, international cohort, interstitial lung disease, single-specificity, systemic sclerosis

1. Introduction

Systemic sclerosis (SSc) is a heterogeneous disease with varying degrees of skin and organ involvement, and can be classified by extent of skin involvement (limited or diffuse cutaneous SSc), and also by serological subtype. Common SSc-specific autoantibodies, such as anticentromere (ACA), antitopoisomerase I (ATA), and anti-RNA polymerase III (ARNAP) antibodies, have been associated with specific clinical features. In recent years, less common SSc-associated autoantibodies have been studied and their clinical correlates characterized. A potential limitation of some of those studies is the confounding introduced by the presence of overlapping antibodies. The study of distinct autoantibodies in the absence of other SSc-related autoantibodies, which we will refer to as single-specificity, has allowed us to understand specific clinical correlates of individual autoantibodies. For example, Ro52/TRIM21 autoantibodies were found to be independently associated with the presence of interstitial lung disease (ILD) and poor survival in SSc,^[1] and distinct associations were found for single-specificity anti-PM75, anti-PM100, and anti-PM-1 α antibodies.^[2,3]

Autoantibodies directed against Ku have been reported in a small percentage of SSc sera. The Ku (p70/p80) antigen is a DNA-binding protein involved in doubled-stranded DNA repair, through the nonhomologous end-joining pathway.^[4–8] It combines with a DNA-dependent protein kinase catalytic subunit (DNA-PKcs) and regulates the phosphorylation of many nuclear proteins, including nuclear enzymes and transcription factors.^[9] It also plays a role in V(D)J recombination of receptor genes on B and T lymphocytes,^[4–8] immunoglobulin class switching,^[10] telomere protection,^[11] and development of the central nervous system.^[12]

The prevalence of anti-Ku autoantibodies in SSc varies from 1.5% to 16%,^[13–25] depending primarily on the detection immunoassay, and on the genetic and geographical background of the subjects studied.^[26] They were first described in 1981 by Mimori et al^[19] as a marker of scleroderma-polymyositis overlap syndrome, but have since been reported in a variety of other autoimmune disorders, including systemic lupus erythematosus (SLE) (0.7%–27%), idiopathic inflammatory myopathies (up to 26%), mixed connective tissue disease and undifferentiated connective tissue disease (up to 8.3%), rheumatoid arthritis (up to 16%), and Sjögren syndrome (<1%–20%), in isolation or as part of overlap syndromes,^[13,14,19,23–25,27–47] and only rarely in healthy controls.^[19,23,25] In SSc, these autoantibodies have been associated with myositis^[14,17,19,22,32,42,48] and ILD,^[14,42] and also limited cutaneous involvement,^[14,19,22] arthritis,^[14,22] and less vascular involvement.^[14,20–22] However, results have been conflicting,^[13,22,24,25] and conclusions have been limited by small numbers of subjects studied and potentially confounded by the co-presence of other SSc-related autoantibodies.

The objective of this study was therefore to identify the demographic, clinical, and serological characteristics of SSc

subjects with single-specificity anti-Ku antibodies in a large international, multicenter cohort.

2. Methods

An international (Canada, Australia, USA, Mexico) retrospective cohort of 2140 SSc subjects was formed, demographic and clinical variables were harmonized, and sera were tested for anti-Ku using a line immunoassay (LIA). Associations between single-specificity anti-Ku antibodies (i.e., in isolation of other SSc-related antibodies), baseline characteristics, and mortality were investigated.

2.1. Sources of data

The study subjects were SSc patients enrolled in the Canadian Scleroderma Research Group (CSRG), the Australian Scleroderma Interest Group (ASIG), or the American Genetics versus Environment in Scleroderma Outcome Study (GENISOS) cohorts. Briefly, subjects in the CSRG are recruited from 15 sites across Canada and Mexico, and must have a diagnosis of SSc verified by an experienced rheumatologist, be >18 years of age, and be fluent in English, French, or Spanish. Over 98% of the cohort meets the 2013 American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) classification criteria for SSc.^[49] Loss to follow-up in the CSRG cohort is 2.5%. Subjects in the ASIG are recruited by investigators from 12 Australian centers specializing in the care of patients with SSc, according to similar inclusion criteria. All subjects fulfill either the 1980 preliminary ACR criteria for classification of SSc, or the Medsger criteria for limited SSc.^[50] Estimated loss to follow-up in the ASIG cohort is 7%. The GENISOS cohort is a longitudinal cohort of subjects with early SSc. Subjects are enrolled within 5 years of disease onset as determined by the first non-Raynaud phenomenon symptom from 3 University of Texas institutions at Houston, San Antonio, and Galveston. All enrolled subjects fulfill the 2013 ACR/EULAR classification criteria for SSc.^[51] Estimated loss to follow-up in the GENISOS cohort is 2.5%.

Ethics committee approval for this study was obtained at McGill University (Montreal, Canada) and at all participating CSRG, ASIG, and GENISOS study sites. All subjects provided informed written consent to participate in the study. The subjects included in this study were those whose baseline visits were between September 2004 and June 2014 for CSRG, between January 2007 and March 2013 for ASIG, and between January 1998 and September 2012 for GENISOS, and who had complete serological profiles for anti-Ku antibodies as detected by the methods described below.

2.2. Clinical variables

Subjects recruited into this study underwent standardized medical evaluation including medical histories, physical

examinations and laboratory investigations, according to the protocols from their respective cohorts, and the following clinical variables were harmonized to create a single dataset with common variable definitions. All study variables were collected at the baseline study visit, except creatine kinase (CK) and mortality, which were also available during follow-up.

Demographic information regarding age, sex, and ethnicity was collected by patient self-report. Disease duration was recorded by study physicians and defined as the interval between the onset of the first non-Raynaud disease manifestation and baseline study visit.

Skin involvement was assessed using the modified Rodnan Skin Score (mRSS), a widely used clinical assessment where the examining rheumatologist records the degree of skin thickening ranging from 0 (no involvement) to 3 (severe thickening) in 17 areas (total score range 0–51). Limited cutaneous disease was defined as skin involvement distal to the elbows and knees with or without facial involvement; diffuse cutaneous disease was defined as skin involvement proximal to the elbows and knees with or without truncal involvement. Those with a clinical diagnosis of SSc but no skin involvement were included with the limited cutaneous subset.

History of inflammatory myositis, calcinosis, inflammatory arthritis, scleroderma renal crisis, and malignancy was recorded by a study physician. The presence of telangiectasias, digital pits, and digital ulcers on physical examination was also recorded by a study physician. CK levels were measured by local laboratories.

To assess gastrointestinal involvement, subjects answered yes/no to 6 questions concerning gastroesophageal reflux disease, dysphagia, antibiotics for bacterial overgrowth, episodes of pseudo-obstruction, fecal incontinence, and hyperalimentation.

The presence of ILD was determined using a clinical decision rule that was recently published.^[52] This algorithm considers ILD to be present if a high-resolution computed tomography (HRCT) scan of the lung was interpreted by an experienced radiologist as showing ILD, or, in the case where no HRCT is available, if either a chest x-ray was reported as showing increased interstitial markings (not thought to be due to congestive heart failure) or fibrosis, and/or if a study physician reported the presence of typical “velcro-like crackles” on physical examination. Pulmonary function tests were performed at local respiratory physiology laboratories.

Pulmonary hypertension was defined as an estimated systolic pulmonary artery pressure (sPAP) ≥ 45 mm Hg measured using the Doppler flow measurement of the tricuspid regurgitant jet on cardiac echocardiography (an estimate that correlates strongly with right heart catheter studies)^[53] for CSRG and GENISOS subjects, or mean pulmonary artery pressure (mPAP) > 25 mm Hg with a pulmonary capillary wedge pressure (PCWP) < 15 mm Hg on right heart catheterization for ASIG subjects.

In addition, disease overlap with SLE and Sjögren syndrome, history of Raynaud phenomenon, trigeminal neuralgia, and autoimmune thyroid disease, and presence of capillaroscopic alterations on dermatoscopic examination recorded by a study physician were also available in the CSRG dataset.

All subjects were assessed, followed, and classified in the same way, on a similar platform, regardless of anti-Ku antibody status and outcomes. Frequency of missing data was recorded.

2.3. Serology

Autoantibody analysis of the CSRG and GENISOS cohorts were performed in a central laboratory—Mitogen Advanced

Diagnostics Laboratory, University of Calgary—and the ASIG analyses were performed using an identical immunoassay kit and protocol. Serum aliquots were stored at -80°C until needed for diagnostic assays. Antinuclear antibodies (ANAs) were detected by indirect immunofluorescence (IIF) performed on HEp-2 cells (ImmunoConcepts, Sacramento, CA). Anti-Ku, centromere (CENP-A and CENP-B), topoisomerase I, RNA polymerase III (RP11 and RP155), fibrillar, NOR-90, Th/To, Ro52/TRIM21, PDGFR, PM75, and PM100 antibodies were detected by Euroline SSc profile LIA (Euroimmun GmbH, Luebeck, Germany) according to manufacturer’s instructions. With the intent of optimizing specificity, antibodies were reported as absent (negative, equivocal, and low titers) and present (moderate and high titers). Data on ANA titers and patterns were also available for subjects from the CSRG cohort.

2.4. Statistical analysis

Subjects were grouped according to anti-Ku status, either positive (further subdivided into single-specificity or overlapping with other SSc antibodies) or negative at baseline visit. Descriptive statistics were used to summarize the baseline demographic and clinical characteristics of the subjects. Given the exploratory nature of the analysis and the small samples in the subgroups, clinically relevant numerical differences between subgroups were considered informative. Exploratory statistical analyses were performed using chi-square tests, Fisher exact tests, and Mann–Whitney *U* tests, as indicated. $P < 0.05$ was considered statistically significant. Bonferroni correction for multiple testing was calculated for statistically significant findings. Missing data, selection bias, and information bias were addressed qualitatively.

Kaplan–Meier analysis and Cox proportional-hazard models adjusting for baseline differences in age, ethnicity, and sex were used to compare survival between autoantibody subsets. Multivariate logistic regression adjusting for baseline differences in age and ethnicity was used to determine the association between anti-Ku antibody groups and ILD. *P* values < 0.05 were considered statistically significant.

All statistical analyses were performed with SAS v.9.2 (SAS Institute, Cary, NC).

3. Results

All cohort subjects were tested for anti-Ku antibodies and were eligible for inclusion. Of the 2140 SSc subjects included in this study, 24 (1.1%) had anti-Ku antibodies. Thirteen (0.6%) had single-specificity anti-Ku antibodies (i.e., in isolation of other SSc-related antibodies), 11 (0.5%) had overlapping anti-Ku antibodies, and 2116 (98.9%) were negative for anti-Ku antibodies (Table 1). Individual clinical and serological characteristics of single-specificity and overlapping anti-Ku-positive subjects are presented in Tables 2 and 3, respectively.

3.1. Clinical correlates of single-specificity anti-Ku-positive subjects

Subjects with single-specificity anti-Ku antibodies tended to be older at disease onset (mean age 51.5 vs 45.3 years), of Hispanic ethnicity (30% vs 7%), and with limited cutaneous disease (77% vs 63%); and less likely to be of white ethnicity (70% vs 81%), have digital pitting (20% vs 49%), digital ulcers (0% vs 15%), and calcinosis (8% vs 25%), compared with anti-Ku-negative subjects.

Table 1

Baseline characteristics of the study cohort, as a group and according to anti-Ku antibody status.

	Whole group (N = 2140)			Anti-Ku-positive (n = 24)		Single-specificity anti-Ku-positive (n = 13)		Overlapping anti-Ku- positive (n = 11)		Anti-Ku- negative (n = 2116)	
	N or mean	% or SD	Missing N (%)	n or mean	% or SD	n or mean	% or SD	n or mean	% or SD	n or mean	% or SD
Sociodemographics											
Female	1845	86%	0 (0%)	21	88%	12	92%	9	82%	1824	86%
Ethnicity			98 (5%)								
White	1653	81%		14	67%	7	70%	7	64%	1639	81%
Black	75	4%		0	0%	0	0%	0	0%	75	4%
Hispanic	139	7%		5	23%	3	30%	2	18%	134	7%
Asian	62	3%		1	5%	0	0%	1	9%	61	3%
Age, y	55.1	12.6	2 (0%)	55.5	15.1	59.1	12.4	51.3	17.5	55.1	12.6
Disease duration, y	9.7	9.4	15 (1%)	10.2	9.9	8.7	8.5	11.8	11.4	9.7	9.4
Age at disease onset, y	45.3	13.7	33 (2%)	45.8	15.3	51.5	14.0	39.4	14.7	45.3	13.7
Clinical variables											
Modified Rodnan Skin Score (0–51)	10.7	10.0	56 (3%)	9.8	9.7	6.5	8.2	13.5	10.2	10.7	10.0
Limited cutaneous disease	1343	63%	7 (0%)	17	71%	10	77%	7	64%	1326	63%
Inflammatory myositis	180	9%	180 (8%)	2	9%	1	8%	1	10%	178	9%
Calcinosis	523	25%	21 (1%)	5	21%	1	8%	4	36%	518	25%
Inflammatory arthritis	595	29%	66 (3%)	7	30%	3	25%	4	36%	588	29%
Telangiectasias	1500	72%	60 (3%)	16	70%	9	69%	7	70%	1484	72%
Digital pitting	892	49%	321 (15%)	8	40%	2	20%	6	60%	884	49%
Digital ulcers	266	14%	306 (14%)	3	15%	0	0%	3	30%	263	15%
Gastrointestinal disease											
GERD/reflux	1741	82%	6 (0%)	20	83%	12	92%	8	73%	1721	82%
Dysphagia	1124	53%	27 (1%)	13	57%	7	58%	6	55%	1111	53%
Antibiotics for bacterial overgrowth	123	6%	196 (9%)	1	5%	1	10%	0	0%	122	6%
Episodes of pseudo-obstruction	63	3%	17 (1%)	0	0%	0	0%	0	0%	63	3%
Fecal incontinence	326	19%	399 (19%)	1	6%	0	0%	1	10%	325	19%
Hyperalimentation	39	3%	692 (32%)	0	0%	0	0%	0	0%	39	3%
Number of GI symptoms (0–6)	1.6	1.0	1 (0%)	1.5	0.7	1.5	0.7	1.4	0.7	1.6	1.0
Scleroderma renal crisis	76	4%	24 (1%)	0	0%	0	0%	0	0%	76	4%
Pulmonary hypertension	250	14%	379 (18%)	2	13%	2	25%	0	0%	248	14%
Interstitial lung disease*	717	34%	49 (2%)	13	57%	7	58%	6	55%	704	34%
Malignancy†	163	8%	0 (0%)	4	17%	1	8%	3	27%	159	8%
Mortality	375	18%	7 (0%)	4	17%	3	23%	1	9%	371	18%
Follow-up duration, y	5.0	3.1	0 (0%)	5.5	3.8	5.4	4.0	5.7	3.7	5.0	3.1
Serology											
Antinuclear antibody (ANA) positivity	1859	95%	181 (8%)	23	100%	12	100%	11	100%	1836	95%
Overlapping with other antibodies											
Anticentromere	741	35%	0 (0%)	5	21%	0	0%	5	45%	736	35%
Antitopoisomerase I	357	17%	0 (0%)	3	13%	0	0%	3	27%	354	17%
Anti-RNA polymerase III	300	14%	0 (0%)	2	8%	0	0%	2	18%	298	14%
Anti-Ro52/TRIM21	561	26%	0 (0%)	5	21%	0	0%	5	45%	556	26%
Laboratory tests											
CK at baseline visit	105.8	148.0	327 (15%)	122.9	106.5	136.7	144.1	107.4	40.8	105.6	148.4
% with CK >3x normal at baseline‡	20	1%	423 (20%)	1	6%	1	11%	0	0%	19	1%
Highest CK during follow-up	145.4	349.2	89 (4%)	145.2	350.9	150.6	130.4	173.6	92.2	145.2	350.9
% with CK >3x normal during follow-up	41	2%	206 (10%)	1	5%	1	10%	0	0%	40	2%
Pulmonary function tests											
DLCO, % predicted	69.0	20.8	472 (22%)	65.3	22.1	61.8	23.5	68.9	21.9	69.0	20.8
FVC, % predicted	91.2	20.4	270 (13%)	88.1	18.1	89.4	9.8	87.0	23.8	91.2	20.4
TLC, % predicted	94.0	23.8	641 (30%)	86.8	17.7	86.5	19.0	87.0	18.0	94.0	23.9

All variables represent study baseline characteristics, except for mortality and CK levels observed during follow-up. Variable comparisons with $P < 0.05$ include the following as described in the footnotes. The single-specificity anti-Ku antibody positive group was exclusive of anticentromere, topoisomerase I, RNA polymerase III, fibrillarin, NOR90, Th/To, Ro52/TRIM21, PDGFR, PM75, and PM100 antibodies. CI = confidence interval, CK = creatine kinase, DLCO = diffusion capacity for carbon monoxide, FVC = forced vital capacity, GERD = gastroesophageal reflux disease, GI = gastrointestinal, NOR = nucleolar organizer region, OR = odds ratio, PDGFR = platelet derived growth factor receptor, SD = standard deviation, TLC = total lung capacity.

* Interstitial lung disease: anti-Ku-positive versus anti-Ku-negative (OR 2.5, 95% CI 1.1–5.8, $P = 0.03$); after Bonferroni correction (95% CI 0.7–9.1). Single-specificity anti-Ku-positive versus anti-Ku-negative (OR 2.7, 95% CI 0.9–8.6, $P = 0.09$).

† Malignancy: overlapping anti-Ku-positive versus anti-Ku-negative (OR 4.6, 95% CI 1.2–17.6, $P = 0.02$); after Bonferroni correction (95% CI 0.6–36.3).

‡ Per cent with CK >3x normal at baseline: single-specificity anti-Ku-positive versus anti-Ku-negative (OR 11.1, 95% CI 1.3–92.9, $P = 0.03$); after Bonferroni correction (95% CI 0.4–294.5).

Table 2

Clinical and serological characteristics of single-specificity anti-Ku-positive subjects.

	ANA titer (1:x)	ANA pattern	Sex (M/F)	Ethnicity	Age, y	Age at disease onset, y	Limited skin disease	Myositis	Arthritis	Calcinosis	Digital			SLE		Death
											ulcers	ILD	PH	Malignancy	overlap	
1	N/A	N/A	F	White	60	56	N	N	N	N	N	Y	N	N	N/A	Y (cause N/A)
2	5120	Sp	F	White	75	63	N	N	N	N	N	Y	Y	N	N	N
3	1280; 640	Nu; Sp	F	N/A	78	76	Y	N	N	N	N	Y	Y	N	N	N
4	1280; 320	H/Sp; Nu	F	White	57	53	Y	N	N	N	N	Y	N	N	N	N
5	N/A	N/A	F	Hispanic	48	42	Y	N	N	N	N	Y	N/A	N	N/A	Y (cause N/A)
6	N/A	N/A	F	Hispanic	53	49	Y	N	Y	N	N	Y	N/A	N	N/A	N
7	N/A	N/A	F	White	51	46	Y	N	Y	N	N	Y	N	N	Y	N
8*	1280; 320	H/Sp; Cyto	F	N/A	45	N/A	Y	N	N/A	Y	N	N	N/A	N	Y	Y (cause N/A)
9	5120; 320; 160	Nu; Sp; Cyto	F	Hispanic	40	38	N	N	Y	N	N	N	N/A	N	N	N
10	N/A	N/A	F	White	79	76	Y	N/A	N	N	N	N	N	N	N/A	N
11	N/A	N/A	F	White	57	43	Y	N	N	N	N	N	N	N	N/A	N
12	N/A	N/A	F	White	65	41	Y†	N	N	N	N	N	N	Skin SCC	N/A	N
13	320	Sp	M	N/A	61	34	Y	Y	N	N	N	N/A	N/A	N	N	N

All variables represent study baseline characteristics, except for death.

ANA = antinuclear antibody, cyto = cytoplasmic, F = female, H/Sp = homogeneous and speckled, ILD = interstitial lung disease, M = male, N = no, N/A = not available, Nu = nucleolar, PH = pulmonary hypertension, SCC = squamous cell carcinoma, SLE = systemic lupus erythematosus, Sp = speckled, Y = yes.

* Subject 8 is also positive for anti-U1RNP and anti-Sm, but is classified as single-specificity as these antibodies were not excluded from the definition of single-specificity anti-Ku, and were not routinely tested in all subjects.

† Subject 12 had no skin involvement (SSc sine scleroderma).

Table 3

Clinical and serological characteristics of overlapping anti-Ku-positive subjects.

	Overlap antibodies	ANA titer (1:x)	ANA pattern	Sex (M/F)	Ethnicity	Age, y	Age at disease onset, y	Limited skin disease	Myositis	Arthritis	Calcinosis	Digital			SLE		Death
												ulcers	ILD	PH	overlap	Malignancy	
1	ACA	N/A	N/A	F	White	69	68	Y	N/A	N	Y	N	Y	N	N/A	Breast cancer	N
2	ACA	2560	Cen	F	White	44	21	N	N	N	Y	Y	N	N/A	N	N	N
3	ACA-Ro52	N/A	N/A	F	White	80	51	Y	N	N	N	N	N	N	N/A	Skin SCC	N
4	ACA-Ro52	2560	Sp	F	Asian	40	37	Y*	N	Y	N	N	N	N	N	N	N
5	ACA-ARNAP-NOR90	N/A	N/A	F	Hispanic	17	15	N	N	N	Y	N/A	N	N/A	N/A	N	N
6	ARNAP	N/A	N/A	F	White	72	48	Y	N	Y	N	N	N	N	N/A	Melanoma	N
7	ATA	5120	Sp	F	Hispanic	46	45	N	N	N	N	N	Y	N	N	N	N
8	ATA	1280; 1280	H/Sp;Nu	M	White	44	31	Y	N	Y	N	Y	Y	N	N	N	N
9	ATA-Ro52	5120	Sp	F	White	54	29	N	N	N	N	Y	Y	N	N	N	N
10	Ro52	N/A	N/A	F	White	45	44	Y	N	N	N	N	Y	N	N/A	N	N
11	PMScl-NOR90-Ro52	2560; 1280; 640	Nu; Sp; O	M	Aboriginal	54	44	Y	Y	Y	Y	N	Y	N	N	N (at baseline visit)	Metastatic rectal cancer

All variables represent study baseline characteristics, except for death.

ACA = anticentromere antibody, ANA = antinuclear antibody, ARNAP = anti-RNA polymerase III antibody, ATA = antitopoisomerase I antibody, Cen = centromere, F = female, H/Sp = homogeneous and speckled, ILD = interstitial lung disease, M = male, N = no, N/A = not available, NOR = nucleolar organizer region, Nu = nucleolar, O = others, PH = pulmonary hypertension, SCC = squamous cell carcinoma, SLE = systemic lupus erythematosus, Sp = speckled, Y = yes.

* Patient 4 had no skin involvement (SSc sine scleroderma).

Table 4

Multivariate logistic model to estimate the association between the presence of anti-Ku antibodies and ILD, adjusting for baseline demographic differences.

	Odds ratio	95% confidence intervals		P
White ethnicity	0.77	0.61	0.98	0.030
Age	1.02	1.01	1.03	0.002
Single-specificity versus negative anti-Ku subjects	2.69	0.75	9.59	0.127
Overlapping versus negative anti-Ku subjects	2.35	0.71	7.79	0.162

ILD = interstitial lung disease.

Table 5

Cox proportional-hazard model to estimate the association between the presence of anti-Ku antibodies and mortality, adjusting for baseline demographic differences.

	Hazard ratio	95% confidence intervals		P
White ethnicity	0.64	0.49	0.84	<0.001
Age	1.03	1.02	1.04	<0.001
Female	0.52	0.40	0.69	<0.001
Single-specificity versus negative anti-Ku subjects	0.97	0.24	3.92	0.968
Overlapping versus negative anti-Ku subjects	0.56	0.08	4.03	0.567

Interstitial lung disease was also more common in single-specificity anti-Ku-positive subjects than in anti-Ku-negative subjects (58% vs 34%; odds ratio [OR] 2.7, 95% confidence interval [CI] 0.9–8.6, $P=0.09$; in logistic regression analysis adjusting for differences in baseline demographic characteristics: OR 2.69, 95% CI 0.75–9.59, $P=0.13$) (Tables 1 and 4).

Pulmonary hypertension was numerically more common in single-specificity anti-Ku-positive subjects compared with anti-Ku-negative subjects (25% vs 14%; OR 2.0, 95% CI 0.4–10.0, $P=0.39$).

Although there was no difference in inflammatory myositis prevalence (8% vs 9%), subjects with single-specificity anti-Ku antibodies were more likely to have significantly elevated CK levels ($>3\times$ normal) at baseline (11% vs 1%; OR 11.1, 95% CI 1.3–92.9, $P=0.03$) and during follow-up (10% vs 2%).

Inflammatory arthritis was not more frequent in anti-Ku-positive subjects.

In a survival analysis adjusted for differences in baseline characteristics, subjects with single-specificity anti-Ku antibodies were not found to be at significantly increased risk of death compared with subjects without anti-Ku antibodies (mean [SD] follow-up of 5.0 [3.1] years) (Table 5 and Supplementary Figure, <http://links.lww.com/MD/B244>).

3.2. Exploratory findings in anti-Ku-positive subjects

Interestingly, subjects with overlapping anti-Ku antibodies were more likely to have a history of malignancy at baseline visit compared with anti-Ku-negative subjects (27% vs 8%; OR 4.6, 95% CI 1.2–17.6, $P=0.03$). The subjects with overlapping anti-Ku antibodies and malignancy had melanoma (ARNAP overlap), breast cancer (ACA overlap), and squamous cell skin cancer (ACA and anti-Ro52/TRIM21 overlap), respectively, none of which occurred within 2 years of SSc diagnosis. In comparison, the frequency of malignancy in single-specificity anti-Ku, ARNAP, and ACA-positive SSc subjects were 8.0%, 7.7%, and 8.9%, respectively.

In the CSRG cohort (comprising 7 single-specificity anti-Ku-positive and 1323 anti-Ku-negative subjects), overlap disease with SLE (28.6% vs 3.3%; OR 11.6, 95% CI 2.2–61.6, $P=0.004$) was reported more frequently in single-specificity anti-Ku subjects compared with anti-Ku-negative subjects. Frequency of Sjögren syndrome (0% vs 7.4%), trigeminal neuralgia (0% vs 2.6%), autoimmune thyroid disease (0% vs 12.3%), Raynaud phenomenon (85.7% vs 97.4%), or abnormal capillaroscopy (85.7% vs 76.5%) was not significantly different in single-specificity anti-Ku-positive subjects compared with anti-Ku-negative subjects (Supplementary Table 1, <http://links.lww.com/MD/B244>).

3.3. Serological characteristics of anti-Ku-positive subjects

All subjects with anti-Ku antibodies had positive ANA by IIF (Table 1). In the CSRG cohort, all subjects with single-specificity anti-Ku antibodies ($n=6$) had ANA titers of at least 1:320 with speckled patterns, along with nucleolar patterns in half of subjects and cytoplasmic patterns in a third of subjects (Table 2 and Supplementary Table 2, <http://links.lww.com/MD/B244>).

Among subjects with overlapping antibodies, 6 had 1 more (2 ACA, 2 ATA, 1 ARNAP, 1 Ro52/TRIM21), 3 had 2 more (2 ACA-Ro52/TRIM21, 1 ATA-Ro52/TRIM21), and 2 had 3 more (1 ACA-RNAP-Nor90, 1 PM-Scl-Nor90-Ro52/TRIM21) overlapping antibodies. Interestingly, the majority of subjects who had ILD had either ATA or anti-Ro52/TRIM21 overlapping antibodies, whereas most subjects with calcinosis had ACA overlapping antibodies (Table 3).

Of note, 1 CSRG subject classified as part of the single-specificity anti-Ku-positive group also had positive anti-U1-RNP and anti-Sm autoantibodies (which were not among the antibodies tested for the whole sample, and therefore not excluded from the definition of single-specificity anti-Ku-positivity in this study). This subject was diagnosed with SSc-SLE overlap disease and was the only single-specificity anti-Ku-positive subject who had calcinosis (Table 2).

4. Discussion

We aimed to describe the demographic, clinical, and serological characteristics of SSc subjects with single-specificity anti-Ku antibodies. In this international, multicenter cohort of 2140 subjects, only 24 (1.1%) had anti-Ku antibodies, and 13 (0.6%) had single-specificity anti-Ku antibodies. These numbers are slightly lower than previously reported frequencies using the LIA technique (Table 6) and might be attributed to our intent of optimizing specificity by using a higher cut-off (moderate and high titers only). Single-specificity anti-Ku-positive subjects in this cohort tended to be older, have more limited skin disease, and less vascular digital complications than anti-Ku-negative subjects. ILD was more frequent in anti-Ku-positive subjects in general, and also in single-specificity and overlapping anti-Ku-positive subjects. CK elevations were also more common in subjects with single-specificity anti-Ku antibodies. Serologically, single-specificity anti-Ku-positive subjects had high-titer speckled ANAs, with or without nucleolar staining patterns.

To date, little was known on the clinical correlates of single-specificity anti-Ku antibodies in SSc. In an extensive review of the literature (Table 6, Supplementary Table 4, <http://links.lww.com/MD/B244>), only 2 studies were identified that examined these antibodies in isolation. The first, by Kuwana et al,^[17] reported on

Table 6 Summary of the literature on clinical associations of anti-Ku antibodies in SSC, by method of detection.*

	Line immunoassay		Counterimmuno-electrophoresis		CIE + Immunoblot		Double immunodiffusion		Western blot Rodriguez-Reyna et al ^[21] (2011)	Immunoblot Yanava and Arnett ^[25] (1989)	Dot blot Rigollet et al ^[42] (2012)	Immuno-precipitation Kaji et al ^[48] (2014)
	Current study (2016)	Patterson et al ^[20] (2015)	Graf et al ^[16] (2012)	Villalta et al ^[24] (2012)	Cavazzana et al ^[14] (2013)	Rozman et al ^[22] (2008)	Cavazzana et al ^[13] (2008)	Franoschini et al ^[32] (2002)				
Study of single-specificity anti-Ku	Yes	No	No	No	No	No	No	Yes	No	No	No	Yes
Total study population	2140	505	129	210	560	625	379	275	66	66	N/A	N/A
Subpopulation positive for anti-Ku	24 (1.1%)	14 (2.8%)	6 (5%)	10 (4.7%)	13 (2.3%)	14 (2.2%)	8 (2.1%)	8 (2.9%)	8 (12%)	14 (10.1%)	9 (16.1%)	N/A
Subpopulation positive for single-specificity anti-Ku [†]	13 (0.6%)	3 (0.6%)	2 (1.6%)	6 (2.9%)	10 (1.8%)	10 (1.6%)	6 (1.6%)	7 (2.5%)	0	N/A	3 (5%)	44
Comparison group	Negative for anti-Ku	Negative for anti-Ku	Uncertain	Negative for anti-Ku	Positive for ACA	Negative for anti-Ku, age-sex-matched	Negative for anti-Ku	Negative for anti-Ku	SSc-PM negative for anti-Ku	Negative for anti-Ku	Positive for anti-Ku but not SSC	Positive for anti-PM-Sc1, RuvBL1/2
Sociodemographics												
Male, %	8 vs 14	N/A	40	N/A	23 vs 1	Sex-matched	0 vs 11	0 vs 12	N/A	N/A	22	18
Comparison	↓				↑		↓	↓				
White, %	70 vs 81	N/A	N/A	100 vs 100	100 vs 100	100	N/A	Japanese	Mexican-Mestizo	67W, 11H, 22 Black	N/A	59
Comparison	↓					Age-matched	62 vs 50	30 vs 42	N/A	N/A	N/A	38
Age at disease onset, years	52 vs 45	N/A	50	N/A	64.1		↑	↓				
Comparison	↑		↑									
Clinical variables												
Limited cutaneous disease, %	77 vs 63	N/A	50 [‡]	60	100 vs 99	86 vs 72	N/A	0 [‡]	100 [§]	71	33 [‡]	55
Comparison	↑			↔	↔	↑	N/A	N/A	N/A	↔	N/A	
Skin (mRSS), mean	6.5 vs 10.7	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Comparison	↓											
Inflammatory myositis, %	8 vs 9	N/A	17	↔	85 vs 0	43 vs 7	88	100 vs 10	88	11	N/A	50
Comparison	↔		↑	↔	↑	↑	↑	↑	↔	↔	N/A	N/A
Elevated muscle enzymes, %	11 vs 1	N/A	N/A	N/A	N/A	43 vs 2	N/A	N/A	100	N/A	N/A	N/A
Comparison	↑					↑						
Calcinosis, %	8 vs 25	N/A	N/A	N/A	N/A	21 vs 21	N/A	N/A	N/A	N/A	N/A	N/A
Comparison	↓					↔						
Inflammatory arthritis, %	25 vs 29	N/A	N/A	↔	39 vs 12	43 vs 21	N/A	43 vs 39	13	N/A	N/A	N/A
Comparison	↔			↔	↑	↑	N/A	↔	↓	↔	N/A	N/A
Telangiectasias, %	69 vs 72	71 vs 90	N/A	N/A	↑	36 vs 67	N/A	N/A	N/A	N/A	N/A	N/A
Comparison	↔	↓				↓						
Digital pits, %	20 vs 49	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Comparison	↓											
Digital ulcers, %	0 vs 15	N/A	N/A	N/A	7.7 vs 55	29 vs 54	N/A	0 vs 31	N/A	N/A	N/A	N/A
Comparison	↓				↓	↓	N/A	↓	N/A	↔	N/A	N/A
Gastrointestinal disease, %	92 vs 82 (GERD)	N/A	N/A	↔	69 vs 72 (dysphagia)	71 vs 63 (esophageal)	N/A	71 vs 63 (esophageal)	N/A	↓	N/A	73
Comparison	↑			↔	↔	↑						
Scleroderma renal crisis, %	0 vs 4	N/A	N/A	↔	N/A	0 vs 0	N/A	0 vs 4	N/A	↔	N/A	2
Comparison	↔			↔		↔		↔				
Pulmonary hypertension, %	25 vs 14	N/A	N/A	N/A	N/A	46 vs 26	N/A	0 vs 6	N/A	↔	N/A	14
Comparison	↑					↑		↔	75 vs 23	↑	N/A	
Interstitial lung disease, %	58 vs 34	N/A	17	10	53.8 vs 4.7	57 vs 56 (CXR)	37.5 vs 21.4	57 vs 54	N/A	↑	N/A	43
Comparison	↑		↔	↔	↑	↔	↑	↔		↔		
Cardiac involvement, %	N/A	N/A	N/A	N/A	N/A	7 to 21	N/A	0 vs 6	N/A	↔	N/A	20
Comparison	↔					↔		↔	50 vs 7	↑	N/A	
Mortality, %	23 vs 18	N/A	N/A	N/A	N/A	N/A	37.5 vs 9.7	0	N/A	↔	N/A	N/A

	Line immunoassay		Counterimmunoelectrophoresis		CIE- μ -immunoblot		Double immunodiffusion		Western blot	Immunoblot	Dot blot	Immuno-precipitation		
	Current study (2016)	Patterson et al ^[20] (2015)	Graf et al ^[16] (2012)	Vilalta et al ^[24] (2012)	Cavazzana et al ^[14] (2013)	Rozman et al ^[25] (2008)	Cavazzana et al ^[13] (2008)	Franceschini et al ^[22] (2002)	Kuwana et al ^[17] (1994)	Mimori et al ^[19] (1981)	Rodriguez-Reyna et al ^[21] (2011)	Yaneva and Arnett ^[26] (1989)	Rigolet et al ^[24] (2012)	Kaji et al ^[48] (2014)
Comparison	↔													
Malignancy, %	8 vs 8	N/A	N/A	N/A	N/A	N/A	N/A	N/A	↓	N/A	N/A	N/A	N/A	N/A
Comparison	↔													
Peripheral neuropathy, %	0 vs 3 (TN)	N/A	N/A	N/A	N/A	29 vs 19	N/A	N/A	0 vs 6	N/A	N/A	N/A	29	N/A
Comparison	↔					↑			↔					

CXR = chest x-ray, CIE = counter immunoelectrophoresis, GERD = gastroesophageal reflux disease, H = Hispanic, mRSS = modified Rodnan Skin Score, N/A = not available, TN = trigeminal neuralgia, W = white.
 * Summary of literature was limited to studies reporting at least 5 anti-Ku-positive subjects with SSc.
 † Definition of single-specificity anti-Ku-positive subpopulation for each study is provided in Supplementary Table 4.
 ‡ Limited disease was exclusive of overlap disease, in which cutaneous involvement was not specified.
 § Skin involvement was confined to face and/or extremities; no further detail was provided.

7 Japanese subjects with SSc and single-specificity anti-Ku antibodies; in contrast to our study, they were found to have a younger age of disease onset compared with anti-Ku-negative subjects. All were classified as overlap SSc syndromes, all were associated with skeletal muscle involvement, and none had digital tip ischemia. The second study, by Kaji et al,^[48] compared 44 SSc subjects (from 2 Japanese institutions and the University of Pittsburgh) with single-specificity anti-Ku antibodies with anti-RuvBL1/2 and anti-PM-Scl-positive subjects, all considered as related to SSc/myositis overlap; 50% had inflammatory myositis and 43% had ILD. However, no comparison with triple-negative subjects was provided. Three additional studies from European centers^[13,14,22] reported on anti-Ku-positive SSc subjects in whom the majority (>70%) had single-specificity anti-Ku antibodies. Subjects were found to be older at disease onset,^[13] have more limited cutaneous involvement,^[14,22] inflammatory myositis,^[13,14,22] inflammatory arthritis,^[14,22] and trigeminal neuralgia,^[22] and less digital vascular complications.^[14,22] ILD was also more frequent, but was characterized by mild functional impairment.^[13,14]

On the contrary, lung disease associated with anti-Ku-positive myositis with or without SSc overlap has been reported to be corticosteroid refractory in 75% of subjects by Rigolet et al.^[42] Furthermore, studies of anti-Ku-positive inflammatory myopathy subjects in which at least half had single-specificity autoantibodies have also shown an association with more ILD,^[42,44] inflammatory arthritis,^[36,44] overlap with other rheumatic diseases,^[36,44] and milder inflammatory myopathies, as evidenced by less frequent dermatomyositis rash,^[31,45] modest CK elevations,^[31,42] nonspecific abnormalities on muscle biopsy,^[31] and treatment-responsive, monophasic course of muscle disease.^[42,54]

Our findings are generally consistent with previously reported clinical associations between anti-Ku autoantibodies in SSc and limited cutaneous involvement, ILD, and less vascular complications, and strengthen these findings by showing an association with single-specificity anti-Ku autoantibodies in a large multi-center patient sample.

Pulmonary hypertension has been previously reported to be associated with anti-Ku antibodies. Rodriguez-Reyna et al^[21] found that 73% of anti-Ku-positive SSc subjects had pulmonary arterial hypertension, compared with only 23% of anti-Ku-negative SSc subjects. We also found a higher rate of pulmonary hypertension among the single-specificity anti-Ku subjects (25%), although this observation was based on only 2 subjects.

Other features previously associated with anti-Ku autoantibodies, such as higher rates of myositis overlap, were not clearly observed in our cohort: only 8% of single-specificity anti-Ku-positive subjects had inflammatory myositis, which was similar to other subgroups. However, CK elevations were more common in the single-specificity anti-Ku subjects, although this observation was based on only 1 subject. Interestingly, most of the studies that reported very high rates of myositis overlap (71%–90%) identified anti-Ku-positive subjects through screening of sera positive for autoantibodies to extractable nuclear antigens (ENAs).^[13,14,19,32,42,55] On the contrary, studies that analyzed anti-Ku-positive subjects by screening a SSc population with a LIA technique such as in our study, did not report such a high prevalence of myositis.^[16,24] Furthermore, Cooley et al^[29] previously observed that anti-Ku-positive subjects who met classification criteria for a connective tissue disease tended to meet the minimum number of criteria. This was again demonstrated in the study by Hausmanowa-Petrusewicz

et al,^[36] in which 5 anti-Ku-positive subjects had sclerodactyly and telangiectasias, but only 2 were identified as scleromyositis or SSc-polymyositis overlap, the 2 others remaining “unclassified,” given that their muscle and SSc-spectrum diseases did not meet classification criteria. As such, it is possible that a number of subjects with anti-Ku-positivity and a diagnosis of inflammatory myositis may present milder clinical features of SSc, such as sclerodactyly, puffy fingers, Raynaud phenomenon, or esophageal dysmotility, which are all suggestive of an overlap disease with SSc, but in isolation may not be classified as SSc disease. Therefore, these subjects may not be referred and captured into a cohort of SSc, which could explain the lower frequency of myositis in association with anti-Ku antibodies in SSc cohort studies.

Interestingly, SSc-SLE disease overlap was more frequent in single-specificity anti-Ku-positive subjects in our cohort; anti-Ku antibodies in SLE subjects have not previously been associated with a particular clinical phenotype, except for African ethnicity.^[13,14,25,33] On the contrary, trigeminal neuralgia and autoimmune thyroid disease were not seen in anti-Ku-positive subjects, contrary to what has been observed with anti-Ku in a few case reports and series.^[22,42,56–58]

Of note, clinical characteristics of single-specificity and overlapping subjects tended to be distinct on many levels: age, cutaneous extent, frequency of digital vascular complications, calcinosis and malignancy, and mortality. This dataset highlights the importance of studying single-specificity autoantibodies, as overlapping subjects may present a different, likely mixed phenotype.

All international subjects with anti-Ku antibodies had ANA by IIF, and all CSRG subjects with single-specificity anti-Ku autoantibodies had titers of at least 1:320, all of speckled with or without nucleolar staining patterns (Table 2). This is consistent with anti-Ku’s serological characteristics previously reported in the literature (Supplementary Table 3, <http://links.lww.com/MD/B244>). We also observed that anti-Ku-positive subjects generally had less frequent concomitant ACA (21% vs 35%); this is consistent with findings by Rozman et al,^[22] who found decreased concomitant ACA and ATA autoantibodies in anti-Ku-positive subjects. On the contrary, Graf et al^[16] found increased association between anti-Ku and antifibrillar (or U3RNP) autoantibodies, whereas none of our anti-Ku-positive subjects had this autoantibody (Supplementary Table 3, <http://links.lww.com/MD/B244>). It is acknowledged, however, that the LIA used in our international cohort has a low sensitivity as compared with immunoprecipitation and other immunoassays (M.J. Fritzler, unpublished data, April 2016).

The role of anti-Ku in the pathophysiology of autoimmune diseases is not entirely understood. The autoantibody-binding target Ku is known to be involved in double-stranded DNA repair.^[4–8] Schild-Poulter et al^[23] found that anti-Ku autoantibodies were often associated with autoantibodies directed against other DNA repair proteins, and suggested that B-cell responses to latent or persistent DNA damage may be involved at the onset or during the development of autoimmunity in certain systemic autoimmune rheumatic diseases. Hypoxia has also been reported to induce chromatin modifications, leading to recruitment and activation of the DNA-dependent protein kinase (which includes Ku)^[59]; this is of interest given that vasculopathy and tissue hypoxia are thought to be part of the initial pathophysiology of SSc. Ku is also involved in V(D)J recombination of receptor genes on B and T lymphocytes^[4–8] and in immunoglobulin class switching^[10]; one could hypothesize that defective expression of

the Ku peptide may lead to altered function of the immune system and result in autoimmunity as well. Genetic background seems to play a role in anti-Ku autoimmunity, as evidenced by its positive association with certain human leukocyte antigen (HLA) class II genotypes.^[55,60] Molecular mimicry between the Ku antigen and certain fungal proteins has also been postulated as a potential trigger for anti-Ku autoimmunity in genetically predisposed individuals.^[61]

Finally, overlapping anti-Ku antibodies were found to be associated with a history of malignancy in our study. To our knowledge, this association has not been reported previously (Table 6). Interestingly, malignancy has been hypothesized to act as a trigger of autoimmunity in certain cases of SSc, particularly in the ARNAP-positive subset, via mechanisms involving antitumor immunity, molecular mimicry, and epitope spreading.^[62,63] In fact, tumors associated with ARNAP-positive SSc have been shown to harbor mutated forms of the RNA polymerase III autoantigen.^[64] In the same way, genetic alterations of the Ku antigen in tumor cells could explain the association between malignancy and overlapping anti-Ku antibodies in SSc. Alternatively, overexpression of repair proteins (such as Ku) in response to DNA damage intrinsic to cancer cells, or defective expression of Ku leading to both uncontrolled tumor expansion and immune system dysfunction, constitute other hypotheses to link anti-Ku autoimmunity and malignancy.

This study has some limitations. Inflammatory myositis was not defined using specific criteria. Instead, a study physician reported its presence or absence. However, the fact that all study physicians were experienced rheumatologists supports the validity of this diagnosis. Nevertheless, mild myositis may have been overlooked. Similarly, defining ILD in the context of longitudinal observational cohort studies is very complex, given issues of missing data and verification bias. We defined ILD using a clinical decision rule that was recently published.^[52] Data on right heart catheterization was not systematically collected in all subjects. Nevertheless, in those without right heart catheterization, we defined pulmonary hypertension using a high cut-off for pulmonary systolic pressure on echocardiogram that has been shown to correlate strongly with right heart catheter studies.^[53] Still, we acknowledge that pulmonary hypertension based on echocardiogram is not synonymous with pulmonary arterial hypertension, and that some of those with pulmonary hypertension based on echocardiogram may have had other causes of pulmonary hypertension such as left heart disease or parenchymal lung disease. Thus, measurement error may have contributed to some of the negative findings of the study.

In addition, the LIA used in this study to detect anti-Ku does not distinguish reactivity to the p70 and/or p80 subunits. In an international cohort study of 73 anti-Ku-positive subjects with different connective tissue diseases, 21 of whom had SSc, Lakota et al^[39] found a positive association between females with anti-Ku-p70 and joint/bone features (defined as synovitis, joint contractions, erosive arthritis, and acroosteolysis), but a negative association between females with anti-Ku-p80 and joint/bone features. Furthermore, in SSc, 38% had isolated anti-Ku-p70, 10% had isolated anti-Ku-p80, and 43% had both.^[39] Yaneva and Arnett^[25] also found that anti-Ku-p86 levels were highest at the onset of disease and decreased and plateaued over the following years, whereas anti-Ku-p70 antibody levels remained fairly constant. The clinical phenotype described in our study could thus still represent a mixture of 2 clinical phenotypes. Further stratifying by reactivity to the 2 subunits could lead to the identification of a “purer” phenotype associated with anti-Ku

subunits, similar to the findings that anti-PM-75 and anti-PM-100 are associated with distinct phenotypes.^[2] However, the rarity of these antibodies poses an enormous challenge to further stratification.

Furthermore, subjects identified as having “single-specificity” anti-Ku antibodies may in fact have had other autoantibodies that were not detected by the LIA employed in this study. This might include some associated with connective tissue disease-related ILD such as anti-Jo1, or markers of overlapping disease such as U1RNP, which were only available for a subset of the cohort. However, in the CSRG cohort, we have previously reported a very low prevalence of anti-Jo1 antibodies (approximately 1%),^[63] and previous studies have not reported any association between anti-Ku and anti-Jo1 or other antisynthetase antibodies in SSc.^[14,22,42] As for U1RNP, 1 case of overlapping Ku and U1RNP autoantibodies out of 13 anti-Ku-positive subjects (8%) was detected in the CSRG cohort, which is similar to the 7% to 13% frequency reported elsewhere.^[17,19,22] This subject was classified as having single-specificity anti-Ku antibodies, and was reported to have calcinosis, SLE overlap, and a fatal outcome (cause unknown). If this subject were to be removed from analyses, the single-specificity anti-Ku-positive group would display absence of calcinosis, increased frequency of SLE (17% vs 3%), and similar mortality (15% vs 18%) as anti-Ku-negative subjects, which is no different from current conclusions. Thus, the presence of these autoantibodies is unlikely to have influenced the results of this study in a meaningful manner.

Additional limitations in cohort studies include missing data (>10% for certain variables, such as pulmonary hypertension, CK levels, and pulmonary function test results) and loss to follow-up. However, given that data collection and follow-up were performed on a similar platform irrespective of anti-Ku status, missing data and loss to follow-up could be considered to be missing completely at random. Also, this cohort was composed predominantly of ambulatory patients with mean disease duration of 9.7 years. Thus, it lacks some generalizability for patients with early-onset disease, for those who may have died earlier in the course of their disease, and for seriously ill patients requiring hospitalization. Nevertheless, about a third of the cohort had disease duration of 5 years or less, and the whole cohort is representative of the majority of SSc patients seen in clinical practice. A final limitation of this study is the fact that most reported associations did not reach statistical significance. Given the exploratory nature of the analysis and the small samples in the subgroups, clinically relevant numerical differences were considered informative. Still, it remains possible that some of our findings occurred by chance alone.

On the contrary, when dealing with uncommon serological profiles (there were only 0.6% of subjects with single-specificity anti-Ku antibodies), large well-phenotyped cohorts are required to begin to fill important gaps in knowledge. In the end, the limitations of our data are counter-balanced by its strengths, which include large sample size and detailed clinical phenotypic data.

In conclusion, this is the largest cohort to date focusing on the prevalence and disease characteristics of single-specificity anti-Ku antibodies in subjects with SSc. In our international cohort, anti-Ku antibodies were rare, being found in only 1.1% of subjects. Nevertheless, as a clinician, if faced with a SSc patient who presents a milder cutaneous, vascular and possibly muscular disease phenotype, and who has strongly positive speckled ANA with or without nucleolar pattern, but has an otherwise negative

panel for other tested SSc-specific antibodies anti-Ku could be suspected and tested for. If positive, increased clinical vigilance for ILD screening may be warranted. On the other hand, the usually mild clinical features and lack of survival difference associated with this autoantibody would be reassuring in terms of prognosis. Due to the very rare presence of these antibodies and thus the small size of the single-specificity anti-Ku sample, these results need to be interpreted with caution. International collaborations are key to understanding the clinical correlates of uncommon serological profiles in SSc.

Acknowledgments

Investigators of the Canadian Scleroderma Research Group: J. Pope, London, Ontario; M. Baron, Montreal, Quebec; J. Markland, Saskatoon, Saskatchewan; D. Robinson, Winnipeg, Manitoba; N. Jones, Edmonton, Alberta; N. Khalidi, Hamilton, Ontario; P. Docherty, Moncton, New Brunswick; E. Kaminska, Calgary, Alberta; A. Masetto, Sherbrooke, Quebec; E. Sutton, Halifax, Nova Scotia; J-P. Mathieu, Montreal, Quebec; M. Hudson, Montreal, Quebec; S. Ligier, Montreal, Quebec; T. Grodzicky, Montreal, Quebec; S. LeClereq, Calgary, Alberta; C. Thorne, Newmarket, Ontario; G. Gyger, Montreal, Quebec; D. Smith, Ottawa, Ontario; P.R. Fortin, Quebec, Quebec; M. Larché, Hamilton, Ontario; T.S. Rodriguez-Reyna, Mexico City, Mexico; A.R. Cabral, Mexico City, Mexico; M.J. Fritzler, Advanced Diagnostics Laboratory, Calgary, Alberta.

Investigators of the Australian Scleroderma Interest Group: C. Hill, Adelaide, South Australia; S. Lester, Adelaide, South Australia; P. Nash, Sunshine Coast, Queensland; M. Nikpour, Melbourne, Victoria; J. Roddy, Perth, Western Australia; K. Patterson, Adelaide, South Australia; S. Proudman, Adelaide, South Australia; M. Rischmueller, Adelaide, South Australia; J. Sahhar, Melbourne, Victoria; W. Stevens, Melbourne, Victoria; J. Walker, Adelaide, South Australia; J. Zochling, Hobart, Tasmania.

Investigators of GENISOS: Shervin Assassi, Houston, Texas; Maureen D. Mayes, Houston, Texas; Terry A. McNearney, Galveston, Texas; Gloria Salazar, Houston, Texas.

References

- [1] Wodkowski M, Hudson M, Proudman S, et al. Monospecific anti-Ro52/TRIM21 antibodies in a tri-nation cohort of 1574 systemic sclerosis subjects: evidence of an association with interstitial lung disease and worse survival. *Clin Exp Rheumatol* 2015;33(4 suppl 91):S131–135.
- [2] Wodkowski M, Hudson M, Proudman S, et al. Clinical correlates of monospecific anti-PM75 and anti-PM100 antibodies in a tri-nation cohort of 1574 systemic sclerosis subjects. *Autoimmunity* 2015; 48:542–51.
- [3] D’Aoust J, Hudson M, Tatibouet S, et al. Clinical and serologic correlates of anti-PM/Scl antibodies in systemic sclerosis: a multicenter study of 763 patients. *Arthr Rheumatol (Hoboken, NJ)* 2014;66:1608–15.
- [4] Boubnov NV, Hall KT, Wills Z, et al. Complementation of the ionizing radiation sensitivity, DNA end binding, and V(D)J recombination defects of double-strand break repair mutants by the p86 Ku autoantigen. *Proc Natl Acad Sci USA* 1995;92:890–4.
- [5] Mimori T. Clinical significance of anti-Ku autoantibodies: a serologic marker of overlap syndrome? *Internal Med (Tokyo, Japan)* 2002; 41:1096–8.
- [6] Rathmell WK, Chu G. Involvement of the Ku autoantigen in the cellular response to DNA double-strand breaks. *Proc Natl Acad Sci USA* 1994;91:7623–7.
- [7] Smider V, Rathmell WK, Lieber MR, et al. Restoration of X-ray resistance and V(D)J recombination in mutant cells by Ku cDNA. *Science (New York, NY)* 1994;266:288–91.
- [8] Taccioli GE, Gottlieb TM, Blunt T, et al. Ku80: product of the XRCC5 gene and its role in DNA repair and V(D)J recombination. *Science (New York, NY)* 1994;265:1442–5.

- [9] Gottlieb TM, Jackson SP. The DNA-dependent protein kinase: requirement for DNA ends and association with Ku antigen. *Cell* 1993;72:131–42.
- [10] Casellas R, Nussenzweig A, Wuerffel R, et al. Ku80 is required for immunoglobulin isotype switching. *EMBO J* 1998;17:2404–11.
- [11] Boulton SJ, Jackson SP. Components of the Ku-dependent non-homologous end-joining pathway are involved in telomeric length maintenance and telomeric silencing. *EMBO J* 1998;17:1819–28.
- [12] Bakalkin G, Yakovleva T, Hurd YL, et al. Autoantigen Ku in the brain. Developmentally regulated expression and subcellular localization. *Neuroreport* 1998;9:2147–51.
- [13] Cavazzana I, Ceribelli A, Quinzanini M, et al. Prevalence and clinical associations of anti-Ku antibodies in systemic autoimmune diseases. *Lupus* 2008;17:727–32.
- [14] Cavazzana I, Fredi M, Taraborelli M, et al. A subset of systemic sclerosis but not of systemic lupus erythematosus is defined by isolated anti-Ku autoantibodies. *Clin Exp Rheumatol* 2013;31(2 suppl 76):118–21.
- [15] Chang WS, Schollum J, White DH, et al. A cross-sectional study of autoantibody profiles in the Waikato systemic sclerosis cohort, New Zealand. *Clin Rheumatol* 2015;34:1921–7.
- [16] Graf SW, Hakendorf P, Lester S, et al. South Australian Scleroderma Register: autoantibodies as predictive biomarkers of phenotype and outcome. *Int J Rheum Dis* 2012;15:102–9.
- [17] Kuwana M, Kaburaki J, Okano Y, et al. Clinical and prognostic associations based on serum antinuclear antibodies in Japanese patients with systemic sclerosis. *Arthr Rheum* 1994;37:75–83.
- [18] Low AH, Wong S, Thumboo J, et al. Evaluation of a new multi-parallel line immunoassay for systemic sclerosis-associated antibodies in an Asian population. *Rheumatology* (Oxford, England) 2012;51:1465–70.
- [19] Mimori T, Akizuki M, Yamagata H, et al. Characterization of a high molecular weight acidic nuclear protein recognized by autoantibodies in sera from patients with polymyositis-scleroderma overlap. *J Clin Invest* 1981;68:611–20.
- [20] Patterson KA, Roberts-Thomson PJ, Lester S, et al. Interpretation of an extended autoantibody profile in a well-characterized Australian systemic sclerosis (scleroderma) cohort using principal components analysis. *Arthr Rheumatol* (Hoboken, NJ) 2015;67:3234–44.
- [21] Rodriguez-Reyna TS, Hinojosa-Azaola A, Martinez-Reyes C, et al. Distinctive autoantibody profile in Mexican Mestizo systemic sclerosis patients. *Autoimmunity* 2011;44:576–84.
- [22] Rozman B, Cucnik S, Sodin-Semrl S, et al. Prevalence and clinical associations of anti-Ku antibodies in patients with systemic sclerosis: a European EUSTAR-initiated multi-centre case-control study. *Ann Rheum Dis* 2008;67:1282–6.
- [23] Schild-Poulter C, Su A, Shih A, et al. Association of autoantibodies with Ku and DNA repair proteins in connective tissue diseases. *Rheumatology* (Oxford, England) 2008;47:165–71.
- [24] Villalta D, Imbastaro T, Di Giovanni S, et al. Diagnostic accuracy and predictive value of extended autoantibody profile in systemic sclerosis. *Autoimmun Rev* 2012;12:114–20.
- [25] Yaneva M, Arnett FC. Antibodies against Ku protein in sera from patients with autoimmune diseases. *Clin Exp Immunol* 1989;76:366–72.
- [26] Belizna C, Henrion D, Beucher A, et al. Anti-Ku antibodies: clinical, genetic and diagnostic insights. *Autoimmun Rev* 2010;9:691–4.
- [27] Aguila LA, Lopes MR, Pretti FZ, et al. Clinical and laboratory features of overlap syndromes of idiopathic inflammatory myopathies associated with systemic lupus erythematosus, systemic sclerosis, or rheumatoid arthritis. *Clin Rheumatol* 2014;33:1093–8.
- [28] Chinoy H, Fertig N, Oddis CV, et al. The diagnostic utility of myositis autoantibody testing for predicting the risk of cancer-associated myositis. *Ann Rheum Dis* 2007;66:1345–9.
- [29] Cooley HM, Melny BJ, Gleeson R, et al. Clinical and serological associations of anti-Ku antibody. *J Rheumatol* 1999;26:563–7.
- [30] Cruellas MG, Viana Vdos S, Levy-Neto M, et al. Myositis-specific and myositis-associated autoantibody profiles and their clinical associations in a large series of patients with polymyositis and dermatomyositis. *Clinics* (Sao Paulo, Brazil) 2013;68:909–14.
- [31] Fernandez C, Bardin N, De Paula AM, et al. Correlation of clinicoserologic and pathologic classifications of inflammatory myopathies: study of 178 cases and guidelines for diagnosis. *Medicine* 2013;92:15–24.
- [32] Franceschini F, Cavazzana I, Generali D, et al. Anti-Ku antibodies in connective tissue diseases: clinical and serological evaluation of 14 patients. *J Rheumatol* 2002;29:1393–7.
- [33] Fredi M, Cavazzana I, Quinzanini M, et al. Rare autoantibodies to cellular antigens in systemic lupus erythematosus. *Lupus* 2014;23:672–7.
- [34] Ghirardello A, Rampudda M, Ekholm L, et al. Diagnostic performance and validation of autoantibody testing in myositis by a commercial line blot assay. *Rheumatology* (Oxford, England) 2010;49:2370–4.
- [35] Ghirardello A, Zampieri S, Tarricone E, et al. Clinical implications of autoantibody screening in patients with autoimmune myositis. *Autoimmunity* 2006;39:217–21.
- [36] Hausmanowa-Petrusewicz I, Kowalska-Oledzka E, Miller FW, et al. Clinical, serologic, and immunogenetic features in Polish patients with idiopathic inflammatory myopathies. *Arthr Rheum* 1997;40:1257–66.
- [37] Hirakata M, Mimori T, Akizuki M, et al. Autoantibodies to small nuclear and cytoplasmic ribonucleoproteins in Japanese patients with inflammatory muscle disease. *Arthr Rheum* 1992;35:449–56.
- [38] Koening M, Fritzler MJ, Targoff IN, et al. Heterogeneity of autoantibodies in 100 patients with autoimmune myositis: insights into clinical features and outcomes. *Arthr Res Ther* 2007;9:R78.
- [39] Lakota K, Thallinger GG, Sodin-Semrl S, et al. International cohort study of 73 anti-Ku-positive patients: association of p70/p80 anti-Ku antibodies with joint/bone features and differentiation of disease populations by using principal-components analysis. *Arthr Res Ther* 2012;14:R2.
- [40] Lega JC, Fabien N, Reynaud Q, et al. The clinical phenotype associated with myositis-specific and associated autoantibodies: a meta-analysis revisiting the so-called antisynthetase syndrome. *Autoimmun Rev* 2014;13:883–91.
- [41] Reeves WH. Use of monoclonal antibodies for the characterization of novel DNA-binding proteins recognized by human autoimmune sera. *J Exp Med* 1985;161:18–39.
- [42] Rigolet A, Musset L, Dubourg O, et al. Inflammatory myopathies with anti-Ku antibodies: a prognosis dependent on associated lung disease. *Medicine* 2012;91:95–102.
- [43] Selva-O'Callaghan A, Labrador-Horrillo M, Solans-Laqué R, et al. Myositis-specific and myositis-associated antibodies in a series of eighty-eight Mediterranean patients with idiopathic inflammatory myopathy. *Arthr Rheum* 2006;55:791–8.
- [44] Suzuki S, Yonekawa T, Kuwana M, et al. Clinical and histological findings associated with autoantibodies detected by RNA immunoprecipitation in inflammatory myopathies. *J Neuroimmunol* 2014;274:202–8.
- [45] Troyanov Y, Targoff IN, Payette MP, et al. Redefining dermatomyositis: a description of new diagnostic criteria that differentiate pure dermatomyositis from overlap myositis with dermatomyositis features. *Medicine* 2014;93:318–32.
- [46] Vancsa A, Gergely L, Ponyi A, et al. Myositis-specific and myositis-associated antibodies in overlap myositis in comparison to primary dermatopolymyositis: Relevance for clinical classification: retrospective study of 169 patients. *Joint Bone Spine* 2010;77:125–30.
- [47] Wang J, Satoh M, Kabir F, et al. Increased prevalence of autoantibodies to ku antigen in African American versus white patients with systemic lupus erythematosus. *Arthr Rheum* 2001;44:2367–70.
- [48] Kaji K, Fertig N, Medsger TAJr, et al. Autoantibodies to RuvBL1 and RuvBL2: a novel systemic sclerosis-related antibody associated with diffuse cutaneous and skeletal muscle involvement. *Arthr Care Res* 2014;66:575–84.
- [49] Alhajeri H, Hudson M, Fritzler M, et al. 2013 American College of Rheumatology/European League against rheumatism classification criteria for systemic sclerosis outperform the 1980 criteria: data from the Canadian Scleroderma Research Group. *Arthr Care Res* 2015;67:582–7.
- [50] Nikpour M, Hissaria P, Byron J, et al. Prevalence, correlates and clinical usefulness of antibodies to RNA polymerase III in systemic sclerosis: a cross-sectional analysis of data from an Australian cohort. *Arthr Res Ther* 2011;13:R211.
- [51] van den Hoogen F, Khanna D, Fransen J, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League against Rheumatism collaborative initiative. *Arthr Rheum* 2013;65:2737–47.
- [52] Steele R, Hudson M, Lo E, et al. Clinical decision rule to predict the presence of interstitial lung disease in systemic sclerosis. *Arthr Care Res* 2012;64:519–24.
- [53] Hsu VM, Moreyra AE, Wilson AC, et al. Assessment of pulmonary arterial hypertension in patients with systemic sclerosis: comparison of noninvasive tests with results of right-heart catheterization. *The Journal of rheumatology* 2008;35:458–65.
- [54] Troyanov Y, Targoff IN, Tremblay JL, et al. Novel classification of idiopathic inflammatory myopathies based on overlap syndrome features and autoantibodies: analysis of 100 French Canadian patients. *Medicine* 2005;84:231–49.

- [55] Hirakata M, Suwa A, Kuwana M, et al. Association between autoantibodies to the Ku protein and DPB1*. *Arthr Rheum* 2005;52:668–9.
- [56] Gryga K, Milewski M, Zolcinski M, et al. Anti-Ku autoantibodies: series of 5 cases. *Polskie Archiwum Medycyny Wewnetrznej* 2009;119:95–7.
- [57] Kamei N, Yamane K, Yamashita Y, et al. Anti-Ku antibody-positive scleroderma-dermatomyositis overlap syndrome developing Graves' disease and immune thrombocytopenic purpura. *Internal Med (Tokyo, Japan)* 2002;41:1199–203.
- [58] Sutter R, Mengiardi B, Benz D, et al. Facial nerve palsy and anti-Ku autoantibodies. *J Neurol* 2012;259:361–3.
- [59] Bouquet F, Ousset M, Biard D, et al. A DNA-dependent stress response involving DNA-PK occurs in hypoxic cells and contributes to cellular adaptation to hypoxia. *J Cell Sci* 2011;124(Pt 11):1943–51.
- [60] O'Hanlon TP, Carrick DM, Targoff IN, et al. Immunogenetic risk and protective factors for the idiopathic inflammatory myopathies: distinct HLA-A, -B, -Cw, -DRB1, and -DQA1 allelic profiles distinguish European American patients with different myositis autoantibodies. *Medicine* 2006;85:111–27.
- [61] Guarneri F, Guarneri B, Vaccaro M, et al. The human Ku autoantigen shares amino acid sequence homology with fungal, but not bacterial and viral, proteins. *Immunopharmacol Immunotoxicol* 2011;33:329–33.
- [62] Shah AA, Casciola-Rosen L, Rosen A. Review: cancer-induced autoimmunity in the rheumatic diseases. *Arthr Rheumatol (Hoboken, NJ)* 2015;67:317–26.
- [63] Shah AA, Rosen A, Hummers L, et al. Close temporal relationship between onset of cancer and scleroderma in patients with RNA polymerase I/III antibodies. *Arthr Rheum* 2010;62:2787–95.
- [64] Joseph CG, Darrach E, Shah AA, et al. Association of the autoimmune disease scleroderma with an immunologic response to cancer. *Science (New York, NY)* 2014;343:152–7.
- [65] Hudson M, Pope J, Mahler M, et al. Clinical significance of antibodies to Ro52/TRIM21 in systemic sclerosis. *Arthr Res Ther* 2012;14:R50.

BRIEF REPORT

IRF4 Newly Identified as a Common Susceptibility Locus for Systemic Sclerosis and Rheumatoid Arthritis in a Cross-Disease Meta-Analysis of Genome-Wide Association Studies

Elena López-Isac,¹ Jose-Ezequiel Martín,¹ Shervin Assassi,² Carmen P. Simeón,³ Patricia Carreira,⁴ Norberto Ortego-Centeno,⁵ Mayka Freire,⁶ Emma Beltrán,⁷ Javier Narváez,⁸ Juan J. Alegre-Sancho,⁹ the Spanish Scleroderma Group, Benjamín Fernández-Gutiérrez,¹⁰ Alejandro Balsa,¹¹ Ana M. Ortiz,¹² Miguel A. González-Gay,¹³ Lorenzo Beretta,¹⁴ Alessandro Santaniello,¹⁴ Chiara Bellocchi,¹⁴ Claudio Lunardi,¹⁵ Gianluca Moroncini,¹⁶ Armando Gabrielli,¹⁶ Torsten Witte,¹⁷ Nicolas Hunzelmann,¹⁸ Jörg H. W. Distler,¹⁹ Gabriella Riekemasten,²⁰ Annette H. van der Helm-van Mil,²¹ Jeska de Vries-Bouwstra,²¹ Cesar Magro-Checa,²¹ Alexandre E. Voskuyl,²² Madelon C. Vonk,²³ Øyvind Molberg,²⁴ Tony Merriman,²⁵ Roger Hesselstrand,²⁶ Annika Nordin,²⁷ Leonid Padyukov,²⁷ Ariane Herrick,²⁸ Steve Eyre,²⁸ Bobby P. C. Koeleman,²⁹ Christopher P. Denton,³⁰ Carmen Fonseca,³⁰ Timothy R. D. J. Radstake,³¹ Jane Worthington,²⁸ Maureen D. Mayes,² and Javier Martín¹

Objective. Systemic sclerosis (SSc) and rheumatoid arthritis (RA) are autoimmune diseases that have similar clinical and immunologic characteristics. To date, several shared SSc–RA genetic loci have been identified independently. The aim of the current study was to systematically search for new common SSc–RA

loci through an inter-disease meta-genome-wide association (meta-GWAS) strategy.

Methods. The study was designed as a meta-analysis combining GWAS data sets of patients with SSc and patients with RA, using a strategy that allowed identification of loci with both same-direction and opposite-direction allelic effects. The top single-nucleotide

Supported by a grant from the Ministerio de Educación, Cultura y Deporte through the program FPU (to Dr. López-Isac), grant 115565 from the EU/EFPIA Innovative Medicines Initiative Joint Undertaking PRECISEADS (ref. no. 115565) and BIO-1395 from the Junta de Andalucía, grant PI-0590-2010 from the Consejería de Salud y Bienestar Social, Junta de Andalucía, Spain (to Dr. Ortego-Centeno), a VIDÍ laureate from the Dutch Association of Research and Dutch Arthritis Foundation (to Dr. Radstake), and grant SAF2012-34435 from the Spanish Ministry of Economy and Competitiveness (to Dr. J. Martín). Dr. Assassi's work was supported by grants KL2-RR-024149-04 and K23-AR-061436 from the NIH, grant 3-UL1-RR-024148 from the NIH National Center for Research Resources, and grant U01-1U01AI09090 from the NIH National Institute of Allergy and Infectious Diseases. Dr. Mayes' work was supported by grant P50-AR-054144 from the NIH National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) Centers of Research Translation, grant N01-AR-0-2251 from the NIAMS SSc Family Registry and DNA Repository, grant PR-1206877 from the Department of Defense, and grant R01-AR-055258 from the NIAMS.

¹Elena López-Isac, MSc, Jose-Ezequiel Martín, PhD, Javier Martín, MD, PhD: Institute of Parasitology and Biomedicine López-Neyra CSIC, Granada, Spain; ²Shervin Assassi, MD, Maureen D. Mayes, MD, MPH: University of Texas Health Science Center, Houston; ³Carmen P. Simeón, MD, PhD: Valle de Hebrón Hospital, Barcelona, Spain; ⁴Patricia Carreira, MD: 12 de Octubre University Hospital, Madrid, Spain; ⁵Norberto Ortego-Centeno, MD: Clinic University Hospital, Granada, Spain; ⁶Mayka Freire, MD, PhD: Complejo Hospitalario Universitario de Vigo, Vigo, Spain; ⁷Emma Beltrán, MD: Hospital General Universitario de Valencia, Valencia, Spain; ⁸Javier Narváez, MD, PhD: Hospital Universitari de Bellvitge,

Barcelona, Spain; ⁹Juan J. Alegre-Sancho, MD, PhD: Hospital Universitari Doctor Peset, Valencia, Spain; ¹⁰Benjamín Fernández-Gutiérrez, MD, PhD: Hospital Clínico San Carlos, Madrid, Spain; ¹¹Alejandro Balsa, MD, PhD: Hospital Universitario La Paz, Instituto de Investigación Sanitaria La Paz, Madrid, Spain; ¹²Ana M. Ortiz, MD, PhD: Hospital Universitario La Princesa, Instituto de Investigación Sanitaria La Princesa, Madrid, Spain; ¹³Miguel A. González-Gay, MD, PhD: University of Cantabria, Santander, Spain; ¹⁴Lorenzo Beretta, MD, Alessandro Santaniello, MD, Chiara Bellocchi, MD: Referral Center for Systemic Autoimmune Diseases, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico di Milano, Milan, Italy; ¹⁵Claudio Lunardi, MD: Università degli Studi di Verona, Verona, Italy; ¹⁶Gianluca Moroncini, MD, PhD, Armando Gabrielli, MD, PhD: Università Politecnica delle Marche and Ospedali Riuniti, Ancona, Italy; ¹⁷Torsten Witte, MD, PhD: Hannover Medical School, Hannover, Germany; ¹⁸Nicolas Hunzelmann, MD: University of Cologne, Cologne, Germany; ¹⁹Jörg H. W. Distler, MD: University of Erlangen-Nuremberg, Erlangen, Germany; ²⁰Gabriella Riekemasten, MD, PhD: University of Lübeck, Lübeck, Germany; ²¹Annette H. van der Helm-van Mil, MD, PhD, Jeska de Vries-Bouwstra, MD, PhD, Cesar Magro-Checa, MD: Leiden University Medical Center, Leiden, The Netherlands; ²²Alexandre E. Voskuyl, MD: VU University Medical Center, Amsterdam, The Netherlands; ²³Madelon C. Vonk, MD, PhD: Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands; ²⁴Øyvind Molberg, MD, PhD: Oslo University Hospital Rikshospitalet and Institute of Clinical Medicine, University of Oslo, Oslo, Norway; ²⁵Tony Merriman, PhD: University of Otago, Otago, New Zealand; ²⁶Roger Hesselstrand, MD, PhD: Lund University,

polymorphisms were followed up in independent SSc and RA case-control cohorts. This allowed an increase in the sample size to a total of 8,830 patients with SSc, 16,870 patients with RA, and 43,393 healthy controls.

Results. This cross-disease meta-analysis of the GWAS data sets identified several loci with nominal association signals ($P < 5 \times 10^{-6}$) that also showed evidence of association in the disease-specific GWAS scans. These loci included several genomic regions not previously reported as shared loci, as well as several risk factors that were previously found to be associated with both diseases. Follow-up analyses of the putatively new SSc-RA loci identified *IRF4* as a shared risk factor for these 2 diseases ($P_{\text{combined}} = 3.29 \times 10^{-12}$). Analysis of the biologic relevance of the known SSc-RA shared loci identified the type I interferon and interleukin-12 signaling pathways as the main common etiologic factors.

Conclusion. This study identified a novel shared locus, *IRF4*, for the risk of SSc and RA, and highlighted the usefulness of a cross-disease GWAS meta-analysis strategy in the identification of common risk loci.

Genome-wide association studies (GWAS) and immune-focused fine-mapping studies have revolutionized our understanding of the genetic component of complex autoimmune diseases by facilitating the identification of thousands of susceptibility loci associated with autoimmunity (1). The vast majority of these loci are shared risk factors for at least 2 autoimmune diseases, pointing to a common genetic background underlying these autoimmune processes. This genetic overlap was suspected some time ago, given the high rate of co-occurrence of autoimmune diseases and the well-established familial aggregation reported for these immune disorders (1).

Systemic sclerosis (SSc) and rheumatoid arthritis (RA) are complex autoimmune diseases that have

similar clinical and immunologic features. Both diseases are rheumatic connective tissue disorders, characterized by an exacerbated inflammatory response, deregulation of innate and adaptive immunity, including autoantibody production, and systemic complications. Because of the establishment of large consortiums and international collaborations, the number of confirmed RA susceptibility factors has increased up to a total of 101 loci associated with the disease at the genome-wide significance level (2). With regard to SSc, GWAS, Immunochip, and candidate gene studies have clearly identified various genetic regions involved in susceptibility to SSc (3). However, the knowledge of the genetic predisposition to this disease is relatively limited, in part due to its low prevalence, which impairs the recruitment of large cohorts required to reach a high statistical power and to effectively detect association signals. Interestingly, a considerable proportion of the SSc susceptibility factors also represent RA risk loci (2,3). In addition, although not very common, co-familiarity and co-occurrence between these 2 rheumatic conditions have been observed (4). These observations provide evidence of a genetic overlap of both diseases. Thus, it is expected that additional shared risk factors remain to be discovered.

One approach that has been developed for the identification of common loci in a cost-effective manner is to perform a combined-phenotype GWAS, that is, to combine genome-wide genotype data for 2 autoimmune diseases. This strategy has been successfully applied to the study of not only closely related phenotypes but also nonrelated phenotypes, and thus far the results have been encouraging (5).

Taking into account all of these considerations, the purpose of the present study was to systematically identify new common risk loci for SSc and RA by applying the combined-phenotype GWAS strategy, followed by replication testing in independent case-control data sets.

PATIENTS AND METHODS

Study population. The first stage of the present study, the discovery phase, included 6,537 patients with either SSc or RA and 8,741 healthy controls. The SSc GWAS panel comprised 4 case-control sets from Spain, Germany, The Netherlands, and the US (2,716 cases and 5,666 controls), whose data had been obtained in previous studies (5–7). The RA case-control GWAS panel included 2 previously published RA GWAS cohorts (the Wellcome Trust Case Control Consortium [WTCCC] and the Epidemiological Investigation of Rheumatoid Arthritis study cohort) from the UK and Sweden (3,821 cases and 3,075 controls) (8).

Subjects included in the second stage of the study, the replication phase, were drawn from independent SSc and RA case-control sets of individuals European ancestry. The SSc

Lund, Sweden; ²⁷Annika Nordin, MD, PhD, Leonid Padyukov, PhD: Karolinska University Hospital, Karolinska Institutet, Stockholm, Sweden; ²⁸Ariane Herrick, MD, PhD, Steve Eyre, MD, PhD, Jane Worthington, MD, PhD: University of Manchester, Manchester Academic Health Science Centre, Manchester, UK; ²⁹Bobby P. C. Koeleman, PhD: University Medical Center Utrecht, Utrecht, The Netherlands; ³⁰Christopher P. Denton, PhD, FRCP, Carmen Fonseca, MD, PhD: Centre for Rheumatology, Royal Free and University College Medical School, London, UK; ³¹Timothy R. D. J. Radstake, MD, PhD: University Medical Center Utrecht, Utrecht, The Netherlands. Members of the Spanish Scleroderma Group are shown in Appendix A.

Address correspondence to Elena López-Isac, MSc, or to Javier Martín, MD, PhD, IPBLN-CSIC, Parque Tecnológico Ciencias de la Salud, Avenida del Conocimiento s/n 18016-Armilla, Granada, Spain. E-mail: elenalopezisac@ipb.csic.es or martin@ipb.csic.es.

Submitted for publication March 30, 2016; accepted in revised form April 19, 2016.

replication cohort included 6,114 cases and 8,744 healthy controls from 8 different countries (Spain, Germany, Italy, the UK, The Netherlands, Sweden, Norway, and the US). The healthy controls from the UK and US partially overlapped with the control sets of previously published cohorts (the WTCCC and the second North American Rheumatoid Arthritis Consortium [NARAC2]) (8). The RA replication cohort included 9 case-control collections from North America (US and Canada), Spain, The Netherlands, the UK, Sweden, France, and New Zealand, and comprised a total of 13,049 RA cases and 25,908 healthy controls. Of these, 9,711 cases and 24,253 healthy controls were obtained from several previously published studies, including the Brigham Rheumatoid Arthritis Sequential Study, NARAC1, CANADA, studies from the Rheumatoid Arthritis Consortium International (RACI-US, RACI-i2b2, RACI-UK, RACI-SE-U, and RACI-NL), Consortium of Rheumatology Researchers of North America, Vanderbilt, Dutch studies (Amsterdam Medical Center, Treatment Strategies for RA [BeSt Study], Leiden University Medical Center, and Dutch Rheumatoid Arthritis Monitoring Registry), Research in Active Rheumatoid Arthritis, and the Anti-TNF Response to Therapy collection (ACR-REF: BRAGGSS, BRAGGSS2, ERA, KI, and TEAR) (2). All of the patients with SSc and patients with RA fulfilled previously described classification criteria for each disease (2,5). All individuals enrolled in the present study provided written informed consent, and approval from the local ethics committees was obtained from all of the centers in accordance with the tenets of the Declaration of Helsinki.

Study design. We performed a 2-stage study to systematically identify SSc-RA shared risk factors, with the first stage being the discovery phase for GWAS meta-analysis of each disease separately and a combined-phenotype GWAS meta-analysis, and the second stage being the replication phase (Figure 1).

Discovery phase. We performed GWAS analysis for each disease separately and in a combined-phenotype GWAS analysis. Two different tests were considered for the combined analysis (5). In the first test, in order to detect common signals for SSc and RA with same-direction allelic effects, the meta-analysis considering both diseases was performed as usual. Those SNPs that showed an association at $P < 5 \times 10^{-6}$ in the combined-phenotype analysis (referred to as P_{combined}) and also showed nominal significance in the association study for each disease ($P < 0.05$) were selected for follow-up in the replication phase.

In the second test, in order to identify common signals with opposite-direction allelic effects, we flipped the direction of association (1/odds ratio [OR]) in the RA data set for the combined-phenotype meta-analysis. To select SNPs for replication, the same selection criteria as stated above were followed.

For both sorts of meta-analyses, we only considered for follow-up those SNPs that had not been previously reported as genetic risk factors for SSc and RA, or those that had been reported for one disease but not reported for the other.

Replication phase. The SNPs selected were followed-up in independent replication cohorts. Subsequently, we performed a meta-analysis of the initial GWAS screening and replication stages. The SNP signals that 1) reached the genome-wide significance level for association ($P_{\text{combined}} < 5 \times 10^{-8}$) in the combined-phenotype meta-analysis (GWAS + Replication phases), and that 2) showed, for each disease separately,

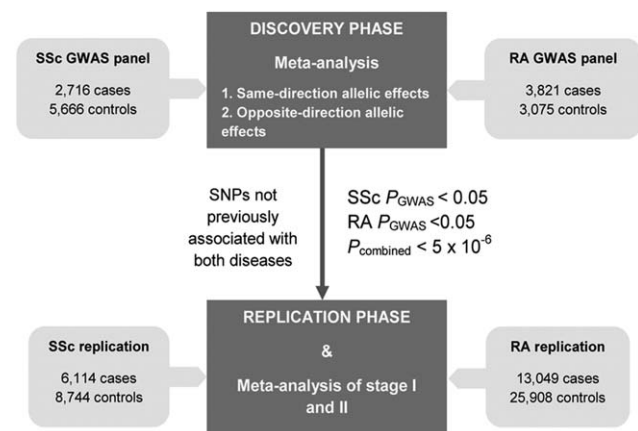


Figure 1. Overall work flow of the present study. SSc = systemic sclerosis; GWAS = genome-wide association study; RA = rheumatoid arthritis; SNPs = single-nucleotide polymorphisms.

nominal significant associations ($P < 0.05$) in the replication phase as well as significant associations ($P < 5 \times 10^{-3}$) in the GWAS + Replication meta-analysis were considered shared risk factors for the 2 analyzed diseases.

Quality control and genotype imputation of GWAS data. We applied stringent quality control criteria in all of the GWAS data sets. Cutoff values for the sample call rate and the SNP call rate were set as 95%. Markers with allele distributions deviating from Hardy-Weinberg equilibrium (HWE) ($P < 0.001$) in controls from any of the populations analyzed separately were excluded. Markers with minor allele frequencies lower than 1% were filtered out. After quality control, we performed whole-genome genotype imputation with IMPUTE2 software (9) using as reference panels the CEU (Utah residents with northern and western European ancestry from the CEPH collection) and TSI (Toscani in Italy) populations of the HapMap Phase 3 project (available at <http://www.hapmap.org>). Imputed SNP quality was assessed by establishing a probability threshold for merging genotypes at 0.9. Subsequently, stringent quality control was applied to the imputed data using the same criteria as stated above. Thereafter, genome-wide genotyping data were available for a total of 219,756 SNPs.

The first 5 principal components were estimated, and individuals deviating more than 6 SDs from the cluster centroids were considered outliers. In addition, duplicate pairs or highly related individuals among data sets were also removed on the basis of pairwise comparisons, using the Genome function in Plink version 1.7 (see <http://pngu.mgh.harvard.edu/purcell/plink/>) (Pi-HAT threshold of 0.5).

Follow-up genotyping. The genotyping of the replication cohorts was performed with either TaqMan SNP genotyping technology in a LightCycler 480 Real-Time polymerase chain reaction system (Roche Applied Science) or the GWAS and Immunochip platforms.

For the SSc study, all cases were genotyped using the TaqMan genotyping system, with TaqMan 5' allele discrimination predesigned assays with Applied Biosystems. The genotyping call rate was >95% for the 3 SNPs. The control samples were also genotyped using this technology, with the exception of the UK and US cohorts. For these 2 control

cohorts, genotyping data were obtained from previously published genome-wide genotyping data sets (from the WTCCC and NARAC2) (8).

RA cases from Spain and New Zealand and the Spanish controls were genotyped by TaqMan technology. Genotype data for the New Zealand healthy controls partially overlapped with those from a previous GWAS report (10). For the remaining RA case-control sets, genotype frequencies and association data were obtained from a previously published study (2). The genotype methods used in these studies were described in detail in the study by Okada et al (2). For those cohorts in which genotyping was performed using the Illumina ImmunoChip platform, only data for the rs9328192 SNP of the interferon regulatory factor 4 (IRF-4) gene (*IRF4*) were available.

Statistical analysis. All data were analyzed using Plink software. To test for association, we performed logistic regression analysis in each of the SSc and RA GWAS cohorts separately. The first 5 principal components were included as covariates to control for any potential population stratification effects. The replication cohorts were also analyzed by logistic regression analysis. The meta-analyses were performed with the inverse-variance method based on population-specific logistic regression results. Heterogeneity of the ORs across studies was assessed using Cochran's Q test. HWE was tested in all of the validation cohorts genotyped by TaqMan technology (in HWE analyses, $P < 0.01$ was considered to show significant deviation from equilibrium). None of the included control cohorts showed significant deviation from HWE, with the exception of *HNFL1A* rs10774577. The cohorts in which HWE was not observed were excluded from the analysis of this specific SNP. The statistical power of the combined-phenotype analysis and the analysis for each disease separately is shown in Supplementary Table 1 (available on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39730/abstract>).

RESULTS

Discovery analysis. In the first phase of this study, we conducted a cross-disease meta-analysis in order to systematically identify new putatively shared loci between SSc and RA. The overall workflow of the study is illustrated in Figure 1.

The meta-analysis combining both data sets identified various SNPs from 7 distinct genomic regions that showed a significant association at the level of $P < 5 \times 10^{-6}$, as well as a nominal signal of association ($P < 0.05$) in the disease-specific analyses. The strongest association was found in the well-accepted SSc- and RA-associated locus *IRF5* ($P_{\text{combined}} = 8.44 \times 10^{-17}$; for SSc, $P_{\text{GWAS}} = 1.14 \times 10^{-16}$; for RA, $P_{\text{GWAS}} = 7.86 \times 10^{-4}$). Three additional known SSc-RA loci, namely *PTPN22*, *ATG5*, and *BLK*, were also identified at the initial discovery stage (Figure 2) (see also Supplementary Table 2 and Supplementary Figure 1, available on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39730/abstract>). The remaining SNPs

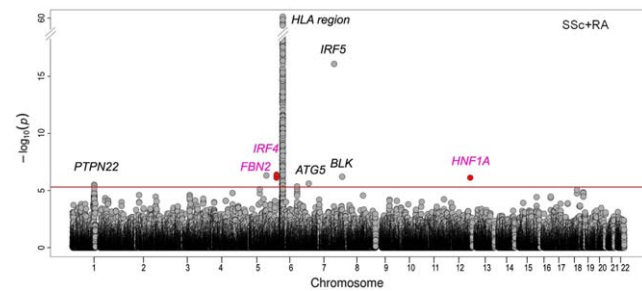


Figure 2. Manhattan plot showing the results of the cross-disease meta-genome-wide association study. The $-\log_{10}$ of the combined-phenotype meta-analysis P values are plotted against their physical chromosomal position. The plot displays the $-\log_{10} P$ values from the same-direction meta-analysis of associations with systemic sclerosis (SSc) and rheumatoid arthritis (RA). The signals from the opposite-direction meta-analysis that reached the selection criteria are also plotted (red points). The red line represents the threshold of significance at $P < 5 \times 10^{-6}$. Those loci with single-nucleotide polymorphisms that reached the selection criteria for the replication phase are plotted (loci selected for follow-up are highlighted in pink).

were located in 3 different loci, including *FBN2* and *HNFL1A*, neither of which has been previously reported as a genetic risk factor for SSc and RA, and *IRF4*, which has been found to be associated with RA in previous studies (Table 1 and Figure 2) (see also Supplementary Figure 2 on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39730/abstract>). Interestingly, the regional association plots of the *FBN2*, *IRF4*, and *HNFL1A* loci showed that the top SNPs in the combined-phenotype analysis were also the top SNPs in the analyses for SSc and RA separately, or at least were in high linkage disequilibrium with the top signal observed for each disease (see Supplementary Figure 2).

These new putatively shared SNPs were selected for follow-up in additional SSc and RA replication cohorts. For *IRF4*, 3 SNPs met our criteria for being selected for validation in the replication phase. In this case, we selected the SNP with the lowest P value for association (see Supplementary Table 2).

Replication phase and meta-analysis. According to the established thresholds discussed above in Patients and Methods, we identified 1 new association signal shared between SSc and RA at SNP rs9328192 of *IRF4* ($P_{\text{combined}} = 3.29 \times 10^{-12}$). Furthermore, this *IRF4* SNP almost reached genome-wide significance in the meta-analysis for each disease separately (for SSc, $P_{\text{GWAS} + \text{Replication}} = 2.78 \times 10^{-7}$, OR 0.90; for RA, $P_{\text{GWAS} + \text{Replication}} = 1.44 \times 10^{-6}$, OR 1.08) (Table 1).

Regarding the *HNFL1A* and *FBN2* genetic variants, despite the initial suggestive association signals found in the first stage, these loci did not show genome-wide significance in our combined-phenotype meta-analysis.

Table 1. Results of the cross-disease meta-GWAS of disease associations with 3 selected SNPs*

Locus (Chr.)	SNP	Ref. allele	<i>P</i> for SSc-RA GWAS	SSc						RA						SSc-RA GWAS + Replication	Current status
				GWAS		Replication		GWAS + Replication (meta-GWAS)		GWAS		Replication		GWAS + Replication (meta-GWAS)			
				<i>P</i>	OR†	<i>P</i>	OR†	<i>P</i>	OR†	<i>P</i>	OR†	<i>P</i>	OR†	<i>P</i>	OR†		
<i>FBN2</i> (5)	rs6897611	T	4.79×10^{-7}	2.85×10^{-3}	1.16	0.641	0.98	0.165	1.04	3.15×10^{-5}	1.24	0.684	0.99	0.650‡	1.02‡	0.018	–
<i>IRF4</i> (6)	rs9328192	G	4.06×10^{-7}	8.86×10^{-6}	0.86	1.89×10^{-3}	0.93	2.78×10^{-7}	0.90	7.26×10^{-3}	1.10	5.22×10^{-5}	1.07	1.44×10^{-6}	1.08	3.29×10^{-12}	SSc-RA; RA
<i>HNF1A</i> (12)	rs10774577§	T	7.53×10^{-7}	8.62×10^{-4}	0.89	0.036	0.94	1.64×10^{-4}	0.91	2.50×10^{-4}	1.14	0.290	1.03	0.208‡	1.05‡	1.59×10^{-6}	–

* GWAS = genome-wide association study; SNPs = single-nucleotide polymorphisms; Chr. = chromosome.

† Represents the odds ratio (OR) for the reference (Ref.) allele.

‡ *P* values and ORs were derived from meta-analysis under random effects due to heterogeneity of the ORs among cohorts.

§ The rheumatoid arthritis (RA) and systemic sclerosis (SSc) replication cohorts from Spain, and the SSc replication cohorts from Italy and The Netherlands were excluded from the analysis of rs10774577 due to issues with Hardy-Weinberg equilibrium.

Nevertheless, *HNFL1A* rs10774577 showed suggestive evidence of association in the meta-analysis performed in the SSc data set ($P_{\text{Replication}} = 0.036$, OR 0.94; $P_{\text{GWAS} + \text{Replication}} = 1.64 \times 10^{-4}$, OR 0.91), and showed an association at the level of $P = 1.59 \times 10^{-6}$ in the combined-phenotype meta-analysis. Considering that this SNP was not included in those cohorts that were genotyped with Immunochip, the present study had a lower statistical power for the analysis of this genomic region. Therefore, the possibility of a slight or modest genetic effect of *HNFL1A* rs10774577 on both diseases cannot be ruled out, and further studies will be required to establish whether this locus is a shared SSc-RA risk factor.

DISCUSSION

In the present study, we identified a novel non-HLA susceptibility locus that is shared between SSc and RA, namely *IRF4*, using a combined-phenotype GWAS strategy in large case-control cohorts of patients with SSc and those with RA. This locus, *IRF4*, was already reported to be involved in RA susceptibility, but had not been previously associated with SSc (2).

The cross-disease meta-analysis performed with the SSc and RA GWAS data sets identified various SNPs from 7 different loci that met our stringent selection criteria for the replication phase ($P_{\text{combined}} < 5 \times 10^{-6}$; for SSc and for RA, each $P_{\text{GWAS}} < 0.05$). Four of the 7 SNPs were already known risk factors for SSc and RA (*PTPN22*, *ATG5*, *IRF5*, and *BLK*), thus providing support for the effectiveness of this strategy in the identification of shared risk loci (2,3). It is worth mentioning that these loci were detected by the 2 different tests used in the first phase, which were performed in order to detect both same-direction and opposite-direction allelic effects. In fact, the shared *IRF4* SNP newly identified in this study showed opposite effects for SSc and for RA (protective effect and risk effect, respectively). This discrepancy might be attributable to the fact that the actual causal variants for the associations in each disease could be different, and that *IRF4* rs9328192 is tagging them. This discordant phenomenon is particularly common between autoimmune diseases (1). However, to completely understand these discordant effects, the interaction with other genetic variants contributing to disease susceptibility should be considered, in addition to analyzing the precise biologic impact of the associations.

The associated *IRF4* SNP (rs9328192) showed modest effect sizes for SSc and RA. However, we were able to capture this association in our meta-analysis because of the large cohort used in this study, together with the combined-phenotype approach, which allowed

us to increase the statistical power. This highlights the capability of the combined-phenotype approach in the identification of shared variants with low penetrance, whose associations might have been missed in disease-specific GWAS due to a lack of power (11).

IRF-4 belongs to the IRF family of transcription factors and plays a pivotal role in the development and function of several autoimmune-associated cells (12). Various genetic and functional studies have pointed to IRF-4 as a master regulator for autoimmunity (12,13). It has been demonstrated that IRF-4 is a crucial factor for the editing and L-chain rearrangements of the B cell receptor, and for pre-B cell expansion, both of which are processes directly related to the development of autoimmunity (14). In addition, IRF-4 is a critical controller of Th17 cell differentiation and the production of interleukin-17 (IL-17) and IL-21 (12), which are components of the immune system that play a key role in the pathogenesis of SSc and RA.

The results of the present study add another *IRF* to the list of *IRFs* associated with SSc (*IRF4*, *IRF5*, *IRF7*, and *IRF8*) and RA (*IRF4*, *IRF5*, and *IRF8*) (2,3), thus providing genetic support for the type I interferon (IFN) signature described in patients with SSc and those with RA (15). Moreover, our pathway enrichment analysis also identified the type I IFN signaling pathway as one of the most relevant common pathways between SSc and RA on the basis of their common genetic background (see Supplementary Methods, Supplementary Table 3, and Supplementary Figure 3, available on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39730/abstract>). Therefore, deregulation of this signaling pathway might be a biologic process that underlies the onset of these 2 autoimmune rheumatic conditions.

In summary, through a cross-disease meta-analysis of GWAS for SSc and RA, we were able to identify *IRF4* as a new shared susceptibility locus for these 2 autoimmune diseases. The results of the present study, taken together with the findings from previous studies, reinforce the idea of a common genetic background between SSc and RA. The identification of these pleiotropic autoimmunity loci may point to common pathogenic pathways, which ultimately may represent a clinical advantage in that it may provide support for drug repositioning on the basis of the true understanding of the pathogenic mechanisms of SSc and RA.

ACKNOWLEDGMENTS

We thank Sofia Vargas, Sonia García, and Gema Robledo for their excellent technical assistance, and all of the

patients and control donors for their essential collaboration. We thank National DNA Bank Carlos III (University of Salamanca, Spain) for supplying part of the control DNA samples. We also would like to thank the following organizations: the European League Against Rheumatism Scleroderma Trials and Research Group, the German Network of Systemic Sclerosis, the US Scleroderma Foundation, and the US Raynauds & Scleroderma Association.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Drs. López-Isac and J. Martín had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Fonseca, Radstake, Worthington, Mayes, J. Martín.

Acquisition of data. Assassi, Simeón, Carreira, Ortego-Centeno, Freire, Beltrán, Narváez, Alegre-Sancho, Fernández-Gutiérrez, Balsa, Ortiz, González-Gay, Beretta, Santaniello, Bellocchi, Lunardi, Moroncini, Gabrielli, Witte, Hunzelmann, Distler, Riekemasten, van der Helm-van Mil, de Vries-Bouwstra, Magro-Checa, Voskuyl, Vonk, Molberg, Merriman, Hesselstrand, Nordin, Padyukov, Herrick, Eyre, Koeleman, Denton, Fonseca, Radstake, Worthington, Mayes, J. Martín.

Analysis and interpretation of data. López-Isac, J.-E. Martín.

REFERENCES

- Parkes M, Cortes A, van Heel DA, Brown MA. Genetic insights into common pathways and complex relationships among immune-mediated diseases. *Nat Rev Genet* 2013;14:661–73.
- Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* 2014;506:376–81.
- Bossini-Castillo L, Lopez-Isac E, Martin J. Immunogenetics of systemic sclerosis: defining heritability, functional variants and shared-autoimmunity pathways. *J Autoimmun* 2015;64:53–65.
- Elhai M, Avouac J, Kahan A, Allanore Y. Systemic sclerosis at the crossroad of polyautoimmunity. *Autoimmun Rev* 2013;12:1052–7.
- Martin JE, Assassi S, Diaz-Gallo LM, Broen JC, Simeon CP, Castellvi I, et al. A systemic sclerosis and systemic lupus erythematosus pan-meta-GWAS reveals new shared susceptibility loci. *Hum Mol Genet* 2013;22:4021–9.
- Radstake TR, Gorlova O, Rueda B, Martin JE, Alizadeh BZ, Palomino-Morales R, et al. Genome-wide association study of systemic sclerosis identifies CD247 as a new susceptibility locus. *Nat Genet* 2010;42:426–9.
- Martin JE, Broen JC, Carmona FD, Teruel M, Simeon CP, Vonk MC, et al. Identification of CSK as a systemic sclerosis genetic risk factor through genome wide association study follow-up. *Hum Mol Genet* 2012;21:2825–35.
- Stahl EA, Raychaudhuri S, Remmers EF, Xie G, Eyre S, Thomson BP, et al. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat Genet* 2010;42:508–14.
- Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* 2009;5:e1000529.
- Consortium TAAZMSG. Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. *Nat Genet* 2009;41:824–8.
- Festen EA, Goyette P, Green T, Boucher G, Beauchamp C, Trynka G, et al. A meta-analysis of genome-wide association scans identifies IL18RAP, PTPN2, TAGAP, and PUS10 as shared risk loci for Crohn's disease and celiac disease. *PLoS Genet* 2011;7:e1001283.
- Xu WD, Pan HF, Ye DQ, Xu Y. Targeting IRF4 in autoimmune diseases. *Autoimmun Rev* 2012;11:918–24.
- Biswas PS, Gupta S, Stirzaker RA, Kumar V, Jessberger R, Lu TT, et al. Dual regulation of IRF4 function in T and B cells is required for the coordination of T-B cell interactions and the prevention of autoimmunity. *J Exp Med* 2012;209:581–96.
- Zouali M. Receptor editing and receptor revision in rheumatic autoimmune diseases. *Trends Immunol* 2008;29:103–9.
- Ronnblom L, Eloranta ML. The interferon signature in autoimmune diseases. *Curr Opin Rheumatol* 2013;25:248–53.

APPENDIX A: MEMBERS OF THE SPANISH SCLERODERMA GROUP

Members of the Spanish Scleroderma Group include the following: Raquel Ríos and Jose Luis Callejas (Hospital Clínico Universitario San Cecilio, Granada), José Antonio Vargas Hitos (Hospital Virgen de las Nieves, Granada), Rosa García Portales (Hospital Virgen de la Victoria, Málaga), María Teresa Camps (Hospital Carlos Haya, Málaga), Antonio Fernández-Nebro (Hospital Carlos Haya, Málaga), María F. González-Escribano (Hospital Virgen del Rocío, Seville), Francisco José García-Hernández and M^a Jesús Castillo (Hospital Virgen del Rocío, Seville), M^a Angeles Aguirre and Inmaculada Gómez-Gracia (Hospital Reina Sofía/IMIBIC, Córdoba), Luis Rodríguez-Rodríguez (Hospital Clínico San Carlos, Madrid), Paloma García de la Peña (Madrid Norte Sanchinarro Hospital, Madrid), Esther Vicente (Hospital La Princesa, Madrid), José Luis Andreu and Mónica Fernández de Castro (Hospital Puerta de Hierro Majadahonda, Madrid), Francisco Javier López-Longo and Lina Martínez (Hospital Universitario Gregorio Marañón, Madrid), Vicente Fonollosa and Alfredo Guillén (Hospital Valle de Hebrón, Barcelona), Iván Castellví (Santa Creu i Sant Pau University Hospital, Barcelona), Gerard Espinosa (Hospital Clinic, Barcelona), Carlos Tolosa (Hospital Parc Tauli, Sabadell), Anna Pros (Hospital Del Mar, Barcelona), Mónica Rodríguez Carballeira (Hospital Universitari Mútua Terrasa, Barcelona), Francisco Javier Narváez (Hospital Universitari de Bellvitge, Barcelona), Manel Rubio Rivas (Hospital Universitari de Bellvitge, Barcelona), Vera Ortiz-Santamaría (Hospital General de Granollers, Granollers), Ana Belén Madroñero (Hospital General San Jorge, Huesca), Bernardino Díaz and Luis Trapiella (Hospital Central de Asturias, Oviedo), Adrián Sousa (Hospital Xeral-Complexo Hospitalario Universitario de Vigo, Vigo), María Victoria Egurbide (Hospital Universitario Cruces, Barakaldo), Patricia Fanlo Mateo (Hospital Virgen del Camino, Pamplona), Luis Sáez-Comet (Hospital Universitario Miguel Servet, Zaragoza), Federico Díaz and Vanesa Hernández (Hospital Universitario de Canarias, Tenerife), Emma Beltrán (Hospital General Universitario de Valencia, Valencia), José Andrés Román-Ivorra and Elena Grau (Hospital Universitari i Politècnic La Fe, Valencia), Juan José Alegre-Sancho (Hospital Universitari Doctor Peset, Valencia), Francisco J. Blanco García and Natividad Oreiro (INIBIC-Hospital Universitario A Coruña, La Coruña).

RESEARCH ARTICLE

Investigating the Causal Relationship of C-Reactive Protein with 32 Complex Somatic and Psychiatric Outcomes: A Large-Scale Cross-Consortium Mendelian Randomization Study

Bram. P. Prins^{1,2*}, Ali Abbasi^{1,3,4}, Anson Wong^{5,6}, Ahmad Vaez^{1,7}, Ilja Nolte¹, Nora Franceschini⁸, Philip E. Stuart⁹, Javier Gutierrez Achury¹⁰, Vanisha Mistry^{11,12}, Jonathan P. Bradfield¹³, Ana M. Valdes¹⁴, Jose Bras¹⁵, Aleksey Shatunov¹⁶, PAGE Consortium⁹, International Stroke Genetics Consortium¹⁷, Systemic Sclerosis consortium^{18,19}, Treat OA consortium²⁰, DIAGRAM Consortium²¹, CARDIoGRAMplusC4D Consortium²², ALS consortium¹⁶, International Parkinson's Disease Genomics Consortium²³, Autism Spectrum Disorder Working Group of the Psychiatric Genomics Consortium²⁴, CKDGen consortium²⁵, GERAD1 Consortium²⁶, International Consortium for Blood Pressure²⁷, Schizophrenia Working Group of the Psychiatric Genomics Consortium²⁸, Inflammation Working Group of the CHARGE Consortium¹, Chen Lu²⁹, Buhm Han³⁰, Soumya Raychaudhuri^{30,31,32,33,34}, Steve Bevan¹⁷, Maureen D. Mayes¹⁸, Lam C. Tsoi^{9,35}, Evangelos Evangelou^{20,36}, Rajan P. Nair⁹, Struan F. A. Grant^{13,37}, Constantin Polychronakos³⁸, Timothy R. D. Radstake³⁹, David A. van Heel¹², Melanie L. Dunstan²⁶, Nicholas W. Wood²³, Ammar Al-Chalabi^{16,40}, Abbas Dehghan⁴¹, Hakon Hakonarson^{13,37}, Hugh S. Markus¹⁷, James T. Elder⁹, Jo Knight^{5,42,43,44}, Dan E. Arking²⁴, Timothy D. Spector⁴⁵, Bobby P. C. Koeleman⁴⁶, Cornelia M. van Duijn⁴¹, Javier Martin¹⁹, Andrew P. Morris^{21,47}, Rinse K. Weersma⁴⁸, Cisca Wijmenga¹⁰, Patricia B. Munroe^{27,49}, John R. B. Perry³, Jennie G. Pouget^{5,42,43}, Yalda Jamshidi⁵⁰, Harold Snieder¹, Behrooz Z. Alizadeh^{1,48*}



OPEN ACCESS

Citation: Prins B.P, Abbasi A, Wong A, Vaez A, Nolte I, Franceschini N, et al. (2016) Investigating the Causal Relationship of C-Reactive Protein with 32 Complex Somatic and Psychiatric Outcomes: A Large-Scale Cross-Consortium Mendelian Randomization Study. *PLoS Med* 13(6): e1001976. doi:10.1371/journal.pmed.1001976

Academic Editor: Phillipa J. Hay, Western Sydney University, AUSTRALIA

Received: June 21, 2015

Accepted: February 3, 2016

Published: June 21, 2016

Copyright: © 2016 Prins et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All our data which have been used for this study may be found in S1 Data file.

Funding: Please refer to file S1 Financial Disclosure for full information with regard to funding and financial support. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

1 Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands, **2** Department of Human Genetics, Wellcome Trust Sanger Institute, Hinxton, United Kingdom, **3** MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, United Kingdom, **4** Department of Internal Medicine, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands, **5** Institute of Medical Sciences, University of Toronto, Toronto, Ontario, Canada, **6** Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada, **7** School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran, **8** Department of Epidemiology, University of North Carolina, Chapel Hill, North Carolina, United States of America, **9** Department of Dermatology, Veterans Affairs Hospital, University of Michigan, Ann Arbor, Michigan, United States of America, **10** Department of Genetics, University of Groningen, University Medical Centre Groningen, Groningen, the Netherlands, **11** Metabolic Research Laboratories, Wellcome Trust–MRC Institute of Metabolic Science, Addenbrooke's Hospital, University of Cambridge, Cambridge, United Kingdom, **12** Blizzard Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, United Kingdom, **13** Center for Applied Genomics, Division of Human Genetics, Children's Hospital of Philadelphia, Philadelphia, United States of America, **14** Department of Academic Rheumatology, University of Nottingham, Nottingham, United Kingdom, **15** Department of Molecular Neuroscience, Institute of Neurology, London, United Kingdom, **16** Department of Basic and Clinical Neuroscience, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom, **17** Neurology Unit, Department of Clinical Neurosciences, University of Cambridge, Cambridge, United Kingdom, **18** Division of Rheumatology and Clinical Immunogenetics, University of Texas Health Science Center at Houston, Houston, Texas, United States of America, **19** Instituto de Parasitología y Biomedicina Lopez-Neyra, Consejo Superior de Investigaciones Científicas, Granada, Spain, **20** Department of Hygiene and Epidemiology, University of Ioannina Medical School, Ioannina, Greece, **21** Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom, **22** Institut für Integrative und Experimentelle Genomik, Universität zu Lübeck, Lübeck, Germany, **23** Department of Molecular

Abbreviations: CAD, coronary artery disease; CCGC, CRP Coronary Heart Disease Genetics Collaboration; CRP, C-reactive protein; CRP_{PRS}, C-reactive protein polygenic risk score; DBP, diastolic blood pressure; eGFR_{Cr}, estimated glomerular filtration rate from serum creatinine; eQTL, expression quantitative trait locus; GRS, genetic risk score; GWAS, genome-wide association study; IBD, inflammatory bowel disease; IV, instrumental variable; lnCRP, natural log of CRP level; MR, Mendelian randomization; OR, odds ratio; PC, principal component; PGC, Psychiatric Genomics Consortium; SBP, systolic blood pressure; SNP, single nucleotide polymorphism; s%, symmetric percentage.

Neuroscience, UCL Institute of Neurology, London, United Kingdom, **24** McKusick-Nathans Institute of Genetic Medicine and Department of Medicine, Division of Cardiology, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America, **25** NHLBI's Framingham Heart Study, Center for Population Studies and Harvard Medical School, Framingham, Massachusetts, United States of America, **26** Institute of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, United Kingdom, **27** NIHR Barts Cardiovascular Biomedical Research Unit, William Harvey Research Institute, Queen Mary University of London, London, United Kingdom, **28** MRC Centre for Neuropsychiatric Genetics and Genomics, Institute of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, United Kingdom, **29** Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, United States of America, **30** Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, Massachusetts, United States of America, **31** Division of Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, **32** Division of Rheumatology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, **33** Partners HealthCare Center for Personalized Genetic Medicine, Boston, Massachusetts, United States of America, **34** Faculty of Medical and Human Sciences, University of Manchester, Manchester, United Kingdom, **35** Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, United States of America, **36** Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, United Kingdom, **37** Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, United States of America, **38** Endocrine Genetics Research Institute, McGill University Health Center, Montreal, Quebec, Canada, **39** Department of Rheumatology & Clinical Immunology and Laboratory of Translational Immunology, University Medical Center Utrecht, Utrecht, the Netherlands, **40** Complex Disease Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, United States of America, **41** Department of Epidemiology, Erasmus University Rotterdam, University Medical Centre Rotterdam, Rotterdam, the Netherlands, **42** Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, Ontario, Canada, **43** Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada, **44** Biostatistics Division, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada, **45** Department of Twin Research and Genetic Epidemiology, King's College London, London, United Kingdom, **46** Complex Genetic Section, Department of Medical Genetics, University Medical Center Utrecht, Utrecht, the Netherlands, **47** Department of Biostatistics, University of Liverpool, Liverpool, United Kingdom, **48** Department of Gastroenterology and Hepatology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands, **49** Clinical Pharmacology, William Harvey Research Institute, Barts and the London School of Medicine, Queen Mary University of London, London, United Kingdom, **50** Cardiogenetics Lab, Cardiovascular and Cell Sciences Institute, St George's Hospital Medical School, London, United Kingdom

✉ These authors contributed equally to this work.
 * bram.prins@sanger.ac.uk (BPP); b.z.alizadeh@umcg.nl (BZA)

Abstract

Background

C-reactive protein (CRP) is associated with immune, cardiometabolic, and psychiatric traits and diseases. Yet it is inconclusive whether these associations are causal.

Methods and Findings

We performed Mendelian randomization (MR) analyses using two genetic risk scores (GRSs) as instrumental variables (IVs). The first GRS consisted of four single nucleotide polymorphisms (SNPs) in the CRP gene (GRS_{CRP}), and the second consisted of 18 SNPs that were significantly associated with CRP levels in the largest genome-wide association study (GWAS) to date (GRS_{GWAS}). To optimize power, we used summary statistics from GWAS consortia and tested the association of these two GRSs with 32 complex somatic and psychiatric outcomes, with up to 123,865 participants per outcome from populations of European ancestry. We performed heterogeneity tests to disentangle the pleiotropic effect of IVs. A Bonferroni-corrected significance level of less than 0.0016 was considered

statistically significant. An observed p -value equal to or less than 0.05 was considered nominally significant evidence for a potential causal association, yet to be confirmed.

The strengths (F -statistics) of the IVs were 31.92–3,761.29 and 82.32–9,403.21 for GRS_{CRP} and GRS_{GWAS} , respectively. CRP GRS_{GWAS} showed a statistically significant protective relationship of a 10% genetically elevated CRP level with the risk of schizophrenia (odds ratio [OR] 0.86 [95% CI 0.79–0.94]; $p < 0.001$). We validated this finding with individual-level genotype data from the schizophrenia GWAS (OR 0.96 [95% CI 0.94–0.98]; $p < 1.72 \times 10^{-6}$). Further, we found that a standardized CRP polygenic risk score (CRP_{PRS}) at p -value thresholds of 1×10^{-4} , 0.001, 0.01, 0.05, and 0.1 using individual-level data also showed a protective effect (OR < 1.00) against schizophrenia; the first CRP_{PRS} (built of SNPs with $p < 1 \times 10^{-4}$) showed a statistically significant ($p < 2.45 \times 10^{-4}$) protective effect with an OR of 0.97 (95% CI 0.95–0.99). The CRP GRS_{GWAS} showed that a 10% increase in genetically determined CRP level was significantly associated with coronary artery disease (OR 0.88 [95% CI 0.84–0.94]; $p < 2.4 \times 10^{-5}$) and was nominally associated with the risk of inflammatory bowel disease (OR 0.85 [95% CI 0.74–0.98]; $p < 0.03$), Crohn disease (OR 0.81 [95% CI 0.70–0.94]; $p < 0.005$), psoriatic arthritis (OR 1.36 [95% CI 1.00–1.84]; $p < 0.049$), knee osteoarthritis (OR 1.17 [95% CI 1.01–1.36]; $p < 0.04$), and bipolar disorder (OR 1.21 [95% CI 1.05–1.40]; $p < 0.007$) and with an increase of 0.72 (95% CI 0.11–1.34; $p < 0.02$) mm Hg in systolic blood pressure, 0.45 (95% CI 0.06–0.84; $p < 0.02$) mm Hg in diastolic blood pressure, 0.01 ml/min/1.73 m² (95% CI 0.003–0.02; $p < 0.005$) in estimated glomerular filtration rate from serum creatinine, 0.01 g/dl (95% CI 0.0004–0.02; $p < 0.04$) in serum albumin level, and 0.03 g/dl (95% CI 0.008–0.05; $p < 0.009$) in serum protein level. However, after adjustment for heterogeneity, neither GRS showed a significant effect of CRP level (at $p < 0.0016$) on any of these outcomes, including coronary artery disease, nor on the other 20 complex outcomes studied. Our study has two potential limitations: the limited variance explained by our genetic instruments modeling CRP levels in blood and the unobserved bias introduced by the use of summary statistics in our MR analyses.

Conclusions

Genetically elevated CRP levels showed a significant potentially protective causal relationship with risk of schizophrenia. We observed nominal evidence at an observed $p < 0.05$ using either GRS_{CRP} or GRS_{GWAS} —with persistence after correction for heterogeneity—for a causal relationship of elevated CRP levels with psoriatic osteoarthritis, rheumatoid arthritis, knee osteoarthritis, systolic blood pressure, diastolic blood pressure, serum albumin, and bipolar disorder. These associations remain yet to be confirmed. We cannot verify any causal effect of CRP level on any of the other common somatic and neuropsychiatric outcomes investigated in the present study. This implies that interventions that lower CRP level are unlikely to result in decreased risk for the majority of common complex outcomes.

Introduction

Emerging evidence suggests that persistent dysregulation of the inflammatory response is linked to a plethora of complex somatic and neuropsychiatric disorders [1–18]. Epidemiological

studies have shown that C-reactive protein (CRP), a well-studied biomarker of inflammation, is associated with and exhibits reliable predictive value for cardiovascular disease [19,20], type 2 diabetes [21], immunity-related disorders such as inflammatory bowel disease (IBD) [22], rheumatoid arthritis [23], and all-cause mortality [20,24]. Nevertheless, the evidence for a causal involvement of CRP in these outcomes from traditional experimental or observational studies remains controversial [25,26], fueling the debate surrounding whether CRP contributes to the chain of causality in disease mechanisms [27]. The use of genetically informed instrumental variables (IVs), termed Mendelian randomization (MR), is a complementary approach to epidemiological observations and allows investigation of whether the effect of an exposure (i.e., CRP level) on observed outcome phenotypes is likely to be causal [28].

Recent large-scale MR studies, focusing mainly on cardiovascular disease and metabolic traits, failed to show a causal association between CRP level and these outcomes (S1 Table). This has led to the notion that elevated CRP levels do not causally contribute to these traits and disorders. However, these studies used a single CRP-associated single nucleoid polymorphism (SNP) or a very limited set of CRP-associated SNPs (S1 Table). Common SNPs serving as proxies for CRP level represent only a small effect on CRP level per se and thus require a large enough sample size to detect causal effects on the outcome. Moreover, most studies have generally included a limited range of common complex diseases, often not more than two or three outcomes, or they have been performed in a single or small population, yielding inadequate study power (S1 Table). In other words, the evidence for a causal relationship between CRP and a broad range of common traits or diseases remains inconclusive. This is mostly due to the lack of well-powered MR studies that use optimally informative genetic IVs for CRP. Here, we sought to comprehensively examine the hypothesis that genetically determined CRP level directly contributes to common somatic and psychiatric outcomes. To optimize IV power, we applied a MR approach using summary statistics from large-scale genome-wide association study (GWAS) consortia of 32 somatic and psychiatric phenotypes for the four *CRP* variants representing 98% of the common variation in the *CRP* gene and for the largest known set of independent SNPs known to be associated with CRP. We further aimed to confirm the identified association between CRP and schizophrenia using a CRP polygenic risk score (CRP_{PRS}) from individual-level genotype data from the largest consortium of schizophrenia to date. We performed an *in silico* pathway analysis (see Discussion) to provide insights into the possible mechanism underlying the observed association of CRP level with schizophrenia.

Methods

Study Design and Rationale

The present MR study consists of two key components. First, we used established gene variants associated with CRP level and combined them to build two genetic risk scores (GRSs) for CRP. The first GRS consisted of only four SNPs in the *CRP* gene (GRS_{CRP}) selected from the largest recent MR study of CRP [29], and the second consisted of 18 SNPs that were associated with CRP level at a genome-wide significance level in the largest GWAS for CRP to date (GRS_{GWAS}) [30]. Second, we obtained summary association statistics from GWAS consortia for a panel of 32 common somatic and psychiatric outcomes (Table 1). BPP and BZA selected the studies, and contacted each consortium with a standardized request for study data, including the name of the study or consortium, the number of cases and controls, the number of available CRP SNPs for GRS_{CRP} and GRS_{GWAS} , and the estimated effect for each SNP (or its proxy) on outcome, i.e., the per allele regression coefficient with standard error or the odds ratio (OR) and corresponding 95% confidence interval. Data were available for 32 different outcomes in five broad disease/trait classes (autoimmune/inflammatory, cardiovascular, metabolic,

Table 1. Diseases and traits included in this study.

Disease or Trait	Cases	Controls	Total	Reference
Autoimmune/inflammatory				
Celiac disease	4,533	10,750	15,283	[31]
IBD (all types)	13,020	34,774	47,794	[32,33]
Crohn disease	6,333	15,056	21,389	[32]
Ulcerative colitis	6,687	19,718	26,405	[33]
Psoriasis vulgaris	4,007	4,934	8,941	[34,35]
Psoriatic arthritis	1,946	4,934	6,880	[34,35]
Cutaneous psoriasis	1,363	3,517	4,880	[34,35]
Rheumatoid arthritis	5,538	20,167	25,705	[36]
Systemic lupus erythematosus	1,311	3,340	4,651	[37]
Systemic sclerosis	2,356	5,187	7,543	[38]
Type 1 diabetes	9,934	16,956	26,890	[39]
Knee osteoarthritis	5,755	18,505	24,260	[40]
Cardiovascular				
Coronary artery disease	60,801	123,504	184,305	[41]
Systolic blood pressure	—	—	69,368	[42]
Diastolic blood pressure	—	—	69,372	[42]
Ischemic stroke (all types)	3,548	5,972	9,520	[43]
Ischemic stroke (cardioembolic)	790	5,972	6,762	[43]
Ischemic stroke (large vessel)	844	5,972	6,816	[43]
Ischemic stroke (small vessel)	580	5,972	6,522	[43]
Metabolic				
Body mass index	—	—	123,865	[44]
Type 2 diabetes	6,698	15,872	22,570	[45]
Chronic kidney disease	6,271	68,083	74,354	[46]
eGFR _{cr}	—	—	74,354	[46]
Serum albumin level	—	—	53,189	[47]
Serum protein level	—	—	25,537	[47]
Neurodegenerative				
Amyotrophic lateral sclerosis	4,133	8,130	12,663	[48]
Alzheimer disease	4,663	8,357	13,020	[49]
Parkinson disease	5,333	12,019	17,352	[50]
Psychiatric				
Autism	90	1,476	1,566	[51]
Bipolar disorder	7,481	9,250	16,731	[52]
Major depressive disorder	9,240	9,519	18,759	[53]
Schizophrenia	34,241	45,604	79,845	[54]

eGFR_{cr}, estimated glomerular filtration rate from serum creatinine.

doi:10.1371/journal.pmed.1001976.t001

neurodegenerative, and psychiatric), including at least 1,566, and up to 184,305, participants per outcome from populations of European ancestry (Table 1). These outcomes were selected based on the following two inclusion criteria: (i) the outcome having been associated with CRP level in epidemiological studies and (ii) availability of large meta-GWAS analyses for the outcome (Table 1).

Genetic Instruments

Weak IVs yielding insufficient statistical power may have hampered estimation of causal effects of CRP on the outcomes in previous analyses (S1 Table). Our MR approach, by using GWAS data and combining multiple independent SNPs into a GRS (i.e., IV), has the potential to greatly increase power. The selected SNPs have been described elsewhere [30,55,56] and are further detailed in S2–S4 Tables. These IVs were used to test the combined effect of the associations of CRP-level-influencing alleles with the outcomes. Our approach was implemented in such a way that the effects of both independent SNPs in the *CRP* gene (GRS_{CRP}) [55,56] (S1 Methods) and independent SNPs known to be genome-wide significantly associated with CRP levels (GRS_{GWAS}) [30], as well as pleiotropic effects of SNPs, could be discriminated [57]. Pleiotropy exists if CRP SNPs influence exposures (risk factors) other than CRP level and therefore violate one of the key MR assumptions.

Statistical Analysis

All analyses were done using the GRS function implemented in the `grs.summary` module of the R package Genetics ToolboX (version 2.15.1 for Windows). The `grs.summary` module approximates the regression of an outcome onto an additive GRS, using only single SNP association summary statistics extracted from GWAS results. The method is described in more detail elsewhere [58]. In brief, we performed MR analyses using GRS IVs in two steps. First, we used four individual *CRP* gene SNPs (i.e., IVs) associated with CRP level [56,59] (S2 and S3 Tables) to create a weighted GRS, named GRS_{CRP} , corresponding to the joint effect of the four SNPs within the *CRP* gene [55]. We extracted ω (the estimated coefficient, or weight) for individual SNPs from association results reported by the CRP Coronary Heart Disease Genetics Collaboration (CCGC) [29,55]; ω represents a one-unit (in mg/l) increase of the natural log of CRP level ($\ln CRP$) per dose of the coded allele. The four tagging SNPs represent 98% of the common variation in the *CRP* gene, assuming a minor allele frequency of ≥ 0.05 and an r^2 threshold of ≥ 0.8 , and aggregately explain $\sim 2\%$ of the total variation (i.e., phenotypic variance) in serum CRP level in populations of European descent [55,59]. Second, we constructed a multilocus GRS, named GRS_{GWAS} , that combined 18 SNPs associated with serum CRP level at a genome-wide significance level ($p < 5 \times 10^{-8}$; S2 and S3 Tables), derived from a large meta-GWAS analysis of CRP conducted by the CHARGE (Cohorts for Heart and Aging Research in Genomic Epidemiology) Consortium [30]. This multilocus GRS explains approximately $\sim 5\%$ of the total variation in serum CRP level [30].

We integrated ω for each CRP SNP from the reference data of CCGC [55] or meta-analysis of GWASs [30] for CRP level with the summary association statistics extracted from the GWAS consortium data for each outcome (S1 Data; S2 Methods). This MR approach using meta-GWAS summary statistics data is equivalent to an inverse-variance-weighted meta-analysis and has previously been validated in comparison to individual-level data [57,60]. To estimate the causal effect of CRP level on an outcome, we obtained the β values (estimated effects from regression analysis) for the effects of CRP SNPs on the outcome, with standard errors, se_{β} , from the corresponding GWAS results. Where no summary statistics for a CRP SNP in the GRS IVs were available in the look-up dataset, we chose the proxy SNP that had the highest linkage disequilibrium with the initial SNP ($r^2 > 0.9$ in HapMap release 22; S3 Table). If several proxy SNPs had the exact same r^2 value, we chose the proxy nearest to the original SNP in the instrument. Separate regressions of outcomes on GRSs were performed to calculate α_{IV} estimators (i.e., causal IV estimators) for each outcome. Correspondingly, the value of a GRS is the sum of the ω values multiplied by the allele dosage (i.e., 0, 1, or 2) for each CRP SNP in the CCGC or in the CHARGE Consortium data [30,55]. For uncorrelated SNPs, when maximizing

the likelihood function, the α_{IV} value and its standard error, se_{α} , can be approximated with the formula $\alpha \cong (\Sigma\omega \times \beta \times se_{\beta}^{-2}) / (\Sigma\omega^2 \times se_{\beta}^{-2})$, with $se_{\alpha} \cong \sqrt{1 / \Sigma\omega^2 \times se_{\beta}^{-2}}$. InCRP was used as the outcome in reference studies [30,55], so in obtaining the ω values (i.e., effect sizes) for each of the CRP SNPs, a unit increase in InCRP equals a 10 symmetric percentage (%) increase in CRP level, which corresponds to a unit change in the level of a continuous outcome or logit of risk estimate (i.e., beta coefficient) for a dichotomous outcome [61]. The α_{IV} value (i.e., causal estimate) for each CRP SNP is, therefore, presented for each outcome as corresponding to a 10-s% increase in actual CRP level. During the course of this study, an updated, larger GWAS dataset for coronary artery disease (CAD) became publicly available (CARDIoGRAMplusC4D Consortium, release 2015 [41]); we therefore redid the analysis for CAD using the release 2015 data.

To assess which SNPs might have violated the key MR assumption regarding pleiotropy, we performed goodness-of-fit tests to correct both GRSs for the heterogeneity of their corresponding SNPs' effects on each outcome. Heterogeneity, which indicates the potential presence of pleiotropy, was measured using the Q statistic and was considered statistically significant at a conservative uncorrected p -value of <0.05 . Although heterogeneity could be an indicator of pleiotropy, there are other factors that could introduce heterogeneity in the analyses. Even though the adjustments for heterogeneity that we have made could be overconservative, we have used this method in order to minimize false positives. After stepwise removal of SNPs with potential pleiotropic effects, we repeated the analyses until significant heterogeneity was no longer observed.

To further ensure the strength of these two GRSs as IVs, we generated an F -statistic for each outcome. We used variance in InCRP explained by each set of CRP SNPs (2% and 5%, respectively, for GRS_{CRP} and GRS_{GWAS}) to calculate the F -statistic using the formula $F\text{-statistic} = [R^2 \times (n - 1 - K)] / [(1 - R^2) \times K]$, where R^2 represents the proportion of variability in CRP level that is explained by the GRS, n represents sample size, and K represents the number of IVs included in model (i.e., for this study $K = 1$) [62]. As a rule of thumb, an F -value above ten indicates that a causal estimate is unlikely to be biased due to weak instruments [57].

Multiple Testing

The present study included 32 independent sample sets. For each sample set, we did one statistical test, for which a global nominal significance level of ≤ 0.05 was considered as satisfactory to derive conclusions. The need for correction for multiple testing is debatable. Nevertheless, to ensure the validity of our conclusions, we took a conservative approach and applied a Bonferroni-corrected significance threshold calculated as 0.05 divided by 32 (i.e., 0.0016). We considered a statistical test with an observed p -value more than 0.05 as a definitely nonsignificant result, i.e., no association; an observed p -value equal to or less than 0.05 as nominally significant evidence for a potential, but yet to be confirmed, causal association; and an observed p -value equal to or less than 0.0016 as statistically significant evidence for a causal association.

CRP Polygenic Risk Score and Schizophrenia Using Individual-Level Data

In an ancillary follow-up study, inspired by comments by the editors and the reviewers, we aimed further to determine whether GRS_{GWAS} was causally associated with schizophrenia using individual-level data retrieved from the Psychiatric Genomics Consortium (PGC) schizophrenia dataset (S3 Methods) [54]. This dataset consisted of 36 independent cohorts with a combined 25,629 cases and 30,976 controls for which we had ethics approval (S4 Methods). Three family-based samples of European ancestry (1,235 parent-affected offspring trios) were

excluded from our analysis. To evaluate whether the observed protective causal association between GRS_{GWAS} and schizophrenia was persistent, we investigated whether the CRP_{PRS} was also protectively associated with schizophrenia. Briefly, CRP_{PRS} values were calculated for each individual by summing the total effect of the SNP dosages by their effect size. In addition to the 18 genome-wide significant CRP SNPs, we grouped subthreshold CRP-associated SNPs at the following p -value thresholds: 1×10^{-4} , 0.001, 0.01, 0.05, and 0.1. Standardized CRP_{PRS} values were tested for association with schizophrenia case status in each cohort with adjustment for ten principal components (PCs). A fixed effects inverse-variance-weighted meta-analysis was performed across all 36 cohorts to obtain the overall effect size estimate as explained in [S4 Methods](#) and elsewhere [63]. The variance in schizophrenia case status explained by CRP_{PRS} was estimated using the deviation in Nagelkerke's pseudo- R^2 between a null model (which included ten PCs) and the full model (which included GRS in addition to the ten PCs), calculated in R using the Functions for Medical Statistics Book with Some Demographic Data (fmsb) R package ([S3 Methods](#)). Similar to previous studies, the statistical significance of CRP_{PRS} values was estimated based on their logistic regression coefficient [64], and reported CRP_{PRS} ORs correspond to a 1-SD increase in CRP_{PRS} [65].

Results

Using GRS_{CRP} , we first tested whether a CRP -gene-determined increase in $\ln CRP$ was associated with each outcome. In [Table 2](#), the causal effects of $\ln CRP$ estimated for each outcome are summarized. We found no heterogeneity in the IV analyses ($p_{\text{heterogeneity}} \geq 0.11$ for all outcomes), and GRS_{CRP} was a strong instrument ($F \geq 31$). IV analyses provided nominal evidence for potential causal relationships of $\ln CRP$ with risk of Crohn disease (OR 0.78 [95% CI 0.65–0.94]; $p < 0.009$), psoriatic arthritis (1.45 [1.04–2.04]; $p < 0.03$), and schizophrenia (0.90 [0.82–0.99]; $p < 0.03$), and with an increase in systolic blood pressure (SBP) (mean increase 1.23 mm Hg per 10-s% increase in CRP level [95% CI 0.45–2.01]; $p < 0.002$) and diastolic blood pressure (DBP) (0.70 [0.20–1.19]; $p < 0.006$). GRS_{CRP} showed no significant effect on any of the other outcomes ([Table 2](#); [S1 Fig](#)).

GRS_{GWAS} showed a statistically significant protective effect of $\ln CRP$ on the risk of schizophrenia (per 10-s% increase in CRP level, OR 0.86 [95% CI 0.79–0.94]; $p < 0.0010$) ([Figs 1 and S1](#); [Table 3](#)). In a follow-up analysis using the individual-level PGC data, we found that a GRS incorporating the same 18 CRP SNPs used to construct the GRS_{GWAS} was again significantly associated with a lower risk of schizophrenia (OR 0.96 [95% CI 0.94–0.98]; $p < 1.72 \times 10^{-6}$). This signal persisted when we included all SNPs meeting a less stringent p -value threshold of 1×10^{-4} (OR 0.97 [95% CI 0.95–0.99]; $p < 2.45 \times 10^{-4}$). At less stringent p -value thresholds, less variance was explained by the logistic model, and the protective effect of CRP risk scores became less significant, but across all p -value thresholds, the direction of the effect was consistently protective ([Figs 2 and 3](#)). To ensure that the association between risk alleles for CRP and schizophrenia was not driven by a small number of genome-wide significant SNPs, we performed a leave-one-out sensitivity analysis of the 18 genome-wide SNPs. In the 18 sets of 17 SNPs, the variance explained (Nagelkerke's pseudo- R^2) ranged from 0.012% to 0.034%, with p -values ranging from 9.3×10^{-5} to 1.6×10^{-2} , suggesting that the protective effect observed between risk alleles for CRP and schizophrenia was not driven by a small number of SNPs with large effects.

GRS_{GWAS} also showed moderate but nominally significant effects of $\ln CRP$ on the risk of IBD (OR 0.85 [95% CI 0.74–0.98]; $p < 0.03$), Crohn disease (0.81 [0.70–0.94]; $p < 0.005$), psoriatic arthritis (1.36 [1.00–1.84]; $p < 0.049$), knee osteoarthritis (1.17 [1.01–1.36]; $p < 0.04$), and bipolar disorder (1.21 [1.05–1.40]; $p < 0.007$), while its effect was statistically significant

Table 2. The effect of the CRP genetic risk score instrument of four SNPs in CRP (GRS_{CRP}) with somatic and neuropsychiatric outcomes.

Disease or Trait	M	N	Effect Size (95% CI) ¹	Goodness-of-Fit Test p-Value	p-Het	F-Value
Autoimmune/inflammatory						
Celiac disease	3	15,283	0.96 (0.77 to 1.21)	0.750	0.19	311.86
IBD (all types)	3	47,794	0.97 (0.84 to 1.13)	0.700	0.30	975.35
Crohn disease	4	21,389	0.78 (0.65 to 0.94)	0.009	0.25	436.47
Ulcerative colitis	4	26,405	1.10 (0.92 to 1.31)	0.290	0.92	538.84
Psoriasis vulgaris	4	8,941	1.23 (0.96 to 1.57)	0.110	0.95	182.43
Psoriatic arthritis	4	6,880	1.45 (1.04 to 2.04)	0.030	0.92	140.37
Cutaneous psoriasis	4	4,880	1.10 (0.76 to 1.59)	0.620	0.60	99.55
Rheumatoid arthritis	4	25,702	0.94 (0.77 to 1.15)	0.550	0.17	524.55
Systemic lupus erythematosus	3	4,651	1.20 (0.80 to 1.81)	0.380	0.19	94.88
Systemic sclerosis	3	7,518	1.07 (0.78 to 1.45)	0.680	0.85	153.90
Type 1 diabetes	2	26,890	1.15 (0.90 to 1.47)	0.260	0.34	548.73
Knee osteoarthritis	4	24,260	0.94 (0.78 to 1.13)	0.500	0.23	495.06
Cardiovascular						
CAD	4	184,305	1.00 (0.93 to 1.07)	0.965	0.65	1,775.37
SBP ²	4	69,372	1.23 (0.45 to 2.01)	0.002	0.51	1,415.63
DBP ²	4	69,368	0.70 (0.2x to 1.19)	0.006	0.68	1,415.71
Ischemic stroke (all types)	4	9,520	1.19 (0.93 to 1.53)	0.160	0.93	194.24
Ischemic stroke (cardioembolic)	4	6,762	1.02 (0.65 to 1.58)	0.940	0.96	137.96
Ischemic stroke (large vessel)	4	6,816	1.44 (0.93 to 2.21)	0.100	0.31	139.06
Ischemic stroke (small vessel)	4	6,552	1.18 (0.71 to 1.95)	0.520	0.36	133.06
Metabolic						
Body mass index ³	4	123,864	-0.017 (-0.06 to 0.02)	0.410	0.50	2,527.82
Type 2 diabetes	4	22,570	1.11 (0.94 to 1.32)	0.230	0.50	460.57
Chronic kidney disease	4	74,354	1.04 (0.88 to 1.22)	0.670	0.90	1,517.39
eGFR _{cr} ⁴	4	74,354	0.004 (-0.01 to 0.02)	0.400	0.88	1,517.39
Serum albumin level ⁵	4	53,189	-0.002 (-0.02 to 0.01)	0.770	0.88	1,085.45
Serum protein level ⁵	4	25,537	0.008 (-0.02 to 0.04)	0.640	0.12	521.12
Neurodegenerative						
Amyotrophic lateral sclerosis	2	12,263	0.79 (0.60 to 1.04)	0.090	0.23	258.39
Alzheimer disease	2	13,020	1.26 (0.89 to 1.78)	0.200	0.11	265.67
Parkinson disease	3	17,352	1.00 (0.85 to 1.17)	0.960	0.33	354.08
Psychiatric						
Autism	3	1,566	1.02 (0.97 to 1.07)	0.380	0.69	31.92
Bipolar disorder	4	16,731	1.17 (0.97 to 1.42)	0.110	0.49	341.41
Major depressive disorder	3	18,759	0.98 (0.81 to 1.18)	0.810	0.86	382.80
Schizophrenia	3	79,845	0.90 (0.82 to 0.99)	0.030	0.79	1,629.45

¹Effect size (95% CI) per 1-mg/l increase in lnCRP. For risk of disease, effect size is given as an OR, otherwise given in the specific units in which the outcome was measured. Derived from the IV causal estimator α .

²Effect size unit is millimeters of mercury per 1-mg/l increase in lnCRP.

³Effect size unit is standard deviations per 1-mg/l increase in lnCRP (the body mass index results were inverse normal transformed to a distribution with $\mu = 0$ and $\sigma = 1$).

⁴Effect size unit is milliliters/minute/1.73 m² per 1-mg/l increase in lnCRP.

⁵Effect size unit is grams/deciliter per 1-mg/l increase in lnCRP.

eGFR_{cr}, estimated glomerular filtration rate from serum creatinine; F-value, F-statistic value for the genetic instrument; M, number of markers used in the genetic instrument; N, number of samples in the disease/trait meta-analysis; p-het, p-value of heterogeneity of effect test.

doi:10.1371/journal.pmed.1001976.t002

CRP_{GWAS} schizophrenia

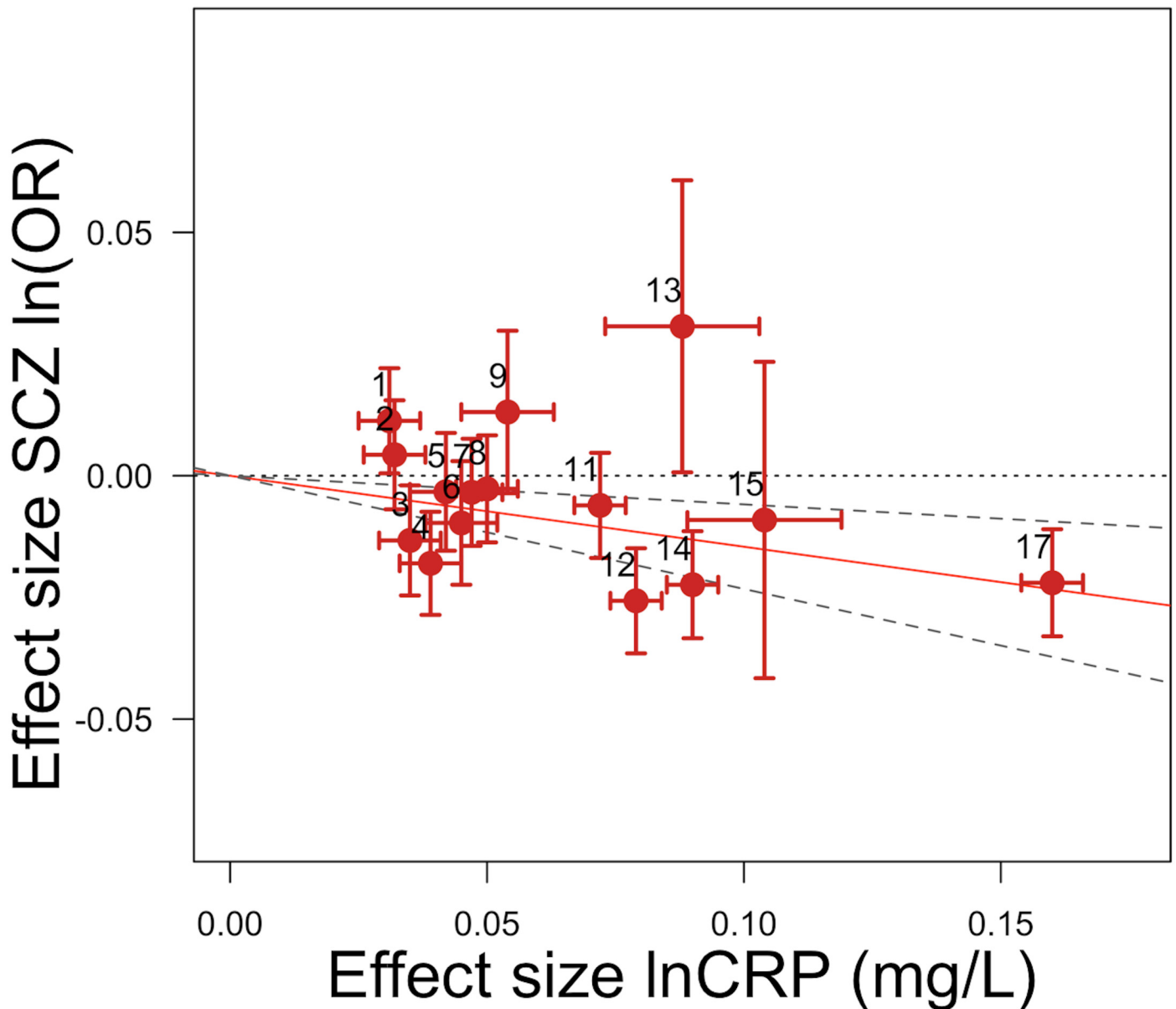


Fig 1. Genetic risk score GRS_{GWAS} for schizophrenia. The x-axis shows the effect size for the 15 SNPs for which data were available in the PGC schizophrenia dataset comprising the GRS_{GWAS} influencing levels of CRP, with corresponding standard error bars. The y-axis shows the log OR of the GRS_{GWAS} SNPs for schizophrenia (SCZ) with corresponding standard error bars. The effect estimate of CRP level on disease risk is represented by the red solid line, with gradient α . The 95% CI of this α estimate is represented by the grey dashed lines. The included SNPs are shown by Arabic numbering: #1, rs2847281 (gene: *PTPN2*; chromosome: 18; basepair position: 12811593); #2, rs340029 (*RORA*; 15; 58682257); #3, rs6901250 (*GPRC6A*; 6; 117220718); #4, rs10745954 (*ASCL1*; 12; 102007224); #5, rs4705952 (*IRF1*; 5; 131867517); #6, rs12037222 (*PABPC4*; 1; 39837548); #7, rs12239046 (*NLRP3*; 1; 245668218); #8, rs6734238 (*IL1F10*; 2; 113557501); #9, rs13233571 (*BCL7B*; 7; 72609167); #11, rs1260326 (*GCKR*; 2; 27584444); #12, rs4129267 (*IL6R*; 1; 152692888); #13, rs1800961 (*HNF4A*; 20; 42475778); #14, rs4420065 (*LEPR*; 1; 5934049); #15, rs10521222 (*SALL1*; 16; 49716211); #17, rs2794520 (*CRP*; 1; 157945440). The three SNPs of #10, rs9987289 (*PPP1R3B*; 8; 9220768); #16, rs1183910 (*HNF1A*); and #18, rs4420638 (*APOC1*; 19; 50114786) were not present in the data of the PGC.

doi:10.1371/journal.pmed.1001976.g001

Table 3. The effect of the CRP genetic risk score instrument of 18 SNPs associated with CRP (GRS_{GWAS}) on somatic and neuropsychiatric outcomes.

Disease or Trait	<i>M</i>	Effect Size (95% CI) ¹	Goodness-of-Fit Test <i>p</i> -Value	<i>p</i> -Het	<i>F</i> -Value
Autoimmune/inflammatory					
Celiac disease	18	0.99 (0.85 to 1.16)	0.930	7.2×10^{-4}	804.26
IBD (all types)	15	0.85 (0.74 to 0.98)	0.030	1.4×10^{-5}	2,515.37
Crohn disease	17	0.81 (0.70 to 0.94)	0.005	4.4×10^{-7}	1,125.63
Ulcerative colitis	17	1.05 (0.91 to 1.21)	0.490	0.01	1,389.63
Psoriasis vulgaris	17	1.12 (0.90 to 1.40)	0.310	0.19	470.47
Psoriatic arthritis	17	1.36 (1.00 to 1.84)	0.049	0.04	362.00
Cutaneous psoriasis	17	1.00 (0.72 to 1.39)	0.990	0.16	256.74
Rheumatoid arthritis	18	0.93 (0.80 to 1.08)	0.350	1.8×10^{-6}	1,352.79
Systemic lupus erythematosus	11	1.06 (0.71 to 1.58)	0.780	0.27	244.68
Systemic sclerosis	11	0.84 (0.62 to 1.14)	0.280	0.63	396.89
Type 1 diabetes	15	1.10 (0.92 to 1.31)	0.310	3.47×10^{-3}	1,415.16
Knee osteoarthritis	18	1.17 (1.01 to 1.36)	0.040	0.10	1,276.74
Cardiovascular					
CAD	18	0.88 (0.84 to 0.94)	2.4×10^{-5}	7.5×10^{-12}	9,403.21
SBP ²	18	0.72 (0.11 to 1.34)	0.020	0.14	3,650.84
DBP ²	18	0.45 (0.06 to 0.84)	0.020	0.02	3,651.05
Ischemic stroke (all types)	18	1.06 (0.87 to 1.29)	0.570	0.37	500.95
Ischemic stroke (cardioembolic)	18	0.98 (0.69 to 1.39)	0.920	0.35	355.79
Ischemic stroke (large vessel)	18	1.30 (0.92 to 1.82)	0.140	0.97	358.63
Ischemic stroke (small vessel)	18	0.85 (0.58 to 1.25)	0.420	0.76	343.16
Metabolic					
Body mass index ³	18	-0.005 (-0.03 to 0.02)	0.740	0.11	6,519.11
Type 2 diabetes	18	1.090 (0.95 to 1.24)	0.210	1.8×10^{-3}	1,187.79
Chronic kidney disease	18	0.960 (0.84 to 1.09)	0.500	0.07	3,913.26
eGFR _{cr} ⁴	18	0.011 (0.003 to 0.02)	0.005	7.2×10^{-9}	3,913.26
Serum albumin level ⁵	18	0.011 (0.0004 to 0.02)	0.041	2.3×10^{-18}	2,799.32
Serum protein level ⁵	18	0.031 (0.008 to 0.05)	0.009	0.03	1,343.95
Neurodegenerative					
Amyotrophic lateral sclerosis	8	1.01 (0.79 to 1.29)	0.960	0.56	666.37
Alzheimer disease	11	1.26 (0.99 to 1.61)	0.060	0.23	685.16
Parkinson disease	10	1.06 (0.90 to 1.25)	0.500	0.50	913.16
Psychiatric					
Autism	9	0.89 (0.70 to 1.13)	0.350	0.99	82.32
Bipolar disorder	18	1.21 (1.05 to 1.40)	0.007	0.15	880.47
Major depressive disorder	15	1.14 (0.96 to 1.36)	0.140	0.84	987.21
Schizophrenia	15	0.86 (0.79 to 0.94)	0.001	0.66	4,202.26

¹Effect size (95% CI) per 1-mg/l increase in lnCRP. For risk of disease, effect size is given as an OR, otherwise given in the specific units in which the outcome was measured. Derived from the IV causal estimator α .

²Effect size unit is millimeters of mercury per 1-mg/l increase in lnCRP.

³Effect size unit is standard deviations per 1-mg/l increase in lnCRP (the body mass index results were inverse normal transformed to a distribution with $\mu = 0$ and $\sigma = 1$).

⁴Effect size unit is milliliters/minute/1.73 m² per 1-mg/l increase in lnCRP.

⁵Effect size unit is grams/deciliter per 1-mg/l increase in lnCRP.

eGFR_{cr}, estimated glomerular filtration rate from serum creatinine; *F*-value, *F*-statistic value for the genetic instrument; *M*, number of markers used in the genetic instrument; *p*-het, *p*-value of heterogeneity of effect test.

doi:10.1371/journal.pmed.1001976.t003

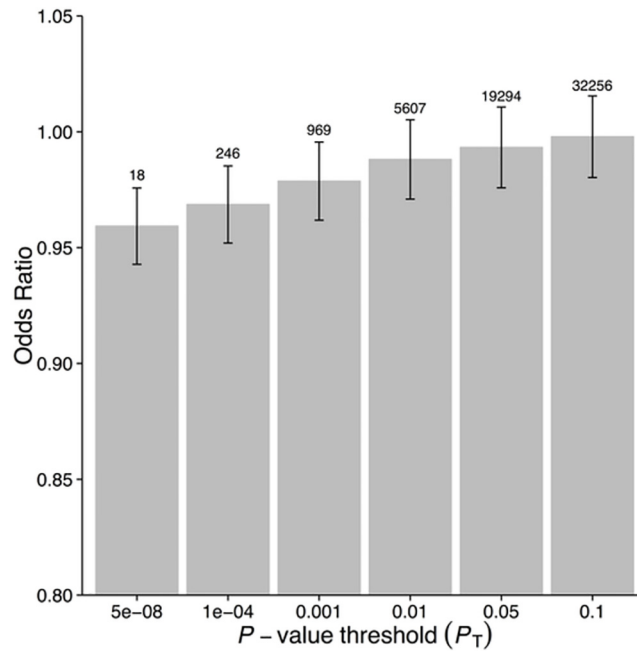


Fig 2. Polygenic risk scores for elevated CRP level and protective effect on schizophrenia, using individual-level genetic data.

doi:10.1371/journal.pmed.1001976.g002

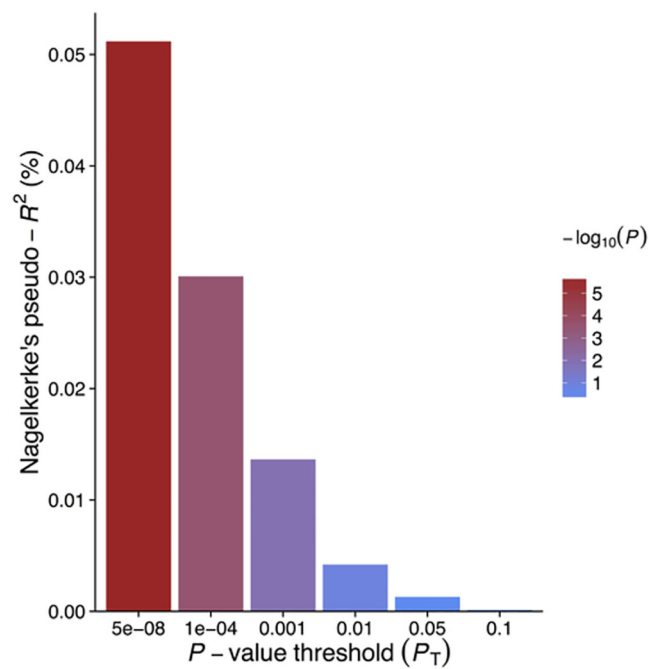


Fig 3. Polygenic risk scores for elevated CRP level and explained variance of schizophrenia using individual-level genetic data.

doi:10.1371/journal.pmed.1001976.g003

for CAD (0.88 [0.84–0.94]; $p < 2.4 \times 10^{-5}$) (Table 3; Figs 4 and S1). GRS_{GWAS} revealed a nominally significant effect of \ln CRP on blood pressure: an increase of 0.72 (95% CI 0.11–1.34; $p < 0.02$) and 0.45 (0.06–0.84; $p < 0.02$) mm Hg in SBP and DBP, respectively (Table 3; S1 Fig). Likewise, a genetically determined 10-s% increase in CRP level was nominally associated with a 0.01 ml/min/1.73 m² (95% CI 0.003–0.02; $p < 0.005$) higher estimated glomerular filtration rate from serum creatinine (eGFR_{cr}), a 0.01 g/dl (0.0004–0.02; $p < 0.04$) higher serum albumin level, and a 0.03 g/dl (0.008–0.05; $p < 0.009$) higher serum protein level. The remaining outcomes tested for causal associations using GRS_{GWAS} did not reach statistical significance, though the corresponding GRS_{GWAS} proved to be a strong IV, with F -values ≥ 82 (Table 3; S1 Fig).

Using GRS_{GWAS} , there was no significant evidence of heterogeneity of the effect size for knee osteoarthritis, bipolar disorder, schizophrenia, or SBP, while the heterogeneity test was statistically significant for psoriatic arthritis, IBD, Crohn disease, CAD, DBP, eGFR_{cr}, serum albumin, and serum protein. These heterogeneities in the effects of GRS_{GWAS} may be attributable to pleiotropic effects of the SNPs used to build the GRS_{GWAS} . We subsequently performed a stepwise removal of SNPs from GRS_{GWAS} until no significant heterogeneity remained (Table 4). This adjustment in the GRS_{GWAS} resulted in the removal of three SNPs from the GRS_{GWAS} for IBD (in *GCKR*, *IRF1*, and *PTPN2*), five SNPs from the GRS_{GWAS} for Crohn disease (in *GCKR*, *IL6R*, *IRF1*, *PABPC4*, and *PTPN2*), one SNP from the GRS_{GWAS} for psoriatic arthritis (in *IRF1*), three SNPs for CAD (in *APOC1*, *HNF1A*, and *IL6R*), one SNP from the GRS_{GWAS} for DBP (in *PABPC4*), two SNPs from the GRS_{GWAS} for eGFR_{cr} (in *LEPR* and *GCKR*), six SNPs from the GRS_{GWAS} for serum albumin level (in *APOC1*, *BCL7B*, *GCKR*, *PPP1R3B*, *PTPN2*, and *IRF1*), and one SNP from the GRS_{GWAS} for serum protein level (in *GCKR*). After removal of these variants from the GRS_{GWAS} , we found no statistically significant (at $p < 0.0016$) association between genetically increased \ln CRP level and any of these outcomes (Table 4). However, the effect estimate of CRP on DBP, serum albumin, and psoriatic arthritis showed nominal association at $p < 0.05$. For example, for DBP, 17 SNPs remained in the GRS_{GWAS} and yielded a slightly lower causal estimate (compared to the values before adjustment) of a 0.39 (95% CI –0.01 to 0.78) mm Hg increase in DBP per 10-s% increase in \ln CRP level, with a nominal significance of $p < 0.05$.

Likewise, we hypothesized that the fact that GRS_{GWAS} showed a nonsignificant effect of CRP on celiac disease, ulcerative colitis, rheumatoid arthritis, type 1 diabetes, and type 2 diabetes can be to some extent explained by the significant heterogeneity observed for these outcomes (Table 3). The stepwise adjustment in the GRS_{GWAS} resulted in the removal of two SNPs from the GRS_{GWAS} for celiac disease (in *PABPC4* and *PTPN2*), one SNP from the GRS_{GWAS} for ulcerative colitis (in *GCKR*), five SNPs from the GRS_{GWAS} for rheumatoid arthritis (in *HNF4A*, *IL6R*, *SALL1*, *NLRP3*, and *PTPN2*), one SNP from the GRS_{GWAS} for type 1 diabetes (in *PTPN2*), and one SNP from the GRS_{GWAS} for type 2 diabetes (in *APOC1*). After adjusting for heterogeneity, the association of GRS_{GWAS} with these outcomes remained statistically nonsignificant (Table 4).

Discussion

In this large-scale cross-consortium MR study of 32 complex outcomes, we found evidence for a potential protective causal relationship between elevated CRP level and schizophrenia in both genetic IVs (i.e., GRS_{CRP} and GRS_{GWAS}) and confirmed this protective relationship in follow-up analyses using individual-level genotype data from the schizophrenia GWAS. We also found a statistically significant association of CRP level with CAD, and nominally significant evidence for a predisposing causal association of CRP level with IBD, Crohn disease, psoriatic

CRP_{GWAS} bipolar disorder

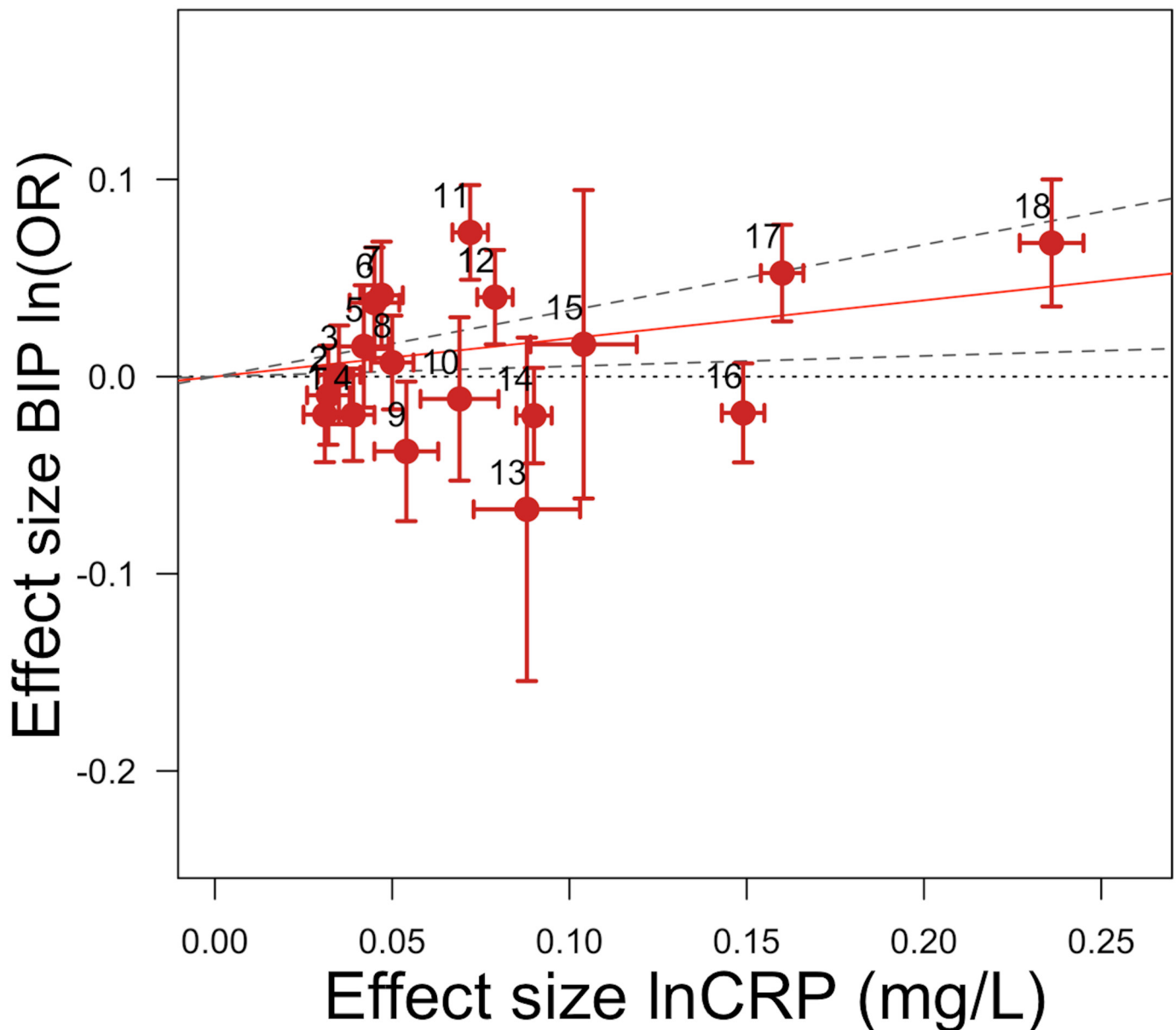


Fig 4. Genetic risk score GRS_{GWAS} for bipolar disorder. The x-axis shows the effect size for the 18 SNPs comprising the GRS_{GWAS} influencing levels of CRP, with corresponding standard error bars. The y-axis shows the log OR of the GRS_{GWAS} SNPs for bipolar disorder (BIP) with corresponding standard error bars. The effect estimate of CRP level on disease risk is represented by the red solid line, with gradient α . The 95% CI of this α estimate is represented by the grey dashed lines. The included SNPs are shown by Arabic numbering: #1, rs2847281 (gene: *PTPN2*; chromosome: 18; basepair position: 12811593); #2, rs340029 (*RORA*; 15; 58682257); #3, rs6901250 (*GPRC6A*; 6; 117220718); #4, rs10745954 (*ASCL1*; 12; 102007224); #5, rs4705952 (*IRF1*; 5; 131867517); #6, rs12037222 (*PABPC4*; 1; 39837548); #7, rs12239046 (*NLRP3*; 1; 245668218); #8, rs6734238 (*IL1F10*; 2; 113557501); #9, rs13233571 (*BCL7B*; 7; 72609167); #10, rs9987289 (*PPP1R3B*; 8; 9220768); #11, rs1260326 (*GCKR*; 2; 27584444); #12, rs4129267 (*IL6R*; 1; 152692888); #13, rs1800961 (*HNF4A*; 20; 42475778); #14, rs4420065 (*LEPR*; 1; 5934049); #15, rs10521222 (*SALL1*; 16; 49716211); #16, rs1183910 (*HNF1A*; 12; 119905190); #17, rs2794520 (*CRP*; 1; 157945440); #18, rs4420638 (*APOC1*; 19; 50114786).

doi:10.1371/journal.pmed.1001976.g004

Table 4. The effect of the CRP genetic risk score instrument of 18 SNPs associated with CRP (GRS_{GWAS}) on somatic and neuropsychiatric outcomes after correcting for heterogeneity.

Disease or Trait	<i>M</i>	Effect Size (95% CI) ¹	Goodness-of-Fit Test <i>p</i> -Value	<i>p</i> -Het
Autoimmune/inflammatory				
Celiac disease	16	1.05 (0.90 to 1.23)	0.56	0.10
IBD	12	0.92 (0.79 to 1.06)	0.24	0.14
Crohn disease	12	0.93 (0.79 to 1.08)	0.34	0.12
Ulcerative colitis	16	1.11 (0.96 to 1.28)	0.16	0.12
Psoriatic arthritis	16	1.42 (1.05 to 1.94)	0.02	0.14
Rheumatoid arthritis	13	0.83 (0.71 to 0.97)	0.02	0.09
Type 1 diabetes	14	1.06 (0.89 to 1.27)	0.52	0.07
Cardiovascular				
CAD	15	0.98 (0.91 to 1.06)	0.65	0.20
DBP ²	17	0.385 (0.008 to 0.78)	0.05	0.09
Metabolic				
Type 2 diabetes	17	0.95 (0.82 to 1.10)	0.52	0.09
eGFR _{cr} ³	16	0.001 (−0.007 to 0.01)	0.74	0.11
Serum albumin level ⁴	12	−0.017 (−0.03 to −0.004)	0.01	0.07
Serum protein level ⁴	17	0.021 (−0.002 to 0.05)	0.07	0.31

¹Effect size (95% CI) per 1-mg/l increase in lnCRP. For risk of disease, effect size is given as an OR, otherwise given in the specific units in which the outcome was measured. Derived from the IV causal estimator α .

²Effect size unit is millimeters of mercury per 1-mg/l increase in lnCRP.

³Effect size unit is milliliters/minute/1.73 m² per 1-mg/l increase in lnCRP.

⁴Effect size unit is grams/deciliter per 1-mg/l increase in lnCRP.

M, number of markers used in the genetic instrument; *p*-het, *p*-value of heterogeneity of effect test.

doi:10.1371/journal.pmed.1001976.t004

arthritis, knee osteoarthritis, SBP, DBP, eGFR_{cr}, serum albumin level, serum protein level, and bipolar disorder, using GRS_{GWAS} as an IV. However, after adjustment for heterogeneity, neither GRS showed a significant effect (at $p < 0.0016$) of CRP level on any of these outcomes, including CAD, nor on the 20 other common somatic and psychiatric outcomes we investigated, including celiac disease, ulcerative colitis, psoriasis (all types), rheumatoid arthritis, systemic lupus erythematosus, systemic sclerosis, type 1 and 2 diabetes, stroke (all types), body mass index, chronic kidney disease, amyotrophic lateral sclerosis, Alzheimer disease, Parkinson disease, autism, and major depressive disorder.

CRP Protection against Schizophrenia

Strikingly, as opposed to the current literature and previous inconclusive small-scale studies [66–68], our findings suggest that genetically elevated levels of CRP are not predisposing but in fact protective for schizophrenia. The significant causal protective role of CRP for schizophrenia was consistent in both IVs using summary statistics, i.e., GRS_{CRP} and GRS_{GWAS} . When incorporating 18 genome-wide CRP-associated SNPs using individual-level data, we confirmed a modest, but significant, protective effect of CRP level for schizophrenia. This signal persisted when we included all SNPs meeting a less stringent *p*-value threshold of 1×10^{-4} . Notably, the leave-one-out sensitivity analysis revealed that the genetic overlap between CRP level and schizophrenia we observed at genome-wide and 1×10^{-4} significance thresholds was not driven by a few major SNPs. In contrast, others have previously shown that CRP levels are significantly elevated in patients with schizophrenia [69,70], with a recent meta-analysis concluding

that the association between elevated CRP and schizophrenia is indeed robust [71]. Given that clinical studies report elevated CRP levels in schizophrenia, one would expect to find that alleles for elevated CRP would confer an increased risk for schizophrenia. The fact that we found the completely opposite effect—in a cohort of over 25,000 cases and 30,000 controls—should give one pause when deriving clinical meaning from these results. Our observation that a genetically determined marginal increase in the level of CRP is likely to be protective for schizophrenia may fuel the debate about whether the observed CRP elevation in schizophrenia is a by-product of the pathogenesis of schizophrenia or directly contributing to clinical manifestations of the disorder [6]. Our finding may also point out potential biases in previous studies regarding the causes of elevated CRP levels in patients with schizophrenia, such as reverse causality and/or pleiotropic effects within chosen instruments.

The exact mechanism for how elevated CRP levels are linked to schizophrenia requires a well-defined experimental analysis. In addition to CRP variants, other recent studies have identified several inflammatory genetic variants associated with schizophrenia and bipolar disorder, which include variants in the major histocompatibility complex (MHC) region on Chromosome 6p21 [72]—harboring many cytokine genes [54,73–76]—and in the *IL10* promoter [77], *TNF* promoter [78], *IL1B* [79], and C4 [80].

Biological Annotation

Following comments made by the reviewers, we explored the possible underlying pathways that may explain the potential protective causal association between CRP and schizophrenia. We performed a follow-up in silico functional pathway analysis using a previously reported approach [81] as summarized in S5 Methods and S4–S13 Tables. In brief, our results show that pathways associated with the interferon response are significantly enriched amongst genes harbored by CRP loci and their associated expression quantitative trait loci (eQTLs) and that there are differentially expressed genes between schizophrenia cases and controls. Previous studies showed that the induction of T cell IFN cytokine release stimulates microglia and astrocytes to facilitate glutamate clearance in neuronal cells without evoking inflammatory mediators [82,83]. One could speculate that CRP-interferon pathways may induce neuroprotection by contributing to glutamate clearance, leading to the protection of neurons against the oxidative stress associated with an excess of glutamate [84,85], and thereby offering a protective effect against schizophrenia.

CRP GRS_{GWAS} Association with Bipolar Disorder

As for bipolar disorder, we found a nominal effect of a 1.21-fold increase in risk for bipolar disorder with a 10-s% increase in CRP level. Though this nominal predisposing effect needs to be confirmed, our finding corroborates epidemiological observations suggesting that elevated CRP is associated with the disease and supports a potential causal influence of general inflammation in bipolar disorder [86]. We note that, though it may be biologically sensible, this result failed to pass multiple testing correction. Confirmation by replication in independent cohorts, functional follow-up analyses, or the use of a stronger CRP GRS_{GWAS} in upcoming studies is required to draw a definitive conclusion.

CRP GRS_{GWAS} Association with Blood Pressure and Hypertension

We found nominally significant evidence for an up to ~0.70-mm Hg increase in blood pressure with a 10-s% increase in CRP level and no evidence of heterogeneity for SBP. Additionally, there was nominally borderline significance for a causal association between CRP and DBP after adjustment for heterogeneity. These nominally significant findings, on the one hand, are

in line with numerous epidemiological studies that have highlighted an association between elevated CRP and an increased risk of hypertension. For instance, one study found an association between CRP loci and hypertension in Asian individuals [87]. An additional line of support for a possible causal association of CRP and blood pressure comes from an experimental study in which an increase in *CRP* gene expression in mice, and subsequently CRP protein levels, led to a rise in SBP particularly [88]. Moreover, an ex vivo study by Zhou et al. showed that combining IL6 treatment and mechanical strain leads to a consistent increase in CRP expression at the protein and mRNA levels in smooth muscle cells [89]. Both inflammatory factors and local mechanical strains are abundant in blood vessels and are well-known risk factors for high blood pressure. Our finding did not reach a statistically significant level after correction for multiple testing; thus, it may echo previous MR studies that have failed to find a causal relationship between CRP level and blood pressure or hypertension in Europeans [90,91]. However, our systematic literature review showed that previous studies had some limitations (S1 Table). For instance, no study used a refined GWAS set of 18 CRP-associated SNPs; instead, they tested single or a limited set of CRP SNPs. Using such instruments might have led to biased estimates as their corresponding effects on CRP levels have been found to be small [30,57]. A combination of weak instruments and small sample sizes might have led to type II error [28,57] and hence to a conclusion of no causal association between CRP and blood pressure traits in previous studies. When all of the evidence is taken together, a direct link between CRP and blood pressure remains to be elucidated, though our nominal associations between GRS_{CRP} and GRS_{GWAS} and blood pressure do add to a line of findings from experimental studies suggesting a potential causal relationship between CRP and blood pressure.

CRP GRS_{GWAS} Association with Osteoarthritis

Our nominally significant finding that CRP might be a potential causal factor for knee osteoarthritis (using GRS_{GWAS}) should be interpreted with caution. In line with our findings, we have previously shown that levels of CRP were higher in women with early radiological knee osteoarthritis (i.e., Kellgren-Lawrence grade 2+) and in women whose disease progressed [92]. Additionally, another study showed that genetically elevated CRP levels contribute to osteoarthritis severity [93]. However, other studies have found contrasting results [71,72,94]. One systematic review provided evidence that the relationship between CRP and osteoarthritis does exist but is dependent on body mass index [95]. It remains to be further investigated whether weight gain over the lifetime mediates the potential causal association between genetically elevated CRP and knee osteoarthritis.

CRP GRS_{GWAS} Shows No Association with Other Remaining Outcomes

The present study was able to calculate nominal causal estimates for IBD, Crohn disease, psoriatic arthritis, CAD, $eGFR_{cr}$, serum albumin level, and serum protein level using CRP GRS_{GWAS} , but the estimates were altered by removal of SNPs from GRS_{GWAS} based on heterogeneity tests, resulting in nominal or nonsignificant associations. These outcomes appeared therefore to have heterogeneity in the causal estimates, suggesting that these observed estimates were biased, likely due to pleiotropic effects of CRP loci. These results corroborate negative findings of previous studies (S1 Table), suggesting that a causal role of CRP in these traits and diseases is unlikely.

Methodological Concerns and Advantages

Pleiotropic biases in Mendelian randomization analyses using CRP GRS_{GWAS} . A detailed evaluation of pleiotropic SNPs in our study showed that the method applied to identify

heterogeneity sources was able to indicate and exclude several already known pleiotropic loci from the GRS_{GWAS} IV. For instance, the use of a SNP in *IL6R* (rs4129267), amongst others, resulted in heterogeneity of effects on CAD risk. The same variant contributed to heterogeneity of effects for Crohn disease in our study, and it has been shown that this SNP is associated with levels of biomarkers other than CRP [56]. Further, a MR study found that *IL6R* SNPs, specifically the nonsynonymous SNP rs8192284, are associated with CAD risk and CRP levels [96]. Our selected *IL6R* SNPs, namely rs4537545 and rs4129267, are in extremely high linkage disequilibrium with rs8192284 ($r^2 \geq 0.96$ for both SNPs in HapMap data, CEU population). Carriers of the risk allele of rs8192284 have higher CRP, IL6, and fibrinogen levels [96]. Fibrinogen is also a well-known risk factor for CAD. Therefore, it is unclear so far which biomarker(s) mediates the effect of *IL6R* SNPs on CAD. Besides the *IL6* locus, *APOC1* and *PABPC4* have been indicated as pleiotropic in three out of 32 our investigated outcomes, and *PTPN2* and *GCKR* in six. With this information taken together, we were able to disentangle at least part of the pleiotropy regarding the causal estimates of CRP for outcomes. Again, we found no significant association of CRP GRS_{GWAS} with IBD, Crohn disease, psoriatic arthritis, CAD, $eGFR_{cr}$, serum albumin level, and serum protein level after adjustment for heterogeneity.

Using summary statistics of large-scale consortia. It is of utmost interest whether the observed effect of CRP as a risk predictor for human disease is causal, and thus whether reduction of CRP levels will lower the risk of disease. Here, we investigated the causality of CRP in 32 phenotypes by leveraging very large sample sizes collected by GWAS consortia, an approach that was much better powered than most previous MR studies. We found that genetically elevated CRP levels approximated by powerful instruments did not appear to contribute directly to most of the studied somatic and psychiatric outcomes. Our findings are consistent with previous MR studies reporting null associations of genetically elevated CRP levels with inflammation-related outcomes including CAD [56,59,97], type 2 diabetes [98], high body mass index [99], Alzheimer disease, and depression [100]. All previous MR studies were substantially limited to a single or a few outcomes, used only SNPs in the *CRP* gene, or had sample sizes much smaller than that of the present study (S1 Table). In addition to these studies, the current GWAS data do not corroborate epidemiological observations suggesting that elevated CRP levels are associated with amyotrophic lateral sclerosis [101], Alzheimer disease [102], Parkinson disease [103], and major depressive disorder [104]. Furthermore, patients with immunity-related disorders frequently have a very high CRP level (as high as 100 mg/l) due to their disease status. Our findings may therefore more favorably indicate reverse causality. Taken together, these results show that CRP is highly unlikely to contribute causally to most of the major common somatic and neuropsychiatric outcomes that were investigated in the present study, with the possible exception of schizophrenia.

Strength of instrumental variables. The results presented in Table 2 show that our GRS_{CRP} is not a weak instrument, as indicated by its high *F*-values owing to the large sample sizes of available outcomes from GWASs for the phenotypes under study. The strength of our instrument increased considerably in all disease classes when we used variants of multiple loci associated with CRP in GWASs. However, the variants comprising the CRP GRS_{GWAS} explain on average only a moderate ~5% of the total variance in baseline CRP levels [30]. Moreover, the possibility of effect modification by nongenetic CRP-related factors on the outcomes remains to be investigated. We may be able to create even stronger instruments based on ongoing efforts to identify additional variation influencing CRP levels. Even if larger sample sizes and stronger instruments can be realized, the overwhelming lack of causal effects observed for most outcomes in our study implies that therapies targeted at lowering CRP will not directly result in decreased risk of the investigated outcomes, or in better symptom management [105,106].

Using summary statistics instead of individual-level data. Here we used summary association statistics obtained from previously conducted meta-GWASs in order to maximize our study power. One may argue this may induce bias compared to when one uses individual-level data. Nevertheless, previous studies showed high agreement in results from MR methods using GWAS summary data and individual-level data [60,107]; Furthermore, our analyses of individual-level data for schizophrenia led to the same conclusion as our analyses using summary statistics data, confirming the robustness of our methodological approach.

Other potential sources of bias. An important rationale for MR is that the gene variants do not change over time and are inherited randomly. Thus, the genetic variants are considered free from confounding and reverse causation [108]. However, one cannot completely control for the possibility of confounding of genotype–intermediate phenotype–disease associations. For instance, there could be a confounding effect by ethnic/racial group (i.e., population stratification), but this is unlikely to be a major problem in most situations [108]. In the present study, we included summary statistics data from highly credible results of meta-GWASs. All the original meta-GWASs corrected for population stratification in cohort-level analyses and at meta-GWAS level.

Another caveat of MR is that developmental compensation might occur, through a genotype being expressed during fetal development that in turn buffers the effects of either environmental or genetic factors, a process called canalization [108,109]. Therefore, buffering mechanisms could hamper the associations between genetic variants and the outcome of interest. As opposed to this, a lifetime exposure to a risk factor may enhance its effects on the disease [109]. However, it is not clear to what extent genetically determined small changes in any given exposure would be sufficient to induce compensation [108].

All 32 of the meta-GWASs from which instrument summary estimates were taken were performed in individuals of European descent in Europe and the US and included thousands of samples for each outcome (S1 Table), which was also the case for our previous CRP meta-GWAS from which we chose the CRP-associated SNPs to calculate GRS_{GWAS} . Therefore, the results of this MR study are applicable to individuals of European descent and are not necessarily generalizable to other ethnic groups.

Conclusion

We showed that elevated CRP levels driven by genetic factors are causally associated with protection against schizophrenia, suggesting that CRP may be one important puzzle piece that leads to an improved understanding of the pathogenesis of schizophrenia. We observed nominal evidence that genetically elevated CRP is causally associated with SBP, DBP, knee osteoarthritis, and bipolar disorder. Based on current GWAS data, we cannot verify any causal effect of CRP on the other 27 common somatic and neuropsychiatric outcomes investigated in the present study. Therefore, disease-associated rise in CRP levels may be a response to the disease process rather than a cause for these 27 outcomes. This implies that interventions to lower CRP levels are unlikely to result in decreased risk for the majority of common complex outcomes.

Supporting Information

S1 Consortia. Consortia coauthors and collaborators.

(DOCX)

S1 Data. Individual association summary statistics of CRP lead SNPs and/or proxies with traits and diseases.

(XLSX)

S1 Fig. GRS_{CRP} and GRS_{GWAS} for each studied outcome.
(DOCX)

S1 Financial Disclosure. Authors' funding information.
(PDF)

S1 Methods. Linkage disequilibrium of the four GRS_{CRP} SNPs.
(DOCX)

S2 Methods. CRP GRS_{GWAS} for Alzheimer disease and body mass index.
(DOCX)

S3 Methods. Web links.
(DOCX)

S4 Methods. CRP polygenic risk score (CRP_{PRS}) for schizophrenia.
(DOCX)

S5 Methods. In silico (gene) pathway analyses highlight the role of interferon in the causal pathway between CRP and schizophrenia.
(DOCX)

S1 Table. Previous Mendelian randomization analyses using CRP variants as instruments.
(XLSX)

S2 Table. CRP lead variants used in the genetic risk scores as instrumental variables.
(XLSX)

S3 Table. Proxy SNPs of CRP lead variants used in the genetic risk scores as instrumental variables.
(XLSX)

S4 Table. Biologically prioritized candidate gene set associated with CRP used as the input query to the gene set enrichment analysis.
(XLSX)

S5 Table. Pathway enrichment results for the biologically prioritized candidate gene set associated with CRP used as the input query to the gene set enrichment analysis.
(XLSX)

S6 Table. Genes ($n = 144$) that were significantly differentially expressed between schizophrenia and unaffected controls in the hippocampus.
(XLSX)

S7 Table. Pathway enrichment results for 144 genes that were significantly differentially expressed between schizophrenia and unaffected controls in the hippocampus.
(XLSX)

S8 Table. Biologically prioritized candidate gene set of CRP from Vaez et al. [81] (from [S5 Table](#), "CRP genes," in blue) and differentially expressed genes in schizophrenia cases versus controls from Hwang et al. [110] (from [S7 Table](#), "SCZ expr genes," in red) used as the input query to the pathway analysis.
(XLSX)

S9 Table. Pathway enrichment results for the combined set of the biologically prioritized candidate gene set of CRP from Vaez et al. [81] and differentially expressed genes in

schizophrenia cases versus controls from Hwang et al. 2013.
(XLSX)

S10 Table. List of genes at 108 genome-wide significant loci associated with schizophrenia.
(XLSX)

S11 Table. Brain and blood eQTL for credible sets of SNPs of the 108 schizophrenia loci.
(XLSX)

S12 Table. List of genes at 108 genome-wide significant loci associated with schizophrenia (yellow), brain eQTL (red), and blood eQTL (blue).
(XLSX)

S13 Table. Pathway enrichment results for the list of genes at 108 genome-wide significant loci associated with schizophrenia (yellow) and those associated with brain eQTL (red) and blood eQTL (blue).
(XLSX)

Acknowledgments

We thank for Dan Arking for providing the genome-wide summary statistics data on CRP polymorphisms for the autism GWAS. We would also like to acknowledge the rheumatoid arthritis and systemic lupus erythematosus consortia, the Genetic Investigation of Anthropometric Traits (GIANT) consortium, and the PGC for making publicly available the data from which we extracted summary association statistics for CRP SNPs. The PAGE Consortium is grateful to Alex Tsoi for adding the consortium information and membership list.

The Type 1 Diabetes Genetics Consortium would like to thank all participants and their families, and also thank Smari Kristinsson, Larus Arni Hermannsson, and Asbjörn Krisbjörnsson of Raförninn for their extensive software design and contribution. We would like to thank the people with motor neuron disease and their families for their participation in this project. We acknowledge sample management undertaken by Biobanking Solutions funded by the Medical Research Council at the Centre for Integrated Genomic Medical Research, University of Manchester.

Consortia Coauthors

Please refer to [S1 Consortia](#) for the full names and affiliations of coauthors from the following five consortia: the Schizophrenia Working Group of the Psychiatric Genomics Consortium, the Autism Spectrum Disorder Working Group of the Psychiatric Genomics Consortium, the GERAD1 Consortium, the International Consortium for Blood Pressure Genome-Wide Association Studies, and the PAGE Consortium.

Author Contributions

Conceived and designed the experiments: BPP AA HS BZA CVD. Performed the experiments: NF JGA VM PES JPB AMV JB AS. Analyzed the data: BPP AA AV AM IN NF JGA VM PES JPB AMV JB AS AD BH SR SB MDM LT EE RPN SG CP TR DVH MD JN NW DA AAC BK HH HM JE TS CVD JM APM RW JP CW PM JGP YJ HS BZA JK. Contributed reagents/materials/analysis tools: CL AAC AD JK CVD. Wrote the first draft of the manuscript: BPP AA BZA. Contributed to the writing of the manuscript: BPP AA AV AW JGA VM PES AMV DEA BK TS JM RW JP JGP YJ HS BZA JK. Agree with the manuscript's results and conclusions: BZA BPP AA AW AV IN NF PES JGA VM JPB AMV JB AS CL BH SR SB MDM LT EE RPN

SG CP TR DVH MD NW AAC AD HH HM JE JK DA TS BK CVD JM APM RW CW PM JP JGP YJ HS. All authors have read, and confirm that they meet, ICMJE criteria for authorship.

References

1. Stockinger B. Immunology: cause of death matters. *Nature*. 2009; 458:44–45. doi: [10.1038/458044a](https://doi.org/10.1038/458044a) PMID: [19262664](https://pubmed.ncbi.nlm.nih.gov/19262664/)
2. Baumgart DC, Carding SR. Inflammatory bowel disease: cause and immunobiology. *Lancet*. 2007; 369:1627–1640. doi: [10.1016/S0140-6736\(07\)60750-8](https://doi.org/10.1016/S0140-6736(07)60750-8) PMID: [17499605](https://pubmed.ncbi.nlm.nih.gov/17499605/)
3. Bruunsgaard H, Pedersen M, Pedersen BK. Aging and proinflammatory cytokines. *Curr Opin Hematol*. 2001; 8:131–136. PMID: [11303144](https://pubmed.ncbi.nlm.nih.gov/11303144/)
4. Cesari M, Penninx BWJH, Newman AB, Kritchevsky SB, Nicklas BJ, Sutton-Tyrrell K, et al. Inflammatory markers and cardiovascular disease (The Health, Aging and Body Composition [Health ABC] Study). *Am J Cardiol*. 2003; 92:522–528. PMID: [12943870](https://pubmed.ncbi.nlm.nih.gov/12943870/)
5. Couzin-Frankel J. Inflammation bares a dark side. *Science*. 2010; 330:1621. doi: [10.1126/science.330.6011.1621](https://doi.org/10.1126/science.330.6011.1621) PMID: [21163993](https://pubmed.ncbi.nlm.nih.gov/21163993/)
6. Fan X, Goff DC, Henderson DC. Inflammation and schizophrenia. *Expert Rev Neurother*. 2007; 7:789–796. doi: [10.1586/14737175.7.7.789](https://doi.org/10.1586/14737175.7.7.789) PMID: [17610386](https://pubmed.ncbi.nlm.nih.gov/17610386/)
7. Grimaldi MP, Vasto S, Balistreri CR, di Carlo D, Caruso M, Incalcaterra E, et al. Genetics of inflammation in age-related atherosclerosis: its relevance to pharmacogenomics. *Ann N Y Acad Sci*. 2007; 1100:123–131. doi: [10.1196/annals.1395.010](https://doi.org/10.1196/annals.1395.010) PMID: [17460170](https://pubmed.ncbi.nlm.nih.gov/17460170/)
8. Hanson DR, Gottesman II. Theories of schizophrenia: a genetic-inflammatory-vascular synthesis. *BMC Med Genet*. 2005; 6:7. doi: [10.1186/1471-2350-6-7](https://doi.org/10.1186/1471-2350-6-7) PMID: [15707482](https://pubmed.ncbi.nlm.nih.gov/15707482/)
9. Johnson TE. Recent results: biomarkers of aging. *Exp Gerontol*. 2006; 41:1243–1246. doi: [10.1016/j.exger.2006.09.006](https://doi.org/10.1016/j.exger.2006.09.006) PMID: [17071038](https://pubmed.ncbi.nlm.nih.gov/17071038/)
10. Kiecolt-Glaser JK, Glaser R. Depression and immune function: central pathways to morbidity and mortality. *J Psychosom Res*. 2002; 53:873–876. PMID: [12377296](https://pubmed.ncbi.nlm.nih.gov/12377296/)
11. Kiecolt-Glaser JK, McGuire L, Robles TF, Glaser R. Emotions, morbidity, and mortality: new perspectives from psychoneuroimmunology. *Annu Rev Psychol*. 2002; 53:83–107. doi: [10.1146/annurev.psych.53.100901.135217](https://doi.org/10.1146/annurev.psych.53.100901.135217) PMID: [11752480](https://pubmed.ncbi.nlm.nih.gov/11752480/)
12. Lynch MA, Mills KHG. Immunology meets neuroscience—opportunities for immune intervention in neurodegenerative diseases. *Brain Behav Immun*. 2012; 26:1–10. doi: [10.1016/j.bbi.2011.05.013](https://doi.org/10.1016/j.bbi.2011.05.013) PMID: [21664452](https://pubmed.ncbi.nlm.nih.gov/21664452/)
13. Meyer U, Schwarz MJ, Müller N. Inflammatory processes in schizophrenia: a promising neuroimmunological target for the treatment of negative/cognitive symptoms and beyond. *Pharmacol Ther*. 2011; 132:96–110. doi: [10.1016/j.pharmthera.2011.06.003](https://doi.org/10.1016/j.pharmthera.2011.06.003) PMID: [21704074](https://pubmed.ncbi.nlm.nih.gov/21704074/)
14. Pardo CA, Vargas DL, Zimmerman AW. Immunity, neuroglia and neuroinflammation in autism. *Int Rev Psychiatry*. 2005; 17:485–495. doi: [10.1080/02646830500381930](https://doi.org/10.1080/02646830500381930) PMID: [16401547](https://pubmed.ncbi.nlm.nih.gov/16401547/)
15. Sansoni P, Vescovini R, Fagnoni F, Biasini C, Zanni F, Zanlari L, et al. The immune system in extreme longevity. *Exp Gerontol*. 2008; 43:61–65. doi: [10.1016/j.exger.2007.06.008](https://doi.org/10.1016/j.exger.2007.06.008) PMID: [17870272](https://pubmed.ncbi.nlm.nih.gov/17870272/)
16. Hotamisligil GS. Inflammation and metabolic disorders. *Nature*. 2006; 444:860–867. doi: [10.1038/nature05485](https://doi.org/10.1038/nature05485) PMID: [17167474](https://pubmed.ncbi.nlm.nih.gov/17167474/)
17. Mantovani A, Allavena P, Sica A, Balkwill F. Cancer-related inflammation. *Nature*. 2008; 454:436–444. doi: [10.1038/nature07205](https://doi.org/10.1038/nature07205) PMID: [18650914](https://pubmed.ncbi.nlm.nih.gov/18650914/)
18. Dantzer R, O'Connor JC, Freund GG, Johnson RW, Kelley KW. From inflammation to sickness and depression: when the immune system subjugates the brain. *Nat Rev Neurosci*. 2008; 9:46–56. doi: [10.1038/nrn2297](https://doi.org/10.1038/nrn2297) PMID: [18073775](https://pubmed.ncbi.nlm.nih.gov/18073775/)
19. Emerging Risk Factors Collaboration, Kaptoge S, Di Angelantonio E, Lowe G, Pepys MB, Thompson SG, et al. C-reactive protein concentration and risk of coronary heart disease, stroke, and mortality: an individual participant meta-analysis. *Lancet*. 2010; 375:132–140. doi: [10.1016/S0140-6736\(09\)61717-7](https://doi.org/10.1016/S0140-6736(09)61717-7) PMID: [20031199](https://pubmed.ncbi.nlm.nih.gov/20031199/)
20. Emerging Risk Factors Collaboration, Kaptoge S, Di Angelantonio E, Pennells L, Wood AM, White IR, et al. C-reactive protein, fibrinogen, and cardiovascular disease prediction. *N Engl J Med*. 2012; 367:1310–1320. doi: [10.1056/NEJMoa1107477](https://doi.org/10.1056/NEJMoa1107477) PMID: [23034020](https://pubmed.ncbi.nlm.nih.gov/23034020/)
21. Wang X, Bao W, Liu J, Ouyang Y-Y, Wang D, Rong S, et al. Inflammatory markers and risk of type 2 diabetes: a systematic review and meta-analysis. *Diabetes Care*. 2013; 36:166–175. doi: [10.2337/dc12-0702](https://doi.org/10.2337/dc12-0702) PMID: [23264288](https://pubmed.ncbi.nlm.nih.gov/23264288/)

22. Henriksen M, Jahnsen J, Lygren I, Stray N, Sauar J, Vatn MH, et al. C-reactive protein: a predictive factor and marker of inflammation in inflammatory bowel disease. Results from a prospective population-based study. *Gut*. 2008; 57:1518–1523. doi: [10.1136/gut.2007.146357](https://doi.org/10.1136/gut.2007.146357) PMID: [18566104](https://pubmed.ncbi.nlm.nih.gov/18566104/)
23. Rhodes B, Merriman ME, Harrison A, Nissen MJ, Smith M, Stamp L, et al. A genetic association study of serum acute-phase C-reactive protein levels in rheumatoid arthritis: implications for clinical interpretation. *PLoS Med*. 2010; 7:e1000341. doi: [10.1371/journal.pmed.1000341](https://doi.org/10.1371/journal.pmed.1000341) PMID: [20877716](https://pubmed.ncbi.nlm.nih.gov/20877716/)
24. Ridker PM. High-sensitivity C-reactive protein as a predictor of all-cause mortality: implications for research and patient care. *Clin Chem*. 2008; 54:234–237. doi: [10.1373/clinchem.2007.099465](https://doi.org/10.1373/clinchem.2007.099465) PMID: [18223130](https://pubmed.ncbi.nlm.nih.gov/18223130/)
25. Dehghan A, Kardys I, de Maat MPM, Uitterlinden AG, Sijbrands EJG, Bootsma AH, et al. Genetic variation, C-reactive protein levels, and incidence of diabetes. *Diabetes*. 2007; 56:872–878. doi: [10.2337/db06-0922](https://doi.org/10.2337/db06-0922) PMID: [17327459](https://pubmed.ncbi.nlm.nih.gov/17327459/)
26. Lee CC, Adler AI, Sandhu MS, Sharp SJ, Forouhi NG, Erqou S, et al. Association of C-reactive protein with type 2 diabetes: prospective analysis and meta-analysis. *Diabetologia*. 2009; 52:1040–1047. doi: [10.1007/s00125-009-1338-3](https://doi.org/10.1007/s00125-009-1338-3) PMID: [19326095](https://pubmed.ncbi.nlm.nih.gov/19326095/)
27. Danesh J, Pepys MB. C-reactive protein and coronary disease: is there a causal link? *Circulation*. 2009; 120:2036–2039. doi: [10.1161/CIRCULATIONAHA.109.907212](https://doi.org/10.1161/CIRCULATIONAHA.109.907212) PMID: [19901186](https://pubmed.ncbi.nlm.nih.gov/19901186/)
28. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med*. 2008; 27:1133–1163. doi: [10.1002/sim.3034](https://doi.org/10.1002/sim.3034) PMID: [17886233](https://pubmed.ncbi.nlm.nih.gov/17886233/)
29. C Reactive Protein Coronary Heart Disease Genetics Collaboration (CCGC), Wensley F, Gao P, Burgess S, Kaptoge S, Di Angelantonio E, et al. Association between C reactive protein and coronary heart disease: Mendelian randomisation analysis based on individual participant data. *BMJ*. 2011; 342:d548. doi: [10.1136/bmj.d548](https://doi.org/10.1136/bmj.d548) PMID: [21325005](https://pubmed.ncbi.nlm.nih.gov/21325005/)
30. Dehghan A, Dupuis J, Barbalic M, Bis JC, Eiriksdottir G, Lu C, et al. Meta-analysis of genome-wide association studies in >80 000 subjects identifies multiple loci for C-reactive protein levels. *Circulation*. 2011; 123:731–738. doi: [10.1161/CIRCULATIONAHA.110.948570](https://doi.org/10.1161/CIRCULATIONAHA.110.948570) PMID: [21300955](https://pubmed.ncbi.nlm.nih.gov/21300955/)
31. Dubois PCA, Trynka G, Franke L, Hunt KA, Romanos J, Curtotti A, et al. Multiple common variants for celiac disease influencing immune gene expression. *Nat Genet*. 2010; 42:295–302. doi: [10.1038/ng.543](https://doi.org/10.1038/ng.543) PMID: [20190752](https://pubmed.ncbi.nlm.nih.gov/20190752/)
32. Franke A, McGovern DPB, Barrett JC, Wang K, Radford-Smith GL, Ahmad T, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet*. 2010; 42:1118–1125. doi: [10.1038/ng.717](https://doi.org/10.1038/ng.717) PMID: [21102463](https://pubmed.ncbi.nlm.nih.gov/21102463/)
33. Anderson CA, Boucher G, Lees CW, Franke A, D'Amato M, Taylor KD, et al. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet*. 2011; 43:246–252. doi: [10.1038/ng.764](https://doi.org/10.1038/ng.764) PMID: [21297633](https://pubmed.ncbi.nlm.nih.gov/21297633/)
34. Nair RP, Duffin KC, Helms C, Ding J, Stuart PE, Goldgar D, et al. Genome-wide scan reveals association of psoriasis with IL-23 and NF-kappaB pathways. *Nat Genet*. 2009; 41:199–204. doi: [10.1038/ng.311](https://doi.org/10.1038/ng.311) PMID: [19169254](https://pubmed.ncbi.nlm.nih.gov/19169254/)
35. Ellinghaus E, Ellinghaus D, Stuart PE, Nair RP, Debrus S, Raelson JV, et al. Genome-wide association study identifies a psoriasis susceptibility locus at TRAF3IP2. *Nat Genet*. 2010; 42:991–995. doi: [10.1038/ng.689](https://doi.org/10.1038/ng.689) PMID: [20953188](https://pubmed.ncbi.nlm.nih.gov/20953188/)
36. Stahl EA, Raychaudhuri S, Remmers EF, Xie G, Eyre S, Thomson BP, et al. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat Genet*. 2010; 42:508–514. doi: [10.1038/ng.582](https://doi.org/10.1038/ng.582) PMID: [20453842](https://pubmed.ncbi.nlm.nih.gov/20453842/)
37. Hom G, Graham RR, Modrek B, Taylor KE, Ortmann W, Garnier S, et al. Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. *N Engl J Med*. 2008; 358:900–909. doi: [10.1056/NEJMoa0707865](https://doi.org/10.1056/NEJMoa0707865) PMID: [18204098](https://pubmed.ncbi.nlm.nih.gov/18204098/)
38. Radstake TRDJ, Gorlova O, Rueda B, Martin J-E, Alizadeh BZ, Palomino-Morales R, et al. Genome-wide association study of systemic sclerosis identifies CD247 as a new susceptibility locus. *Nat Genet*. 2010; 42:426–429. doi: [10.1038/ng.565](https://doi.org/10.1038/ng.565) PMID: [20383147](https://pubmed.ncbi.nlm.nih.gov/20383147/)
39. Bradfield JP, Qu H-Q, Wang K, Zhang H, Sleiman PM, Kim CE, et al. A genome-wide meta-analysis of six type 1 diabetes cohorts identifies multiple associated loci. *PLoS Genet*. 2011; 7:e1002293. doi: [10.1371/journal.pgen.1002293](https://doi.org/10.1371/journal.pgen.1002293) PMID: [21980299](https://pubmed.ncbi.nlm.nih.gov/21980299/)
40. Ramos YFM, Metrustry S, Arden N, Bay-Jensen AC, Beekman M, de Craen AJM, et al. Meta-analysis identifies loci affecting levels of the potential osteoarthritis biomarkers sCOMP and uCTX-II with genome wide significance. *J Med Genet*. 2014; 51:596–604. doi: [10.1136/jmedgenet-2014-102478](https://doi.org/10.1136/jmedgenet-2014-102478) PMID: [25057126](https://pubmed.ncbi.nlm.nih.gov/25057126/)

41. Nikpey M, Goel A, Won H, Hall LM, Willenborg C, Kanoni S, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet.* 2015; 47:1121–1130. doi: [10.1038/ng.3396](https://doi.org/10.1038/ng.3396) PMID: [26343387](https://pubmed.ncbi.nlm.nih.gov/26343387/)
42. International Consortium for Blood Pressure Genome-Wide Association Studies, Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature.* 2011; 478:103–109. doi: [10.1038/nature10405](https://doi.org/10.1038/nature10405) PMID: [21909115](https://pubmed.ncbi.nlm.nih.gov/21909115/)
43. International Stroke Genetics Consortium (ISGC), Wellcome Trust Case Control Consortium 2 (WTCCC2), Bellenguez C, Bevan S, Gschwendtner A, Spencer CCA, et al. Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke. *Nat Genet.* 2012; 44:328–333. doi: [10.1038/ng.1081](https://doi.org/10.1038/ng.1081) PMID: [22306652](https://pubmed.ncbi.nlm.nih.gov/22306652/)
44. Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet.* 2010; 42:937–948. doi: [10.1038/ng.686](https://doi.org/10.1038/ng.686) PMID: [20935630](https://pubmed.ncbi.nlm.nih.gov/20935630/)
45. Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, Welch RP, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet.* 2010; 42:579–589. doi: [10.1038/ng.609](https://doi.org/10.1038/ng.609) PMID: [20581827](https://pubmed.ncbi.nlm.nih.gov/20581827/)
46. Köttgen A, Pattaro C, Böger CA, Fuchsberger C, Olden M, Glazer NL, et al. New loci associated with kidney function and chronic kidney disease. *Nat Genet.* 2010; 42:376–384. doi: [10.1038/ng.568](https://doi.org/10.1038/ng.568) PMID: [20383146](https://pubmed.ncbi.nlm.nih.gov/20383146/)
47. Franceschini N, van Rooij FJA, Prins BP, Feitosa MF, Karakas M, Eckfeldt JH, et al. Discovery and fine mapping of serum protein loci through transethnic meta-analysis. *Am J Hum Genet.* 2012; 91:744–753. doi: [10.1016/j.ajhg.2012.08.021](https://doi.org/10.1016/j.ajhg.2012.08.021) PMID: [23022100](https://pubmed.ncbi.nlm.nih.gov/23022100/)
48. Shatunov A, Mok K, Newhouse S, Weale ME, Smith B, Vance C, et al. Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study. *Lancet Neurol.* 2010; 9:986–994. doi: [10.1016/S1474-4422\(10\)70197-6](https://doi.org/10.1016/S1474-4422(10)70197-6) PMID: [20801717](https://pubmed.ncbi.nlm.nih.gov/20801717/)
49. Hollingworth P, Harold D, Sims R, Gerrish A, Lambert J-C, Carrasquillo MM, et al. Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. *Nat Genet.* 2011; 43:429–435. doi: [10.1038/ng.803](https://doi.org/10.1038/ng.803) PMID: [21460840](https://pubmed.ncbi.nlm.nih.gov/21460840/)
50. International Parkinson Disease Genomics Consortium, Nalls MA, Plagnol V, Hernandez DG, Sharma M, Sheerin U-M, et al. Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet.* 2011; 377:641–649. doi: [10.1016/S0140-6736\(10\)62345-8](https://doi.org/10.1016/S0140-6736(10)62345-8) PMID: [21292315](https://pubmed.ncbi.nlm.nih.gov/21292315/)
51. Weiss LA, Arking DE, Gene Discovery Project of Johns Hopkins & the Autism Consortium, Daly MJ, Chakravarti A. A genome-wide linkage and association scan reveals novel loci for autism. *Nature.* 2009; 461:802–808. doi: [10.1038/nature08490](https://doi.org/10.1038/nature08490) PMID: [19812673](https://pubmed.ncbi.nlm.nih.gov/19812673/)
52. Psychiatric GWAS Consortium Bipolar Disorder Working Group. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet.* 2011; 43:977–983. doi: [10.1038/ng.943](https://doi.org/10.1038/ng.943) PMID: [21926972](https://pubmed.ncbi.nlm.nih.gov/21926972/)
53. Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium, Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM, et al. A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry.* 2013; 18:497–511. doi: [10.1038/mp.2012.21](https://doi.org/10.1038/mp.2012.21) PMID: [22472876](https://pubmed.ncbi.nlm.nih.gov/22472876/)
54. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature.* 2014; 511:421–427. doi: [10.1038/nature13595](https://doi.org/10.1038/nature13595) PMID: [25056061](https://pubmed.ncbi.nlm.nih.gov/25056061/)
55. C Reactive Protein Coronary Heart Disease Genetics Collaboration (CCGC), Danesh J, Hingorani A, Wensley F, Casas JP, Smeeth L, Samani NJ, et al. Collaborative pooled analysis of data on C-reactive protein gene variants and coronary disease: judging causality by Mendelian randomisation. *Eur J Epidemiol.* 2008; 23:531–540. doi: [10.1007/s10654-008-9249-z](https://doi.org/10.1007/s10654-008-9249-z) PMID: [18425592](https://pubmed.ncbi.nlm.nih.gov/18425592/)
56. Elliott P, Chambers JC, Zhang W, Clarke R, Hopewell JC, Peden JF, et al. Genetic loci associated with C-reactive protein levels and risk of coronary heart disease. *JAMA.* 2009; 302:37–48. doi: [10.1001/jama.2009.954](https://doi.org/10.1001/jama.2009.954) PMID: [19567438](https://pubmed.ncbi.nlm.nih.gov/19567438/)
57. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol.* 2013; 37:658–665. doi: [10.1002/gepi.21758](https://doi.org/10.1002/gepi.21758) PMID: [24114802](https://pubmed.ncbi.nlm.nih.gov/24114802/)
58. Dastani Z, Hivert M-F, Timpson N, Perry JRB, Yuan X, Scott RA, et al. Novel loci for adiponectin levels and their influence on type 2 diabetes and metabolic traits: a multi-ethnic meta-analysis of 45,891 individuals. *PLoS Genet.* 2012; 8:e1002607. doi: [10.1371/journal.pgen.1002607](https://doi.org/10.1371/journal.pgen.1002607) PMID: [22479202](https://pubmed.ncbi.nlm.nih.gov/22479202/)

59. Zacho J, Tybjaerg-Hansen A, Jensen JS, Grande P, Silleesen H, Nordestgaard BG. Genetically elevated C-reactive protein and ischemic vascular disease. *N Engl J Med*. 2008; 359:1897–1908. doi: [10.1056/NEJMoa0707402](https://doi.org/10.1056/NEJMoa0707402) PMID: [18971492](https://pubmed.ncbi.nlm.nih.gov/18971492/)
60. Voight BF, Peloso GM, Orho-Melander M, Frikke-Schmidt R, Barbalic M, Jensen MK, et al. Plasma HDL cholesterol and risk of myocardial infarction: a Mendelian randomisation study. *Lancet*. 2012; 380:572–580. doi: [10.1016/S0140-6736\(12\)60312-2](https://doi.org/10.1016/S0140-6736(12)60312-2) PMID: [22607825](https://pubmed.ncbi.nlm.nih.gov/22607825/)
61. Cole TJ. Sympercents: symmetric percentage differences on the 100 log(e) scale simplify the presentation of log transformed data. *Stat Med*. 2000; 19:3109–3125. PMID: [11113946](https://pubmed.ncbi.nlm.nih.gov/11113946/)
62. Rice JA. *Mathematical statistics and data analysis*. 2nd ed. Belmont (California): Duxbury Press; 1995.
63. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007; 81:559–575. PMID: [17701901](https://pubmed.ncbi.nlm.nih.gov/17701901/)
64. Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet*. 2013; 381:1371–1379. doi: [10.1016/S0140-6736\(12\)62129-1](https://doi.org/10.1016/S0140-6736(12)62129-1) PMID: [23453885](https://pubmed.ncbi.nlm.nih.gov/23453885/)
65. Power RA, Steinberg S, Bjornsdottir G, Rietveld CA, Abdellaoui A, Nivard MM, et al. Polygenic risk scores for schizophrenia and bipolar disorder predict creativity. *Nat Neurosci*. 2015; 18:953–955. doi: [10.1038/nn.4040](https://doi.org/10.1038/nn.4040) PMID: [26053403](https://pubmed.ncbi.nlm.nih.gov/26053403/)
66. Zakharyan R, Chavushyan A, Khoyetsyan A, Stahelova A, Arakelyan A, Boyajyan A, et al. Genetic variants of the inflammatory C-reactive protein and schizophrenia in Armenian population: a pilot study. *Int J Immunogenet*. 2010; 37:407–410. doi: [10.1111/j.1744-313X.2010.00942.x](https://doi.org/10.1111/j.1744-313X.2010.00942.x) PMID: [21182750](https://pubmed.ncbi.nlm.nih.gov/21182750/)
67. Singh B, Chaudhuri TK. Role of C-reactive protein in schizophrenia: an overview. *Psychiatry Res*. 2014; 216:277–285. doi: [10.1016/j.psychres.2014.02.004](https://doi.org/10.1016/j.psychres.2014.02.004) PMID: [24565000](https://pubmed.ncbi.nlm.nih.gov/24565000/)
68. Wiium-Andersen MK, Ørsted DD, Nordestgaard BG. Elevated C-reactive protein associated with late- and very-late-onset schizophrenia in the general population: a prospective study. *Schizophr Bull*. 2014; 40:1117–1127. doi: [10.1093/schbul/sbt120](https://doi.org/10.1093/schbul/sbt120) PMID: [23996346](https://pubmed.ncbi.nlm.nih.gov/23996346/)
69. Miller BJ, Culppepper N, Rapaport MH. C-reactive protein levels in schizophrenia: a review and meta-analysis. *Clin Schizophr Relat Psychoses*. 2014; 7:223–230. doi: [10.3371/CSRP.MICU.020813](https://doi.org/10.3371/CSRP.MICU.020813) PMID: [23428789](https://pubmed.ncbi.nlm.nih.gov/23428789/)
70. Dickerson F, Stallings C, Origoni A, Vaughan C, Khushalani S, Yang S, et al. C-reactive protein is elevated in schizophrenia. *Schizophr Res*. 2013; 143:198–202. doi: [10.1016/j.schres.2012.10.041](https://doi.org/10.1016/j.schres.2012.10.041) PMID: [23218564](https://pubmed.ncbi.nlm.nih.gov/23218564/)
71. Fernandes BS, Steiner J, Bernstein H-G, Dodd S, Pasco JA, Dean OM, et al. C-reactive protein is increased in schizophrenia but is not altered by antipsychotics: meta-analysis and implications. *Mol Psychiatry*. 2016; 21:554–564. doi: [10.1038/mp.2015.87](https://doi.org/10.1038/mp.2015.87) PMID: [26169974](https://pubmed.ncbi.nlm.nih.gov/26169974/)
72. Lewis CM, Levinson DF, Wise LH, DeLisi LE, Straub RE, Hovatta I, et al. Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: schizophrenia. *Am J Hum Genet*. 2003; 73:34–48. doi: [10.1086/376549](https://doi.org/10.1086/376549) PMID: [12802786](https://pubmed.ncbi.nlm.nih.gov/12802786/)
73. International Schizophrenia Consortium, Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*. 2009; 460:748–752. doi: [10.1038/nature08185](https://doi.org/10.1038/nature08185) PMID: [19571811](https://pubmed.ncbi.nlm.nih.gov/19571811/)
74. Shi J, Levinson DF, Duan J, Sanders AR, Zheng Y, Pe'er I, et al. Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature*. 2009; 460:753–757. doi: [10.1038/nature08192](https://doi.org/10.1038/nature08192) PMID: [19571809](https://pubmed.ncbi.nlm.nih.gov/19571809/)
75. Stefansson H, Ophoff RA, Steinberg S, Andreassen OA, Cichon S, Rujescu D, et al. Common variants conferring risk of schizophrenia. *Nature*. 2009; 460:744–747. doi: [10.1038/nature08186](https://doi.org/10.1038/nature08186) PMID: [19571808](https://pubmed.ncbi.nlm.nih.gov/19571808/)
76. Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. Genome-wide association study identifies five new schizophrenia loci. *Nat Genet*. 2011; 43:969–976. doi: [10.1038/ng.940](https://doi.org/10.1038/ng.940) PMID: [21926974](https://pubmed.ncbi.nlm.nih.gov/21926974/)
77. Al-Asmari SM, Kadasah S, Arfin M, Tariq M, Al-Asmari A. Genetic variants of interleukin-10 gene promoter are associated with schizophrenia in Saudi patients: A case-control study. *N Am J Med Sci*. 2014; 6:558–565. doi: [10.4103/1947-2714.145466](https://doi.org/10.4103/1947-2714.145466) PMID: [25535603](https://pubmed.ncbi.nlm.nih.gov/25535603/)
78. Saviouk V, Chow EWC, Bassett AS, Brzustowicz LM. Tumor necrosis factor promoter haplotype associated with schizophrenia reveals a linked locus on 1q44. *Mol Psychiatry*. 2005; 10:375–383. doi: [10.1038/sj.mp.4001582](https://doi.org/10.1038/sj.mp.4001582) PMID: [15340354](https://pubmed.ncbi.nlm.nih.gov/15340354/)

79. Hänninen K, Katila H, Saarela M, Rontu R, Mattila KM, Fan M, et al. Interleukin-1 beta gene polymorphism and its interactions with neuregulin-1 gene polymorphism are associated with schizophrenia. *Eur Arch Psychiatry Clin Neurosci*. 2008; 258:10–15. doi: [10.1007/s00406-007-0756-9](https://doi.org/10.1007/s00406-007-0756-9) PMID: [17901998](https://pubmed.ncbi.nlm.nih.gov/17901998/)
80. Sekar A. A natural allelic series of complex structural variants and its influence on the risk of lupus and schizophrenia. 2014 [cited 19 May 2016]. Digital Access to Scholarship at Harvard. Available: <http://dash.harvard.edu/handle/1/13070061>.
81. Vaez A, Jansen R, Prins BP, Hottenga J-J, de Geus EJC, Boomsma DI, et al. In silico post genome-wide association studies analysis of C-reactive protein loci suggests an important role for interferons. *Circ Cardiovasc Genet*. 2015; 8:487–497. doi: [10.1161/CIRCGENETICS.114.000714](https://doi.org/10.1161/CIRCGENETICS.114.000714) PMID: [25752597](https://pubmed.ncbi.nlm.nih.gov/25752597/)
82. Shaked I, Tchoresh D, Gersner R, Meiri G, Mordechai S, Xiao X, et al. Protective autoimmunity: interferon-gamma enables microglia to remove glutamate without evoking inflammatory mediators. *J Neurochem*. 2005; 92:997–1009. doi: [10.1111/j.1471-4159.2004.02954.x](https://doi.org/10.1111/j.1471-4159.2004.02954.x) PMID: [15715651](https://pubmed.ncbi.nlm.nih.gov/15715651/)
83. Garg SK, Banerjee R, Kipnis J. Neuroprotective immunity: T cell-derived glutamate endows astrocytes with a neuroprotective phenotype. *J Immunol*. 2008; 180:3866–3873. PMID: [18322194](https://pubmed.ncbi.nlm.nih.gov/18322194/)
84. Javitt DC. Glutamatergic theories of schizophrenia. *Isr J Psychiatry Relat Sci*. 2010; 47:4–16. PMID: [20686195](https://pubmed.ncbi.nlm.nih.gov/20686195/)
85. Marsman A, van den Heuvel MP, Klomp DWJ, Kahn RS, Luijten PR, Hulshoff Pol HE. Glutamate in schizophrenia: a focused review and meta-analysis of 1H-MRS studies. *Schizophr Bull*. 2013; 39:120–129. doi: [10.1093/schbul/sbr069](https://doi.org/10.1093/schbul/sbr069) PMID: [21746807](https://pubmed.ncbi.nlm.nih.gov/21746807/)
86. Chung K-H, Huang S-H, Wu J-Y, Chen P-H, Hsu J-L, Tsai S-Y. The link between high-sensitivity C-reactive protein and orbitofrontal cortex in euthymic bipolar disorder. *Neuropsychobiology*. 2013; 68:168–173. doi: [10.1159/000353613](https://doi.org/10.1159/000353613) PMID: [24051690](https://pubmed.ncbi.nlm.nih.gov/24051690/)
87. Hong EP, Kim DH, Suh JG, Park JW. Genetic risk assessment for cardiovascular disease with seven genes associated with plasma C-reactive protein concentrations in Asian populations. *Hypertens Res*. 2014; 37:692–698. doi: [10.1038/hr.2014.56](https://doi.org/10.1038/hr.2014.56) PMID: [24671014](https://pubmed.ncbi.nlm.nih.gov/24671014/)
88. Vongpatanasin W, Thomas GD, Schwartz R, Cassis LA, Osborne-Lawrence S, Hahner L, et al. C-reactive protein causes downregulation of vascular angiotensin subtype 2 receptors and systolic hypertension in mice. *Circulation*. 2007; 115:1020–1028. doi: [10.1161/CIRCULATIONAHA.106.664854](https://doi.org/10.1161/CIRCULATIONAHA.106.664854) PMID: [17283257](https://pubmed.ncbi.nlm.nih.gov/17283257/)
89. Zhou H, Li Y, Huang G, Gu X, Zeng J, Li Y, et al. Interleukin 6 augments mechanical strain-induced C-reactive protein synthesis via the stretch-activated channel-nuclear factor κ B signal pathway. *Heart*. 2013; 99:570–576. doi: [10.1136/heartjnl-2012-303355](https://doi.org/10.1136/heartjnl-2012-303355) PMID: [23257175](https://pubmed.ncbi.nlm.nih.gov/23257175/)
90. Timpson NJ, Lawlor DA, Harbord RM, Gaunt TR, Day INM, Palmer LJ, et al. C-reactive protein and its role in metabolic syndrome: Mendelian randomisation study. *Lancet*. 2005; 366:1954–1959. doi: [10.1016/S0140-6736\(05\)67786-0](https://doi.org/10.1016/S0140-6736(05)67786-0) PMID: [16325697](https://pubmed.ncbi.nlm.nih.gov/16325697/)
91. Davey Smith G, Lawlor DA, Harbord R, Timpson N, Rumley A, Lowe GDO, et al. Association of C-reactive protein with blood pressure and hypertension: life course confounding and Mendelian randomization tests of causality. *Arterioscler Thromb Vasc Biol*. 2005; 25:1051–1056. doi: [10.1161/01.ATV.0000160351.95181.d0](https://doi.org/10.1161/01.ATV.0000160351.95181.d0) PMID: [15731495](https://pubmed.ncbi.nlm.nih.gov/15731495/)
92. Spector TD, Hart DJ, Nandra D, Doyle DV, Mackillop N, Gallimore JR, et al. Low-level increases in serum C-reactive protein are present in early osteoarthritis of the knee and predict progressive disease. *Arthritis Rheum*. 1997; 40:723–727. PMID: [9125256](https://pubmed.ncbi.nlm.nih.gov/9125256/)
93. Bos SD, Suchiman HED, Kloppenburg M, Houwing-Duistermaat JJ, le Graverand MPH, Seymour AB, et al. Allelic variation at the C-reactive protein gene associates to both hand osteoarthritis severity and serum high sensitive C-reactive protein levels in the GARP study. *Ann Rheum Dis*. 2008; 67:877–879. doi: [10.1136/ard.2007.079228](https://doi.org/10.1136/ard.2007.079228) PMID: [18055473](https://pubmed.ncbi.nlm.nih.gov/18055473/)
94. Vlad SC, Neogi T, Aliabadi P, Fontes JDT, Felson DT. No association between markers of inflammation and osteoarthritis of the hands and knees. *J Rheumatol*. 2011; 38:1665–1670. doi: [10.3899/jrheum.100971](https://doi.org/10.3899/jrheum.100971) PMID: [21572158](https://pubmed.ncbi.nlm.nih.gov/21572158/)
95. Kerkhof HJM, Bierma-Zeinstra SMA, Castano-Betancourt MC, de Maat MP, Hofman A, Pols HAP, et al. Serum C reactive protein levels and genetic variation in the CRP gene are not associated with the prevalence, incidence or progression of osteoarthritis independent of body mass index. *Ann Rheum Dis*. 2010; 69:1976–1982. doi: [10.1136/ard.2009.125260](https://doi.org/10.1136/ard.2009.125260) PMID: [20511616](https://pubmed.ncbi.nlm.nih.gov/20511616/)
96. Interleukin-6 Receptor Mendelian Randomisation Analysis (IL6R MR) Consortium. The interleukin-6 receptor as a target for prevention of coronary heart disease: a Mendelian randomisation analysis. *Lancet*. 2012; 379:1214–1224. doi: [10.1016/S0140-6736\(12\)60110-X](https://doi.org/10.1016/S0140-6736(12)60110-X) PMID: [22421340](https://pubmed.ncbi.nlm.nih.gov/22421340/)
97. Anand SS, Yusuf S. C-reactive protein is a bystander of cardiovascular disease. *Eur Heart J*. 2010; 31:2092–2096. doi: [10.1093/eurheartj/ehq242](https://doi.org/10.1093/eurheartj/ehq242) PMID: [20675658](https://pubmed.ncbi.nlm.nih.gov/20675658/)

98. Brunner EJ, Kivimäki M, Witte DR, Lawlor DA, Davey Smith G, Cooper JA, et al. Inflammation, insulin resistance, and diabetes—Mendelian randomization using CRP haplotypes points upstream. *PLoS Med.* 2008; 5:e155. doi: [10.1371/journal.pmed.0050155](https://doi.org/10.1371/journal.pmed.0050155) PMID: [18700811](https://pubmed.ncbi.nlm.nih.gov/18700811/)
99. Timpson NJ, Nordestgaard BG, Harbord RM, Zacho J, Frayling TM, Tybjaerg-Hansen A, et al. C-reactive protein levels and body mass index: elucidating direction of causation through reciprocal Mendelian randomization. *Int J Obes* 2005. 2011; 35:300–308. doi: [10.1038/ijo.2010.137](https://doi.org/10.1038/ijo.2010.137)
100. Wium-Andersen MK, Orsted DD, Nordestgaard BG. Elevated C-reactive protein, depression, somatic diseases, and all-cause mortality: a Mendelian randomization study. *Biol Psychiatry.* 2014; 76:249–257. doi: [10.1016/j.biopsych.2013.10.009](https://doi.org/10.1016/j.biopsych.2013.10.009) PMID: [24246360](https://pubmed.ncbi.nlm.nih.gov/24246360/)
101. Ryberg H, An J, Darko S, Lustgarten JL, Jaffa M, Gopalakrishnan V, et al. Discovery and verification of amyotrophic lateral sclerosis biomarkers by proteomics. *Muscle Nerve.* 2010; 42:104–111. doi: [10.1002/mus.21683](https://doi.org/10.1002/mus.21683) PMID: [20583124](https://pubmed.ncbi.nlm.nih.gov/20583124/)
102. Kok EH, Alanne-Kinnunen M, Isotalo K, Luoto T, Haikonen S, Goebeler S, et al. CRP gene variation affects early development of Alzheimer's disease-related plaques. *J Neuroinflammation.* 2011; 8:96. doi: [10.1186/1742-2094-8-96](https://doi.org/10.1186/1742-2094-8-96) PMID: [21831326](https://pubmed.ncbi.nlm.nih.gov/21831326/)
103. Song I-U, Chung S-W, Kim J-S, Lee K-S. Association between high-sensitivity C-reactive protein and risk of early idiopathic Parkinson's disease. *Neurol Sci.* 2011; 32:31–34. doi: [10.1007/s10072-010-0335-0](https://doi.org/10.1007/s10072-010-0335-0) PMID: [20532580](https://pubmed.ncbi.nlm.nih.gov/20532580/)
104. Pasco JA, Nicholson GC, Williams LJ, Jacka FN, Henry MJ, Kotowicz MA, et al. Association of high-sensitivity C-reactive protein with de novo major depression. *Br J Psychiatry.* 2010; 197:372–377. doi: [10.1192/bjp.bp.109.076430](https://doi.org/10.1192/bjp.bp.109.076430) PMID: [21037214](https://pubmed.ncbi.nlm.nih.gov/21037214/)
105. Prasad K. C-reactive protein (CRP)-lowering agents. *Cardiovasc Drug Rev.* 2006; 24:33–50. doi: [10.1111/j.1527-3466.2006.00033.x](https://doi.org/10.1111/j.1527-3466.2006.00033.x) PMID: [16939632](https://pubmed.ncbi.nlm.nih.gov/16939632/)
106. Wray NR, Yang J, Hayes BJ, Price AL, Goddard ME, Visscher PM. Pitfalls of predicting complex traits from SNPs. *Nat Rev Genet.* 2013; 14:507–515. doi: [10.1038/nrg3457](https://doi.org/10.1038/nrg3457) PMID: [23774735](https://pubmed.ncbi.nlm.nih.gov/23774735/)
107. Vimalaewaran KS, Berry DJ, Lu C, Tikkanen E, Pilz S, Hiraki LT, et al. Causal relationship between obesity and vitamin D status: bi-directional Mendelian randomization analysis of multiple cohorts. *PLoS Med.* 2013; 10:e1001383. doi: [10.1371/journal.pmed.1001383](https://doi.org/10.1371/journal.pmed.1001383) PMID: [23393431](https://pubmed.ncbi.nlm.nih.gov/23393431/)
108. Smith GD, Ebrahim S. Mendelian randomization: prospects, potentials, and limitations. *Int J Epidemiol.* 2004; 33:30–42. doi: [10.1093/ije/dyh132](https://doi.org/10.1093/ije/dyh132) PMID: [15075143](https://pubmed.ncbi.nlm.nih.gov/15075143/)
109. Jansen H, Samani NJ, Schunkert H. Mendelian randomization studies in coronary artery disease. *Eur Heart J.* 2014; 35:1917–1924. doi: [10.1093/eurheartj/ehu208](https://doi.org/10.1093/eurheartj/ehu208) PMID: [24917639](https://pubmed.ncbi.nlm.nih.gov/24917639/)
110. Hwang Y, Kim J, Shin JY, Kim JI, Seo JS, Webster MJ, et al. Gene expression profiling by mRNA sequencing reveals increased expression of immune/inflammation-related genes in the hippocampus of individuals with schizophrenia. *Transl Psychiatry.* 2013; 3:e321. doi: [10.1038/tp.2013.94](https://doi.org/10.1038/tp.2013.94) PMID: [24169640](https://pubmed.ncbi.nlm.nih.gov/24169640/)

Editors' Summary

Background

Inflammation is an important part of the human immune response, the network of cells and molecules that protects the body from attack by pathogens (infectious organisms) and from harmful substances and foreign particles (for example, splinters). When human cells are attacked by pathogens or injured by trauma or chemicals, molecules called inflammatory mediators induce fluid leakage from the blood vessels into the damaged tissue and attract “phagocytes” (a type of immune cell) to the site of infection or injury to “eat” the germs and dead or damaged cells. The end result is inflammation, which is characterized by swelling, redness, heat, and pain. The inflammatory response, although unpleasant, limits the damage caused by foreign invaders or chemicals by preventing further contact with body tissues. Sometimes, however, inflammation can be harmful. Persistent dysregulation of the inflammatory response is implicated in numerous somatic disorders (diseases that affect the body, such as cardiovascular disease) and neuropsychiatric disorders (mental disorders attributable to diseases of the nervous system, such as schizophrenia).

Why Was This Study Done?

Observational studies suggest that increased blood levels of C-reactive protein (CRP, an inflammatory protein) are associated with certain somatic and neuropsychiatric disorders. But observational studies cannot prove that changes in CRP levels actually cause any of these disorders. It could be that the individuals who develop a specific disease and who have a high CRP level also share another unknown characteristic that is actually responsible for disease development (confounding). Alternatively, it could be that the disease itself increases CRP levels (reverse causation). It is important to know whether CRP is causally involved in the development of specific diseases because it might then be possible to prevent or treat these diseases using drugs that control CRP levels. Here, the researchers undertake a Mendelian randomization study to determine whether CRP has a causal relationship with 32 common complex somatic and neuropsychiatric outcomes. Because gene variants are inherited randomly, they are not prone to confounding and are free from reverse causation. So, if CRP levels actually cause a specific somatic or neuropsychiatric disease, genetic variants that affect CRP levels should be associated with an altered risk for that disease.

What Did the Researchers Do and Find?

The researchers used data collected by several consortia involved in large genome-wide association studies (studies that ask whether specific genetic changes across the whole human genome, or blueprint, are associated with specific diseases) to look for associations between 32 somatic and neuropsychiatric outcomes and two genetic risk scores (GRSs) for CRP level. GRS_{CRP} consisted of four single nucleotide polymorphisms (SNPs; a type of genetic variant) in the gene encoding CRP; GRS_{GWAS} consisted of 18 SNPs that were associated with CRP level in a genome-wide association study. The researchers report that a genetically increased CRP level was significantly associated with a reduced risk of schizophrenia (a significant association is one unlikely to have arisen by chance). In addition, they found a nominally significant association (an association that needs to be confirmed) between genetically increased CRP levels and an increased risk of knee osteoarthritis,

raised diastolic and systolic blood pressure, and bipolar disorder. Notably, there was no evidence for an effect of genetically increased CRP levels on any of the other 27 outcomes studied.

What Do These Findings Mean?

These findings suggest that genetically raised levels of CRP are causally associated with protection against schizophrenia, an unexpected finding given other recent studies that suggest that raised CRP levels and brain inflammation predispose individuals to schizophrenia. The findings also provide preliminary evidence that genetically raised levels of CRP may be causally associated with an increased risk of raised blood pressure, knee arthritis, and bipolar disorder. The lack of any association between genetically raised levels of CRP and the other outcomes studied suggests, however, that many previously identified disease-associated rises in CRP levels might be a response to the disease process rather than a cause of these diseases. Like all Mendelian randomization studies, the reliability of these findings depends on the validity of several assumptions made by the researchers and on the ability of the GRSs used in the study to explain variations in CRP level. Importantly, however, these findings suggest that interventions designed to lower CRP level are unlikely to decrease the risk of people developing the majority of common complex somatic and neuropsychiatric outcomes.

Additional Information

This list of resources contains links that can be accessed when viewing the PDF on a device or via the online version of the article at <http://dx.doi.org/10.1371/journal.pmed.1001976>.


- Wikipedia has pages on [inflammation](#), [C-reactive protein](#), and [Mendelian randomization](#) (note: Wikipedia is a free online encyclopedia that anyone can edit; available in several languages)
- The MedlinePlus encyclopedia has a page on [C-reactive protein](#) (in English and Spanish)
- The American Heart Association provides a short article on [inflammation and heart disease](#)
- A UK National Health Service “Behind the Headlines” article explains a recent study that found [an association between immune activity in the brain and schizophrenia](#)

RESEARCH ARTICLE

Open Access



Genetic susceptibility loci of idiopathic interstitial pneumonia do not represent risk for systemic sclerosis: a case control study in Caucasian patients

Minghua Wu^{1*} , Shervin Assassi¹, Gloria A. Salazar¹, Claudia Pedroza², Olga Y. Gorlova³, Wei V. Chen³, Julio Charles¹, Miranda L. Taing¹, Kelley Liao¹, Fredrick M. Wigley⁴, Laura K. Hummers⁴, Ami A. Shah⁴, Monique Hinchcliff⁵, Dinesh Khanna⁶, Elena Schiopu⁶, Kristine Phillips⁶, Daniel E. Furst⁷, Virginia Steen⁸, Murray Baron⁹, Marie Hudson⁹, Xiaodong Zhou¹, Janet Pope¹⁰, Niall Jones¹¹, Peter Docherty¹², Nader A. Khalidi¹³, David Robinson¹⁴, Robert W. Simms¹⁵, Richard M. Silver¹⁶, Tracy M. Frech¹⁷, Barri J. Fessler¹⁸, Marvin J. Fritzler¹⁹, Jerry A. Molitor²⁰, Barbara M. Segal²⁰, Malahat Movahedian¹, Javier Martín²¹, John Varga⁵ and Maureen D. Mayes¹

Abstract

Background: Systemic sclerosis (SSc)-related interstitial lung disease (ILD) has phenotypic similarities to lung involvement in idiopathic interstitial pneumonia (IIP). We aimed to assess whether genetic susceptibility loci recently identified in the large IIP genome-wide association studies (GWASs) were also risk loci for SSc overall or severity of ILD in SSc.

Methods: A total of 2571 SSc patients and 4500 healthy controls were investigated from the US discovery GWAS and additional US replication cohorts. Thirteen IIP-related selected single nucleotide polymorphisms (SNPs) were genotyped and analyzed for their association with SSc.

Results: We found an association of SSc with the SNP rs6793295 in the *LRRC34* gene (OR = 1.14, CI 95 % 1.03 to 1.25, p value = 0.009) and rs11191865 in the *OBFC1* gene (OR = 1.09, CI 95 % 1.00 to 1.19, p value = 0.043) in the discovery cohort. Additionally, rs7934606 in *MUC2* (OR = 1.24, CI 95 % 1.01 to 1.52, p value = 0.037) was associated with SSc-ILD defined by imaging. However, these associations failed to replicate in the validation cohort. Furthermore, SNPs rs2076295 in *DSP* (β = -2.29, CI 95 % -3.85 to -0.74, p value = 0.004) rs17690703 in *SPPL2C* (β = 2.04, CI 95 % 0.21 to 3.88, p value = 0.029) and rs1981997 in *MAPT* (β = 2.26, CI 95 % 0.35 to 4.17, p value = 0.02) were associated with percent predicted forced vital capacity (FVC%) even after adjusting for the anti-topoisomerase (ATA)-positive subset. However, these associations also did not replicate in the validation cohort.

Conclusions: Our results add new evidence that SSc and SSc-related ILD are genetically distinct from IIP, although they share phenotypic similarities.

Keywords: Idiopathic interstitial pneumonia (IIP), SSc-ILD, Genetic susceptibility

* Correspondence: Minghua.wu@uth.tmc.edu

¹Division of Rheumatology and Clinical Immunogenetics, Department of Internal Medicine, University of Texas McGovern Medical School at Houston, 6431 Fannin Street, Houston, TX 77030, USA

Full list of author information is available at the end of the article

Background

Systemic sclerosis or scleroderma (SSc) is a complex autoimmune disease characterized by vasculopathy, autoantibody production and fibrosis in the skin and internal organs. The etiology of SSc remains unknown; effective treatments that target the underlying pathophysiology of SSc are unavailable and the disease-related mortality remains high [1]. Specifically, pulmonary fibrosis and pulmonary arterial hypertension (PAH) account for the majority of disease-related deaths in SSc [2]. SSc-associated interstitial lung disease (SSc-ILD) is usually characterized by a histologic pattern of nonspecific interstitial pneumonia (NSIP) or, less frequently, usual interstitial pneumonia (UIP). SSc-ILD has clinical and radiologic similarities to idiopathic interstitial pneumonia (IIP). A polymorphism in the MUC5B promoter region was strongly associated with familial and idiopathic pulmonary fibrosis (IPF; the most common IIP type) [3], but it has not been found to be a susceptibility locus for SSc or SSc-associated interstitial lung disease (SSc-ILD) [4–6].

Recently, large genome-wide association studies (GWASs) in fibrotic IIP (n = 1616 patients) and IPF (n = 542 patients and two validation cohorts n = 544 and n = 324) identified/confirmed susceptibility single nucleotide polymorphisms (SNPs) in *TERT*, *AZGP1*, *MUC2*, *IVD*, *DSP*, *MAPT*, *DPP9*, *LRRC34*, *FAM13A*, *OBFC1*, *CSMD1*, *ATP11A* [7] and *TOLLIP*, *MDGA2*, *SPPL2C* [8]. Given the clinical and radiologic similarities between SSc-ILD and IIP, we examined the association of the 13 SNPs among the listed genes, identified in the above IIP and IPF GWASs [7, 8] (*TOLLIP* and *MDGA2* genes were excluded because the related SNPs were not present on the Illumina BeadChip utilized in SSc GWAS) with SSc as a single disease entity, with SSc-ILD by imaging, or SSc-ILD severity (as determined by percent predicted forced vital capacity (FVC%)) in two large SSc patient samples.

Methods

Study population

Two non-Hispanic white populations (discovery and replication cohorts) were investigated. For the discovery cohort, we utilized data from our previously published SSc GWAS study consisting of 1486 SSc cases (patients from the US) and 3477 unaffected race- and ethnicity-matched controls [9]. Selected polymorphisms were genotyped in an independent replication cohort consisting of 1085 additional SSc cases (patients from the US and Canada) and 1023 additional unaffected controls. Patients were recruited at the University of Texas – Houston and from the following sites: the participating Canadian Scleroderma Research Group (CSRG) sites, University of California Los Angeles, University of Michigan, Georgetown University, Boston University, Medical University of South Carolina, Johns Hopkins University, University

of Utah, Northwestern University, University of Alabama at Birmingham and University of Minnesota. All patients were enrolled in the National Scleroderma Family Registry and DNA Repository. All patients with SSc fulfilled the 1980 American College of Rheumatology classification criteria for SSc or had at least three of the five CREST (Calcinosis, Raynaud's phenomenon, Esophageal dysmotility, Sclerodactyly, Telangiectasias) features (Table 1).

The genotypes of unaffected controls for the discovery cohort were obtained from the Cancer Genetic Markers of Susceptibility (CGEMS; non-cancer healthy controls) studies and Illumina iControlDB database (www.illumina.com/iControlDB, Illumina, San Diego, CA, USA). The unaffected controls for the replication cohort were recruited through a nationwide effort by the Scleroderma Family Registry and DNA Repository.

Collection of blood samples and clinical information from case and control subjects was undertaken with fully informed consent and relevant ethical review board approval from each contributing center in accordance with the tenets of the Declaration of Helsinki.

SNP selection and genotyping assay

In the discovery cohort, genotyping was performed using the Illumina Bead-Array GWAS platform. Specifically, patients were genotyped using Illumina Human 610-Quad BeadChip and controls were genotyped on Illumina Hap550K-BeadChip [9]. The 13 single nucleotide polymorphisms (SNPs) identified/confirmed to be associated with IIP [7], rs2736100 (*TERT*), rs2076295 (*DSP*), rs4727443 (*AZGP1*), rs7934606 (*MUC2*), rs2034650 (*IVD*), rs1981997 (*MAPT*), rs12610495 (*DPP9*), rs6793295 (*LRRC34*), rs2609255 (*FAM13A*), rs11191865 (*OBFC1*), rs1278769 (*ATP11A*), rs1379326 (*CSMD1*) and an additional SNP rs17690703 (*SPPL2C*) [8], were investigated in the discovery cohort. We also investigated association of 13 SNPs with anti-topoisomerase 1 antibody (ATA) and anti-centromere

Table 1 Demographic and clinical features of SSc patients in the discovery and replication cohorts

	Discovery (n = 1486)	Replication (n = 1085)
Female	1307 (88.0 %)	934 (86.1 %)
Mean age ± years	54.5 ± 12.9 y	56.1 ± 12.8 y
Diffuse SSc	505 (34.0 %)	403 (37.1 %)
Disease duration	10.1 ± 8.7 y	10.9 ± 9.6 y
ACA	425 (28.6 %)	340 (31.3 %)
ATA	238 (16.0 %)	180 (16.6 %)
SSc-ILD		
Yes	185 (65.6 %)	313 (61.9 %)
No	97 (34.4 %)	193 (38.1 %)

ACA anti-centromere antibody, ATA anti-topoisomerase 1 antibody, SSc systemic sclerosis, ILD interstitial lung disease

antibody (ACA) positive subgroup patients. The SNPs reaching a nominal level of significance ($p < 0.05$) in the discovery cohort (rs6793295 and rs11191865) were genotyped in the replication cohort using TaqMan allele discrimination assays in a 7900HT fast real-time PCR system (Applied Biosystems, Foster City, CA, USA).

SSc-ILD and severity of SSc-ILD

We also investigated the association of the above SNPs with SSc-ILD by imaging. We compared the frequency of the above SNPs in patients with SSc-ILD to unaffected controls. SSc-ILD was defined by chest high-resolution computed tomography (HRCT)/chest X-ray images. SSc patients (total = 498; $n = 185$ in the discovery cohort and $n = 313$ in the replication cohort) were considered to have interstitial lung disease (ILD) by imaging if they had: (1) ground glass opacity or increased interstitial markings on chest HRCT; or (2) increased basilar reticular marking on chest X-ray (in total, 498 patients, 185 in the discovery cohort and 313 in the replication cohort had ILD). The number of patients with imaging-proven ILD is low because HRCT imaging results were available only $n = 212$ in the discovery cohort and $n = 347$ in the replication cohort patients. The case-case comparison in regard to presence of ILD was compared by the low number of patients with available HRCT imaging in the discovery cohort.

Furthermore, an association with severity of ILD was investigated. Percent predicted forced vital capacity (FVC%) (measured in a total of 1954 patients; $n = 1072$ in the discovery cohort and $n = 882$ in the replication cohort) as a continuous variable was used as a surrogate for severity of SSc-ILD as this has been demonstrated to be a validated outcome measure for severity of ILD in randomized controlled studies of patients with SSc [10]. SNPs reaching the nominal significance level ($p < 0.05$) in these two analyses were also genotyped in the replication cohort (rs7934606, rs2076295, rs17690703 and rs1981997).

Statistical analysis

We followed the same genetic inheritance modes for each specific SNP that was utilized in the IIP GWAS [7]. Specifically, all SNPs except for rs1379326 (*CSMD1*) were investigated in an additive model. Similar to the IIP GWAS [7], rs1379326 (*CSMD1*) was investigated in a recessive model. The additive model corresponds to the risk or protective effect conferred by the rarer (minor) allele (0, 1 or 2 copies) as a predictor of phenotypic status; and the recessive model corresponds to the effect conferred by only homozygous status for the rarer allele. Genotype data quality was verified for each SNP by testing for Hardy-Weinberg equilibrium (HWE). Hardy-Weinberg equilibrium was assessed by an χ^2 test or

Fisher's exact test. None of the included cohorts showed significant deviation from HWE for all the genotyped SNPs. Logistic regression that included both patients and controls was utilized to examine the association of the above SNPs with SSc overall and SSc-ILD. A linear regression model was used to investigate the association of these genetic variants with FVC% predicted (data was normally distributed) as a continuous variable, among patients only. FVC% predicted was investigated as a continuous rather than a dichotomous variable to increase our power to detect a difference in FVC levels conditional on the genotype. Anti-topoisomerase 1 antibody (ATA) status was included as a potential confounder in the multivariable regression model. The combined analysis of the discovery and the replication cohorts was performed via a random-effects meta-analysis model.

Results

Discovery cohort

We first investigated whether the 13 IIP-associated polymorphisms were associated with risk of SSc overall or with SSc-ILD. Several SNPs showed nominally significant associations with SSc in the discovery cohort (Table 2). Specifically, *LRRC34* rs6793295 (OR = 1.14, CI 95 % 1.03 to 1.25, p value = 0.009) and *OBFC1* rs11191865 (OR = 1.09, CI 95 % 1.00 to 1.19, p value = 0.043) were associated with SSc compared to controls (Table 2). There were no significant associations observed when the SNP frequencies in SSc antibody subgroups (ATA- or ACA-positive patients) were compared to controls (Table 3).

Next, we investigated the association in the discovery cohort between the above polymorphisms and SSc-ILD by imaging or FVC% predicted in order to investigate the relationship with SSc-ILD presence or severity. We also performed case-case comparison by presence of ILD with no ILD in SSc. Interestingly, rs7934606 in *MUC2* (OR = 1.24, CI 95 % 1.01 to 1.52, p value = 0.037) was associated with SSc-ILD by imaging when compared to controls (Table 4). Furthermore, three polymorphisms *DSP* rs2076295 ($\beta = -2.29$, CI 95 % -3.85 to -0.74, p value = 0.004) *SPPL2C* rs17690703 ($\beta = 2.04$, CI 95 % 0.21 to 3.88, p value = 0.029) and *MAPT* rs1981997 ($\beta = 2.26$, CI 95 % 0.35 to 4.17, p value = 0.02) were associated with FVC% predicted (Table 5). Of note, the minor allele (A) of *MAPT* rs1981997 and (T) of *SPPL2C* rs17690703 were associated with higher FVC% predicted, which is consistent with the IIP GWAS results with the minor allele being protective in that study (OR < 1) [7]. Even after adjustment for ATA status, the above three polymorphisms showed nominally significant associations with ILD severity in SSc patients (Table 5). *SPPL2C* rs17690703 was associated with SSc-ILD compared to SSc with no ILD as determined by imaging in the discovery cohort (Table S1 in Additional file 1).

Table 2 Association between the investigated genotypes in SSc-versus-control comparisons in the discovery cohort

SNP	Gene	Minor allele*	MAF (case)**	OR (95 % CI)	p value
rs2736100	<i>TERT</i>	T	0.50	1.04 (0.95; 1.13) [†]	0.393
rs2076295	<i>DSP</i>	G	0.44	0.96 (0.88; 1.05) [†]	0.387
rs4727443	<i>AZGP1</i>	A	0.41	1.03 (0.95; 1.13) [†]	0.487
rs7934606	<i>MUC2</i>	A	0.41	1.08 (0.99; 1.18) [†]	0.084
rs2034650	<i>IVD</i>	G	0.49	1.03 (0.95; 1.13) [†]	0.458
rs1981997	<i>MAPT</i>	A	0.22	0.97 (0.88; 1.08) [†]	0.636
rs12610495	<i>DPP9</i>	G	0.28	1.00 (0.91; 1.10) [†]	0.967
rs6793295	<i>LRRC34</i>	C	0.29	1.14 (1.03; 1.25)[†]	0.009
rs2609255	<i>FAM13A</i>	G	0.22	1.01 (0.91; 1.12) [†]	0.851
rs11191865	<i>OBFC1</i>	G	0.52	1.09 (1.00; 1.19)[†]	0.043
rs1278769	<i>ATP11A</i>	A	0.24	0.95 (0.86; 1.06) [†]	0.365
rs1379326	<i>CSMD1</i>	G	0.26	0.86 (0.69; 1.08) [‡]	0.199
rs17690703	<i>SPPL2C</i>	T	0.26	1.03 (0.95; 1.11)	0.546

Gene names: *TERT*: telomerase reverse transcriptase; *DSP*: desmoplakin; *AZGP1*: alpha-2-glycoprotein 1, zinc-binding; *MUC2*: mucin 2; *IVD*: isovaleryl-CoA dehydrogenase; *MAPT*: microtubule-associated protein tau; *DPP9*: dipeptidyl-peptidase 9; *LRRC34*: leucine-rich repeat-containing 34; *FAM13A*: family with sequence similarity 13, member A; *OBFC1*: oligonucleotide/oligosaccharide-binding fold containing 1; *ATP11A*: ATPase, class VI, type 11A; *CSMD1*: CUB and Sushi multiple domains 1; *SPPL2C*: signal peptide peptidase-like 2C

SSc systemic sclerosis, SNP single nucleotide polymorphism, MAF minor allele frequency

*The minor allele is defined as the minor allele in the combined case and control group; **MAF is defined as the minor allele frequency in cases and controls;

[†]Odds ratio for the minor allele in the additive model; [‡]Odds ratio based on the recessive model

Replication cohort and combined analysis

We genotyped the samples from the validation set for the above six SNPs, which either showed nominally significant associations ($p < 0.05$) with risk of SSc overall (*LRRC34* rs6793295 and *OBFC1* rs11191865) or with

SSc-ILD (rs7934606 in *MUC2*) and FVC% predicted (*DSP* rs2076295, *SPPL2C* rs17690703 and *MAPT* rs1981997). However, these associations were not found in the replication cohort (Table 4 and upper part of Table 6). Contrary to the discovery cohort, *MAPT*

Table 3 Association studies of investigated genotypes in comparison of ATA- or ACA-positive SSc patients to controls in the discovery cohort

SNP	Gene	ATA (+)		ACA (+)	
		OR (95 % CI)	p value	OR (95 % CI)	p value
rs2736100*	<i>TERT</i>	1.01 (0.84; 1.22)	0.881	1.14 (0.99; 1.31)	0.069
rs2076295*	<i>DSP</i>	1.01 (0.83; 1.21)	0.945	0.94 (0.82; 1.09)	0.423
rs4727443*	<i>AZGP1</i>	1.12 (0.93; 1.36)	0.226	0.96 (0.83; 1.11)	0.609
rs7934606*	<i>MUC2</i>	1.00 (0.83; 1.20)	0.992	1.07 (0.93; 1.23)	0.925
rs2034650*	<i>IVD</i>	0.99 (0.82; 1.19)	0.915	1.04 (0.90; 1.20)	0.584
rs1981997*	<i>MAPT</i>	0.92 (0.73; 1.16)	0.495	1.15 (0.97; 1.36)	0.100
rs12610495*	<i>DPP9</i>	0.98 (0.80; 1.20)	0.847	1.05 (0.90; 1.22)	0.572
rs6793295*	<i>LRRC34</i>	1.15 (0.94; 1.42)	0.172	1.15 (0.98; 1.35)	0.080
rs2609255*	<i>FAM13A</i>	1.04 (0.83; 1.30)	0.736	1.00 (0.84; 1.19)	1.000
rs11191865*	<i>OBFC1</i>	1.21 (1.00; 1.45)	0.048	0.99 (0.86; 1.14)	0.898
rs1278769*	<i>ATP11A</i>	1.09 (0.89; 1.35)	0.394	0.96 (0.81; 1.14)	0.650
rs1379326 [†]	<i>CSMD1</i>	0.78 (0.46; 1.31)	0.340	0.94 (0.65; 1.36)	0.736
rs17690703	<i>SPPL2C</i>	0.97 (0.76; 1.26)	0.870	1.14 (0.94; 1.38)	0.184

Gene names: *TERT*: telomerase reverse transcriptase; *DSP*: desmoplakin; *AZGP1*: alpha-2-glycoprotein 1, zinc-binding; *MUC2*: mucin 2; *IVD*: isovaleryl-CoA dehydrogenase; *MAPT*: microtubule-associated protein tau; *DPP9*: dipeptidyl-peptidase 9; *LRRC34*: leucine-rich repeat-containing 34; *FAM13A*: family with sequence similarity 13, member A; *OBFC1*: oligonucleotide/oligosaccharide-binding fold containing 1; *ATP11A*: ATPase, class VI, type 11A; *CSMD1*: CUB and Sushi multiple domains 1; *SPPL2C*: signal peptide peptidase-like 2C

ATA anti-topoisomerase 1 antibody, ACA anti-centromere antibody, SSc systemic sclerosis, SNP single nucleotide polymorphism

*Additive model; [†]Recessive model

Table 4 Association between the investigated genotypes in SSc-ILD (by imaging) patients compared to controls in the discovery and replication cohort

SNP (coded allele)	Gene	SSc-ILD vs control (discovery cohort)		SSc-ILD vs control (replication cohort)	
		OR (95 % CI)	p value	OR (95 % CI)	p value
rs2736100 (T)*	<i>TERT</i>	1.14 (0.93; 1.40)	0.199		
rs2076295 (G)*	<i>DSP</i>	1.13 (0.92; 1.39)	0.243		
rs4727443 (A)*	<i>AZGP1</i>	0.89 (0.71; 1.10)	0.270		
rs7934606 (A)*	<i>MUC2</i>	1.24 (1.01; 1.52)	0.037	1.14 (0.81; 1.63)	0.444
rs2034650 (C)*	<i>IVD</i>	0.86 (0.70; 1.06)	0.152		
rs1981997 (A)*	<i>MAPT</i>	0.81 (0.62; 1.06)	0.123		
rs12610495 (G)*	<i>DPP9</i>	1.17 (0.93; 1.46)	0.179		
rs6793295 (C)*	<i>LRRC34</i>	1.02 (0.80; 1.30)	0.863		
rs2609255 (G)*	<i>FAM13A</i>	0.99 (0.77; 1.27)	0.919		
rs11191865 (G)*	<i>OBFC1</i>	1.08 (0.87; 1.32)	0.496		
rs1278769 (A)*	<i>ATP11A</i>	1.00 (0.79; 1.28)	0.987		
rs1379326 (G)†	<i>CSMD1</i>	0.68 (0.37; 1.27)	0.224		
rs17690703 (T)*	<i>SPPL2C</i>	1.23 (0.87; 1.72)	0.238		

Gene names: *TERT*: telomerase reverse transcriptase; *DSP*: desmoplakin; *AZGP1*: alpha-2-glycoprotein 1, zinc-binding; *MUC2*: mucin 2; *IVD*: isovaleryl-CoA dehydrogenase; *MAPT*: microtubule-associated protein tau; *DPP9*: dipeptidyl-peptidase 9; *LRRC34*: leucine-rich repeat-containing 34; *FAM13A*: family with sequence similarity 13, member A; *OBFC1*: oligonucleotide/oligosaccharide-binding fold containing 1; *ATP11A*: ATPase, class VI, type 11A; *CSMD1*: CUB and Sushi multiple domains 1; *SPPL2C*: signal peptide peptidase-like 2C

SSc systemic sclerosis, *ILD* interstitial lung disease, *SNP* single nucleotide polymorphism

*Additive model; †Recessive model

rs1981997 (OR = 0.79; CI 95 % 0.68 to 0.91; *p* value = 0.002) and *SPPL2C* rs17690703 (OR = 0.64; CI 95 % 0.49 to 0.83; *p* value = 0.001) were associated with SSc susceptibility but not with ILD severity in the replication cohort (Table 6). Furthermore, *SPPL2C* rs17690703 was

not associated with SSc-ILD when compared to SSc without ILD by imaging in the validation cohort (Table S1 in Additional file 1).

In the combined discovery and replication cohorts (lower part of Table 6), only the *LRRC34* rs6793295 was

Table 5 Association between the investigated genotypes and FVC% predicted in SSc patients in the discovery cohort

SNP (Coded allele)	Gene	FVC%		FVC% (adjusted for ATA status)	
		β (95 % CI)	p value	β (95 % CI)	p value
rs2736100 (T)*	<i>TERT</i>	0.51 (-1.04; 2.05)	0.521	0.76 (-0.81; 2.34)	0.343
rs2076295 (G)*	<i>DSP</i>	-2.29 (-3.85; -0.74)	0.004	-2.52 (-4.09; -0.95)	0.002
rs4727443 (A)*	<i>AZGP1</i>	-0.27 (-1.85; 1.32)	0.742	0.08 (-1.53; 1.70)	0.919
rs7934606 (A)*	<i>MUC2</i>	-0.55 (-2.16; 1.05)	0.499	-0.75 (-2.37; 0.86)	0.360
rs2034650 (C)*	<i>IVD</i>	0.67 (-0.90; 2.25)	0.402	0.37 (-1.22; 1.97)	0.646
rs1981997 (A)*	<i>MAPT</i>	2.26 (0.35; 4.17)	0.020	3.19 (1.23; 5.14)	0.001
rs12610495 (G)*	<i>DPP9</i>	0.57 (-1.18; 2.32)	0.523	1.05 (-0.72; 2.82)	0.246
rs6793295 (C)*	<i>LRRC34</i>	-0.46 (-2.19; 1.27)	0.599	-1.05 (-2.82; 0.71)	0.242
rs2609255 (G)*	<i>FAM13A</i>	1.72 (-0.10; 3.55)	0.064	1.61 (-0.18; 3.41)	0.079
rs11191865 (G)*	<i>OBFC1</i>	-0.73 (-2.27; 0.82)	0.356	0.02 (-1.55; 1.60)	0.977
rs1278769 (A)*	<i>ATP11A</i>	-0.62 (-2.45; 1.20)	0.503	-0.74 (-2.60; 1.11)	0.433
rs1379326 (G)†	<i>CSMD1</i>	3.35 (-0.84; 7.54)	0.117	3.12 (-1.08; 7.31)	0.145
rs17690703 (T)*	<i>SPPL2C</i>	2.04 (0.21; 3.88)	0.029	2.11 (0.30; 3.91)	0.022

Gene names: *TERT*: telomerase reverse transcriptase; *DSP*: desmoplakin; *AZGP1*: alpha-2-glycoprotein 1, zinc-binding; *MUC2*: mucin 2; *IVD*: isovaleryl-CoA dehydrogenase; *MAPT*: microtubule-associated protein tau; *DPP9*: dipeptidyl-peptidase 9; *LRRC34*: leucine-rich repeat-containing 34; *FAM13A*: family with sequence similarity 13, member A; *OBFC1*: oligonucleotide/oligosaccharide-binding fold containing 1; *ATP11A*: ATPase, class VI, type 11A; *CSMD1*: CUB and Sushi multiple domains 1; *SPPL2C*: signal peptide peptidase-like 2C

FVC% percent predicted forced vital capacity, SSc systemic sclerosis, *ATA* anti-topoisomerase 1 antibody, *SNP* single nucleotide polymorphism

*Additive model; †Recessive model

Table 6 Association analysis of investigated genotypes in SSc versus control comparison as well as FVC% in the replication and combined cohorts

	SNPs	Gene	Allele frequency		FVC%	FVC%–adjusted for ATA		
			OR (95 % CI)	<i>p</i> value	β (95 % CI)	<i>p</i> value	β (95 % CI)	<i>p</i> value
Replication	rs2076295 (G) [†]	<i>DSP</i>	0.93 (0.82; 1.05)	0.252	0.32 (-1.49; 2.13)	0.729	0.39 (-1.40; 2.17)	0.671
	rs1981997 (A)[†]	<i>MAPT</i>	0.79 (0.68; 0.91)	0.002	0.15 (-2.05; 2.35)	0.894	-0.01 (-2.17; 2.15)	0.994
	rs6793295 (C) [†]	<i>LRRC34</i>	1.06 (0.92; 1.21)	0.435	-0.54 (-2.53; 1.45)	0.593	-0.55 (-2.51; 1.41)	0.581
	rs11191865 (G) [†]	<i>OBFC1</i>	0.90 (0.80; 1.01)	0.093	-0.34 (-2.11; 1.44)	0.710	-0.30 (-2.05; 1.45)	0.736
	rs17690703 (T)[†]	<i>SPPL2C</i>	0.64 (0.49; 0.83)	0.001	2.20 (-1.38; 5.78)	0.227	3.11 (-0.45; 6.66)	0.086
Combined	rs2076295 (G) [†]	<i>DSP</i>	0.95 (0.89; 1.02)	0.171	-1.02 (-3.58; 1.54)	0.579	-1.07 (-3.94; 1.80)	0.079
	rs1981997 (A) [†]	<i>MAPT</i>	0.89 (0.78; 1.02)	0.091	1.26 (-0.81; 3.32)	0.445	1.59 (-1.56; 4.74)	0.503
	rs6793295 (C)[†]	<i>LRRC34</i>	1.11 (1.03; 1.20)	0.010	-0.50 (-1.80; 0.81)	0.590	-0.81 (-2.12; 0.50)	0.440
	rs11191865 (G) [†]	<i>OBFC1</i>	1.01 (0.90; 1.13)	0.912	-0.55 (-1.72; 0.61)	0.524	-0.13 (-1.30; 1.04)	0.868
	rs17690703 (T)[†]	<i>SPPL2C</i>	1.04 (0.97; 1.13)	0.221	2.32 (0.68; 3.96)	0.006	-1.03 (-4.67; 2.60)	0.575

Gene names: *TERT*: telomerase reverse transcriptase; *DSP*: desmoplakin; *AZGP1*: alpha-2-glycoprotein 1, zinc-binding; *MUC2*: mucin 2; *IVD*: isovaleryl-CoA dehydrogenase; *MAPT*: microtubule-associated protein tau; *DPP9*: dipeptidyl-peptidase 9; *LRRC34*: leucine-rich repeat-containing 34; *FAM13A*: family with sequence similarity 13, member A; *OBFC1*: oligonucleotide/oligosaccharide-binding fold containing 1; *ATP11A*: ATPase, class VI, type 11A; *CSMD1*: CUB and Sushi multiple domains 1; *SPPL2C*: signal peptide peptidase-like 2C

SSc systemic sclerosis, FVC % percent predicted forced vital capacity, ATA anti-topoisomerase 1 antibody, SNP single nucleotide polymorphism

[†]Additive model

associated with risk of SSc overall (OR = 1.11; CI 95 % 1.03 to 1.2; *p* value = 0.01). However, the observed significance level was not stronger in the combined cohort than in the discovery cohort and this association did not withstand correction for multiple comparisons. Specifically, the *p* value = 0.13 in the combined cohort after Bonferroni correction for multiple comparisons. Furthermore, *LRRC34* rs6793295 was not associated with FVC% predicted in the combined cohort. *SPPL2C* rs17690703 was associated with FVC% predicted in the combined cohort but not after ATA adjustment or with overall SSc risk (Table 6). The observed association of *SPPL2C* rs17690703 with FVC% did not withstand correction for multiple comparison (*p*_{corr} = 0.078).

Discussion

The aim of this study was to assess genetic components of SSc and SSc-related ILD, specifically to determine if SSc-ILD and IIP share common genetic risk factors. We analyzed 13 SNPs which showed robust association with IIP in a previous GWAS report [7, 8]. However, we failed to replicate these associations with SSc or SSc-related ILD in a large North American cohort.

Interstitial lung disease (ILD) is the most common pulmonary manifestation in patients with SSc [11] and is the most frequent cause of SSc disease-related death [12]. ILD occurs with greater frequency and increased severity in patients with diffuse cutaneous SSc (dc-SSc) and in ATA-positive patients. In SSc-ILD, loss of lung function and extent of fibrosis are the most important prognostic factors. It is known that genetic factors contribute not only to SSc susceptibility, but also to

predisposition to SSc clinical phenotypes including disease type (limited versus diffuse cutaneous disease) and autoantibody status [13]. Recently, several lines of evidence have suggested that, even though SSc-ILD has phenotypic similarities to IIP, the genetic risk factors for these two conditions are quite different. For example the *MUC5B* promoter region polymorphism (rs868903), which was strongly associated with familial and idiopathic pulmonary fibrosis [3, 7], was not identified as a susceptibility locus for SSc or SSc-associated interstitial lung disease [4–6]. In addition, the two large SSc GWAS reports (each with independent discovery and validation cohorts) identified the major histocompatibility complex (MHC) class II region as being the most strongly associated region with SSc [9, 14], whereas MHC loci were not found to be risk factors in the IIP GWAS [7].

The present study has some limitations. In our study, only a subgroup of SSc patients had undergone HRCT. We performed a case-case comparison in regard to presence of SSc-ILD versus no SSc-ILD in the low number of patients because HRCT results were not available in the remainder of patients. The genotype results for *TOLLIP* and *MDGA2* genes were not available on our GWAS platform and could not be investigated. Furthermore, this is a cross-sectional study, thus, we cannot examine whether the investigated genetic loci have predictive significance for ILD progression. However, mean disease duration of this study was more than 10 years and most of the SSc-ILD cases had been already established. Therefore, we believe the cross-sectional FVC% is a reasonable surrogate for severity of ILD.

Conclusions

In this study, we confirm that genetic susceptibility loci for IIP are not risk loci for SSc or severity of SSc-ILD in two validation SSc cohorts of non-Hispanic white ethnic background. Our findings and those of previous genetic susceptibility studies in SSc [15] implicate pathways in innate and adaptive immunity, whereas susceptibility loci for IIP relate to epithelial cell injury/dysfunction and abnormal wound healing [16]. Future challenges will be to identify the functional relevance of these variants for the final common pathway that results in the phenotype of fibrotic lung disease. This study represents an important step forward toward a better understanding of the complex genetic association of SSc particularly with lung involvement. We add new evidence that SSc and SSc-ILD are genetically distinct from IIP.

Additional file

Additional file 1: Table S1. Association between the investigated genotypes in SSc-ILD (by imaging) patients compared to SSc-no ILD (by imaging) in the discovery and replication cohort. (DOCX 19 kb)

Abbreviations

ACA: Anti-centromere antibody; ATA: Anti-topoisomerase 1 antibody; CGEMS: Cancer Genetic Markers of Susceptibility; CREST: Calcinosis, Raynaud's phenomenon, Esophageal dysmotility, Sclerodactyly, Telangiectasias; dcSSc: Diffuse cutaneous SSc; FVC%: Percent predicted forced vital capacity; GWAS: Genome-wide association study; HRCT: High-resolution computed tomography; HWE: Hardy-Weinberg equilibrium; IIP: Idiopathic interstitial pneumonia; ILD: Interstitial lung disease; IPF: Idiopathic pulmonary fibrosis; MHC: Major histocompatibility complex; NSIP: Nonspecific interstitial pneumonia; PAH: Pulmonary arterial hypertension; SNP: Single nucleotide polymorphisms; SSc: Systemic sclerosis; SSc-ILD: SSc-associated interstitial lung disease; UIP: Usual interstitial pneumonia.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

MW, SA, JV, and MDM contributed to study conception and design, MW, SA, GAS, JC, MLT, KL, FMW, LKH, AAS, MH, DK, ES, KP, DEF, VS, MB, MH, XZ, JP, NJ, PD, NAK, DR, RWS, RMS, TMF, BJF, MJF, JAM, BMS, MM, JM, JV, and MDM contributed to acquisition of data. MW, SA, GAS, CP, OYG, WVC, JC, MLT, KL, FMW, LKH, AAS, MH, DK, ES, KP, DEF, VS, MB, MH, XZ, JP, NJ, PD, NAK, DR, RWS, RMS, TMF, BJF, MJF, JAM, BMS, MM, JM, JV, and MDM contributed to analysis and interpretation of data. MW had full access to all of the data in the study and takes responsibility for the integrity of the data and accuracy of the data analysis. All authors were involved in drafting the manuscript or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication.

Acknowledgements

The authors thank Marilyn Perry and Yuxiao Du for their work as Registry coordinators and we are grateful for the generous participation of our subjects. This work was supported by the Scleroderma Foundation New Investigator Award to Dr. Wu, and by the National Institutes of Health (NIH) National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) Centers of Research Translation (CORT) grant P50AR054144 to Dr. Mayes, NIH grant K23AR061436 to Dr. Assassi, NIH/NIAMS Scleroderma Family Registry and DNA Repository grant N01-AR02251 to Dr. Mayes, NIH/NIAMS grant AR055258 to Dr. Mayes, NIH National Center for Clinical and Translational Sciences grant 3UL1RR024148, Department of Defense Congressionally

Directed Medical Research Program W81XWH-13-1-0452, Proposal number PR120687 to Dr. Mayes.

Author details

¹Division of Rheumatology and Clinical Immunogenetics, Department of Internal Medicine, University of Texas McGovern Medical School at Houston, 6431 Fannin Street, Houston, TX 77030, USA. ²Department of Pediatrics, The University of Texas McGovern Medical School at Houston, 6431 Fannin Street, Houston, TX 77030, USA. ³Department of Community and Family Medicine, Geisel School of Medicine at Dartmouth, 1 Rope Ferry Road, Hanover, NH 03755, USA. ⁴Division of Rheumatology, Johns Hopkins University School of Medicine, 5501 Hopkins Bayview circle, Baltimore, MD 21224, USA. ⁵Division of Rheumatology, Northwestern University Feinberg School of Medicine, 240 East Huron Street, Chicago, IL 60611, USA. ⁶Department of Internal Medicine, Division of Rheumatology, University of Michigan Health Center, 300 North Ingalls Street, Ann Arbor, MI 48109, USA. ⁷Division of Rheumatology, University of California Los Angeles, 1000 Veterans Avenue, Los Angeles, CA 90024, USA. ⁸Division of Rheumatology, Georgetown University Medical Center, 3800 Reservoir Road, Washington, MD 20007, USA. ⁹Division of Rheumatology, McGill University, 1650 Cedar Avenue, Montreal, QC H3G 1A4, Canada. ¹⁰Division of Rheumatology, St. Joseph's Health Care, University of Western Ontario, 268 Grosvenor Street, London, ON, Canada. ¹¹Division of Rheumatology, University of Alberta, 11405-87 Avenue, Edmonton, AB T6G 1C9, Canada. ¹²Division of Rheumatology, Moncton Hospital, 135 MacBeath Avenue, Moncton, NB E1C 6Z8, Canada. ¹³Division of Rheumatology, McMaster University at Hamilton, 1280 Main Street West, Hamilton, ON L8S 4L8, Canada. ¹⁴Division of Rheumatology, University of Manitoba, 820 Sherbrook Street, Winnipeg, MB R3A 1R9, Canada. ¹⁵Division of Rheumatology, Boston University School of Medicine, 72 East Concord Street, Boston, MA 02118, USA. ¹⁶Division of Rheumatology, Medical University of South Carolina, 171 Ashley Avenue, Charleston, SC 29425, USA. ¹⁷Division of Rheumatology, University of Utah, 30 North 1900 East, Salt Lake City, UT 84132, USA. ¹⁸Division of Clinical Immunology and Rheumatology, University of Alabama at Birmingham, 1825 University Boulevard, Birmingham, AB 35294, USA. ¹⁹Division of Rheumatology, University of Calgary Cumming School of Medicine, 3330 Hospital Drive NW, Calgary, AB T2N 4N1, Canada. ²⁰Division of Rheumatology, University of Minnesota, 420 Delaware Street SE, Minneapolis, MN 55455, USA. ²¹Instituto de Parasitología y Biomedicina López-Neyra, Consejo Superior de Investigaciones Científicas, C/Ventana 11, 18001 Granada, Spain.

Received: 8 July 2015 Accepted: 6 January 2016

Published online: 20 January 2016

References

- Elhai M, Meune C, Avouac J, Kahan A, Allanore Y. Trends in mortality in patients with systemic sclerosis over 40 years: a systematic review and meta-analysis of cohort studies. *Rheumatology* (Oxford). 2012;51:1017–26.
- Tyndall AJ, Bannert B, Vonk M, Airo P, Cozzi F, Carreira PE, et al. Causes and risk factors for death in systemic sclerosis: a study from the EULAR Scleroderma Trials and Research (EUSTAR) database. *Ann Rheum Dis*. 2010;69:1809–15.
- Seibold MA, Wise AL, Speer MC, Steele MP, Brown KK, Loyd JE, et al. A common MUC5B promoter polymorphism and pulmonary fibrosis. *N Engl J Med*. 2011;364:1503–12.
- Stock CJ, Sato H, Fonseca C, Banya WA, Molyneux PL, Adamali H, et al. Mucin 5B promoter polymorphism is associated with idiopathic pulmonary fibrosis but not with development of lung fibrosis in systemic sclerosis or sarcoidosis. *Thorax*. 2013;68:436–41.
- Borie R, Crestani B, Dieude P, Nunes H, Allanore Y, Kannengiesser C, et al. The MUC5B variant is associated with idiopathic pulmonary fibrosis but not with systemic sclerosis interstitial lung disease in the European Caucasian population. *PLoS One*. 2013;8:e70621.
- Peljto AL, Steele MP, Fingerlin TE, Hinchcliff ME, Murphy E, Podlasky S, et al. The pulmonary fibrosis-associated MUC5B promoter polymorphism does not influence the development of interstitial pneumonia in systemic sclerosis. *Chest*. 2012;142:1584–8.
- Fingerlin TE, Murphy E, Zhang W, Peljto AL, Brown KK, Steele MP, et al. Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. *Nat Genet*. 2013;45:613–20.

8. Noth I, Zhang Y, Ma SF, Flores C, Barber M, Huang Y, et al. Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: a genome-wide association study. *Lancet Respir Med*. 2013;1:309–17.
9. Radstake TR, Gorlova O, Rueda B, Martin JE, Alizadeh BZ, Palomino-Morales R, et al. Genome-wide association study of systemic sclerosis identifies CD247 as a new susceptibility locus. *Nat Genet*. 2010;42:426–9.
10. Furst D, Khanna D, Matucci-Cerinic M, Clements P, Steen V, Pope J, et al. Systemic sclerosis - continuing progress in developing clinical measures of response. *J Rheumatol*. 2007;34:1194–200.
11. Ramirez A, Varga J. Pulmonary arterial hypertension in systemic sclerosis: clinical manifestations, pathophysiology, evaluation, and management. *Treat Respir Med*. 2004;3:339–52.
12. Steen VD, Medsger TA. Changes in causes of death in systemic sclerosis, 1972-2002. *Ann Rheum Dis*. 2007;66:940–4.
13. Gorlova O, Martin JE, Rueda B, Koeleman BP, Ying J, Teruel M, et al. Identification of novel genetic markers associated with clinical phenotypes of systemic sclerosis through a genome-wide association strategy. *PLoS Genet*. 2011;7:e1002178.
14. Allanore Y, Saad M, Dieude P, Avouac J, Distler JH, Amouyel P, et al. Genome-wide scan identifies TNIP1, PSORS1C1, and RHOB as novel risk loci for systemic sclerosis. *PLoS Genet*. 2011;7:e1002091.
15. Wu M, Mayes MD. Insights into the genetic basis of systemic sclerosis: immunity in human disease and in mouse models. *Adv Genomics Genet*. 2014;4:143–51.
16. Spagnolo P, Grunewald J, du Bois RM. Genetic determinants of pulmonary fibrosis: evolving concepts. *Lancet Respir Med*. 2014;2:416–28.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit



Dissecting the Heterogeneity of Skin Gene Expression Patterns in Systemic Sclerosis

Shervin Assassi,¹ William R. Swindell,² Minghua Wu,¹ Filemon D. Tan,¹ Dinesh Khanna,² Daniel E. Furst,³ Donald P. Tashkin,³ Richard R. Jahan-Tigh,¹ Maureen D. Mayes,¹ Johann E. Gudjonsson,² and Jeffrey T. Chang¹

Objective. To examine the heterogeneity of global transcriptome patterns in systemic sclerosis (SSc) skin in a large sample of patients with SSc and control subjects.

Methods. Skin biopsy specimens obtained from 61 patients enrolled in the Genetics versus Environment in Scleroderma Outcome Study (GENISOS) cohort and 36 unaffected control subjects with a similar demographic background were examined by Illumina HumanHT-12 bead arrays. Followup experiments using quantitative polymerase chain reaction and immunohistochemical analysis were also performed.

Results. We identified 2,754 differentially expressed transcripts in SSc patients compared with controls. Clus-

tering analysis revealed 2 prominent transcriptomes in SSc patients: the keratin and fibroinflammatory signatures. Higher keratin transcript scores were associated with shorter disease duration and interstitial lung disease, while higher fibroinflammatory scores were associated with diffuse cutaneous involvement, a higher skin score at the biopsy site, and a higher modified Rodnan skin thickness score. A subgroup of patients with significantly longer disease duration had a normal-like transcript pattern. Analysis of cell type-specific signature scores revealed remarkable heterogeneity across patients. Significantly higher scores were calculated for fibroblasts (72% of patients), microvascular cells (61%), macrophages (54%), and dendritic cells (DCs) (49%). The majority of samples with significantly higher fibroblast scores (35 of 44 [80%]) had significantly increased macrophage and/or DC scores. Further analysis and immunohistochemical staining indicated that the keratin signature was not a general marker of keratinocyte activation but was in fact associated with an activation pattern in hair and adnexal structures.

Conclusion. Prominent fibroinflammatory and keratin signatures are present in SSc skin. Expression profiles of SSc skin show significant heterogeneity, and this finding might be useful for stratifying patients for targeted therapies or predicting the response to immunosuppression.

Systemic sclerosis (SSc; scleroderma) is a multi-system autoimmune disease associated with high morbidity and mortality (1). Global gene expression profiling with microarrays allows an unbiased genome-

Supported by the NIH (grants KL2-RR-024149 and K23-AR-061436 to Dr. Assassi, K08-AR-060802 to Dr. Gudjonsson, and UL1-TR-000371), the Scleroderma Foundation (SCORE Grant to Dr. Assassi), the A. Alfred Taubman Medical Research Institute (Kenneth and Frances Eisenberg Emerging Scholar Award to Dr. Gudjonsson), the Doris Duke Foundation (grant 2013106 to Dr. Gudjonsson), and the US Department of Defense (contract W81XWH-13-1-0452 to Dr. Mayes).

¹Shervin Assassi, MD, MS, Minghua Wu, MD, PhD, Filemon D. Tan, MD, PhD, Richard R. Jahan-Tigh, MD, Maureen D. Mayes, MD, MPH, Jeffrey T. Chang, PhD: University of Texas Health Science Center at Houston; ²William R. Swindell, PhD, MS, Dinesh Khanna, MD, MS, Johann E. Gudjonsson, MD, PhD: University of Michigan, Ann Arbor; ³Daniel E. Furst, MD, Donald P. Tashkin, MD: University of California, Los Angeles.

Dr. Khanna has received consulting fees from Actelion, Bayer, Bristol-Myers Squibb, Cytori, Roche/Genentech, InterMune, Lycera, Merck, and Seattle Genetics (less than \$10,000 each). Dr. Furst has received consulting fees, speaking fees, and/or honoraria from AbbVie, Actelion, Bristol-Myers Squibb, Amgen, Gilead, GlaxoSmithKline, Novartis, Pfizer, Roche/Genentech, UCB, Cytori, and Janssen (less than \$10,000 each). Dr. Mayes has received consulting fees, speaking fees, and/or honoraria from Practice Point Communications, Medtelligence, Boehringer-Ingelheim, and Cytori (less than \$10,000 each), receives royalties from Oxford University Press for *The Scleroderma Book*, from Harry Stewart Talks for a video of a prerecorded lecture on scleroderma, and from the British Medical Journal for a monograph on scleroderma, and has served as a paid consultant to the Gerson-Lehrman Group and to Leerink Partners, investment analysis companies, regarding systemic sclerosis unmet needs. Dr. Gudjonsson has received consulting fees,

speaking fees, and/or honoraria from Novartis and AnaptysBio (less than \$10,000 each) and owns stock or stock options in NuMedii.

Address correspondence to Shervin Assassi, MD, MS, Division of Rheumatology, University of Texas Health Science Center at Houston, 6431 Fannin Street, MSB 5.266, Houston, TX 77030. E-mail: shervin.assassi@uth.tmc.edu.

Submitted for publication June 15, 2014; accepted in revised form July 14, 2015.

wide assessment of the transcript dysregulation in a given tissue. This technology is now increasingly used to fingerprint pathologic processes, stratify diseases at the molecular level, and predict disease outcome (2–6). Development of effective treatment options in SSc has been hampered by a lack of sufficient understanding of its pathophysiology. Global gene expression studies in SSc at the end-organ level (7–10) or in peripheral blood cells (11–15) have indicated the presence of distinct transcript patterns in the majority of patients.

Skin is a prominently affected and easily accessible end-organ in SSc. Previous global gene expression studies have shown that SSc skin has a distinct gene expression profile, with inflammatory as well as fibrotic signatures (7,10). In a larger study involving 24 patients with SSc, the subgroup of patients with diffuse cutaneous SSc (dcSSc) could be divided into 3 distinct groups and the patients with limited cutaneous SSc (lcSSc) into 2 groups based on the intrinsic gene expression profiles observed in their skin biopsy specimens. A subgroup of patients including those with dcSSc and those with lcSSc showed an inflammatory pattern. Another subgroup of SSc patients with diffuse skin involvement showed a proliferative gene expression profile. A third subgroup of patients had a “normal-like” gene expression signature (9). Two other studies by the same group of investigators confirmed the presence of these 3 intrinsic subsets in independent samples (inflammatory, proliferative, and normal-like) (16,17).

In the present study, we investigated the heterogeneity of SSc transcript profiles in a large, well-characterized sample using a comprehensive gene expression profiling platform. First, we examined the impact of skin status (affected versus unaffected) on the SSc gene expression profile. Second, we identified 2 gene expression signatures in our large data set and examined the relationship between these signatures and SSc clinical features and previously described intrinsic transcript signatures (9). Last, we investigated the specific contribution of different cell types present in the skin (e.g., fibroblasts or macrophages) to the observed transcript heterogeneity in SSc skin.

PATIENTS AND METHODS

Patients and control subjects. Sixty-one patients were recruited from the Genetics versus Environment in Scleroderma Outcome Study (GENISOS) (18) or at the baseline visit in an investigator-initiated, open-label phase I/IIa study of imatinib (19). In addition to samples obtained at baseline, followup samples were also obtained from 5 of the 61 patients enrolled in the GENISOS cohort. All patients fulfilled the American College of Rheumatology/European League Against Rheuma-

tism classification criteria for SSc (20). We also investigated 36 control subjects who had a similar demographic background. The study protocol was approved by the institutional review boards of the participating institutions, and all participants provided fully informed, voluntary consent.

Disease type was defined based on the extent of skin involvement (21). Patients who had diffuse skin involvement at any time during their disease course were categorized as having dcSSc. The modified Rodnan skin thickness score (MRSS) (22) was used to assess skin involvement, and the skin score at the biopsy site was also recorded (0–3 scale, where 0 = not involved and 3 = severe thickening). The presence of moderate to severe interstitial lung disease (ILD) was defined as a forced vital capacity of <70% predicted and findings indicative of pulmonary fibrosis on high-resolution computed tomography of the chest. Patients were considered as being treated with immunosuppressive agents if they had received immunosuppressive medication within 4 weeks of skin biopsy (with the exception of prednisone at a dose of ≤ 5 mg or hydroxychloroquine).

Skin biopsy and quantitative reverse transcription-polymerase chain reaction (qRT-PCR) analysis. Three-millimeter punch biopsy samples of skin were obtained from the arms of the study subjects and were immediately immersed in RNAlater solution (Qiagen) and stored at -80°C . RNA was extracted using RNeasy Fibrous Tissue kits (Qiagen). Global gene expression was assessed using Illumina HumanHT-12 bead arrays. All microarray experiments were performed in a single batch (see Supplementary Methods, available on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39289/abstract>). Microarray data from this study are available from NCBI GEO (<http://www.ncbi.nlm.nih.gov/geo/>) using accession no. GSE58095. There is overlap with accession no. GSE47162, but the previous data set does not include any control samples.

Quantitative RT-PCR was also performed for 2 prominently overexpressed keratin transcripts, keratin 25 (type I keratin) and keratin 85 (type II keratin) to confirm the microarray results. In these experiments, the expression values were normalized to those of GAPDH. Relative quantification was performed using the C_t method, where $\Delta\Delta C_t$ values were calculated based on GAPDH and transcript levels in controls.

Immunohistochemical analysis. Immunohistochemical analysis for 2 general markers of epidermis activation (keratin 6 and keratin 16) and keratin 85 (which was prominently overexpressed in our samples) was performed in skin biopsy samples obtained from 5 SSc patients with the keratin transcript signature and from unaffected control subjects (matched for age, sex, and ethnicity), as well as in a biopsy sample from psoriatic skin (positive control) (see Supplementary Methods, available on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39289/abstract>).

Microarray data analysis. Raw data were analyzed with BRB ArrayTools (<http://linus.nci.nih.gov/BRB-ArrayTools.html>). Data were normalized according to the quantile method. Genes whose log intensity variance was in the bottom 75th percentile were filtered out, and 11,819 transcripts met this criterion. Differentially expressed genes were detected using Significance Analysis of Microarrays (SAM), at a false discovery rate of <5% (23). The sets of differentially expressed genes were also modeled using Ingenuity Pathways Analysis (IPA) software, and upstream regulators were identified. The goal of Upstream Regulator Analysis in IPA is to identify upstream

Table 1. Clinical correlates of keratin, fibroinflammatory, and normal-like signatures*

Clinical variable	Keratin score		Fibroinflammatory score			Clustering with controls	
	r_s	P^\dagger	r	b (95% CI)	P	b or OR (95% CI)	P
Male sex	NA	0.658	NA	0.15 (−0.15, 0.46)	0.319	1.22 (0.34, 4.37)	0.820
Diffuse cutaneous involvement	NA	0.179	NA	0.39 (0.12, 0.67)	0.006	0.46 (0.14, 1.54)‡	0.212
Immunosuppressive agents	NA	0.902	NA	0.11 (−0.41, 0.2)	0.488	0.68 (0.18, 2.62)‡	0.513
Interstitial lung disease	NA	0.017	NA	0.14 (−0.14, 0.41)	0.33	0.43 (0.11, 1.65)‡	0.183
Affected skin at biopsy site	NA	0.658	NA	0.54 (0.29, 0.78)	<0.001	0.46 (0.14, 1.51)‡	0.197
Disease duration	−0.32	0.02	−0.19	−0.02 (−0.05, 0.01)	0.165	3.22 (0.11, 6.34)‡	0.043
Concurrent MRSS	0.03	0.85	0.66	0.03 (0.02, 0.04)	<0.001	−5.02 (−11.28, 1.24)	0.114
Skin score at biopsy site	0.01	0.98	0.55	0.31 (0.18, 0.43)	<0.001	−0.47 (−1.02, 0.09)	0.1

* 95% CI = 95% confidence interval; NA = not applicable; MRSS = modified Rodnan skin thickness score.

† Calculated using Wilcoxon's rank sum test, because the keratin transcript scores were strongly right skewed.

‡ Odds ratio (OR).

regulators and predict whether they are activated or inhibited, given the observed gene expression changes in the experimental data set. IPA uses a Z score algorithm to make predictions. The Z score algorithm is designed to reduce the chance that random data will generate significant predictions. Upstream Regulator Analysis is based on expected causal effects between upstream regulators and targets; the expected causal effects are derived from the literature compiled in Ingenuity Knowledge Base.

Composite transcript scores for the keratin and fibroinflammatory signatures were also calculated. For this purpose, we normalized each transcript included in the profiles to mean = 0 and variance = 1. To derive the composite score, we averaged the expression values of the positively correlated genes with the inverse values of the negatively correlated genes. We also calculated a composite score for predicted activation of the Wnt/ β -catenin pathway based on a previously published data set, using the above algorithm (24). The statistical methods used to examine the relationship between these transcript composite scores and clinical outcomes depended on the distribution of composite scores (dependent variable). The distribution of the keratin scores was right skewed. Therefore, nonparametric methods were used for analysis of the keratin scores. Specifically, Spearman's rho was used if the independent variable was continuous, and Wilcoxon's rank sum test was used if the independent variable was categorical. Parametric tests were used for analysis of the composite fibroinflammatory scores (Pearson's correlation and linear regression).

Analysis of cell type-specific expression. Skin consists of a complex set of cell types, each of which contributes to the aggregate expression data generated from a given biopsy. We performed cell type-specific expression analysis according to the approach previously described by our group (25), to investigate the potential contribution of these cell types to the SSc skin transcript profile (see Supplementary Methods, available on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39289/abstract>).

RESULTS

Sixty-one patients with SSc and 36 control subjects were assessed in this study. The demographic and clinical characteristics of the study population are shown in Supplementary Table 1, available on the *Arthritis &*

Rheumatology web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39289/abstract>.

Skin thickening at the biopsy site is a prominent source of heterogeneity. The transcriptome of the majority of SSc patients differed from that of control subjects. There were 2,754 differentially expressed genes in SSc patients compared with controls (see Additional Tables 1 and 2, available on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39289/abstract>). IPA revealed that the top 3 represented canonical pathways belonged to hepatic fibrosis, agranulocyte adhesion/diapedesis, and granulocyte adhesion/diapedesis. The top significantly activated upstream transcription regulators were interferon (IFN) regulatory factor 7 (IRF-7), NF- κ B, IRF-1, STAT-1, and tumor protein 53, while the top upstream activated cytokines/growth factors were transforming growth factor β 1 (TGF β 1), IFN α , oncostatin M, IFN α 2, and IFN γ (see also Supplementary Figure 1 and Additional Table, available on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39289/abstract>). Next, we examined the SSc gene profile based on the status of skin involvement at the biopsy site (Figure 1); a complete list of differentially expressed transcripts in all comparison groups is shown in Additional Table 1 (available on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39289/abstract>).

A comparison of patients with clinically affected skin at the biopsy site and control subjects revealed 4,184 differentially expressed transcripts belonging to hepatic fibrosis, agranulocyte adhesion/diapedesis, and dendritic cell (DC) maturation, while a comparison of SSc patients with clinically unaffected skin at the biopsy site and controls resulted in only 24 differentially expressed genes (the small number of differentially expressed transcripts in this comparison precluded a meaningful pathway analysis). Furthermore, we detected prominent gene expression differences between SSc

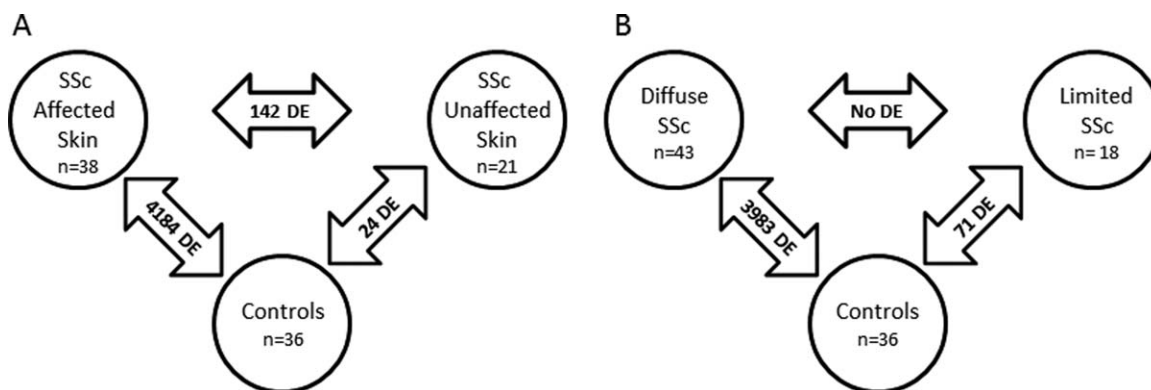


Figure 1. Numbers of differentially expressed (DE) genes in systemic sclerosis (SSc) patients with affected skin at the biopsy site compared with those with unaffected skin at the biopsy site and control subjects (A) and in patients with diffuse SSc compared with those with limited SSc and control subjects (B).

patients with clinically affected skin at the biopsy site and those with unaffected skin at the biopsy site. This comparison resulted in 142 differentially expressed transcripts belonging to the top 3 overrepresented canonical pathways.

We also performed a parallel analysis in order to investigate whether the heterogeneity based on the status of skin at the biopsy site is driven by disease type. In a comparison between patients with dcSSc and controls, there were 3,983 differentially expressed transcripts belonging to hepatic fibrosis, DC maturation, and graft-versus-host signaling, while 71 differentially expressed transcripts were detected in a comparison between patients with lcSSc and controls, with overrepresentation of the hepatic fibrosis pathway. Of note, there were no differentially expressed transcripts when patients with dcSSc were compared with those with lcSSc, indicating that the status of skin at the biopsy site is a more prominent source of heterogeneity than is disease type. Of note, a pairwise comparison between affected and unaffected skin within the same individual SSc patients, as previously performed (9,10,16,17), could not be conducted in our study, because biopsy samples obtained from other typically uninvolved anatomic areas such as the buttock area were not performed.

Differentially expressed gene expression profiles in SSc patients. Figure 2 shows the unsupervised hierarchical clustering of 2,754 differentially expressed genes identified in the comparison between SSc samples and controls (for additional information see Supplementary Figure 2, available on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39289/abstract>). Two partially overlapping gene clusters were overexpressed in SSc patients (Figure 2). The first gene cluster was highly enriched with keratins

and keratin-associated proteins (mainly hair and adnexal structure keratins), with 77 of 93 transcripts (82.8%) being keratin-related (see Additional Table, available on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39289/abstract>). Thirty-four patients (57.6%), including those with clinically affected skin and those with unaffected skin (see Supplementary Figure 3, available on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39289/abstract>), as well as both patients with lcSSc and patients with dcSSc, clustered in the group with overexpression of the keratin signature (data not shown). The keratin signature was significantly less common among control subjects (present in only 8 controls [22.2%]; $P = 0.002$). We also investigated levels of 2 representative keratin transcripts, keratin 25 and keratin 85, by qRT-PCR. As shown in Supplementary Figure 4 (available on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39289/abstract>), levels of both keratin 25 (type I keratin) and keratin 85 (type II keratin) were significantly higher in patients compared with controls (median fold changes 21.1 [$P = 0.015$] and 15.7 [$P = 0.003$], respectively).

The second overexpressed cluster was enriched with genes involving inflammatory (including IFN-inducible genes) and fibrotic pathways, such as *THY1*, *COL1A1*, *COMP*, *OAS1*, and *CCL2* (Figure 2; see also Additional Table, available on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39289/abstract>). As shown in Supplementary Figure 5 (available on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39289/abstract>), this fibroinflammatory signature was present in 44 patients (74.6%), while only 5 control subjects had this signature ($P < 0.001$). IPA of this cluster

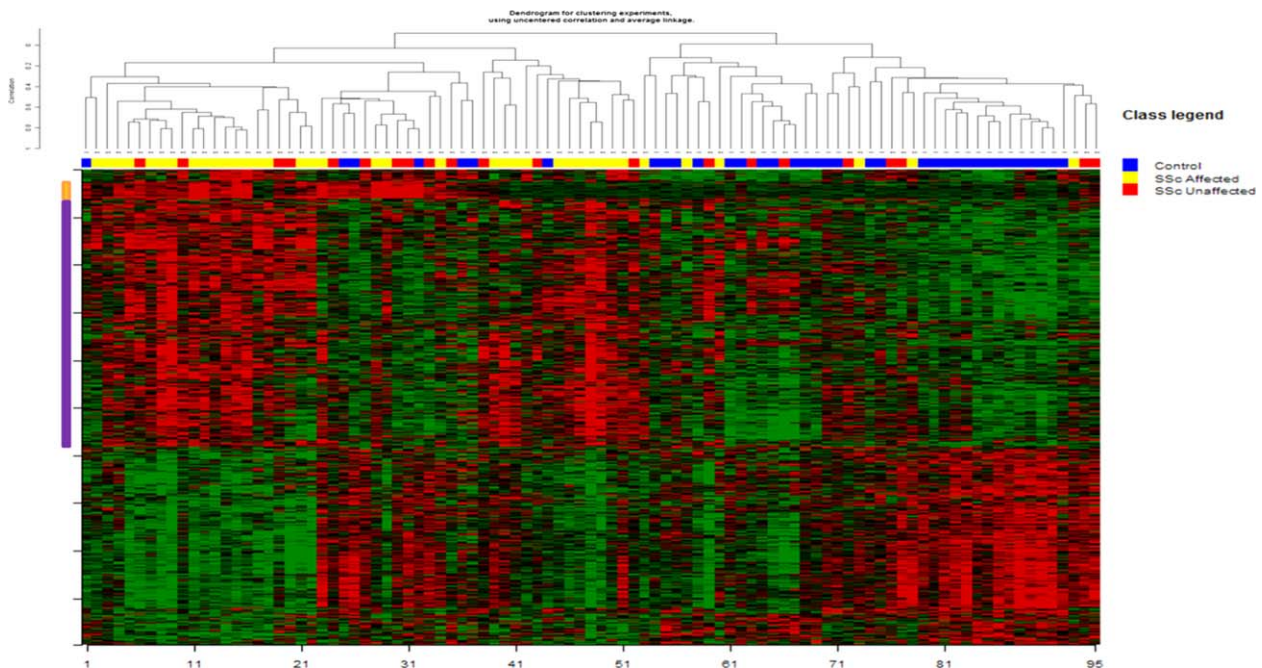


Figure 2. Unsupervised hierarchical clustering of skin samples based on the 2,754 transcripts that were differentially expressed in clinically affected systemic sclerosis (SSc) skin, unaffected SSc skin, and control skin. The orange bars and purple bars on the left indicate keratin and fibroinflammatory gene clusters, respectively.

revealed that the top 3 overrepresented canonical pathways were hepatic fibrosis, agranulocyte adhesion/diapedesis, and DC maturation. In this cluster, genes belonging to inflammatory and fibrotic pathways were coexpressed. As shown in Figure 2, this gene cluster partially overlapped with the keratin signature. Similar to previously published data (9), 14 SSc samples (23.7%) clustered with control samples (Figure 2; see also Supplementary Figure 2, available on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39289/abstract>).

Cross comparison with the inflammatory and proliferative intrinsic subsets. We mapped the inflammatory intrinsic subset (60 genes corresponding to 93 probes) and the proliferative intrinsic subset (56 genes corresponding to 86 probes) previously described by Milano et al (9) to probes present on Illumina HumanHT-12 arrays. As shown in Supplementary Figure 6 (available on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39289/abstract>), some patient samples showed increased coexpression of transcripts in the inflammatory intrinsic subset, but we did not observe homogeneous overexpression of proliferative intrinsic genes among patient samples (see Supplementary Figure 7, available on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39289/abstract>). We also investigated the overlap between the above-mentioned

fibroinflammatory cluster and these 2 intrinsic gene subsets. Twenty inflammatory intrinsic genes (23 probes) were also present in our fibroinflammatory cluster, whereas only 5 genes were in common between the proliferative intrinsic gene subset and our fibroinflammatory gene cluster.

We also performed a hierarchical clustering analysis using unbiased selection of genes that deviated at least 2-fold from the mean in at least 5 samples in order to parallel the analytic approach described in previous SSc skin gene expression studies (9,10). Following this approach, 1,870 genes were identified. As shown in Supplementary Figure 8 (available on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39289/abstract>), unsupervised hierarchical clustering using this gene list demonstrated that a subgroup of SSc samples (especially those from patients with affected skin) clustered together. Furthermore, there were clusters of genes that corresponded to the samples with an intrinsic inflammatory signature (9) and the above-described keratin signature, but samples with a prominent intrinsic proliferative signature (9) and their corresponding genes could not be detected. Furthermore, 13 of 58 genes in the inflammatory intrinsic signature overlapped with our list of 1,870 transcripts ($P < 0.001$, odds ratio [OR] 3.48, 95% confidence interval [95% CI] 1.89–6.41), while only 1 of 55 genes in the proliferative intrinsic signature overlapped with this

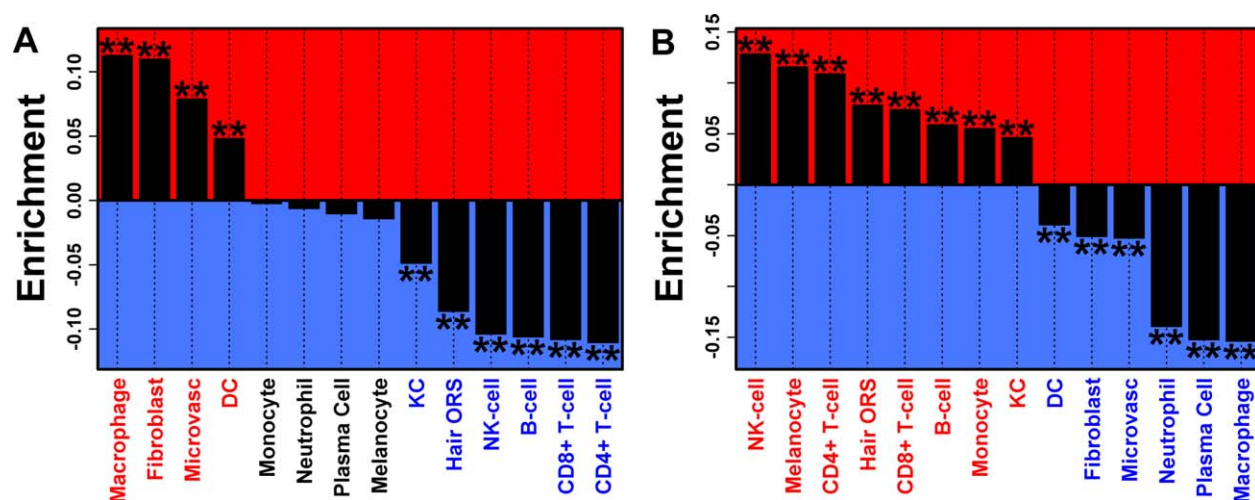


Figure 3. Genes differentially expressed in systemic sclerosis (SSc) skin samples (versus control samples). Enrichment with genes specifically expressed in certain cell types is shown. Genes were ranked according to the degree to which they were specifically expressed in each of 14 cell types. Enrichment statistics quantify the degree to which genes that were significantly increased (A) or decreased (B) in SSc skin samples are among the genes specifically expressed with respect to each cell type. Positive enrichment statistics denote enrichment of differentially expressed genes with respect to genes most specifically expressed in a given cell type. Negative enrichment statistics denote enrichment of differentially expressed genes with respect to genes showing specifically low expression in a given cell type. ** = $P < 0.001$ by Wilcoxon's rank sum test. DC = dendritic cell; KC = keratinocyte; ORS = outer root sheath; NK = natural killer.

gene list ($P = 0.127$, OR 0.32, 95% CI 0.06–1.61). This indicates that the previously described inflammatory intrinsic signature (9) had significant overlap with our list of 1,870 genes, while the proliferative intrinsic genes did not show a significant overlap with the genes that have highly varying levels in our data set.

Correlates of keratin and fibroinflammatory signatures. Next, we calculated composite scores for the keratin and fibroinflammatory clusters. As shown in Supplementary Figure 9 (available on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39289/abstract>), both clusters showed higher composite scores in patients compared with controls ($P < 0.001$ for the keratin cluster and $P < 0.001$ for the fibroinflammatory clusters).

Table 1 shows clinical correlates of these 2 composite transcript scores. A higher composite keratin score was associated with shorter disease duration and the presence of ILD, while a higher composite fibroinflammatory score was associated with a higher MRSS, local skin score, and diffuse cutaneous involvement. Composite scores were not associated with SSc-related antibodies (data not shown) or treatment with immunosuppressive agents.

Examination of clinical correlates in the 14 SSc patients with a “normal-like signature” revealed that they had a significantly longer disease duration. In addition, a composite score for predicted activation of the Wnt/ β -cat-

enin pathway in our data set based on a previously published gene list (24) was calculated. Both the keratin and fibroinflammatory composite scores were significantly correlated with the Wnt/ β -catenin composite scores ($r_s = 0.29$ [$P = 0.004$] and $r = 0.56$ [$P < 0.001$], respectively).

Analysis of cell type signature scores. As shown in Figure 3, the 1,604 overexpressed transcripts in SSc skin were enriched in genes specifically expressed in fibroblasts, macrophages, microvascular tissue, and DCs. Similarly, the 1,150 underexpressed transcripts were enriched in genes specifically expressed in lymphocytes (NK cells, CD4+ and CD8+ T cells, and B cells), and epidermis-related cell types (melanocytes, keratinocytes, and hair outer root sheath [ORS] cells). Again, calculation of cell type signature scores revealed significant heterogeneity across patients.

At the individual patient level (Figure 4), significantly higher scores were more frequently calculated for fibroblasts (72% of patients), indicating that genes specifically expressed in fibroblasts were increased in most of the biopsy specimens from patients with SSc (for additional information, see Supplementary Figure 10, available on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39289/abstract>). Similarly, in the majority of patients, we calculated significantly higher signature scores for microvascular and macrophage-related transcripts. Consistent with coexpression of fibrotic and inflammatory transcripts in our

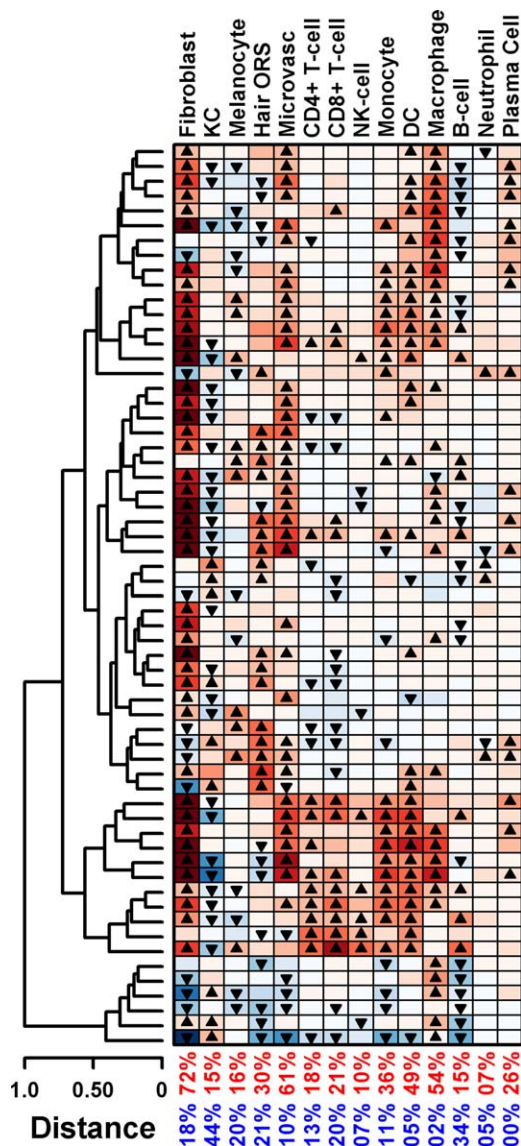


Figure 4. Hierarchical clustering of patients with systemic sclerosis based on cell type signature scores. Each row represents a patient sample. Transcript scores for each patient were calculated with respect to 14 cell types. Scores were calculated based on fold-change estimates for 125 signature genes per cell type (expression in patient samples/average expression in 36 control samples). Triangles denote scores that are significantly high (\blacktriangle) or low (\blacktriangledown) compared with all other human genes. Values in the 2 bottom rows are the percentages of patients with significantly high (red) or low (blue) scores. KC = keratinocyte; ORS = outer root sheath; NK = natural killer; DC = dendritic cell.

unsupervised hierarchical clustering experiments, the majority of samples with significantly higher fibroblast scores (35 of 44 [80%]) had significant macrophage and/or DC scores. Furthermore, the top cell-based profiles correlating with the fibroinflammatory signature in the initial clustering (Figure 2) were macrophage genes

($r = 0.75$, $P < 0.005$) and fibroblast genes ($r = 0.68$, $P < 0.001$). There was also a subgroup of patients with elevated expression of hair ORS-specific genes; the top cell-based profiles correlating with the keratin signature in the initial clustering (Figure 2) were hair ORS genes ($r_s = 0.44$, $P < 0.001$) and keratinocyte genes ($r_s = 0.38$, $P < 0.001$).

Longitudinal progression of the keratin signature. In an exploratory investigation, we examined the longitudinal progression of the above-mentioned signatures (Figure 2) in 5 patients. The keratin signature increased numerically over time in the 2 patients with early disease (disease duration < 2 years), while it declined numerically in the remainder of the patients (see Supplementary Figure 11, available on the *Arthritis & Rheumatology* web site at <http://onlinelibrary.wiley.com/doi/10.1002/art.39289/abstract>). The fibroinflammatory score showed less variation over time, although a downward trend over time was observed in the majority of samples (see Supplementary Figure 11). None of the observed changes reached statistical significance, which might be attributable to the small sample size ($n = 5$).

Immunohistochemical studies of the keratin signature. As shown in Figure 5, keratin 6 and keratin 16 staining in SSc skin samples (5 with a keratin signature) and matched controls was confined to hair follicles

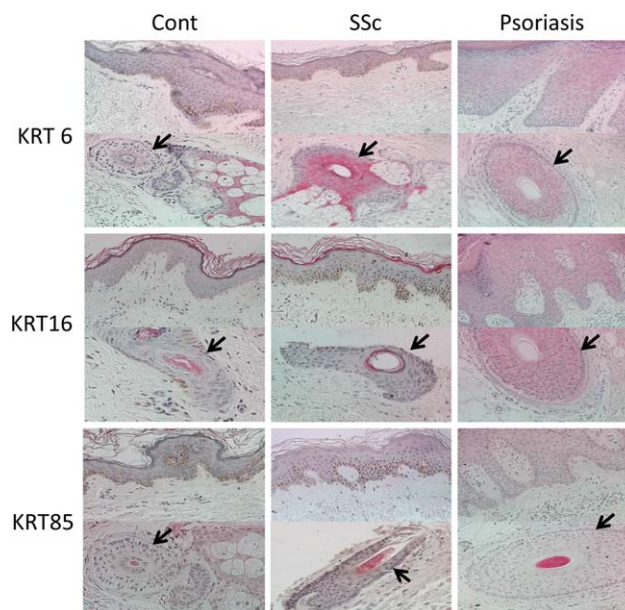


Figure 5. Representative images showing keratin 6 (KRT6), KRT16, and KRT85 staining of skin tissue from patients with systemic sclerosis (SSc) with a keratin transcript signature ($n = 5$) and unaffected control subjects ($n = 5$). Psoriatic skin was used as a positive control for KRT6 and KRT16. Arrows indicate hair follicles. Original magnification $\times 400$.

and eccrine sweat glands, whereas there was diffuse staining of the epidermis in a sample of psoriatic skin (positive controls). The keratin 85 staining was confined to hair matrix/pre cortex and hair cuticle in all 3 sample types (SSc, psoriasis, and controls) and was not present in the epidermis.

DISCUSSION

In the present study, we investigated the global gene expression profile in a large sample of patients with SSc and control subjects, using a comprehensive microarray platform to dissect the heterogeneity of transcriptome patterns in affected SSc skin. SSc skin samples showed prominent fibroinflammatory and keratin transcript profiles that correlated with certain disease features. Furthermore, we performed a cell-based modular analysis that showed substantial heterogeneity in the inflammatory profile of SSc skin. These findings may have important implications for identification of therapeutic targets and development of biomarkers.

A prominent fibroinflammatory signature correlating with the MRSS and the local skin score was present in the majority of SSc skin samples. Hierarchical clustering and cell type-specific signature scores indicated that inflammatory and fibrotic signatures coexist in most patients. As expected, the majority of patients in the current study (72%) displayed significant up-regulation of the fibroblast cell-type signature, while the overall inflammatory profile of SSc patients showed significant heterogeneity. This heterogeneity might provide important information for stratifying patients for targeted therapies and/or responses to general immunosuppression. The DC/macrophage modules were the most commonly up-regulated inflammatory modules (>50%), whereas NK cell, CD4, and CD8 T cell modules were up-regulated in only 10%, 18%, and 21% of cases, respectively. A similar analysis in psoriatic plaques revealed a more inflammatory and less fibrotic profile. In patients with psoriasis, the majority of samples displayed heightened macrophage, DC, NK cell, and CD8 T cell signatures, while only 9% of samples showed up-regulation of the fibroblast signature (25). This finding is also clinically plausible, because psoriatic plaques are more responsive to immunosuppression than is SSc skin.

Full-thickness skin biopsy specimens consist of a heterogeneous collection of cell types, each of which contributes to the aggregate expression measurement for any individual gene. However, few previous global gene expression studies in SSc skin have included analyses to relate the observed transcript heterogeneity to cell-specific signatures. Whitfield et al (10) measured

gene expression in 11 different cell lines grown in culture that represent cell types likely to be present in skin. A comparison of the transcript profile of these cell lines with skin samples revealed that SSc skin had prominent dysregulation of fibroblast- and endothelial-related genes. Gardner et al (7) examined the gene expression profile of SSc skin and concomitantly collected explanted passage 4 fibroblasts. A comparison of skin and fibroblast transcriptomes showed that a subgroup of differentially expressed transcripts are likely to be of fibroblast origin, while other cell types were also required for full expression of the SSc phenotype. Composite scores for the cell-specific signatures were not calculated in either of those studies.

Pendergrass et al (17) also performed a cell type-specific analysis based on 2 previously published data sets (26,27). Similar to our analysis, a composite score was calculated for cell-based gene signatures. However, we have here calculated a rank-based signature score statistic while using a novel and large collection of data samples to identify signature genes for each cell type (25,28,29). Pendergrass et al also observed significant heterogeneity in the SSc skin inflammatory profile. However, a cross-comparison between their specific cell type signatures and our results is difficult, because the cell types were defined differently. Similar to our results, a prominent fibroblast signature in patient samples was observed. The most prominent inflammatory cell signature in the Pendergrass study was the granulocyte signature, while the macrophage signature was the most prominent inflammatory signature in our study. Those investigators also reported a combined signature for myeloid cells that was present in the majority of patients.

We also investigated the presence of previously described intrinsic inflammatory and fibrotic signatures (9,16,17) in our data set. A subset of SSc samples showed an increased coexpression pattern for the intrinsic inflammatory transcripts, but this was not observed for the intrinsic proliferative transcripts. In the present study, we not only included a larger number of patients but also investigated a substantially higher number of control subjects with a similar demographic background (36 control subjects in the present study versus 6–10 in previous studies [9,16,17]). This larger number increased the power to detect differentially expressed transcripts and better reflects the heterogeneity within each study population (patients and controls) during the clustering analysis. Furthermore, we used a platform that includes 54% more genes (30,500 versus 19,800), allowing us to perform a more comprehensive investigation of SSc skin. For example, only 55.1% of keratin signature genes were

present in the previously used platform (16,17). Based on our results, it seems prudent to continue broader molecular profiling of SSc skin to characterize the transcript dysregulations most relevant for understanding disease mechanisms, creating subsets, identifying therapeutic targets, and developing biomarkers.

We observed a prominent keratin signature in SSc skin, which is a novel finding. Previous research mainly focused on molecular dysregulations in the dermal layer of SSc skin, although phenotypic changes in the skin such as hypopigmentation or hyperpigmentation also imply involvement of the epidermis (30). Previous studies have also indicated overexpression of the key cytokines TGF β (31), monocyte chemoattractant protein 1 (32), vascular endothelial growth factor (33), and interleukin-2 receptor (34) in SSc epidermis. A proteomic analysis of lesional scleroderma skin showed prominent dysregulation of proteins specific to epidermal differentiation in addition to those involved in extracellular matrix production and myofibroblast contractility (35). In a followup study, epidermal keratinocyte maturation was delayed, and an activation pattern with up-regulation of keratin 6 and keratin 16 was observed in both clinically involved and uninvolved skin of patients with early dcSSc. Furthermore, coculture of epidermis from SSc patients and normal human fibroblasts promoted fibroblast contractility to the extent observed with TGF β , while epidermis from healthy controls did not have a similar effect (36).

Similar to the above-mentioned study (36), the keratin transcript signature was present in involved and uninvolved skin in the present study. However, our immunohistochemical staining results did not indicate that the keratin signature that we observed was a marker of general activation of keratinocytes, as seen in psoriatic skin (25). Specifically, keratin 6 and keratin 16, both of which are usually expressed only in hair and eccrine sweat glands but show broad epidermal staining in diseases with general activation of keratinocytes (e.g., psoriasis), did not show a broad epidermal staining pattern in SSc patients with the keratin signature. The keratin signature consisted mainly of up-regulation of hair- and adnexal structure-related keratins. Consistent with this notion, keratin 85 staining was confined to hair follicles and eccrine sweat glands. Of note, we do not believe that the observed keratin signature is due to the higher number of hair follicles in SSc patients, because the same anatomic site was biopsied in patients and controls. Furthermore, there was no association between the keratin signature and male sex.

The observed association between the keratin signature and shorter disease duration supports the

notion that this transcript profile is a time-dependent occurrence. The observed keratin signature might be a response of the hair follicle to the molecular changes in SSc skin. Hair follicle morphogenesis and growth are regulated through complex and reciprocal epithelial mesenchymal interactions. The dermal papilla is essential for follicle formation in the embryo and for initiating the next period of hair growth in adults (37). Activation of Wnt/ β -catenin signaling is an important stimulator of new hair growth and differentiation (38–40). In our data set, the keratin signature correlated with the predicted activation of the Wnt/ β -catenin pathway, supporting the notion that the up-regulation of this pathway in SSc (41) might contribute to the observed keratin signature. The observed association of the keratin signature with ILD might also stem from common upstream regulators for ILD and this transcript signature. Further studies are needed to investigate the potential role of this signature in the pathogenesis of SSc. In general, our results provide further evidence for the presence of molecular dysregulation in several skin compartments (not only dermis) in SSc.

Similar to what was observed in previous studies, a subset of patients showed normal-like gene expression profiles (7,9,16,17). These patients had a significantly longer disease duration, and a trend for association of a normal-like signature with longer disease duration was observed in a previous SSc skin data set (17). This finding is also supported by the decreasing composite fibroinflammatory scores observed in our pilot longitudinal study (see Supplementary Figure 11, available on the *Arthritis & Rheumatology* web site at <http://online.library.wiley.com/doi/10.1002/art.39289/abstract>). Patients with a normal-like transcript profile also tended to have milder disease (lower MRSS and lower skin score at the biopsy site), although these findings did not reach statistical significance.

Patients in whom skin at the biopsy site was affected were more likely to have a distinct gene expression profile and a higher composite fibroinflammatory score. However, we were unable to compare the gene expression profile of skin biopsy specimens obtained from 1 affected area (e.g., the arm) and 1 unaffected area (usually the buttock or back) in the same individual, because biopsies of the buttocks area were not performed in our study. Furthermore, the number of differentially expressed genes in clinically affected versus unaffected SSc skin was substantially lower than the number of transcripts in affected SSc skin versus control skin, indicating substantial heterogeneity in the group with unaffected SSc skin. It is likely that a subgroup of

SSc patients with clinically unaffected skin have histologic changes characteristic of SSc.

In conclusion, this large global SSc skin gene expression study indicated the presence of prominent fibroinflammatory and keratin profiles. A subset of patients with a longer disease duration had a normal-like gene expression profile. Furthermore, analysis of cell type-specific signature scores revealed significant heterogeneity in the inflammatory profile of SSc skin, which might provide important information for substratifying patients to enable targeted therapies and to predict the response to immunosuppression.

ACKNOWLEDGMENTS

We thank Dr. Samuel Theodore for his assistance with the data collection and Julio Charles for performing the laboratory experiments. We are also grateful to Professor Lutz Langbein (University of Heidelberg, Germany) for providing the KRT85 antibodies.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Assassi had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Assassi, Tan, Furst, Mayes, Gudjonsson, Chang.

Acquisition of data. Assassi, Tan, Khanna, Furst, Tashkin, Mayes.

Analysis and interpretation of data. Assassi, Swindell, Wu, Tan, Jahan-Tigh, Mayes, Gudjonsson, Chang.

REFERENCES

- Elhai M, Meune C, Avouac J, Kahan A, Allanore Y. Trends in mortality in patients with systemic sclerosis over 40 years: a systematic review and meta-analysis of cohort studies. *Rheumatology (Oxford)* 2011;51:1017–26.
- Bullinger L, Dohner K, Bair E, Frohling S, Schlenk RF, Tibshirani R, et al. Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. *N Engl J Med* 2004;350:1605–16.
- Baechler EC, Batliwalla FM, Karypis G, Gaffney PM, Ortmann WA, Espe KJ, et al. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. *Proc Natl Acad Sci U S A* 2003;100:2610–5.
- Beer DG, Kardias SL, Huang CC, Giordano TJ, Levin AM, Misek DE, et al. Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med* 2002;8:816–24.
- Van Baarsen LG, Vosslander S, Tijssen M, Baggen JM, van der Voort LF, Killestein J, et al. Pharmacogenomics of interferon- β therapy in multiple sclerosis: baseline IFN signature determines pharmacological differences between patients. *PLoS One* 2008;3:e1927.
- Kandath C, Schultz N, Cherniack AD, Akbani R, Liu Y, Shen H, et al. Integrated genomic characterization of endometrial carcinoma. *Nature* 2013;497:67–73.
- Gardner H, Shearstone JR, Bandaru R, Crowell T, Lynes M, Trojanowska M, et al. Gene profiling of scleroderma skin reveals robust signatures of disease that are imperfectly reflected in the transcript profiles of explanted fibroblasts. *Arthritis Rheum* 2006;54:1961–73.
- Hsu E, Shi H, Jordan RM, Lyons-Weiler J, Pilewski JM, Feghali-Bostwick CA. Lung tissues in patients with systemic sclerosis have gene expression patterns unique to pulmonary fibrosis and pulmonary hypertension. *Arthritis Rheum* 2011;63:783–94.
- Milano A, Pendergrass SA, Sargent JL, George LK, McCalmont TH, Connolly MK, et al. Molecular subsets in the gene expression signatures of scleroderma skin. *PLoS One* 2008;3:e2696.
- Whitfield ML, Finlay DR, Murray JI, Troyanskaya OG, Chi JT, Pergamenschikov A, et al. Systemic and cell type-specific gene expression patterns in scleroderma skin. *Proc Natl Acad Sci U S A* 2003;100:12319–24.
- Assassi S, Mayes MD, Arnett FC, Gourh P, Agarwal SK, McNearney TA, et al. Systemic sclerosis and lupus: points in an interferon-mediated continuum. *Arthritis Rheum* 2010;62:589–98.
- Christmann RB, Hayes E, Pendergrass S, Padilla C, Farina G, Affandi AJ, et al. Interferon and alternative activation of monocyte/macrophages in systemic sclerosis-associated pulmonary arterial hypertension. *Arthritis Rheum* 2011;63:1718–28.
- Higgs BW, Liu Z, White B, Zhu W, White WI, Morehouse C, et al. Patients with systemic lupus erythematosus, myositis, rheumatoid arthritis and scleroderma share activation of a common type I interferon pathway. *Ann Rheum Dis* 2011;70:2029–36.
- Streicher K, Morehouse CA, Groves CJ, Rajan B, Pilataxi F, Lehmann KP, et al. The plasma cell signature in autoimmune disease. *Arthritis Rheumatol* 2014;66:173–84.
- Tan FK, Zhou X, Mayes MD, Gourh P, Guo X, Marcum C, et al. Signatures of differentially regulated interferon gene expression and vasculotrophism in the peripheral blood cells of systemic sclerosis patients. *Rheumatology (Oxford)* 2006;45:694–702.
- Hinchcliff M, Huang CC, Wood TA, Mahoney JM, Martyanov V, Bhattacharyya S, et al. Molecular signatures in skin associated with clinical improvement during mycophenolate treatment in systemic sclerosis. *J Invest Dermatol* 2013;133:1979–89.
- Pendergrass SA, Lemaire R, Francis IP, Mahoney JM, Lafyatis R, Whitfield ML. Intrinsic gene expression subsets of diffuse cutaneous systemic sclerosis are stable in serial skin biopsies. *J Invest Dermatol* 2012;132:1363–73.
- Assassi S, Sharif R, Lasky RE, McNearney TA, Estrada YM, Draeger HT, et al. Predictors of interstitial lung disease in early systemic sclerosis: a prospective longitudinal study of the GENI-SOS cohort. *Arthritis Res Ther* 2010;12:R166.
- Khanna D, Saggari R, Mayes MD, Abtin F, Clements PJ, Maranian P, et al. A one-year, phase I/IIa, open-label pilot trial of imatinib mesylate in the treatment of systemic sclerosis-associated active interstitial lung disease. *Arthritis Rheum* 2011;63:3540–6.
- Van den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum* 2013;65:2737–47.
- 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.
- 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.
- Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 2001;98:5116–21.
- Kenny PA, Enver T, Ashworth A. Receptor and secreted targets of Wnt-1/ β -catenin signalling in mouse mammary epithelial cells. *BMC Cancer* 2005;5:3.
- Swindell WR, Johnston A, Voorhees JJ, Elder JT, Gudjonsson JE. Dissecting the psoriasis transcriptome: inflammatory- and

- cytokine-driven gene expression in lesions from 163 patients. *BMC Genomics* 2013;14:527.
26. Palmer C, Diehn M, Alizadeh AA, Brown PO. Cell-type specific gene expression profiles of leukocytes in human peripheral blood. *BMC Genomics* 2006;7:115.
 27. Haider AS, Lowes MA, Suarez-Farinas M, Zaba LC, Cardinale I, Blumenberg M, et al. Cellular genomic maps help dissect pathology in human skin disease. *J Invest Dermatol* 2008;128:606–15.
 28. Swindell WR, Johnston A, Xing X, Voorhees JJ, Elder JT, Gudjonsson JE. Modulation of epidermal transcription circuits in psoriasis: new links between inflammation and hyperproliferation. *PLoS One* 2013;8:e79253.
 29. Swindell WR, Stuart PE, Sarkar MK, Voorhees JJ, Elder JT, Johnston A, et al. Cellular dissection of psoriasis for transcriptome analyses and the post-GWAS era. *BMC Med Genomics* 2014;7:27.
 30. Tabata H, Hara N, Otsuka S, Yamakage A, Yamazaki S, Koibuchi N. Correlation between diffuse pigmentation and keratinocyte-derived endothelin-1 in systemic sclerosis. *Int J Dermatol* 2000;39:899–902.
 31. Rudnicka L, Varga J, Christiano AM, Iozzo RV, Jimenez SA, Uitto J. Elevated expression of type VII collagen in the skin of patients with systemic sclerosis: regulation by transforming growth factor- β . *J Clin Invest* 1994;93:1709–15.
 32. Distler O, Pap T, Kowal-Bielecka O, Meyringer R, Guiducci S, Landthaler M, et al. Overexpression of monocyte chemoattractant protein 1 in systemic sclerosis: role of platelet-derived growth factor and effects on monocyte chemotaxis and collagen synthesis. *Arthritis Rheum* 2001;44:2665–78.
 33. Davies CA, Jeziorska M, Freemont AJ, Herrick AL. The differential expression of VEGF, VEGFR-2, and GLUT-1 proteins in disease subtypes of systemic sclerosis. *Hum Pathol* 2006;37:190–7.
 34. Distler JH, Jungel A, Kowal-Bielecka O, Michel BA, Gay RE, Sprott H, et al. Expression of interleukin-21 receptor in epidermis from patients with systemic sclerosis. *Arthritis Rheum* 2005;52:856–64.
 35. Aden N, Shiwen X, Aden D, Black C, Nuttall A, Denton CP, et al. Proteomic analysis of scleroderma lesional skin reveals activated wound healing phenotype of epidermal cell layer. *Rheumatology (Oxford)* 2008;47:1754–60.
 36. Aden N, Nuttall A, Shiwen X, de Winter P, Leask A, Black CM, et al. Epithelial cells promote fibroblast activation via IL-1 α in systemic sclerosis. *J Invest Dermatol* 2010;130:2191–200.
 37. Hardy MH. The secret life of the hair follicle. *Trends Genet* 1992;8:55–61.
 38. Lo CC, Prowse DM, Watt FM. Transient activation of β -catenin signalling in adult mouse epidermis is sufficient to induce new hair follicles but continuous activation is required to maintain hair follicle tumours. *Development* 2004;131:1787–99.
 39. Lowry WE, Blanpain C, Nowak JA, Guasch G, Lewis L, Fuchs E. Defining the impact of β -catenin/Tcf transactivation on epithelial stem cells. *Genes Dev* 2005;19:1596–611.
 40. Van Mater D, Kolligs FT, Dlugosz AA, Fearon ER. Transient activation of β -catenin signaling in cutaneous keratinocytes is sufficient to trigger the active growth phase of the hair cycle in mice. *Genes Dev* 2003;17:1219–24.
 41. Wei J, Fang F, Lam AP, Sargent JL, Hamburg E, Hinchcliff ME, et al. Wnt/ β -catenin signaling is hyperactivated in systemic sclerosis and induces Smad-dependent fibrotic responses in mesenchymal cells. *Arthritis Rheum* 2012;64:2734–45.



Published in final edited form as:

Semin Arthritis Rheum. 2015 June ; 44(6): 680–686. doi:10.1016/j.semarthrit.2014.11.006.

Antinuclear Antibody Negative Systemic Sclerosis

GA Salazar^a, S Assassi^a, F Wigley^b, L Hummers^b, J Varga^c, M Hinchcliff^c, D Khanna^d, E Schiopu^d, K Phillips^d, DE Furst^e, V Steen^f, M Baron^g, M Hudson^g, SS Taillefer^g, J Pope^h, N Jonesⁱ, P Docherty^j, NA Khalidi^k, D Robinson^l, R Simms^m, R Silverⁿ, TM Frech^o, B Fessler^p, J Molitor^q, M Fritzler^r, B Segal^q, F Al-Kassab^s, M Perry^a, J Yang^a, S Zamanian^a, JD Reveille^a, FC Arnett^a, C Pedroza^a, and MD Mayes^a

^aUniversity of Texas Medical School at Houston, Division of Rheumatology. 6431 Fannin St. MSB 5.270, Houston TX 77030, USA

^bJohns Hopkins University. 5501 Hopkins Bayview Circle. Asthma and Allergy Center. Baltimore, MD 21224, USA

^cFeinberg School of Medicine. Northwestern University. Dept. of Medicine. 251 E Huron St Galter Suite 3–150. Chicago, IL 60611. USA

^dUniversity of Michigan. Division of Rheumatology. Taubman Center Floor 3 Reception A. 1500 E Medical Center Dr SPC 5370. Ann Arbor, MI 48109. USA

^eUniversity of California Los Angeles. Division of Rheumatology. Box 951670, 32–59. Rehab. Los Angeles, CA 90095-1670. USA

^fGeorgetown University Medical Center. Division of Rheumatology. 4000 Reservoir Road NW, Washington, DC 20007. USA

^gMcGill University, Division of Rheumatology. Jewish General Hospital. Suite A 725 3755 Cote St, Catherine Rd Montreal QC H3T 1E2. Canada

^hUniversity of Western Ontario. Monsignor Roney Bldg/Rheum. St. Joseph Health Care London. 268 Grosvenor St. London ON N6A 4V2. Canada

ⁱUniversity of Alberta, Canada. Rheumatology 10839 124th St. NW Edmonton AB T5M 0H4. Canada

^jThe Moncton Hospital. South East Regional Health Authority. 135 Macbeath Avenue. Moncton NB E1C 6Z8. Canada

^kMcMaster University. Internal Medicine/Rheumatology. 25 Charlton Ave East, Suite 708. Hamilton ON L8N 1Y2. Canada

© 2014 Elsevier Inc. All rights reserved.

CORRESPONDING AUTHOR: Gloria Salazar, MD, 6431 Fannin St. MSB 5.270, Houston TX 77030, Telephone: 713-500-6894
Fax: 713-500-0580, gloria.salazarcintora@uth.tmc.edu.

DISCLOSURES: The authors report no conflicts of interest.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

^lUniversity of Manitoba. Arthritis Center. RR149-800 Sherbrook. Winnipeg MB R3A 1M4 Canada

^mBoston University School of Medicine. Arthritis Center/Rheumatology. 72 E Concord St, Evans 501. Boston, MA 02118. USA

ⁿMedical University of South Carolina. Division of Rheumatology. 96 Jonathan Lucas St, Suite 912 P.O. Box 250637. Charleston, SC 29425. USA

^oUniversity of Utah School of Medicine, Division of Rheumatology. 30 N 1900 E Rm 4B200. Salt Lake City, UT 84132-2101. USA

^pUniversity of Alabama Birmingham. Division of Rheumatology. FOT 844, 1720 2nd Ave S. Birmingham AL 35294-3408. USA

^qUniversity of Minnesota. Division of Rheumatic/Autoimmune Diseases. MCC 108, 420 Delaware St SE. Minneapolis, MN 55455. USA

^rThe University of Calgary, Heritage Medical Research Building, Room 414. 3330 Hospital Drive NW, Calgary, Alberta. Canada T2N 4N1

^sUniversity of Massachusetts Medical School. Department of Medicine, Division of Rheumatology. Worcester, MA. USA

Abstract

Objective—To examine the demographic and clinical characteristics of systemic sclerosis (SSc) patients without antinuclear antibodies (ANA) compared to ANA positive patients.

Methods—SSc patients enrolled in the Scleroderma Family Registry and DNA Repository were included. Relevant demographic and clinical data were entered by participating sites or obtained by chart review. ANA and SSc related antibodies were determined in all investigated patients using commercially available kits at our laboratories.

Results—This study included 3249 patients, of whom 208 (6.4%) were ANA negative. The proportion of male patients was higher in the ANA negative group (OR 1.65; $p=0.008$). ANA negative patients experienced less vasculopathic manifestations of SSc. The percent predicted diffusing capacity of carbon monoxide (DLco) was higher in ANA negative patients ($p=0.03$). Pulmonary arterial hypertension (PAH) per right heart catheterization was less common in the ANA negative group (OR= 0.28; $p=0.03$). Furthermore, patients with negative ANA had a lower prevalence of telangiectasias and digital ulcers/pits (OR= 0.59; $p=0.03$ and OR=0.38; $p=0.01$, respectively). Although diffuse cutaneous involvement was more common, the modified Rodnan Skin Score (mRSS) was lower in the ANA negative group (2.4 points lower, $p=0.05$). Furthermore, they experienced more malabsorption ($p=0.05$). There was no difference in the frequency of pulmonary fibrosis or scleroderma renal crisis. All-cause mortality was not different between the two groups ($p=0.28$).

Conclusions—In conclusion, the results of this study suggest that SSc patients who are ANA negative constitute a distinct subset of SSc with less vasculopathy (less PAH, digital ulcers and fewer telangiectasias), a greater proportion of males and possibly, more frequent lower gastrointestinal involvement.

1. Introduction

Systemic Sclerosis (SSc) is an autoimmune disease characterized by fibrosis of skin and internal organs, as well as vasculopathy and immune dysregulation. SSc is a clinically heterogeneous disease that can range from limited skin involvement and minimal internal organ disease to rapidly progressive organ involvement and skin fibrosis resulting in premature death.

Autoantibody formation is one of the hallmarks of SSc. Several studies have shown that the autoantibodies found in patients with SSc carry considerable value in diagnosis and in predicting various clinical outcomes [1–4]. Although SSc related autoantibodies are associated with specific genotypes as well as characteristic clinical manifestations, the role of ANA antibodies and its subsets in the pathogenesis of SSc is unclear. While the great majority of patients with SSc have circulating antinuclear antibodies (ANA) (90–95%), a small percentage of patients are ANA negative (5–10%) [1, 2]. Although the typical clinical presentations of the different subsets of ANA positive patients have been extensively examined, the detailed demographic and clinical characteristics of patients without detectable ANA have not been clearly explored.

The purpose of this study was exploratory and to describe the clinical manifestations of this SSc subgroup by determining their clinical and demographic differences compared to ANA positive patients. Our hypothesis was that ANA negative patients are a subgroup of SSc with a distinct clinical presentation.

2. Patients and Methods

2.1 Study population

Patient information was obtained from the Scleroderma Family Registry and DNA Repository[5] database. Patients were recruited at the University of Texas – Houston and from the following participating sites including: the participating Canadian Scleroderma Research Group (CSRG) sites, University of California Los Angeles, University of Michigan, Georgetown University, Boston University, Medical University of South Carolina, Johns Hopkins University, University of Utah, Northwestern University, University of Alabama Birmingham and University of Minnesota. All patients that agreed to be enrolled in the National Scleroderma Family Registry and DNA Repository at the participating sites were included in the current study. Of note, 390 of the Canadian patients included in our study were also investigated in a recently published study that investigated the frequency of autoantibody negative SSc patients constituting a small overlap of 12% in the study population between these two studies [6].

All study patients fulfilled the 1983 American College of Rheumatology preliminary criteria for SSc [7] or had 3 of the 5 clinical features of the CREST syndrome (Calcinosis, Raynaud’s phenomenon, Esophageal dysfunction, Sclerodactyly or Telangiectasias) with sclerodactyly being mandatory [8].

2.2 Autoantibodies

Presence of antinuclear antibodies (ANA) was investigated in all patients at the time of enrollment using indirect immunofluorescence on HEp-2 cells as the antigen substrate in the Rheumatology laboratory of the University of Texas Health Science Center at Houston. A titer of >1:80 was considered positive. All ANA titers and patterns were determined by the same investigator (FCA). Anticentromere antibodies (ACA) were determined by the pattern of immunofluorescence staining on HEp-2 cells. Antitopoisomerase antibodies (anti-topo), anti-U1-RNP (RNP), anti-SSA (anti-Ro60) and anti-SSB (anti-La) were determined by passive immunodiffusion against calf thymus extract with commercial kits (Inova Diagnostics, San Diego, CA, USA). Anti-RNA polymerase III (RNA Pol-III) antibodies were determined by enzyme-linked immunosorbent assay (MBL, Co. Ltd, Nagoya, Japan). To be considered ANA negative both the ANA and all other autoantibodies listed above had to be negative. The comparison group was defined as patients with a positive ANA.

A repeat ANA testing was performed in a subgroup of ANA negative patients from whom a follow-up serum sample was available (n=19). Only one patient became positive after repeat testing, this patient was excluded from the analysis.

2.3 Clinical manifestations

Cross-sectional demographic and clinical data were entered directly or captured from medical records utilizing a standard abstract form. Clinical manifestations were entered based on the findings of latest clinic visit.

Age, sex, disease type, disease duration (calculated from the onset of both Raynaud's and the first non-Raynaud's phenomenon symptom attributable to SSc), and modified Rodnan skin score (mRSS) at the time of study entry, were recorded. Disease type (limited cutaneous versus diffuse cutaneous SSc) [8] was defined based on the extent of skin involvement as assessed by the physician upon enrollment.

Malabsorption was defined as diarrhea associated with >10% loss of body weight. Scleroderma renal crisis was defined as accelerated hypertension, rapidly progressive renal insufficiency and/or microangiopathic hemolysis [9]. Electrocardiography and 2-dimensional echocardiography findings were recorded and were used to determine the presence of pericarditis or clinically significant pericardial effusion.

Interstitial lung disease (ILD) was defined as imaging changes consistent with scleroderma-related fibrosis including honeycombing, increased interstitial markings or ground glass opacity on chest radiograph, chest computed tomography (CT) or high resolution CT (HRCT) of the chest and restrictive lung disease indicative of ILD was determined when the forced vital capacity (FVC) predicted value was < 70%. Only 35 patients or 3% of the 1098 that were classified as having ILD by these means had an FEV1 (forced expiratory volume) to FVC ratio below 0.7 indicative of a possible obstructive process. These patients were excluded from the ILD analyses to prevent misclassification.

Pulmonary arterial hypertension (PAH) was defined by mean pulmonary artery pressure 25 mm Hg on right heart catheterization (RHC) with pulmonary wedge pressure <15.

Pulmonary Hypertension (PH) was defined as a right ventricular systolic pressure (RVSP) 50 mm Hg on 2-dimensional echocardiography when RHC was not available; this cut-off value was chosen because previous studies have shown RVSP >50 corresponds most reliably to pulmonary hypertension on RHC [10, 11].

Not all enrolled patients underwent all diagnostic studies reflecting the practice pattern of participating sites and individual physicians. Specifically, pulmonary function test results were available for analysis in 2624 cases, whereas echocardiogram was available for 1356 patients. When more than one echocardiogram was available the highest measurement was recorded. Chest imaging (chest X-ray, chest CT, or HRCT) was available in 756, 400 and 754 patients respectively. Right heart catheterization (RHC) was performed when clinically indicated in 415 patients, 381 were ANA positive and 31 were ANA negative.

2.4 Vital Status

A vital status search was performed only in patients that were enrolled at the University of Texas - Houston for whom personal identifiers were available. The vital status of these 1475 patients was determined by the National Death Index (NDI), at Centers for Disease Control and Prevention (CDCP) and the online Social Security Death Index (SSDI). The ascertainment of vital status was complete based on the available information in the above mentioned data bases. The censoring date was January 30th 2013.

All study subjects provided written informed consent and the study was approved by all participating sites.

3. Data Analysis

Analysis was carried out using STATA 12 (Statacorp LP, College Station, TX, USA) statistical package. Clinical manifestations were considered the outcome (dependent) variables and ANA status was the independent variable. Initial comparisons were conducted by chi-square or t-test depending on whether the outcome being analyzed was categorical or continuous, respectively.

Multivariable analysis adjusting for potential confounders, specifically age at enrollment, disease duration, disease type (limited or diffuse cutaneous disease) and gender was performed. The same confounding variables were used for all multivariable analysis. Linear regression was used for continuous variables while logistic regression was used for the categorical. The results were adjusted for multiple comparisons using the False Discovery Rate (FDR) or Benjamini-Hochberg <5% adjustment method. The FDR method was used because several independent variables such as markers of pulmonary hypertension were not independent from each other.

All-cause mortality analysis with disease onset as the starting point was performed by Cox proportional-hazards regression and was graphically depicted by Kaplan-Meier survival curve.

4. Results

4.1 Demographic characteristics

A total of 3249 patients were included in this study, of whom 208 (6.4%) were ANA negative. Table 1 shows the demographic and basic clinical characteristics of patients with and without ANA antibodies. There was no difference for mean age at disease onset which was 44 years for both groups. When considering the onset of Raynaud's phenomenon as the starting point for disease onset there was still no significant difference for mean age at disease onset. There were more males in the ANA negative group. No differences were observed in the ethnic composition between the two groups.

4.2 Clinical Characteristics

The mean disease duration from non-Raynaud's phenomenon onset to diagnosis was 1.9 years for ANA negative and 2.2 years for ANA positive patients. The mean disease duration from Raynaud's phenomenon onset to diagnosis and the average disease duration from SSc onset (non-Raynaud's phenomenon) to study entry were not different between the two groups (Table 1). Similar findings for this analysis were observed if the disease duration was measured from Raynaud's phenomenon onset (12 years for ANA positive and 13 for ANA negative, $p=0.20$). ANA negative patients had a higher proportion of diffuse skin involvement than their ANA positive counterparts (Table 1). The severity of skin involvement as evaluated by the modified Rodnan skin score (mRSS) was not different in the initial comparison (Table 2). Although ANA negatives had a higher percentage of patients with diffuse skin involvement, the multivariable analysis adjusting for potential confounders including disease type (plus age at enrollment, disease duration, and gender) showed that the fibrosis severity as measured by mRSS was lower in the ANA negative group (Table 3).

The FVC % predicted was lower in ANA negative patients; however, the multivariable analysis adjusting for the above mentioned confounders failed to show a significant difference (Table 3). The DLco % predicted was higher in the ANA negative patients (Table 2). Pulmonary arterial hypertension (PAH) verified by right heart catheterization (RHC) was significantly less common in the ANA negative group (Table 2). Only 7 ANA negative patients had PAH compared to 213 ANA positive patients.

When patients who had PAH diagnosed per RHC and patients that did not have a RHC but had an RVSP>50 mmHg per echocardiogram were included in the analysis, a total of 2040 patients were analyzed (1907 were ANA positive and 133 were ANA negative). In this analysis, we again observed a statistically significant difference between the two groups for pulmonary hypertension (PH) (Table 2).

After adjusting for potential confounders (disease type, age at enrollment, disease duration, and gender), the difference in prevalence of PAH by RHC between the two groups remained statistically significant (Table 3). When centromere antibody status was added to the model, the relationship was still significant ($p=0.009$). PH (per RHC or RVSP>50) was also statistically significant in the multivariable model but the level of significance was not sustained after adjusting for multiple comparisons (Table 3). ANA negative patients were

found to have a lower prevalence of telangiectasias and digital ulcers/pits (Table 2). As shown in Table 3, these differences remained significant in the multivariable model.

ANA negative patients experienced more malabsorption (Table 2). This finding remained significant in the multivariable analysis (Table 3).

There was no difference in creatine kinase (CK) levels, fibrotic changes per chest imaging or ejection fraction by echocardiogram (LVEF<50%) between the two groups. Also no difference was observed for the presence of Raynaud's phenomenon, GERD or calcinosis. Of note, there were no cases of primary biliary cirrhosis (PBC) in the ANA negative group and 12 in the ANA positive group but this finding did not reach statistical significance.

Ten patients had scleroderma renal crisis (SRC) in the ANA negative group. While the ANA status (negative versus positive) was not associated with higher occurrence of SRC in the overall cohort, ANA negative patients were more likely to develop SRC than patients with anti-topo or ACA (OR= 3.83, CI 1.49–9.84, $p=0.005$ and OR= 11.86, CI 3.68–38.22, $p<0.001$, respectively). Only RNA-polymerase III, which is a well-known risk for SRC, was associated with a higher risk compared to the ANA negative group (OR 4.87; CI=3.19–7.44; $p<0.001$) after adjusting for confounders).

4.3 Survival Analysis

Among 1475 patients in whom a death search could be performed, 300 had died. The ANA status was not associated with overall mortality either in the univariable or the multivariable analysis. The univariable analysis showed a hazard ratio (HR) 0.96 (CI 0.62–1.49, $p=0.87$). Multivariable analysis correcting for potential confounders resulted in a HR=0.83 (CI 0.53–1.30, $p=0.42$). Kaplan-Meier survivor function is shown in figure 1.

4.4 ANA follow up

Nineteen ANA negative patients in our database had a follow-up serum sample. The repeat ANA was performed at a mean of 4.5 years after the initial determination (SD 4.3 years). Out of these, only one patient who was initially ANA negative seroconverted to ANA positive upon follow-up testing. The pattern was atypical (few bright speckles).

5. Discussion

This study is the first to specifically focus on the demographic and clinical features of ANA negative SSc in a large and diverse sample of patients. Our study was of sufficient size for statistical analysis and our survival cohort was also larger than published cohorts to date. We observed that ANA negative patients represent a distinct subset of patients who have the fibrotic features of SSc but are less likely to have the vasculopathic features of the disease.

We observed that PAH and clinical markers of vasculopathy such as digital ulcerations and telangiectasias were less common in ANA negative patients.

Our findings show that, although SSc is relatively uncommon in males, ANA negative patients in our study were more commonly male. The ANA negative patients might also have a higher frequency of malabsorption.

Only a few studies have previously described some of the clinical characteristics of ANA negative patients. Hamaguchi et al delineated the clinical manifestations of SSc patients based on antibody subsets and included 203 Japanese SSc patients, 10 of whom were ANA negative (5% of their sample) [12]. They reported lower severity of skin involvement, rare incidence of PAH, ILD and pitting scars (one patient each) in the ANA negative group.

The German network for SSc published a study focusing on the correlation of clinical features and the different antibody subsets [13]. Within their sample of 863 patients, 50 (5.8%) were ANA negative. In their study, ANA negative cases had significantly less frequent PAH, digital ulcers and Raynaud's phenomenon.

More recently Hudson et al published their observations on autoantibody negative SSc in the CSRG cohort [6]. Their sample included 874 patients. They divided their patients into groups based on their ANA and/or ENA (Extractable Nuclear Antigens) status as positive (+) versus negative (-). ENA antibodies also included cytoplasmic antibodies. This resulted in 3 different study groups (group 1: ANA+/ENA+, group 2: ANA+/ENA- and group 3: ANA-/ENA-). Their sample included 15 patients (1.7%) who were autoantibody negative (group 3). Three hundred ninety of the patients examined in our study were also studied by Hudson et al, indicating a small overlap of 12% in the study population between our studies. Only univariable comparisons without correction for potential confounders were performed due to sample size limitations and were done between the ANA+/ENA- group and the remaining two groups separately. They reported lower prevalence of Raynaud's phenomenon and telangiectasias in the autoantibody negative group. They did not observe any other clinical differences, possibly due to sample size limitations or differences in the study design.

Despite the small overlap with this last study (390 patients), thanks to our multi-institutional international effort our cohort was able to support a more in depth analysis and represents the largest sample of ANA negative SSc patients described in the literature to date. The data collected from this endeavor enabled us to conduct for the first time, robust clinical correlations given the significantly larger sample size, wider geographic catchment area and multi-ethnic population. Encouragingly the results of the above mentioned studies support several of our findings. Similar to our results, these studies indicate that ANA negative patients have less often vasculopathic features of disease.

The association of SSc vasculopathic features with ANA positivity might have pathophysiological implications. This relationship was still present after correcting for ACA which has been associated with susceptibility to PAH. Therefore any circulating ANA is potentially a risk for PAH compared to the absence of ANA. Our analysis showed a significant difference for overall markers of vasculopathy in ANA positive compared to ANA negative patients including PAH, telangiectasias and digital ulcers/pits. A correlation between PAH and telangiectasias was previously described by Shah et al which supports the notion that vasculopathic features tend to occur together and that there is a subgroup of patients who have more vasculopathy [14].

Several studies have described the relationship of specific antinuclear antibodies including ACA, RNP and Th/To to the presence of PAH. More recently, Nihtyanova et al also reported an association between PAH and RNA-polymerase III antibody positivity.[15] The relationship of ACA with digital ulcers, telangiectasias and microvascular damage per capillaroscopy has also been observed [1–3, 16–18]. However, the pathophysiologic role of these antibodies has not been determined (direct cause of vascular damage, or an epiphenomenon). It is also possible that ANA negative patients have other antibodies that are not currently detected by our traditional assays causing the observed clinical differences or that the ANAs are associated with other antibodies that are causing the vascular damage. Studies designed to understand the pathophysiology of this important manifestation of SSc are needed.

Several studies have linked autoantibody production with genetics in SSc. Multiple polymorphisms in the human leukocyte (HLA) regions have been linked to SSc and the association of some of the class II haplotypes with the different ANA subsets has been described [19–21]. Also familial SSc studies have shown a concordance of autoantibodies between family members [22, 23]. Although the relationship between ANA and the genetic background has been studied, the genetics of the ANA negative subgroup are not known. The discovery that these patients are more commonly male underlines the importance of exploring the genetics of this subset. If genetics do not play a role than other causes for the observed clinical differences such as hormonal changes or differentially expressed epigenetic mechanisms by gender should be considered.

ANA negative patients have a lower frequency of PAH which is a common cause of mortality for SSc patients; however their all-cause mortality was not different. This could be due to severe GI involvement or ILD. Dedicated longitudinal prospective studies are needed to analyze this finding in more detail. We considered the possibility of detection bias in our study given that ANA negative SSc is rare and can be easily missed unless clinically evident, but this would have biased toward female sex, longer disease duration before diagnosis and diffuse disease. We observed more diffuse disease in the ANA negative group. However, the severity of skin fibrosis as measured by mRSS was lower in the multivariable analysis that included adjustment for disease type. Interestingly, the ANA negative group did not have longer time to diagnosis as we expected, and had a higher proportion of males. It is also important to mention that although RHC was obtained for only a subgroup of patients which could arguably bias the results of our analysis, we observed that the proportion of ANA negative and positive patients was not statistically different between the patients that had a RHC versus the patients that did not have one (7% vs 6% in ANA negatives and 92% vs 93% in ANA positives, $p=0.4$).

Taken together, these data support the concept that ANA negative SSc patients are a subgroup of patients with distinct demographic and clinical manifestations. Further prospective clinical studies as well as genetic studies examining this subpopulation of patients are warranted. Also studies examining the pathophysiologic link between ANAs and vasculopathy in SSc should be considered.

The ANA negativity appears stable over time, as almost all patients in a follow-up subset of the original cohort that had an initial negative determination remained negative for up to 14 years when retested. Only one patient that was initially negative became positive upon follow up. The repeated ANA had an IIF pattern of few bright speckles also known as NSp-I or “multiple nuclear dots” (MND).[24, 25] Among a variety of potential MND targets, this pattern was seen in patients with sp100 antibodies, which are very specific for primary biliary cirrhosis (PBC) [26, 27]. However this particular patient has not developed PBC up until now. We acknowledge that longitudinal data were only available in a small subgroup of ANA negative patients. Nevertheless, we believe this limited longitudinal study provides important pilot data on stability of ANA negativity over time, which needs to be verified in larger follow-up studies.

The present study has some limitations. Reflecting the practice patterns of treating physicians, the data on clinical manifestations and diagnostic studies were not available in all patients. Furthermore, the diagnostic tests and basic laboratory studies were done at different centers. Although the data collection and patient inclusion for this study was performed before the 2013 ACR/EULAR classification criteria were published, we set out to determine how many of our patients fulfilled these criteria. Based on the available data, we were able to classify 97.5 percent of our sample. However our data does not include capillaroscopy information for all the participating sites therefore we cannot make assumptions on the 80 remaining patients which were not classified.

It is possible that some of the observed associations are false-positive findings; however, almost all the statistically significant differences observed remained significant after adjusting our analysis for multiple comparisons. Furthermore, we observed negative associations with several independent vasculopathic features supporting the notion that these associations are true biological findings rather than random spurious results.

As this is a cross-sectional study, the information on highest ever mRSS was not available. We could not investigate cause-specific mortality and prevalence of malignancy as this information was not captured in our data base. The paucity of data on chemical studies to clearly define malabsorption is also a limitation. Although we realize that our definition of malabsorption was relatively non-specific and did not include measures of bacterial overgrowth, stool fat measurements or laboratory tests for malabsorption such as carotene, we feel that clinically significant weight loss along with diarrhea is a reasonable surrogate. Moreover although the survival analysis was performed only on patients who were enrolled at UT Houston (due to lack of personal identifiers in our data base for those enrolled at other centers), we do not think this introduces a systemic bias as the percentage of ANA negative patients was not different between the patients who had mortality data available and the patients who did not (6.8% vs 5.9%, $p=0.3$). Finally, we cannot report on percentage of patients who were ANA positive at a lower titer (i.e. 1:40) because this was not examined in our laboratory.

The strengths of the current study are that the ANA testing was centrally performed and read by the same investigator and our large sample which is multi-ethnic and representative of

the SSc population across a wide geographic area in the United States and Canada. Finally, this sample includes the largest ANA negative SSc sample reported to date.

6. Conclusion

In conclusion, the results of this study suggest that SSc patients who are ANA negative constitute a distinct subset of SSc with less vasculopathy (less PAH, digital ulcers and fewer telangiectasias), a greater proportion of males and possibly, more frequent lower gastrointestinal involvement.

It is important to understand the clinical characteristics and genetics of ANA negative patients with SSc because this will allow further understanding the role of ANA in the pathophysiology of SSc.

Acknowledgments

We would like to thank Tony Mattar for his assistance in the database management, Julio Charles for performing the IIF laboratory studies and all the participating sites including: the Canadian Scleroderma Research Group, University of California Los Angeles, University of Michigan, Georgetown University, Boston University, Medical University of South Carolina, Johns Hopkins University, University of Utah, Northwestern University, University of Alabama Birmingham and University of Minnesota. We would also like to thank and acknowledge the Scleroderma Research Foundation, the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health and the Department of Defense for their support of our centers to accomplish this study. NIH/NIAMS K23 AR061436 (Assassi); NIH/NIAMS N01-AR-0-2251 (Mayes); NIH/NIAMS K24 AR063120-02 (Khanna); NIH/NIAMS-RO1-AR055258; NIH/NIAMS P50-AR05414; U01AI09090-01 (Mayes) and the Department of Defense Congressionally Directed Medical Research Programs (W81XWH-07-01-0111-Mayes). Finally, we would like to extend a special acknowledgement to the late Dr. Janet Markland who contributed to this project. We would like to recognize her not only for her participation in this study but also for being a part of the effort to advance scleroderma research.

References

1. Ho KT, Reveille JD. The clinical relevance of autoantibodies in scleroderma. *Arthritis Res Ther.* 2003; 5:80–93. [PubMed: 12718748]
2. Hamaguchi Y. Autoantibody profiles in systemic sclerosis: predictive value for clinical evaluation and prognosis. *J Dermatol.* 2010; 37:42–53. [PubMed: 20175839]
3. Steen VD. Autoantibodies in systemic sclerosis. *Semin Arthritis Rheum.* 2005; 35:35–42. [PubMed: 16084222]
4. Arnett FC. Is scleroderma an autoantibody mediated disease? *Curr Opin Rheumatol.* 2006; 18:579–81. [PubMed: 17053501]
5. Mayes MD. The establishment and utility of a population-based registry to understand the epidemiology of systemic sclerosis. *Curr Rheumatol Rep.* 2000; 2:512–6. [PubMed: 11123106]
6. Hudson M, Satoh M, Chan JY, Tatibouet S, Mehra S, Baron M, et al. Prevalence and clinical profiles of ‘autoantibody-negative’ systemic sclerosis subjects. *Clin Exp Rheumatol.* 2013
7. Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Preliminary criteria for the classification of systemic sclerosis (scleroderma). *Arthritis Rheum.* 1980; 23:581–90. [PubMed: 7378088]
8. 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. [PubMed: 3361530]
9. Steen VD, Mayes MD, Merkel PA. Assessment of kidney involvement. *Clin Exp Rheumatol.* 2003; 21:S29–S31. [PubMed: 12889219]
10. Galie N, Hoeper MM, Humbert M, Torbicki A, Vachiery JL, Barbera JA, et al. Guidelines for the diagnosis and treatment of pulmonary hypertension: the Task Force for the Diagnosis and

Treatment of Pulmonary Hypertension of the European Society of Cardiology (ESC) and the European Respiratory Society (ERS), endorsed by the International Society of Heart and Lung Transplantation (ISHLT). *Eur Heart J*. 2009; 30:2493–537. [PubMed: 19713419]

11. Hsu VM, Moreyra AE, Wilson AC, Shinnar M, Shindler DM, Wilson JE, et al. Assessment of pulmonary arterial hypertension in patients with systemic sclerosis: comparison of noninvasive tests with results of right-heart catheterization. *J Rheumatol*. 2008; 35:458–65. [PubMed: 18203320]
12. Hamaguchi Y, Hasegawa M, Fujimoto M, Matsushita T, Komura K, Kaji K, et al. The clinical relevance of serum antinuclear antibodies in Japanese patients with systemic sclerosis. *Br J Dermatol*. 2008; 158:487–95. [PubMed: 18205876]
13. Mierau R, Moinzadeh P, Riemekasten G, Melchers I, Meurer M, Reichenberger F, et al. Frequency of disease-associated and other nuclear autoantibodies in patients of the German Network for Systemic Scleroderma: correlation with characteristic clinical features. *Arthritis Res Ther*. 2011; 13:R172. [PubMed: 22018289]
14. Shah AA, Wigley FM, Hummers LK. Telangiectases in scleroderma: a potential clinical marker of pulmonary arterial hypertension. *J Rheumatol*. 2010; 37:98–104. [PubMed: 19955048]
15. Nihtyanova SI, Schreiber BE, Ong VH, Rosenberg D, Moinzadeh P, Coghlan JG, et al. Prediction of pulmonary complications and long term survival in systemic sclerosis. *Arthritis Rheumatol*. 2014
16. Steen VD, Lucas M, Fertig N, Medsger TA Jr. Pulmonary arterial hypertension and severe pulmonary fibrosis in systemic sclerosis patients with a nucleolar antibody. *J Rheumatol*. 2007; 34:2230–5. [PubMed: 17937469]
17. Kampolis C, Plastiras S, Vlachoyiannopoulos P, Moysakakis I, Tzelepis G. The presence of anti-centromere antibodies may predict progression of estimated pulmonary arterial systolic pressure in systemic sclerosis. *Scand J Rheumatol*. 2008; 37:278–83. [PubMed: 18612928]
18. Herrick AL, Moore TL, Murray AK, Whidby N, Manning JB, Bhushan M, et al. Nail-fold capillary abnormalities are associated with anti-centromere antibody and severity of digital ischaemia. *Rheumatology (Oxford)*. 2010; 49:1776–82. [PubMed: 20538680]
19. Arnett FC, Gourh P, Shete S, Ahn CW, Honey RE, Agarwal SK, et al. Major histocompatibility complex (MHC) class II alleles, haplotypes and epitopes which confer susceptibility or protection in systemic sclerosis: analyses in 1300 Caucasian, African-American and Hispanic cases and 1000 controls. *Ann Rheum Dis*. 2010; 69:822–7. [PubMed: 19596691]
20. Mayes MD. The genetics of scleroderma: looking into the postgenomic era. *Curr Opin Rheumatol*. 2012; 24:677–84. [PubMed: 23026857]
21. Agarwal SK, Reveille JD. The genetics of scleroderma (systemic sclerosis). *Curr Opin Rheumatol*. 2010; 22:133–8. [PubMed: 20090527]
22. Feghali-Bostwick C, Medsger TA Jr, Wright TM. Analysis of systemic sclerosis in twins reveals low concordance for disease and high concordance for the presence of antinuclear antibodies. *Arthritis Rheum*. 2003; 48:1956–63. [PubMed: 12847690]
23. Assassi S, Arnett FC, Reveille JD, Gourh P, Mayes MD. Clinical, immunologic, and genetic features of familial systemic sclerosis. *Arthritis Rheum*. 2007; 56:2031–7. [PubMed: 17530643]
24. Fritzler MJ, Valencia DW, McCarty GA. Speckled pattern antinuclear antibodies resembling anticentromere antibodies. *Arthritis Rheum*. 1984; 27:92–6. [PubMed: 6197978]
25. Cozzani E, Drosera M, Riva S, Parodi A. Analysis of a multiple nuclear dots pattern in a large cohort of dermatological patients. *Clin Lab*. 2012; 58:329–32. [PubMed: 22582508]
26. Assassi S, Fritzler MJ, Arnett FC, Norman GL, Shah KR, Gourh P, et al. Primary biliary cirrhosis (PBC), PBC autoantibodies, and hepatic parameter abnormalities in a large population of systemic sclerosis patients. *J Rheumatol*. 2009; 36:2250–6. [PubMed: 19723904]
27. Zuchner D, Sternsdorf T, Szostecki C, Heathcote EJ, Cauch-Dudek K, Will H. Prevalence, kinetics, and therapeutic modulation of autoantibodies against Sp100 and promyelocytic leukemia protein in a large cohort of patients with primary biliary cirrhosis. *Hepatology*. 1997; 26:1123–30. [PubMed: 9362351]

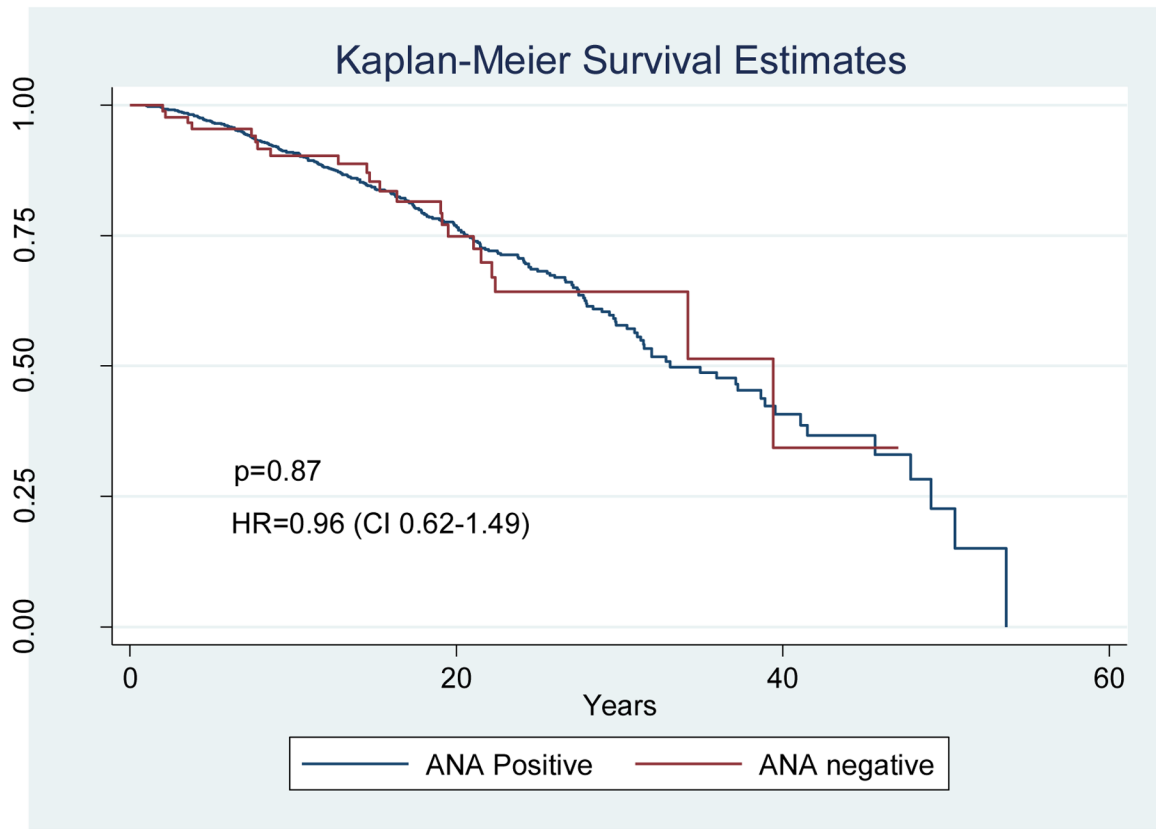


Figure 1. Kaplan-Meier survivor function of ANA negative SSc patients compared to ANA positive patients

Table 1

Analysis of demographic and basic clinical parameters in systemic sclerosis (SSc) patients who are ANA negative compared with ANA positive patients.

	ANA negative (n=208)	ANA positive (n=3041)	<i>p</i>
Age at disease onset, mean +SD years	44.7+13.7	44.6+13.5	0.97
Disease duration to diagnosis, mean +SD	1.9+3.6	2.2+4.0	0.43
Disease duration to enrollment, median (IQR)	7.8(11.64)	6.9(10.99)	0.10
Gender (male),%	41(19.7)	402(13.2)	0.008
<i>Ethnicity</i>			
Caucasian, %	182(88.3)	2516(83.7)	0.07
African American, %	12(5.8)	221(7.3)	0.41
Hispanic, %	10(4.8)	204(6.7)	0.28
Asian, %	1(0.4)	42(1.4)	0.27
American Indian, %	1(0.4)	23(0.7)	0.65
Unknown or mixed, %	2(0.09)	35 (1.2)	0.78
Disease type (diffuse), %	105(49.8)	1240(40.1)	0.007

Table 2

Comparisons of clinical parameters in systemic sclerosis (SSc) patients who are ANA negative versus ANA positive.

	ANA negative (n=208)	ANA positive (n=3041)	OR ^a	b coef ^a	95% CI	p	Adjusted P [*]
mRSS, mean +SD	12 (10)	11 (10)	N/A	0.25	-2.25, 2.77	0.84	0.85
Telangectasias,%	71/114(62)	1505/2061(73)	0.60	N/A	0.41, 0.91	0.01	0.03
Digital ulcers and pits,%	41/105(39)	1088/1925(56)	0.49	N/A	0.32, 0.73	0.001	0.008
Calcinosis,%	26/95(27)	518/1791(28)	0.92	N/A	0.58, 1.47	0.74	0.85
FVC%, mean +SD	80.6+21.7	84.1+19.4	N/A	-3.5	-6.69, -0.44	0.02	0.04
DLCO%, mean +SD	72+25	68+22	N/A	4.02	0.25, 7.8	0.03	0.05
Fibrosis per imaging,%	50/67(74)	716/1102(64)	1.66	N/A	0.93, 2.96	0.08	0.13
MPAP per RHC, mean +SD	23.4+18	29.4+13	N/A	-6.20	-11.30, -1.10	0.01	0.03
PAH per RHC,%	7/31(22.5)	221/405(55.4)	0.23	N/A	0.09, 0.55	<0.001	0.008
PH (per RHC and/or RVSP>50),%	13/133(9.7)	332/1907(17.4)	0.51	N/A	0.28, 0.92	0.02	0.04
GERD,%	116/130(89)	1834/2144(85)	1.40	N/A	0.79, 2.47	0.24	0.36
Malabsorption,%	48/94(51)	576/1597(36)	1.84	N/A	1.21, 2.80	0.004	0.02
CK, mean +SD	249 (309)	209 (704)	N/A	39.34	-380, 429	0.85	0.85
Renal crisis,%	10/208(4.8)	103/3041(3.3)	1.44	N/A	0.74, 2.8	0.28	0.38
Low LVEF(<50), %	4/106(3.7)	39/1614(2.4)	1.58	N/A	0.55, 4.51	0.39	0.48

Abbreviations: ANA= antinuclear antibodies, mRSS= modified Rodnan skin score, FVC= Forced vital capacity, DLCO= diffusing capacity of carbon monoxide, PAH= Pulmonary arterial hypertension, PH= Pulmonary Hypertension, MPAP= Mean pulmonary artery pressure, RHC= Right heart catheterization, RVSP= Right ventricular systolic pressure, GERD= Gastroesophageal reflux disease, CK= creatine kinase, LVEF= Left ventricular ejection fraction. OR =odds ratio; b coef= b coefficient; 95% CI=95% confidence interval; SD= standard deviation.

^aLinear regression was used for continuous data; logistic regression was used for categorical data

^{*} After False Discovery Rate (FDR) or Benjamini-Hochberg adjustment method for multiple comparisons.

Table 3

Multivariable analysis of clinical parameters in systemic sclerosis (SSc) patients who are ANA negative compared with ANA positive patients.

	OR ^a	b	coef ^a	95% CI	P	Adjusted P
mRSS	N/A	-2.48		-4.54, -0.42	0.018	0.05
Telangectasias	0.59	N/A		0.38, 0.91	0.01	0.03
Digital ulcers and pits	0.38	N/A		0.24, 0.59	<0.001	0.01
Calcinosis	0.63	N/A		1.08, 0.37	0.09	0.15
FVC%	N/A	-2.18		-5.33, 0.95	0.17	0.25
DLCO%	N/A	5.16		1.33, 8.99	0.008	0.03
Fibrosis per imaging	1.46	N/A		0.81, 2.64	0.20	0.27
MPAP per RHC	N/A	-5.58		-10.85, -0.32	0.04	0.07
PAH per RHC	0.28	N/A		0.11, 0.70	0.006	0.03
PH (per RHC and/or RVSP>50)	0.52	N/A		0.28, 0.94	0.03	0.06
GERD	1.32	N/A		0.75, 2.35	0.33	0.41
Malabsorption	1.66	N/A		1.08, 2.56	0.02	0.05
CK	N/A	26.64		-394, 447	0.90	0.90
Renal crisis	1.23	N/A		0.62, 2.43	0.53	0.61
Low LVEF (<50)	1.31	N/A		0.45, 3.83	0.61	0.65

Abbreviations: ANA= antinuclear antibodies, mRSS= modified Rodnan skin score, FVC= Forced vital capacity, DLCO= diffusing capacity of carbon monoxide, PAH= Pulmonary arterial hypertension, PH= Pulmonary Hypertension, MPAP= Mean pulmonary artery pressure, RHC= Right heart catheterization, RVSP= Right ventricular systolic pressure, GERD= Gastroesophageal reflux disease, CK= creatine kinase, LVEF= Left ventricular ejection fraction, OR=odds ratio; b coef= b coefficient; 95% CI=95% confidence interval; SD= standard deviation.

^aLinear regression was used for continuous data; logistic regression was used for categorical data. The multivariable analysis was adjusted for age, disease duration, gender and disease type.

* After False Discovery Rate (FDR) or Benjamini-Hochberg adjustment method for multiple comparisons.

Detection of anti-U3-RNP/fibrillarin IgG antibodies by line immunoblot assay has comparable clinical significance to immunoprecipitation testing in systemic sclerosis

Lisa K. Peterson¹ · Troy D. Jaskowski² · Maureen D. Mayes³ · Anne E. Tebo^{1,2}

Published online: 14 October 2015
© Springer Science+Business Media New York 2015

Abstract The aim of this study was to evaluate the performance and clinical relevance of a commercially available line immunoblot assay (LIA) for detecting anti-U3-RNP/fibrillarin (anti-U3-RNP), against immunoprecipitation (gold standard). This study involved a multi-ethnic cohort of 1000 American systemic sclerosis (SSc) patients and 50 healthy controls. Antinuclear antibodies and centromere antibodies were detected by indirect immunofluorescent antibody test, anti-topo I by immunodiffusion and anti-RNAP III by ELISA. The presence of anti-U3-RNP in select serum samples was detected by immunoprecipitation (IP) and LIA. By IP, U3-RNP antibody was detected in 75 (7.5 %) patients with SSc. Overall agreement between LIA and IP was very good ($\kappa = 0.966$). Analytic sensitivity and specificity of the U3-RNP LIA was 100 and 94.7 %, respectively. Clinical features associated with positivity for the anti-U3-RNP antibody include diffuse cutaneous SSc and increased prevalence of renal crisis, consistent with previous studies that used IP. Testing for U3-RNP antibodies is only performed by a small number of laboratories due to the complexity of both performance and interpretation of the IP. LIA is faster and less complex than IP.

Excellent agreement between IP and LIA demonstrates that LIA is an acceptable and attractive alternative to IP for anti-U3-RNP detection.

Keywords Systemic sclerosis · Autoantibodies · Anti-fibrillarin · Anti-U3 small nucleolar ribonucleoprotein · Antinuclear antibodies

Introduction

Systemic sclerosis (SSc) is an autoimmune connective tissue disease characterized by fibrosis of the skin and internal organs, Raynaud's phenomenon and the presence of antinuclear antibodies (ANA). Anti-centromere, anti-topoisomerase I (anti-topo-1 or anti-Scl-70) and anti-RNA polymerase III (anti-RNAP III) are the most common ANA-positive patients with SSc, but only account for the ANA in 45–90 % of patients [1–5]. An increasing number of SSc-specific autoantibodies and their corresponding autoantigens have been identified. Detection of these autoantibodies aids in diagnosis, prognosis and therapeutic decisions [1–3, 6–9].

Anti-fibrillarin was first described as a clumpy nucleolar pattern on ANA IFA using sera from patients with progressive SSc [10]. It was subsequently named anti-fibrillarin based on reactivity of SSc patient sera with fibrillar regions of nucleoli [11]. It recognizes a highly conserved 34 kDa protein that is a component of U3 ribonucleoprotein (U3-RNP) [12–16]. Antibodies against U3-RNP/fibrillarin (anti-U3-RNP) are considered to be a very specific marker for SSc and are associated with shorter mean disease duration at time of diagnosis, younger age of onset, severe diffuse cutaneous SSc (dcSSc) and internal organ involvement [9, 15, 17, 18]. Antibodies reactive with

✉ Anne E. Tebo
anne.tebo@hsc.utah.edu

¹ Department of Pathology, University of Utah, Salt Lake City, UT, USA

² ARUP Institute for Clinical and Experimental Pathology, Salt Lake City, UT, USA

³ Department of Internal Medicine, Division of Rheumatology and Clinical Immunogenetics, University of Texas Health Science Center, Houston, TX, USA

U3-RNP are found in 5–8 % of SSc patients [4, 12, 19]. U3-RNP antibodies are more prevalent in male SSc patients and among African American patients with SSc compared to patients from other ethnic groups [20–25]. Moreover, anti-U3-RNP is the second most common ANA in African American patients with SSc [21], and is found in 12–48 % of ANoA-positive SSc patients [6, 10, 11, 26, 27]. Thus, although infrequent in the general population, U3-RNP antibodies are a marker for severe SSc, especially in male and African American patients.

Testing for U3-RNP antibodies is only performed by a small number of laboratories due to the complexity of both performance and interpretation of the immunoprecipitation (IP) [4, 19]. Autoantibodies to U3-RNP are not currently included in the classification criteria for SSc, but may be incorporated if testing became more widely available [28]. Recently, an LIA was developed for the detection of 12 SSc-associated antibodies including U3-RNP. Several studies have evaluated the performance of the LIA as a whole on unselected SSc patient populations, but this approach yielded minimal information about the performance for markers with low prevalence, such as U3-RNP [4, 29–32]. Thus, the concordance between LIA and IP for the detection of anti-U3-RNP is not known. In this study, we evaluated the performance of the more rapid and less complex LIA assay as an alternative to the gold standard IP method for the detection of anti-U3-RNP in a multi-ethnic cohort of American patients with the two main forms of SSc, dcSSc and limited cutaneous (lcSSc).

Materials and methods

Serum samples

From a multi-racial cohort of 1000 SSc patients (Scleroderma Family Registry and DNA Repository, SFRDR), 98 sera [ANoA alone ($n = 94$) or ANoA and nuclear speckled patterns ($n = 4$)] negative for antibodies to Scl-70 by immunodiffusion, to RNA polymerase III (RNAP III) by ELISA and to centromere (ACA) by indirect fluorescent antibody were identified for autoantibody characterization by IP and U3-RNP LIA. Sera were tested by IP by Dr. Minoru Satoh at the University of Florida. The patients were characterized clinically based on extent of skin involvement, age at disease onset, disease duration at blood draw, gender and ethnicity. The study was approved by the Institutional Review Boards at the University of Utah School of Medicine, Salt Lake City, and the University of Texas Health Science Center at Houston. Written informed consents from all subjects according to the declaration of Helsinki were obtained at the University of Texas Health Science Center at Houston.

Detection of SSc-associated antibodies by immunoblot

Sera from 98 samples previously tested for anti-U3-RNP by IP and 50 “self-proclaimed” healthy adults were used in this study. Sera were tested for autoantibodies against U3-RNP using the Systemic Sclerosis (Nucleoli) Profile (IgG) Immunoblot (Euroimmun AG, Lübeck, Germany) according to the manufacturer’s instructions. The blot strips were digitized using a camera and band intensities determined using the EUROLIne-Scan software (Euroimmun AG). Signal strengths of <6 Units (U) were considered negative, 6–10 Units (U) were considered borderline, and results >10 U were considered positive according to the manufacturer’s recommendations.

Statistical analysis

Agreement between the assays was quantified using Cohen’s κ statistics for pair-wise comparisons. Kappa coefficients >0.75 signify substantial agreement. Significance of associations between sets of categorical data was determined using the Chi-square test. Statistical analysis was performed using Prism 5.0 (GraphPad Software). Differences were considered statistically significant for p values <0.05.

Results

Anti-U3-RNP IP and LIA results concurred in 98.9 % of patients, yielding a κ coefficient 0.966 (Fig. 1a). Two U3-RNP IP positive sera were borderline by LIA, one of which reacted with Ro-52 by LIA and the other did not react with any of the antigens on the LIA (data not shown). Both of these samples were from patients with dcSSc. One LIA U3-RNP positive sera was not detected by IP and another was borderline by IP. Interestingly, both of these samples reacted with Th/To by IP. The sample that was positive by LIA, but negative by IP was from a patient with lcSSc. The sample that was positive for U3-RNP by LIA but borderline by IP was from a patient with dcSSc. Four samples that were anti-U3-RNP negative by IP had borderline results by LIA, all from patients with dcSSc. Two additional samples had borderline results for U3-RNP by IP, one of which was negative by LIA and the other was also borderline by LIA. Both of these samples were from patients with dcSSc. Comparison of U3-RNP LIA results in patients with the two main forms of SSc, dcSSc and lcSSc, and healthy controls confirmed the diagnostic performance of the LIA assay (Fig. 1b). In addition, LIA demonstrated that the prevalence of anti-U3-RNP was significantly higher in the dcSSc group (80.6 %) compared to the lcSSc group (55.2 %;

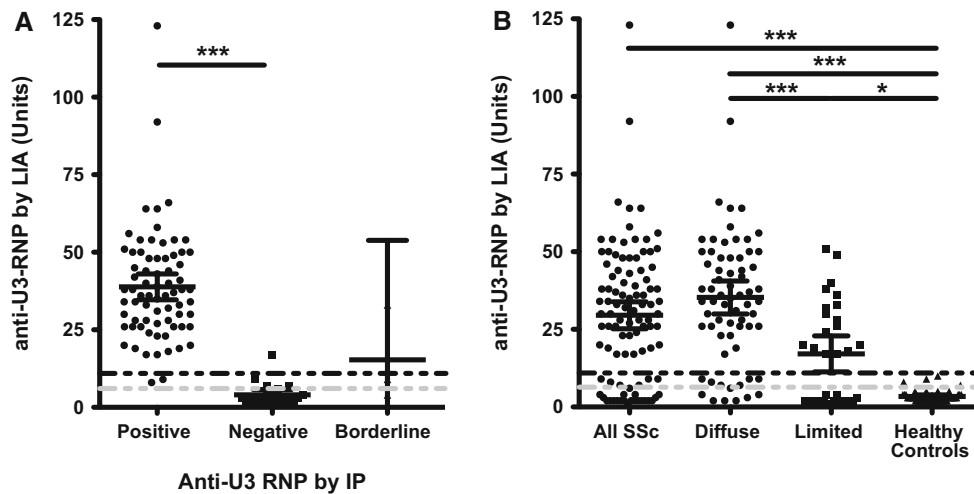


Fig. 1 Performance characteristics of the U3-RNP line immunoassay. **a** Correlation between line immunoblot assay (LIA) and immunoprecipitation (IP) for the detection of anti-U3-RNP in sera from SSc patients divided into 3 groups based on the detection of U3-RNP antibodies by IP. **b** Anti-U3-RNP expression in the sera of 98 patients with SSc and 50 healthy controls (Controls). **a, b** Error bars

represent the mean and 95 % CI. The *black dotted lines* indicate the cutoff level for positivity according to the manufacturer’s recommendation. The *gray dotted lines* indicate the cutoff level for samples considered borderline according to the manufacturer’s recommendation. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$

$p = 0.0005$). Based on the semi-quantitative values generated by the LIA, patients with dcSSc also had significantly higher levels of antibodies to U3-RNP [mean of 35.3 U (95 % CI 30–40.6 U)] compared to those with lcSSc [mean of 17.1 U (95 % CI 11.3–22.9 U)] ($p < 0.0001$).

Table 1 summarizes the demographic and clinical information for SSc patients with concordant results by LIA and IP. There was no difference between these

methods for associations between anti-U3-RNP and demographic or clinical features. There was a trend toward a greater proportion of men (1:4.83 vs 1:9; $p = 0.7255$), a younger age at the time of diagnosis (mean age 35.9 vs 43.0 years; $p = 0.651$) and a higher prevalence of renal crisis (22.9 vs 12.5 %; $p = 0.6130$) in anti-U3-RNP positive patients, but these associations did not reach statistical significance. There was no difference in disease duration at

Table 1 Demographics and clinical features of ANoA-positive SSc patients positive or negative for anti-U3-RNP by both LIA and IP

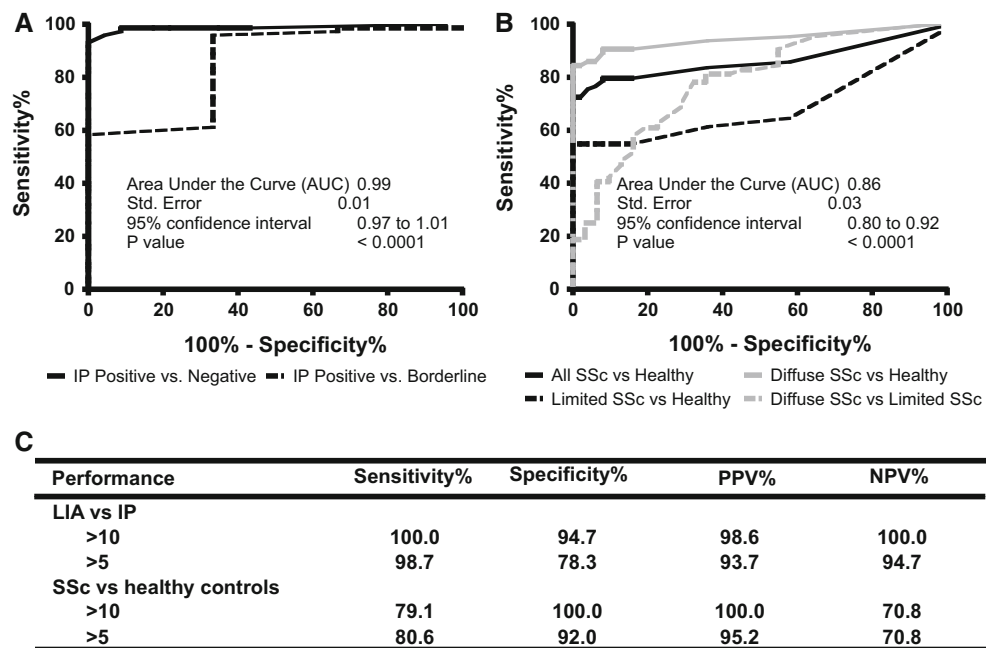
Demographic features and disease classification	Anti-U3-RNP positive ($n = 70$)	Anti-U3-RNP negative ($n = 18$)	p value
Age at onset (years)	35.9 ± 15.0	43.0 ± 11.2	ns
Age at onset (years, median)	35.7	43.0	ns
Disease duration (years, mean ± SD)	5.7 ± 6.0	5.9 ± 4.0	ns
Disease duration (years, median)	3.4	5.7	ns
Female ($n = 74$)	58 (82.9)	16 (88.9)	ns
Male ($n = 14$)	12 (17.1)	2 (11.1)	ns
Hispanic ($n = 27$)	16 (22.9)	11 (61.1)	0.003
Caucasian ($n = 11$)	8 (11.4)	3 (16.7)	ns
African American ($n = 51$)	47 (67.1)	4 (22.2)	0.001
Diffuse ($n = 58$)	54 (77.1)	4 (22.2)	<0.0001
	54/58 (93.1) ^a	4/58 (6.9) ^a	<0.0001 ^a
Limited ($n = 30$)	16 (22.9)	14 (77.8)	<0.0001
	16/30 (53.3) ^b	14/30 (47.7) ^b	<0.0001 ^b
Renal crisis ($n = 5$)	4/14 (28.6)	1/8 (12.5)	ns

ANoA anti-nucleolar antibody, SSc systemic sclerosis, U3-RNP U3 ribonucleoprotein, LIA line immunoblot assay, IP immunoprecipitation, ns no significant difference, SD standard deviation

^a Comparison of anti-U3-RNP production in patients with diffuse cutaneous SSc

^b Comparison of anti-U3-RNP production in patients with limited cutaneous SSc

Fig. 2 ROC curves, sensitivity, specificity, positive predictive value and negative predictive value of the U3-RNP line immunoblot assay (LIA). **a** ROC curves for anti-U3-RNP detected by LIA compared with IP. **b** ROC curves for anti-U3-RNP/fibrillarin detected by LIA in patients with diffuse Systemic Sclerosis (SSc) and/or limited SSc versus healthy controls and for diffuse SSc versus limited SSc. **c** Performance characteristics of LIA calculated using the positivity cutoff (>10), borderline cutoff (6–10), and negative cutoff (<6) recommended by the manufacturer



the time of blood draw between patients that were positive vs negative for anti-U3-RNP. Anti-U3-RNP positivity was associated with a higher proportion of African American patients (67.1 vs 22.2 %; $p = 0.001$) and a lower proportion of Hispanic patients (22.9 vs 61.1 %; $p = 0.003$).

The U3-RNP LIA had a sensitivity of 100 %, specificity of 94.7 %, positive predictive value of 98.6 % and negative predictive value of 100 % compared to IP, with an AUC of 0.99 (Figs. 1a, 2a). ROC analysis comparing SSc patients with healthy controls demonstrated a sensitivity of 79.1 %, specificity of 100 %, positive predictive value of 100 % and negative predictive value of 70.8 %, with an AUC of 0.86 (Fig. 2b). DcSSc patients also demonstrated a significant ($p < 0.0001$) AUC of 0.94 (95 % CI 0.90–0.99) when compared to healthy controls. In contrast, ROC analysis comparing lcSSc patients with healthy controls demonstrated an AUC of 0.6845 (95 % CI 0.55–0.82). Similarly, ROC analysis comparing dcSSc patients to lcSSc patients demonstrated a significant AUC of 0.7848 (95 % CI 0.69–0.88). Thus, anti-U3-RNP measured by LIA is clinically useful for discriminating between SSc patients and healthy controls (Fig. 2), particularly for dcSSc.

Discussion

The increased prevalence of anti-U3-RNP in African American patients with dcSSc compared to other ethnic groups has been recognized in a number of sero-epidemiological studies [9, 15, 17, 18, 20–25]. In addition, the presence of anti-U3-RNP antibodies is considered a

specific marker for SSc and found in 5–8 % of these patients [4, 12, 19]. Despite the significance of this antibody in the diagnostic evaluation and management of SSc, its availability in routine clinical laboratories remains restricted due to methodologic limitations [4, 12, 19]. In this study, the diagnostic and clinical implications for testing U3-RNP antibodies in a multi-ethnic cohort of SSc patients by a conventional immunoassay (LIA) and the IP were comparable. The LIA is faster, less subjective and more cost-effective than IP enabling incorporation of anti-U3-RNP testing in a routine laboratory setting.

Of clinical significance, the presence of U3-RNP antibodies by LIA was associated with high prevalence in African Americans as well as dcSSc. Furthermore, the observation that anti-U3-RNP positive patients have a higher prevalence of renal crisis was supported by the trend toward a higher prevalence of renal crisis found in this study [9, 15, 17, 18], but data regarding kidney involvement was only available for 23.5 % of the cohort, which may explain why statistical significance was not achieved. However, of the patients with renal crisis included in this study, 80 % were positive for anti-U3-RNP. Previous studies have shown an association between anti-U3-RNP positivity and male gender and a younger age at disease onset [9, 15, 17–19]. This study demonstrated a similar trend, but failed to reach statistical significance likely due to the cross-sectional nature of the study, potential selection bias and the small number of men in the study. Of the men included in this study, 80 % were positive for anti-U3-RNP. Anti-U3-RNP positivity is associated with a nucleolar pattern [10, 11]. In our study, all of the patients were ANoA positive, but 4 (4 %) patients had ANoA in

combination with a nuclear speckled pattern. All four of these patients were negative for anti-U3-RNP by both IP and LIA (data not shown).

Detection of anti-U3-RNP using the EUROLINE LIA in samples that were negative for anti-U3-RNP by IP has previously been reported [18]. However, this reactivity was of low signal intensity and overlapped with that detected in controls. Similar results were observed in this study, but it remains to be determined whether this low-level reactivity is due to contaminants in the antigen preparation or the LIA being more sensitive than IP.

This study is not without limitations. Like most studies of this nature, this study is cross-sectional and limited by the paucity of clinical data for internal organ involvement. Secondly, the cohort is highly selected and does not represent routine patient population. Another potential setback in the study design is the limited ethnic diversity of the healthy control population which was comprised of predominantly Caucasian donors. Lastly, although a limited number of U3-RNP antibody negative SSc patients by IP were evaluated in the LIA, the analytical specificity in the context of other autoimmune diseases with a high incidence of myositis commonly associated with this antibody was not tested. This is particularly important as neither this study, nor others using EUROLINE systemic sclerosis LIA and IP have investigated the specificity of the assay for the detection of anti-U3-RNP with respect to other autoimmune diseases [4, 29–32].

Despite these limitations, the concordance between LIA and the IP assay considered the gold standard for detecting anti-U3-RNP antibodies represent a positive outlook for autoantibody testing in SSc. However, testing for anti-U3-RNP antibodies should only be considered in SSc patients who have positive ANA with a nucleolar staining pattern, particularly in African American patients, dcSSc and/or renal crisis. Overall, this study shows excellent agreement between the Euroline LIA and IP for the detection of antibodies against U3-RNP, making it an acceptable and attractive alternative to IP.

Acknowledgments The authors thank Dr. Minoru Satoh for testing the samples by IP at the University of Florida. EUROLINE Systemic Sclerosis (Nucleoli) Profile (IgG) kits were provided by EURO-IMMUN US, which was not involved in the study design or the writing of this manuscript.

References

- Koenig M, Dieudé M, Sénécal JL. Predictive value of antinuclear autoantibodies: the lessons of the systemic sclerosis autoantibodies. *Autoimmun Rev*. 2008;7(8):588–93.
- Mahler M, Fritzler MJ. Epitope specificity and significance in systemic autoimmune diseases. *Ann N Y Acad Sci*. 2010;1183:267–87.
- Mehra S, Walker J, Patterson K, Fritzler MJ. Autoantibodies in systemic sclerosis. *Autoimmun Rev*. 2013;12(3):340–54.
- Villalta D, Imbastro T, Di Giovanni S, Lauriti C, Gabini M, Turi MC, Bizzaro N. Diagnostic accuracy and predictive value of extended autoantibody profile in systemic sclerosis. *Autoimmun Rev*. 2012;12(2):114–20.
- Ceribelli A, Cavazzana I, Franceschini F, Airò P, Tincani A, Cattaneo R, Pauley BA, Chan EK, Satoh M. Anti-Th/To are common antinuclear autoantibodies in Italian patients with scleroderma. *J Rheumatol*. 2010;37(10):2071–5.
- Steen VD, Powell DL, Medsger TA Jr. Correlates between autoantibodies to nucleolar antigens and clinical features in patients with systemic sclerosis (scleroderma). *Arthritis Rheum*. 1988;31(4):525–32.
- Steen VD. Autoantibodies in systemic sclerosis. *Semin Arthritis Rheum*. 2005;35(1):35–42.
- Meyer O. Prognostic markers for systemic sclerosis. *Joint Bone Spine*. 2006;73(5):490–4.
- Hamaguchi Y. Autoantibody profiles in systemic sclerosis: predictive value for clinical evaluation and prognosis. *J Dermatol*. 2010;37(1):42–53.
- Bernstein RM, Steigerwald JC, Tan EM. Association of antinuclear and antinucleolar antibodies in progressive systemic sclerosis. *Clin Exp Immunol*. 1982;48(1):43–51.
- Ochs RL, Lischwe MA, Spohn WH, Busch H. Fibrillarin: a new protein of the nucleolus identified by autoimmune sera. *Biol Cell*. 1985;54(2):123–33.
- Reimer G, Pollard KM, Penning CA, Ochs RL, Lischwe MA, Busch H, Tan EM. Monoclonal autoantibody from a (New Zealand black x New Zealand white) F₁ mouse and some human scleroderma sera target an M_r 34,000 nucleolar protein of the U3 RNP particle. *Arthritis Rheum*. 1987;30:793–800.
- Baserga SJ, Yang XD, Steitz JA. An intact Box C sequence in the U3 snRNA is required for binding of fibrillarin, the protein common to the major family of nucleolar snRNPs. *EMBO J*. 1991;10(9):2645–51.
- Lischwe MA, Ochs RL, Reddy R, Cook RG, Yeoman LC, Tan EM, Reichlin M, Busch H. Purification and partial characterization of a nucleolar scleroderma antigen (Mr = 34,000; pI, 8.5) rich in NG, NG-dimethylarginine. *J Biol Chem*. 1985;260(26):14304–10.
- Pollard KM, Hultman P. Fibrillarin autoantibodies. In: Peter JB, Shoenfeld Y, editors. *Autoantibodies*. Amsterdam: Elsevier; 2007. p. 317–23.
- Yang JM, Hildebrandt B, Luderschmidt C, Pollard KM. Human scleroderma sera contain autoantibodies to protein components specific to the U3 small nucleolar RNP complex. *Arthritis Rheum*. 2003;48(1):210–7.
- Reimer G, Steen VD, Penning CA, Medsger TA Jr, Tan EM. Correlates between autoantibodies to nucleolar antigens and clinical features in patients with systemic sclerosis (scleroderma). *Arthritis Rheum*. 1988;31(4):525–32.
- Arnett FC, Reveille JD, Goldstein R, Pollard KM, Leaird K, Smith EA, Leroy EC, Fritzler MJ. Autoantibodies to fibrillarin in systemic sclerosis (scleroderma). An immunogenetic, serologic, and clinical analysis. *Arthritis Rheum*. 1996;39(7):1151–60.
- Aggarwal R, Lucas M, Fertig N, Oddis CV, Medsger TA Jr. Anti-U3 RNP autoantibodies in systemic sclerosis. *Arthritis Rheum*. 2009;60(4):1112–8.
- Tormey VJ, Bunn CC, Denton CP, Black CM. Anti-fibrillarin antibodies in systemic sclerosis. *Rheumatology (Oxford)*. 2001;40(10):1157–62.
- Sharif R, Fritzler MJ, Mayes MD, Gonzalez EB, McNearney TA, Draeger H, Baron M, Canadian Scleroderma Research Group, Furst DE, Khanna DK, del Junco DJ, Molitor JA, Schiopu E, Phillips K, Seibold JR, Silver RM, Simms RW, GENISOS Study Group, Perry M, Rojo C, Charles J, Zhou X, Agarwal SK,

- Reveille JD, Assassi S, Arnett FC. Anti-fibrillar antibody in African American patients with systemic sclerosis: immunogenetics, clinical features, and survival analysis. *J Rheumatol*. 2011;38(8):1622–30.
22. Okano Y, Steen VD, Medsger TA Jr. Autoantibody to U3 nucleolar ribonucleoprotein (fibrillar) in patients with systemic sclerosis. *Arthritis Rheum*. 1992;35(1):95–100.
23. Kuwana M, Okano Y, Kaburaki J, Tojo T, Medsger TA Jr. Racial differences in the distribution of systemic sclerosis-related serum antinuclear antibodies. *Arthritis Rheum*. 1994;37(6):902–6.
24. Reveille JD, Fischbach M, McNearney T, Friedman AW, Aguilar MB, Lisse J, Fritzler MJ, Ahn C, Arnett FC, GENISOS Study Group. Systemic sclerosis in 3 US ethnic groups: a comparison of clinical, sociodemographic, serologic, and immunogenetic determinants. *Semin Arthritis Rheum*. 2001;30(5):332–46.
25. Krzyszcak ME, Li Y, Ross SJ, Ceribelli A, Chan EK, Bubb MR, Sobel ES, Reeves WH, Satoh M. Gender and ethnicity differences in the prevalence of scleroderma-related autoantibodies. *Clin Rheumatol*. 2011;30(10):1333–9.
26. Kipnis RJ, Craft J, Hardin JA. The analysis of antinuclear and antinucleolar autoantibodies of scleroderma by radioimmuno-precipitation assays. *Arthritis Rheum*. 1990;33(9):1431–7.
27. Harvey G, Black C, Maddison P, McHugh N. Characterization of antinucleolar antibody reactivity in patients with systemic sclerosis and their relatives. *J Rheumatol*. 1997;24(3):477–84.
28. van den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, Matucci-Cerinic M, Naden RP, Medsger TA Jr, Carreira PE, Riemekasten G, Clements PJ, Denton CP, Distler O, Allanore Y, Furst DE, Gabrielli A, Mayes MD, van Laar JM, Seibold JR, Czirjak L, Steen VD, Inanc M, Kowal-Bielecka O, Müller-Ladner U, Valentini G, Veale DJ, Vonk MC, Walker UA, Chung L, Collier DH, Csuka ME, Fessler BJ, Guiducci S, Herrick A, Hsu VM, Jimenez S, Kahaleh B, Merkel PA, Sierakowski S, Silver RM, Simms RW, Varga J, Pope JE. 2013 classification criteria for systemic sclerosis: an American college of rheumatology/European league against rheumatism collaborative initiative. *Arthritis Rheum*. 2013;65(11):2737–47.
29. Bonroy C, Van Praet J, Smith V, Van Steendam K, Mimori T, Deschepper E, Deforce D, Devreese K, De Keyser F. Optimization and diagnostic performance of a single multiparameter lineblot in the serological workup of systemic sclerosis. *J Immunol Methods*. 2012;379(1–2):53–60.
30. Wielosz E, Dryglewska M, Majdan M. Serological profile of patients with systemic sclerosis. *Postepy Hig Med Dosw (Online)*. 2014;68:987–91.
31. Low AH, Wong S, Thumboo J, Ng SC, Lim JY, Ng X, Earnest A, Fong KY. Evaluation of a new multi-parallel line immunoassay for systemic sclerosis-associated antibodies in an Asian population. *Rheumatology (Oxford)*. 2012;51(8):1465–70.
32. Graf SW, Hakendorf P, Lester S, Patterson K, Walker JG, Smith MD, Ahern MJ, Roberts-Thomson PJ. South Australian Scleroderma Register: autoantibodies as predictive biomarkers of phenotype and outcome. *Int J Rheum Dis*. 2012;15(1):102–9.




Clinical correlates of monospecific anti-PM75 and anti-PM100 antibodies in a tri-nation cohort of 1574 systemic sclerosis subjects

Michael Wodkowski, Marie Hudson, Susanna Proudman, Jennifer Walker, Wendy Stevens, Mandana Nikpour, Shervin Assassi, Maureen D. Mayes, Solène Tatibouet, Mianbo Wang, Canadian Scleroderma Research Group (CSRG), Australian Scleroderma Interest Group (ASIG), Genetics versus Environment in Scleroderma Outcome Study (GENISOS), Murray Baron & Marvin J. Fritzler


To cite this article: Michael Wodkowski, Marie Hudson, Susanna Proudman, Jennifer Walker, Wendy Stevens, Mandana Nikpour, Shervin Assassi, Maureen D. Mayes, Solène Tatibouet, Mianbo Wang, Canadian Scleroderma Research Group (CSRG), Australian Scleroderma Interest Group (ASIG), Genetics versus Environment in Scleroderma Outcome Study (GENISOS), Murray Baron & Marvin J. Fritzler (2015) Clinical correlates of monospecific anti-PM75 and anti-PM100 antibodies in a tri-nation cohort of 1574 systemic sclerosis subjects, *Autoimmunity*, 48:8, 542-551, DOI: [10.3109/08916934.2015.1077231](https://doi.org/10.3109/08916934.2015.1077231)

To link to this article: <http://dx.doi.org/10.3109/08916934.2015.1077231>

 View supplementary material 

 Published online: 31 Aug 2015.

 Submit your article to this journal 

 Article views: 272

 View related articles 

 View Crossmark data 

ORIGINAL ARTICLE

Clinical correlates of monospecific anti-PM75 and anti-PM100 antibodies in a tri-nation cohort of 1574 systemic sclerosis subjects

Michael Wodkowski¹, Marie Hudson^{1,2,3}, Susanna Proudman^{4,5}, Jennifer Walker⁶, Wendy Stevens⁷, Mandana Nikpour^{7,8}, Shervin Assassi⁹, Maureen D. Mayes⁹, Solène Tatibouet³, Mianbo Wang³, Canadian Scleroderma Research Group (CSRG)*, Australian Scleroderma Interest Group (ASIG)**, Genetics versus Environment in Scleroderma Outcome Study (GENISOS)***, Murray Baron^{1,2,3}, and Marvin J. Fritzler¹⁰

¹Department of Medicine, McGill University, Montréal, Quebec, Canada, ²Division of Rheumatology, Jewish General Hospital, Montréal, Quebec, Canada, ³Lady Davis Institute, Jewish General Hospital, Montréal, Quebec, Canada, ⁴Rheumatology Unit, Royal Adelaide Hospital, Adelaide, Australia, ⁵Discipline of Medicine, University of Adelaide, Bedford Park, Australia, ⁶Department of Allergy and Immunology, Flinders Medical Centre, Bedford Park, Australia, ⁷Department of Rheumatology, St. Vincent's Hospital Melbourne, Fitzroy, Victoria, Australia, ⁸Department of Medicine, The University of Melbourne at St. Vincent's Hospital, Melbourne, Victoria, Australia, ⁹Division of Rheumatology and Immunogenetics, University of Texas Health Science Centre at Houston, Houston, TX, USA, and ¹⁰Faculty of Medicine, University of Calgary, Calgary, Alberta, Canada

Abstract

Objective: Autoantibodies directed against the two principal antigens of the human exosome complex, PM75 and PM100, are present in systemic sclerosis (SSc) sera and have been associated with myositis and calcinosis. However, there is a paucity of data on the clinical correlates of these autoantibodies separately and in the absence of other SSc-specific antibodies. The aim of this study was to assess the clinical correlates of monospecific anti-PM75 and anti-PM100 in SSc. **Methods:** A tri-nation cohort of 1574 SSc subjects was formed, clinical variables were harmonized and sera were tested for anti-PM75 and anti-PM100 antibodies using a line immunoassay. **Results:** Forty-eight (3.0%) subjects had antibodies against PM75 and 18 (1.1%) against PM100. However, only 16 (1%) had monospecific anti-PM75 antibodies and 11 (0.7%) monospecific anti-PM100 antibodies (i.e. in isolation of each other and other SSc-specific antibodies). Monospecific profiles of each autoantibody included more calcinosis. An increased frequency of myositis was only seen in subjects positive for both anti-PM75 and anti-PM100 antibodies. Lung disease was only associated with anti-PM75 and subjects with anti-PM100 antibodies had better survival compared to other antibody subsets. **Conclusion:** The prevalence of monospecific anti-PM75 and anti-PM100 antibodies in this large SSc cohort was low. Disease features associated with anti-PM/Scl antibodies may depend on particular and possibly multiple antigen specificities. However, due to the small samples, these results need to be interpreted with caution. International collaborations are key to understanding the clinical correlates of uncommon serological profiles in SSc.

Keywords

Autoantibodies, anti-PM/Scl antibodies, autoimmunity, observational cohort, systemic sclerosis

History

Received 7 January 2015
Revised 2 June 2015
Accepted 19 July 2015
Published online 31 August 2015

*Investigators of the Canadian Scleroderma Research Group: J. Pope, London, Ontario; M. Baron, Montreal, Quebec; J. Markland, Saskatoon, Saskatchewan; D. Robinson, Winnipeg, Manitoba; N. Jones, Edmonton, Alberta; N. Khalidi, Hamilton, Ontario; P. Docherty, Moncton, New Brunswick; E. Kaminska, Calgary, Alberta; A. Masetto, Sherbrooke, Quebec; E. Sutton, Halifax, Nova Scotia; J-P. Mathieu, Montreal, Quebec; M. Hudson, Montreal, Quebec; S. Ligier, Montreal, Quebec; T. Grodzicky, Montreal, Quebec; S. LeClercq, Calgary, Alberta; C. Thorne, Newmarket, Ontario; G. Gyger, Montreal, Quebec; D. Smith, Ottawa, Ontario; P.R. Fortin, Quebec, Quebec; M. Larché, Hamilton, Ontario; M. Fritzler, Advanced Diagnostics Laboratory, Calgary, Alberta.

**Investigators of the Australian Scleroderma Interest Group: C. Hill, Adelaide, South Australia; S. Lester, Adelaide, South Australia; P. Nash, Sunshine Coast, Queensland; M. Nikpour, Melbourne, Victoria; J. Roddy, Perth, Western Australia; K. Patterson, Adelaide, South Australia; S. Proudman, Adelaide, South Australia; M. Rischmueller, Adelaide, South Australia; J. Sahhar, Melbourne, Victoria; W. Stevens, Melbourne, Victoria; J. Walker, Adelaide, South Australia; J. Zochling, Hobart, Tasmania.

***Investigators of GENISOS: Shervin Assassi, Houston, Texas; Maureen D. Mayes, Houston, Texas; Terry A. McNearney, Galveston, Texas; Gloria Salazar, Houston, Texas

Correspondence: Dr. Marie Hudson, Jewish General Hospital, Room A-725, 3755 Côte Sainte-Catherine Road, Montreal, Quebec, H3T 1E2, Canada. Tel: +514-340-8222 ext. 3476. Fax: +514-340-7906. E-mail: marie.hudson@mcgill.ca

Introduction

The wide spectrum of serum autoantibodies in systemic sclerosis (SSc) has been extensively studied in the context of the diagnostic uncertainty, varying clinical phenotypes and unpredictable course of the disease [1]. SSc-specific antibodies, namely anti-topoisomerase I (ATA), anti-centromere (ACA) and anti-RNA polymerase III (ARNAP), and a number of SSc-associated antibodies (anti-Ro52/TRIM21, U1RNP and PM/Scl antibodies) are now understood to be biomarkers of diagnosis and prognosis and may contribute to disease pathogenesis [2,3].

Among these, autoantibodies against the human exosome complex were discovered as early as 1977 [4]. They were designated as anti-PM/Scl since they were commonly found in patients with polymyositis and scleroderma overlap syndrome (20–55%) in contrast to SSc alone (3–11%) [5]. To date, anti-PM/Scl antibodies have been associated with a variety of clinical profiles, of which the associations with inflammatory myositis and calcinosis have been the most consistently reported [6,7].

The PM/Scl complex consists of as many as 16 proteins that are located primarily in the nucleolus and are involved in RNA processing and degradation. The major B cell targets of the PM/Scl complex have been identified as PM100 (and its major epitope called PM1-Alpha) and PM75, both named for their molecular masses [6]. Commercial assays to detect anti-PM100 and anti-PM75 antibodies are available and a few studies have reported distinct clinical associations with these autoantibodies [8,9]. However, to our knowledge, there have been no studies to date that have examined the clinical profiles of SSc patients with anti-PM75 and anti-PM100 antibodies present in isolation (i.e. monospecific) of each other and of other SSc-specific antibodies (ACA, ATA or ARNAP), possibly because of the paucity of subjects with these serological profiles. It is possible that clinical characteristics associated with other SSc autoantibodies may obscure the distinct phenotypes associated with anti-PM75 and anti-PM100 antibodies. The aim of this study was therefore to identify the demographic, clinical and serological characteristics of SSc subjects with monospecific anti-PM75 and anti-PM100 antibodies in a large international, multi-centered cohort.

Methods

Sources of data

The study subjects were SSc patients enrolled in the Canadian Scleroderma Research Group (CSRG), the Australian Scleroderma Interest Group (ASIG) or the American GENISOS cohorts. Briefly, patients in the CSRG are recruited from 15 sites across Canada, must have a diagnosis of SSc verified by an experienced rheumatologist, be >18 years of age and be fluent in English or French. Over 98% of the cohort meet the 2013 ACR/EULAR classification criteria for SSc [10]. Patients in the Australian Scleroderma Cohort Study (ASCS) are recruited by ASIG investigators from 12 Australian centers specializing in the care of patients with SSc, according to similar inclusion criteria. All patients fulfilled either the 1980 preliminary ACR criteria for classification of SSc, or the Medsger criteria for limited

SSc [11]. The Genetics versus Environment in Scleroderma Outcome Study (GENISOS) cohort is a longitudinal cohort of patients with early SSc. Patients were enrolled within 5 years of disease onset as determined by the first non-Raynaud's phenomenon symptom from three University of Texas institutions at Houston, San Antonio, and Galveston. All enrolled patients fulfill the 2013 ACR/EULAR classification criteria for SSc [10].

Ethics committee approval for this study was obtained at McGill University (Montreal, Canada) and at all participating CSRG, ASIG, and GENISOS study sites. All subjects provided informed written consent to participate in the data collection protocol. The CSRG subjects included in this study were those whose baseline visits were between September 2004 and February 2011, for ASIG between December 2007 and September 2011 and GENISOS between January 1998 and January 2013, and who had complete serological profiles for anti-PM75 and anti-PM100 antibodies as detected by the methods described below.

Clinical variables

Patients recruited into this study underwent standardized medical evaluation including medical histories, physical examinations and laboratory investigations. Demographic information regarding age, sex and ethnicity was collected by patient self-report. Disease duration was recorded by study physicians and defined as the interval between the onset of the first non-Raynaud disease manifestation and baseline study visit.

Skin involvement was assessed using the modified Rodnan skin score, a widely used clinical assessment where the examining rheumatologist records the degree of skin thickening ranging from 0 (no involvement) to 3 (severe thickening) in 17 areas (total score range 0–51). Limited cutaneous disease (lcSSc) was defined as skin involvement distal to the elbows and knees with or without facial involvement; diffuse cutaneous disease (dcSSc) was defined as skin involvement proximal to the elbows and knees, with or without truncal involvement. Those with a clinical diagnosis of SSc but no skin involvement were included with the lcSSc subset.

A history of inflammatory myositis, calcinosis, inflammatory arthritis and scleroderma renal crisis was recorded by a study physician. Creatine kinase (CK) was measured by local laboratories.

To assess gastrointestinal involvement, patients answered yes/no to six questions concerning gastroesophageal reflux disease, dysphagia, antibiotics for bacterial overgrowth, episodes of pseudo-obstruction, fecal incontinence and hyperalimentation.

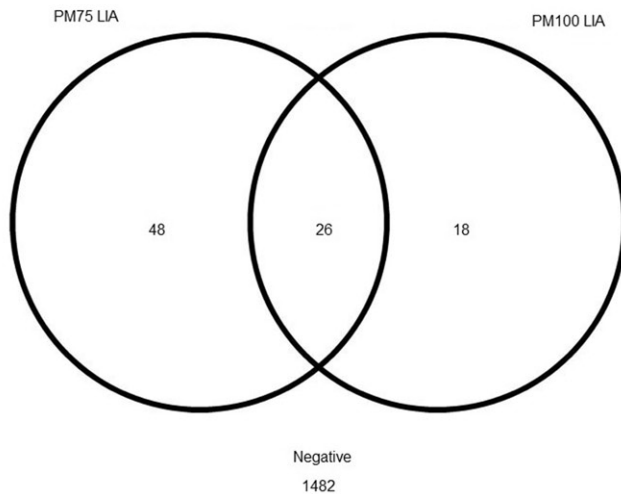
The presence of interstitial lung disease (ILD) was determined using a clinical decision rule that was recently published [12]. Using this algorithm, ILD was considered present if a high resolution computed tomography (HRCT) scan of the lung was interpreted by an experienced radiologist as showing ILD or, in the case where no HRCT is available, if either a chest X-ray was reported as showing either increased interstitial markings (not thought to be due to congestive heart failure) or fibrosis, and/or if a study physician reported the presence of typical "velcro-like crackles" on physical examination.

Table 1. Baseline characteristics according to antibody status (N = 1574). Subjects in the anti-PM75, anti-PM100, ACA, ATA and ARNAP groups have monospecific antibodies, i.e. exclusive of each other.

	PM75 (N=16)		PM100 (N=11)		ACA (N=487)		ATA (N=251)		ARNAP (N=206)		Others (N=603)	
	% (N) or mean ± SD	NA	% (N) or mean ± SD	NA	% (N) or mean ± SD	NA	% (N) or mean ± SD	NA	% (N) or mean ± SD	NA	% (N) or mean ± SD	NA
Sociodemographic variables												
Females	81.3% (13)	0	90.9% (10)	0	94.1% (458)	0	79.3% (199)	0	83.0% (171)	0	83.6% (504)	0
Age, years	56.6 ± 11.7	0	47.5 ± 13.3	0	59.4 ± 11.6	1	52.1 ± 13.0	0	53.6 ± 12.2	0	53.4 ± 13.0	1
Race/ethnicity		0		1		18		10		1		22
White	93.8% (15)		100.0% (10)		88.7% (416)		76.8% (185)		84.4% (173)		76.6% (445)	
Black	0.0% (0)		0.0% (0)		1.1% (5)		5.8% (14)		2.9% (6)		7.9% (46)	
Hispanic	0.0% (0)		0.0% (0)		3.6% (17)		5.8% (14)		4.4% (9)		8.6% (50)	
Disease duration, years	10.9 ± 10.9	0	11.7 ± 9.8	0	11.6 ± 9.9	2	8.3 ± 8.6	4	7.2 ± 7.3	2	9.1 ± 9.1	4
Age at disease onset	45.7 ± 12.9	0	35.8 ± 14.1	0	47.8 ± 13.8	3	43.7 ± 13.8	4	46.4 ± 12.0	2	44.3 ± 13.8	5
Disease subsets												
Limited cutaneous disease	75.0% (12)	0	90.9% (10)	0	87.3% (425)	0	46.8% (117)	1	20.1% (41)	2	59.6% (358)	2
Diffuse cutaneous disease	25.0% (4)	0	9.1% (1)	0	12.7% (62)	0	53.2% (133)	1	79.9% (163)	2	40.4% (243)	2
Clinical variables												
Modified Rodnan skin score	10.8 ± 9.6	0	4.4 (3.3)	0	7.3 ± 6.8	20	14.0 ± 10.0	7	20.3 ± 11.4	6	10.8 ± 10.1	16
Inflammatory myositis	7.7% (1)	3	0.0% (0)	3	4.7% (20)	62	10.0% (22)	30	9.8% (18)	23	12.4% (68)	54
Calcinosis	37.5% (6)	0	36.4% (4)	0	31.6% (151)	9	22.3% (56)	0	27.7% (56)	4	22.1% (132)	6
Arthritis	31.3% (5)	0	36.4% (4)	0	22.6% (110)	0	32.3% (81)	0	34.8% (71)	2	28.7% (172)	3
GI symptoms												
GERD	87.5% (14)	0	54.6% (6)	0	86.2% (418)	2	81.2% (203)	1	83.4% (171)	1	77.5% (466)	2
Dysphagia	68.8% (11)	0	45.5% (5)	0	59.1% (282)	10	46.4% (115)	3	43.1% (88)	2	51.9% (307)	11
Antibiotics for bacterial overgrowth	14.3% (2)	2	0.0% (0)	1	7.7% (35)	30	6.6% (15)	24	7.1% (13)	23	5.9% (29)	108
Episodes of pseudo-obstruction	0.0% (0)	0	0.0% (0)	0	4.1% (20)	1	2.0% (5)	2	2.9% (6)	2	1.8% (11)	6
Fecal incontinence	28.6% (4)	2	12.5% (1)	3	26.8% (113)	66	9.0% (18)	50	16.2% (27)	39	13.1% (54)	192
Hyperalimantation	0.0% (0)	11	0.0% (0)	5	1.5% (4)	215	3.5% (5)	109	4.0% (5)	81	2.7% (9)	267
Number of GI symptoms (0–6)	1.9 ± 0.9	0	1.1 ± 0.9	0	1.8 ± 1.0	0	1.4 ± 0.9	1	1.5 ± 1.0	0	1.5 ± 0.9	2
Scleroderma renal crisis	0.0% (0)	1	0.0% (0)	0	0.0% (0)	5	2.0% (5)	2	15.3% (31)	3	4.0% (24)	7
Interstitial lung disease	50.0% (8)	0	20.0% (2)	1	17.8% (84)	16	56.1% (139)	3	37.3% (76)	2	40.7% (239)	16
Pulmonary hypertension	21.4% (3)	2	0.0% (0)	2	16.2% (68)	68	13.3% (25)	49	13.4% (23)	34	11.9% (53)	156
Laboratory tests												
CK at baseline visit	90.5 ± 51.0	3	69.1 ± 27.5	4	88.0 ± 70.6	81	107.0 ± 80.5	32	107.2 ± 121.7	34	128.7 ± 241.3	83
Percentage with CK > 3x normal at baseline	0.0% (0)	3	0.0% (0)	4	0.8% (3)	100	0.0% (0)	47	1.9% (3)	45	2.4% (12)	106
Highest CK during follow up	134.3 ± 151.3	1	81.5 ± 39.3	0	121.4 ± 166.8	9	143.7 ± 124.9	8	140.9 ± 134.3	5	188.7 ± 611.3	13
Percentage with CK > 3x normal during follow up	0.0% (0)	1	0.0% (0)	0	2.0% (9)	31	1.8% (4)	24	2.7% (5)	18	2.8% (16)	37
Serology												
ANA positive	86.7% (13)	1	100.0% (11)	0	99.4% (484)	0	98.0% (246)	0	96.1% (197)	1	88.4% (533)	0

ACA: anti-centromere antibodies; ATA: anti-topoisomerase I antibodies; ARNAP: anti-RNA polymerase III antibodies; SD: standard deviation; NA: not available; GI: gastrointestinal; GERD: gastroesophageal reflux disease; HRCT: high resolution computed tomography scans of the lungs; CXR: chest X-ray; CK: creatinine kinase; ANA: anti-nuclear antibody.

(A) Subjects with non-exclusive PM75 and PM100 (i.e. possible overlap with ACA, ATA, ARNAP)



(B) Subjects with monospecific PM75 and PM100 (i.e. no overlap with ACA, ATA, ARNAP)

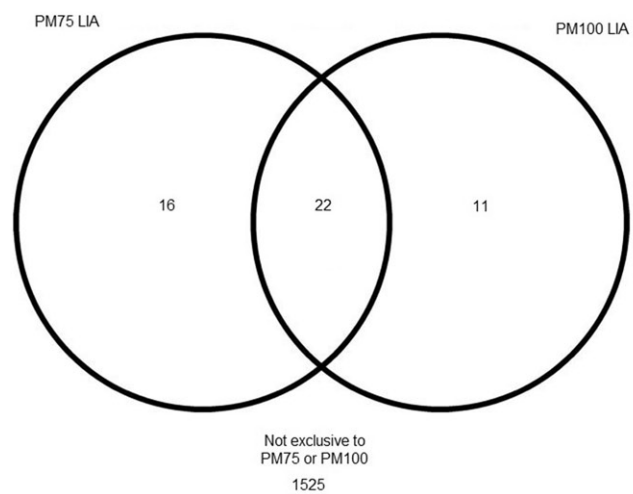


Figure 1. Venn diagrams showing overlap between anti-PM75 and anti-PM100 antibodies. (A) Subjects with non-exclusive PM75 and PM100 (i.e. possible overlap with ACA, ATA, ARNAP). (B) Subjects with monospecific PM75 and PM100 (i.e. no overlap with ACA, ATA, ARNAP). ACA: anticentromere antibodies; ATA: anti-topoisomerase I antibodies; ARNAP: anti-RNA polymerase III antibodies.

Pulmonary hypertension was defined as an estimated systolic pulmonary artery pressure (sPAP) ≥ 45 mmHg measured using the Doppler flow measurement of the tricuspid regurgitant jet on cardiac echocardiography (an estimate that correlates strongly with right heart catheter studies) [13] for CSRG and GENISOS subjects, or mean pulmonary artery pressure (mPAP) >25 mmHg with a pulmonary capillary wedge pressure (PCWP) <15 mmHg on right heart catheterization for ASIG subjects.

Serology

Autoantibody analysis of the CSRG and GENISOS cohorts were performed in a central laboratory, Mitogen Advanced Diagnostics Laboratory, University of Calgary and the ASIG analyses were performed using an identical immunoassay kit and protocols. Serum aliquots were stored at -80°C until needed for diagnostic assays. Antinuclear antibodies were detected by indirect immunofluorescence (IIF) performed on HEp-2 cells (ImmunoConcepts, Sacramento, CA, USA). Anti-PM75 and anti-PM100, ACA (CENP A and CENP B), ATA and ARNAP (RP11 and RP155) antibodies were detected by Euroline systemic sclerosis profile line immunoassay (LIA) (Euroimmun, Luebeck, Germany) according to manufacturer's instructions (<http://www.euroimmun.com/index.php?id=anwendungen&L=1>). With the intent of optimizing specificity, antibodies were reported as absent (negative, equivocal and low titres) and present (moderate and high titres).

Statistical analysis

Descriptive statistics were used to summarize the baseline demographic and clinical characteristics of the patients. Given the exploratory nature of the analysis and the small samples in the sub-groups, clinically relevant numerical differences between sub-groups were considered informative. Exploratory statistical analyses were also performed using Chi-square tests, Fisher's exact tests and Mann-Whitney's U tests, as indicated. Kaplan Meier analysis was used to compare survival between autoantibody subsets and log

rank p values were reported. No correction for multiple testing was done and $p < 0.05$ was considered statistically significant. All statistical analyses were performed with SAS v.9.2 (SAS Institute, Cary, NC).

Results

This study included 1574 SSc subjects who had complete serological profiles for anti-PM75 and anti-PM100 antibodies (Table 1). Of these, 48 (3.0%) had antibodies detected against PM75, 18 (1.1%) against PM100 and 26 (1.7%) against both PM75 and PM100 (Figure 1A). However, only 16 (1%) had monospecific anti-PM75 antibodies and 11 (0.7%) monospecific anti-PM100 antibodies (i.e. in isolation of each other and other SSc-specific antibodies; Figure 1B). In addition, there were 487 (30.9%) subjects with monospecific ACA, 251 (15.9%) monospecific ATA and 206 (13.1%) monospecific ARNAP. Considering the small sample number in the anti-PM75 and anti-PM100 groups, sociodemographic characteristics were generally similar between the various antibody groups, except for the fact that the anti-PM100 group had a younger age at disease onset (35.8 years) in comparison to anti-PM75 (45.7 years) and to the other antibody groups (43.7–47.8 years; Table 1 and Supplementary Table 1).

The majority of anti-PM75 and anti-PM100 patients had lcSSc (75.0% and 90.9%, respectively), which was similar to ACA patients (87.3%) but higher than those with ATA (46.8%) and ARNAP (20.1%) antibodies (Table 1 and Supplementary Table 1). The modified Rodnan skin score was 10.8 ± 9.6 and 4.4 ± 3.3 in the anti-PM75 and anti-PM100 groups, compared to 7.3 ± 6.8 , 14.0 ± 10.0 and 20.3 ± 11.4 in the ACA, ATA and ARNAP groups, respectively (Table 1 and Supplementary Table 1).

The prevalence of inflammatory myositis was low in the patient groups with monospecific anti-PM/Scl antibodies (1/13 (7.7%) in the anti-PM75 group and 0/8 in the anti-PM100 group) and was lying in the same range as in those patients with

Table 2. Baseline characteristic all subjects with anti-PM75 and anti-PM100 antibodies, including those with antibodies overlapping with ACA, ATA and ARNAP.

	Anti-PM75 (N = 48)		Anti-PM100 (N = 18)		Anti-PM75 and anti-PM100 positive (N = 26)		Anti-PM75 and anti-PM100 Negative (N = 1482)	
	% (N) or mean ± SD	NA	% (N) or mean ± SD	NA	% (N) or mean ± SD	NA	% (N) or mean ± SD	NA
Socio-demographics								
Females	85.4% (41)	0	88.9% (16)	0	84.6% (22)	0	86.1% (1276)	0
Age, years	55.3 ± 12.8	0	48.2 ± 13.7	0	54.8 ± 12.3	0	55.2 ± 12.8	2
Race/ethnicity		1		1		1		49
White	89.4% (42)		94.1% (16)		84.0% (21)		81.3% (1165)	
Black	0.0% (0)		0.0% (0)		4.0% (1)		4.9% (70)	
Hispanic	2.1% (1)		0.0% (0)		8.0% (2)		6.1% (87)	
Disease duration, years	10.8 ± 9.7	0	11.2 ± 10.3	0	10.4 ± 11.8	0	9.5 ± 9.1	12
Age at disease onset	44.4 ± 13.1	0	37.1 ± 12.7	0	44.4 ± 15.0	0	45.7 ± 13.7	14
Disease subsets								
Limited, N (%)	68.8% (33)	0	83.3% (15)	0	69.2% (18)	0	60.7% (897)	5
Diffuse, N (%)	31.3% (15)	0	16.7% (3)	0	30.8% (8)	0	39.3% (580)	
Clinical variables								
Modified Rodnan skin score, mean ± SD	10.6 ± 8.8	2	8.1 ± 11.3	0	9.1 ± 9.0	0	11.5 ± 10.3	47
Inflammatory myositis	5.3% (2)	10	0.0% (0)	4	36.0% (9)	1	8.9% (118)	160
Calcinosis	36.2% (17)	1	33.3% (6)	0	42.3% (11)	0	25.3% (371)	18
Arthritis	27.1% (13)	0	27.8% (5)	0	42.3% (11)	0	28.0% (414)	5
GI symptoms								
GERD/reflux	85.4% (41)	0	61.1% (11)	0	57.7% (15)	0	82.0% (1211)	6
Dysphagia	57.4% (27)	1	50.0% (9)	0	32.0% (8)	1	52.4% (764)	24
Antibiotics for bacterial overgrowth	4.5% (2)	4	0.0% (0)	1	4.3% (1)	3	7.0% (91)	180
Episodes of pseudo-obstruction	8.3% (4)	0	0.0% (0)	0	0.0% (0)	0	2.6% (38)	11
Fecal incontinence	26.2% (11)	6	13.3% (2)	3	11.1% (2)	8	17.6% (202)	335
Hyperalimantation	0.0% (0)	36	0.0% (0)	9	0.0% (0)	7	2.7% (23)	636
Number of GI symptoms (0–6)	1.8 ± 1.0	0	1.2 ± 0.9	0	1.0 ± 1.0	0	1.6 ± 1.0	3
Scleroderma renal crisis	2.1% (1)	1	5.6% (1)	0	4.0% (1)	1	3.9% (57)	16
Interstitial lung disease	48.9% (23)	1	11.8% (2)	1	42.3% (11)	0	35.4% (512)	36
Pulmonary hypertension	17.5% (7)	8	0.0% (0)	3	14.3% (3)	5	13.6% (162)	295
Lab tests								
Baseline visit	96.0 ± 52.5	11	84.5 ± 35.5	6	176.8 ± 232.4	5	108.8 ± 167.0	215
Percentage with CK > 3x normal baseline	0.0% (0)	12	0.0% (0)	6	5.0% (1)	6	1.4% (17)	281
Highest CK during follow up	160.1 ± 288.2	1	99.8 ± 54.8	1	159.2 ± 139.3	3	153.4 ± 405.2	31
Percentage with CK > 3x normal in f/u	2.2% (1)	2	0.0% (0)	1	4.5% (1)	4	2.3% (32)	104
Serology								
ANA positive	93.6% (44)	1	100.0% (18)	0	100.0% (26)	0	94.3% (1396)	1

ACA: anti-centromere antibodies; ATA: anti-topoisomerase I antibodies; ARNAP: anti-RNA polymerase III antibodies; SD: standard deviation; NA: not available; GI: gastrointestinal; GERD: gastroesophageal reflux disease; HRCT: high resolution computed tomography scans of the lungs; CXR: chest X-ray; CK: creatinine kinase; ANA: anti-nuclear antibody.

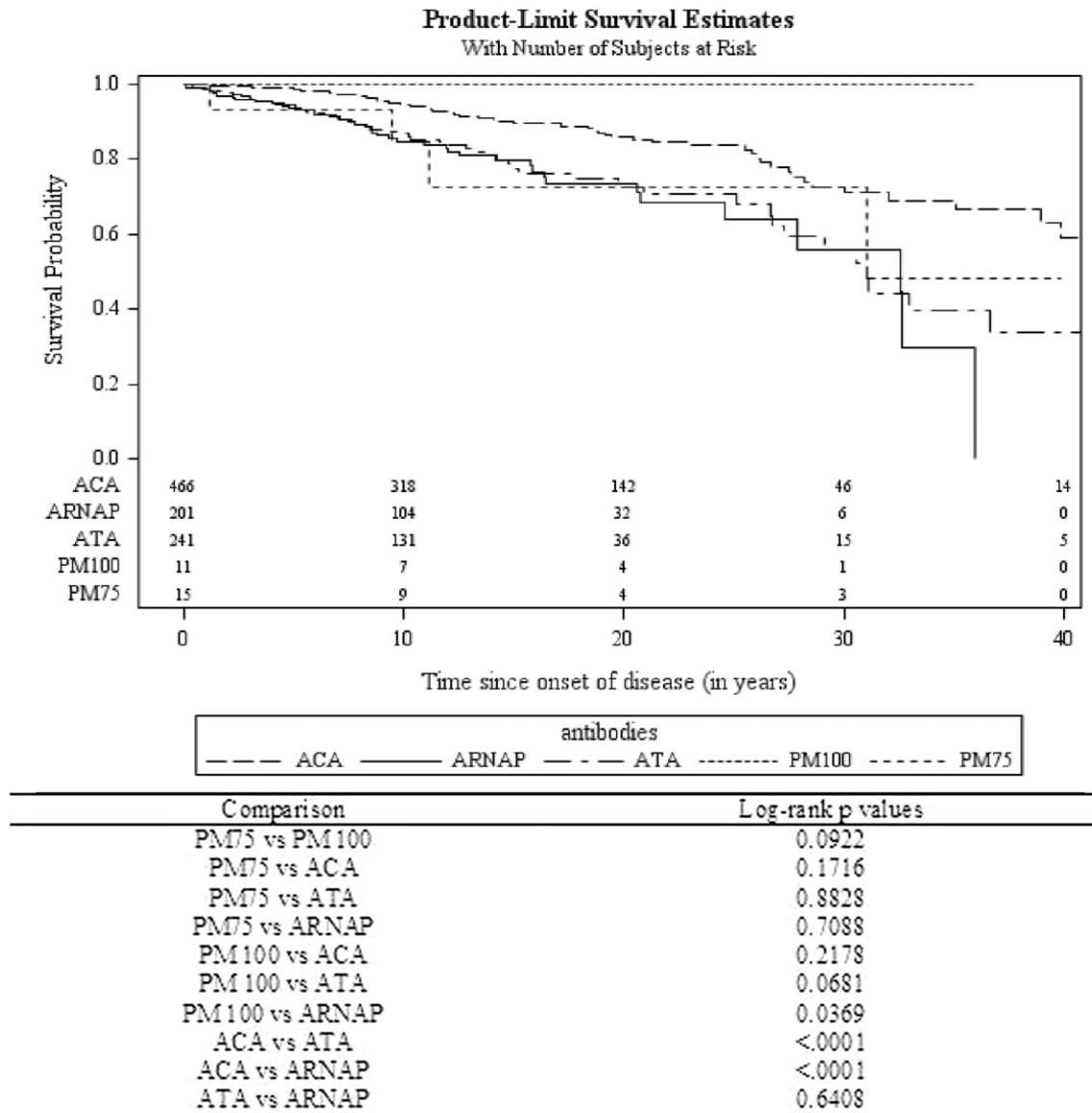


Figure 2. Kaplan–Meier survival analysis comparing subjects with monospecific anti-PM75, anti-PM100, ACA, ATA and ARNAP.

ACA (20/425, 4.7%), ATA (22/221, 10.0%) and ARNAP (18/183, 9.8%) (Table 1; differences between monospecific anti-PM/Scl groups and all other groups not significant, Supplementary Table 1). On the other hand, the highest prevalence of inflammatory myositis (9/25, 36%) was among the subjects with both anti-PM75 and anti-PM100 antibodies (vs. 120/1374 [8.7%]) of all other patients, odds ratio 5.9, 95% confidence interval 2.5, 13.6, $p = 0.0002$; Table 2).

In the monospecific anti-PM75 group, a trend towards relatively high rates of calcinosis (37.5%) and gastrointestinal symptoms (gastroesophageal reflux disease [GERD] 87.5%, dysphagia 68.8%, antibiotics for bacterial overgrowth 14.3% and fecal incontinence 28.6%) was seen (Table 1). ILD was also common (50.0%), second only to the ATA group (56.1%), and pulmonary hypertension (21.4%) was most frequent in this antibody group. However, probably due to the low numbers, few of the comparisons between the monospecific anti-PM75 and all other groups reached statistical significance (Supplementary Table 1).

The monospecific anti-PM100 group also showed a relatively high rate of calcinosis (36.4%) but low rates of

gastrointestinal symptoms (mean number of symptoms 1.1, range 0–6). ILD was less common compared to the anti-PM75 and other antibody groups, and none of the subjects had pulmonary hypertension. Again, only a few of the comparisons reached statistical significance (Supplementary Table 1).

In unadjusted survival analysis (Figure 2), anti-PM100 was associated with a significantly better survival compared to ARNAP (log rank $p = 0.0369$) and there was a trend toward better survival compared to ATA (log rank $p = 0.0681$) and anti-PM75 (log rank $p = 0.0922$).

Discussion

We sought to describe the distinct clinical phenotypes associated with anti-PM75 and anti-PM100 antibodies in SSc patients, autoantibodies directed against different antigens on a common molecular target. In this international cohort of 1574 subjects, only 16 (1%) had monospecific anti-PM75 antibodies and 11 (0.7%) monospecific anti-PM100 antibodies. Monospecific anti-PM75 and anti-PM100 had common (e.g. calcinosis) and distinct (anti-PM75 more

Table 3. Summary of the literature on clinical associations of anti-Pm/Scl antibodies in SSC.

Monospecific Detection method	Anti-PM75			Anti-PM100			Anti-PM/Scl			Anti-PM1-alpha	
	Tri-nation (2014)	Hanke et al. [8]	Villalta et al. [9]	Tri-nation (2014)	Hanke et al. [8]	Villalta et al. [9]	Koschik et al. [7]	Mierau et al. [14]	Marguerie et al. [15]	Oddis et al. [16]	D'Aoust et al. [17]
	Yes LJA	No LJA	No LJA	Yes LJA	No LJA	No LJA	No Double immunodiffusion	No ID/ELISA/LJA	No Counterimmunoelectrophoresis	No Double immunodiffusion	Yes ELISA
Total study population	1574	280	210	1574	280	210	2425	863	879	400	763
Subpopulation positive for antibody [monospecific]	16 (1%)	29 (10.4%)	14 (6.7%)	11 (0.7%)	20 (7.1%)	5 (2.4%)	76 (3.1%)	42 (4.9%)	32 (3.5)	16 (4.0%)	26 (3.4%)
Comparison group (percentage of total)	Negative for anti-PM75 and anti-PM100	Negative for anti-PM75	NA	Negative for anti-PM75 and anti-PM100	Negative for anti-PM100	NA	Negative for anti-PM/Scl	Percentage of total study population	No comparison available	PM/Scl negative group	PM1-Alpha antibody negative
Diagnosis	4	18	NA	1	7	NA	12	2	14	1	6
Diffuse SSC	12	3	NA	10	3	NA	28	16	17	13	20
Limited SSC	0	6	0	0	8	0	36	22	0	10	7
Overlap	45.7 vs 45.7	44.2 vs 50.4	NA	35.8 vs 45.7	48.3 vs 49.8	NA	37.6 vs 42.7	47.6 vs 51.1 (%)	33 (median)	36.9	37.9 vs 40.6
Age at onset of disease	↔	↓	NA	↓	↔	NA	↓	↔	NA	NA	↓
Comparison	7.7 vs 8.9	65.5 vs 44.2	NA	0.0 vs 8.9	35.0 vs 11.5	NA	51 vs 14	47.6 vs 49.4	88	75 vs 5	35 vs 10
Muscle (%)	↔	↑	NA	↓	↑	NA	↑	↔	NA	↑	↑
Comparison	37.5 vs 25.3	NA	NA	36.4 vs 25.3	NA	NA	NA	NA	47	NA	58 vs 30
Calcinosis (%)	↑	NA	NA	↑	NA	NA	NA	NA	NA	NA	↑
Comparison	10.8 vs 11.5	9.0 vs 7.2	NA	4.4 vs 11.5	NA	NA	6.0 vs 15.9	33.3 vs 39.2	97	NA	8.8 vs 10.2
Skin (mRSS)	↔	↔	NA	↓	NA	NA	↓	↓	NA	NA	↓
Comparison	50 vs 35.4	55.2 vs 32.7	NA	20.0 vs 35.4	55.0 vs 33.5	NA	50 vs 37	38.1 vs 33.3	78	37.5 vs 33	15 vs 34
Lung (%)	↑	↑	NA	↓	↑	NA	↑	↔	NA	↔	↓
Comparison	87.5 vs 82.0	24.1 vs 44.6	NA	54.6 vs 82.0	55.0 vs 78.5	NA	52 vs 79	33.3 vs 62.0	78	NA	63 vs 70
Gastrointestinal (%)	↑	↓	NA	↓	↓	NA	↓	↓	NA	NA	↓
Comparison	31.3 vs 28.0	NA	NA	36.4 vs 28.0	NA	NA	84 vs 79	19.0 vs 18.8	97	75 vs 36	58 vs 32
Arthritis (%)	↔	NA	NA	↑	NA	NA	↑	↔	NA	↑	↑
COMPARISON	↔	NA	NA	↑	NA	NA	↑	↔	NA	↑	↑

Tri-nation refers to the current study.

commonly associated with GI and lung disease than anti-PM100 antibodies; anti-PM100 associated with better survival) clinical profiles. Other features previously associated with anti-PM/Scl antibodies such as higher rates of myositis were only observed in subjects with both anti-PM75 and anti-PM100 antibodies. These results suggest that various disease features associated with anti-PM/Scl antibodies may in fact depend on particular and possibly multiple antigen specificities. However, due to the small samples, these results need to be interpreted with caution.

Of particular importance, this dataset highlights the very low prevalence of the monospecific anti-PM75 and anti-PM100 antibodies in SSc. This underscores the importance of international collaborations to assemble sufficiently large cohorts to understand the clinical profiles of relatively uncommon antibodies in SSc.

In addition, little was known to date on the distinct clinical correlates of anti-Pm/Scl antibodies in SSc. We performed an extensive review of the literature (Table 3). All except one of the reported studies [14–16] of anti-PM/Scl antibodies in SSc examined the clinical correlates of those antibodies without excluding subjects with concomitant SSc-specific antibodies, in particular ACA, ATA and ARNAP. Thus, the variable profiles associated with those antibodies may in fact have been confounded by the presence of other antibodies known to have distinct clinical profiles. The only study examining monospecific anti-PM/Scl antibodies was recently reported and was done on a subset of the subjects included in the current study but using a different assay, an ELISA with the synthetic peptide PM1-Alpha, PM100's major epitope [17]. In that study, monospecific anti-PM1-Alpha antibodies were present in 3.4% of the cohort and were associated with younger age at disease onset, more limited skin disease, skeletal muscle involvement, calcinosis, inflammatory arthritis and overlap disease, and less ILD and gastrointestinal symptoms (Table 4). It is noteworthy that this profile is similar to that of anti-PM100 in the current study, with the additional feature that anti-PM1-Alpha antibodies were also associated with muscle disease (whereas anti-PM100 was only associated with muscle disease in the presence of anti-PM75 antibodies). This suggests that testing both anti-PM75 and anti-PM100 antibodies or anti-PM1-Alpha antibody alone is necessary to cover the spectrum of the anti-PM/Scl antibodies. Of note, a novel anti-PM/Scl-100 antibody immunoassay has recently been developed and found to have good performance characteristics [18]. This immunoassay was used to test 223 Japanese SSc sera and found to be positive in only one (0.4%), a patient with lcSSc and pulmonary hypertension. The presence of concomitant SSc-specific antibodies in this patient was not reported, although the patient was said to have a positive ANA (titre 1:320) with nucleolar staining. Although anti-Pm/Scl antibodies are thought to be rare in Japanese patients, that proportion is consistent with the one reported for anti-PM100 in this study.

This study is not without limitations. In particular, variables including inflammatory myositis and calcinosis were not defined using specific criteria. Instead, a study physician reported their presence or absence. However, the fact that all study physicians were experienced rheumatologists supports the validity of these diagnoses. Similarly,

defining ILD in the context of a longitudinal observational cohort studies is very complex, given issues of missing data and verification bias. We defined ILD using a clinical decision rule that was recently published [12]. Data on right heart catheterization was not systematically collected in all patients. Nevertheless, in those without right heart catheterization, we defined pulmonary hypertension using a high cut-off for pulmonary systolic pressure on echocardiogram that has been shown to correlate strongly with right heart catheter studies [13]. Still, we acknowledge that pulmonary hypertension based on echocardiogram is not synonymous with pulmonary arterial hypertension and that some of those with pulmonary hypertension based on echocardiogram may have had other causes of pulmonary hypertension such as left heart disease. Thus, measurement error may have contributed to some of the negative findings of the study. On the other hand, when dealing with relatively uncommon autoantibodies (there were only 16 subjects with monospecific anti-PM75 and 11 subjects with monospecific anti-PM100 antibodies in a cohort of 1574 subjects), large well-phenotyped cohorts are required. In the end, the limitations of our data are counter-balanced by its strengths, which include large sample size and detailed clinical phenotypic data.

This study aimed to identify clinical correlates of monospecific anti-PM75 and anti-PM100 antibodies. These antibodies were uncommon in isolation and this points to the need for international collaborations to fill important knowledge gaps in SSc. Monospecific profiles of each autoantibody included more calcinosis. An increased frequency of myositis was only seen in subjects positive for both anti-PM75 and anti-PM100 antibodies. Lung disease was only associated with anti-PM75 and subjects with anti-PM100 antibodies had better survival compared to other antibody subsets. Disease features associated with anti-PM/Scl antibodies may depend on particular and possibly multiple antigen specificities. However, due to the small numbers of subjects with monospecific antibodies, these results need to be interpreted with caution and replicated in other studies.

Declaration of interest

The authors report no conflicts of interest.

The CSRG received funds and/or gifts in kind from the Canadian Institutes of Health Research (CIHR) (grant #FRN 83518), the Scleroderma Society of Canada and its provincial Chapters, Scleroderma Society of Ontario, Sclérodémie Québec, Cure Scleroderma Foundation, INOVA Diagnostics Inc. (San Diego, CA, USA), Euroimmun (Lubeck, Germany), Fonds de la recherche en santé du Québec (FRSQ), the Canadian Arthritis Network (CAN), the Arthritis Society Research Chair (University of Calgary) and the Lady Davis Institute of Medical Research of the Jewish General Hospital, Montreal, QC. The CSRG has also received educational grants from Pfizer and Actelion pharmaceuticals. Dr. Hudson is funded by the *Fonds de la recherche en Santé du Québec*. Dr. Nikpour holds an NHMRC research fellowship (APP1071735). ASIG receives unrestricted grants from Actelion, Pfizer and GSK, and is also supported by Scleroderma Australia. Dr. Assassi is funded by the NIH/NIAMS K23AR061436. Dr. Mayes is funded by NIH/NIAMS

AR055258. The GENISOS cohort receives funding from DoDW81XWH-13-1-0452. The funding sources had no role in the design of the study, analysis of the data, preparation of the manuscript and decision to submit for publication.

References

1. Mehra, S., J. Walker, K. Patterson, and M. J. Fritzler. 2013. Autoantibodies in systemic sclerosis. *Autoimmun. Rev.* 12: 340–354.
2. Steen, V. D. 2005. Autoantibodies in systemic sclerosis. *Semin. Arthritis Rheum.* 35: 35–42.
3. Castellino, F. V., and J. Varga. 2013. Current status of systemic sclerosis biomarkers: applications for diagnosis, management and drug development. *Exp. Rev. Clin. Immunol.* 9: 1077–1090.
4. Wolfe, F. A. E., and G. Sharp. 1977. Antinuclear antibody with distinct specificity for polymyositis. *J. Clin. Investigat.* 59: 176–178.
5. Iaccarino, L., M. Gatto, S. Bettio, et al. 2013. Overlap connective tissue disease syndromes. *Autoimmun. Rev.* 12: 363–373.
6. Mahler, M., and R. Raijmakers. 2007. Novel aspects of autoantibodies to the PM/Scl complex: clinical, genetic and diagnostic insights. *Autoimmun. Rev.* 6: 432–437.
7. Koschik II R. W., N. Fertig, M. R. Lucas, et al. 2012. Anti-PM-Scl antibody in patients with systemic sclerosis. *Clin. Exp. Rheumatol.* 30(2 Suppl 71): S12–S16.
8. Hanke, K., C. S. Bruckner, C. Dahnrich, et al. 2009. Antibodies against PM/Scl-75 and PM/Scl-100 are independent markers for different subsets of systemic sclerosis patients. *Arthritis Res. Ther.* 11: R22.
9. Villalta, D., T. Imbustaro, S. Di Giovanni, et al. 2012. Diagnostic accuracy and predictive value of extended autoantibody profile in systemic sclerosis. *Autoimmun. Rev.* 12: 114–120.
10. Alhajeri, H., M. Hudson, M. Fritzler, et al. 2015. 2013 American College of Rheumatology/European league against rheumatism classification criteria for systemic sclerosis outperform the 1980 criteria: data from the Canadian Scleroderma Research Group. *ACR (Hoboken).* 67: 582–587.
11. Nikpour, M., P. Hissaria, J. Byron, et al. 2011. Prevalence, correlates and clinical usefulness of antibodies to RNA polymerase III in systemic sclerosis: a cross-sectional analysis of data from an Australian cohort. *Arthritis Res. Ther.* 13: R211.
12. Steele, R., M. Hudson, E. Lo, et al. 2012. Clinical decision rule to predict the presence of interstitial lung disease in systemic sclerosis. *Arthritis Care Res.* 64: 519–524.
13. Hsu, V. M., A. E. Moreyra, A. C. Wilson, et al. 2008. Assessment of pulmonary arterial hypertension in patients with systemic sclerosis: comparison of noninvasive tests with results of right-heart catheterization. *J. Rheumatol.* 35: 458–465.
14. Mierau, R., P. Moinzadeh, G. Riemekasten, et al. 2011. Frequency of disease-associated and other nuclear autoantibodies in patients of the German Network for Systemic Scleroderma: correlation with characteristic clinical features. *Arthritis Res. Ther.* 13: R172.
15. Marguerie, C., C. C. Bunn, J. Copier, et al. 1992. The clinical and immunological features of patients with autoantibodies to the nucleolar antigen PM-Scl. *Medicine* 71: 327–336.
16. Oddis, C. V., Y. Okana, W. A. Rudert, et al. 1992. Serum autoantibody to the nucleolar antigen PM-Scl. *Arthritis Rheum.* 35: 1211–1217.
17. D'Aoust, J., M. Hudson, S. Tatibouet, et al. 2014. Clinical and serologic correlates of anti-PM/Scl antibodies in systemic sclerosis: a multicenter study of 763 patients. *Arthritis Rheumatol.* 66: 1608–1615.
18. Muro, Y., Y. Hosono, K. Sugiura, et al. 2015. Anti-PM/Scl antibodies are found in Japanese patients with various systemic autoimmune conditions besides myositis and scleroderma. *Arthritis Res. Ther.* 17: 57.

Supplementary material available online
Supplementary Table 1

Monospecific anti-Ro52/TRIM21 antibodies in a tri-nation cohort of 1574 systemic sclerosis subjects: evidence of an association with interstitial lung disease and worse survival

M. Wodkowski¹, M. Hudson¹⁻³, S. Proudman^{4,5}, J. Walker⁶, W. Stevens⁷, M. Nikpour^{7,8}, S. Assassi⁹, M.D. Mayes⁹, M. Wang³, M. Baron¹⁻³, M.J. Fritzler¹⁰, and the Canadian Scleroderma Research Group (CSRG), the Australian Scleroderma Cohort Study (ASCS), and the Genetics versus Environment in Scleroderma Outcome Study (GENISOS)

¹Dept. of Medicine, McGill University,

²Division of Rheumatology and ³Lady Davis Institute, Jewish General Hospital, Montréal, Canada; ⁴Rheumatology Unit, Royal Adelaide Hospital, Adelaide,

South Australia; ⁵Discipline of Medicine, University of Adelaide; ⁶Department of Allergy and Immunology, Flinders Medical Centre, South Australia;

⁷Department of Rheumatology, St Vincent's Hospital Melbourne, Victoria, Australia;

⁸Department of Medicine, The University of Melbourne at St. Vincent's Hospital, Melbourne, Victoria, Australia; ⁹Division of Rheumatology and Immunogenetics, University of Texas Health Science Centre at Houston, Houston, Texas, USA;

¹⁰Faculty of Medicine, University of Calgary, Calgary, Alberta, Canada.

Michael Wodkowski, MD

Marie Hudson, MD, MPH

Susanna Proudman, MD, MPH

Jennifer Walker, MD

Wendy Stevens, MD

Mandana Nikpour, MD

Shervin Assassi, MD

Maureen D. Mayes, MD

Mianbo Wang, MSc

Murray Baron, MD

Marvin J. Fritzler, MD, PhD

Please address correspondence and reprint requests to:

Dr Marie Hudson,

Jewish General Hospital, Room A-725,

3755 Côte Sainte-Catherine Road,

Montreal, Quebec H3T 1E2, Canada.

E-mail: marie.hudson@mcgill.ca

Received on March 17, 2015; accepted in revised form on July 6, 2015.

Clin Exp Rheumatol 2015; 33 (Suppl. 91): S131-S135.

© Copyright CLINICAL AND EXPERIMENTAL RHEUMATOLOGY 2015.

Key words: systemic sclerosis, autoantibodies, anti-Ro52/TRIM21 antibodies, interstitial lung disease, cohort study

Funding information and competing interests on page S-135.

ABSTRACT

Objective. Autoantibodies directed against Ro52/TRIM21 are common in systemic sclerosis (SSc) but their clinical significance remains uncertain. The aim of this study was to assess the clinical correlates and survival of subjects with monospecific anti-Ro52/TRIM21 antibodies, i.e. anti-Ro52/TRIM21 antibodies in the absence of other SSc-related antibodies.

Methods. A tri-nation (Canada, Australia, USA) cohort of 1574 SSc subjects was formed, demographic and clinical variables were harmonised and sera were tested using a common diagnostic platform. Statistical analyses were performed to determine associations between the presence of monospecific anti-Ro52/TRIM21 antibodies and outcomes of interest, including interstitial lung disease (ILD) and survival.

Results. 103 (6.5%) had monospecific anti-Ro52/TRIM21 antibodies, 324 (20.6%) had anti-Ro52/TRIM21 antibodies overlapping with other SSc-related antibodies and 1147 (72.9%) were negative for anti-Ro52/TRIM21 antibodies. Monospecific subjects were less likely to be White compared to negative subjects (68% vs. 82%, odds ratio (OR) 0.48, 95% confidence interval (CI) 0.30–0.75, $p=0.0011$). ILD was the only clinical variable significantly associated with monospecific anti-Ro52/TRIM21 antibodies compared to negative subjects (adjusted OR 2.70, 95% CI 1.75–4.14, $p<0.0001$). Subjects with monospecific anti-Ro52/TRIM21 antibodies were at significantly increased risk of death compared to subjects without anti-Ro52/TRIM21 antibodies (log rank $p=0.0003$; adjusted hazard ratio (HR) 1.87, 95% CI 1.24–2.82, $p=0.0029$).

Conclusion. The results obtained from this unique tri-nation cohort represent the strongest evidence to date that anti-Ro52/TRIM21 antibodies are independently associated with the presence of ILD and poor survival in SSc. These data provide strong support for the predictive and prognostic value of this serological biomarker in SSc.

Introduction

Two main types of SS-A/Ro autoantibodies have been described. One is directed at a 60 kDa protein known as SS-A/Ro60, which is a component of a small cytoplasmic ribonucleoprotein (scRNP) macromolecular complexes. Another, that often coexists with SS-A/Ro60 autoantibodies, is directed against a 52 kDa (Ro52) protein that is not normally part of the scRNP complex but is an E3 ubiquitin ligase and member of the tripartite motif (TRIM) family of proteins known as TRIM21 (1, 2); hence, the preferred terminology of Ro52/TRIM21 will be used in this report. Anti-Ro52/TRIM21 antibodies have been reported in a wide variety of autoimmune diseases, often overlapping with other autoantibodies (3-5). Hence, they have often been considered non-specific markers of autoimmune inflammation. The fact that these autoantibodies have also been detected in sera of patients with neoplasia (6), viral infections or even healthy individuals who later developed auto-immune diseases (7) has provided further support for this. In SSc, a recent report on a Spanish cohort of 132 consecutive SSc patients did not find any clinical associations with anti-Ro52/TRIM21 (8). On the other hand, anti-Ro52/TRIM21 antibodies have been reported to be as-

sociated with interstitial lung disease (ILD) in various autoimmune diseases (9, 10), in particular in association with anti-Jo1 antibodies (11), which are well known to be associated with ILD. We previously reported an association between anti-Ro52/TRIM21 antibodies and ILD in SSc where, if present, ILD was 1.5 times more likely (12). However, the relationship between anti-Ro52/TRIM21 and ILD in SSc (and other autoimmune diseases) may have been confounded by the presence of concomitant antibodies known to be associated with ILD, particularly anti-topoisomerase I in SSc (13). We have previously shown that overlap with SSc-specific autoantibodies can confound the associations with other autoantibodies (14). The aim of this study was therefore to assess the clinical correlates of *monospecific* anti-Ro52/TRIM21 antibodies, *i.e.* anti-Ro52/TRIM21 antibodies in the absence of other SSc-related antibodies.

Materials and methods

The Tri-Nation cohort comprises SSc subjects included in the Canadian Scleroderma Research Group (CSRG), the Australian Scleroderma Cohort Study (ASCS) and the American Genetics *versus* Environment in Scleroderma Outcome Study (GENISOS) cohorts. Ethics committee approval for this study was obtained at McGill University (Montreal, Canada) and at all participating CSRG, ASCS, and GENISOS study sites. All subjects provided informed written consent to participate in the study. Selection of study subjects in and harmonisation of clinical variables between the 3 study cohorts have been described (15). Briefly, over 98% of the CSRG (16) and ASCS subjects, and all GENISOS subjects meet the 2013 ACR/EULAR classification criteria for SSc (17). Demographic information regarding age, sex and ethnicity was collected by subject self-report. Disease duration was recorded by study physicians and defined as the interval between the onset of the first non-Raynaud disease manifestation and baseline study visit. Skin involvement was assessed using the modified Rodnan skin score. Limited cutaneous disease (lcSSc) was defined

as skin involvement distal to the elbows and knees with or without facial involvement; diffuse cutaneous disease (dcSSc) was defined as skin involvement proximal to the elbows and knees and/or of the trunk. A history of inflammatory myositis, calcinosis, inflammatory arthritis and scleroderma renal crisis was recorded by a study physician. To assess gastrointestinal involvement, subjects answered yes/no to 6 questions concerning gastroesophageal reflux disease, dysphagia, antibiotics for bacterial overgrowth, episodes of pseudo-obstruction, fecal incontinence and hyperalimentation. The presence of interstitial lung disease (ILD) was determined using a clinical decision rule that was recently published (18). Using this algorithm, ILD was considered present if a high resolution computed tomography (HRCT) scan of the lung was interpreted by an experienced radiologist as showing ILD or, in the case where no HRCT was available, if either a chest x-ray was reported as showing either increased interstitial markings (not thought to be due to congestive heart failure) or fibrosis, and/or if a study physician reported the presence of typical “velcro-like crackles” on physical examination. Pulmonary hypertension was defined as an estimated systolic pulmonary artery pressure (sPAP) ≥ 45 mmHg measured using the Doppler flow measurement of the tricuspid regurgitant jet on cardiac echocardiography (an estimate that correlates strongly with right heart catheter studies) (19) for CSRG and GENISOS subjects, or mean pulmonary artery pressure (mPAP) > 25 mmHg with a pulmonary capillary wedge pressure (PCWP) < 15 mmHg on right heart catheterisation for ASCS subjects.

Serology

Autoantibody analyses of the CSRG and GENISOS cohorts were performed in a central laboratory, Mitogen Advanced Diagnostics Laboratory, University of Calgary and the ASCS analyses were performed in Australia using an identical immunoassay kit and protocols. Serum aliquots were stored at -80°C until needed for diagnostic assays. Antibodies against Ro52/TRIM21, centromere (CENP A and CENP B), topoi-

somerase I, RNA polymerase III (RP11 and RP155), fibrillarin, Nor90, Th/To, Ku, PDGFR, PM75 and PM100 were detected and digitally quantified by the Euroline systemic sclerosis profile line immunoassay (LIA) (Euroimmun, Luebeck, Germany) according to the manufacturer's instructions.

Statistical analysis

Subjects were divided into those exclusively positive for anti-Ro52/TRIM21 antibodies (*i.e.* monospecific anti-Ro52/TRIM21 antibodies subjects), those with anti-Ro52/TRIM21 antibodies overlapping with other measured antibodies (*i.e.* overlapping anti-Ro52/TRIM21 antibodies subjects), and those altogether negative for anti-Ro52/TRIM21 antibodies. Descriptive statistics were used to compare 20 selected variables between 1) monospecific anti-Ro52/TRIM21 antibody positive *versus* negative subjects and 2) overlapping anti-Ro52/TRIM21 antibodies *versus* negative subjects. Adjusting for multiple comparisons, $p < 0.00125$ was considered statistically significant. Multivariate logistic regression adjusting for baseline differences in age and ethnicity was used to determine the association between anti-Ro52/TRIM21 antibody groups and ILD. Kaplan Meier analysis and Cox proportional hazard models adjusting for baseline differences in age and ethnicity were used to compare survival between autoantibody subsets. p -values < 0.05 were considered statistically significant for these 3 latter analyses. All statistical analyses were performed with SAS v.9.2 (SAS Institute, USA).

Results

A total of 1574 SSc subjects were included in this study, of whom 103 (6.5%) had monospecific anti-Ro52/TRIM21 antibodies, 324 (20.6%) had anti-Ro52/TRIM21 antibodies overlapping with other SSc-related antibodies and 1147 (72.9%) were negative for anti-Ro52/TRIM21 antibodies (Table I). Monospecific subjects were less likely to be White compared to negative subjects (68% *vs.* 82%, odds ratio (OR) 0.48, 95% confidence interval (CI) 0.30–0.75, $p = 0.0011$). Subjects with overlapping anti-Ro52/TRIM21 anti-

Table I. Baseline characteristics of the study cohort, as a group and according to anti-Ro52/TRIM21 antibody status. The monospecific anti-Ro52/TRIM21 antibody positive group was exclusive of anti-CENP, topoisomerase I, RNA polymerase III, fibrillarin, NOR90, Th/To, Ku, PDGFR, PM75 and PM100 antibodies. Adjusting for multiple comparisons, $p < 0.00125$ was considered statistically significant.

	Whole group (n=1574)		Monospecific anti-Ro52/ TRIM21 positive (n=103)		Overlapping anti-Ro52/ TRIM21 positive (n=324)		Anti-Ro52/ TRIM21 negative (n=1147)		Monospecific vs. Negative			Overlapping vs. Negative		
	% or mean	n. or SD	% or mean	n. or SD	% or mean	n. or SD	% or mean	n. or SD	OR	CI	p-value	OR	CI	p-value
<i>Sociodemographics</i>														
Female	86%	1355	82%	84	90%	290	86%	981			0.2776			0.0664
White	82%	1244	68%	66	86%	273	82%	905	0.48	0.30, 0.75	0.0011			0.0665
Age, years	55.1	12.8	53.2	13.1	58.7	12.6	54.2	12.6			0.4180	1.03	1.02, 1.04	<0.0001
Disease duration, years	9.5	9.2	7.9	7.7	10.3	9.5	9.5	9.2			0.1069			0.1358
Age at disease onset, years	45.5	13.7	45.2	13.2	48.4	13.9	44.7	13.6			0.7070	1.02	1.01, 1.03	<0.0001
<i>Clinical variables</i>														
Modified Rodnan skin score (0-51)	11.4	10.2	11.3	9.4	10.9	9.8	11.6	10.5			0.7969			0.3333
Limited cutaneous disease	61%	963	52%	53	68%	221	60%	689			0.1021			0.0094
Inflammatory myositis	9%	129	11%	10	8%	24	9%	95			0.5727			0.6272
Calcinosis	26%	405	18%	18	30%	95	26%	292			0.0736			0.1396
Inflammatory arthritis	28%	443	27%	28	30%	97	28%	318			0.9364			0.4554
<i>Gastrointestinal disease</i>														
GERD/reflux	82%	1278	75%	76	87%	279	81%	923			0.1960			0.0135
Dysphagia	52%	808	51%	50	56%	177	51%	581			0.8892			0.1196
Antibiotics for bacterial overgrowth	7%	94	6%	5	8%	24	7%	65			0.8393			0.4145
Episodes of pseudo-obstruction	3%	42	1%	1	4%	12	3%	29			0.3473			0.2688
Fecal incontinence	18%	217	9%	6	21%	56	18%	155			0.0766			0.2726
Hyperalimentation	3%	23	9%	5	2%	4	2%	14			0.0079			0.9323
Number of GI symptoms (0-6)	1.6	1	1.4	0.9	1.7	1	1.5	1			0.2130			0.0092
Scleroderma renal crisis	4%	60	6%	6	3%	11	4%	4			0.3060			0.7489
Pulmonary hypertension	14%	172	8%	6	20%	56	12%	110			0.2459	1.81	1.27, 2.58	0.0011
Interstitial lung disease	36%	548	57%	57	36%	116	34%	375	2.63	1.74, 3.98	<0.0001			0.3303

CENP: centromere proteins; CI: confidence interval; GERD: gastro-esophageal reflux disease; GI: gastrointestinal; NOR: nucleolar organiser; PDGFR: platelet derived growth factor; OR: odds ratio; SD: standard deviation; TRIM: tripartite motif.

bodies were significantly older than the negative subjects (58.7 years vs. 54.2, OR 1.03, 95% CI 1.02–1.04, $p < 0.0001$). In univariate analysis, ILD was the only clinical variable significantly associated with monospecific anti-Ro52/TRIM21 antibodies compared to negative subjects (OR 2.63, 95% CI 1.74–3.98, $p < 0.0001$; Table I). In logistic regression analysis adjusting for differences in baseline demographic characteristics, subjects with monospecific anti-Ro52/TRIM21 antibodies were almost 3 times more likely to have ILD compared to those without those antibodies (OR 2.70, 95% CI 1.75–4.14, $p < 0.0001$; Table II). Of note, subjects with overlapping anti-Ro52/TRIM21 antibodies did not have a higher frequency of ILD either in univariate or multivariate analysis.

In unadjusted survival analysis (Fig. 1),

subjects with monospecific anti-Ro52/TRIM21 antibodies were at increased risk of death compared to negative subjects (log rank $p = 0.0003$). Again, after adjusting for differences in baseline demographic characteristics, subjects with monospecific anti-Ro52/TRIM21 antibodies were still at significantly increased risk of death compared to subjects without anti-Ro52/TRIM21 antibodies (hazard ratio (HR) 1.87, 95% CI 1.24–2.82, $p = 0.0029$; Table III).

Discussion

Although anti-Ro52/TRIM21 is the second most common autoantibody in SSc sera (12, 20), the prevalence of monospecific anti-Ro52/TRIM21 antibodies in this large SSc cohort was less than 10%. Nonetheless, leveraging this large unique tri-nation dataset using a common serological platform, we

found strong evidence that monospecific anti-Ro52/TRIM21 antibodies are independently associated with ILD and increased mortality in SSc. Currently, there are few robust clinical biomarkers in SSc-ILD aside from C-reactive protein, which has been shown to be associated with worse pulmonary function (21) and anti-topoisomerase I (Scl-70) with ILD (22) and worsening forced vital capacity (23). Our data provide evidence for a novel predictive and prognostic biomarker in SSc. Of note, though, subjects with overlapping anti-Ro52/TRIM21 antibodies did not have a higher frequency of ILD. It is possible that the presence of other SSc-related antibodies modifies the association between anti-Ro52/TRIM21 and ILD. The role of anti-Ro52/TRIM21 in the pathophysiology of autoimmune diseases remains largely unknown.

Table II. Logistic regression model to estimate the association between the presence of anti-Ro52/TRIM21 antibodies and ILD, adjusting for baseline demographic differences.

	β	Odds ratio	95% CI		<i>p</i> -value
White	-0.23	0.80	0.60	1.06	0.1144
Age	0.01	1.01	1.01	1.02	0.0016
Monospecific vs. negative anti-Ro52/TRIM21 subjects	0.99	2.70	1.75	4.14	<.0001
Overlapping vs. negative anti-Ro52/TRIM21 subjects	0.05	1.05	0.80	1.37	0.7165

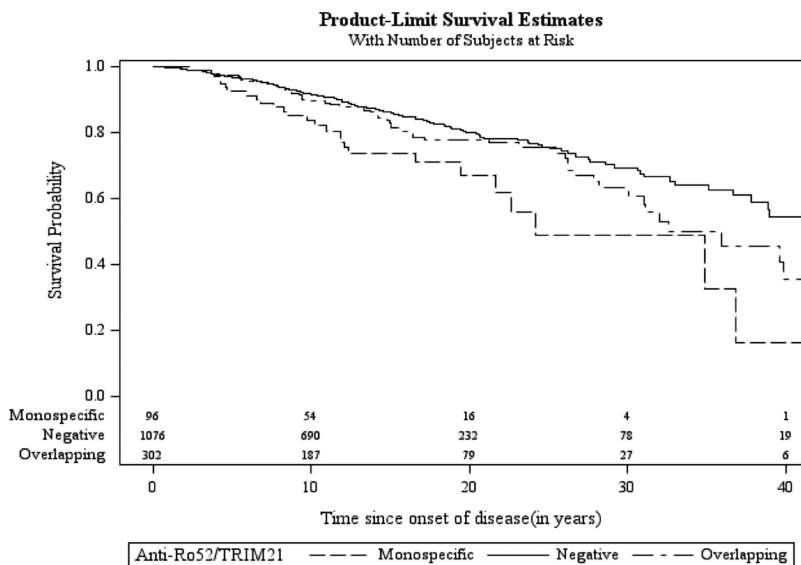


Fig. 1. Kaplan Meier curve to compare survival in the anti-Ro52/TRIM21 monospecific, overlapping and negative subjects. Log rank *p*-values: monospecific vs. negative subjects *p*=0.0003; overlapping vs. negative *p*=0.1106; monospecific vs. overlapping subjects *p*=0.0210.

Table III. Cox proportional hazard model to estimate the association between the presence of anti-Ro52/TRIM21 antibodies and mortality, adjusting for baseline demographic differences.

	β	Hazard ratio	95% CI		<i>p</i> -value
White	-0.51	0.60	0.49	0.81	0.0006
Age	0.03	1.03	1.02	1.04	<.0001
Monospecific vs. negative anti-Ro52/TRIM21 subjects	0.63	1.87	1.24	2.82	0.0029
Overlapping vs. negative anti-Ro52/TRIM21 subjects	0.28	1.33	0.99	1.79	0.0598

Nevertheless, some reports suggest a pathogenic role. In general, the autoantibody binding target, Ro52/TRIM21, is a regulator of type I interferon (IFN) and proinflammatory cytokine production (2). In turn, IFN α upregulates Ro52/TRIM21 and promotes its nuclear translocation (24). This self-perpetuating process has the potential to contribute to the inflammatory cascade. In tissue, Ro52/TRIM21 expression is increased in cutaneous lupus erythematosus and ultraviolet light-induced skin lesions and translocation to apoptotic blebs has been hypothesised as a mechanism for its immunogenicity (25). Evidence also exists to support a similar mechanism

occurring during cardiomyocyte apoptosis, as well as direct cross-reactivity with cardiac membrane proteins involved in the control of electric signal generation and/or conduction, as in congenital heart block (26, 27). Finally, in primary Sjögren syndrome, a single nucleotide polymorphism in the Ro52 gene has been shown to be associated with anti-Ro52/TRIM21 autoantibodies (28). Although anti-Ro52/TRIM21 autoantibodies have been shown to be associated with severe disease refractory to steroids in auto-immune hepatitis, the pathogenic mechanisms, if any, are not known (29). Similarly, the pathogenic role of anti-Ro52/TRIM21 and other au-

toantibodies associated with interstitial lung disease is not known (30).

This study is not without limitations. In particular, defining ILD in the context of longitudinal observational cohort studies is very complex, given issues of missing data and verification bias. We defined ILD using a clinical decision rule that was recently published (18). Still, measurement error may have contributed to some of the negative findings of the study. On the other hand, when dealing with relatively uncommon serological profiles (there were only 6.5% of subjects with monospecific anti-Ro52/TRIM21 antibodies), large well-phenotyped cohorts are required to obtain robust estimates. Thus, the limitations of our data are counter-balanced by its strengths, which include large sample size and detailed clinical phenotypic data. Finally, subjects identified as having “monospecific” anti-Ro52/TRIM21 antibodies may in fact have had other autoantibodies that are undetected by the immunoassays employed in this study. This might include some associated with connective tissue disease-related ILD such as anti-Jo1, which were not included among those tested for this study. However, we have previously reported a very low prevalence of anti-Jo1 antibodies in the CSRG SSc cohort (approximately 1%) (12). Thus, the presence of these autoantibodies is unlikely to have influenced the results of this study in a meaningful manner.

We found that monospecific anti-Ro52/TRIM21 antibodies were strongly associated with ILD and an independent predictor of mortality in this large SSc cohort. This provides the strongest evidence to date for the predictive and prognostic value of this serological biomarker in SSc and contributes important clinically meaningful data.

Investigators of the Canadian Scleroderma Research Group

- J. Pope, London, Ontario
- M. Baron, Montreal, Quebec
- J. Markland, Saskatoon, Saskatchewan
- D. Robinson, Winnipeg, Manitoba
- N. Jones, Edmonton, Alberta
- N. Khalidi, Hamilton, Ontario
- P. Docherty, Moncton, New Brunswick
- E. Kaminska, Calgary, Alberta
- A. Masetto, Sherbrooke, Quebec
- E. Sutton, Halifax, Nova Scotia
- J-P. Mathieu, Montreal, Quebec

M. Hudson, Montreal, Quebec
 S. Ligier, Montreal, Quebec
 T. Grodzicky, Montreal, Quebec
 S. LeClercq, Calgary, Alberta
 C. Thorne, Newmarket, Ontario
 G. Gyger, Montreal, Quebec
 D. Smith, Ottawa, Ontario
 P.R. Fortin, Quebec, Quebec
 M. Larché, Hamilton, Ontario
 T.S. Rodriguez-Reyna, Mexico City
 A.R. Cabral, Mexico City, Mexico
 M.J. Fritzler, Cumming School of
 Medicine, University of Calgary, Alberta

Investigators of the Australian Scleroderma Cohort Study

C. Hill, Adelaide, South Australia
 S. Lester, Adelaide, South Australia
 P. Nash, Sunshine Coast, Queensland
 M. Nikpour, Melbourne, Victoria
 J. Roddy, Perth, Western Australia
 K. Patterson, Adelaide, South Australia
 S. Proudman, Adelaide, South Australia
 M. Rischmueller, Adelaide, S. Australia
 J. Sahhar, Melbourne, Victoria
 W. Stevens, Melbourne, Victoria
 J. Walker, Adelaide, South Australia
 J. Zochling, Hobart, Tasmania

Investigators of GENISOS

S. Assassi, Houston, Texas
 M.D. Mayes, Houston, Texas
 T.A. McNearney, Galveston, Texas
 G. Salazar, Houston, Texas

Funding

The CSRG received funds and/or gifts in kind from the Canadian Institutes of Health Research (CIHR) (grant n.FRN 83518), the Scleroderma Society of Canada and its provincial Chapters, Scleroderma Society of Ontario, Sclérodemie Québec, Cure Scleroderma Foundation, Scleroderma Society of Saskatchewan, INOVA Diagnostics Inc. (San Diego, CA), Euroimmun (Lübeck, Germany), Fonds de la recherche en santé du Québec (FRSQ), the Canadian Arthritis Network (CAN), the Arthritis Society Research Chair (University of Calgary) and the Lady Davis Institute for Medical Research of the Jewish General Hospital, Montreal, QC. The CSRG has also received educational grants from Pfizer and Actelion pharmaceuticals. ASCS receives unrestricted grants from Actelion, Pfizer and GSK, and is also supported by Scleroderma Australia. The GENISOS cohort receives funding from DoDW81XWH-13-1-0452.

M. Hudson is funded by the Fonds de la recherche en Santé du Québec.

N. Nikpour holds an NHMRC research fellowship (APP1071735).

S. Assassi is funded by the NIH/NIAMS K23AR061436.

M.D. Mayes is funded by NIH/NIAMS AR055258.

M. Fritzler is a paid consultant, has received honoraria or has received gifts in kind from Inova Diagnostics (San Diego, CA) and Euroimmun GmgH (Lübeck, Germany).

The funding sources had no role in the design of the study, analysis of the data, preparation of the manuscript and decision to submit for publication.

References

1. RACANELLI V, PRETE M, MUSARAJ G, DAMMACCO F, PEROSA F: Autoantibodies to intracellular antigens: generation and pathogenic role. *Autoimmun Rev* 2011; 10: 503-8.
2. OKE V, WAHREN-HERLENIUS M: The immunobiology of Ro52 (TRIM21) in autoimmunity: a critical review. *J Autoimmun* 2012; 39: 77-82.
3. SCHULTE-PELKUM J, FRITZLER M, MAHLER M: Latest update on the Ro/SS-A autoantibody system. *Autoimmun Rev* 2009; 8: 632-7.
4. GRANITO A, MURATORI P, MURATORI L *et al.*: Antibodies to SS-A/Ro-52kD and centromere in autoimmune liver disease: a clue to diagnosis and prognosis of primary biliary cirrhosis. *Aliment Pharmacol Ther* 2007; 26: 831-8.
5. RUTJES SA, VETTM: Anti-Ro52 antibodies frequently co-occur with anti-Jo-1 antibodies in sera from patients with idiopathic inflammatory myopathy. *Clin Exp Immunol* 1997; 109: 8.
6. MENENDEZ A, GÓMEZ J, ESCANLAR E, CAMINAL-MONTERO L, MOZO L: Clinical associations of anti-SSA/Ro60 and anti-Ro52/TRIM21 antibodies: Diagnostic utility of their separate detection. *Autoimmunity* 2013; 46: 32-9.
7. HEINLEN LD, MCCLAIN MT, RITTERHOUSE LL *et al.*: 60 kD Ro and nRNP A frequently initiate human lupus autoimmunity. *PLoS one* 2010; 5: e9599.
8. SANCHEZ-MONTALVA A, FERNANDEZ-LUQUE A, SIMEON CP *et al.*: Anti-SSA/Ro52 autoantibodies in scleroderma: results of an observational, cross-sectional study. *Clin Exp Rheumatol* 2014; 32: S-177-82.
9. GHILLANI P, ANDRE C, TOLY C *et al.*: Clinical significance of anti-Ro52 (TRIM21) antibodies non-associated with anti-SSA 60kDa antibodies: results of a multicentric study. *Autoimmun Rev* 2011; 10: 509-13.
10. FERREIRA JP, ALMEIDA I, MARINHO A, CERVEIRA C, VASCONCELOS C: Anti-ro52 antibodies and interstitial lung disease in connective tissue diseases excluding scleroderma. *ISRN Rheumatol* 2012; 2012: 415272.
11. MARIE I, HATRON PY, DOMINIQUE S *et al.*: Short-Term and Long-Term Outcome of Anti-Jo1-Positive Patients with Anti-Ro52 Antibody. *Semin Arthritis Rheum* 2012; 41: 890-9.
12. HUDSON M, POPE J, MAHLER M *et al.*: Clinical significance of antibodies to Ro52/TRIM21 in systemic sclerosis. *Arthritis Res Ther* 2012; 14: R50.
13. ASSASSI S, SHARIF R, LASKY RE *et al.*: Predictors of interstitial lung disease in early systemic sclerosis: a prospective longitudinal study of the GENISOS cohort. *Arthritis Res Ther* 2010; 12: R166.
14. D'AOUST J, HUDSON M, MAHLER M, BARON M, FRITZLER MJ: Additional reasons to measure anti-PM1-Alpha antibodies in systemic sclerosis. *Arthritis Rheum* (Hoboken, NJ) 2014 (in press).
15. WODKOWSKI M, HUDSON M, PROUDMAN S *et al.*: Clinical correlates of monospecific anti-PM75 and anti-PM100 antibodies in an international cohort of 1574 systemic sclerosis subjects. 2014 (under review).
16. ALHAJERI H, HUDSON M, FRITZLER M *et al.*: The 2013 ACR/EULAR Classification Criteria for Systemic Sclerosis Out-perform the 1980 Criteria. Data from the Canadian Scleroderma Research Group. *Arthritis Care Res* 2015; 67: 582-7.
17. VAN DEN HOOGEN F, KHANNA D, FRANSEN J *et al.*: Classification Criteria for Systemic Sclerosis: An ACR-EULAR Collaborative Initiative. *Arthritis Rheum* 2013; 65: 2737-47.
18. STEELE R, HUDSON M, LO E, BARON M, CANADIAN SCLERODERMA RESEARCH G: Clinical decision rule to predict the presence of interstitial lung disease in systemic sclerosis. *Arthritis Care Res* 2012; 64: 519-24.
19. HSU VM, MOREYRA AE, WILSON AC *et al.*: Assessment of pulmonary arterial hypertension in patients with systemic sclerosis: comparison of noninvasive tests with results of right-heart catheterization. *J Rheumatol [Comparative Study]* 2008; 35: 458-65.
20. MEHRA S, WALKER J, PATTERSON K, FRITZLER MJ: Autoantibodies in systemic sclerosis. *Autoimmun Rev* 2013; 12: 340-54.
21. MUANGCHAN C, HARDING S, KHIMDAS S, BONNER A, BARON M, POPE J: Association of C-reactive protein with high disease activity in systemic sclerosis: results from the Canadian Scleroderma Research Group. *Arthritis Care Res* 2012; 64: 1405-14.
22. ZHANG XJ, BONNER A, HUDSON M, BARON M, POPE J: Association of gastroesophageal factors and worsening of forced vital capacity in systemic sclerosis. *J Rheumatol* 2013; 40: 850-8.
23. ASSASSI S, SHARIF R, LASKY RE *et al.*: Predictors of interstitial lung disease in early systemic sclerosis: a prospective longitudinal study of the GENISOS cohort. *Arthritis Res Ther* 2010; 12: R166.
24. STRANDBERG L, AMBROSI A, ESPINOSA A *et al.*: Interferon-alpha induces up-regulation and nuclear translocation of the Ro52 autoantigen as detected by a panel of novel Ro52-specific monoclonal antibodies. *J Clin Immunol* 2008; 28: 220-31.
25. OKE V, VASSILAKI I, ESPINOSA A *et al.*: High Ro52 expression in spontaneous and UV-induced cutaneous inflammation. *J Invest Dermatol* 2009; 129: 2000-10.
26. CLANCY RM, BUYON JP, IKEDA K *et al.*: Maternal antibody responses to the 52-kd SSA/RO p200 peptide and the development of fetal conduction defects. *Arthritis Rheum* 2005; 52: 3079-86.
27. AMBROSI A, WAHREN-HERLENIUS M: Congenital heart block: evidence for a pathogenic role of maternal autoantibodies. *Arthritis Res Ther* 2012; 14: 208.
28. NAKKEN B, JONSSON R, BOLSTAD AI: Polymorphisms of the Ro52 gene associated with anti-Ro 52-kd autoantibodies in patients with primary Sjögren's syndrome. *Arthritis Rheum* 2001; 44: 638-46.
29. MONTANO-LOZA AJ, SHUMS Z, NORMAN GL, CZAJA AJ: Prognostic implications of antibodies to Ro/SSA and soluble liver antigen in type I autoimmune hepatitis. *Liver Int* 2012; 32: 85-92.
30. WELLS AU, DENTON CP: Interstitial lung disease in connective tissue disease-mechanisms and management. *Nat Rev Rheum* 2014; 10: 728-39.