Clinical Immunology 279 (2025) 110540

Contents lists available at ScienceDirect

Clinical Immunology

journal homepage: www.elsevier.com/locate/yclim

Genetic variability associates with ancestry, age at disease onset, organ involvement and disease severity in juvenile-onset systemic lupus erythematosus

Valentina Natoli [a,b,c,1], Amandine Charras [a,1], Megan S.R. Hasoon [d,e,1], Andrea L. Jorgensen [f], Eve M.D. Smith [a,g], Eva Caamaño Gutiérrez [d,e], Michael W. Beresford [a,g], Christian M. Hedrich [a,g,*], on behalf of the UK JSLE Cohort Study Group [2]

[a] Department of Women's & Children's Health, Institute of Life Course and Medical Sciences, University of Liverpool, Liverpool, UK
[b] Dipartimento di Neuroscienze, Riabilitazione, Oftalmologia, Genetica e Scienze Materno-Infantili, Università degli Studi di Genova, Genoa, Italy
[c] Rheumatology and Autoinflammatory Diseases Unit, IRCCS Istituto Giannina Gaslini, Genoa, Italy
[d] Computational Biology Facility, Liverpool Shared Research Facilities, Faculty of Health and Life Sciences, University of Liverpool, Liverpool L69 7ZB, UK
[e] Department of Biochemistry & Systems Biology, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool L69 7ZB, UK
[f] Department of Health Data Science, Institute of Population Health, University of Liverpool, Liverpool, UK
[g] Department of Paediatric Rheumatology, Alder Hey Children's NHS Foundation Trust Hospital, Liverpool, UK

ARTICLE INFO

Keywords:
Lupus
SLE
Genetic
Variability
Ancestry
Phenotype

ABSTRACT

Juvenile-onset systemic lupus erythematosus (jSLE) is a complex autoimmune/inflammatory disease in which genetic factors likely contribute to pathophysiology and clinical expression. This study explored associations between general (alternate allele counts; AAC) and gene-specific (alternate allele scores; GAAS) sequence variability, age at onset, sex, ancestry, disease activity/severity, organ involvement and treatments in jSLE. 289 participants from the UK JSLE Cohort Study underwent panel sequencing of 62 genes/genomic regions. Weighted AAC and GAAS were calculated. Correlation analyses and generalized linear models assessed associations between genetic burden, ancestry, age at diagnosis and clinical variables. AAC inversely correlated with age at diagnosis ($R = -0.15$, $p = 0.01$), primarily driven by South Asians ($R = -0.28$, $p < 0.001$). African/Caribbean patients exhibited higher AAC ($p < 0.001$). Clinical variables, including severity of renal involvement (*ACP5*,

* Corresponding author at: Institute in the Park, Alder Hey Children's NHS Foundation Trust Hospital, East Prescot Road, Liverpool L14 5AB, UK.
E-mail address: christian.hedrich@liverpool.ac.uk (C.M. Hedrich).

[1] These authors contributed equally to the work presented.

[2] Carla Roberts, Department of Women's & Children's Health, Institute of Life Course and Medical Sciences, University of Liverpool.
Eslam Al-Abadi, Department of Rheumatology, Birmingham Children's Hospital, Birmingham.
Kate Armon, Department of Paediatric Rheumatology, Cambridge University Hospitals, Cambridge.
Kathryn Bailey, Department of Paediatric Rheumatology, Oxford University Hospitals NHS Foundation Trust, Oxford.
Coziana Ciurtin, Centre for Adolescent Rheumatology, University College London, London.
Janet Gardner-Medwin, Department of Child Health, University of Glasgow, Glasgow.
Kirsty Haslam, Department of Paediatrics, Bradford Royal Infirmary, Bradford.
Daniel P Hawley, Department of Paediatric Rheumatology, Sheffield Children's Hospital, Sheffield.
Alice Leahy, Department of Paediatric Rheumatology, Southampton General Hospital, Southampton.
Valentina Leone, Department of Paediatric Rheumatology, Leeds Children Hospital, Leeds.
Flora McErlane, Paediatric Rheumatology, Great North Children's Hospital, Royal Victoria Infirmary, Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne.
Gita Modgil, Department of Paediatrics, Musgrove Park Hospital, Taunton.
Clarissa Pilkington, Department of Paediatric Rheumatology, Great Ormond Street Hospital, London.
Athimalaipet V Ramanan, University Hospitals Bristol NHS Foundation Trust & Bristol Medical School, University of Bristol, Bristol.
Satyapal Rangaraj, Department of Paediatric Rheumatology, Nottingham University Hospitals, Nottingham.
Phil Riley, Department of Paediatric Rheumatology, Royal Manchester Children's Hospital, Manchester.
Arani Sridhar, Department of Paediatrics, Leicester Royal Infirmary, Leicester, UK.

https://doi.org/10.1016/j.clim.2025.110540
Received 27 March 2025; Received in revised form 4 June 2025; Accepted 6 June 2025
Available online 15 June 2025

ITGAM, *LYN*, $p < 0.001$; *TNFAIP3*, $p = 0.007$), associated with GAAS. Genetic variability likely contributes to early disease expression and severity in jSLE, supporting patient stratification and personalised care.

1. Introduction

Juvenile-onset systemic lupus erythematosus (jSLE) is a pathophysiologically complex systemic autoimmune/inflammatory disease that can affect any organ [1,2]. It can cause organ damage and failure, significantly impacting on the well-being of children affected [1,2]. Compared to patients with disease-onset during adulthood, children with SLE generally have higher disease activity, greater organ damage, increased mortality, and require more aggressive treatment [3]. Although significant progress has been made, the exact pathophysiology of SLE remains unknown [4], involving both inherited and acquired mechanisms [1,2]. Genetic factors are required to develop SLE, with several studies reporting (ultra-)rare genetic causes of SLE (or lupus-like disease) and common risk alleles [5,6]. Based on observations from previous studies, sub-groups of SLE patients exist which are differentially represented across age groups [4]. "Genetic SLE" affects a small fraction of SLE patients (estimated 1–4 % across age groups) and is caused by single or a combination of gene mutations with high functional impact, most commonly affecting type 1 interferon (IFN) pathways and/or the complement system [1,2,6]. The majority of SLE patients carry gene variants, so-called risk-alleles, that increase an individual's risk for the development of SLE but are not strong enough to confer disease [6]. Additional factors, such as hormonal changes and environmental impact (viral infections, medications, etc.), are necessary to cause disease expression [1,2,4,6].

In most children and young people with SLE, a combination of risk alleles (notably, a higher number than in adult-onset patients) may contribute to early disease onset, its clinical variability, and more severe disease phenotypes than those seen in adults [1,6,7]. This study investigated overall patterns of genetic variability around previously reported SLE-associated genes and risk loci in a multi-ancestral cohort, with a focus on cumulative alternate allele burden, assessing the relationship between alternate allele counts (AAC), age at disease onset, sex, ancestry, organ involvement, clinical severity, and treatments [8]. It furthermore explored associations between gene-level alternate allele scores (GAAS), organ involvement and disease severity.

2. Methods

This study follows the STREGA (STrengthening the REporting of Genetic Association studies) reporting guidelines [9].

2.1. Study cohort

A total of 315 patients from the UK JSLE Cohort Study were initially enrolled in this study [8]. The UK JSLE Cohort Study includes two distinct cohorts: an established JSLE cohort, comprising patients diagnosed with JSLE from 1995 to the present day (retrospective data), and a prospective JSLE cohort, encompassing all newly diagnosed JSLE patients (since 2006) enrolled during the ongoing study period (prospective data). For this specific study, patient data were collected up to September 2022. All participants fulfilled the American College of Rheumatology (ACR) 1997 classification criteria for SLE ($\geq$ 4 items) [10] and were diagnosed before their 18th birthday. We were not able to perform a formal power analysis prior to this study, due to the absence of studies using a comparable design in this specific disease area. Furthermore, sample size was pre-defined accessing a large national cohort in a rare disease area. Written patient assent/consent and/or parental consent was obtained as appropriate. The study received ethical approval from the National Research Ethics Service Northwest (REC_06/Q1502/77). Research was carried out in accordance with the declaration of Helsinki.

2.2. Data collected

Comprehensive patient data were collected and analysed in the following areas: 1) Demographic information, encompassing age at diagnosis, sex, self-reported ancestry, and family history of autoimmune/inflammatory diseases in first-grade relatives (SLE, systemic connective tissue diseases/CTD, rheumatoid arthritis/RA, endocrinopathies); 2) Disease activity, assessed at each study visit using the paediatric version of the British Isles Lupus Assessment Grade (pBILAG)-2004 [11,12] and the Systemic Lupus Erythematosus Disease Activity Index (SLEDAI) score, and damage at the last visit through the Systemic Lupus International Collaborating Clinics (SLICC)/ACR standardised damage index (SDI) [13]; 3) Prospectively collected clinical data, including malar rash, discoid lupus, photosensitivity, oral and/or nasal ulcers, non-erosive arthritis, serositis, nephritis, neurological, haematological, and/or immunological involvement, as per the 1997 ACR criteria for SLE; and 4) treatments. Patients were categorised in the "intensive treatment" group if, throughout the disease course, they received either cyclophosphamide, rituximab or belimumab, and/or two or more conventional disease-modifying anti-rheumatic drugs (DMARDs), excluding hydroxychloroquine, simultaneously for $\geq$2 consecutive visits, and in the "non-intensive treatment" group, when a maximum of one conventional DMARD, including hydroxychloroquine, was used at a time (Table 1, Supplementary Table 1). Ancestry inference was not performed in this study primarily because the available data were limited to 62 genes/genomic target regions (exonic regions, exon:intron junctions), which represent only a small fraction of the genome and are typically under strong functional constraint. A more detailed rationale is provided in the discussion.

Considering all patient visits, we assigned a "severity" value to each pBILAG organ/system domain based on the highest score ever recorded: "severe involvement" was attributed to any organ or system that received a score of 'A' at any visit, "moderate involvement" was assigned if the domain's highest recorded score was 'B', "mild involvement" was determined if the most severe score ever given was 'C'. Lastly, "never involved" was the designation for any organ or system that received an 'E' score across all visits.

2.3. Gene panel selection

Sequencing targets were selected based on a literature search (2018) targeting: 1) genes associated with known Mendelian forms of SLE/SLE-like disease, and 2) SLE-associated risk alleles previously identified through Genome-wide association studies (GWAS). As a result, 62 genes/genomic regions were selected for sequencing, including exons and exon:intron junctions, as well as previously reported SLE-associated risk alleles, as described previously [14].

2.4. Target sequencing and variant identification

Sequence capture probes (NimbleGen/Roche) were designed to target exonic regions and exon:intron junctions of pre-selected genes (Supplementary Table 2). Sequencing libraries were prepared from genomic DNA, hybridized to the probes and then sequenced with 150 bp paired-end reads using Illumina MiSeq technology (Illumina). Demultiplexing, adaptor and quality trimming (Cutadapt v1.2.1, Sickle v1.2) of reads was performed [15,16]. Polymerase Chain Reaction (PCR) duplicates were identified and excluded from the dataset using Picard [17]. Sequencing data were aligned to the human reference genome (hg38)

Table 1
Demographic and clinical information on the study cohort.

	JSLE patients included in the alternate allele count analysis (n = 238)	JSLE patients included in the gene-level score analysis (n = 289)
Age at diagnosis, years (median [IQR])	13.0 [10.7–14.5]	12.8 [10.3–14.5]
Sex, n (%)		
Female	198 (83.2)	242 (83.7)
Male	40 (16.8)	47 (16.3)
Ancestry, n (%)		
European	114 (47.9)	138 (47.7)
African/Caribbean	38 (16.0)	48 (16.6)
South Asian	57 (23.9)	70 (24.2)
East Asian	9 (3.8)	10 (3.5)
Other Asian	20 (8.4)	23 (8.0)
Family history autoimmune diseases, n (%)	48 (20.2)	53 (18.3)
SLE	18 (37.5)	21 (39.6)
Systemic CTD	22 (45.8)	25 (47.2)
Endocrinopathies	19 (39.6)	21 (39.6)
RA	11 (22.9)	12 (22.6)
Highest SLEDAI score during follow-up (median [IQR])	20 [17–24]	20 [17–24]
Highest numerical pBILAG score during follow-up (median [IQR])	21 [13−32]	21 [13−31]
SLICC SDI at last visit (median [IQR])	0 [0–1]	0 [0–1]
Treatments, n (%)*		
Intensive	109 (45.8)	129 (44.6)
Non-intensive	129 (54.2)	160 (55.4)

JSLE, juvenile systemic lupus erythematosus; IQR, interquartile range; CTD, connective tissue disease; RA, rheumatoid arthritis; SLEDAI, Systemic Lupus Erythematosus Disease Activity Index; pBILAG, paediatric British Isles Lupus Assessment Grade 2004; SLICC SDI, SLICC/ACR Damage Index.

* "Intensive treatment": patients, throughout the disease course, received either cyclophosphamide, rituximab or belimumab, and/or two or more conventional DMARDs, excluding hydroxychloroquine, simultaneously for ≥2 consecutive visits. "Non-intensive treatment": a maximum of one conventional DMARD, including hydroxychloroquine, was used at a time.

using bwa [18] and variants were subsequently detected with Genome Analysis Toolkit (GATK) Software [18,19]. Variant calling was performed jointly across samples using the GATK Haplotype Caller. GATK base quality score recalibration was applied. Variants were then filtered using the GATK Variant Filtration tool. Single nucleotide polymorphisms (SNPs) were removed if they met any of the following criteria: Quality by Depth (QD) < 5.0, Quality Score (QUAL) < 30.0, Strand Odds Ratio (SOR) > 3.0, (Fisher Strand) FS > 60.0, Mapping Quality (MQ) < 40.0, Mapping Quality Rank Sum (MQRankSum) < −12.5, Read Position Rank Sum (ReadPosRankSum) < −8.0. Variants passing the filtering thresholds were annotated using SNP Effect (SnpEff) [20] and database of SNP (dbSNP) [21]. Finally, data were extracted from the Variant Call Format (vcf) file using the R packages *VariantAnnotation* [22] and *vcfR* [23]. Human Leukocyte Antigen (HLA) class I regions were excluded from this project, as the high genetic variability within these regions undermines the accuracy of standard variant calling methods. Genotype (GT) data were extracted and converted into a matrix where genotypes were numerically encoded based on alternate allele counts: homozygous reference (0/0) = 0, heterozygous (0/1 or 1/0) = 1, and homozygous alternate (1/1) = 2. Missing genotypes (./.) were excluded from downstream analysis.

2.5. Alternate allele counts

Genotype calls were converted to reflect the number of alternate alleles per SNP (0, 1, or 2), and samples with missing genotype data across any of the 4100 exonic SNPs were excluded to ensure complete-case analysis. This meant 238/289 (82.4 %) patients were retained for this analysis as they had genotype information for all 4100 SNPs contained in the genome fractions studied and complete clinical information.

Functional annotations (ANN), extracted from the INFO field of the VCF file, were used to assign biological impact weights: MODIFIER and LOW = 1, MODERATE = 2, and HIGH = 3. For each individual, the alternate allele count at each SNP was multiplied by its corresponding ANN weight, and these values were summed to produce a weighted AAC score. The weighted AAC scores were merged with the following 21 demographic and clinical variables: age at diagnosis; severity of constitutional, mucocutaneous, neuropsychiatric, musculoskeletal, cardiorespiratory, gastrointestinal, ophthalmic, renal and haematological involvement (pBILAG domains); presence of malar rash, discoid lupus, photosensitivity, oral or nasal ulcers, non-erosive arthritis, serositis, nephritis, neurologic, haematological and immunological disorders (1997 ACR classification criteria) and treatment received ("non-intensive" versus "intensive"). As mentioned above, for each pBILAG organ/system domain, patients were stratified into four mutually exclusive severity categories based on the highest score ever recorded during follow-up: "severe involvement" (maximum score = A), "moderate involvement" (maximum score = B), "mild involvement" (maximum score = C), and "never involved" (only E scores across all visits). For the 1997 ACR classification criteria, patients were categorised as "yes" (feature ever present) or "no" (feature never present) (Supplementary Table 1).

To identify associations between AAC and clinical phenotypes, generalized linear models (GLMs) with a Poisson distribution were used. Each model included ancestry, sex, and family history of autoimmune disease as covariates. *P*-values from the GLMs were adjusted for multiple testing using the Benjamini-Hochberg false discovery rate (FDR) method. In addition, Pearson correlation analyses were performed to evaluate relationships between AAC and continuous clinical outcomes, including age at diagnosis, the highest SLEDAI and numeric BILAG scores ever recorded during follow-up, and the SLICC-SDI at the latest visit. Correlation tests were also stratified by ancestry to explore subgroup-specific associations, with FDR correction applied across tests. All methods can be consulted in detail and fully reproduced in the publicly available R Markdown document https://github.com/CBFLivUni/jSLE_paper/blob/main/scripts/AAC_calculation_and_analysis.Rmd.

2.6. Gene-level alternate allele scores

For each SNP and individual, a mutation burden score was calculated by multiplying the alternate allele count by the assigned annotation severity value (as described above). Mutation burden scores were then aggregated at the gene level by summing scores across all variants within a gene for each individual. This resulted in a matrix of 62 Gene-level alternate alelle scores (GAAS) for all patients, where rows corresponded to individuals and columns to genes. Note that SNPs present across all patients were 3764 of the sequenced 4100 SNPs (91.8 %), and the total number of patients included was 289 (those with complete clinical data). To visualize sample stratification, principal component analysis (PCA) was performed on the transposed gene-level score matrix using the prcomp() function. The first two principal components (PC1 and PC2) were visualized using ggplot2, with patient samples coloured by ancestry. To identify gene-level associations with clinical features, generalized linear models (GLMs) with a Poisson distribution were fitted for each gene against a panel of 21 curated clinical variables, corresponding to the same set of demographic and clinical features used in the AAC analyses. Each model included additional covariates to adjust for ancestry, sex, and family history of autoimmune diseases. Genes previously identified as pseudogenes, antisense transcripts, or non-coding RNAs (n = 26) were excluded from testing. Statistical significance of the clinical predictor was assessed using chi-square tests from the model

ANOVA. *P*-values were corrected for multiple testing using the Benjamini and Hochberg method. Models with adjusted *p*-values <0.05 were considered significant. All model estimates, standard errors, and p-values were exported into structured Excel spreadsheets for review and downstream interpretation. For full transparency and reproducibility, the results are provided in Supplementary Table 3 and can be reproduced following the publicly available script https://github.com/CBFLivUni/jSLE_paper/blob/main/scripts/GAAS_caculation_and_analysis.Rmd.

2.7. Comparison with alternate allele counts generated from previously published SLE-associated SNPs

To compare findings from this study with previously reported genetic variants, a curated list of 330 genome-wide significant lupus-predisposing SNPs was extracted from a recent comprehensive literature review [24]. These SNPs were then cross-referenced with those sequenced in our dataset. The intersection of common SNPs, named "subset AAC", was used to undertake the same AAC methodology described above. This subset contained only 15 overlapped SNPs (Supplementary Table 4). Notably, many of the previously reported SNPs: 1) mapped to genomic regions (e.g. intronic or intergenic) that were not covered by the sequencing panel used in this study, 2) had been reported exclusively in East Asian cohorts and were therefore excluded in this study due to the low representation of East Asian patients in the UK JSLE cohort, and 3) were identified after 2018, when the panel for this study was designed. The correlation between the original AAC and the subset AAC was evaluated, along with the relationship between the subset AAC and diagnosis age across ethnic groups. Due to the limited number of overlapping SNPs between this dataset and the previously reported SNPs, gene-level analyses for the subset AAC could not be performed.

3. Results

3.1. Cohort characteristics

Of the 315 jSLE patients sequenced, 289 (91.7 %) were included in the GAAS analysis by excluding SNPs lacking sequencing data across all patients as well as patients with incomplete demographic/clinical information. For the AAC analysis, only patients with complete sequencing information for all 4100 SNPs contained in the genome portions studied were included, resulting in 238/289 (82.4 %) patients being analysed (Fig. 1A).

Demographic and clinical characteristics of the two sub-cohorts were comparable (Table 1; Supplementary Table 1). The subsequent description of demographics focuses on the larger cohort subjected to the GAAS analysis.

The median age at diagnosis in this cohort was 12.8 years (interquartile range/IQR 10.3–14.5) (Fig. 1B), and most patients were female (83.7 %). Patients of European ancestry represented 47.7 % of the study population; 24.2 % were South Asian, 16.6 % African/Caribbean, 3.5 % East Asian, and 8.0 % were of "other" Asian ancestry (Fig. 1B). 18 % of participants had a family history of autoimmune/inflammatory diseases in first-degree relatives, including SLE (39.6 %), systemic CTD (47.2 %), endocrinopathies (39.6 %) and RA (22.6 %). Notably, damage measured by SLICC-SDI was low across the cohort, with a median score of 0 (IQR 0–1). Almost half of the participants (44.6 %) received "intensive" treatment throughout their disease course (Table 1).

3.2. AAC inversely correlate with age at diagnosis

Based on previous studies suggesting that jSLE patients with early disease onset experience a higher genetic burden [25], we investigated the possible relationship between AAC and age at diagnosis. A weak inverse correlation between age at diagnosis and AAC was noted ($R = -0.15$, $p = 0.012$, Fig. 2A). Correlation analyses within each ancestral

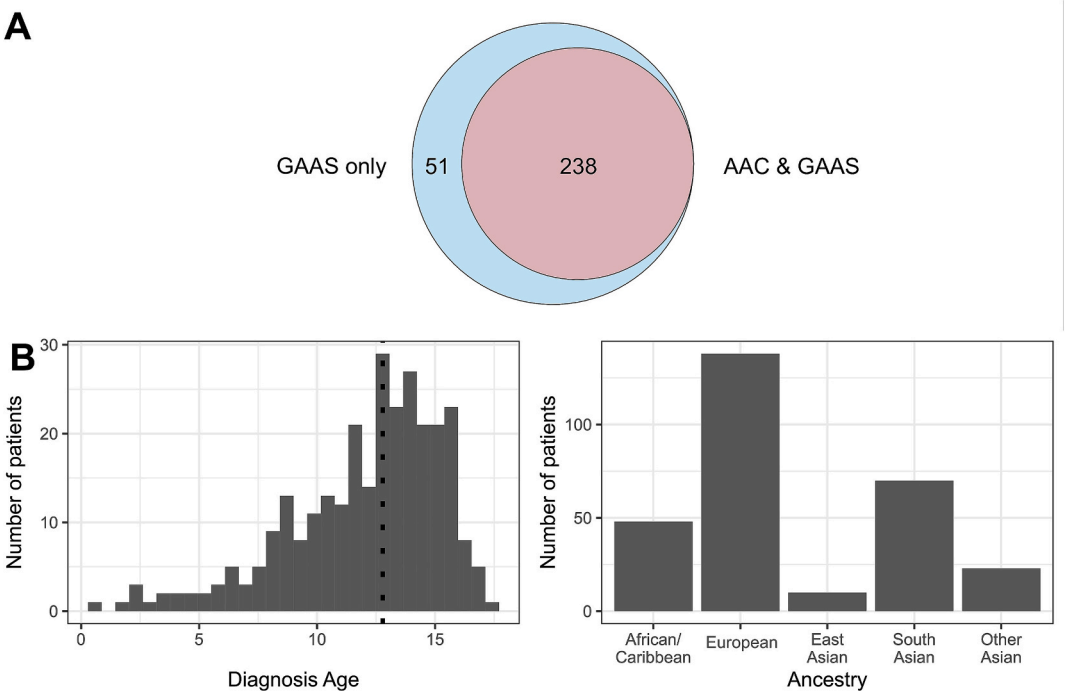

Fig. 1. Demographic characteristics of the jSLE patient sub-cohorts. A) Venn diagram displaying the overlap of study participants between the gene-level alternate allele scores (GAAS) (blue circle) and alternate allele counts (AAC) (pink circle) analysis groups. Out of the 289 patients included in the GAAS analysis, (238, 82.4 %) were included in both GAAS and AAC analyses, and no patients were included exclusively in the AAC analysis. B) Demographic characteristics of jSLE patients included in the GAAS. The left panel illustrates the age distribution in years at the time of diagnosis; the dotted line indicates the median age at diagnosis. The right panel displays the ancestral composition. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

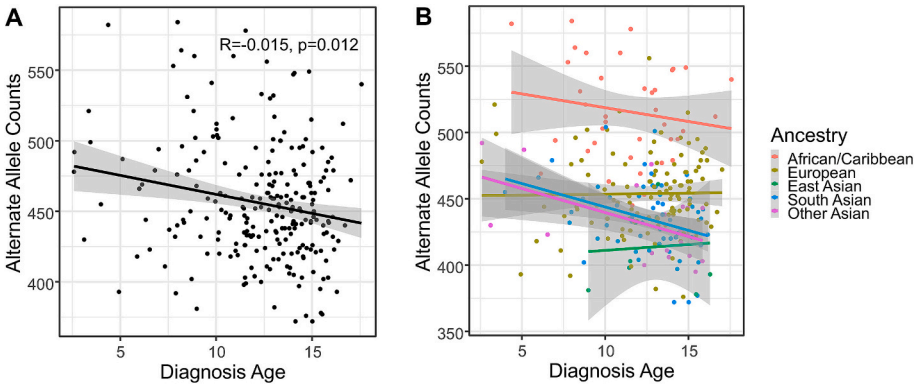

Fig. 2. Alternate allele counts (AAC) inversely correlate with age at disease onset. A) In a multi-ancestral jSLE cohort, age at diagnosis (in years) and AAC are inversely correlated (R = -0.15, $p = 0.012$). Data points represent the number of alternate alleles in individual jSLE patients; a trend line indicates the direction and strength of the relationship. The grey shaded area indicates the 95 % confidence interval. B) Analysis of the relationship between AAC and age at disease onset across the five different ancestral groups, reveals a moderate inverse correlation between age at onset and AAC among South Asian jSLE patients (R = -0.34, $p = 0.047$). Each data point represents an individual sample, colour-coded by ancestry: African/Caribbean (red), European (gold), East Asian (green), South Asian (blue), and "Other" Asian (purple). Trend lines indicate the direction and strength of the relationships. The grey shaded areas represent the 95 % confidence intervals. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

group revealed a moderate inverse correlation between age at diagnosis and AAC among South Asian participants (R = -0.34, $p = 0.047$) but no significant correlation was found within the remaining ancestral groups (Fig. 2B).

3.3. AAC are higher in patients of African/Caribbean ancestry

When comparing AAC across ancestries, higher scores were recorded in patients of African/Caribbean ancestry compared to patients of other ancestral groups ($p < 0.001$) (Fig. 3A). No significant differences were detected between male and female participants across ancestral groups. PCA of GAAS across all 62 genes/loci included in this study showed that patients of African/Caribbean ancestry cluster separately from other ancestries (Fig. 3B).

3.4. AAC associate with organ involvement but not with disease activity or damage

GLM analyses (Supplementary Tables 5 and 6) identified relationships between AAC and severity of pBILAG-2004-defined constitutional ($p < 0.001$), renal ($p = 0.0016$), haematological (p = 0.0016) and neuropsychiatric ($p = 0.03$) involvement, and the presence of non-erosive arthritis ($p = 0.004$) and malar rash ($p = 0.03$). We did not observe correlations between AAC and highest SLEDAI scores ($R = 0.044$, $p = 0.5$; Supplementary Fig. 1A), highest pBILAG scores ($R = 0.041$, $p = 0.53$; Supplementary Fig. 1B) or SLICC-SDI at last visit ($R = 0.032$, $p = 0.62$; Supplementary Fig. 1C).

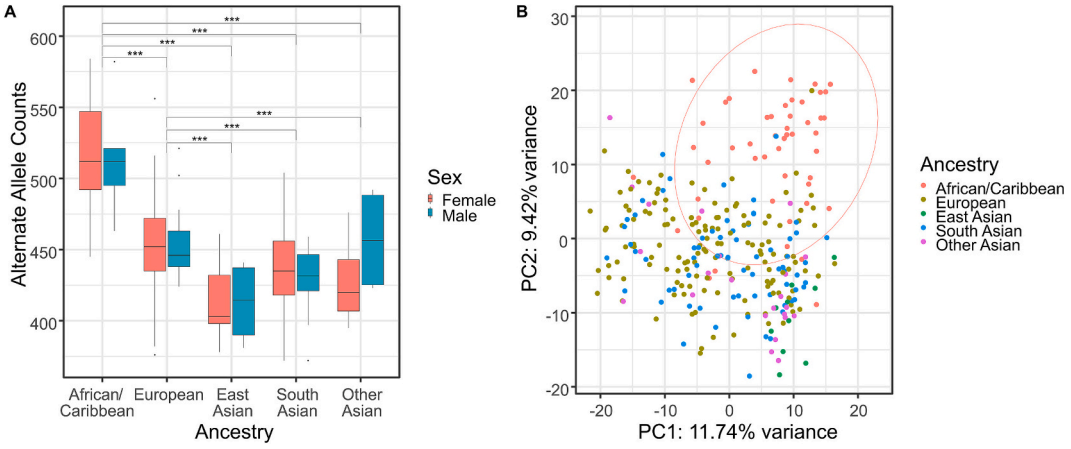

Fig. 3. Alternate allele counts (AAC) reveal differences among ancestries. A) Distribution of AAC stratified by ancestry and sex, showing higher AAC in patients of African/Caribbean ancestry, with no differences observed between male and female participants across ancestral groups. Sex distribution across ancestry groups: African/Caribbean females 33 (87 %), males 5 (13 %); European females 101 (89 %), males 13 (11 %); East Asian females 5 (56 %), males 4 (44 %); South Asian females 43 (75 %), males 14 (25 %); "other" Asian background females 16 (80 %), males 4 (20 %). Adjusted p-values are displayed (***$p \leq 0.001$). Box plots display interquartile ranges (IQRs) and median values. Whiskers extend to the minimum and maximum values within 1.5 times the IQR, with data points outside this range plotted individually as outliers. Ancestral groups are labelled along the x-axis: African/Caribbean, European, East Asian, South Asian and "Other" Asian. Sexes are separated by colour as indicated. B) Principal Component Analysis (PCA) of GAAS highlighting the distribution across different ancestral groups. Each dot signifies an individual data sample, with colour coding ancestry African/Caribbean, European, South Asian, and other Asian background. The first principal component (PC1) accounts for 11.74 % of the total variance, while the second principal component (PC2) captures 9.42 %. The points of African/Caribbean participants cluster largely separated from the other ancestries. The distinct clustering of African/Caribbean participants is encircled in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.5. GAAS associate with severity of organ involvement

We found associations between 18 clinical variables and 13 genes (Supplementary Tables 3, 7–9, Fig. 4). Neuropsychiatric severity associated with GAAS of the genes *Acid Phosphatase 5 (ACP5)* ($p < 0.001$), *Rat Sarcoma Guanyl Releasing Protein 3* (*RASGRP3)* ($p = 0.04$), *Ribonuclease H2 subunit A* (*RNASEH2 A)* ($p < 0.001$) and *Tyrosine Kinase 2* (*TYK2)* ($p < 0.001$) (Fig. 5A). Renal severity associated with GAAS of *ACP5* ($p < 0.001$), *Integrin Subunit Alpha M* (*ITGAM)* ($p < 0.001$), *Lck/Yes Novel Tyrosine Kinase (LYN)* ($p < 0.001$), and *Tumor Necrosis Factor Alpha-Induced Protein 3* (*TNFAIP3)* ($p = 0.007$) (Fig. 5B). Haematological involvement severity associated with GAAS of *ACP5* ($p = 0.003$), *Deoxyribonuclease 1* (*DNASE1)* ($p = 0.003$), *RNASEH2 A* ($p < 0.001$), and *Ubiquitin Conjugating Enzyme E2 L3* (*UBE2L3)* ($p = 0.03$) (Supplementary Fig. 2A). Cardiorespiratory severity associated with GAAS of *Interferon Regulatory Factor 8* (*IRF8)* ($p = 0.02$), *ITGAM* ($p < 0.001$) and *UBE2L3* ($p = 0.04$) (Supplementary Fig. 2B), while gastrointestinal severity associated with *ITGAM* ($p < 0.001$) and *PX Domain Containing Serine/Threonine Kinase Like* (*PXK)* ($p < 0.001$) (Supplementary Fig. 2C).

3.6. AAC and GAAS associate with treatment intensity

Overall, patients in the "intensive" treatment group had significantly higher AAC compared to participants receiving "non-intensive" treatment ($p < 0.001$, Fig. 6A). Participants from all ancestral groups, except European patients, receiving "intensive" treatment exhibited higher AAC compared to those receiving "non-intensive" treatment, although differences were only significant for African/Caribbean participants ($p = 0.001$) (Fig. 6B), consistent with their overall higher disease burden. Additionally, GAAS of *ITGAM* ($p < 0.001$), *LYN* ($p < 0.001$), *PXK* ($p < 0.001$) and *RNASEH2A* ($p < 0.001$) associated with increased treatment intensity (Fig. 6C).

3.7. AAC correlate with the subset AAC generated from previously published SNPs

A strong correlation was observed between the originally calculated AACs and the subset AACs, which were calculated using the 15 previously reported SLE-associated SNPs that were also captured by our sequencing panel ($R = 0.47$, $p < 0.001$; Supplementary Fig. 3A, Supplementary Table 4). However, when including previously reported SNPs only, an inverse correlation between AAC and age at diagnosis was not observed ($R = -0.08$, $p = 0.964$; Supplementary Fig. 3B). When stratifying by ancestry, European and African/Caribbean ancestral groups followed similar trends when comparing the two AACs, though significance was lost ($R = 0.05$, $p = 1$ and $R = -0.08$, $p = 1$, respectively). In East Asians, a moderate but non-significant inverse correlation emerged ($R = -0.48$, $p = 0.965$), while no correlation was observed in South Asian patients ($R = -0.05$, $p = 1$) or those of "other" Asian ancestry ($R = -0.07$, $p = 1$) (Supplementary Fig. 3C).

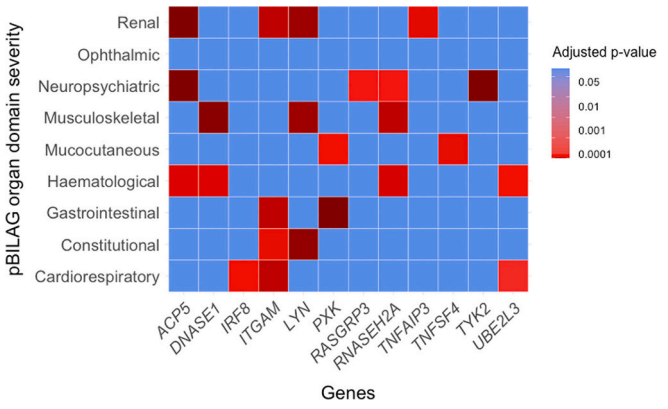

Fig. 4. Associations between pBILAG organ domain severity and gene-level alternate allele scores (GAAS), tested using generalized linear models. P-values were adjusted for multiple testing using the Benjamini–Hochberg method. Heatmap showing adjusted p-values of associations between clinical severity and GAAS, including the 12 genes displaying correlations between GAAS and pBILAG domain severity. Each cell represents a specific gene (x-axis) and pBILAG organ domain severity (y-axis), with colour intensity indicating the strength of the association. Darker red shades denote lower p-values, while blue shades indicate non-significant associations. The adjusted p-value thresholds are categorised as follows: 0.0001 (darkest red), 0.001, 0.01, 0.05 (lightest red) and > 0.05 (blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4. Discussion

In adult-onset SLE, associations between genetic variability, age at onset, ancestry, and disease severity have been established [6,26,27]. Preliminary reports suggest that, also in jSLE, increased genetic impact associates with early disease onset and high disease activity, especially in patients of African/Caribbean ancestry [6,25]. This study represents the most comprehensive analysis currently available in the age group, linking genetic variability across 62 SLE-associated genes/genomic regions with demographic and clinical variables, as well as treatment. Although all patients were recruited from study centers within a single nation, the cohort was multi-ancestral, with an expected "SLE-typical" over-representation of "non-European" ancestry, when compared to national UK census data [7,14] which allowed investigation of genetic disparities across ancestral groups.

The identified inverse correlation between AAC and age at diagnosis is in agreement with previous smaller studies, adding weight to the hypothesis that genetic burden contributes to early disease onset in jSLE [6,25,26]. In this study, the inverse correlation between AAC and age at onset was mainly driven by participants of South Asian descent. South Asian jSLE patients represented 24.2 % of the study population, while other non-European ancestries were less represented. Therefore, the relatively small sample size among these ancestries likely limited the statistical power to detect similar correlations. For example, a trend was observed in the African/Caribbean group, though it did not reach significance.

While Webb et al. reported an inverse correlation between genetic impact, age at disease-onset and severity in jSLE patients of African descent in the United States of America [6], we failed to identify associations between AAC and age at onset in the UK's African/Caribbean sample population. Differences may be explained by the more extensive sequencing panel approach chosen here (4100 vs. 19 SNPs) and the relatively small sample sizes across both cohorts (238 vs. 111 jSLE patients). Nevertheless, this study confirmed increased overall AAC in jSLE patients of African/Caribbean ancestry (compared to other ancestral groups) which may contribute to a lower threshold for SLE development, an overall earlier age of onset, more severe disease and less favourable outcomes [28]. An association between ancestry, disease severity and the need for intensive treatment, particularly among African/Caribbean patients, has been previously established in cohorts consisting of mostly adult-onset SLE patients [28,29]. However, findings from this study not only confirm this pattern in a paediatric multi-ancestral cohort, but they also link ancestry and phenotype-related differences with increased overall genetic variability around SLE-related genes/loci. Although expected to some extent based on previous studies reporting SNP-associations [30,31], here reported associations between overall genetic variability, organ involvement (especially renal involvement and *ACP5*, *ITGAM*, *LYN* and *TNFAIP3*) and treatment intensity underscore the importance of understanding genetic factors and their role in shaping organ-specific disease features for the development of individualised therapeutic approaches. Lastly, although SLE patients of Asian ancestry have previously been suggested to experience an increased genetic risk in adult-onset SLE cohorts [32], this study, when compared

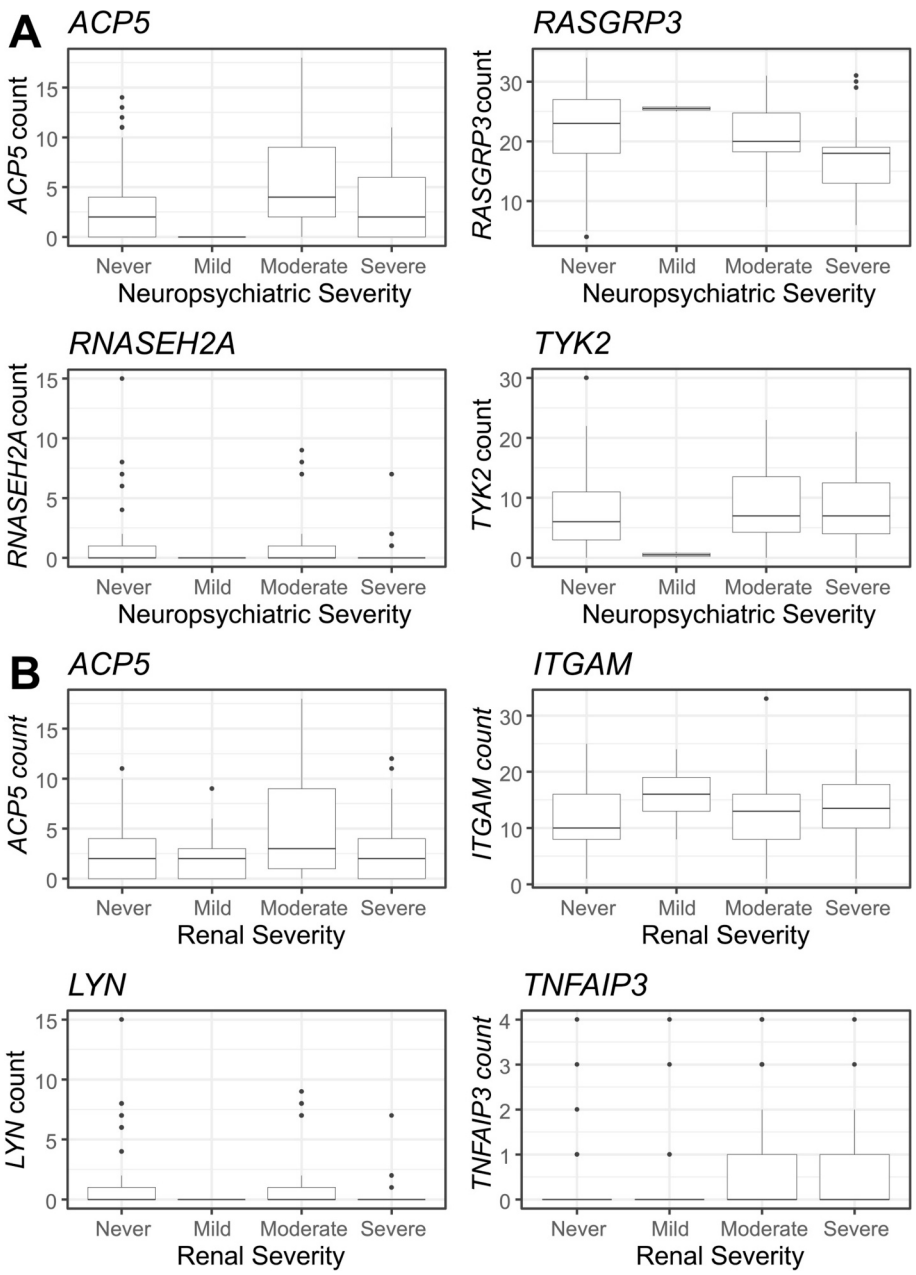

Fig. 5. Associations between severity of neuropsychiatric and renal involvement, and gene-level alternate allele scores (GAAS). A) Neuropsychiatric severity significantly associates to *ACP5* ($p < 0.001$), *RASGRP3* ($p = 0.04$), *RNASEH2A* ($p < 0.001$) and *TYK2* ($p < 0.001$) GAAS. B) Renal severity significantly associates to *ACP5* ($p < 0.001$), *ITGAM* ($p < 0.001$), *LYN* ($p < 0.001$), *TNFAIP3* ($p = 0.007$) GAAS. Box plots display interquartile ranges (IQR) and median values. Whiskers extend to the minimum and maximum values within 1.5 times the IQR, with data points outside this range plotted individually as outliers.

to other ancestries, did not detect higher AAC in children of Asian descent. This may be due to the relatively small sample size, differences in the sequencing panels used across studies, and/or possible differences between juvenile- and adult-onset SLE [33].

In addition to age at onset, AAC associated with the presence of certain clinical features. High AAC associated with severity of renal involvement. This is of particular interest, because lupus nephritis (LN) centrally impacts on overall disease severity, associates with the use of "intensive" treatments, and contributes to SLE-associated mortality [26,34]. Notably, a recent GWAS also reported a correlation between the development of LN and the number of SLE risk variants [26]. Moreover, a large Canadian study including both juvenile- and adult-onset SLE patients (1237 participants/572 jSLE) found associations between SLE susceptibility loci and the risk of developing LN [35].

Although previous studies in predominantly adult-onset SLE cohorts suggested increased damage accrual in patients with high genetic risk scores [27], this study did not detect correlations between AAC and SLICC-SDI scores. This may be the result of overall low SLICC-SDI scores across the study cohort (median: 0, IQR: 0–1), which could be attributed to SLICC-SDI having been developed and validated for adult SLE patients, emphasizing aspects of potentially lower relevance to children (including malignancy and diabetes mellitus) [36].

Key goals of genetic profiling across autoimmune/inflammatory diseases are the prediction of organ involvement, disease outcomes and/or treatment responses [33]. Among various associations identified, we found organ domains uniquely associated to GAAS of specific genes, such as *TNFAIP3* with renal severity and *TYK2* with neuropsychiatric severity. Additionally, in some cases, shared associations were observed; for example, GAAS of *ACP5* were associated with neuropsychiatric, renal and haematological severity.

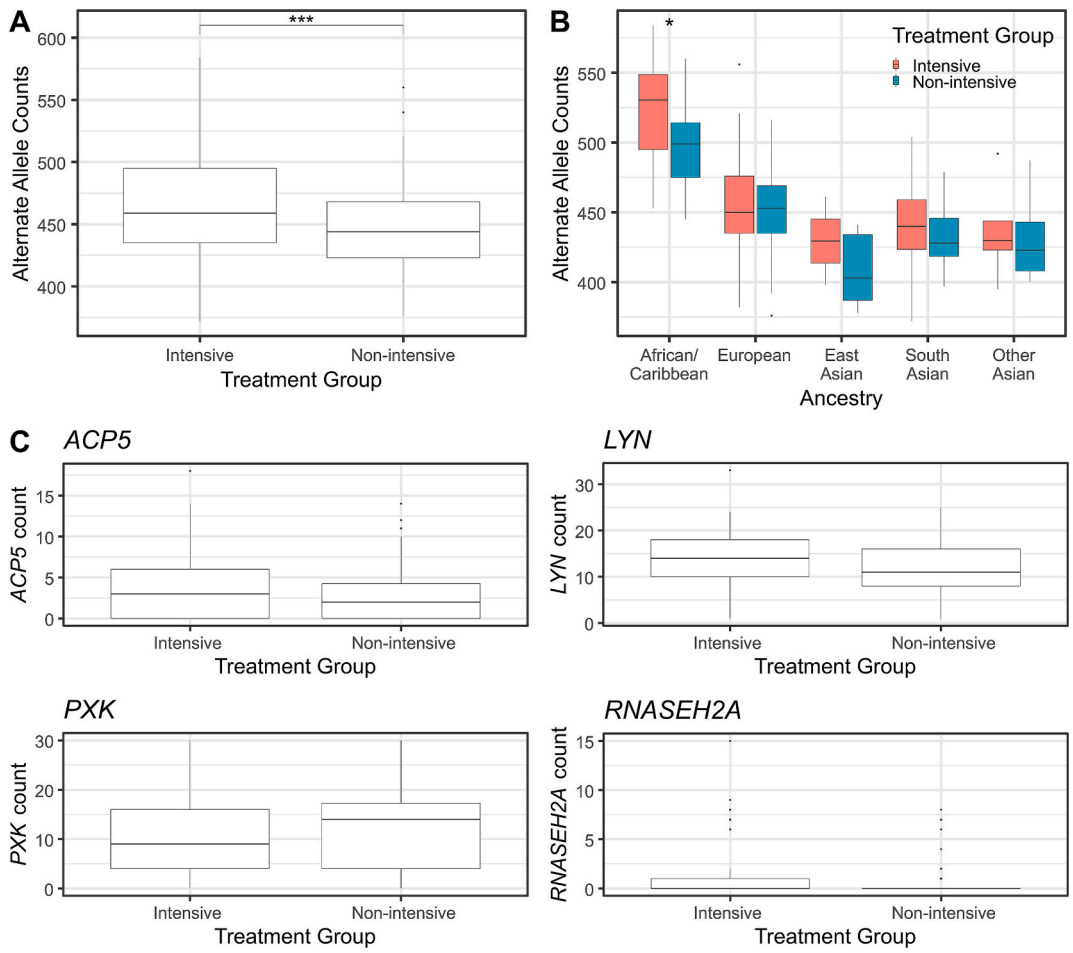

Fig. 6. Alternate allele counts (AAC) and gene-level alternate allele scores (GAAS) associate and with treatment intensity. A) Relationship between AAC and treatment intensity across ancestries, showing significantly higher AAC in the "intensive treatment" group compared to the "non-intensive treatment" group (***p ≤ 0.001). B) Box plots illustrating the relationship between alternate allele count (AAC) and two different treatment groups («intensive» and «non-intensive») across the five ancestral categories, showing higher AAC in patients requiring intensive treatment (in red) compared to those receiving non-intensive treatment regimens (in blue) among African/Caribbean jSLE patients (*$p \leq 0.05$). C) Box plots showing the distribution of significant associations between GAAS and treatment choices. Treatment intensity was significantly associated with *ITGAM* ($p < 0.001$), *LYN* ($p < 0.001$), *PXK* ($p < 0.001$), and *RNASEH2A* ($p < 0.001$) GAAS. Box plots display interquartile ranges (IQRs) and median values. Whiskers extend to the minimum and maximum values within 1.5 times the IQR, with data points outside this range plotted individually as outliers. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Several genes associated with specific organ involvement patterns are of particular interest as they have previously been linked with genetically determined SLE-like diseases and/or are targets of already available molecular interventions. For instance, recessive mutations in *ACP5*, encoding for the tartrate-resistant acid phosphatase (TRAP), associate with spondyloenchondrodysplasia (SPENCD). In line with organ domains associated with GAAS in this study, SPENCD is characterised by increased IFN expression, central nervous system involvement, haematological manifestations, vasculitis and renal involvement [37]. Furthermore, genetic variability affecting *RNASEH2A* associated with neurological and haematological involvement and severity. This gene has previously been linked with the neurodegenerative disease Aicardi-Goutières syndrome (AGS), another interferonopathy characterised by early-onset encephalopathy and lupus-like features [38]. Notably, both SPENCD and AGS patients show (at least limited) improvement in response to JAK/STAT inhibitors that are used to control IFN expression [39], suggesting that genetic profiling may not only predict organ involvement but also inform future treatment choices. Additionally, the tyrosine kinase encoding *TYK2* gene is another IFN signalling related gene associated with neuropsychiatric severity and the occurrence of discoid lupus. The pivotal role of TYK2 in the JAK/STAT pathway and IFN expression prompted the development of TYK2 modulators (deucravacitinib, brepocitinib). While approval for the treatment of SLE is pending, deucravacitinib was effective and safe in adult SLE patients [40]. The tyrosine kinase encoding *LYN* gene was associated with disease severity affecting the renal pBILAG organ domain. The Lyn kinase plays a key role in regulating B lymphocyte signalling; mice carrying gain-of-function variants in *Lyn* exhibit circulating autoantibodies and severe autoimmune glomerulonephritis [41]. A genome-wide association study in IgA nephropathy identified *LYN* as a risk locus [42], and a recent study in severe COVID-19 found *LYN* to be associated with progression of kidney damage [43]. Notably, the Lyn/Bcr-Abl tyrosine kinase inhibitor bafetinib was efficacious and safe in acute lymphocytic leukaemia, refractory/relapsing B cell lymphocytic leukaemia and prostate cancer [44]. Finally, the observed association between *TNFAIP3* GAAS and renal severity is particularly interesting because heterozygous loss-of-function mutations in *TNFAIP3* have recently been linked with haploinsufficiency A20 [45]. Patients affected by this autoinflammatory condition can exhibit SLE-like phenotypes [14,46], and some can develop LN in the presence of autoantibodies and a pronounced type I IFN signature [46]. Findings from this and another recent study [47] therefore suggest that variants in the *TNFAIP3* may affect renal involvement and outcomes in SLE. Taken together, associations between GAAS and organ involvement/severity may inform

future personalised jSLE management, including genetic risk assessment and patient stratification towards individualised treatment and care. Such an approach could, for example, entail increased monitoring frequency for renal function, proteinuria and blood pressure elevation or the choice of more aggressive immunosuppressive treatment in jSLE patients with elevated GAAS in *ACP5*, *ITGAM*, *LYN*, or *TNFAIP3*. Indeed, similar approaches have already been introduced in cancer medicine [48].

This is the first study to show that high genetic risk scores (AAC) in jSLE patients associate with the need for "intensive" treatment. While only reaching statistical significance among African/Caribbean jSLE patients, this association remained across all ancestral groups, except European participants. While the reason for this remains unclear and requires further investigation, genetic risk has been reported "lower" among European patients as compared to other ancestries [5,32], which also associates with "milder" disease courses and better outcomes [49].

Differences between allele counts and age at disease onset observed between AAC, measuring overall genetic variation, and the subset AAC generated from previously reported SLE-associated SNPs, highlight the impact of (ancestry- and potentially age-related biased) SNP selection in genetic studies. A large proportion of previously published SLE-associated SNPs, more precisely 148/330 (44.8 %), were exclusively identified in East Asian populations [24]. Previous studies overall had limited representation of, e.g., South Asian or African/Caribbean jSLE patients, notably, two of the three most represented ancestries in the dataset presented here (respectively 23.9/24.2 % and 16/16.6 % of the cohort). Analyses restricted to a subset of previously published SLE-associated SNPs may therefore have resulted in a loss of signal, particularly in underrepresented populations (such as South Asian patients). This underscores the importance of inclusive and appropriately matched genomic datasets, as limiting analyses to previously reported variants may skew associations, particularly in multi-ancestral cohorts.

A key strength of this study is its focus on genetic variability across SLE-associated genes and regions rather than limiting to specific previously identified risk alleles (usually from GWAS). While GWAS have been instrumental in identifying individual SNPs associated with SLE risk, or protection from disease or complications, the approach taken in this study provides a broader and, potentially, less biased assessment of genetic burden by capturing the combined effect of multiple alternate alleles. This perspective may offer insights into disease heterogeneity and severity, particularly in multi-ancestral populations, and could complement existing analytic approaches by providing an alternative or additional measure of variability and its impact on disease expression and/or phenotype. This is reflected in the design of the sequencing panel, which was focused on exons, exon–intron junctions, and regions around several previously reported risk alleles [14]. Consequently, a large proportion of sequences analysed here are located in coding or nearby regulatory regions, rather than in intergenic areas where many GWAS hits are found. This choice aligned with our aim to assess gene-level variation rather than replicate known SNP-level associations.

While this study highlights the potential of genetic assessment in guiding and personalising clinical management for jSLE patients, several limitations require to be acknowledged. Despite representing one of the largest jSLE cohorts available, the overall sample size remains relatively small, especially when compared to studies in adult-onset SLE [25,27]. This limitation is particularly evident in subgroup analyses comparing sex across ancestries, where the number of male patients, especially those of African or East Asian descent, was low. Additionally, although the multi-ancestral composition of the cohort supports broader relevance, variability in genetic backgrounds, healthcare systems, drug availability and socio-economic factors may influence the generalizability of findings to other populations. Furthermore, the selection of genes/loci was based on a literature review conducted in 2018 [14]. Consequently, SLE-associated genes identified more recently, such as *SAT1, P2RY8* and *DOCK11* [50–52], were not included. Moreover, the *APOL1* gene, which has been associated with severe renal involvement and damage [53–55], as well as atherosclerosis [56] in adult SLE patients (particularly those of African American descent), was not included in this analysis. Another potential limitation may be related to the use of self-reported ancestry data. While this approach may introduce some degree of misclassification, it remains widely accepted when comprehensive genomic data are not available, as it has been shown to have good concordance with genetically inferred ancestry [57]. The panel used in this study included only 62 exonic and exon–intron regions, with a strong focus on immune-related genes associated with SLE. These regions are typically under strong functional constraint and, as such, are unlikely to contain ancestry-informative markers (AIMs), which are more often located in neutral, non-functional regions [58]. Furthermore, the exon panel in question comprises genes associated with SLE, many of which are involved in immune function and have been shaped by population-specific selective pressures. For example, *IRF5* polymorphisms are associated with IFN pathway activation and show varying allele frequencies across populations, likely due to historical pathogen-driven selection [59]. Similarly, *TYK2* variants have been identified as protective against SLE in European populations but are rare or absent in indigenous populations, suggesting that these alleles were introduced through European admixture and may have been subject to negative selection in populations with high exposure to infectious diseases [60]. This indicates that the distribution of these alleles may be influenced more by selection pressures than by neutral drift. Using SNPs chosen for disease association—especially in immune-related genes—introduces the risk of confounding disease susceptibility with ancestry. Effective ancestry inference requires either genome-wide data or panels of AIMs selected specifically for their ability to capture population divergence through neutral evolution (for example through microhaplotypes, clusters of closely linked SNPs [61], which were not the focus of this study). For these reasons, ancestry analysis was not pursued with this exon-limited, disease-biased gene set. Lastly, studies linking individual or the combination of variants with gene function are necessary, for example, to inform future treatment choices.

In conclusion, this study underscores the important role of genetic variability in jSLE, demonstrating its key contribution to early disease expression - particularly in jSLE patients of South Asian ancestry - as well as organ involvement, disease severity and the need for more intensive treatment. By capturing a broader spectrum of genetic variability rather than focusing on previously reported SLE-associated risk alleles, the approach taken in this study may provide new insights into how cumulative genetic impact can influence disease heterogeneity. Prospective studies in larger, unrelated multi-ancestral cohorts are required to validate these findings and further assess the need for more aggressive treatments in patients exhibiting high GAAS in genes associated with specific manifestations.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.clim.2025.110540.

Funding sources

The UK JSLE Cohort Study and Repository and the sequencing project are supported by LUPUS UK (to CMH and MWB), the UK's Experimental Arthritis Treatment Centre for Children (supported by Versus Arthritis, Alder Hey Children's NHS Foundation Trust, the Alder Hey Charity and the University of Liverpool) (to CMH and MWB), the Alder Hey NIHR Clinical Research Facility (CRF), and the NIHR Great Ormond Street Biomedical Research Centre (BRC) (to CMH). The bioinformatics and statistical analyses were supported by the Computational Biology Facility through Liverpool Shared Research Facilities, University of Liverpool.

CRediT authorship contribution statement

Valentina Natoli: Conceptualization, Data curation, Investigation, Project administration, Writing – original draft. **Amandine Charras:**

Conceptualization, Data curation, Investigation, Project administration, Supervision, Writing – original draft, Writing – review & editing. **Megan S.R. Hasoon:** Formal analysis, Methodology, Project administration, Software, Visualization, Writing – original draft, Writing – review & editing. **Andrea L. Jorgensen:** Writing – review & editing. **Eve M.D. Smith:** Resources, Writing – review & editing. **Eva Caamaño Gutiérrez:** Conceptualization, Formal analysis, Investigation, Methodology, Resources, Software, Supervision, Validation, Visualization, Writing – review & editing. **Michael W. Beresford:** Conceptualization, Project administration, Resources, Writing – review & editing. **Christian M. Hedrich:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Writing – review & editing.

Declaration of competing interest

CMH received unrestricted grant support from Novartis (to study effector T cells in psoriasis) and Merck (to study spatial transcriptional profiles in SLE). The authors declare that the here presented research was conducted in the absence of any commercial or financial relationships that could be construed as a potential competing interest.

Acknowledgements

The authors wish to acknowledge all patients and their parents who consented to participate in the UK JSLE Cohort Study, medical staff caring for lupus patients and all members of the teams within each of the paediatric centres who take part to the UK JSLE Study Group.

Data availability

Because not all results from this study have been published yet, raw sequencing data and processed dcc files will be made available upon reasonable request. Code to reproduce the analysis and figures is available at https://github.com/CBFLivUni/jSLE_paper.

References

[1] A. Charras, Systemic lupus erythematosus in children and young people, Curr. Rheumatol. Rep. 15 (2021).
[2] T. Alexander, C.M. Hedrich, Systemic lupus erythematosus – are children miniature adults? Clin. Immunol. 234 (2022 Jan) 108907.
[3] N. Ambrose, T.A. Morgan, J. Galloway, Y. Ionnoau, M.W. Beresford, D.A. Isenberg, Differences in disease phenotype and severity in SLE across age groups, Lupus 25 (14) (2016 Dec) 1542–1550.
[4] C.M. Hedrich, E.M.D. Smith, M.W. Beresford, Juvenile-onset systemic lupus erythematosus (jSLE) – pathophysiological concepts and treatment options, Best Pract. Res. Clin. Rheumatol. 31 (4) (2017 Aug) 488–504.
[5] S. Guga, Y. Wang, D.C. Graham, T.J. Vyse, A review of genetic risk in systemic lupus erythematosus, Expert Rev. Clin. Immunol. 19 (10) (2023) 1247–1258.
[6] R. Webb, J.A. Kelly, E.C. Somers, T. Hughes, K.M. Kaufman, E. Sanchez, et al., Early disease onset is predicted by a higher genetic risk for lupus and is associated with a more severe phenotype in lupus patients, Ann. Rheum. Dis. 70 (1) (2011) 151–156.
[7] J.S. Massias, E.M.D. Smith, E. Al-Abadi, K. Armon, K. Bailey, C. Ciurtin, et al., Clinical and laboratory characteristics in juvenile-onset systemic lupus erythematosus across age groups, Lupus 29 (5) (2020 Apr) 474–481.
[8] UK JSLE Cohort Study. https://www.liverpool.ac.uk/translational-medicine/research/ukjsle/jsle/, 2025.
[9] J. Little, J.P.T. Higgins, J.P.A. Ioannidis, D. Moher, F. Gagnon, E. von Elm, et al., Strengthening the reporting of genetic association studies (STREGA)–an extension of the STROBE statement, Genet. Epidemiol. 33 (7) (2009 Nov) 581–598.
[10] E.A.R. Hartman, A. van Royen-Kerkhof, J.W.G. Jacobs, P.M.J. Welsing, R.D. E. Fritsch-Stork, Performance of the 2012 systemic lupus international collaborating clinics classification criteria versus the 1997 American College of Rheumatology classification criteria in adult and juvenile systemic lupus erythematosus. A systematic review and meta-analysis, Autoimmun. Rev. 17 (3) (2018 Mar) 316–322.
[11] D.A. Isenberg, A. Rahman, E. Allen, V. Farewell, M. Akil, I.N. Bruce, et al., BILAG 2004. Development and initial validation of an updated version of the British Isles Lupus Assessment Group's disease activity index for patients with systemic lupus erythematosus, Rheumatology (Oxford) 44 (7) (2005 Jul) 902–906.
[12] S.D. Marks, C. Pilkington, P. Woo, M.J. Dillon, The use of the British Isles Lupus Assessment Group (BILAG) index as a valid tool in assessing disease activity in childhood-onset systemic lupus erythematosus, Rheumatology (Oxford) 43 (9) (2004 Sep) 1186–1189.
[13] D. Gladman, E. Ginzler, C. Goldsmith, et al., The development and initial validation of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology damage index for systemic lupus erythematosus, Arthritis Rheum 39 (3) (1996) 363–369, https://doi.org/10.1002/art.1780390303.
[14] A. Charras, S. Haldenby, E.M.D. Smith, N. Egbivwie, L. Olohan, J.G. Kenny, et al., Panel Sequencing Links Rare, Likely Damaging Gene Variants with Distinct Clinical Phenotypes and Outcomes in Juvenile-Onset SLE. Rheumatology (Oxford) 62(SI2), 2023, pp. SI210–SI225.
[15] M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads, EMBnet J. 17 (1) (2011 May 2) 10–12.
[16] J.N. Fass, N.A. Joshi, J.N. Fass, Sickle: A Sliding-Window, Adaptive, Quality-Based Trimming Tool for FastQ Files (Version 1.33) [Software], Available at, https://github.com/najoshi/sickle, 2011.
[17] Picard Toolkit, Broad Institute, GitHub Repository. https://broadinstitute.github.io/picard/, 2019.
[18] H. Li, R. Durbin, Fast and accurate short read alignment with burrows-wheeler transform, Bioinformatics 25 (14) (2009 Jul 15) 1754–1760.
[19] G.A. Van der Auwera, M.O. Carneiro, C. Hartl, R. Poplin, G. Del Angel, A. Levy-Moonshine, et al., From FastQ data to high confidence variant calls: the genome analysis toolkit best practices pipeline, Curr. Protoc. Bioinformatics 43 (2013), 11.10.1-11.10.33.
[20] P. Cingolani, A. Platts, L.L. Wang, M. Coon, T. Nguyen, L. Wang, et al., A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3, Fly (Austin) 6 (2) (2012 Jun) 80–92.
[21] S.T. Sherry, M.H. Ward, M. Kholodov, J. Baker, L. Phan, E.M. Smigielski, et al., dbSNP: the NCBI database of genetic variation, Nucleic Acids Res. 29 (1) (2001 Jan 1) 308–311.
[22] V. Obenchain, M. Lawrence, V. Carey, S. Gogarten, P. Shannon, M. Morgan, VariantAnnotation: a Bioconductor package for exploration and annotation of genetic variants, Bioinformatics 30 (14) (2014 Jul 15) 2076–2078.
[23] B.J. Knaus, N.J. Grünwald, VcfR: A Package to Manipulate and Visualize VCF Format Data in R [Internet], bioRxiv, 2016, p. 041277 [cited 2023 Oct 20]. Available from: https://doi.org/10.1101/041277v1.
[24] V. Laurynenka, J.B. Harley, The 330 risk loci known for systemic lupus erythematosus (SLE): a review, Front. Lupus [Internet] 2 (2024 May 24), https://doi.org/10.3389/flupu.2024.1398035/full. Available from:. [cited 2025 Mar 7].
[25] D. Dominguez, S. Kamphuis, J. Beyene, J. Wither, J.B. Harley, I. Blanco, et al., Relationship between genetic risk and age of diagnosis in systemic lupus erythematosus, J. Rheumatol. 48 (6) (2021 Jun 1) 852–858.
[26] L. Chen, Y.F. Wang, L. Liu, A. Bielowka, R. Ahmed, H. Zhang, et al., Genome-wide assessment of genetic risk for systemic lupus erythematosus and disease severity, Hum. Mol. Genet. 29 (10) (2020 Jun 27) 1745–1756.
[27] S. Reid, A. Alexsson, M. Frodlund, D. Morris, J.K. Sandling, K. Bolin, et al., High Genetic Risk Score is Associated with Early Disease Onset, Damage Accrual and Decreased Survival in Systemic Lupus Erythematosus. Systemic Lupus Erythematosus, Ann Rheum Dis 79 (3) (2020) 363–369.
[28] M.J. Lewis, A.S. Jawad, The effect of ethnicity and genetic ancestry on the epidemiology, clinical features and outcome of systemic lupus erythematosus, Rheumatology (Oxford) 56 (suppl_1) (2017 Apr 1) i67–i77.
[29] S. Slight-Webb, K. Thomas, M. Smith, C.A. Wagner, S. Macwana, A. Bylinska, et al., Ancestry-based differences in the immune phenotype are associated with lupus activity, JCI Insight 8 (16) (2023 Aug 22) e169584.
[30] K. Song, L. Liu, X. Zhang, X. Chen, An update on genetic susceptibility in lupus nephritis, Clin. Immunol. 210 (2020 Jan 1) 108272.
[31] K.A. Owen, K.A. Bell, A. Price, P. Bachali, H. Ainsworth, M.C. Marion, et al., Molecular pathways identified from single nucleotide polymorphisms demonstrate mechanistic differences in systemic lupus erythematosus patients of Asian and European ancestry, Sci. Rep. 13 (1) (2023 Apr 1) 5339.
[32] G.N. Goulielmos, M.I. Zervou, V.M. Vazgiourakis, Y. Ghodke-Puranik, A. Garyfallos, T.B. Niewold, The genetics and molecular pathogenesis of systemic lupus erythematosus (SLE) in populations of different ancestry, Gene 668 (2018 Aug 20) 59–72.
[33] S. Eyre, G. Orozco, J. Worthington, The genetics revolution in rheumatology: large scale genomic arrays and genetic mapping, Nat. Rev. Rheumatol. 13 (7) (2017 Jul) 421–432.
[34] G.A. Bello, M.A. Brown, J.A. Kelly, A. Thanou, J.A. James, C.G. Montgomery, Development and validation of a simple lupus severity index using ACR criteria for classification of SLE, Lupus Sci. Med. 3 (1) (2016 Mar 10) e000136.
[35] D. Webber, J. Cao, D. Dominguez, D.D. Gladman, D.M. Levy, L. Ng, et al., Association of systemic lupus erythematosus (SLE) genetic susceptibility loci with lupus nephritis in childhood-onset and adult-onset SLE, Rheumatology 59 (1) (2020 Jan 1) 90–98.
[36] D. Gladman, E. Ginzler, C. Goldsmith, P. Fortin, M. Liang, M. Urowitz, et al., Systemic lupus international collaborative clinics: development of a damage index in systemic lupus erythematosus, J. Rheumatol. 19 (11) (1992 Nov) 1820–1821.
[37] T.A. Briggs, G.I. Rice, N. Adib, L. Ades, S. Barete, K. Baskar, et al., Spondyloenchondrodysplasia due to mutations in ACP5: a comprehensive survey, J. Clin. Immunol. 36 (3) (2016 Apr) 220–234.
[38] G. Ramantani, J. Kohlhase, C. Hertzberg, A.M. Innes, K. Engel, S. Hunger, et al., Expanding the phenotypic spectrum of lupus erythematosus in Aicardi-Goutières syndrome, Arthritis Rheum. 62 (5) (2010 May) 1469–1477.

[39] W. Li, W. Wang, W. Wang, L. Zhong, L. Gou, C. Wang, et al., Janus kinase inhibitors in the treatment of type I interferonopathies: a case series from a single center in China, Front. Immunol. 13 (2022 Mar 28) 825367.
[40] E. Morand, M. Pike, J.T. Merrill, R. van Vollenhoven, V.P. Werth, C. Hobar, et al., Deucravacitinib, a tyrosine kinase 2 inhibitor, in systemic lupus erythematosus: a phase II, randomized, double-blind, placebo-controlled trial, Arthritis Rheumatol. 75 (2) (2023 Feb) 242–252.
[41] M.L. Hibbs, K.W. Harder, J. Armes, N. Kountouri, C. Quilici, F. Casagranda, et al., Sustained activation of Lyn tyrosine kinase in vivo leads to autoimmunity, J. Exp. Med. 196 (12) (2002 Dec 16) 1593–1604.
[42] K. Kiryluk, E. Sanchez-Rodriguez, X.J. Zhou, F. Zanoni, L. Liu, N. Mladkova, et al., Genome-wide association analyses define pathogenic signaling pathways and prioritize drug targets for IgA nephropathy, Nat. Genet. 55 (7) (2023 Jul) 1091–1105.
[43] Z. Chen, C. Chen, F. Chen, R. Lan, G. Lin, Y. Xu, Bioinformatics analysis of potential pathogenesis and risk genes of immunoinflammation-promoted renal injury in severe COVID-19, Front. Immunol. 13 (2022) 950076.
[44] F.P.S. Santos, H. Kantarjian, J. Cortes, A. Quintas-Cardama, Bafetinib, a dual Bcr-Abl/Lyn tyrosine kinase inhibitor for the potential treatment of leukemia, Curr. Opin. Investig. Drugs 11 (12) (2010 Dec) 1450–1465.
[45] F.A. Aeschlimann, E.D. Batu, S.W. Canna, E. Go, A. Gül, P. Hoffmann, et al., A20 haploinsufficiency (HA20): clinical phenotypes and disease course of patients with a newly recognised NF-kB-mediated autoinflammatory disease, Ann. Rheum. Dis. 77 (5) (2018 May) 728–735.
[46] C. Zhang, X. Han, L. Sun, S. Yang, J. Peng, Y. Chen, et al., Novel loss-of-function mutations in TNFAIP3 gene in patients with lupus nephritis, Clin. Kidney J. 15 (11) (2022 Nov) 2027–2038.
[47] J.S. Bates, C.J. Lessard, J.M. Leon, T. Nguyen, L.J. Battiest, J. Rodgers, et al., Meta-analysis and imputation identifies a 109 kb risk haplotype spanning TNFAIP3 associated with lupus nephritis and hematologic manifestations, Genes Immun. 10 (5) (2009 Jul) 470–477.
[48] C. Huntley, B. Torr, A. Sud, C.F. Rowlands, R. Way, K. Snape, et al., Utility of polygenic risk scores in UK cancer screening: a modelling analysis, Lancet Oncol. 24 (6) (2023 Jun) 658–668.
[49] E. Krishnan, H.B. Hubert, Ethnicity and mortality from systemic lupus erythematosus in the US, Ann. Rheum. Dis. 65 (11) (2006 Nov) 1500–1505.
[50] L. Xu, J. Zhao, Q. Sun, X. Xu, L. Wang, T. Liu, et al., Loss-of-function variants in SAT1 cause X-linked childhood-onset systemic lupus erythematosus, Ann. Rheum. Dis. 81 (12) (2022 Dec) 1712–1721.
[51] Y. He, A.E. Gallman, C. Xie, Q. Shen, J. Ma, F.D. Wolfreys, et al., P2RY8 variants in lupus patients uncover a role for the receptor in immunological tolerance, J. Exp. Med. 219 (1) (2022 Jan 3) e20211004.
[52] C. Boussard, L. Delage, T. Gajardo, A. Kauskot, M. Batignes, N. Goudin, et al., DOCK11 deficiency in patients with X-linked actinopathy and autoimmunity, Blood 141 (22) (2023 Jun 1) 2713–2726.
[53] C.P. Larsen, M.L. Beggs, M. Saeed, P.D. Walker, Apolipoprotein L1 risk variants associate with systemic lupus erythematosus-associated collapsing glomerulopathy, J. Am. Soc. Nephrol. 24 (5) (2013 Apr) 722–725.
[54] B.I. Freedman, C.D. Langefeld, K.K. Andringa, J.A. Croker, A.H. Williams, N. E. Garner, et al., End-stage renal disease in African Americans with lupus nephritis is associated with APOL1, Arthritis Rheumatol. 66 (2) (2014 Feb) 390–396.
[55] T. Kofman, C. Narjoz, Q. Raimbourg, M.A. Loriot, A. Karras, M. Roland, et al., Collapsing glomerulopathy associated lupus in a black female with homozygous APOL1 mutation, Lupus 21 (13) (2012 Nov) 1459–1462.
[56] A. Blazer, B. Wang, D. Simpson, T. Kirchhoff, S. Heffron, R.M. Clancy, et al., Apolipoprotein L1 risk variants associate with prevalent atherosclerotic disease in African American systemic lupus erythematosus patients, PLoS One 12 (8) (2017) e0182483.
[57] Y. Banda, M.N. Kvale, T.J. Hoffmann, S.E. Hesselson, D. Ranatunga, H. Tang, et al., Characterizing race/ethnicity and genetic ancestry for 100,000 subjects in the genetic epidemiology research on adult health and aging (GERA) cohort, Genetics 200 (4) (2015 Aug) 1285–1295.
[58] N.A. Rosenberg, L.M. Li, R. Ward, J.K. Pritchard, Informativeness of genetic markers for inference of ancestry, Am. J. Hum. Genet. 73 (6) (2003 Dec) 1402–1422.
[59] Y. Nédélec, J. Sanz, G. Baharian, Z.A. Szpiech, A. Pacis, A. Dumaine, et al., Genetic ancestry and natural selection drive population differences in immune responses to pathogens, Cell 167 (3) (2016 Oct 20) 657–669.e21.
[60] C. Contreras-Cubas, H. García-Ortiz, R. Velázquez-Cruz, F. Barajas-Olmos, P. Baca, A. Martínez-Hernández, et al., Catalytically impaired TYK2 variants are protective against childhood- and adult-onset systemic lupus erythematosus in Mexicans, Sci. Rep. 9 (1) (2019 Aug 21) 12165.
[61] K.K. Kidd, W.C. Speed, A.J. Pakstis, M.R. Furtado, R. Fang, A. Madbouly, et al., Progress toward an efficient panel of SNPs for ancestry inference, Forensic Sci. Int. Genet. 10 (2014 May) 23–32.