



OPEN ACCESS

EDITED BY

Ying Zhou,
Tongji University, China

REVIEWED BY

Humberto Garcia-Ortiz,
National Institute of Genomic Medicine
(INMEGEN), Mexico
Ilias C. Papanikolaou,
General Hospital of Corfu, Greece

*CORRESPONDENCE

Natalia V. Rivera
✉ natalia.rivera@ki.se

RECEIVED 27 December 2022

ACCEPTED 10 April 2023

PUBLISHED 11 May 2023

CITATION

Xiong Y, Kullberg S, Garman L, Pezant N,
Ellinghaus D, Vasila V, Eklund A, Rybicki BA,
Iannuzzi MC, Schreiber S,
Müller-Quernheim J, Montgomery CG,
Grunewald J, Padyukov L and Rivera NV
(2023) Sex differences in the genetics
of sarcoidosis across European and African
ancestry populations.
Front. Med. 10:1132799.
doi: 10.3389/fmed.2023.1132799

COPYRIGHT

© 2023 Xiong, Kullberg, Garman, Pezant,
Ellinghaus, Vasila, Eklund, Rybicki, Iannuzzi,
Schreiber, Müller-Quernheim, Montgomery,
Grunewald, Padyukov and Rivera. This is an
open-access article distributed under the terms
of the [Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted which
does not comply with these terms.

Sex differences in the genetics of sarcoidosis across European and African ancestry populations

Ying Xiong¹, Susanna Kullberg^{1,2}, Lori Garman³, Nathan Pezant³,
David Ellinghaus⁴, Vasiliki Vasila¹, Anders Eklund²,
Benjamin A. Rybicki⁵, Michael C. Iannuzzi⁶, Stefan Schreiber^{4,7},
Joachim Müller-Quernheim⁸, Courtney G. Montgomery³,
Johan Grunewald^{1,2,9}, Leonid Padyukov^{9,10} and
Natalia V. Rivera^{1,9,10*}

¹Respiratory Medicine Division, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden, ²Department of Respiratory Medicine and Allergy, Theme Inflammation and Ageing, Karolinska University Hospital, Stockholm, Sweden, ³Genes and Human Disease, Oklahoma Medical Research Foundation, Oklahoma City, OK, United States, ⁴Institute of Clinical Molecular Biology, Christian-Albrechts-University of Kiel, Kiel, Germany, ⁵Department of Public Health Sciences, Henry Ford Health System, Detroit, MI, United States, ⁶Zucker School of Medicine, Staten Island University Hospital, Northwell/Hofstra University, Staten Island, NY, United States, ⁷Clinic for Internal Medicine I, University Hospital Schleswig-Holstein, Kiel, Germany, ⁸Department of Pneumology, University Medical Center Freiburg, Faculty of Medicine, University of Freiburg, Freiburg im Breisgau, Germany, ⁹Center for Molecular Medicine, Karolinska Institutet, Stockholm, Sweden, ¹⁰Division of Rheumatology, Department of Medicine Solna, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden

Background: Sex differences in the susceptibility of sarcoidosis are unknown. The study aims to identify sex-dependent genetic variations in two clinical sarcoidosis phenotypes: Löfgren's syndrome (LS) and non-Löfgren's syndrome (non-LS).

Methods: A meta-analysis of genome-wide association studies was conducted on Europeans and African Americans, totaling 10,103 individuals from three population-based cohorts, Sweden ($n = 3,843$), Germany ($n = 3,342$), and the United States ($n = 2,918$), followed by an SNP lookup in the UK Biobank (UKB, $n = 387,945$). A genome-wide association study based on ImmunoChip data consisting of 141,000 single nucleotide polymorphisms (SNPs) was conducted in the sex groups. The association test was based on logistic regression using the additive model in LS and non-LS sex groups independently. Additionally, gene-based analysis, gene expression, expression quantitative trait loci (eQTL) mapping, and pathway analysis were performed to discover functionally relevant mechanisms related to sarcoidosis and biological sex.

Results: We identified sex-dependent genetic variations in LS and non-LS sex groups. Genetic findings in LS sex groups were explicitly located in the extended Major Histocompatibility Complex (xMHC). In non-LS, genetic differences in the sex groups were primarily located in the MHC class II subregion and *ANXA11*. Gene-based analysis and eQTL enrichment revealed distinct sex-specific gene expression patterns in various tissues and immune cell types. In LS sex groups, a pathway map related to antigen presentation machinery by IFN-gamma. In non-LS, pathway maps related to immune response lectin-induced complement pathway in males and related to maturation and migration of dendritic cells in skin sensitization in females were identified.

Conclusion: Our findings provide new evidence for a sex bias underlying sarcoidosis genetic architecture, particularly in clinical phenotypes LS and non-LS. Biological sex likely plays a role in disease mechanisms in sarcoidosis.

KEYWORDS

sarcoidosis, genetics, genome-wide association study (GWAS), meta-analysis, single nucleotide polymorphisms, immunogenetics and HLA

Background

Sarcoidosis is a multi-system disease with unknown etiology, characterized by non-caseating granulomas (1). In 90% of cases, sarcoidosis affects the lungs and lymphatic system (2).

Sarcoidosis is prevalent worldwide, commonly affecting individuals between 20 and 50 years of age, with males diagnosed earlier than females (3, 4). On a global scale, sarcoidosis incidence and prevalence vary based on sex, age, geography, and ethnicity (4). In the United States, the prevalence is 141 per 100,000 in African-Americans and 49.8 per 100,000 in Caucasians (5). In Nordic countries such as Sweden, the prevalence of sarcoidosis was reported between 152 and 215 per 100,000 (6), whereas in Denmark was reported at 77 per 100,000 (7).

The disease is heterogeneous, as different manifestations and clinical outcomes have been observed in clinical settings, particularly among sex groups (3, 8, 9). Interestingly, sarcoidosis is more common in females, has a late-onset, has a variable disease course depending on the affected organ, and has a higher mortality rate than men (3). Sex differences are also seen in extra-pulmonary phenotypes (10, 11), including cardiac (12, 13) and skin sarcoidosis.

Sex hormones have been shown to play a role in sarcoidosis (14, 15). For instance, sarcoidosis has a low disease activity in pregnancy or goes into remission. However, the ameliorating effect is lost as flares occur after delivery (16, 17).

In the lungs, sarcoidosis is characterized by Löfgren's syndrome (LS) and non-Löfgren's syndrome (NLS). LS is an acute form of the disease characterized by periarticular swelling around the ankles, erythema nodosum, and bilateral hilar lymphadenopathy (18). Typically, females with LS exhibit erythema nodosum (EN), while males with LS have symptoms of bilateral ankle arthritis. Patients with LS usually have a good prognosis (19). LS often occurs in individuals of European ancestry and is relatively rare in individuals of African origin. In NLS, sarcoidosis is characterized by a heterogeneous disease course, often with an insidious onset, unrelenting disease course, and a high risk for clinical organ impairment, predominantly pulmonary fibrosis (18, 20). Patients with fibrotic sarcoidosis have markedly decreased survival compared with the general population (21, 22).

The causes for clinical differences and disease course in sarcoidosis among sex groups are unknown. However, evidence-based factors, including genetics, epigenetics, and environmental exposures, have been postulated. Regarding genetics, few studies have addressed sex differences in the genetic variation in disease, despite most studies adjusting for sex in their analyses (23). An exception is a genome-wide admixture scan conducted in African Americans stratified by sex, identifying several genetic variants associated with sarcoidosis exclusively in females (24).

In this work, we sought to investigate the genetic associations in the sex groups and identify differences and commonalities in European and African population ancestries by a meta-analysis approach and high-density mapping.

Materials and methods

Samples and study design

The samples were from three independent sarcoidosis population-based cohorts from Sweden, Germany, and the United States. All participants provided written informed consent for the study. Study protocols of all studies had been approved by respective local institutional boards.

The Swedish cohort

The case-control study consisted of 4,133 individuals. The study was approved by the local institutional review board in Stockholm, Sweden. All participants provided written informed consent and were permitted to use their DNA for research purposes. Baseline clinical and demographic characteristics were measured when the participants were enrolled in the study.

Patients were enrolled at the time of disease investigation at the Sarcoidosis Centrum, Karolinska University Hospital Solna, Sweden. The diagnosis was established on radiographic manifestations, findings at bronchoscopy with bronchoalveolar lavage (BAL), including an elevated CD4/CD8 ratio > 3.5, and positive biopsies and in accordance with the criteria outlined by the World Association of Sarcoidosis and Other Granulomatous Disorders [WASOG, Statement on sarcoidosis (27)]. Sarcoidosis patients were further characterized into two clinical phenotypes, LS and NLS. LS was defined by typical clinical manifestations with an acute onset of the disease, including fever, bilateral hilar lymphadenopathy on chest X-ray, bilateral ankle arthritis, and/or erythema nodosum. NLS was defined as a heterogeneous disease course with an insidious onset, unrelenting disease course, and a

Abbreviations: LS, Löfgren's syndrome; NLS, non-Löfgren's syndrome; SNPs, single nucleotide polymorphisms; eQTL, expression quantitative trait loci; UK Biobank, United Kingdom Biobank; *ANXA11*, Annexin A11 gene; GWAS, Genome-wide association study; US-AA, United States African Americans; PC, Principal component; PCA, principal component analysis; IVW, inverse variance weighting; OR, odds ratio; CI, confidence interval.

high risk for clinical organ impairment, such as developing fibrosis in the lung. Healthy controls included 3,085 individuals who were recruited from two large-scale epidemiological cohorts. Mainly, 2,025 individuals were from the Environmental Investigation of Rheumatoid arthritis (EIRA) study described in Klareskog et al. (25) and 1,060 individuals were from the Epidemiological investigation of risk factors for Multiple Sclerosis (EIMS) study described in Hedstrom et al. (26).

The German cohort

The case-control study consisted of 4,975 individuals, of which 413 were non-LS (27). The description and inclusion of these patients are described elsewhere (28, 29). Briefly, German control subjects ($n = 4,498$) were derived from Popgen ($n = 2,485$) (30), and the Heinz Nixdorf RECALL (HNR) study ($n = 1,499$) (31). Additionally, 304 control individuals of South German origin were recruited from the Bavarian Red Cross, and 210 control individuals were recruited from the Charité - Universitätsmedizin Berlin. The mean age and male percentage were 62.5 years (SD = 12.3) and 40% male in non-LS and 57.8 years (SD = 12.2), and 51% male in controls.

The USA African American cohort

The case-control study consisted of 1,657 individuals. The sample included 781 sarcoidosis cases characterized as non-LS and 876 healthy controls. All individuals were taken from an extensive cohort of African-American (AA) sarcoidosis patients and controls assembled from various studies. Further details are available in Refs. (32–35).

Genotyping and quality control

Genotyping for sarcoidosis patients in the Swedish cohort was performed using the Illumina ImmunoChip platform and was performed at The SNP&SEQ Technology Platform in Uppsala University, Sweden, with Illumina Infinium assay using the Immuno Bead Chip (ImmunoChip version 1) as described in Rivera et al. (23). Healthy controls in the EIRA cohort ($n = 2,086$) were genotyped on the same platform at the Genome Institute of Singapore, as described previously (36, 37). Healthy controls in the EIMS cohort ($n = 1,060$) were genotyped with the same SNP array described elsewhere (26). Briefly, quality control filtering thresholds were applied using tools implemented in PLINK v1.09b software (38, 39). SNPs at call rate < 95% with minor allele frequency (MAF) > 1% and call rate < 99% with MAF ≤ 5% were filtered out. Moreover, SNPs that had Hardy-Weinberg Equilibrium (HWE) $P < 1 \times 10^{-7}$ (in the control group) were also excluded. Individuals with missing genotypes < 97% were also removed. Quality control (QC) resulted in 141,151 SNPs and 3,842 individuals on the Illumina ImmunoChip array.

Genotypes for the German cohort were quality controlled, as described in Fischer et al. (40). Genotyping for the African American cohort (US-AA) was performed at the OMRF using the Illumina HumanOmni1-Quad array for ~1.1M variants across the genome. Of these, 121,988 were in common with the ImmunoChip platform and carried forward for analysis. Further details on genotyping and quality control filtering are described in Adrianto et al. (41).

Statistical analysis

Association analysis of LS and non-LS in sex groups

In each sex group, we examined the association of single nucleotide polymorphisms (SNPs) in LS and non-LS. For assessing the association in LS, we applied an additive model using logistic regression adjusted for age and four principal components (PCs). PCs were derived from principal component analysis (PCA) performed using EIGENSTRAT (42) software on a pruned genotyped data set. The pruning of genotypes was performed using the pruning function and default parameters implemented in PLINK 1.90 beta software (39). For assessing the association in non-LS, we applied an additive model using logistic regression without covariate adjustment due to the lack of age information across the cohorts. In LS and non-LS, a significance threshold was defined by genomic control-corrected $P < 3.5 \times 10^{-7}$ (0.05/141,151 quality-controlled SNPs) considering Bonferroni correction. A suggestive $P < 5 \times 10^{-5}$ was also considered to identify potential signals (43–46). Furthermore, to account for high linkage disequilibrium (LD) ($r^2 \geq 0.8$) among associated SNPs, LD pruning was performed using PriorityPruner version 0.01.4 software¹ with default parameters, which identified and prioritized SNPs in genomic loci and thus referred as tag-SNPs.

As an exploratory analysis, to assess the effect of sex on genetics, we implemented an interaction model where the disease outcome was regressed on SNP adjusted for sex and the interaction term (SNP × sex) as covariates. A significant threshold for the interaction term was set at $P < 0.05$. The interaction analysis was performed using the glm function implemented in PLINK 2.0 software (39).

Additionally, we also performed a heterogeneity test to assess the effects of association between males and females. Heterogeneity statistics i.e., Cochran's Q-statistic, I^2 statistic (the percentage of variability in the effect sizes), and tau² (the between-study variance in our meta-analysis) were computed applying a random effects model and using the rma function implemented in the metafor R package (47).

Meta-analysis of significant SNPs

Meta-analysis was performed using SNP association results in non-LS. Herein, the meta-analysis was conducted using the inverse variance weighting (IVW) method implemented in the METAL software (48). For each SNP, the combined genetic effect size (defined by Beta), standard error (SE), meta- P -value (P_{meta}), total variability in effect size, also known as the heterogeneity index (I^2), and heterogeneity P -value were calculated. SNPs at $P_{meta} < 5 \times 10^{-8}$ were defined as genome-wide significant. Two meta-analyses were conducted, (1) a meta-analysis considering SNP associations in the European ancestry cohorts (Sweden and Germany) and (2) a meta-analysis considering SNP associations in the multi-ethnic (Sweden, Germany, and US African American).

Besides METAL computing heterogeneity index (I^2), we also performed an independent heterogeneity test applying the random effects model using the rma function implemented in the metafor

¹ <http://prioritypruner.sourceforge.net/>

R package (47). Summary of heterogeneity statistics include Cochran's Q-statistic, I^2 statistic (the percentage of variability in the effect sizes), and τ^2 (the between-study variance in our meta-analysis).

Lookup of significant SNPs in the UK Biobank

Using the UKB resource, we examined SNP associations at $P < 5 \times 10^{-5}$ using summary statistics of a genome-wide association study (GWAS) on "doctor-diagnosed sarcoidosis" (Data-Field ID: 22133), consisting of 91,787 individuals (395 cases and 91,392 controls) in the UKBB (49). Chiefly, sex-stratified summary statistics GWAS on males (189 cases and 41,045 controls) and females (206 cases and 50,347 controls) were obtained from https://github.com/Nealelab/UK_Biobank_GWAS. GWAS analysis was conducted using an additive model and logistic regression adjusted by the first 20 PCs, age, and age². A lookup significance threshold was set at $P < 5 \times 10^{-8}$.

Gene-based analysis

A gene-based analysis was conducted using SNP associations at $P < 5 \times 10^{-5}$ and MAGMA (50). Gene-based significance was set to $P < 2.0 \times 10^{-6}$ based on Bonferroni correction (0.05/24,769 genes). The number of genes was adopted from the VEGAS2 software definition, where 24,769 unique genes on the 22 autosomes were identified (51).

SNP enrichment analysis via expression quantitative trait loci (eQTLs) and pathway analysis

Expression quantitative trait loci enrichment was performed using SNP associations at $P < 5 \times 10^{-5}$ and gene expression data from 53 human tissues from the Genotype-Tissue Expression (GTEx) project version 8 available on the Functional Mapping and Annotation of genetic associations (FUMA) web tool (52). Additional eQTL enrichment using gene expression of lung tissue, whole blood, and immune cell types was conducted using EUGENE software (53, 54). A significant threshold for n SNP correlated with gene expression (defined as eQTL-SNP or eSNP) was set at $P < 0.05$.

Pathway analysis was conducted based on SNP associations at $P < 5 \times 10^{-8}$ to elucidate disease mechanisms using MetaCoreTM software.

Results

The summary of phenotypes and the number of samples in the sex groups for cohorts investigated are shown in Table 1.

Association results

The genome-wide association analysis using high-density mapping ImmunoChip array was conducted in males and females independently in Swedish, German, and US-African American cohorts.

TABLE 1 Sample size in association analysis in Sweden, Germany, and United States (A); sample size in replication lookup in the UK Biobank (B).

(A)				
Cohorts	LS		Non-LS	
	Males	Females	Males	Females
Swedish	166 cases vs. 849 HC	132 cases vs. 2,236 HC	263 cases vs. 855 HC	181 cases vs. 2,245 HC
German			470 cases vs. 1,750 HC	732 cases vs. 1,692 HC
USA-AA			322 cases vs. 497 HC	951 cases vs. 1,148 HC
(B)				
Males			Females	
344 cases vs. 166,644 controls			361 cases and 193,792 controls	

HC, healthy controls.

LS sex groups

In the Swedish cohort (Supplementary Tables 1, 2), 542 SNPs in LS males and 617 SNPs in LS females were identified at $P < 3.5 \times 10^{-7}$. Rs2187668 (OR = 4.715, SE = 0.1983, $P = 6.02 \times 10^{-14}$) located in *HLA-DQA1* was the top signal identified in males, and rs9268219 (OR = 3.667, SE = 0.1586, $P = 3.33 \times 10^{-15}$) located in *C6orf10* was the top signal identified in females. After considering linkage disequilibrium (LD) among associated signals, 77 SNPs in males and 101 SNPs in females remained, which were defined as tag-SNPs. The top results for LS sex groups (Table 2) are illustrated in the forest plot (Figure 1). Manhattan and QQ plots are shown in Supplementary Figure 1. Heterogeneity statistics are available in Supplementary Table 2X.

Conditioning for the top SNP rs2187668 in males identified 12 SNPs at $P < 5 \times 10^{-5}$. The most significant SNP was rs9296079 (OR = 7.476, SE = 0.410, $P = 9.51 \times 10^{-7}$), located at 7.8 kb 5' of *HLA-DPB2*. Similarly, conditioning for rs9268219 in females identified 8 SNPs at $P < 5 \times 10^{-5}$. The most significant was rs29243 (OR = 10.980, SE = 0.514, $P = 6.04 \times 10^{-6}$) located in *GABBR1*. Non-MHC SNPs at $P < 5 \times 10^{-5}$ were identified near *LRRTM4* in males and *CAST*, *LNPEP*, *TNIP1*, *CSMD1*, *B4GALNT1*, *SBNO2*, and *ABCG1* in females (Supplementary Table 3).

Non-LS sex groups

In the Swedish cohort (Supplementary Tables 4, 5), 5 SNPs in males at $P < 3.5 \times 10^{-7}$ and 1 SNP in females at $P = 6.73 \times 10^{-7}$ were identified. The top signals were rs1049550 (OR = 0.5231, SE = 0.1126, $P = 1.82 \times 10^{-8}$) located in *ANXA11* in males and rs1964995 (OR = 0.5314, SE = 0.1198, $P = 6.73 \times 10^{-7}$) located at 36 kb of the 3' of *HLA-DRB5*, in females. After LD assessment, 3 SNPs in males defined as tag-SNPs remained.

In the German cohort (Supplementary Tables 6, 7), rs1964995 (OR = 0.5452, SE = 0.08223, $P = 3.40 \times 10^{-11}$) in males and rs4502931 (OR = 0.5899, SE = 0.06549, $P = 4.36 \times 10^{-14}$) in females were identified as the top signals. In the US African American cohort (Supplementary Tables 8, 9), rs9271640 (OR = 1.771, SE = 0.1097, $P = 3.62 \times 10^{-7}$) in males and rs1964995 (OR = 0.5452, SE = 0.08223, $P = 3.40 \times 10^{-11}$) in females were identified as the top signals.

TABLE 2 The top tag-SNPs ($P_{GC} < 3.5 \times 10^{-7}$) associated with LS males and females in the Swedish cohort.

SNP	CHR	BP (hg19)	Alleles	RefSeq genes	Sex group	CA	CAF	OR	95% CI	P_{GC}
rs2187668	6	32,605,884	A/G	<i>HLA-DQA1</i>	LS-M	A	0.3557	4.7150	(3.1966, 6.9547)	6.02E-14
					LS-F	A	0.3557	3.4630	(2.5643, 4.6768)	6.57E-15
rs3130288	6	32,096,001	A/C	<i>ATF6B</i>	LS-M	A	0.3456	4.4560	(3.0489, 6.5124)	1.29E-13
					LS-F		0.3456	3.5090	(2.58, 4.7725)	1.39E-14
rs2105902	6	32,395,698	A/T	<i>12kb 5' of HLA-DRA</i>	LS-M	A	0.4010	4.2330	(2.9272, 6.1214)	1.87E-13
					LS-F		0.4010	3.1160	(2.32, 4.1851)	3.76E-13
rs3135382	6	32,383,441	C/A	<i>8.5kb 5' of BTNL2</i>	LS-M	C	0.4446	3.9120	(2.749, 5.567)	3.47E-13
					LS-F		0.4460	2.9310	(2.2128, 3.8823)	5.50E-13
rs3131643	6	31,442,782	A/G	<i>2.6kb 3' of HCG26</i>	LS-M	A	0.3788	3.9640	(2.7471, 5.7201)	1.59E-12
					LS-F		0.3788	2.8890	(2.1476, 3.8864)	1.57E-11
rs4143332	6	31,348,365	A/G	<i>19kb 5' of MICA</i>	LS-M	A	0.3389	4.1600	(2.8303, 6.1145)	3.32E-12
					LS-F		0.3389	3.3970	(2.5095, 4.5985)	2.66E-14
rs2071278	6	32,165,444	G/A	<i>NOTCH4</i>	LS-M	G	0.3725	3.8000	(2.6308, 5.4888)	8.55E-12
					LS-F		0.3725	3.1160	(2.3032, 4.2156)	1.36E-12
rs3094049	6	30,359,360	A/G	<i>45kb 3' of RPP21</i>	LS-M	A	0.2903	4.1510	(2.7873, 6.1819)	1.78E-11
					LS-F		0.2903	3.2150	(2.3198, 4.4557)	1.53E-11
rs2524069	6	31,244,789	T/A	<i>4.9kb 5' of HLA-C</i>	LS-M	T	0.3540	3.6090	(2.5168, 5.1752)	2.11E-11
					LS-F		0.3540	2.7470	(2.0481, 3.6845)	8.72E-11
rs2736157	6	31,600,820	G/A	<i>PRRC2A</i>	LS-M	G	0.4161	3.4380	(2.4292, 4.8657)	2.22E-11
					LS-F		0.4161	2.7580	(2.0587, 3.6949)	6.19E-11
rs9268219	6	32,284,108	C/A	<i>C6orf10</i>	LS-M	C	0.3356	4.4090	(3.0003, 6.4792)	4.10E-13
					LS-F		0.3356	3.6670	(2.6872, 5.004)	3.33E-15
rs1634721	6	30,977,680	A/G	<i>PBMUCL1</i>	LS-M	A	0.3221	3.4810	(2.372, 5.1085)	9.45E-10
					LS-F		0.3221	3.5930	(2.6382, 4.8934)	6.08E-15
rs3132449	6	31,626,013	A/G	<i>25bp 3' of APOM</i>	LS-M	A	0.3418	3.9780	(2.7203, 5.8173)	8.08E-12
					LS-F		0.3418	3.4680	(2.5569, 4.7038)	1.47E-14
rs1064627	6	30,698,541	G/A	<i>FLOT1</i>	LS-M	G	0.354	2.9360	(2.0263, 4.2541)	4.61E-08
					LS-F		0.354	3.3110	(2.4618, 4.4532)	2.64E-14
rs9266669	6	31,348,077	A/G	<i>19kb 5' of MICA</i>	LS-M	A	0.3625	3.3530	(2.3232, 4.8393)	5.56E-10
					LS-F		0.3625	3.3570	(2.4872, 4.531)	2.73E-14
rs886423	6	30,782,205	G/C	<i>70kb 5' of DDR1</i>	LS-M	G	0.3322	2.8220	(1.9374, 4.1107)	2.11E-07
					LS-F		0.3322	3.2850	(2.4358, 4.4303)	6.69E-14
rs3129963	6	32,380,208	G/A	<i>5.3kb 5' of BTNL2</i>	LS-M	G	0.4513	3.7770	(2.6521, 5.3791)	1.53E-12
					LS-F		0.4513	3.0010	(2.2666, 3.9734)	1.61E-13
rs2233974	6	31,080,016	C/G	<i>C6orf15</i>	LS-M	C	0.3793	3.3470	(2.325, 4.8184)	4.47E-10
					LS-F		0.3793	3.1680	(2.3592, 4.2542)	1.66E-13
rs3130477	6	31,428,920	G/A	<i>2kb 5' of HCP5</i>	LS-M	G	0.3389	4.0890	(2.7858, 6.0019)	5.00E-12
					LS-F		0.3389	3.2220	(2.3816, 4.359)	2.91E-13

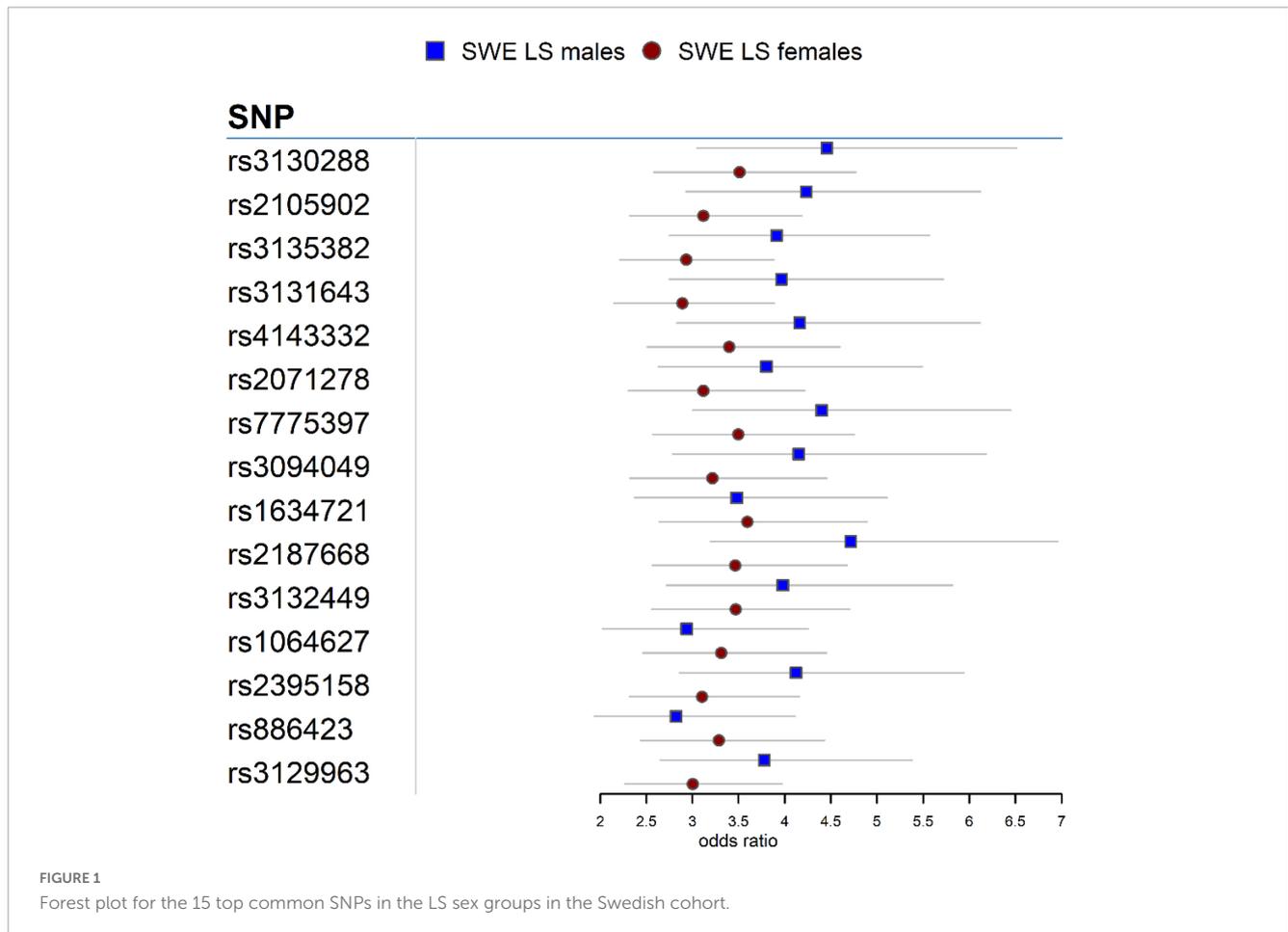
SNP, single nucleotide polymorphism; BP, chromosome base pairs are based on human genome assembly version 19 (hg19); CA, coded allele; CAF, coded allele frequency; OR, odds ratio; SE, standard error of odds ratio; P_{GC} , genomic controlled P -value.

The top association results for each cohort are summarized in [Table 3](#). Manhattan and QQ plots for non-LS sex groups are shown in [Supplementary Figures 2–4](#).

Interaction analysis with sex

Results from the interaction analysis in the LS and non-LS sex groups in the Swedish cohort showed significant

SNPs interacting with the sex variable. In LS sex groups ([Supplementary Table 10](#)), the most significant interacting SNPs were rs2853973 ($P_{int} = 3.8 \times 10^{-4}$) in males and rs1470410 ($P_{int} = 5.44 \times 10^{-4}$) in females. In non-LS sex groups ([Supplementary Table 11](#)), the most significant interacting SNPs were rs10940422 ($P_{int} = 4.18 \times 10^{-4}$) in males and rs12432418 ($P_{int} = 4.85 \times 10^{-5}$) in females.



Meta-analysis in non-LS

Non-LS sex groups

Meta-analysis in the European cohorts (Sweden and Germany) (Supplementary Tables 12, 13) identified 57 SNPs in males and 112 SNPs in females at $P_{meta} < 5 \times 10^{-8}$. Top signal rs1964995 located 36 kb from the 3' of *HLA-DRB5* was the same in males (OR = 1.715, SE = 0.0656, $P_{meta} = 3.92 \times 10^{-18}$) and females (OR = 1.7304, SE = 0.0598, $P_{meta} = 5 \times 10^{-20}$). A forest plot illustrating the top findings is shown in Figure 2. Manhattan and Q-Q plots of the meta-analysis are shown in Supplementary Figure 5. Heterogeneity statistics in the SWE-GER are shown in Supplementary Table 13X.

A comparison of associations at $P_{meta} < 5 \times 10^{-8}$ showed 44 SNPs shared between non-LS males and females, 17 SNPs exclusively associated with non-LS males, and 215 SNPs exclusively associated with non-LS females. Most SNPs were in the MHC region. Non-MHC SNPs include rs694739 (OR = 1.3626, SE = 0.063, $P_{meta} = 9.21 \times 10^{-07}$) located at 7.9 kbp from the 3' of *PRDX5* associated with non-LS males, and rs2573346 (OR = 0.7712, SE = 0.0564, $P_{meta} = 4.19 \times 10^{-06}$) in *ANXA11* associated with non-LS females.

Results from a meta-analysis in the multi-ethnic cohorts (Sweden, Germany, and US-African American) (Supplementary Tables 14, 15) identified 12 SNPs in non-LS males and 49 in non-LS females (Table 3). Top SNPs were rs1964995 (OR = 1.7736,

SE = 0.0587, $P_{meta} = 1.54 \times 10^{-22}$) located 36 kb from 3' of *HLA-DRB5* in males, and rs2395153 (Beta = 0.6565, SE = 0.0477, $P_{meta} = 1.11 \times 10^{-18}$) located at 5.9 kb from 5' of *C6orf10* in females. The top meta-SNPs are shown in Table 4. Forest plots (Figure 2) illustrate the top common meta-SNPs in the European and multi-ancestry cohort groups. Manhattan plots of meta-analysis in multi-ethnic groups are shown in Supplementary Figure 6. Heterogeneity statistics in SWE-GER-USA are shown in Supplementary Table 15X.

A comparison analysis at $P_{meta} < 5 \times 10^{-8}$ showed a shared SNP between non-LS males and females, 11 SNPs exclusively associated with non-LS males, and 48 SNPs exclusively associated with non-LS females. Most associated SNPs were in the MHC region. Non-MHC SNPs were observed in non-LS females and included rs7813186, rs1049550, and rs7133604 (Data not shown).

Lookup of meta-SNPs in the UK Biobank

Genome-wide association study of "Doctor diagnosed sarcoidosis" in the UKB was used as a validation step. SNP-lookup showed several sex-specific SNPs in LS and non-LS at $P < 5 \times 10^{-5}$. Chiefly, 223 SNPs associated with LS males and 209 SNPs associated with LS females were validated at $P < 5 \times 10^{-8}$ (Supplementary Tables 16, 17). In non-LS, 11 non-LS males and 31 SNPs non-LS females in the SWE-GER cohorts were validated

TABLE 3 The top tag-SNPs ($P_{GC} < 3.5 \times 10^{-7}$) associated with non-LS males and females in Swedish, German, and USA-AA cohorts.

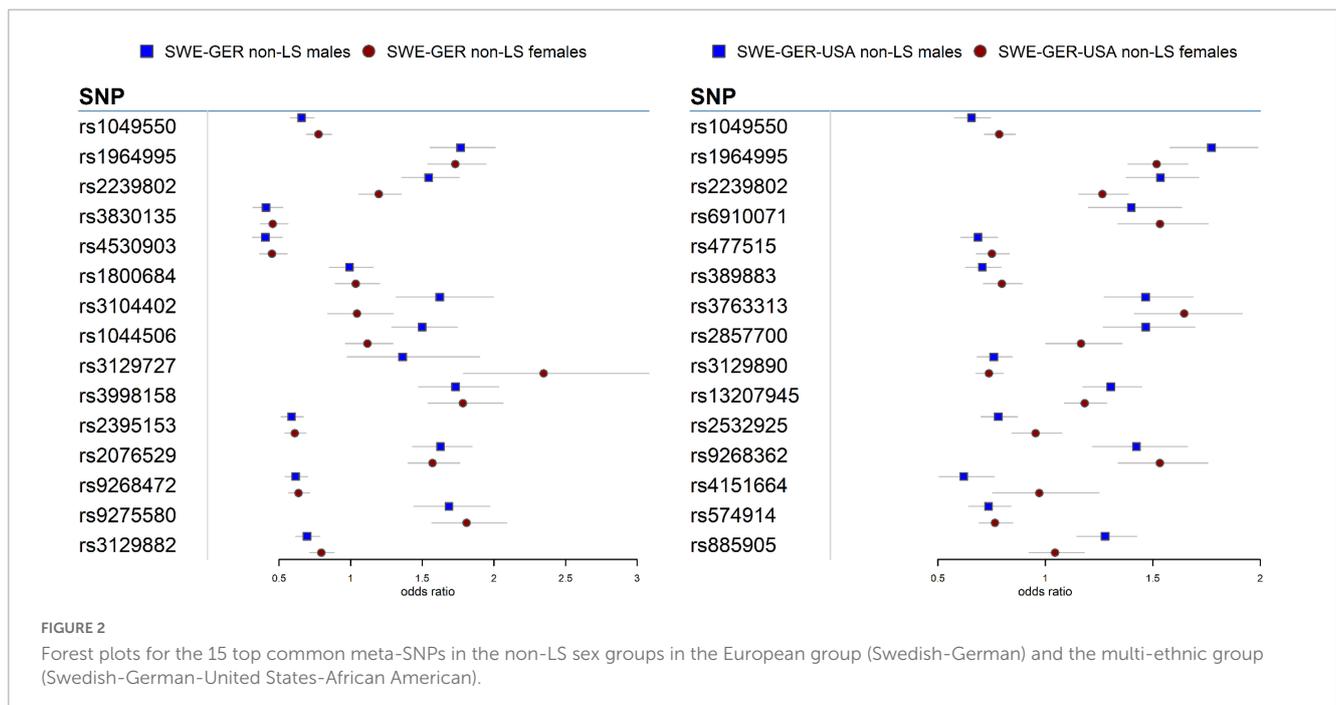
SNP	CHR	BP (hg19)	Alleles	RefSeq genes	Cohort	Males				
						CA	CAF	OR	95% CI	P_{GC}
rs1049550	10	81,926,702	A/G	ANXA11	SWE	A	0.2962	0.5231	(0.4195, 0.6523)	1.82E-08
					GER	A	0.3845	0.7325	(0.6281, 0.8543)	3.63E-04
					USA-AA	A	0.1508	0.7724	(0.5841, 1.0215)	7.70E-02
rs2239802	6	32,411,846	C/G	HLA-DRA	SWE	C	0.3375	1.7790	(1.4295, 2.214)	4.52E-07
					GER	C	0.2283	1.4480	(1.2336, 1.6997)	4.75E-05
					USA-AA	C	0.3321	1.5140	(1.2288, 1.8655)	1.41E-04
rs7197	6	32,412,580	A/G	HLA-DRA	SWE	A	0.2885	1.7610	(1.4056, 2.2063)	1.50E-06
					GER	A	0.1901	1.4210	(1.1976, 1.6862)	2.96E-04
					USA-AA	A	0.2253	1.6350	(1.2895, 2.0731)	7.32E-05
rs1964995	6	32,449,411	G/A	36kb 3' of HLA-DRB5	SWE	G	0.2995	0.6033	(0.4872, 0.747)	5.74E-06
					GER	G	0.3664	0.5452	(0.464, 0.6406)	3.40E-11
					USA-AA	G	0.2314	0.5568	(0.431, 0.7194)	1.21E-05
rs3830135	6	32,548,464	A/G	HLA-DRB1	SWE	A	0.0529	0.3543	(0.2278, 0.5512)	6.71E-06
					GER	A	0.0980	0.4374	(0.3212, 0.5956)	2.40E-06
					USA-AA	C	0.0702	0.5942	(0.3897, 0.906)	1.82E-02
rs3131283	6	32,119,898	A/G	177bp 5' of PRRT1	SWE	A	0.1971	1.8100	(1.4054, 2.3312)	6.85E-06
					GER	A	0.1417	1.3790	(1.1382, 1.6707)	3.21E-03
					USA-AA	C	0.0256	1.7040	(0.9263, 3.1348)	9.41E-02
rs4530903	6	32,581,889	A/G	23kb 5' of HLA-DQA1	SWE	A	0.0529	0.3590	(0.2308, 0.5585)	8.79E-06
					GER	A	0.0973	0.4282	(0.3133, 0.5853)	1.74E-06
					USA-AA	T	0.0702	0.5942	(0.3897, 0.906)	1.82E-02
rs1800684	6	32,151,994	T/A	AGER	SWE	T	0.1982	1.7920	(1.3917, 2.3076)	9.75E-06
					GER	A	0.1428	1.3810	(1.1401, 1.6729)	3.00E-03
					USA-AA	A	0.0293	1.6280	(0.9369, 2.8289)	9.13E-02
rs3104402	6	32,681,676	A/C	27kb 5' of HLA-DQA2	SWE	A	0.1050	2.0860	(1.5013, 2.8984)	1.83E-05
					GER	T	0.0689	1.3750	(1.0539, 1.794)	3.50E-02
					USA-AA	G	0.0153	2.3790	(1.0553, 5.3629)	4.12E-02
rs1044506	6	32,172,065	A/C	NOTCH4	SWE	A	0.2016	1.7870	(1.3878, 2.3011)	1.08E-05
					GER	T	0.1441	1.3560	(1.1203, 1.6413)	4.97E-03
					USA-AA	G	0.0263	1.9640	(1.0695, 3.6067)	3.35E-02
Females										
rs1964995	6	32,449,411	G/A	36kb 3' of HLA-DRB5	SWE	G	0.2995	0.5314	(0.4202, 0.6721)	6.73E-07
					GER	G	0.3611	0.5942	(0.519, 0.6804)	1.65E-12
					USA-AA	G	0.2408	0.8051	(0.6963, 0.9309)	3.77E-03
rs3129727	6	32,679,690	A/G	29kb 5' of HLA-DQA2	SWE	A	0.0474	3.0760	(2.0025, 4.7251)	1.37E-06
					GER	A	0.0268	1.9580	(1.3773, 2.7836)	4.53E-04
					USA-AA	T	0.0217	1.6270	(1.0686, 2.4773)	2.47E-02
rs3998158	6	32,681,992	G/A	27kb 5' of HLA-DQA2	SWE	G	0.1700	0.5047	(0.3771, 0.6755)	1.49E-05
					GER	C	0.1990	0.5811	(0.4908, 0.6881)	3.67E-09
					USA-AA	C	0.1665	1.0890	(0.925, 1.2821)	3.10E-01
rs2395153	6	32,345,595	C/G	5.9kb 5' of C6orf10	SWE	C	0.2748	0.5947	(0.4694, 0.7535)	5.05E-05
					GER	C	0.3391	0.6137	(0.5347, 0.7044)	7.74E-11
					USA-AA	G	0.2175	0.7410	(0.6372, 0.8618)	1.18E-04

(Continued)

TABLE 3 (Continued)

SNP	CHR	BP (hg19)	Alleles	RefSeq genes	Cohort	Females				
						CA	CAF	OR	95% CI	P_{GC}
rs2076529	6	32,363,955	G/A	<i>BTNL2</i>	SWE	G	0.3176	0.6105	(0.4859, 0.7672)	6.66E-05
					GER	C	0.3752	0.6458	(0.5659, 0.7371)	1.22E-09
					USA-AA	C	0.3056	0.8406	(0.7363, 0.9598)	1.11E-02
rs9268472	6	32,355,605	A/G	6.9kb 3' of <i>BTNL2</i>	SWE	A	0.3176	0.6107	(0.486, 0.7674)	6.71E-05
					GER	T	0.3754	0.6446	(0.5648, 0.7357)	1.04E-09
					USA-AA	A	0.3063	0.8471	(0.7421, 0.967)	1.51E-02
rs9275580	6	32,679,462	G/A	30kb 5' of <i>HLA-DQA2</i>	SWE	G	0.1791	0.5543	(0.4192, 0.7329)	9.63E-05
					GER	C	0.2102	0.5519	(0.4666, 0.6528)	7.77E-11
					USA-AA	G	0.2458	1.1700	(1.0138, 1.3503)	3.35E-02
rs3129882	6	32,409,530	G/A	<i>HLA-DRA</i>	SWE	G	0.5146	1.5580	(1.263, 1.922)	9.81E-05
					GER	G	0.4499	1.1700	(1.0356, 1.322)	1.85E-02
					USA-AA	A	0.4586	1.0660	(0.9446, 1.203)	3.05E-01

SNP, single nucleotide polymorphism; BP, chromosome base pairs are based on human genome assembly version 19 (hg19); CA, coded allele; CAF, coded allele frequency; OR, odds ratio; SE, standard error of odds ratio; P_{GC} , genomic controlled P -value.



at $P < 5 \times 10^{-8}$ (Supplementary Tables 18, 19). In the multi-ethnic cohorts (SWE-GER-USA-AA), 2 SNPs in non-LS males and 8 SNPs in non-LS females were observed (Supplementary Tables 20, 21).

Gene-based analysis

The gene-based analysis revealed significant genomic loci associated with sex groups in LS and non-LS (Supplementary Tables 22, 23). The most significant genomic locus was *HLA-DRA* ($P = 4.36 \times 10^{-14}$) in LS males and *FKBP1* ($P = 1.39 \times 10^{-14}$) in LS females at a gene-based $P < 2.0 \times 10^{-6}$ (0.05/24,769 genes).

In the European cohorts (SWE-GER), 15 genes associated with non-LS males and 18 genes associated with non-LS females were identified, where the most significant genomic locus was *HLA-DRA* ($P = 5.16 \times 10^{-15}$) in non-LS males and *HLA-DRB1* ($P = 2.42 \times 10^{-13}$) in non-LS females. In the multi-ethnic cohorts (SWE-GER-USA-AA), 5 genes associated with non-LS males and 8 genes associated with non-LS females were identified. The top genes were *HLA-DRA* ($P = 1.88 \times 10^{-14}$) in non-LS males and *C6orf10* ($P = 6.47 \times 10^{-12}$) in non-LS females.

A comparison of gene-based analysis in LS and non-LS sex groups across cohorts (Supplementary Table 24) revealed differences in gene-based associations. In LS, gene-based associations showed 77 genes shared among sex groups, 11

TABLE 4 The top significant SNPs ($P_{meta} < 5e-8$) of GWAS meta-analysis on non-LS sex groups.

Sex group (Cohorts)	SNP	CHR	BP	Alleles	RefSeq_genes	CA	OR	95% CI	P-value	Direction	Het-I ²	Het-PVal
Non-LS males (SWE-GER)	rs1964995	6	32,449,411	G/A	36kb 3' of HLA-DRB5	A	1.7681	(1.5548, 2.0107)	3.92E-18	++	0	0.4583
	rs2395153	6	32,345,595	C/G	5.9kb 5' of C6orf10	C	0.5861	(0.5134, 0.6692)	2.61E-15	-	0	0.5916
	rs2213585	6	32,413,150	G/A	323bp 3' of HLA-DRA	A	0.6315	(0.5618, 0.71)	1.35E-14	-	0	0.3240
	rs7195	6	32,412,539	A/G	HLA-DRA	A	1.5830	(1.4082, 1.7795)	1.46E-14	++	14.5	0.2795
	rs9271588	6	32,590,953	G/A	14kb 5' of HLA-DQA1	A	1.6153	(1.4291, 1.8258)	1.72E-14	++	0	0.5981
	rs2213586	6	32,413,094	A/G	267bp 3' of HLA-DRA	A	1.5785	(1.4046, 1.7742)	1.93E-14	++	10	0.2918
	rs2227139	6	32,413,459	G/A	632bp 3' of HLA-DRA	A	0.6335	(0.5637, 0.712)	1.93E-14	-	10	0.2918
	rs7192	6	32,411,646	A/C	HLA-DRA	A	1.5763	(1.4023, 1.7721)	2.54E-14	++	0.3	0.3165
	rs4373382	6	32,350,868	C/A	11kb 5' of C6orf10	A	1.6291	(1.4337, 1.8512)	6.84E-14	++	0	0.6069
	rs2076529	6	32,363,955	G/A	BTNL2	A	1.6268	(1.4322, 1.8479)	7.03E-14	++	0	0.5625
Non-LS females (SWE-GER)	rs1964995	6	32,449,411	G/A	36kb 3' of HLA-DRB5	A	1.7305	(1.5391, 1.9457)	5E-20	++	0	0.4192
	rs4502931	6	32,380,782	A/T	5.9kb 5' of BTNL2	A	0.6033	(0.5401, 0.674)	3.82E-19	-	0	0.4968
	rs2395153	6	32,345,595	C/G	5.9kb 5' of C6orf10	C	0.6088	(0.5405, 0.6859)	3.14E-16	-	0	0.8219
	rs9275582	6	32,680,070	A/G	29kb 5' of HLA-DQA2	A	0.5499	(0.4757, 0.6358)	6.15E-16	-	0	0.7935
	rs2647012	6	32,664,458	A/G	30kb 5' of HLA-DQB1	A	1.5502	(1.3899, 1.7291)	3.74E-15	++	0	0.5473
	rs9275393	6	32,669,439	A/G	35kb 5' of HLA-DQB1	A	0.5890	(0.5156, 0.6728)	6.22E-15	-	0	0.3517
	rs2294878	6	32,367,795	A/C	BTNL2	A	0.6563	(0.5872, 0.7337)	1.27E-13	-	0	0.4173
	rs4530903	6	32,581,889	A/G	23kb 5' of HLA-DQA1	A	0.4500	(0.364, 0.5564)	1.68E-13	-	0	0.5299
	rs2858332	6	32,68,1161	A/C	28kb 5' of HLA-DQA2	A	0.6747	(0.6043, 0.7535)	2.83E-12	-	0	0.6205
	rs2858867	6	32,575,325	G/A	18kb 5' of HLA-DRB1	A	0.6810	(0.6105, 0.7597)	5.54E-12	-	0	0.7201
Non-LS males (SWE-GER-USA-AA)	rs1964995	6	32,449,411	G/A	36kb 3' of HLA-DRB5	A	1.7736	(1.5809, 1.9899)	1.54E-22	+++	0	0.7552
	rs2239802	6	32,411,846	C/G	HLA-DRA	C	1.5360	(1.3764, 1.7143)	1.88E-14	+++	19.6	0.2883
	rs4530903	6	32,581,889	A/G	23kb 5' of HLA-DQA1	A	0.4478	(0.36, 0.5571)	5.40E-13	—	27.7	0.2506
	rs477515	6	32,569,691	A/G	12kb 5' of HLA-DRB1	A	0.6866	(0.607, 0.7767)	2.21E-09	—	0	0.4141
	rs389883	6	31,947,460	C/A	STK19	A	0.7073	(0.6302, 0.7939)	4.02E-09	—	0	0.3934

(Continued)

TABLE 4 (Continued)

Sex group (Cohorts)	SNP	CHR	BP	Alleles	RefSeq_genes	CA	OR	95% CI	P-value	Direction	Het-I ²	Het-PVal
	rs3763313	6	32,376,471	C/A	1.6kb 5' of <i>BTNL2</i>	A	1.4667	(1.2747, 1.6877)	9.03E-08	+++	21.6	0.2794
	rs9268835	6	32,428,115	A/G	15kb 3' of <i>HLA-DRA</i>	A	0.7131	(0.6293, 0.8081)	1.13E-07	—	0	0.6909
	rs2857700	6	31,572,481	A/G	11kb 5' of <i>AIF1</i>	A	1.4676	(1.27, 1.696)	2.02E-07	+++	16.1	0.3036
	rs3129890	6	32,414,273	G/A	1.4kb 3' of <i>HLA-DRA</i>	A	0.7603	(0.6831, 0.8464)	5.58E-07	—	64.4	0.0604
	rs2269425	6	32,123,639	A/G	<i>PPT2</i>	A	0.6415	(0.538, 0.765)	7.69E-07	—	0	0.9529
Non-LS females (SWE-GER-USA-AA)	rs2395153	6	32,345,595	C/G	5.9kb 5' of <i>C6orf10</i>	C	0.6565	(0.598, 0.7209)	1.11E-18	—	50.8	0.1311
	rs7195	6	32,412,539	A/G	<i>HLA-DRA</i>	A	1.4402	(1.3254, 1.5651)	7.42E-18	+++	0	0.8684
	rs2294878	6	32,367,795	A/C	<i>BTNL2</i>	A	0.7053	(0.6471, 0.7689)	2.06E-15	—	57.1	0.0971
	rs3129890	6	32,414,273	G/A	1.4kb 3' of <i>HLA-DRA</i>	A	0.7374	(0.6772, 0.8031)	2.61E-12	—	20.5	0.2843
	rs6913309	6	32,339,840	A/T	183bp 5' of <i>C6orf10</i>	A	0.7060	(0.6397, 0.7792)	4.25E-12	—	0	0.5575
	rs2269425	6	32,123,639	A/G	<i>PPT2</i>	A	0.6160	(0.5359, 0.7082)	9.74E-12	—	5.5	0.3471
	rs3129727	6	32,679,690	A/G	29kb 5' of <i>HLA-DQA2</i>	A	2.1069	(1.6768, 2.6473)	1.62E-10	+++	56.6	0.1000
	rs6910071	6	32,282,854	G/A	<i>C6orf10</i>	A	1.5340	(1.3392, 1.7573)	6.61E-10	+++	0	0.4492
	rs3115572	6	32,220,484	C/G	29kb 5' of <i>NOTCH4</i>	C	1.2871	(1.1875, 1.3951)	8.48E-10	+++	14.3	0.3113
	rs3817963	6	32,368,087	G/A	<i>BTNL2</i>	A	1.3835	(1.2453, 1.5371)	1.53E-09	+++	0	0.8790

SNP, single nucleotide polymorphism; CHR, chromosome; BP, base pairs; hg19, human genome assembly version 19; OR, odds ratio; 95% CI, 95% confidence interval; P_{meta} , meta P-value; Direction: (–) reversed strand; (++) forward strand; Het-I², heterogeneity I-squared; Het-PVal, heterogeneity I-squared P-value.

genes exclusively associated with males, and 14 genes exclusively associated with females. In non-LS sex groups, in the European cohorts (SWE-GER), 18 genes were shared among non-LS sex groups, and 8 genes were exclusively associated with non-LS males. In the multi-ethnic cohorts (SWE-GER-USA-AA), 4 genes were shared, one gene was exclusively associated with non-LS males, and 4 were exclusively associated with non-LS females.

Figure 3 shows a circos plot depicting gene-based associations among LS and non-LS sex groups across cohorts.

Gene expression and expression quantitative trait loci (eQTL) enrichment

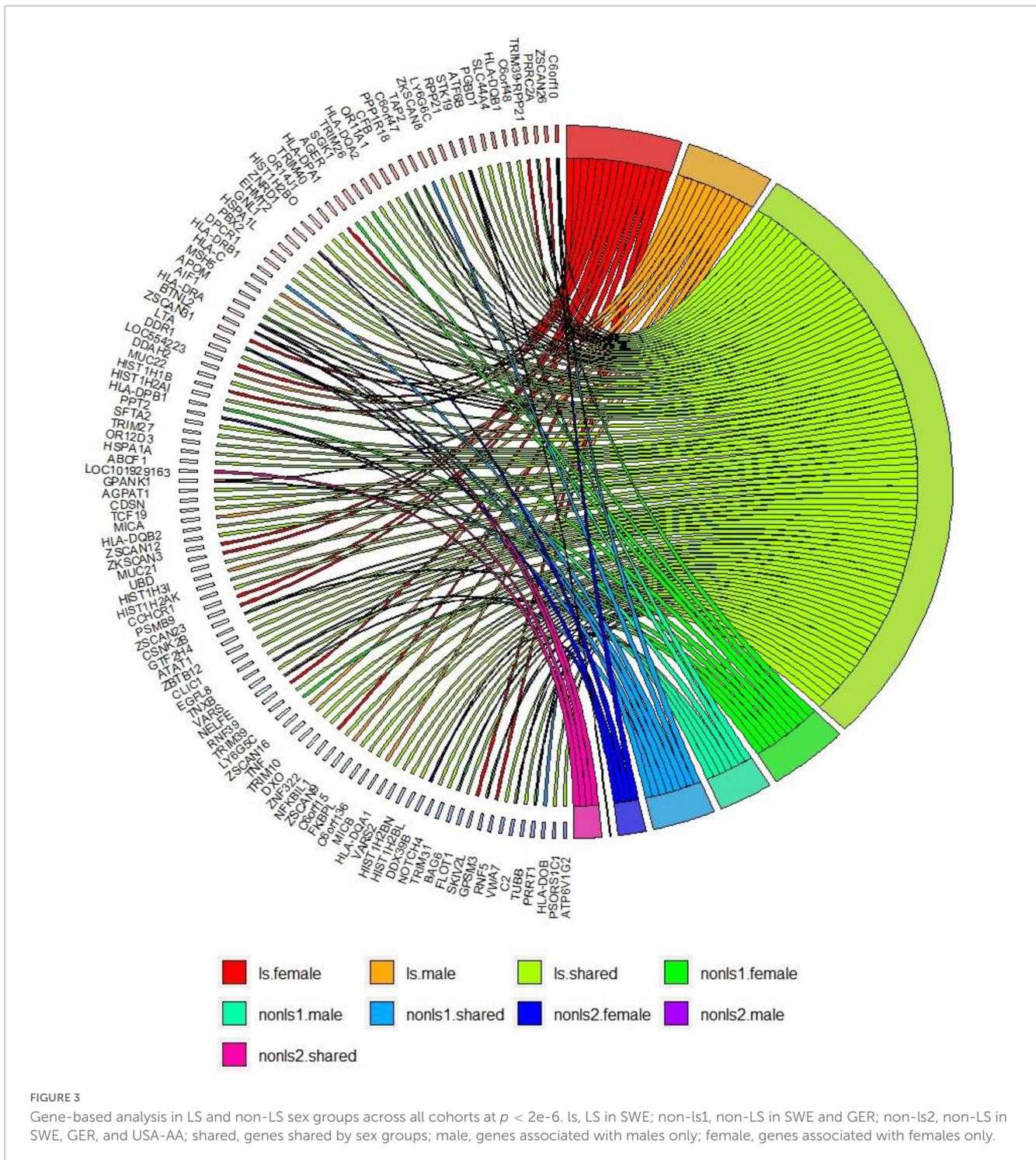
Results from tissue expression analysis using 54 tissue types in GText version 8 and FUMA revealed significant sex-specific gene expression at Bonferroni P -value ($P_{Bon} < 0.05$).

Löfgren's syndrome males and females (Supplementary Figure 7), significant differential gene expression was observed in the spleen, small intestine, lung, and blood.

In the European cohorts (SWE, GER) (Supplementary Figure 8), differential gene expression was observed in the spleen, small intestine terminal ileum, lung, and blood in non-LS males. In contrast, significant differential gene expression was observed in the brain hippocampus, and lung in non-LS females. In the multi-ethnic cohorts (SWE, GER, USA-AA) (Supplementary Figure 8), differentially expressed genes were observed in the spleen, small intestine terminal ileum, lung, and blood in non-LS males and females.

Expression quantitative trait loci enrichment using different eQTL databases across different tissues and cell types (Supplementary Table 25) revealed several significant eQTLs (eSNPs) in sex groups.

In LS sex groups, 311 eSNPs in 134 genes in males and 354 eSNPs in 140 genes in females were identified using the GText v8 database. In SWE-GER cohorts, non-LS sex groups, 32 eSNPs in 77 genes in males and 59 eSNPs in 89 genes in females were identified using the GText v8 database. In the SWE-GER-USA-AA cohorts, non-LS sex groups, two eSNPs in males and 23 eSNPs in females were identified using the GText v8 database.



Similar observations were also observed using eQTLGen and eQTLcatalogue databases.

Moreover, eQTL mapping using immune cell eQTLs (Supplementary Table 26) available in the EUGENE database identified significant eSNPs in relevant immune cell types. In LS sex, significant eSNPs of B cells, T cells, CD4, and CD8 T cells, granulocytes, monocytes, macrophages, neutrophils, and PBMCs were identified. In the SWE-GER-USA-AA non-LS sex groups, eQNPs of T cells, CD4 T cells, fibroblasts, monocytes, macrophages, and neutrophils were observed.

Gene enrichment and pathway analysis

In LS sex groups (Supplementary Table 27), a pathway map related to immune response induction of the antigen presentation machinery by IFN-gamma (FDR = 1.530e-8 in males; FDR = 1.926e-11 in females) were identified. In the SWE-GER non-LS sex groups (Supplementary Table 27), pathway maps related to maturation and migration of dendritic cells in skin sensitization (FDR = 2.085e-9) in males and related to immune response

induction of the antigen presentation machinery by IFN-gamma (FDR = 1.998e-10) in females were identified. In the SWE-GER-USA-AA non-LS sex groups (Supplementary Table 27), pathway maps related to immune response lectin-induced complement pathway (FDR = 1.571e-5) in males and related to maturation and migration of dendritic cells in skin sensitization (FDR = 1.824e-6) in females were noted.

Discussion

The present study investigated sex differences in sarcoidosis, particularly in LS and non-LS phenotypes. This work is the first investigation for independently characterizing the genetic architecture of sarcoidosis in males and females. In brief, we analyzed genetic associations between LS and non-LS males and females conducted in European and multi-ethnic cohorts. The multi-ethnic cohorts comprised three population-based cohorts from Sweden, Germany, and the United States. The cohorts included different ancestries, i.e., white European and black African American.

For both LS and non-LS sex groups, a genetic analysis based on a genome-wide association study was conducted, followed by a meta-analysis, an SNP lookup in the UKB, gene-based analysis, differential gene expression analysis using the GText version 8 resource, eQTL tissue/cell mapping enrichment, and pathway analysis.

Löfgren's syndrome, an acute form of sarcoidosis, is highly prevalent in Sweden (55). However, due to insufficient LS cases in the German cohort and being rare in the United States, genetic analyses in the LS sex groups were only conducted in the Swedish cohort. In contrast, genetic investigations in the non-LS sex groups were performed across all three cohorts.

Association analysis in LS sex groups identified distinct genetic associations clustering in the extended MHC (xMHC) (56), highlighting the role of genes located in this region as main component of LS susceptibility, as previously reported (23). The top LS findings were associations in *HLA-DQA1* in males and *C6orf10* in females. Although these genomic loci are on the same chromosome region (classical class II) and many variants in these genes are in linkage disequilibrium (LD $r^2 > 0.8$), we observed significant differences in the genetic effects (beta estimates) among variants associated with LS males and females (Figure 1). Interestingly, after conditioning for the top MHC signals, we identified non-MHC loci worth investigating in future studies of LS. For instance, a nearby signal close to *LRRTM4* associated with LS males is worth looking into, as *LRRTM4* is a synapse-organizing molecule involved in protein-protein interactions that regulate the nervous and immune systems (57). Another interesting non-MHC association is *LNPEP* associated with LS females. *LNPEP* is an aminopeptidase that regulates hormones (i.e., arginine-vasopressin and oxytocin), and is involved in trimming peptide antigens for cross-presentation to T cells in autoimmune diseases (58).

In the European cohorts (SWE and GER), meta-analysis in non-LS sex groups identified mainly associations in the MHC classical class II subregion and a few non-MHC signals, chiefly in *ANXA11*, *TMEM163*, and nearby *PRDX5* in both sex groups. As in LS, the differences among these signals were the genetic

effects (beta estimates) in the sex groups, as illustrated in the forest plot (Figure 2). Similar findings were observed in the multi-ethnic cohorts (SWE, GER, and USA-AA).

To further understand the genetic differences, the comparison of genetic signals of LS and non-LS sex groups showed that some signals were associated with a specific sex, whereas other signals were common to both males and females; however, their genetic effects differed. These observations were then validated using the UKB resource, using the “doctors diagnosed sarcoidosis” phenotype. Based on this evidence, it is compelling to suggest that genetic differences in the sex groups could be linked to the observed differences in clinical manifestations in sarcoidosis male and female patients, as documented in various works (59–61). Additionally, it is essential to highlight that ancestry also played a role in the differences in genetic effects in the sex groups, as the effects changed when multi-ethnic ancestries were considered.

Results from gene-based analysis showed patterns of genomic associations in the sex groups, such as shared and sex-specific (Figure 3). Notably, the genomic patterns of associations in the LS sex groups were more extensive since the LS susceptibility spanned across the xMHC, which harbors 252 genes and 139 pseudogenes. In the non-LS sex groups, on the contrary, the genomic patterns were limited to a few genes located in the MHC class II region. As expected, genomic patterns decreased as more ancestries were included, stressing the role of ancestry as a modifier of sarcoidosis susceptibility. Ancestry plays a significant role in the genetics of complex diseases (62, 63), which is also relevant in sarcoidosis.

Insights from tissue gene expression and eQTL enrichment showed that SNPs associated with biological sex play a role in gene expression variability across various tissues and cell types. For instance, gene expression in the spleen, small intestine, lung, and whole blood differed in LS and non-LS sex groups. Furthermore, SNP associations correlated with eQTLs of tissues and cell types, particularly immune cells eQTLs (e.g., B cells, T cells, CD4 T cells, CD8 T cells, macrophages, and monocytes), were identified. These findings underscore the functional role of genetic variants on gene expression of immune cells, thus shaping the immunopathogenesis of sarcoidosis in the sex groups. eQTL studies in complex diseases showed that eQTL SNPs act as master regulators of gene expression in several tissues and cell types (64, 65), influencing the expression of multiple genes and acting as gene regulators (66).

Expression quantitative trait loci findings in various immune cell types (Supplementary Table 26) further supports our hypothesis that sarcoidosis is an immune-regulated disease, as previously proposed (67–69).

Indeed, differences in SNP associations in HLA genes (70–75) may result in different immune responses and inflammatory disease course as seen in male and female patients with sarcoidosis. Thus, it is worth keeping in mind that HLA genes play a crucial role in autoimmunity and that their function has a profound sex bias on disease phenotype, as more females are affected by autoimmune (or chronic inflammatory) disorders than the counterparts. Understanding the molecular mechanisms underlying sex selection in the development of autoimmune disorders is an area of active research (76).

Compiling the evidence shown in this work, with a particular focus on the strong presence of HLA genetic associations in sarcoidosis, the involvement of various immune cells types, and the differences in immune response and clinical course in sex groups

(67, 68, 77), the possibility of sarcoidosis being an autoimmune disorder cannot be ignored.

Besides HLA genes, we also identified signals in the *ANXA11* gene in non-LS sex groups with different genetic effects, as denoted by rs1049550 (Figure 2). SNP associations in *ANXA11* were first reported by Hofmann et al. (78) and have been linked to sarcoidosis risk, radiographic phenotype (Scadding stage IV), and to interact with HLA-DRA (79). SNPs in *ANXA11* have also been reported to be associated with autoimmune diseases (80). Indeed, the functional role of *ANXA11* in cell division and apoptosis (81) and the regulation of inflammatory cells (82) make *ANXA11* an attractive biomarker for sarcoidosis.

Moreover, evidence for a pathway related to immune response induction of the antigen presentation machinery by IFN-gamma in LS sex groups offers an incentive to examine the role of IFN-gamma in sarcoidosis (Supplementary Table 27). IFN-gamma is a pro-inflammatory cytokine produced by immune cells and has immune effector functions on many genes involved in tissue homeostasis, immune/inflammatory responses, and tumor immunosurveillance (83). In the non-LS groups, pathway maps identified mainly were related to immune response but not specific to a sex group. Thus, more data is needed to understand the underlying mechanisms of non-LS as this sarcoidosis phenotype is heterogeneous and polygenic.

In summary, we provide new insights into a sex bias underpinning the genetic architecture of sarcoidosis. The differences observed in the genetic susceptibility in the sex groups in LS and non-LS add further information which may explain the well-recognized observations between sex groups, incidence, prevalence, age of onset, symptoms, severity of the disease, and drug reactions observed in male and female sarcoidosis patients (23, 84–87).

It goes without a doubt that further genetic and omics studies ought to be conducted to broaden the characterization of the genetic architecture of sarcoidosis among sex groups. Such efforts, of course, shall consider the inclusion of multi-ancestry populations with large sample sizes and balanced sex ratios. An ongoing endeavor to meet such requirements is the MESARGEN consortium,² an international framework aimed to include multiple ancestry populations and large sample sizes for genetics and genomics studies of sarcoidosis.

Limitations of the study

We used the Illumina ImmunoChip version 2 platform in our analysis—a customized design based on investigating major autoimmune and inflammatory diseases (88). Therefore, we limited our investigations to immune-mediated genomic regions, leaving out other potential loci implicated in sarcoidosis susceptibility. Despite this limitation, we offer substantial evidence that the genetic susceptibility of sarcoidosis and LS and non-LS phenotypes in males and females shared common

genomic loci but also differed in their genetic architectures given the sex group. We observed these differences in gene-based analysis, eQTLs enrichment across different tissues, and pathway analysis. Further investigations with multiple ancestry populations and large sample sizes are needed to unveil discoveries.

Perspectives and significance

In this work, we revealed sex differences in the genetic architecture of sarcoidosis, particularly in clinical phenotypes LS and non-LS. Our work provides further knowledge that the genetic architectures of LS and non-LS are distinct, and that biological sex plays a role in determining the underlying genetic architecture.

Data availability statement

Access to genomic data is limited due to GDPR restrictions because the data contain personal or other sensitive information and cannot be deposited in public databases. All summary statistics from the analyses are accessible via DOI: 10.6084/m9.figshare.21785273.

Ethics statement

This study was approved by the Stockholm Ethical Review Board, Stockholm, Sweden. The patients/participants provided their written informed consent to participate in this study.

Author contributions

NR and LP designed and implemented the study. YX conducted all the genetic analyses in the Swedish cohort and a meta-analysis among all cohorts. YX and NR drafted the manuscript. SK, JG, and AE recruited patients, obtained samples, and clinically evaluated patients for phenotype characterization in the Swedish cohort. DE conducted genetic analyses in the German cohort. SS and JM-Q recruited patients, obtained samples, and clinically evaluated patients for phenotype characterization in the German cohort. BR and MI recruited patients and obtained samples. CM provided genotype data and oversaw analyses conducted by LG and NP of the US-AA cohort. All authors revised the manuscript and provided feedback.

Funding

This work was funded by the Swedish Heart-Lung Foundation awarded to NR (grant nos. 20170664, 20200505, and 20200506), SK (grant no. 20200163), and JG (grant no. 20190478); Karolinska Institutet Foundation awarded to NR (grant no. FS-2018:0007); and Swedish Research Council awarded to LP (grant no. 2018-02884) and JG (grant no. 2019-01034). This work was also funded

² <http://mesargen.org>

by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through grant FI 1935/1-1. The work received further infrastructure support from the DFG under Germany's Excellence Strategy—EXC 2167-390884018 and the PopGen Biobank (Kiel, Germany; Ref Nr. 2018-032). The popgen 2.0 network (P2N) was supported by a grant from the German Federal Ministry of Education and Research (01EY1103). The US-AA sarcoidosis data and specimen collection are derived from the Trans-Omics in Precision Medicine (TOPMed) program was supported by the National Heart, Lung and Blood Institute (NHLBI). Genome Sequencing for “NHLBI TOPMed: African-American Sarcoidosis Genetics Resource” (phs001207.v3.p1) was performed at Baylor College of Medicine Human Genome Sequencing Center (3R01HL113326-04S1); Northwest Genomics Center (HHSN268201600032I); and Broad Institute Genomics Platform (HHSN268201600034I). Core support including centralized genomic read mapping and genotype calling, along with variant quality metrics and filtering were provided by the TOPMed Informatics Research Center (3R01HL-117626-02S1; contract HHSN268201800002I). Core support including phenotype harmonization, data management, sample-identity QC, and general program coordination were provided by the TOPMed Data Coordinating Center (R01HL-120393; U01HL-120393; contract HHSN268201800001I). We gratefully acknowledge the studies and participants who provided biological samples and data for TOPMed. The US-AA cohort investigators are funded via grants from the Foundation for Sarcoidosis Research (Chicago, IL) and the National Institutes of Health (R01HL113326, P30 GM110766, and U54GM104938-06). The computations and genomic data handling were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at the Uppsala Multidisciplinary Center for Advanced Computational Science

(UPPMAX), partially funded by the Swedish Research Council through grant agreement no. 2018-05973.

Acknowledgments

We thank all patients and the personnel involved in this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmed.2023.1132799/full#supplementary-material>

References

- Valeyre D, Prasse A, Nunes H, Uzunhan Y, Brillet P, Müller-Quernheim J. Sarcoidosis. *Lancet*. (2014) 383:1155–67. doi: 10.1016/S0140-6736(13)60680-7
- Sauer W, Stern B, Baughman R, Culver D, Royal W. High-risk sarcoidosis: current concepts and research imperatives. *Ann Am Thorac Soc*. (2017) 14(Suppl_6):S437–44. doi: 10.1513/AnnalsATS.201707-566OT
- Birnbaum A, Rifkin L. Sarcoidosis: sex-dependent variations in presentation and management. *J Ophthalmol*. (2014) 2014:236905. doi: 10.1155/2014/236905
- Thomeer M, Mateyo K. Sarcoidosis around the Globe. *Semin Respir Crit Care Med*. (2017) 38:393–403. doi: 10.1055/s-0037-1602845
- Ma X, Zhu L, Kurche J, Xiao H, Dai H, Wang C. Global and regional burden of interstitial lung disease and pulmonary sarcoidosis from 1990 to 2019: results from the Global Burden of Disease study 2019. *Thorax*. (2022) 77:596–605. doi: 10.1136/thoraxjnl-2020-216732
- Arkema E, Grunewald J, Kullberg S, Eklund A, Asklund J. Sarcoidosis incidence and prevalence: a nationwide register-based assessment in Sweden. *Eur Respir J*. (2016) 48:1690–9. doi: 10.1183/13993003.00477-2016
- Sikjaer M, Hilberg O, Ibsen R, Lokke A. Sarcoidosis: a nationwide registry-based study of incidence, prevalence and diagnostic work-up. *Respir med*. (2021) 187:106548. doi: 10.1016/j.rmed.2021.106548
- Ungprasert P, Crowson C, Matteson E. Influence of gender on epidemiology and clinical manifestations of sarcoidosis: a population-based retrospective cohort study 1976–2013. *Lung*. (2017) 195:87–91. doi: 10.1007/s00408-016-9952-6
- Singha A, Kirkland M, Drake W, Crouser ED. The influence of age and sex in sarcoidosis. *Curr Opin Pulm Med*. (2022) 28:307–13. doi: 10.1097/MCP.0000000000000882
- Lundkvist A, Kullberg S, Arkema E, Cedelund K, Eklund A, Grunewald J, et al. Differences in disease presentation between men and women with sarcoidosis: a cohort study. *Respir Med*. (2022) 191:106688. doi: 10.1016/j.rmed.2021.106688
- James W, Koutroumpakis E, Saha B, Nathani A, Saavedra L, Yucler R, et al. Clinical features of extrapulmonary sarcoidosis without lung involvement. *Chest*. (2018) 154:349–56. doi: 10.1016/j.chest.2018.02.003
- Duvall C, Pavlovic N, Rosen N, Wand A, Griffin J, Okada D, et al. Sex differences in presentation and outcomes of cardiac sarcoidosis. *J Cardiac Fail*. (2022) 28:S16. doi: 10.1016/j.cardfail.2022.03.045
- Williamson K, Rosenbaum A, Kapa S, Blauwet L. Sex differences in cardiac sarcoidosis. *J Am Coll Cardiol*. (2020) 75(11_Supplement_1):999. doi: 10.1016/S0735-1097(20)31626-0
- Porter N, Beynon H, Randevara H. Endocrine and reproductive manifestations of sarcoidosis. *QJM*. (2003) 96:553–61. doi: 10.1093/qjmed/hcg103
- Singha A, Kirkland M, Drake W, Crouser ED. The influence of age and sex in sarcoidosis. *Curr Opin Pulm Med*. (2022) 28:307–13.
- Hadid V, Patenaude V, Oddy L, Abenheim H. Sarcoidosis and pregnancy: obstetrical and neonatal outcomes in a population-based cohort of 7 million births. *J Perinatal Med*. (2015) 43:201–7. doi: 10.1515/jpm-2014-0017
- Cozier Y, Berman J, Palmer J, Boggs D, Wise L, Rosenberg L. Reproductive and hormonal factors in relation to incidence of sarcoidosis in US Black women: the black women's health study. *Am J Epidemiol*. (2012) 176:635–41. doi: 10.1093/aje/kws145
- Grunewald J, Grutters J, Arkema E, Saketkoo L, Moller D, Muller-Quernheim J. Sarcoidosis. *Nat Rev Dis Primers*. (2019) 5:45. doi: 10.1038/s41572-019-0096-x

19. Grunewald J, Eklund A. Löfgren's syndrome. *Am J Respir Crit Care Med.* (2009) 179:307–12. doi: 10.1164/rccm.200807-1082OC
20. Patterson K, Strek M. Pulmonary fibrosis in sarcoidosis. Clinical features and outcomes. *Ann Am Thorac Soc.* (2013) 10:362–70. doi: 10.1513/AnnalsATS.201303-069FR
21. Nardi A, Brillet P, Letoumelin P, Girard F, Brauner M, Uzunhan Y, et al. Stage IV sarcoidosis: comparison of survival with the general population and causes of death. *Eur Respir J.* (2011) 38:1368–73. doi: 10.1183/09031936.00187410
22. Jeny F, Uzunhan Y, Lacroix M, Gille T, Brillet P, Nardi A, et al. Predictors of mortality in fibrosing pulmonary sarcoidosis. *Respir Med.* (2020) 169:105997. doi: 10.1016/j.rmed.2020.105997
23. Rivera N, Ronninger M, Shchetynsky K, Franke A, Nothen M, Muller-Quernheim J, et al. High-density genetic mapping identifies new susceptibility variants in sarcoidosis phenotypes and shows genomic-driven phenotypic differences. *Am J Respir Crit Care Med.* (2016) 193:1008–22. doi: 10.1164/rccm.201507-1372OC
24. Rybicki B, Levin A, McKeigue P, Datta I, Gray-McGuire C, Colombo M, et al. A genome-wide admixture scan for ancestry-linked genes predisposing to sarcoidosis in African-Americans. *Genes Immun.* (2011) 12:67–77. doi: 10.1038/gene.2010.56
25. Klareskog L, Stolt P, Lundberg K, Kallberg H, Bengtsson C, Grunewald J, et al. A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA-DR (shared epitope)-restricted immune reactions to autoantigens modified by citrullination. *Arthritis Rheum.* (2006) 54:38–46. doi: 10.1002/art.21575
26. Hedstrom A, Baarnhielm M, Olsson T, Alfredsson L. Tobacco smoking, but not Swedish snuff use, increases the risk of multiple sclerosis. *Neurology.* (2009) 73:696–701. doi: 10.1212/WNL.0b013e3181b59c40
27. Statement on sarcoidosis. Joint Statement of the American Thoracic Society (ATS), the European Respiratory Society (ERS) and the World Association of Sarcoidosis and Other Granulomatous Disorders (WASOG) adopted by the ATS Board of Directors and by the ERS Executive Committee. *Am J Respir Crit Care Med.* (1999) 160:736–55.
28. Valentonyte R, Hampe J, Huse K, Rosenstiel P, Albrecht M, Stenzel A, et al. Sarcoidosis is associated with a truncating splice site mutation in BTNL2. *Nat Genet.* (2005) 37:357–64. doi: 10.1038/ng1519
29. Hofmann S, Fischer A, Nothnagel M, Jacobs G, Schmid B, Wittig M, et al. Genome-wide association analysis reveals 12q13.3-q14.1 as new risk locus for sarcoidosis. *Eur Respir J.* (2013) 41:888–900. doi: 10.1183/09031936.00033812
30. Krawczak M, Nikolaus S, von Eberstein H, Croucher P, El Mokhtari N, Schreiber S. PopGen: population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. *Commun Genet.* (2006) 9:55–61. doi: 10.1159/000090694
31. Schmermund A, Mohlenkamp S, Stang A, Gronemeyer D, Seibel R, Hirche H, et al. Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: rationale and design of the Heinz Nixdorf RECALL Study. Risk factors, evaluation of coronary calcium and lifestyle. *Am Heart J.* (2002) 144:212–8. doi: 10.1067/mhj.2002.123579
32. Design of a case control etiologic study of sarcoidosis (ACCESS). ACCESS Research Group. *J Clin Epidemiol.* (1999) 52(12):1173–86.
33. Rybicki B, Hirst K, Iyengar S, Barnard J, Judson M, Rose C, et al. A sarcoidosis genetic linkage consortium: the sarcoidosis genetic analysis (SAGA) study. *Sarcoidosis Vasc Diffuse Lung Dis.* (2005) 22:115–22.
34. Iannuzzi M, Malinarik M, Poisson L, Rybicki B. Sarcoidosis susceptibility and resistance HLA-DQB1 alleles in African Americans. *Am J Respir Crit Care Med.* (2003) 167:1225–31. doi: 10.1164/rccm.200209-1097OC
35. Rasmussen A, Sevier S, Kelly J, Glenn S, Aberle T, Cooney C, et al. The lupus family registry and repository. *Rheumatology.* (2011) 50:47–59. doi: 10.1093/rheumatology/keq302
36. Eyre S, Bowes J, Diogo D, Lee A, Barton A, Martin P, et al. High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. *Nat Genet.* (2012) 44:1336–40. doi: 10.1038/ng.2462
37. Padyukov L, Seielstad M, Ong R, Ding B, Ronnelid J, Seddighzadeh M, et al. A genome-wide association study suggests contrasting associations in ACPA-positive versus ACPA-negative rheumatoid arthritis. *Ann Rheum Dis.* (2011) 70:259–65. doi: 10.1136/ard.2009.126821
38. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira M, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* (2007) 81:559–75. doi: 10.1086/519795
39. Chang C, Chow C, Tellier L, Vattikuti S, Purcell S, Lee J. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience.* (2015) 4:7. doi: 10.1186/s13742-015-0047-8
40. Fischer A, Ellinghaus D, Nutsua M, Hofmann S, Montgomery C, Iannuzzi M, et al. Identification of immune-relevant factors conferring sarcoidosis genetic risk. *Am J Respir Crit Care Med.* (2015) 192:727–36. doi: 10.1164/rccm.201503-0418OC
41. Adrianto I, Lin C, Hale J, Levin A, Datta I, Parker R, et al. Genome-wide association study of African and European Americans implicates multiple shared and ethnic specific loci in sarcoidosis susceptibility. *PLoS One.* (2012) 7:e43907. doi: 10.1371/journal.pone.0043907
42. Price A, Patterson N, Plenge R, Weinblatt M, Shadick N, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet.* (2006) 38:904–9. doi: 10.1038/ng1847
43. Steve E, John B, Dorothee D, Annette L, Anne B, Paul M, et al. High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. *Nat Genet.* (2012) 44:1336.
44. Beecham A, Patsopoulos N, Xifara D, Davis M, Kempainen A, Cotsapas C, et al. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nat Genet.* (2013) 45:1353. doi: 10.1038/ng.2770
45. David E, Hansjörg B, Jorge E, Elke R, Anja M, Ingo M, et al. High-density genotyping study identifies four new susceptibility loci for atopic dermatitis. *Nat Genet.* (2013) 45:808.
46. Carmona F, Mackie Sarah I, Martin J, Taylor John C, Vaglio A, Eyre S, et al. A large-scale genetic analysis reveals a strong contribution of the HLA class II region to giant cell arteritis susceptibility. *Am J Hum Genet.* (2015) 96:565–80.
47. Viechtbauer W. Conducting meta-analyses in R with the metafor Package. *J Stat Softw.* (2010) 36:1–48. doi: 10.18637/jss.v036.i03
48. Willer C, Li Y, Abecasis G. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics.* (2010) 26:2190–1. doi: 10.1093/bioinformatics/btq340
49. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* (2015) 12:e1001779. doi: 10.1371/journal.pmed.1001779
50. De Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: Generalized gene-set analysis of GWAS data. *PLoS Comput Biol.* (2015) 11:e1004219.
51. Mishra A, Macgregor S. VEGAS2: software for more flexible gene-based testing. *Twin Res Hum Genet.* (2015) 18:86–91. doi: 10.1017/thg.2014.79
52. Watanabe K, Taskesen E, Van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun.* (2017) 8:1826. doi: 10.1038/s41467-017-01261-5
53. Ferreira M, Jansen R, Willemsen G, Penninx B, Bain L, Vicente C, et al. Gene-based analysis of regulatory variants identifies 4 putative novel asthma risk genes related to nucleotide synthesis and signaling. *J Allergy Clin Immunol.* (2017) 139:1148–57. doi: 10.1016/j.jaci.2016.07.017
54. Franke L, Jansen R. eQTL analysis in humans. *Methods Mol Biol.* (2009) 573:311. doi: 10.1007/978-1-60761-247-6_17
55. Brown F, Modi P, Tanner LS. *Lofgren syndrome*. StatPearls. Treasure Island (FL): StatPearls Publishing Copyright © 2022, StatPearls Publishing LLC. (2022).
56. Horton R, Wilming L, Rand V, Lovering R, Bruford E, Khodiyar V, et al. Gene map of the extended human MHC. *Nat Rev Genet.* (2004) 5:889–99. doi: 10.1038/nrg1489
57. Sinha R, Siddiqui T, Padmanabhan N, Wallin J, Zhang C, Karimi B, et al. LRRTM4: a novel regulator of presynaptic inhibition and ribbon synapse arrangements of retinal bipolar cells. *Neuron.* (2020) 105:1007–17.e5. doi: 10.1016/j.neuron.2019.12.028
58. Paladini F, Fiorillo M, Tedeschi V, Mattorre B, Sorrentino R. The multifaceted nature of aminopeptidases ERAP1, ERAP2, and LNPEP: from evolution to disease. *Front Immunol.* (2020) 11:1576. doi: 10.3389/fimmu.2020.01576
59. Gerke A, Judson M, Cozier Y, Culver D, Koth L. Disease burden and variability in sarcoidosis. *Ann Am Thorac Soc.* (2017) 14(Suppl_6):S421–8. doi: 10.1513/AnnalsATS.201707-564OT
60. Pandit P, Perez R, Roman J. Sex-based differences in interstitial lung disease. *Am J Med Sci.* (2020) 360:467–73. doi: 10.1016/j.amjms.2020.04.023
61. Sharp M, Psoter K, Balasubramanian A, Pulapaka A, Chen E, Brown S, et al. Heterogeneity of lung function phenotypes in sarcoidosis: role of race and sex differences. *Ann Am Thorac Soc.* (2023) 20:30–7. doi: 10.1513/AnnalsATS.202204-328OC
62. Qiu J, Moore J, Darabos C. Studying the genetics of complex disease with ancestry-specific human phenotype networks: the case of type 2 diabetes in East Asian populations. *Genet Epidemiol.* (2016) 40:293–303. doi: 10.1002/gepi.21964
63. Dries D. Genetic ancestry, population admixture, and the genetic epidemiology of complex disease. *Circulation.* (2009) 2:540–3. doi: 10.1161/CIRCGENETICS.109.922898
64. Fairfax B, Makino S, Radhakrishnan J, Plant K, Leslie S, Dilthey A, et al. Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nat Genet.* (2012) 44:502–10. doi: 10.1038/ng.2205

65. Vösa U, Claringbould A, Westra H, Bonder M, Deelen P, Zeng B, et al. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat Genet.* (2021) 53:1300–10.
66. Tian J, Keller M, Broman A, Kendziorski C, Yandell B, Attie A, et al. The dissection of expression quantitative trait locus hotspots. *Genetics.* (2016) 202:1563–74. doi: 10.1534/genetics.115.183624
67. Kaiser Y, Eklund A, Grunewald J. Moving target: shifting the focus to pulmonary sarcoidosis as an autoimmune spectrum disorder. *Eur Respir J.* (2019) 54.
68. Starshinova A, Malkova A, Basantsova N, Zinchenko Y, Kudryavtsev I, Ershov G, et al. Sarcoidosis as an autoimmune disease. *Front Immunol.* (2020) 10:2933. doi: 10.3389/fimmu.2019.02933
69. Song M, Manansala M, Parmar P, Ascoli C, Rubinstein I, Sweiss N. Sarcoidosis and autoimmunity. *Curr Opin Pulm Med.* (2021) 27:448–54. doi: 10.1097/MCP.0000000000000809
70. Rossman MD, Kreider ME. Is chronic beryllium disease sarcoidosis of known etiology? *Sarcoidosis Vasc Diffuse Lung Dis.* (2003) 20(2):104–9.
71. Sharma SK, Balamurugan A, Pandey RM, Saha PK, Mehra NK. Human leukocyte antigen-DR alleles influence the clinical course of pulmonary sarcoidosis in Asian Indians. *Am J Respir Cell Mol Biol.* (2003) 29(2):225–31.
72. Darlington P, Tallstedt L, Padyukov L, Kockum I, Cederlund K, Eklund A, et al. HLA-DRB1* alleles and symptoms associated with Heerfordt's syndrome in sarcoidosis. *Eur Respir J.* (2011) 38:1151–57.
73. Zhou Y, Shen L, Zhang Y, Jiang D, Li H. Human leukocyte antigen-A, -B, and -DRB1 alleles and sarcoidosis in Chinese Han subjects. *Hum Immunol.* (2011) 72:571–5.
74. Wennerstrom A, Pietinalho A, Vauhkonen H, Lahtela L, Palikhe A, Hedman J, et al. HLA-DRB1 allele frequencies and C4 copy number variation in Finnish sarcoidosis patients and associations with disease prognosis. *Hum Immunol.* (2012) 73:93–100.
75. van Moorsel CHM, Petrek M, Rivera, NV. Unravelling the genetic basis of sarcoidosis. *Eur Respir Monogr.* (2022) 96:41–56.
76. Rubtsova K, Marrack P, Rubtsov A. Sexual dimorphism in autoimmunity. *J Clin Invest.* (2015) 125:2187–93. doi: 10.1172/JCI78082
77. Korsten P, Tampe B, König MF, Nikiphorou E. Sarcoidosis and autoimmune diseases: differences, similarities and overlaps. *Curr Opin Pulm Med.* (2018) 24(5):504–12.
78. Hofmann S, Franke A, Fischer A, Jacobs G, Nothnagel M, Gaede KI. Genome-wide association study identifies ANXA11 as a new susceptibility locus for sarcoidosis. *Nat Genet.* (2008) 40:1103–6. doi: 10.1038/ng.198
79. Levin AM, Iannuzzi MC, Montgomery CG, Trudeau S, Datta I, McKeigue P, et al. Association of ANXA11 genetic variation with sarcoidosis in African Americans and European Americans. *Genes Immun.* (2013) 14:13–8.
80. Iaccarino L, Ghirardello A, Canova M, Zen M, Bettio S, Nalotto L, et al. Anti-annexins autoantibodies: their role as biomarkers of autoimmune diseases. *Autoimmun Rev.* (2011) 10:553–8. doi: 10.1016/j.autrev.2011.04.007
81. Moss, SE, Morgan, RO. The annexins. *Genome Biol.* (2004) 5:219. doi: 10.1186/gb-2004-5-4-219
82. Bruschi M, Petretto A, Vaglio A, Santucci L, Candiano G, Ghiggeri G. Annexin A1 and autoimmunity: from basic science to clinical applications. *Int J Mol Sci.* (2018) 19:1348.
83. Castro F, Cardoso A, Gonçalves R, Serre K, Oliveira M. Interferon-gamma at the crossroads of tumor immune surveillance or evasion. *Front Immunol.* (2018) 9:847. doi: 10.3389/fimmu.2018.00847
84. Dempsey O, Paterson E, Kerr K, Denison A. Sarcoidosis. *BMJ.* (2009) 339:b3206. doi: 10.1136/bmj.b3206
85. Miedema J, Kaiser Y, Broos C, Wijsenbeek M, Grunewald J, Kool M. Th17-lineage cells in pulmonary sarcoidosis and Lofgren's syndrome: friend or foe? *J Autoimmun.* (2018) 87:82–96. doi: 10.1016/j.jaut.2017.12.012
86. Grunewald J, Eklund A. Sex-specific manifestations of Lofgren's syndrome. *Am J Respir Crit Care Med.* (2007) 175:40–4. doi: 10.1164/rccm.200608-1197OC
87. Fischer A, Valentonyte R, Nebel A, Nothnagel M, Müller-Quernheim J, Schürmann M, et al. Female-specific association of C-C chemokine receptor 5 gene polymorphisms with Lofgren's syndrome. *J Mol Med.* (2008) 86:553–61. doi: 10.1007/s00109-008-0315-5
88. Cortes A, Brown M. Promise and pitfalls of the immunochip. *Arthritis Res Ther.* (2011) 13:101. doi: 10.1186/ar3204