RMD Open
Rheumatic & Musculoskeletal Diseases

SHORT REPORT

Assessing the *MUC5B* promoter variant in a large cohort of systemic sclerosis-associated interstitial lung disease

Carlos Rosa-Baez ,[1] Carlos Rangel-Pelaez ,[1] Inmaculada Rodriguez-Martin ,[1] Martin Kerick ,[1] Alfredo Guillen-Del-Castillo ,[2] Carmen P Simeon-Aznar,[2] José Luis Callejas,[3] Alexandre E Voskuyl,[4] Alexander Kreuter,[5] Oliver Distler ,[6] Susanna M Proudman,[7] Mandana Nikpour,[8] Nicolas Hunzelmann,[9] Jeska K De Vries-Bouwstra ,[10] Ariane L Herrick ,[11,12] Yannick Allanore,[13] Lorenzo Beretta ,[14] Maureen D Mayes,[15] Christopher P Denton,[16] Shervin Assassi,[15] Javier Martin,[1] Marialbert Acosta-Herrera ,[1] International SSc Group, Australian Scleroderma Interest Group (ASIG)

To cite: Rosa-Baez C, Rangel-Pelaez C, Rodriguez-Martin I, *et al*. Assessing the *MUC5B* promoter variant in a large cohort of systemic sclerosis-associated interstitial lung disease. *RMD Open* 2025;**11**:e005754. doi:10.1136/rmdopen-2025-005754

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/rmdopen-2025-005754).

Received 3 April 2025
Accepted 26 July 2025

For numbered affiliations see end of article.

Correspondence to
Dr Marialbert Acosta-Herrera; m.acostaherrera@ipb.csic.es

ABSTRACT

Objective The common gain-of-function variant rs35705950, located in the promoter of *MUC5B* gene, has been strongly associated with interstitial lung diseases (ILDs) of different aetiology, such as idiopathic pulmonary fibrosis (IPF) and rheumatoid arthritis-associated ILD (RA-ILD). In this study, we aimed to investigate the association of this variant and its nearby single nucleotide polymorphisms (SNPs) in the largest cohort of systemic sclerosis-associated ILD (SSc-ILD) to date.

Methods Samples were collected from blood/saliva, followed by DNA extraction and genotyping using SNP arrays. Data for rs35705950 and additional 903 variants within 100 Kb were obtained using genomic imputation. Subsequently, we tested their association in a meta-analysis to increase the consistency of the results, including 10 European ancestry cohorts comprising 2363 patients with SSc-ILD, 3526 SSc patients without ILD and 15 076 controls.

Results Meta-analysis showed no significant association between rs35705950 and SSc-ILD, either comparing patients with SSc with and without ILD (p value: 0.588, OR: 1.05, 95% CI: 0.87 to 1.27) nor patients with SSc-ILD with controls (p value: 0.061, OR: 1.16, 95% CI: 0.99 to 1.36). Moreover, none of the additional 903 variants tested in the genomic region reached statistical significance.

Conclusion Despite analysing the largest and most statistically powered SSc-ILD cohort to date, we found no evidence of association between the *MUC5B* promoter variant rs35705950 and its surrounding SNPs with SSc-ILD. These results suggest that the pathogenic mechanisms underlying SSc-ILD may only partially overlap with those of other similar ILDs, such as IPF or RA-ILD. This highlights the need for further studies regarding their genetic architecture.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ *MUC5B* promoter variant rs35705950 is a well-established genetic risk factor for idiopathic pulmonary fibrosis (IPF) and rheumatoid arthritis-associated interstitial lung disease (RA-ILD).

⇒ Previous studies in systemic sclerosis-associated ILD (SSc-ILD) have shown no association with this functional variant, yet they might be underpowered due to their limited sample size.

WHAT THIS STUDY ADDS

⇒ This study confirms that the *MUC5B* promoter variant rs35705950 and its surrounding single nucleotide polymorphisms are not associated with SSc-ILD, in a well-powered cohort.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The results of this study highlight the differences in the genetic architecture of SSc-ILD and similar ILDs as IPF and RA-ILD.

⇒ More research is needed regarding the genetics of SSc-ILD and other ILDs, to determine the extent of their genetic overlap.

INTRODUCTION

Systemic sclerosis-associated interstitial lung disease (SSc-ILD) is currently the leading cause of mortality in patients with SSc and is characterised by significant clinical heterogeneity and unpredictable disease progression.[1]

Despite this variability, SSc-ILD shares key clinical features with other ILDs, such as idiopathic pulmonary fibrosis (IPF) and rheumatoid arthritis-associated ILD (RA-ILD), including progressive pulmonary fibrosis, declining lung function, worsening respiratory symptoms and higher mortality.[2] In addition to clinical overlap, these ILDs also exhibit shared genetic architecture. For instance, a variant in the *FAM13A* gene, previously associated with IPF, has recently been reported in patients with SSc-ILD presenting a usual interstitial pneumonia (UIP) pattern.[3] Furthermore, both IPF and RA-ILD have been associated with rare deleterious variants in telomere-related genes such as *TERT* and *RTEL1*.[4]

In this regard, the common gain-of-function variant rs35705950, located in the promoter of the *MUC5B* gene, is the strongest genetic risk factor for IPF and is also strongly associated with RA-ILD.[5 6] *MUC5B* encodes a mucin protein that plays a crucial role in mucus secretion across multiple organs, including the lungs.[7] However, its exact pathogenic mechanism is not well described. Furthermore, despite the similarities among IPF, RA-ILD and SSc-ILD, no significant association has been reported for this gain-of-function variant and SSc-ILD to date.[3 8–11] Nevertheless, previous studies have been limited by small sample sizes (ranging from 109 to 662 patients with SSc-ILD), or the use of suboptimal control groups, such as the general population rather than patients with SSc without ILD. This may have hindered the detection of a potential SSc-ILD specific association.

Given these considerations, we hypothesised that the *MUC5B* promoter variant may contribute to SSc-ILD risk in the largest well-powered cohort of patients analysed to date. To test this, we assessed the association of rs35705950 and its surrounding variants in a cohort comprising 2363 patients with SSc-ILD, 3526 patients with SSc without ILD and 15 076 controls.

MATERIALS AND METHODS

Study population

The 10 cohorts included in this study comprised a total of 2363 patients with SSc-ILD, 3526 patients with SSc without ILD and 15 076 controls. Blood/saliva samples were collected from all individuals, and, after DNA extraction, genome-wide genotyping was carried out using the arrays specified in online supplemental table S1. All individuals had European descent and were recruited from Spain, the UK, the USA, Australia, Italy, Germany, Switzerland, France and the Netherlands. All patients fulfilled the 2013 American College of Rheumatology/the European League Against Rheumatism classification criteria, or the criteria proposed by LeRoy and Medsger for early SSc.[12 13] Clinical characteristics and further information on the 10 cohorts analysed are summarised in online supplemental table S1.

Presence or absence of ILD in patients with SSc was assessed using high-resolution CT (HRCT), or by the presence of radiological abnormalities on chest x-ray or abnormalities on pulmonary function tests (PFTs).

Quality control and imputation of MUC5B relevant region

Standard and stringent single nucleotide polymorphism (SNP) array quality control (QC) measures were applied to avoid potential source of bias. Samples with genotype missingness >0.05 or ambiguous sex annotation were excluded. Furthermore, to eliminate related or duplicated individuals, we estimated identity-by-descent and excluded one individual from each pair of relatives (Pi_Hat>0.4) or duplicates (Pi_Hat>0.99). Additionally, variants with low call rate (< 0.98), deviations from Hardy-Weinberg equilibrium (HWE) (p value<1×10^{-3}) or a minor allele frequency (MAF) <0.01 were filtered out. Palindromic SNPs with an A/T or C/G allele frequency exceeding 40% were also removed. After QC, we conducted genome-wide imputation using the TOPMed Imputation Server (https://imputation.biodatacatalyst.nhlbi.nih.gov/) and focused on the surrounding region of the *MUC5B* gain-of-function promoter variant rs35705950 (100 Kb upstream and 100 Kb downstream), comprising a total of 904 variants. We only considered SNPs with a good quality imputation (squared correlation (Rsq) >0.3) for further analyses and excluded SNPs with MAF <0.01. Finally, to account for population structure, principal component analysis was conducted using PLINK and GCTA64[14 15] for each cohort individually. 10 principal components were calculated for each cohort, and individuals deviating by more than four SD from the cluster centroid of their respective cohort were identified as outliers and removed.

Statistical analysis

Logistic regression analyses for allele dosages were conducted in PLINK for each cohort, adjusting for sex and the first five principal components in the comparison of SSc-ILD with patients with SSc without ILD, and for patients with SSc-ILD with controls. Then inverse variance weighted meta-analysis of the datasets was carried out using Metasoft[16] under a fixed-effect model in variants without evidence of heterogeneity (Cochran's Q p value>0.05), while the random effect model (RE2) was used for SNPs showing heterogeneity (Cochran's Q p value<0.05). Statistical significance threshold was declared in all comparisons at 5.53×10^{-5} for the 904 SNPs within the region analysed, calculated using Bonferroni correction for multiple testing (0.05/904). Power calculations were conducted with Quanto (https://keck.usc.edu/biostatistics/software/) for MAF=0.05–0.45 and allele effects (OR)=1–2 (online supplemental figure 1 and online supplemental figure 2).

RESULTS

The meta-analyses revealed no significant association between the *MUC5B* promoter SNP rs35705950 and SSc-ILD. Specifically, for the comparison of SSc-ILD with SSc without ILD, the p value was 0.588 (OR: 1.05, 95% CI:

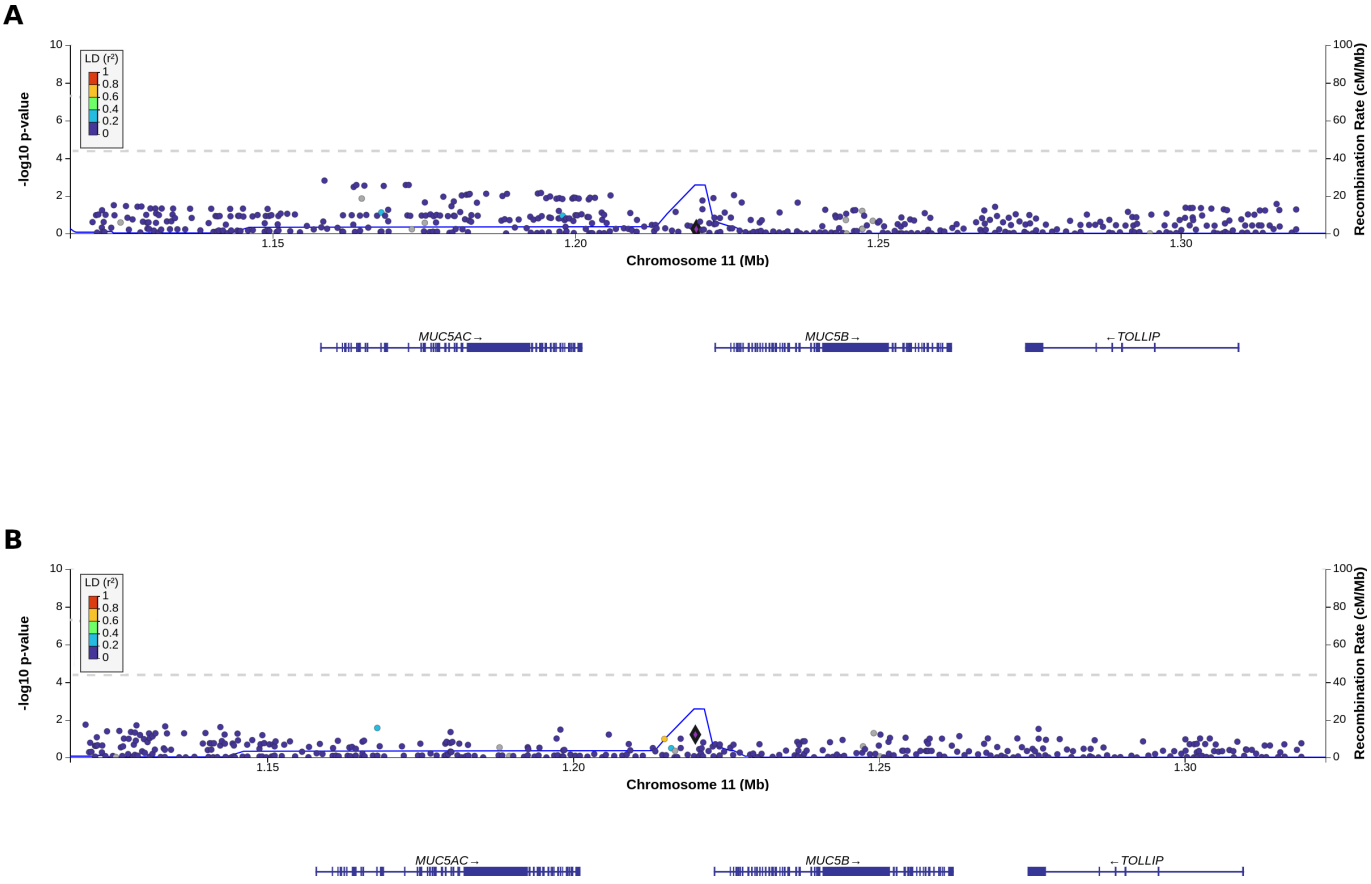

Figure 1 Locus zooms of the region surrounding the *MUC5B* rs35705950 variant**.** A) Locus zoom of the region surrounding rs35705950 for the comparison of SSc-ILD versus patients with SSc without ILD. Variant rs35705950 is represented as a diamond, and other SNPs as dots. The colour of the dots indicates linkage disequilibrium with rs35705950. Dashed line corresponds to 5.53×10^{-5} significance threshold. (B) Locus zoom of the region surrounding rs35705950 for the comparison of SSc-ILD versus controls. Variant rs35705950 is represented as a diamond, and other SNPs as dots. The colour of the dots indicates linkage disequilibrium with rs35705950. The dashed line corresponds to 5.53×10^{-5} significance threshold. LD, linkage disequilibrium; SNP, single nucleotide polymorphism; SSc-ILD, systemic sclerosis-associated interstitial lung disease.

0.87 to 1.27). Similarly, no significant association was observed when comparing patients with SSc-ILD with controls (p value: 0.061, OR: 1.16, 95% CI: 0.99 to 1.36). Forest plots showing the OR and 95% CI of rs35705950 for each cohort separately are available in online supplemental figure 3.

To further investigate potential associations in the genomic region, we analysed an additional 903 variants located within ±100 Kb of the *MUC5B* promoter SNP. None of these variants reached statistical significance after multiple testing correction, suggesting a lack of association in these loci in either comparison (figure 1). The 10 most significant variants in each comparison are included in table 1 and table 2. Association data for the remaining SNPs are included in online supplemental table S2 and S3.

DISCUSSION

In this study, we investigated the association between the *MUC5B* functional promoter variant rs35705950 and SSc-ILD in the largest cohort analysed to date. Our results showed no significant association between this variant, or any of the surrounding SNPs, and SSc-ILD. Notably, this study attained an 80% statistical power to detect associations in variants with MAF of 0.1, which is the case of the SNP rs35705950, and effect sizes of 1.35. Thus, if the promoter variant had a similar effect in SSc-ILD as it does in IPF or RA-ILD (ORs >3.1),[5 6] our study would have been sufficiently powered to detect it.

The findings from this work replicate the results from previous studies in SSc. More specifically, this association has been tested in the participants from Scleroderma Lung Study II clinical trial to assess its effect on SSc-ILD progression and on the clinical course of patients,[9] as well as in SSc cohorts with low sample size (less than 670 patients with SSc-ILD),[3 8 10 11] all of them reporting no association with the disease. These results suggest that, despite the clinical similarities, the pathogenic mechanisms underlying SSc-ILD differ from those of RA-ILD or IPF.

A previously proposed explanation for the differential association of rs35705950 across ILDs relates to the predominant histopathological pattern in each disease. The association with *MUC5B* has been consistently

Table 1 Summary statistics of *MUC5B* rs35705950 promoter variant and 10 most significant variants within 100 Kb for the comparison of SSc-ILD versus SSc without ILD

SNP	Ref allele	Effect allele	Position	Nearest gene	P value	OR (CI 95%)	Model	Population frequency*
rs35705950	**G**	**T**	**1 219 991**	***MUC5B***	**5.88E-01**	**1.05 (0.87 to 1.27)**	**FE**	**0.110**
rs34582600	G	C	1 158 593	*MUC5AC*	1.55E-03	1.06 (1.02 to 1.11)	RE	0.601
rs28463198	G	C	1 158 598	*MUC5AC*	1.56E-03	1.06 (1.02 to 1.11)	RE	0.601
rs28713701	C	G	1 158 590	*MUC5AC*	1.56E-03	1.06 (1.02 to 1.11)	RE	0.601
rs996372487	A	C	1 172 511	*MUC5AC*	2.64E-03	0.96 (0.93 to 0.98)	RE	0.402
rs914303524	C	T	1 172 014	*MUC5AC*	2.67E-03	1.04 (1.02 to 1.07)	RE	0.598
rs28707071	G	C	1 163 859	*MUC5AC*	2.67E-03	1.05 (1.02 to 1.08)	RE	0.590
rs28433367	T	C	1 165 184	*MUC5AC*	2.87E-03	0.95 (0.92 to 0.98)	RE	0.413
rs28404876	C	G	1 168 378	*MUC5AC*	2.98E-03	1.04 (1.01 to 1.07)	RE	0.594
rs28373192	C	T	1 163 418	*MUC5AC*	3.39E-03	1.05 (1.02 to 1.09)	RE	0.598
rs35917282	A	G	1 165 484	*MUC5AC*	3.46E-03	0.95 (0.91 to 0.98)	RE	0.410

In boldface, the *MUC5B* promoter variant rs35705950.
Genomic position is referenced to build hg38.
*Population frequency corresponds to the frequency of the effect allele in European (non-Finnish) population from gnomAD V.4.1.0 database.
FE, fixed effect; ILD, interstitial lung disease; RE, random effect; Ref, reference; SNP, single nucleotide polymorphism; SSc, systemic sclerosis.

observed in diseases like IPF and RA-ILD, all of which are primarily characterised by UIP. However, ILDs for which there is no association with this variant, as SSc-ILD or autoimmune myositis-associated ILD, present other predominant patterns as non-specific interstitial pneumonia.[1 17 18] These findings support the hypothesis that this variant is more likely associated with the UIP pattern rather than with ILDs in general. However, a study focusing on the UIP pattern in SSc-ILD did not find an association either, although its sample size was very limited, so further research is warranted in this subgroup of patients.[3] It is worth mentioning that nominally significant variants in the surrounding region of rs35705950 are located near the *MUC5AC* gene. Given that a previous work in IPF has discussed the independent effect of both genes,[19] we believe that further studies focussing on the role of *MUC5AC* in SSc-ILD are warranted.

Despite the strengths of our study, certain limitations must be acknowledged. First, we lacked the necessary data to stratify our cohort by its histopathological

Table 2 Summary statistics of *MUC5B* rs35705950 promoter variant and 10 most significant variants within 100 Kb for the comparison of patients with SSc-ILD versus controls

SNP	Ref allele	Effect allele	Position	Nearest gene	P value	OR (CI 95%)	Model	Population frequency*
rs35705950	**G**	**T**	**1 219 991**	***MUC5B***	**6.12E-02**	**1.16 (0.99 to 1.36)**	**FE**	**0.110**
rs75194112	G	A	1 120 287	*MUC2*	1.81E-02	0.68 (0.49 to 0.94)	FE	0.028
rs148428215	C	G	1 128 608	*MUC5AC*	1.97E-02	0.68 (0.50 to 0.94)	FE	0.028
rs139058053	T	A	1 133 305	*MUC5AC*	2.24E-02	0.69 (0.50 to 0.95)	FE	0.028
rs118034733	G	C	1 142 329	*MUC5AC*	2.43E-02	0.69 (0.50 to 0.95)	FE	0.028
rs28403537	C	T	1 167 980	*MUC5AC*	2.72E-02	1.31 (1.03 to 1.65)	FE	0.041
rs5744025	C	T	1 276 105	*TOLLIP*	3.03E-02	1.44 (1.04 to 2.00)	FE	0.020
rs34815853	C	A	1 197 927	*MUC5AC*	3.31E-02	1.29 (1.02 to 1.63)	FE	0.044
rs34474233	G	A	1 197 926	*MUC5AC*	3.33E-02	1.29 (1.02 to 1.63)	FE	0.044
rs75900172	A	G	1 125 806	*MUC2*	3.89E-02	0.85 (0.72 to 0.99)	FE	0.126
rs72842479	C	T	1 123 799	*MUC2*	4.10E-02	0.85 (0.72 to 0.99)	FE	0.127

In boldface, the *MUC5B* promoter variant rs35705950.
Genomic position is referenced to build hg38.
*Population frequency corresponds to the frequency of the effect allele in European (non-Finnish) population from gnomAD V.4.1.0 database.
FE, fixed effect; ILD, interstitial lung disease; Ref, reference; SNP, single nucleotide polymorphism; SSc, systemic sclerosis.

pattern. Given that some patients with SSc-ILD exhibit a UIP pattern, such stratification could have allowed us to detect an association specific to this subgroup, especially given the low sample size of the previous work focused on exploring rs37505950 in patients with SSc-ILD with UIP.[3] Additionally, a small percentage of patients included in this study (9.6%) were not diagnosed using HRCT, as during the early period of the study this technique was ordered in the presence of risk factors such as pulmonary symptoms, radiological abnormalities on chest X-ray or physiological abnormalities on PFTs. This may have caused the misclassification of patients with SSc-ILD, mainly by underdiagnosing them.[20] However, as this is a small proportion of the cohort, the potential loss of statistical power may not be substantial.

In summary, despite analysing the largest and most statistically powered SSc-ILD cohort to date, we did not detect an association between the *MUC5B* promoter variant rs35705950 and the surrounding SNPs with SSc-ILD. This suggests that, for this locus, the genetic architecture and pathogenic mechanisms underlying SSc-ILD differ from those of other ILDs such as IPF or RA-ILD. However, shared genetic risk factors may still exist at other loci, underscoring the need for further genetic studies across these conditions.

Author affiliations
[1]Department of Cell Biology and Immunology, Institute of Parasitology and Biomedicine Lopez-Neyra, CSIC, Granada, Spain
[2]Unit of Autoimmune Diseases, Department of Internal Medicine, Hospital Universitario Vall d'Hebron, Barcelona, Spain
[3]Department of Internal Medicine, Hospital San Cecilio, Granada, Spain
[4]Department of Rheumatology and Clinical Immunology, Amsterdam Rheumatology and immunology Center, Amsterdam UMC, Amsterdam, Netherlands
[5]Department of Dermatology, Venereology, and Allergology, Helios St. Elisabeth Klinik Oberhausen, University Witten-Herdecke, Oberhausen, Germany
[6]Department of Rheumatology, University Hospital Zürich Center of Experimental Rheumatology, Zürich, Switzerland
[7]Rheumatology Unit, Royal Adelaide Hospital and University of Adelaide, Adelaide, Victoria, Australia
[8]University of Sydney School of Public Health, Royal Prince Alfred Hospital Sydney, and Sydney MSK Research Flagship Centre, University of Sydney, Sydney, New South Wales, Australia
[9]Department of Dermatology and Venereology, University Hospital Cologne, Cologne, Germany
[10]Department of Rheumatology, Leiden University Medical Center, Leiden, Netherlands
[11]Centre for Musculoskeletal Research, The University of Manchester, Northern Care Alliance NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK
[12]NIHR Manchester Biomedical Research Centre, Manchester University NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK
[13]Department of Rheumatology, Université Paris Cité UFR de Médecine, Paris, France
[14]Referral Center for Systemic Autoimmune Diseases, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico di Milano, Milan, Italy
[15]Division of Rheumatology, Department of internal medicine, The University of Texas Health Science Center at Houston, Houston, Texas, USA
[16]Division of Medicine, University College London, London, UK

X Alfredo Guillen-Del-Castillo @_AlfredoGuillen

Acknowledgements We sincerely thank Sofia Vargas and Gema Robledo for their exceptional technical support and to all the patients and control donors for their invaluable participation and cooperation. This research is part of the doctoral degree awarded to CR-B, within the Biomedicine program from the University of Granada.

Collaborators International SSc Group: P Carreira, Department of Rheumatology, 12 de Octubre University Hospital, Madrid, Spain; I Castellvi, Department of Rheumatology, Santa Creu i Sant Pau University Hospital, Barcelona, Spain; R Ríos, Department of Internal Medicine, San Cecilio Clinic University Hospital, Granada, Spain; N Ortego-Centeno, Department of Medicine, University of Granada, Granada, Spain; R García Portales, Department of Rheumatology, Virgen de la Victoria Hospital, Málaga, Spain; A Fernández-Nebro, Department of Rheumatology, Carlos Haya Hospital, Málaga, Spain; F J García-Hernández, Department of Internal Medicine, Virgen del Rocío Hospital, Sevilla, Spain; M A Aguirre, Department of Rheumatology, Reina Sofía/IMIBIC Hospital, Córdoba, Spain; B Fernández-Gutiérrez, Department of Rheumatology, San Carlos Clinic Hospital, Madrid, Spain; L Rodríguez-Rodríguez, Department of Rheumatology, San Carlos Clinic Hospital, Madrid, Spain; P García de la Peña, Department of Rheumatology, Madrid Norte Sanchinarro Hospital, Madrid, Spain; E Vicente, Department of Rheumatology, La Princesa Hospital, Madrid, Spain; J L Andreu, Department of Rheumatology, Puerta de Hierro Hospital-Majadahonda, Madrid, Spain; M Fernández de Castro, Department of Rheumatology, Puerta de Hierro Hospital-Majadahonda, Madrid, Spain; F J López-Longo, Department of Rheumatology, Gregorio Marañón University Hospital, Madrid, Spain; V Fonollosa, Department of Internal Medicine, Valle de Hebrón Hospital, Barcelona, Spain; G Espinosa, Department of Internal Medicine, Clinic Hospital, Barcelona, Spain; C Tolosa, Department of Internal Medicine, Parc Tauli Hospital, Sabadell, Spain; A Pros, Department of Rheumatology, Hospital Del Mar, Barcelona, Spain; L Tío, Department of Rheumatology, Hospital Del Mar, Barcelona, Spain; E Beltrán, Department of Rheumatology, Hospital Del Mar, Barcelona, Spain; M Rodríguez Carballeira, Department of Internal Medicine, Hospital Universitari Mútua Terrasa, Barcelona, Spain; F J Narváez, Department of Rheumatology, Bellvitge University Hospital, Barcelona, Spain; M Rubio Rivas, Department of Internal Medicine, Bellvitge University Hospital, Barcelona, Spain; V Ortiz-Santamaría, Department of Rheumatology, Granollers Hospital, Granollers, Spain; A B Madroñero, Department of Internal Medicine, Hospital General San Jorge, Huesca, Spain; M A González-Gay, Epidemiology, Genetics and Atherosclerosis Research Group on Systemic Inflammatory Diseases, IDIVAL, University of Cantabria, Santander, Spain; B Díaz, Department of Internal Medicine, Hospital Central de Asturias, Oviedo, Spain; L Trapiella, Department of Internal Medicine, Hospital Central de Asturias, Oviedo, Spain; M V Egurbide, Department of Internal Medicine, Hospital Universitario Cruces, Barakaldo, Spain; P Fanlo-Mateo, Department of Internal Medicine, Hospital Virgen del Camino, Pamplona, Spain; L Saez-Comet, Department of Internal Medicine, Hospital Universitario Miguel Servet, Zaragoza, Spain; F Díaz, Department of Rheumatology, Hospital Universitario de Canarias, Tenerife, Spain; J A Roman-Ivorra, Department of Rheumatology, Hospital Universitari i Politecnic La Fe, Valencia, Spain; J J Alegre Sancho, Department of Rheumatology, Hospital Universitari Doctor Peset, Valencia, Spain; M Freire, Department of Internal Medicine, Thrombosis and Vasculitis Unit, Complexo Hospitalario Universitario de Vigo, Vigo, Spain; F J Blanco Garcia, Department of Rheumatology, INIBIC-Hospital Universitario A Coruña, La Coruña, Spain; N Oreiro, Department of Rheumatology, INIBIC-Hospital Universitario A Coruña, La Coruña, Spain; T Witte, Department of Clinical Immunology, Hannover Medical School, Hannover, Germany; G Riemekasten, Clinic of Rheumatology, University of Lübeck, Lübeck, Germany; P Airo, Service of Rheumatology and Clinic Immunology Spedali Civili, Brescia, Italy; C Magro, Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands; M C Vonk, Department of Rheumatology, Radboud University Nijmegen Medical Center, Nijmegen, Netherlands; R Hesselstrand, Department of Rheumatology, Lund University, Lund, Sweden; A Nordin, Division of Rheumatology, Department of Medicine, Karolinska University Hospital, Karolinska Institute, Stockholm, Sweden; C Lunardi, Department of Medicine, Università degli Studi di Verona, Verona, Italy; G Moroncini, Department of Clinical and Molecular Science, Università Politecnica delle Marche and Ospedali Riuniti, Ancona, Italy; A Gabrielli, Istituto di Clinica Medica Generale, Ematologia ed Immunologia Clinica, Università Politecnica delle Marche, Ancona, Italy; A Hoffmann-Vold, Department of Rheumatology, Oslo University Hospital, Oslo, Norway; J H W Distler, Department of Internal Medicine, Institute for Clinical Immunology, University of Erlangen-Nuremberg, Erlangen, Germany; L Padyukov, Division of Rheumatology, Department of Medicine, Karolinska University Hospital, Karolinska Institute, Stockholm, Sweden; B P C Koeleman, University Medical Center Utrecht, Utrecht, The Netherlands; T R D J Radstake, University Medical Center Utrecht, Utrecht, The Netherlands; G Orozco, Centre for Genetics and Genomics Versus Arthritis, Division of Musculoskeletal and Dermatological Sciences, School of Biological Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester; A Barton, Centre for Genetics and Genomics Versus Arthritis, Division of Musculoskeletal and Dermatological Sciences, School of Biological Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester; C Fonseca, Center for Rheumatology, Royal Free and University College Medical School, London, UK. Australian Scleroderma Interest Group (ASIG): A Apostolopoulos, Department of

Rheumatology, Monash Medical Centre, Melbourne, Victoria, Australia; Department of Medicine, Monash University, Melbourne, Australia; N Ferdowsi, Department of Rheumatology, St Vincent's Hospital, Melbourne, Victoria, Australia; D Hansen, Department of Rheumatology, St Vincent's Hospital, Melbourne, Victoria, Australia; L V Host, Fiona Stanley Hospital, Perth, Western Australia, Australia; G Major, Department of Rheumatology, John Hunter Hospital, New Lambton Heights, New South Wales, Australia; Faculty of Medicine, University of Newcastle, Callaghan NSW Australia; G S Ngian, Department of Rheumatology, Monash Medical Centre, Melbourne, Victoria, Australia; Department of Medicine, Monash University, Melbourne, Australia; A Quinlivan; Department of Rheumatology, St Vincent's Hospital, Melbourne, Victoria, Australia; Department of Medicine, The University of Melbourne, Australia; L Ross, Department of Rheumatology, St Vincent's Hospital, Melbourne, Victoria, Australia; Department of Medicine, The University of Melbourne, Australia; J Sahhar, Department of Rheumatology, Monash Medical Centre, Melbourne, Victoria, Australia; Department of Medicine, Monash University, Melbourne, Australia; W Stevens; Department of Rheumatology, St Vincent's Hospital, Melbourne, Victoria, Australia; M Tabesh; Department of Rheumatology, St Vincent's Hospital, Melbourne, Victoria, Australia; J Walker, Rheumatology unit, Royal Adelaide Hospital, Adelaide, SA, Australia; Flinders Medical Centre, SA, Australia; Flinders University, Adelaide, SA, Australia.

Contributors Conceptualisation: JM, MA-H; Methodology: MA-H; Formal analysis: CR-B; Resources: AG-D-C, CPS-A, JLC, AV, AK, OD, SMP, MN, NH, JDV-B, ALH, YA, LB, MM, CD, SA; Data curation: CR-B, CR-P, IR-M, MK; Writing original draft: CR-B, MA-H; Visualisation: CR-B; Supervision: JM, MA-H; Project Administration: MA-H. All authors have carefully read and approved the final version of the manuscript. MA-H is the guarantor.

Funding This work was supported by grant PID2022-139292OB-I00 funded by MICIU/AEI/10.13039/501100011033 and by ERDF/EU. C.R.B. was supported by the program FPU: [FPU22/01652], funded by the Spanish Ministry of Science Innovation and Universities. C.R.P. was funded by Grant PREP2022-000747 funded by MICIU/AEI /10.13039/501100011033 and by ESF+. I.R.M. was supported by the program FPU: [FPU21/02746], funded by the Spanish Ministry of University. S.A. was supported by R33AR078078, R61AR078078, R01AR081280 and DoD-W81XWH-22-1-0162.

Competing interests SMP has received honoraria from Janssen, Boehringer Ingelheim, MSD. MN has received honoraria or consultancies from Janssen, AstraZeneca, GlaxoSmithKlein, Boehringer Ingelheim and Bristol Myers Squibb. The Australian Scleroderma Cohort Study (ASCS) is supported by Janssen, Boehringer Ingelheim, Scleroderma Australia, Scleroderma Victoria, Arthritis Australia, Musculoskeletal Australia (muscle, bone and joint health), Australian Rheumatology Association and St Vincent's Hospital Melbourne IT Department and Research Endowment Fund (REF).

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by CSIC's ethics committee, and written informed consent was obtained in accordance with the tenets of the Declaration of Helsinki.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Carlos Rosa-Baez http://orcid.org/0000-0002-1257-3338
Carlos Rangel-Pelaez http://orcid.org/0000-0001-7697-1696
Inmaculada Rodriguez-Martin http://orcid.org/0000-0001-5413-1581
Martin Kerick http://orcid.org/0000-0002-6298-4514
Alfredo Guillen-Del-Castillo http://orcid.org/0000-0003-0626-507X
Oliver Distler http://orcid.org/0000-0002-0546-8310
Jeska K De Vries-Bouwstra http://orcid.org/0000-0002-5624-1415
Ariane L Herrick http://orcid.org/0000-0003-4941-7926
Lorenzo Beretta http://orcid.org/0000-0002-6529-6258
Marialbert Acosta-Herrera http://orcid.org/0000-0002-9868-6535

REFERENCES

1 Liakouli V, Ciancio A, Del Galdo F, *et al*. Systemic sclerosis interstitial lung disease: unmet needs and potential solutions. *Nat Rev Rheumatol* 2024;20:21–32.
2 Spagnolo P, Distler O, Ryerson CJ, *et al*. Mechanisms of progressive fibrosis in connective tissue disease (CTD)-associated interstitial lung diseases (ILDs). *Ann Rheum Dis* 2021;80:143–50.
3 Bernstein EJ, Boin F, Elicker B, *et al*. n.d. FAM13A polymorphism is associated with a usual interstitial pneumonia pattern in patients with systemic sclerosis-associated interstitial lung disease. *Rheumatology (Sunnyvale)*.
4 Juge P-A, Borie R, Kannengiesser C, *et al*. Shared genetic predisposition in rheumatoid arthritis-interstitial lung disease and familial pulmonary fibrosis. *Eur Respir J* 2017;49:1602314.
5 Seibold MA, Wise AL, Speer MC, *et al*. A common MUC5B promoter polymorphism and pulmonary fibrosis. *N Engl J Med* 2011;364:1503–12.
6 Juge P-A, Lee JS, Ebstein E, *et al*. MUC5B Promoter Variant and Rheumatoid Arthritis with Interstitial Lung Disease. *N Engl J Med* 2018;379:2209–19.
7 Mathai SK, Newton CA, Schwartz DA, *et al*. Pulmonary fibrosis in the era of stratified medicine. *Thorax* 2016;71:1154–60.
8 Stock CJ, Sato H, Fonseca C, *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.
9 Volkmann ER, Tashkin DP, Roth MD, *et al*. The MUC5B promoter variant does not predict progression of interstitial lung disease in systemic sclerosis. *Semin Arthritis Rheum* 2020;50:963–7.
10 Peljto AL, Steele MP, Fingerlin TE, *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.
11 Borie R, Crestani B, Dieude P, *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.
12 LeRoy EC, Medsger TA. Criteria for the classification of early systemic sclerosis. *J Rheumatol* 2001;28:1573–6.
13 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. *Ann Rheum Dis* 2013;72:1747–55.
14 Chang CC, Chow CC, Tellier LC, *et al*. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 2015;4:7.
15 Yang J, Lee SH, Goddard ME, *et al*. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 2011;88:76–82.
16 Han B, Eskin E. Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. *Am J Hum Genet* 2011;88:586–98.
17 Hallowell RW, Paik JJ. Myositis-associated interstitial lung disease: a comprehensive approach to diagnosis and management. *Clin Exp Rheumatol* 2022;40:373–83.
18 Johnson C, Rosen P, Lloyd T, *et al*. Exploration of the MUC5B promoter variant and ILD risk in patients with autoimmune myositis. *Respir Med* 2017;130:52–4.
19 Lorenzo-Salazar JM, Ma S-F, Jou J, *et al*. Novel idiopathic pulmonary fibrosis susceptibility variants revealed by deep sequencing. *ERJ Open Res* 2019;5:00071-2019.
20 Walsh SLF, Devaraj A, Enghelmayer JI, *et al*. Role of imaging in progressive-fibrosing interstitial lung diseases. *Eur Respir Rev* 2018;27:180073.