Universiteit Leiden
The Netherlands

Sex-specific autosomal susceptibility loci in systemic sclerosis: a genome-wide association study

Rodriguez-Martin, I.; Kerick, M.; Rangel-Peláez, C.; Rosa-Baez, C.; Borrego-Yaniz, G.; Ortiz-Fernández, L.; ... ; Australian Scleroderma Interest Grp

Citation
Rodriguez-Martin, I., Kerick, M., Rangel-Peláez, C., Rosa-Baez, C., Borrego-Yaniz, G., Ortiz-Fernández, L., ... Acosta-Herrera, M. (2026). Sex-specific autosomal susceptibility loci in systemic sclerosis: a genome-wide association study. *The Lancet Rheumatology*, *8*(6), e438-e450. doi:10.1016/S2665-9913(25)00376-5

Version: Publisher's Version
License: Creative Commons CC BY-NC-ND 4.0 license
Downloaded from: https://hdl.handle.net/1887/4308232

Note: To cite this publication please use the final published version (if applicable).

Sex-specific autosomal susceptibility loci in systemic sclerosis: a genome-wide association study

Inmaculada Rodriguez-Martin, Martin Kerick, Carlos Rangel-Peláez, Carlos Rosa-Baez, Gonzalo Borrego-Yaniz, Lourdes Ortiz-Fernández, Alfredo Guillen-Del-Castillo, Carmen P Simeón-Aznar, José Luis Callejas, Oliver Distler, Susanna M Proudman, Mandana Nikpour, Nicolas Hunzelmann, Jeska K de Vries-Bouwstra, Ariane L Herrick, Yannick Allanore, Marta E Alarcón-Riquelme, Lorenzo Beretta, Shervin Assassi, Christopher P Denton, Maureen D Mayes, Javier Martin, Marialbert Acosta-Herrera on behalf of the International SSc Group, the PRECISESADS Clinical Consortium, and the Australian Scleroderma Interest Group

Lancet Rheumatol 2026; 8: e438–50

Published Online April 15, 2026 https://doi.org/10.1016/S2665-9913(25)00376-5

See Comment page e410

Institute of Parasitology and Biomedicine López-Neyra, Spanish National Research Council, Granada, Spain (I Rodriguez-Martin MSc, M Kerick PhD, C Rangel-Peláez MSc, C Rosa-Baez MSc, G Borrego-Yaniz MSc, L Ortiz-Fernández PhD, Prof J Martin MD PhD, M Acosta-Herrera PhD); **Unit of Systemic Autoimmune Diseases, Department of Internal Medicine, Hospital Universitari Vall d'Hebron, Barcelona, Spain** (A Guillen-Del-Castillo MD PhD, C P Simeón-Aznar MD PhD); **Systemic Autoimmune Disease Unit, Hospital Clínico San Cecilio, Instituto de Investigación Biosanitaria Ibs. GRANADA, Granada, Spain** (J L Callejas MD PhD); **Department of Rheumatology, University Hospital Zurich, University of Zurich, Zurich, Switzerland** (Prof O Distler MD); **Rheumatology Unit, Royal Adelaide Hospital and University of Adelaide, Adelaide, SA, Australia** (Prof S M Proudman MBBS); **The University of Sydney, Sydney, NSW, Australia** (Prof M Nikpour MBBS PhD); **Department of Dermatology, University of Cologne, Cologne, Germany** (Prof N Hunzelmann MD); **Department of Rheumatology, Leiden University Medical Center, Leiden, Netherlands** (J K de Vries-Bouwstra MD PhD); **Division of Musculoskeletal and Dermatological Sciences, The University of Manchester, Northern Care Alliance NHS**

Summary

Background Systemic sclerosis is an immune-mediated inflammatory disease with marked sex differences in prevalence and severity. Although genome-wide association studies (GWAS) have advanced the understanding of systemic sclerosis genetics, sex-aware approaches remain scarce. We aimed to address this gap by examining autosomal sex-specific genetic factors in systemic sclerosis.

Methods Based on a chromosomal definition of sex, we conducted a sex-stratified autosomal meta-analysis of GWAS in systemic sclerosis. We retrieved patient-level and control data from a previous GWAS for systemic sclerosis and newly recruited participants from an international multicentre collaboration (data cutoff June 1, 2024). Adult patients (aged ≥18 years) were required to meet the 2013 American College of Rheumatology–European League Against Rheumatism classification criteria or the criteria for early systemic sclerosis proposed by LeRoy and Medsger, 2001. Patients with systemic sclerosis were compared with healthy controls through sex-stratified analyses. Sex-specific loci were functionally annotated and a sex-differential expression analysis was performed to validate the findings. Additionally, we performed a drug repurposing analysis to explore the clinical implications of our findings. People with lived experience of systemic sclerosis were not involved in the study design or conduct.

Findings After quality control filtering, we included 10 653 patients with systemic sclerosis (1556 male and 9097 female) and 18 043 controls (6990 male and 11 053 female). We identified eight sex-specific loci associated with systemic sclerosis: one novel male-specific locus (*BCL11A* odds ratio [OR] 1·32 [95% CI 1·20–1·45], $p=8·39\times10^{-9}$), two novel female-specific loci (*IRF4* OR 0·78 [0·71–0·85], $p=2·36\times10^{-8}$; *RPL3-PDGFB* OR 1·19 [1·12–1·26], $p=3·18\times10^{-8}$), and five previously known loci now characterised as female-specific. Functional analyses implicated 39 candidate genes, of which five (*IL12RB2, ARHGAP31, UBL7, CSK*, and *MPI*) showed sex-differential expression in patients with systemic sclerosis. Two approved drugs, fostamatinib and dasatinib, were identified as potential repurposing candidates for systemic sclerosis sex-aware treatment.

Interpretation We report the first evidence of sex-specific genetic architecture in systemic sclerosis. These findings offer insight into the molecular basis of sex differences in the disease, highlighting both *IL12RB2* gene and interferon signalling. These results also support the development of sex-aware precision medicine approaches.

Funding Instituto de Salud Carlos III, EU's Seventh Framework Program, and European Federation of Pharmaceutical Industries and Associations companies.

Copyright © 2026 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY-NC-ND 4.0 license.

Introduction

Systemic sclerosis is a rare immune-mediated inflammatory disease characterised by immune dysregulation, vasculopathy, and extensive fibrosis of the skin and internal organs.[1] Despite advances in the understanding of its pathogenesis, systemic sclerosis imposes a substantial burden on patients and health-care systems due to its high morbidity and its status as the rheumatic disease with the highest mortality rate.[1,2] Similar to other immune-mediated inflammatory diseases, systemic sclerosis presents a marked sex bias and disproportionately affects female individuals, with a female-to-male ratio of approximately 8:1.[3] Male patients have a more severe clinical course and higher mortality than female patients.[4] These observations suggest that sex has a fundamental role in the pathogenesis, progression, and outcomes of systemic sclerosis.

Multiple factors have been proposed to explain sex differences in systemic sclerosis.[4,5] Differences in the immune response have been widely reported in healthy people, with female individuals displaying more robust activation of innate, humoral, and cellular immune pathways than male individuals.[5] Additional mechanisms relate to sex chromosome effects, such as an escape from

Research in context

Evidence before this study

Systemic sclerosis is a complex, immune-mediated inflammatory disease with one of the highest mortality rates among rheumatic diseases. Female individuals are disproportionately affected (8:1 female-to-male ratio), although male individuals tend to present with more severe disease and poorer prognosis. Multiple biological factors have been proposed to explain sex differences in rheumatic diseases, namely hormonal influences, sex chromosome effects, and well established sex-based differences in immune function. Although genome-wide association studies (GWAS) have identified genetic risk factors for systemic sclerosis, none have systematically assessed its sex-specific autosomal genetic architecture. We searched PubMed, with no language restrictions, for studies up to Sept 15, 2025, using the terms "sex-specific genet* scleroderma", "sex-diff* genet* scleroderma", "sex-bias genet* scleroderma", "sex genetic architecture scleroderma". The search retrieved epidemiological and review studies emphasising the relevance of sex-aware studies and sex differences in the disease. A few studies have examined sex differences at molecular levels, including transcriptome-wide differences in skin, copy number variation of complement component *C4*, and expression of miR-146a, which is known to downregulate the X-linked gene *IRAK1*. These studies were limited in scope and did not provide a genome-wide autosomal assessment of sex-specific genetic effects.

Added value of this study

To our knowledge, this study is the first sex-stratified meta-analysis of genomic data in systemic sclerosis. Using the largest European systemic sclerosis cohort to date, we identified eight sex-specific genetic associations, including three novel loci. Whole-blood transcriptomic analysis functionally validated five of the proposed genes with sex-differential expression in systemic sclerosis. This study highlights *IL12RB2* as a key female-specific gene associated with systemic sclerosis. We also identified that autosomal genetic predisposition in interferon signalling could contribute to systemic sclerosis through a sex-biased mechanism, based on the prioritisation of the interferon-related genes *UBL7*, *IRF4*, and *IRF8*. Additionally, we identified candidate drugs for potential repurposing in a sex-aware manner based on genetic evidence.

Implications of all the available evidence

Although sex differences in systemic sclerosis are well documented clinically, the biological basis for these disparities remains poorly understood. Our findings, in line with previously reported risk conferred by complement component C4, indicate that sex-specific differences arise as early as the level of genetic predisposition, given that we show sex-related autosomal genetic factors contributing to disease susceptibility. Consistent with research in other conditions, the results highlight the need for sex-aware analytical approaches to uncover disease mechanisms that may be overlooked when both sexes are analysed together. Overall, the growing body of evidence supports advancing toward sex-informed biomedical research.

Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK (Prof A L Herrick MD); **Department of Rheumatology, Cochin Hospital, INSERM U1016, Université Paris Cité, Paris, Île-de-France, France** (Prof Y Allanore MD PhD); **Centre for Genomics and Oncological Research, Pfizer, University of Granada/Andalusian Regional Government, Granada, Spain** (Prof M E Alarcón-Riquelme MD PhD); **Referral Center for Systemic Autoimmune Diseases, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico di Milano, Milan, Italy** (L Beretta MD); **Division of Rheumatology, The University of Texas Health Science Center at Houston, Houston, TX, USA** (Prof S Assassi MD MS, Prof M D Mayes MD); **Center for Rheumatology, Royal Free and University College Medical School, London, UK** (Prof C P Denton PhD)

Correspondence to:
Dr Marialbert Acosta-Herrera, Institute of Parasitology and Biomedicine López-Neyra, Spanish National Research Council, Granada 18016, Spain
m.acostaherrera@ipb.csic.es

X chromosome inactivation.[4,5] Hormonal influences and environmental factors, including lifestyle behaviours and occupational exposures, have also been proposed as contributing factors for the observed sex differences.[4] However, the exact mechanisms by which sex influences systemic sclerosis pathogenesis remain unclear.[2]

Autosomal genetic predisposition contributes to systemic sclerosis development, as shown in the largest European-descent genome-wide association study (GWAS) performed to date.[6] However, this study was conducted without evaluating the sex differences in the disease.[6] Meanwhile, emerging evidence highlights the importance of sex-aware genetic assessments to uncover relevant biological insights in sex-biased diseases.[7] In systemic sclerosis, research has been focused on sex differences in gene expression,[8,9] and one study of copy number variation of complement component *C4* identified sex-dependent differences in genetics, gene expression, and serum protein concentrations in patients with systemic sclerosis.[10] These observations suggest that sex-specific genetic mechanisms might have a substantial role in the pathogenesis of the disease, yet they remain largely unexplored[11] due to the absence of sex-stratified GWAS. Guided by this hypothesis, we aimed to investigate the contribution of sex-specific autosomal genetic factors in systemic sclerosis by performing the first sex-aware meta-analysis of GWAS in the disease.

Methods

Study design and participants

This GWAS meta-analysis comprised patients with systemic sclerosis and healthy controls of European descent from 15 different cohorts. European ancestry was established based on geography and genetic similarity. The majority of participants were retrieved and recruited from a previous GWAS.[6] The remaining participants were newly recruited for the present study, all as part of an international, multicentre collaborative effort with a data cutoff of June 1, 2024. Details of the cohorts are provided in the appendix (p 13). All adult patients (ie, aged ≥18 years) with systemic sclerosis included in the study met the 2013 American College of Rheumatology–European League Against Rheumatism classification criteria,[12] which uses a weighted additive score based on skin thickening, vascular and microvascular features, organ involvement, Raynaud's disease, and disease-specific autoantibodies, or the criteria for early systemic sclerosis proposed by LeRoy and Medsger,[13] which includes

See Online for appendix

Raynaud's disease, systemic sclerosis-specific autoantibodies, and characteristic nailfold capillaroscopy abnormalities. Sex was treated as a binary variable and defined through genetic inference from genotype data based on X-chromosome heterozygosity. Although this approach provides an objective and consistent definition for genetic analyses, we acknowledge that sex is not an entirely discrete concept and its biological attributes and expression could vary.[14] Importantly, this genetically inferred definition might not always correspond to sex assigned at birth or self-reported sex, as typically used in epidemiological studies. Accordingly, references throughout the manuscript to sex-specific, sex-differential, or sex-dependent effects refer specifically to this genetically inferred definition. The Ethics Committee of the Spanish National Research Council approved the study, and all participants voluntarily gave their written informed consent in compliance with the principles of the Declaration of Helsinki. There were no people with lived experience included in the study.

Genotyping, quality control, and imputation

Genome-wide genotyping was performed with DNA extracted from saliva or whole blood, collected by clinical investigators. We applied rigorous and established quality-control measures using PLINK version 2.0, as described elsewhere.[6] Briefly, variants were excluded if they had a minor allele frequency lower than 1%, a low call rate (<98%), deviated from Hardy–Weinberg equilibrium ($p<0·001$), or were palindromic single nucleotide polymorphisms (SNPs) with a minor allele frequency of more than 40%. Samples were filtered out if they had a call rate of less than 95% or ambiguous sex annotation. Additionally, identity-by-descent estimation was used to identify and remove one individual from every pair of duplicates (ie, those with pi_hat >0·99) or relatives (pi_hat >0·45).

For more on **PLINK** see https://www.cog-genomics.org/plink/ and https://www.cog-genomics.org/plink/2.0/

To increase genomic coverage, genotypes were imputed (ie, inferred) with the TOPMed Imputation Server and the TOPMed Imputation reference panel (appendix p 13). Variants with low imputation quality ($R^2<0·3$) or minor allele frequency lower than 1% were removed from further analyses.

For more on **GCTA-COJO** see https://yanglab.westlake.edu.cn/software/gcta/#COJO

See **Online** for appendix

To identify individuals with different genetic ancestry and correct for population structure, we performed principal component analysis separately for each cohort (appendix pp 4, 12–13) using PLINK and Genome-wide Complex Trait Analysis (GCTA). Approximately 100 000 independent variants were used to calculate ten principal components per individual, and samples deviating by more than four SDs from the cluster centroid of their respective cohort were discarded from further analyses.

For more on **GCTA** see https://yanglab.westlake.edu.cn/software/gcta/#Overview

Data analysis

We assessed sex-dependent genetic associations through logistic regression on allele dosages in PLINK, including the first five principal components as covariates. Four types of logistic regression models were conducted: sex-stratified analyses in male and female participants separately, a model including a genotype-by-sex interaction term (also adjusted by sex), and a model comparing female and male patients with systemic sclerosis (intracase by sex). We combined the results of the included cohorts through an inverse variance-weighted meta-analysis using METASOFT (Buhm Han and Eleazar Eskin) for each type of logistic regression model. Both the fixed-effects and the Han and Eskin random-effects models were used. Variants were considered significantly associated if the p value was $<5\times10^{-8}$, and suggestively associated if the p value was $<1\times10^{-5}$. The MHC region was excluded from further analyses due to its complex linkage disequilibrium structure. We further evaluated sex differences through Cochran's Q tests of sex-specific effect sizes and a meta-analysis between sexes using the R package ASSET. This subset-based analysis allowed us to identify the best sex combination to explain the observed association for each variant (male only, female only, or both).

A locus was considered sex-specific if it met the following criteria: (1) genome-wide significance in one sex ($p<5\times10^{-8}$); (2) no association in the other sex ($p>0·05$); (3) best fit in the sex-only model in the ASSET subset-based meta-analysis; and (4) high posterior probability of replicability (PPR; >0·9) in that sex but low PPR (<0·1) in the other.

To ensure that female-specific loci were not driven by sample size differences, we retained those that either did not reach genome-wide significance when both sexes were analysed together (ie, the combined dataset [$p>5\times10^{-8}$]) or had all proxy variants in strong linkage disequilibrium ($R^2>0·8$) best fit in the female-only model in the ASSET subset-based meta-analysis.

Finally, we investigated whether significant genomic regions contained more than one independent signal through a joint conditional analysis using GCTA-COJO. This method considers the linkage disequilibrium pattern when testing additional variants in the region. The most associated variant (lead SNP) was included as a covariate and variants within a range of 1·5 Mb of the lead SNP were analysed. Variants were considered secondary signals if the p value was $<1\times10^{-5}$ and were in low linkage disequilibrium with the lead SNP ($R^2<0·2$ and $D'<0·5$).

Gene expression analysis

A subset of the dataset was used to evaluate the effect of sex and systemic sclerosis on the expression levels of candidate genes prioritised by the in silico functional annotation (appendix p 6). This subset was selected based on the availability of whole-blood RNA-seq data, with no additional inclusion or exclusion criteria applied beyond those of the original dataset. Whole blood RNA-seq data were obtained as described by Beretta and colleagues.[15]

Briefly, we removed genes with low expression (mean read count lower than three) and used the edgeRs TMM (trimmed mean of M-values) algorithm to calculate a factor offset, which represents library size and library compositional bias, to be included as a covariate into the final model, which was fit for log2-transformed read counts (plus pseudocount of 0·125). To adjust gene expression by cell type composition, we used fluorescence-activated cell sorting data and Cell-type Identification by Estimating Relative Subsets of RNA Transcripts estimates, considering the following cell types: naive B cells, memory B cells, plasma cells, $CD8^+$ T cells, naive $CD4^+$ T cells, memory resting $CD4^+$ T cells, memory activated $CD4^+$ T cells, regulatory T cells, natural killer cells, monocytes, dendritic cells, mastocytes, neutrophils, eosinophils and basophils.

We conducted a whole-blood differential expression analysis with an interaction term between sex and disease status analysing a subset of patients with systemic sclerosis and healthy controls. A linear model was fitted for each gene, adjusting for cell type composition, age, disease status, sex, sex-by-disease interaction, and the offset term. Next, non-informative non-zero terms were removed through stepwise regression backward elimination. Genes were considered differentially expressed if they had a false discovery rate of <0·05 for the interaction term.

Drug repurposing

To identify potential new treatment options for systemic sclerosis in a sex-aware manner, we performed a drug repurposing analysis. Briefly, we queried the DrugBank database for drugs targeting the proteins encoded by the 39 genes prioritised in the functional annotation analysis. To refine the list, we filtered for drugs that are either approved or with clinical trial evidence for immune-mediated inflammatory diseases, given their relevance to systemic sclerosis. Moreover, we supplemented this information with manual literature curation by querying PubMed and ClinicalTrials.gov for studies linking each candidate drug to systemic sclerosis, fibrotic, or inflammatory conditions, to validate the biological and clinical relevance of the identified candidates (appendix p 7). Full details of additional methodological analyses are provided in the appendix (pp 4–7).

Role of the funding source

The funder of the study had no role in the study design, data collection, data analysis, data interpretation, writing of the report.

Results

In total, 15 cohorts were collected through international collaboration consisting of 10 903 patients with systemic sclerosis (9320 female patients and 1583 male patients) and 18123 healthy controls (11107 female controls and 7016 male controls). After all quality controls, the final

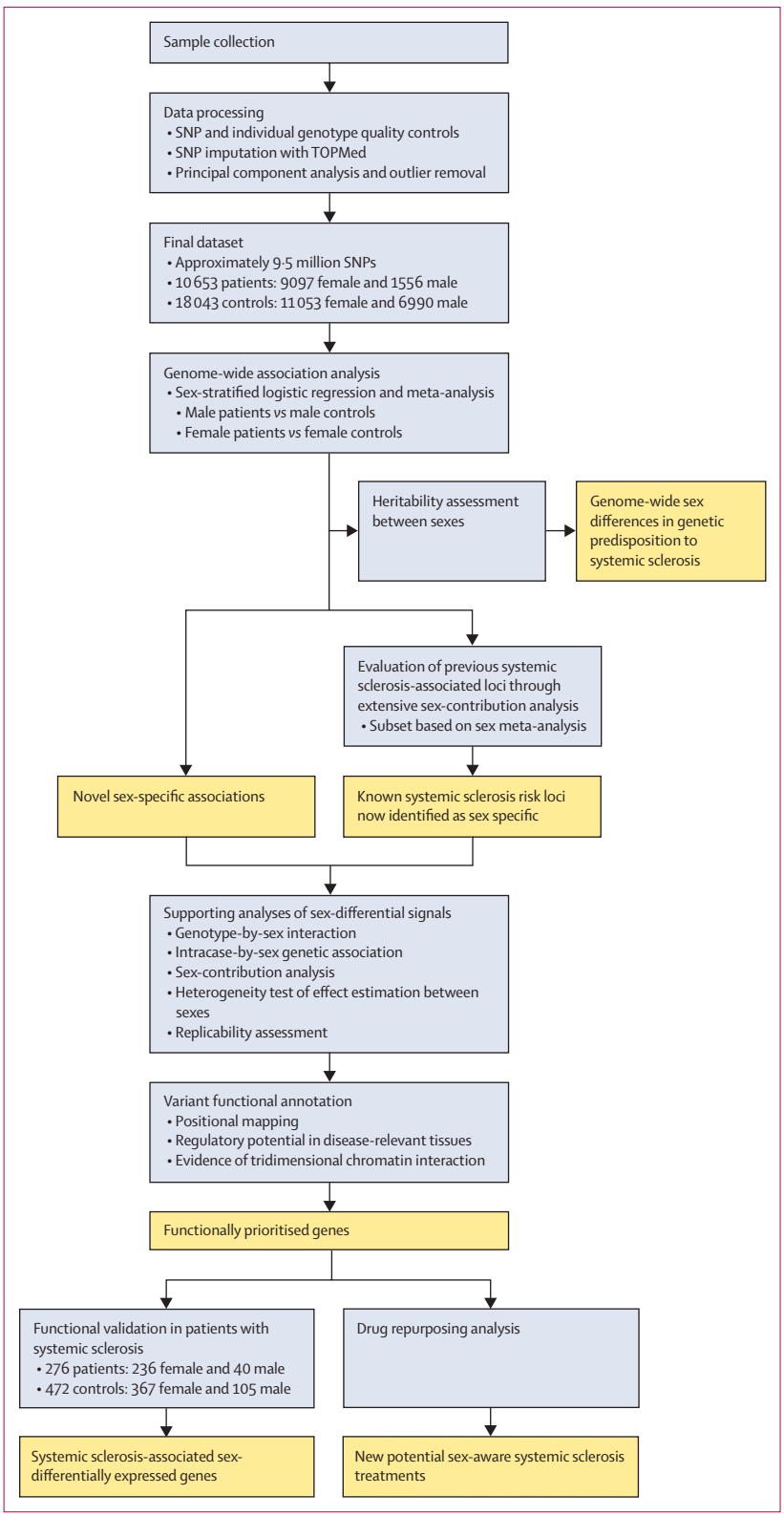


Figure 1: **Overview of the study design**
Main results are shown in yellow. SNP=single nucleotide polymorphism.

	Chromosome	Base pair (hg38)	Nearest gene	Effect allele	Male participants					Female participants				
					p value	OR (95% CI)	Cochran's Q p value	I^2	PPR*	p value	OR (95% CI)	Cochran's Q p value	I^2	PPR*
rs6662198	1	67340524	*IL12RB2*	C	0·64	0·98 (0·88–1·08)	0·18	24·56	0·0016	$8·52×10^{-13}$	0·83 (0·79–0·87)	0·85	0·00	~1·0000
rs2540100	2	59523629	*BCL11A*	A	$8·39×10^{-9}$	1·32 (1·20–1·45)	0·87	0·00	~1·0000	0·82	0·99 (0·95–1·04)	0·38	6·18	0·0003
rs185407974	3	58332737	*PXK*	C	0·086	1·17 (0·98–1·39)	0·77	0·00	0·0134	$1·40×10^{-21}$	1·50 (1·38–1·63)	0·39	5·35	1·0000
rs7349554	3	119384733	*ARHGAP31*	T	0·23	0·93 (0·83–1·05)	0·86	0·00	0·0042	$1·11×10^{-12}$	0·81 (0·76–0·86)	0·051	40·71	~1·0000
rs12203592	6	396321	*IRF4*	T	0·94	1·01 (0·85–1·19)	0·67	0·00	0·0079	$2·36×10^{-8}$	0·78 (0·71–0·85)	0·019	49·24	0·9965
rs2290572	15	74838232	*ULK3*	G	0·39	0·96 (0·87–1·05)	0·49	0·00	0·0023	$7·02×10^{-12}$	0·85 (0·81–0·89)	0·87	0·00	~1·0000
rs34912238	16	85968297	*IRF8*	T	0·41	0·95 (0·83–1·08)	0·90	0·00	0·0042	$1·94×10^{-11}$	0·79 (0·74–0·85)	0·49	0·00	~1·0000
rs66578187	22	39328807	*RPL3-PDGFB*	AT	0·45	1·05 (0·93–1·18)	0·64	0·00	0·0018	$3·18×10^{-8}$	1·19 (1·12–1·26)	0·53	0·00	0·9997

Statistical data for the sex-specific meta-analysis are shown. OR=odds ratio. PPR=posterior probability of replicability. *PPR estimated with the Meta-Analysis Model-based Assessment of Replicability.

Table 1: **Summary statistics for single nucleotide polymorphisms with sex-specific association in systemic sclerosis**

	Nearest gene	Genotype-by-sex interaction		Intracase by sex		p value for heterogeneity between sexes*	Sex contributing†
		p value	OR (95% CI)	p value	OR (95% CI)		
rs6662198	*IL12RB2*	0·0055	0·85 (0·76–0·95)	0·011	0·89 (0·81–0·97)	0·0051	Female
rs2540100	*BCL11A*	$2·14×10^{-7}$	0·76 (0·68–0·84)	$5·44×10^{-6}$	0·82 (0·75–0·89)	$1·67×10^{-7}$	Male
rs185407974	*PXK*	0·0068	1·31 (1·08–1·60)	0·0045	1·25 (1·07–1·45)	0·010	Female
rs7349554	*ARHGAP31*	0·011	0·84 (0·74–0·96)	0·043	0·89 (0·80–1·00)	0·029	Female
rs12203592	*IRF4*	0·0047	0·76 (0·63–0·92)	0·022	0·83 (0·71–0·97)	0·0090	Female
rs2290572	*ULK3*	0·0095	0·87 (0·78–0·97)	0·24	0·95 (0·87–1·04)	0·019	Female
rs34912238	*IRF8*	0·051	0·86 (0·74–1·00)	0·082	0·90 (0·79–1·01)	0·021	Female
rs66578187	*RPL3-PDGFB*	0·049	1·15 (1·00–1·31)	0·27	1·07 (0·95–1·19)	0·070	Female

Statistical data for two additional methods for detecting sex-dependent genetic effects are shown. OR=odds ratio. *p value for the Cochran's Q test between sexes. †Best model to explain the association (male only, female only, or both) based on the subset meta-analysis method with ASSET.

Table 2: **Complementary analyses for single nucleotide polymorphisms with sex-specific association in systemic sclerosis**

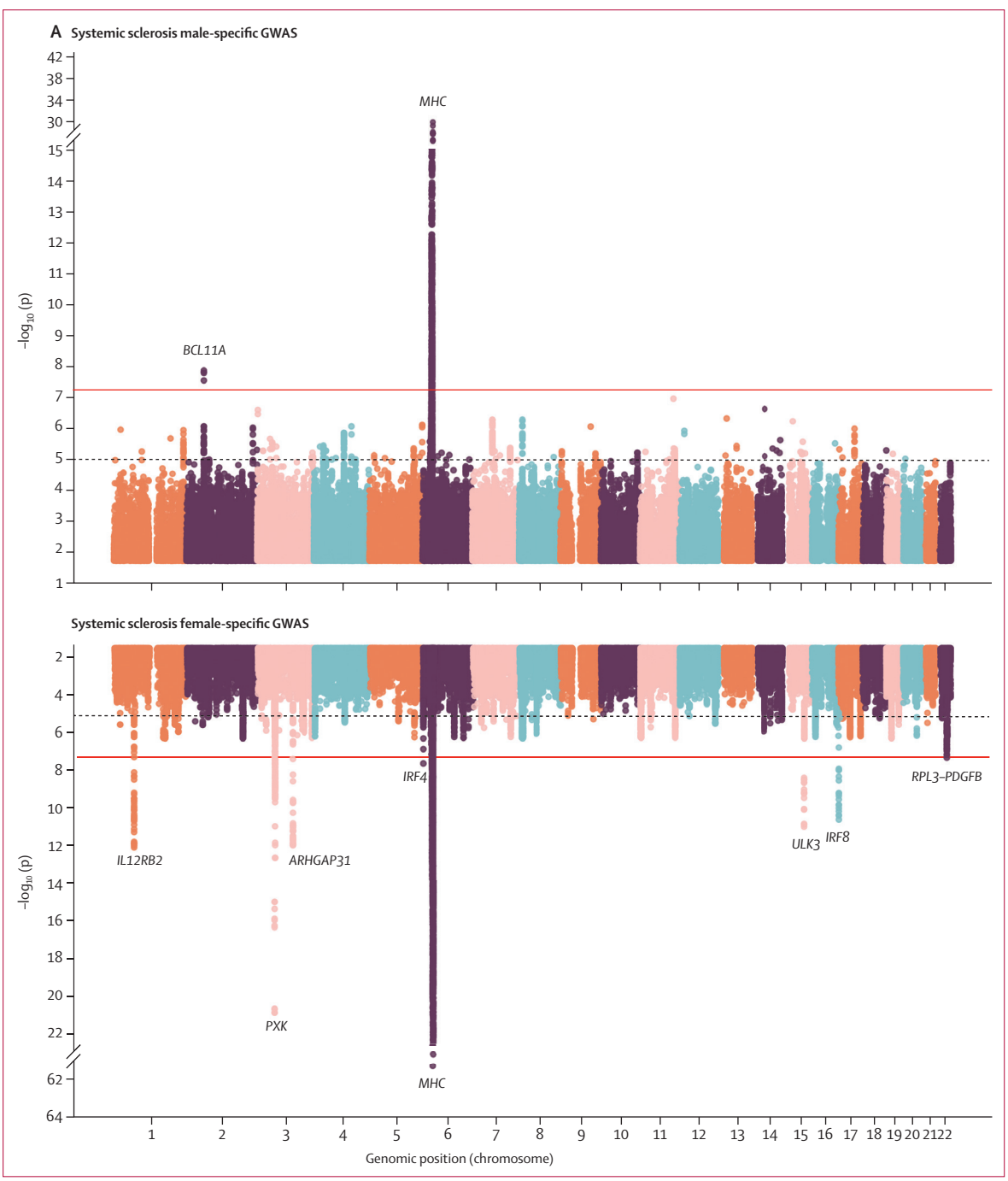
study dataset consisted of 10 653 patients with systemic sclerosis (9097 female patients and 1556 male patients) and 18 043 healthy controls (11 053 female controls and 6990 male controls; appendix p 13). A comprehensive overview of the study design and workflow is provided in figure 1. A total of 26 679 (93·0%) of the study population was included in a previous meta-analysis of GWAS.[6]

Sex-stratified meta-analysis of GWAS analysed 9 541 185 variants in male participants and 9 599 780 in female participants. The absence of systematic bias in the test statistics was confirmed with the genomic inflation factor (λ) in each cohort (appendix p 13). In addition, the linkage disequilibrium score regression (LDSC) intercept was 1·042 (SE 0·008) for male meta-analysis and 1·006 (SE 0·010) for female meta-analysis. Heritability estimates for systemic sclerosis in the liability scale were 2·73% (SE 3·25) for male participants, 7·58% (1·47) for female participants, and 6·32% (1·27) for the combined dataset (sex-agnostic analyses). Consistency across datasets was further confirmed with forest plots to assess effect direction and reproducibility across cohorts (appendix p 8). Minor discrepancies in odds ratios (ORs) in a few cohorts were attributed to differences in allele frequencies, sample sizes, or population structures; however, overall effect direction and magnitude were consistent across most cohorts. The sex-specific results are summarised in tables 1, 2, figure 2, and the appendix (p 9). They were largely consistent between the fixed-effects and the Han and Eskin random-effects models.

We identified one male-specific locus containing four genome-wide significant variants mapping to *BCL11A* (OR 1·32 [95% CI 1·20–1·45]; $p=8·39×10^{-9}$; table 1, figure 2). It represents the first male-specific systemic sclerosis-associated locus and has not previously been linked to the disease. The association was robust across cohorts, with no secondary signals after conditioning on the lead SNP. Replicability was high (PPR>0·99), reinforcing the strength of the finding. Genotype-by-sex interaction and intracase-by-sex meta-analyses showed suggestive associations at this locus (table 2), and significant heterogeneity between male and

female effect sizes was observed (Cochran's *Q* test $p=1{\cdot}67\times10^{-7}$; table 2). The male-only model best explained the signal, indicating that including female participants did not contribute to the association. In contrast, no association or replicability was detected in female participants for this locus despite their larger sample size (OR 0·99 [95% CI 0·95–1·04]; p=0·82; PPR=0·0003; table 1).

In the female-stratified meta-analysis, 1069 variants outside the MHC region reached genome-wide significance. Two novel loci were identified: one within an intronic region of *IRF4* (chromosome 6) and another near *RPL3*, with *PDGFB* located nearby (chromosome 22; figure 2). Both signals were consistent across cohorts, showed no secondary associations after conditioning, and had high replicability (PPR>0·99; table 1). Each



(Figure 2 continues on next page)

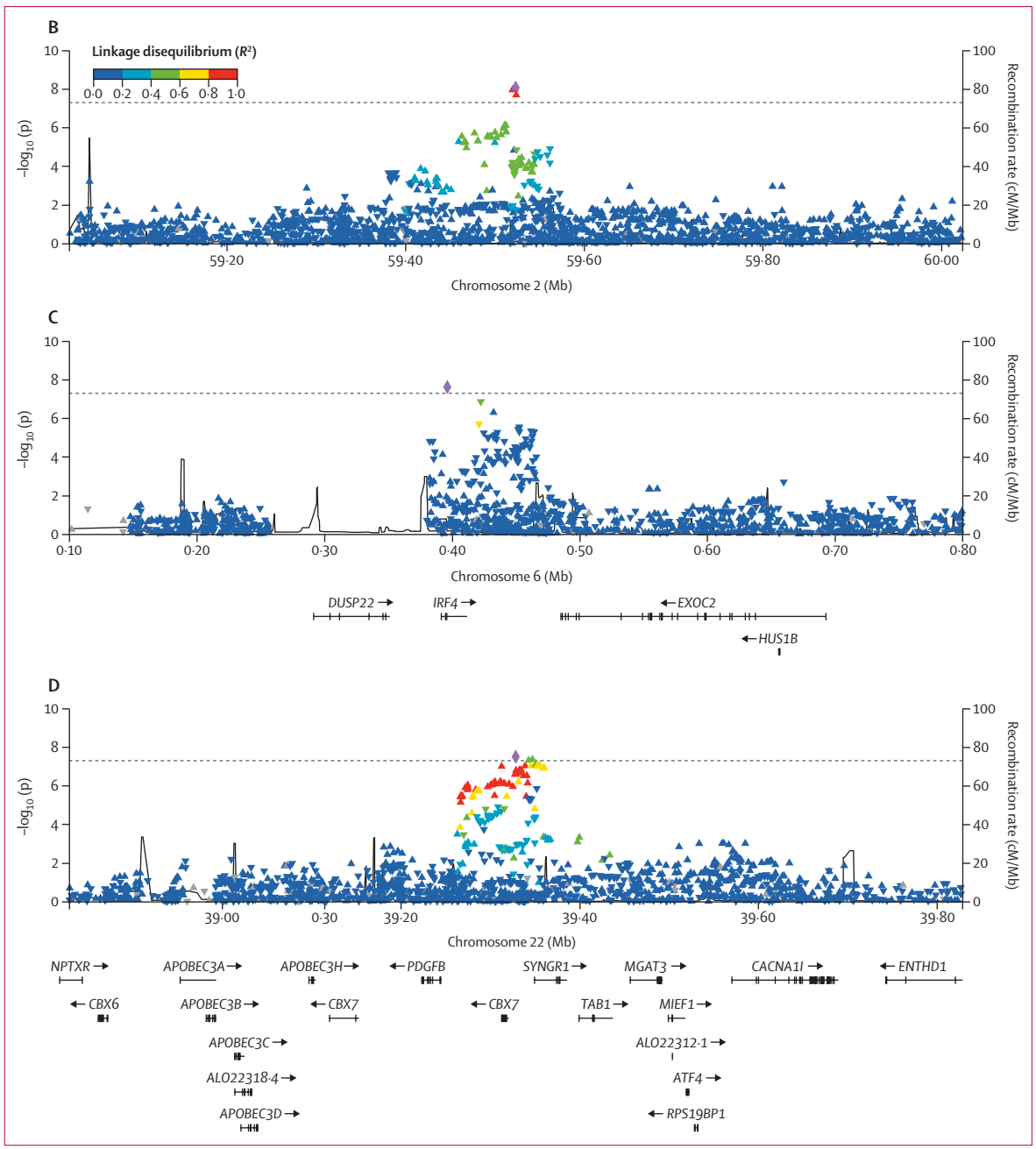


Figure 2: **New sex-specific associations for systemic sclerosis**
(A) Miami plot showing genome-wide association results stratified by sex. Only significant variants meeting the study-defined criteria for sex-specific associations are plotted. The statistical significance of the variants is plotted as $-\log_{10}$ of the p values against their chromosomal position. The red line indicates the threshold for genome-wide significance ($p<5\times10^{-8}$) and the black dashed line represents the suggestive threshold ($p<1\times10^{-5}$). (B) LocusZoom of chromosome 2 association from the male-specific meta-analysis mapping to *BCL11A*. (C) LocusZoom of chromosome 6 association from the female-specific meta-analysis mapping to *IRF4*. (D) LocusZoom of the association signal in chromosome 22 from the female-specific meta-analysis mapping to *RPL3-PGDFB*. (B–D) The lead SNP is represented as a violet diamond. Colour graduation indicates linkage disequilibrium in the European population between the lead SNP and each variant. The genes near the associated variants are depicted in the bottom of the plot. The grey dashed lines indicate the threshold for genome-wide significance ($p<5\times10^{-8}$). GWAS=genome-wide association study. SNP=single nucleotide polymorphism.

For more on **LocusZoom** see http://locuszoom.org/

displayed nominal significance in the genotype-by-sex interaction meta-analysis, with the female-only model providing the best fit (table 2). The *IRF4* variant also showed significant heterogeneity in effect sizes between sexes (p=0·0009; table 2). Neither locus was associated in male participants (both p>0·05 and PPR <0·01; table 1).

Reassessment of previously known systemic sclerosis-risk loci identified five as female specific: *IL12RB2, PXK,*

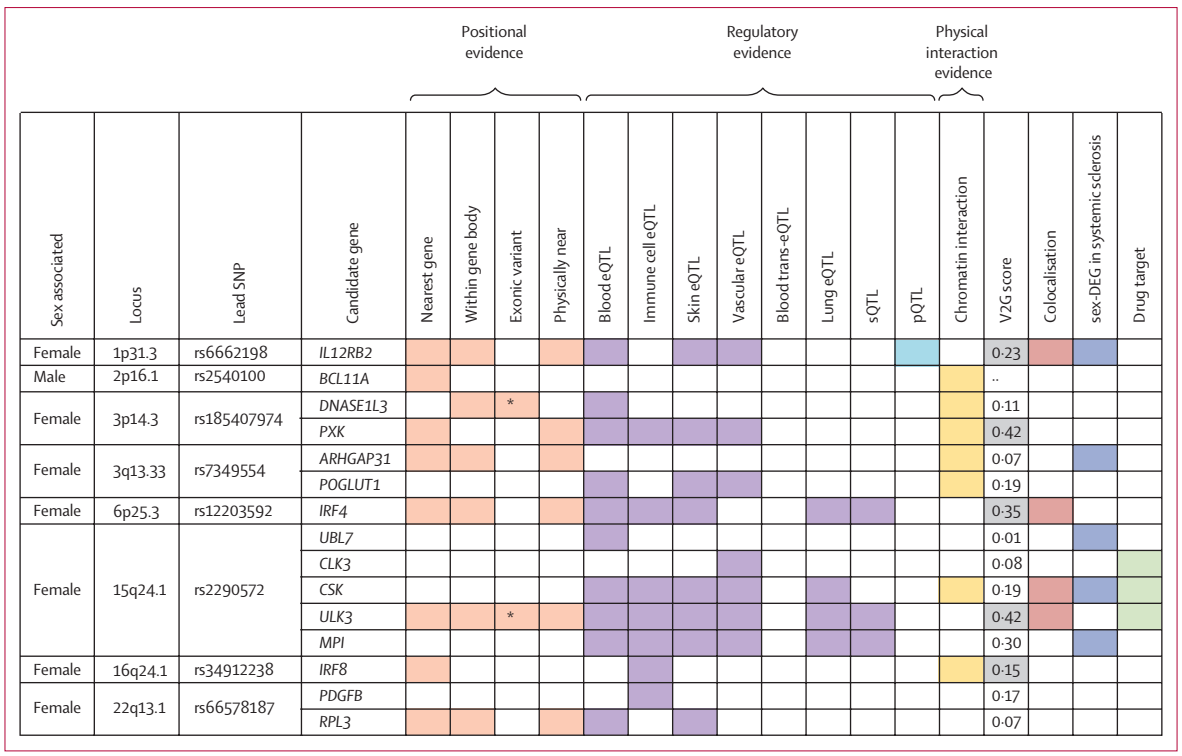


Figure 3: **Summary of candidate genes and functional evidence for sex-specific systemic sclerosis loci**
Shading denotes overlap between the genetic variants or genes and the functional annotation considered. Each colour indicates the layer of information: positional evidence of the variant within the gene (orange), eQTL or sQTL (purple), pQTL (light blue), and chromatin interaction (yellow). The V2G score for each gene is indicated in relation to the lead SNP. The greatest V2G for each locus is highlighted in grey. Colocalisation (red) is annotated when a high posterior probability (>80%) was observed for the associated sex in relevant tissues for the disease. The sex-DEG (dark blue) in systemic sclerosis column indicates whether the gene was differentially expressed in systemic sclerosis by sex (false discovery rate <0·05) in whole blood analysis. The last column shows which genes were targets for approved drugs in our drug repurposing analysis (green). DEG=differentially expressed gene. eQTL=expression quantitative trait loci. pQTL=protein quantitative trait loci. SNP=single nucleotide polymorphism. sQTL=splicing quantitative trait loci. V2G=variant-to-gene. *The lead SNP or a proxy variant has non-synonymous effects for the gene.

ARHGAP31, *ULK3*, and *IRF8*, mapped based on physical distance (table 1; figure 2A; appendix p 9). Conditional analysis of *IRF8* revealed a secondary association that failed to meet the sex-specific criteria and was excluded from subsequent analyses (data not shown). No secondary associations were observed in the remaining four loci. These regions showed nominal significance in two complementary analyses (table 2), were highly replicable in female participants, and were not associated nor replicable in male participants, supporting a sex-differentiated genetic susceptibility pattern in systemic sclerosis.

We identified a total of 39 candidate genes showing at least positional or regulatory evidence of being affected by these sex-specific loci according to FUMA and Open Targets Genetics (appendix p 10) without accounting for sex differences (appendix pp 5–6). All the candidate genes showing the strongest evidence for being causal are summarised in figure 3. We prioritised *BCL11A* as a candidate gene for the male-specific association for being the nearest gene and for showing evidence of chromatin interaction with the associated region in IMR90, a fibroblast cell line derived from lung tissue. For female-specific risk loci, we identified that two of the associated variants were non-synonymous, rs35677470-*DNASE1L3* and rs12898397-*ULK3*. Interestingly, both variants were predicted as deleterious by the SIFT (sorting intolerant from tolerant) score (<0·03) and the *DNASE1L3* SNP was predicted as probably damaging according to PolyPhen-2 score (>0·99). This SNP was identified by Mayes and colleagues[16] and Zochling and colleagues[17] as the candidate causal variant for the association at the *DNASE1L3* locus without accounting for the effect of sex.

After considering regulatory evidence, we prioritised a total of 38 candidate genes for female-specific associations (appendix p 10). Notably, of the 12 genes supported by positional evidence, nine were also regulated by the associated variants in at least one disease-relevant tissue, according to FUMA. This finding is consistent with the recognised utility of positional mapping for identifying functionally relevant genes. Moreover, colocalisation analyses (appendix p 6) showed >80% posterior probability that a single SNP accounts for both genetic and regulatory effects, with sex-specificity identified in seven genes in female

For more on **FUMA** see https://fuma.ctglab.nl/

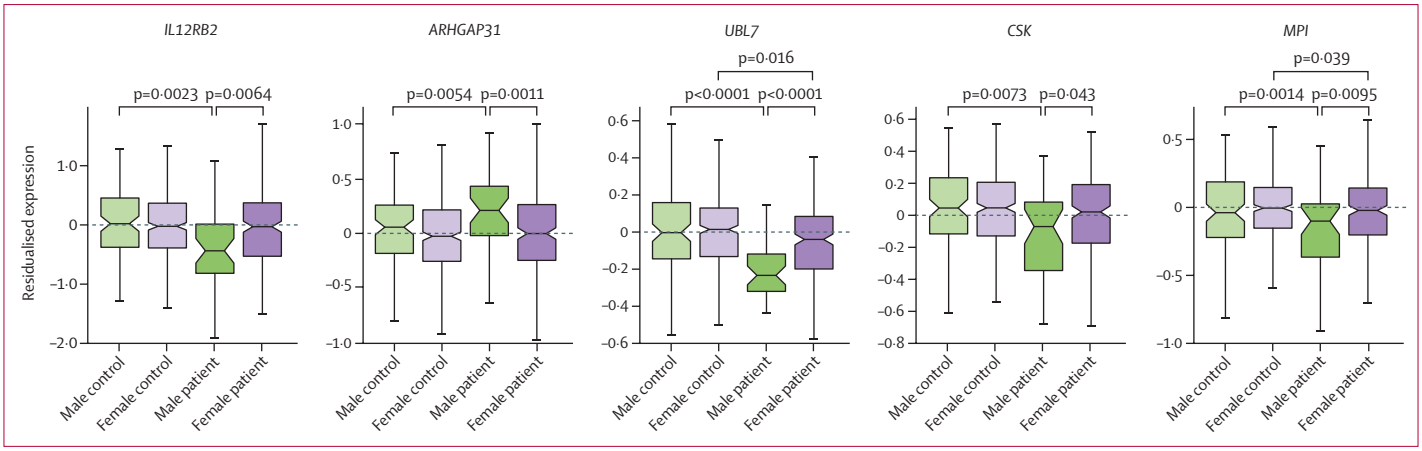


Figure 4: **Boxplots for whole blood sex-differentially expressed genes in systemic sclerosis**
Statistical comparisons between pairs of groups were performed with a *t* test. The central horizontal line in each box indicates the median, and the lower and upper boundaries represent the first and third quartiles, respectively. The whiskers extend to the most extreme datapoints within 1·5-times the IQR from the box boundaries. The groups compared that presented significant differences are shown.

participants, including *IL12RB2*, *IRF4*, *CSK*, and *ULK3* (appendix p 14). Variants mapping to *IL12RB2* influence its expression in blood, skin, and vascular tissue. Remarkably, the lead SNP is a protein quantitative trait locus of IL12RB2 in serum (ie, regulates serum protein concentrations of IL12RB2), establishing this gene as a strong candidate for this female-specific association. Additionally, we identified two other positionally mapped genes, *IRF4* and *ULK3*, for which splicing is affected by the associated variants and that are expression quantitative trait loci in relevant systemic-sclerosis tissues. In addition to the positional candidates, quantitative trait locus analyses identified further genes potentially dysregulated by the associated SNPs (appendix p 10). Notably, variants annotated to *ULK3* were associated with altered expression of *CSK* and *MPI*. All three genes had significant expression quantitative trait locus signals across all the five tissues evaluated for systemic sclerosis, highlighting their putative involvement in female-biased regulatory networks.

Sex-differential gene expression was analysed in a subset of 276 patients with systemic sclerosis (236 female, 40 male) and 472 healthy controls (367 female, 105 male). In these patients, five (13%) of the 39 prioritised genes showed significant sex-by-disease interaction effects in whole blood after correction for multiple testing (*IL12RB2*, *ARHGAP31*, *UBL7*, *CSK*, and *MPI*; figure 4; appendix p 15), highlighting their potential role in sex-biased disease mechanisms. All these genes were upregulated in female patients compared with male patients, except for *ARHGAP31*, which was down-regulated. The robustness of these findings was further confirmed by permutation analyses, with significant empirical p values across the five genes (appendix pp 6–7, 11).

Additionally, we tested for sex-by-disease-by-cell-type interactions (appendix p 6) and identified two significant differentially expressed genes after multiple testing correction: compared with male patients, *IL12RB2* was upregulated in monocytes from female patients and *POGLUT1* was downregulated in plasma B cells from female patients (appendix p 15).

After filtering drugs targeting proteins encoded by the functionally prioritised genes for those with clinical trial evidence in immune-mediated inflammatory diseases, two candidate drugs were identified for further investigation: fostamatinib and dasatinib (appendix p 16). These identifications were based on the targeting of the female-specific prioritised genes *CLK3*, *CSK*, and *ULK3*.

Discussion

This study presents a meta-analysis of GWAS of sex-specific genetic effects in systemic sclerosis, using the largest European dataset reported to date. By leveraging a sex-stratified approach, we identified eight loci outside the MHC region associated with systemic sclerosis in a sex-specific manner: three novel loci (one male-specific [*BCL11A*] and two female-specific [*IRF4* and *RPL3-PDGFB*]) and five previously known systemic sclerosis-risk loci highlighted as female-specific. Functional annotation revealed potential regulatory effects at these loci and identified putative causal genes, some of which are involved in fibroblast function and inflammatory pathways. Notably, we functionally validated five of these genes, which showed sex-differentiated gene expression in a subset of patients with systemic sclerosis, reinforcing their role in sex-biased disease mechanisms. These gene expression analyses were conducted independently of the genetic association signals. These findings provide new insights into the genetic basis of the sex differences in systemic sclerosis and underscore the importance of considering the contribution of sex in genetic studies of immune-mediated inflammatory diseases.

Current research on sex-biased rheumatic diseases has primarily focused on sex chromosomes, hormonal influences, and differences in immune cell composition.[5,18] However, autosomal genetic contribution remains largely unexplored,[7] which could be especially informative in systemic sclerosis due to its genetic background[6,19] and marked sex-differences.[3,4] Here, we estimated the sex-stratified heritability in this condition, which was higher in female participants than in both male participants and the combined datasets. These results indicate the presence of sex differences in the genetic architecture of systemic sclerosis, underscoring the need for sex-stratified analyses and showing that pooling sexes could mask sex-dependent associations.

This study identified rs2540100 (*BCL11A*) as a novel male-specific genetic risk factor in systemic sclerosis. Strikingly, this association was observed despite the substantially smaller sample size of male participants and showed no signal in female participants, reinforcing its specificity. *BCL11A* has been related with haematopoietic stem cell function during ageing and affects plasmacytoid dendritic cells count in mice.[20,21] Given the involvement of plasmacytoid dendritic cells in the pathogenesis of systemic sclerosis and its sex-biased functional profile,[5,22] these findings point to *BCL11A* as a contributor to the observed sex differences in the disease. Supporting this hypothesis, we observed a nominal sex-differential expression of *BCL11A* in dendritic cells in systemic sclerosis. Alternatively, this male-specific association might reflect a genetic marker of disease severity because male patients are more likely to develop severe forms of systemic sclerosis than female patients, including interstitial lung disease.[4] Of note, *BCL11A* has also been associated with another pulmonary affection that is more severe in male patients than in female patients: critical COVID-19.[23,24] In light of the pressing need for prognostic biomarkers in systemic sclerosis,[2] further investigation into *BCL11A* could provide insights into sex-biased disease pathogenesis and progression.

Remarkably, we found seven loci specifically associated with female patients. Gene annotation followed standard GWAS practices, starting from the nearest gene and integrating functional evidence. The subsequent discussion of genes was guided by both biological relevance and functional evidence, giving greater weight to the genes showing sex-differential expression in systemic sclerosis and relevant biological function. This fact does not exclude the potential involvement of other nearby genes. Indeed, sex-differential expression revealed sex-dependent dysregulation across the main pathogenic processes in the disease. Specifically, *IL12RB2*, *MPI*, and *CSK* are involved in immune and inflammatory signalling,[25–28] *UBL7* participates in immune and fibrotic processes,[22,29] and *ARHGAP31* is related to the vascular component,[30] altogether framing the interpretation of the female-specific prioritised genes.

Among these associations, the rs2290572, for which the nearest gene was *ULK3*, emerged as a female-specific locus with three nearby genes showing sex-differentiated expression in whole blood (*MPI*, *CSK*, and *UBL7*). This simultaneous dysregulation raises the possibility of shared regulatory influences, such as chromatin remodelling, that might underlie sex-related transcriptional differences in systemic sclerosis. *UBL7* was of particular interest due to its involvement in a positive feedback loop for type I interferon signalling implicating TRIM21, an autoantigen targeted predominantly in female patients with systemic sclerosis.[29,31] In addition, we identified female associations in two interferon regulatory genes, *IRF4* and *IRF8*.[6,32,33] The interferon signalling pathway has been widely implicated in systemic sclerosis and is known to have distinct roles between sexes,[5,22] mainly attributing these differences to hormonal and X chromosome effects.[5,18] However, to the best of our knowledge, these findings provide the first evidence of sex-differential autosomal genetic architecture affecting this pathway in systemic sclerosis.

One of the most compelling findings in this study was the identification of rs6662198 nearby *IL12RB2* as a female-specific genetic association in systemic sclerosis. Functional analyses revealed multiple regulatory effects at this locus in relevant tissues for the disease, including variations on the protein level depending on the allele in plasma samples. Notably, *IL12RB2* showed significant sex-dependent differential expression in whole blood and monocytes after multiple testing correction and was nominally significant in natural killer cells. This gene encodes a subunit of the IL-12 receptor and was associated with systemic sclerosis in the sex-agnostic GWAS.[6] IL-12 binds to its receptor and activates the JAK–STAT pathway, particularly involving STAT4, a well established risk gene for systemic sclerosis.[6,25] This signalling promotes T-helper-1 (Th1) profile,[25] potentially contributing to the proinflammatory component of the disease. Both cell types identified in the gene expression analysis, monocytes and natural killer cells, are responsive to IL-12 signalling,[34,35] with monocytes producing IL-12 and influencing the function of natural killer cells.[34,35] Further supporting its relevance, sex-differential responses to IL-12/IL-23-targeted therapies have been reported in psoriatic arthritis.[36] These findings suggest that genetically mediated sex differences in IL-12 signalling might contribute to the immune dysregulation underlying this sex-biased disease.

In addition, the novel association in rs66578187 near *PDGFB*, an angiogenic and fibrotic factor,[37,38] was identified as a female-specific locus. Of note, nintedanib, a drug targeting its receptor (PDGFR), caused more mild adverse effects, dose reductions, and treatment interruptions in female patients treated for autoimmune disease-related interstitial lung disease, including systemic sclerosis, than in male patients.[39–41] Given this

finding and in light of our sex-differentiated genetic results, we aimed to propose new potential therapeutic targets by a drug repurposing analysis. This approach allowed us to identify two additional drugs as potential treatments in systemic sclerosis considering the contribution of sex: fostamatinib and dasatinib. Fostamatinib, a spleen tyrosine kinase inhibitor, has completed phase 3 clinical trials in rheumatoid arthritis, another immune-mediated inflammatory disease more prevalent in females.[18] Meanwhile, dasatinib, a tyrosine kinase inhibitor, has completed phase 1 and phase 2 clinical trials for lung disease in systemic sclerosis, one of the sex-biased manifestations of the disease.[4] Both drugs did not show sufficient clinical efficacy in these clinical trials, but did show acceptable safety profiles.[42,43] Considering all of these findings, revisiting clinical trial data with a sex-stratified framework could offer a more refined understanding of drug response than currently known, ultimately contributing to more tailored and effective therapeutic strategies for systemic sclerosis.[41]

Although this study provides valuable insights into sex differences in systemic sclerosis, several limitations should be acknowledged. First, the imbalance of sample sizes between male and female participants in our dataset results in substantially greater statistical power in female participants and should be considered when interpreting the results. This disparity reduces the precision of male-specific estimates and limits the ability to distinguish female-specific associations from shared associations. However, several complementary analyses were conducted to help mitigate this limitation. Given the higher prevalence of systemic sclerosis in female individuals than in male individuals, future studies should make deliberate efforts to include larger male cohorts to improve statistical power and refine sex-specific genetic analyses. Second, neither the MHC region nor the X and Y chromosomes were examined in this study because their analysis requires specific methodological approaches. The application of these approaches in future work might help to address this limitation and reveal additional sex-differential loci. Third, cell-type-specific expression was inferred computationally, further research on directly measured cell expression will provide a more comprehensive understanding of sex bias in the disease. Fourth, we were unable to evaluate how genetic variation interacts with environmental factors in systemic sclerosis. Integrative analyses incorporating epigenomic data to explore sex differences, which might reflect environmental influences, could enhance understanding of this sex-biased condition. Finally, no individuals with lived experience of systemic sclerosis were involved in this study. Despite these limitations, this study identified genetic variants and genes with sex-differential associations in systemic sclerosis, offering new insights into its biological basis. We also identified two drugs with genetic evidence for sex-differential effect in the disease, underscoring the need for sex-aware studies, including clinical trials, in systemic sclerosis and other rheumatic diseases.

Contributors

JM and MA-H contributed to the conception and study design. IR-M, MK, CR-P, CR-B, GB-Y, LO-F, and MA-H contributed to quality control and imputation. IR-M, MK, and MA-H contributed to data analysis and interpretation and have directly accessed and verified the underlying data reported in the manuscript. AG-D-C, CPS-A, JLC, OD, SMP, MN, NH, JKdV-B, ALH, YA, MEA-R, LB, SA, CPD, and MDM contributed to control and patient data collection. IR-M, MK, JM, and MA-H contributed to manuscript drafting. All authors participated in the manuscript revision and approval, and have read and agreed to the published version of the manuscript. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Declaration of interests

CPS-A reports consulting fees from Janssen and Boehringer Ingelheim; honoraria or payments from Janssen, MSD, and Boehringer Ingelheim; support for travel or meeting attendance from Janssen, MSD, and Boehringer Ingelheim; and research funds from Boehringer Ingelheim. OD reports research grants from Kymera, Mitsubishi Tanabe, and Boehringer Ingelheim; consulting fees from 4P-Pharma, AbbVie, Acepodia Biotech, Aera Therapeutics, AnaMar, Anaveon, Argenx, AstraZeneca, Boehringer Ingelheim, BMS, Calluna Pharma, Cantargia, CSL Behring, EMD Serono, Galapagos, Galderma, Gossamer, Hemetron, Innovaderm, Kali Therapeutics, Lilly, Janssen, Mediar, MSD Merck, Nkarta, Novartis, Oorja Bio, Orion, Pilan Therapeutics, Prometheus, Quell Therapeutics, Redxpharma, Scleroderma Research Foundation, Sumitomo, Topadur, and UCB related to scleroderma and its complications; speaker honoraria related to scleroderma and its complications from Janssen and Boehringer Ingelheim; a patent relating to systemic sclerosis (US8247389, EP2331143); serving as Chair of the Executive Committee in the FOREUM Foundation and co-Chair President in ERS/EULAR Guidelines EUSTAR; consulting activities related to the Platform Trial in Scleroderma for the Scleroderma Research Foundation; serving as Member Board of Trustees for the Swiss Clinical Quality Management in Rheumatic Diseases and the Hartmann Müller Foundation; serving as a senate member on the Swiss Academy of Medical Sciences; and being co-founder of Citus. SMP reports research support from Janssen and Boehringer Ingelheim to the Australian Scleroderma Interest Group; honoraria from Janssen and Boehringer Ingelheim; and participation on an advisory board for Boehringer Ingelheim and MSD. MN has received an Investigator Grant from the National Health and Medical Research Council of Australia (NHMRC GNT1176538); research grants from Boehringer Ingelheim and Janssen; consulting fees from AstraZeneca and GSK; honoraria for presentations from AstraZeneca, Boehringer Ingelheim, and GSK; and support for conference attendance from Boehringer Ingelheim. JKdV-B reports grants and contracts from Jansen-Cilag (Johnson & Johnson), ARCH, ReumaNederland, and Boehringer Ingelheim, paid to their institution; consulting fees from Boehringer Ingelheim and AbbVie, paid to their institution; payment or honoraria from Boehringer Ingelheim, Jansen-Cilag (Johnson & Johnson), BMS, Pfizer, GSK, Novartis, and USB, paid to their institution; participation on data safety monitoring and advisory board for AbbVie, Boehringer Ingelheim, the Scleroderma Research Foundation, and the TASER trial, paid to their institution; and a leadership or fiduciary role in ARCH (beneficiary). ALH reports consulting fees from AbbVie, Arena, Boehringer Ingelheim, Janssen, Merck, and ZuraBio, paid to their institution; payment or honoraria from Janssen, paid to their institution; and participation on data safety monitoring and advisory board for Novartis, paid to their institution. LB has received payment or honoraria from Johnson & Johnson and AlfaSigma, and reports participation on data safety monitoring and advisory board for AbbVie. SA reports grants or contracts from aTyr, AbbVie, Boehringer Ingelheim, and the Scleroderma Research Foundation, paid to their institution; consulting fees from AbbVie,

AstraZeneca, aTyr, Boehringer Ingelheim, CSL Behring, Merck, Mitsubishi Tanabe, Regeneron, the Scleroderma Research Foundation, and TeneoFour; payment or honoraria from Boehringer Ingelheim, the Academy for Continued Healthcare Learning, and the Institute for Medical Education; and leadership or fiduciary roles in the Scleroderma Clinical Trial Consortium and the National Scleroderma Foundation (unpaid volunteer). CPD has received research grants from AbbVie, GSK, and Servier, paid to their institution; consultancy fees from AbbVie, Janssen, GSK, Boehringer Ingelheim, Roche, Novartis, Certa, Mitsubishi, and Quell; and honoraria from Boehringer Ingelheim and Janssen. MDM reports receiving grants or contracts as clinical trial investigator from Prometheus-Merck, Mitsubishi Tanabe, Boehringer Ingelheim, AstroZenica, ATYR Pharma, Horizon/Amgen Pharma, and BMS Pharma, paid to their insitution; book royalties from Springer International Publishing and Oxford University Press, and monograph royalties from *The British Medical Journal*; consulting fees to advise clinical trial protocol from Cabaletta Pharma; payment or honoraria from GSK Pharma, AstraZenica, and Novartis for workshop to teach the modified Rodnan Skin Score technique; and serves as an advisory board member for Mitsubishi Tanabe, Boehringer Ingelheim, and EICOS. All other authors declare no competing interests.

Data sharing

Summary statistics of the sex-specific meta-analysis of GWAS analysed in this study will be available upon publication through the National Human Genome Research Institute–European Bioinformatics Institute GWAS Catalog (https://www.ebi.ac.uk/gwas/downloads/summary-statistics; please use "systemic sclerosis" as search term). Individual-level data are not publicly available due to privacy concerns and informed consent restrictions. All other data are provided in the main Article and supplementary materials or are available upon reasonable request to the corresponding author.

Acknowledgments

This work was supported by the Instituto de Salud Carlos III (CP21/00132, PI22/00092 [co-financed by the EU], and Redes de Investigación Cooperativa Orientadas a Resultados en Salud [RD21/0002/0039 and RD24/0007/0016]) and the Innovative Medicines Initiative Joint Undertaking (grant 115565; PRECISESADS project; these resources comprise financial contribution from the EU's Seventh Framework Program [FP7/2007–2013] and European Federation of Pharmaceutical Industries and Associations companies' in-kind contribution). IR-M (FPU21/02746) and CR-B (FPU22/01652) were supported by the programme Formación de Profesorado Universitario funded by the Spanish Ministry of Science, Innovation, and Universities (MICIU). CR-P was supported by grant PREP2022-000747 funded by MICIU/AEI/10.13039/501100011033 and by the European Social Fund Plus (ESF+). GB-Y's contract is part of the grant PREP2022-000712 funded by MICIU/AEI/10.13039/501100011033 and by ESF+. LO-F is a recipient of a Ramon y Cajal fellowship (RYC2022-036635-I) funded by MICIU/AEI/10.13039/501100011033 and by ESF+. The research was also supported by the UK National Institute for Health and Care Research (NIHR)Manchester Clinical Research Facility (NIHR203956). The views expressed are our own and not necessarily those of the NIHR or the UK Department of Health and Social Care. We thank Sofia Vargas and Gema Robledo for their exceptional technical expertise. We also thank all of the participants who provided the data for their essential collaboration and all members of the PRECISESADS consortium. During the preparation of this work we used ChatGPT (GPT 4) to enhance the writing of some sections of the manuscript. After using this tool, we reviewed and edited the content as needed and take full responsibility for the content of the published article. This research is part of the doctoral degree awarded to IR-M within the Biomedicine programme from the University of Granada.

References

1 Volkmann ER, Andréasson K, Smith V. Systemic sclerosis. *Lancet* 2023; **401:** 304–18.
2 Abraham DJ, Black CM, Denton CP, et al. An international perspective on the future of systemic sclerosis research. *Nat Rev Rheumatol* 2025; **21:** 174–87.
3 De Angelis R, Giuggioli D, Bajocchi G, et al, and the Systemic sclerosis PRogression INvestiGation group (SPRING). Sex-related differences in systemic sclerosis: a multicenter cross-sectional study from the National Registry of the Italian Society for Rheumatology. *J Rheumatol* 2022; **49:** 176–85.
4 Hughes M, Pauling JD, Armstrong-James L, Denton CP, Galdas P, Flurey C. Gender-related differences in systemic sclerosis. *Autoimmun Rev* 2020; **19:** 102494.
5 Jiwrajka N, Anguera MC. The X in sex-biased immunity and autoimmune rheumatic disease. *J Exp Med* 2022; **219:** e20211487.
6 López-Isac E, Acosta-Herrera M, Kerick M, et al, and the European Scleroderma Group, and the Australian Scleroderma Interest Group (ASIG). GWAS for systemic sclerosis identifies multiple risk loci and highlights fibrotic and vasculopathy pathways. *Nat Commun* 2019; **10:** 4955.
7 Khramtsova EA, Wilson MA, Martin J, et al. Quality control and analytic best practices for testing genetic models of sex differences in large populations. *Cell* 2023; **186:** 2044–61.
8 Tian Y, Chen S, Yu B, et al. Identification of biomarkers and immune infiltration associated with sexes in SSc: a bioinformatics and machine learning. *Rheumatology* 2025; **64:** 3806–15.
9 Vreca M, Andjelkovic M, Tosic N, et al. Impact of alterations in X-linked IRAK1gene and miR-146a on susceptibility and clinical manifestations in patients with systemic sclerosis. *Immunol Lett* 2018; **204:** 1–8.
10 Kerick M, Acosta-Herrera M, Simeón-Aznar CP, et al, and the International SSc Group, the Australian Scleroderma Interest Group (ASIG), and the PRECISESADS Clinical Consortium. Complement component C4 structural variation and quantitative traits contribute to sex-biased vulnerability in systemic sclerosis. *NPJ Genom Med* 2022; **7:** 57.
11 Sakkas LI, Bogdanos DP, Chikanza IC. Sex bias in systemic sclerosis: from clinical to immunological differences. *Clin Rev Allergy Immunol* 2025; **68:** 51.
12 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.
13 LeRoy EC, Medsger TA Jr. Criteria for the classification of early systemic sclerosis. *J Rheumatol* 2001; **28:** 1573–76.
14 Heidari S, Babor TF, De Castro P, Tort S, Curno M. Sex and gender equity in research: rationale for the SAGER guidelines and recommended use. *Res Integr Peer Rev* 2016; **1:** 2.
15 Beretta L, Barturen G, Vigone B, et al, and the PRECISESADS SSc substudy group, and the PRECISESADS Flow Cytometry study group. Genome-wide whole blood transcriptome profiling in a large European cohort of systemic sclerosis patients. *Ann Rheum Dis* 2020; **79:** 1218–26.
16 Mayes MD, Bossini-Castillo L, Gorlova O, et al, and the Spanish Scleroderma Group. Immunochip analysis identifies multiple susceptibility loci for systemic sclerosis. *Am J Hum Genet* 2014; **94:** 47–61.
17 Zochling J, Newell F, Charlesworth JC, et al. An Immunochip-based interrogation of scleroderma susceptibility variants identifies a novel association at DNASE1L3. *Arthritis Res Ther* 2014; **16:** 438.
18 Fairweather D, Beetler DJ, McCabe EJ, Lieberman SM. Mechanisms underlying sex differences in autoimmunity. *J Clin Invest* 2024; **134:** e180076.
19 Ishikawa Y, Tanaka N, Asano Y, et al. GWAS for systemic sclerosis identifies six novel susceptibility loci including one in the Fcγ receptor region. *Nat Commun* 2024; **15:** 319.
20 Wu X, Satpathy AT, Kc W, Liu P, Murphy TL, Murphy KM. *Bcl11a* controls Flt3 expression in early hematopoietic progenitors and is required for pDC development in vivo. *PLoS One* 2013; **8:** e64800.
21 Wang J, Zhang L, Cui X, et al. *Bcl11a* maintains hematopoietic stem cell function but accelerates inflammation-driven exhaustion during aging. *Sci Immunol* 2025; **10:** eadr2041.
22 Fang D, Chen B, Lescoat A, Khanna D, Mu R. Immune cell dysregulation as a mediator of fibrosis in systemic sclerosis. *Nat Rev Rheumatol* 2022; **18:** 683–93.

23 Kousathanas A, Pairo-Castineira E, Rawlik K, et al, and the GenOMICC investigators, the 23andMe investigators, and the COVID-19 Human Genetics Initiative. Whole-genome sequencing reveals host factors underlying critical COVID-19. *Nature* 2022; **607:** 97–103.

24 Zhang Q, Bastard P, Liu Z, et al, and the COVID-STORM Clinicians, the COVID Clinicians, the Imagine COVID Group, the French COVID Cohort Study Group, the CoV-Contact Cohort, the Amsterdam UMC Covid-19 Biobank, the COVID Human Genetic Effort, and the NIAID-USUHS/TAGC COVID Immunity Group. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science* 2020; **370:** eabd4570.

25 Tait Wojno ED, Hunter CA, Stumhofer JS. The immunobiology of the interleukin-12 family: room for discovery. *Immunity* 2019; **50:** 851–70.

26 Liang R, Ye Z-W, Qin Z, et al. PMI-controlled mannose metabolism and glycosylation determines tissue tolerance and virus fitness. *Nat Commun* 2024; **15:** 2144.

27 Inoue T, Matsumoto Y, Kawai C, et al. Csk restrains BCR-mediated ROS production and contributes to germinal center selection and affinity maturation. *J Exp Med* 2024; **221:** e20231996.

28 Hui X, Xue M, Ren Y, et al. *GPR132* regulates the function of NK cells through the Gαs/CSK/ZAP70/NF-κB signaling pathway as a potential immune checkpoint. *Sci Adv* 2025; **11:** eadr9395.

29 Jiang W, Li X, Xu H, et al. UBL7 enhances antiviral innate immunity by promoting Lys27-linked polyubiquitination of MAVS. *Cell Rep* 2023; **42:** 112272.

30 Caron C, DeGeer J, Fournier P, et al. CdGAP/ARHGAP31, a Cdc42/Rac1 GTPase regulator, is critical for vascular development and VEGF-mediated angiogenesis. *Sci Rep* 2016; **6:** 27485.

31 Martel M-E, Leurs A, Launay D, et al. Prevalence of anti-Ro52-kDa/SSA (TRIM21) antibodies and associated clinical phenotype in systemic sclerosis: data from a French cohort, a systematic review and meta-analysis. *Autoimmun Rev* 2024; **23:** 103536.

32 López-Isac E, Martín J-E, Assassi S, et al, and the Spanish Scleroderma Group. 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. *Arthritis Rheumatol* 2016; **68:** 2338–44.

33 Gorlova OY, Li Y, Gorlov I, et al. Gene-level association analysis of systemic sclerosis: a comparison of African-Americans and White populations. *PLoS One* 2018; **13:** e0189498.

34 Guan F, Wang R, Yi Z, et al. Tissue macrophages: origin, heterogenity, biological functions, diseases and therapeutic targets. *Signal Transduct Target Ther* 2025; **10:** 93.

35 Coma G, Peña R, Blanco J, et al. Treatment of monocytes with interleukin (IL)-12 plus IL-18 stimulates survival, differentiation and the production of CXC chemokine ligands (CXCL)8, CXCL9 and CXCL10. *Clin Exp Immunol* 2006; **145:** 535–44.

36 Eder L, Mylvaganam S, Pardo Pardo J, et al. Sex-related differences in patient characteristics, and efficacy and safety of advanced therapies in randomised clinical trials in psoriatic arthritis: a systematic literature review and meta-analysis. *Lancet Rheumatol* 2023; **5:** e716–27.

37 Liu Z-L, Chen H-H, Zheng L-L, Sun L-P, Shi L. Angiogenic signaling pathways and anti-angiogenic therapy for cancer. *Signal Transduct Target Ther* 2023; **8:** 198.

38 Yao L, Rathnakar BH, Kwon HR, et al. Temporal control of PDGFRα regulates the fibroblast-to-myofibroblast transition in wound healing. *Cell Rep* 2022; **40:** 111192.

39 Distler O, Highland KB, Gahlemann M, et al, and the SENSCIS Trial Investigators. Nintedanib for systemic sclerosis-associated interstitial lung disease. *N Engl J Med* 2019; **380:** 2518–28.

40 Wollin L, Wex E, Pautsch A, et al. Mode of action of nintedanib in the treatment of idiopathic pulmonary fibrosis. *Eur Respir J* 2015; **45:** 1434–45.

41 Volkmann ER, Feghali-Bostwick C. Sex- and gender-based personalized medicine in rheumatology. *Nat Rev Rheumatol* 2025; **21:** 251–52.

42 Weinblatt ME, Genovese MC, Ho M, et al. Effects of fostamatinib, an oral spleen tyrosine kinase inhibitor, in rheumatoid arthritis patients with an inadequate response to methotrexate: results from a phase III, multicenter, randomized, double-blind, placebo-controlled, parallel-group study. *Arthritis Rheumatol* 2014; **66:** 3255–64.

43 Martyanov V, Kim GJ, Hayes W, et al. Novel lung imaging biomarkers and skin gene expression subsetting in dasatinib treatment of systemic sclerosis-associated interstitial lung disease. *PLoS One* 2017; **12:** e0187580.