Research article

Examination of eQTL Polymorphisms Associated with Increased Risk of Progressive Complicated Sarcoidosis in European and African Descent Subjects

Nancy G. Casanova[1], Sara M. Camp[2], Manuel L.Gonzalez-Garay[6], Ken Batai[3], Lori Garman[4], Courtney G. Montgomery[4], Nathan Ellis[5], Rick Kittles[6], Christian Bime[7], Amy P. Hsu[8], Steven Holland[8], Yves A.Lussier[9], Jason Karnes[10], Nadera Sweiss[11], Lisa A. Maier[12], Laura Koth[13], David R. Moller[14], Naftali Kaminski[14], Joe G.N. Garcia[*2]

1. Department of Molecular Medicine, Univeristy of Florida, Scripps, Jupiter FL, USA

2. Center for Inflammation Science and Systems Medicine, University of Florida, Wertheim Scripps Research Institute, Jupiter FL, USA

3. Cancer Prevention & Control, Roswell Park Comprehensive Cancer Center, Buffalo, NY, USA

4. Oklahoma Medical Research Foundation, Oklahoma City, OK, USA

5. University of Arizona Cancer Center, Tucson, AZ, USA

6. Division of Health Equities, Department of Population Sciences, City of Hope, Duarte, California, USA

7. Department of Medicine University of Arizona, Tucson, AZ, USA

8. National Institute of Allergy and Infectious Diseases. National Institutes of Health, USA

9. Department of Biomedical Informatics, University of Utah, Salt Lake City, UT, USA

10. Department of Pharmacology, University of Arizona, College of Pharmacy, Tucson, AZ, USA

11. Department of Medicine University of Illinois, Chicago, IL, USA

12. Department of Medicine National Jewish Health, University of Colorado, Denver, CO, USA

13. Department of Medicine University of California San Francisco, San Francisco , CA , US, USA

14. Department of Medicine Johns Hopkins University School of Medicine, Baltimore Maryland, USA

15. Department of Medicine Yale University School of Medicine, New Haven, CT, USA

*Corresponding author: Joe G. N. Garcia, MD, University of Florida Wertheim Scripps Research Institute, Jupiter, FL 33458, USA

Abstract

Background: A limited pool of SNPs are linked to the development and severity of sarcoidosis, a systemic granulomatous inflammatory disease. By integrating genome-wide association studies (GWAS) data and expression quantitative trait loci (eQTL) single nuclear polymorphisms (SNPs), we aimed to identify novel sarcoidosis SNPs potentially influencing the development of complicated sarcoidosis. Methods: A GWAS (Affymetrix 6.0) involving 209 African-American (AA) and 193 European-American (EA, 75 and 51 complicated cases respectively) and publicly-available GWAS controls (GAIN) was utilized. Annotation of multi-tissue eQTL SNPs present on the GWAS created a pool of ~46,000 eQTL SNPs examined for association with sarcoidosis risk and severity (Logistic Model, Plink). The most significant EA/AA eQTL SNPs were genotyped in a sarcoidosis validation cohort (n=1034) and cross-validated in two independent GWAS cohorts. Results: No single GWAS SNP achieved significance ($p < 1 \times 10^{-8}$), however, analysis of the eQTL/GWAS SNP pool yielded 621 eQTL SNPs ($p < 10^{-4}$) associated with 730 genes that highlighted innate immunity, MHC Class II, and allograft rejection pathways with multiple SNPs validated in an independent sarcoidosis cohort (105 SNPs analyzed) (NOTCH4, IL27RA, BTNL2, ANXA11, HLA-DRB1). These studies confirm significant association of eQTL/GWAS SNPs in EAs and AAs with sarcoidosis risk and severity (complicated sarcoidosis) involving HLA region and innate immunity. Conclusion: Despite the challenge of deciphering the genetic basis for sarcoidosis risk/severity, these results suggest that integrated eQTL/GWAS approaches may identify novel variants/genes and support the contribution of dysregulated innate immune responses to sarcoidosis severity.

Keywords: Sarcoidosis, GWAS, expression quantitative trait loci (eQTL), polymorphism, SNPs

Introduction

Sarcoidosis is a complex multisystemic disease of unknown etiology characterized by the formation of non-necrotizing granulomas in multiple organs but invariably involving lung tissues and thoracic lymph nodes [1]. Although impacting all races, ages, and genders, when compared with non-Hispanic Whites (NHWs), African-Americans (AAs) in the United States have the highest sarcoidosis prevalence, a three-fold higher lifetime risk (2.4%), and greater disease severity and extra-thoracic involvement compared to non-Hispanic Whites (NHWs) [2-4]. The clinical heterogeneity and unpredictable disease course in sarcoidosis represents a challenge to early diagnosis and prognosis as the majority of sarcoidosis patients experience spontaneous recovery (uncomplicated sarcoidosis). However, a third of the patients develop complicated sarcoidosis defined as: i) multiorgan involvement [5,6], including cardiac and nervous system involvement, and ii) progressive fibrotic lung involvement with pulmonary function deterioration. Cardiac and nervous system involvement are thought to be severe and usually require aggressive treatment [7]. Environmental and genetic susceptibility factors have been associated with sarcoidosis [3,8,9] with current concepts invoking sarcoidosis development in genetically-predisposed individuals exposed to yet unknown environmental trigger(s) that result in dysregulated immune responses [8,10,11].

We and others have previously utilized genomic expression profiling of peripheral blood mononuclear cells (PBMCs), including RNA sequencing, microRNA and protein-coding gene expression data, to sub-phenotype subjects with uncomplicated or complicated sarcoidosis [12-19]. These studies highlighted immunological dysregulation in genes/pathways associated with the T-cell receptor, CXCL9 [20] and Jak-STAT signaling pathways involving pro-fibrotic genes/proteins, including HBEG, eNAMPT and ANG-2 (angiopoietin-2), which were identified as potential biomarkers of sarcoidosis severity [21-23]. Importantly, significant progress has been made in identifying genetic variants that contribute to the development of sarcoidosis and complicated sarcoidosis [24, 25] including SNPs in the butyrophillin-like 2 gene (BTNL2), a member of the B7 costimulatory receptor family [26, 27]; and in annexin A11 (ANXA11), a calcium-dependent phospholipid-binding protein involved in cell division and vesicle trafficking [28] which can influence granuloma formation and maintenance [29]. Similarly, the D allele polymorphism in the angiotensin-converting enzyme gene (ACE) [30, 31], and NOD2, which induces constitutive activation of the transcription factor NF-kB, are associated with early-onset sarcoidosis [32]. NOTCH4 was identified as a sarcoidosis-associated risk and severity gene in AA populations, and HLA-DRB1 SNPs as risk/severity SNPs for progressive sarcoidosis in European descent subjects [4, 33] and disease susceptibility in African descent subjects [4].

Multiple linkage and association studies, including the AC-CESS study interrogating genome-wide loci and familial clustering, have shown the HLA locus on the short arm of chromosome 6 to harbor variants that influence sarcoidosis risk and severity [11, 34-38]. Recently HLA- DRB1*03:01 and HLA-DQA1*05:01 alleles were associated with better prognosis in sarcoidosis [39], although MHC-related studies differ significantly among specific populations.

Our study was initially designed as a limited GWAS study of sarcoidosis cohorts comprised of European- and African-American subjects (EA, AA) with either uncomplicated or complicated sarcoidosis. Although this underpowered study failed to yield a single genome-wide significant SNP, by integrating a list of known expression quantitative trait loci (eQTL) expressed in multiple tissues with SNPs present in our GWAS platform, we successfully identified novel SNPs/genes associated with sarcoidosis susceptibility in our AA/EA sarcoidosis cohort which were represented in pathways involved in allograft rejection and MHC Class I and II genes. A number of eQTL/GWAS SNPs were successfully validated in a cohort of ~1000 genotyped sarcoidosis subjects. These studies indicate that despite the challenge of deciphering the genetic basis for sarcoidosis risk/severity, integrated eQTL/GWAS approaches may identify novel variants/genes, particularly HLA region and innate immunity variants/genes, that contribute to dysregulated innate immune responses and the risk of severe sarcoidosis.

Materials And Methods

Genetic analyses of European American and African American sarcoidosis subjects

A genome-wide association study (GWAS) was conducted utilizing the Affymetrix Genome-Wide Human SNP Array 6.0 Platform utilizing DNA samples and data obtained from a discovery population of 402 well-phenotyped European American and African American subjects (193 EA 209 AA) with proven sarcoidosis. DNA specimens were obtained from University of Chicago, University of Illinois at Chicago, National Jewish Health, University of California San Francisco, Johns Hopkins University, and the NHLBI repositories ACCESS and GRADS studies. A dataset from the Genetic Association Information Network (GAIN) (627 AAs, 579 EAs) was used as control (Accession phs000167.v1.p1). All studies had appropriate institutional review board approvals. The clinical characteristics of GWAS (n=402) participants are presented in Table 1. Complicated sarcoidosis was defined as presence of any indicator of parenchymal lung disease, (CT scan or radiographic Scadding stages III and IV; Forced vital capacity (FVC) <50%) and/or cardiac or neurological involvement. Due to limited information in our data sets multiorgan involent was not assessed. The non-complicated group, was defined as sarcoidosis was limited to skin or mediastinal lymphadenopathy. DNA from a second independent cohort of well-phenotyped sarcoidosis subjects (n=1034), was utilized for MassArray-based genotyping (Agena Bioscience) targeting eQTL-SNPs in a replication study.

We excluded 123 cases due to incomplete phenotype or low call rate genotype, leaving 911 subjects included in the analysis. The characteristics of the final number of successfully genotyped subjects in the validation cohort are presented in Table 2. In limited studies, a third distinct sarcoidosis cohort was utilized for further validation of the MAssArray genotyping results. These independent GWAS data set from Adrianto I, et al. included 818 AAs and 1088 controls and two replication independent set of 442 EAs cases and 455 AAs sarcoidosis cases, sample summary details have been previously published [2].

Table 1. Clinical characteristics of the sarcoidosis GWAS Discovery Cohort

	AA (n=209)	EA (n=193)
Gender (n, %)		
Female	143, 68	122, 63
Male	66,32	71, 37
Age (n, %)		
< 39 years old	81, 39	82, 43
40-60 years old	113, 54	95, 49
>61 years old	13, 6.2	14, 7
Non-complicated sarcoidosis (n, %)	134, 64	141, 73
Complicated sarcoidosis (n, %)	75, 36	51, 26

Table 2. Clinical characteristics of the sarcoidosis MassArray Validation Cohort

	AA (n=323)	EA (n=588)
Gender (n,%)		
Female	227, 70	323, 55
Male	96, 28	265, 43
Age (n,%)		
< 39 years old	101, 31	140, 24
40-60 years old	192, 54	355, 60
>61 years old	26, 8	87, 15
Non-complicated sarcoidosis (n, %)	218, 67	398, 68
Complicated sarcoidosis (n, %)	105, 32	190, 32

Sarcoidosis Discovery Cohort GWAS

High-quality DNA was genotyped on the Affymetrix®, 6.0 genotyping platform that included over 906,600 SNPs and more than 946,000 probes for the detection of copy number variation using GeneChip® Array station with the analytical process outlined in Supplemental Figure S2. Briefly, after restriction enzyme digestion, fragments were PCR-amplified, labeled and hybridized to the array. The Birdseed genotype calling algorithm was used as implemented by the Affymetrix Power Tools (APT) software. The APT software and PLINK 1.9 were used for quality control. 402 samples passed call rates > 95%, MAF > 1% and HWE >1e-06. Controls were obtained from a publicly-available database of GWAS data from the GAIN study and provided 627 AA and 579 EA GAIN controls. Genotyping data derived from the healthy control group of the GAIN studies was matched with sarcoidosis cases according to race, age, and gender (3:1 controls/case ratio). When genotyping calling for sarcoidosis cases was performed together with the GAIN controls, overall call rate was reduced (88.7%), with Hardy-Weinberg Equilibrium p= 3.92 x <10-25. Utilization of a genotyping call rate of 0.95 yielded 402,232 SNPs in EAs and 541,309 SNPs in AAs (Supplemental Figure S2). Allele and genotype calling data was formatted to PLINK [40] to determine variant associations with sarcoidosis with each EA and AA GWAS data initially processed together and then separately applying logistic regression and additive genetic models. Variant filtering and annotation was performed adjusting for 10 PCs with VarSeq (Golden Helix, Inc. Bozeman, MT).

Generation of an annotated eQTLs tissue-specific database

An extensive list of expression quantitative locus (eQTLs) was generated by comprehensive cis-eQTL search of the publicly-available database, Genotype-Tissue Expression (GTEx) version 7 (accession phs000424.vN.pN, October 2017) [41]. eQTLs previously identified from a total of 7 tissues frequently affected in sarcoidosis (lung, heart, arteries, spleen, skin, pancreas, brain) and blood were included. The intersection of these annotated eQTLs with the 541,309 non-intergenic SNPs present on the Affymetrix 6.0 array generate a tissue-specific pool of eQTL SNPs: lung (71,698), brain (332,322), heart (113,941), skin (90,737), arteries (212,554), spleen (34,482), pancreas (50,576) and whole blood (71,616). Elimination of redundant eQTLs yielded a pool of~46,900 SNPs that were present in both GTEX and GWAS datasets with required tissue representation in either brain, lung, heart, or blood (Figure 2).

eQTL selection and colocalization analysis

Logistic Model Analysis (PLINK) of the extracted eQTL/GWAS dataset was next performed in each of 4 comparison groups (EA or AA cases vs controls, AA or EA complicated vs uncomplicated sarcoidosis cases) generating eQTL SNPs/group with p<10-4 and prioritized for affecting multiple genes and pathways. The resulting 2000 SNPs were evaluated for linkage disequilibrium (LD) defined as r2 >0.25 in the grouped AA or EA or at least one of the 5 EUR subpopulations or at least one of the 7 AFR subpopulations (1000 Genome). We next identified the genes influenced by these selected eQTLs resided within 25K base upstream/downstream of 5' and 3' UTR of the gene region defined by Genome Reference Consortium Human Build 37 (GRCh37) (https://www.ncbi.nlm.nih.gov/ assembly/ GCF_000001405.13/). The depiction of the steps to obtain the annotated tissue-specific eQTLs in the 4 sarcoidosis comparison groups is shown in Figure 2.

Selection of eQTL SNPs for MassArray genotyping validation

From the 621 eQTL SNPs identified, the top 25 eQTL SNPs with the highest p value in each of the four comparison groups (AA uncomplicated, EA uncomplicated, AA complicated, EA complicated), were chosen for validation by MassArray genotyping in a validation cohort of >900 sarcoidosis subjects. In addition, the MassArray genotyping replication SNP list included SNPs selected from the 46,000 eQTL pool representing JAK-STAT pathway SNPs/genes postulated to be involved in sarcoidosis pathobiology (rs 7549445, rs6749232) [16]. Similarly, SNPs were included for validation which were previously highlighted in published studies of sarcoidosis risk and severity in AAs and EAs [2, 28, 38, 42-45]. Finally, we incorporated SNPs rs16910526 (CLEC7A), and rs5743809 (TLR6), as these were previously identified genetic studies of coccidiomycosis, a granulomatous disorder sharing a number of immunologic and histologic features with sarcoidosis [2, 16]. Thus, a total of 113 SNPs were selected for validation via MassArray genotyping (Supplimentary table 1). Lastly, these SNPs were compared to independent sarcoidosis GWAS datasets that included 818 AAs cases and 1088 AAs controls and two replication-independent sets of 442 EAs and 455 AAs sarcoidosis cases [2].
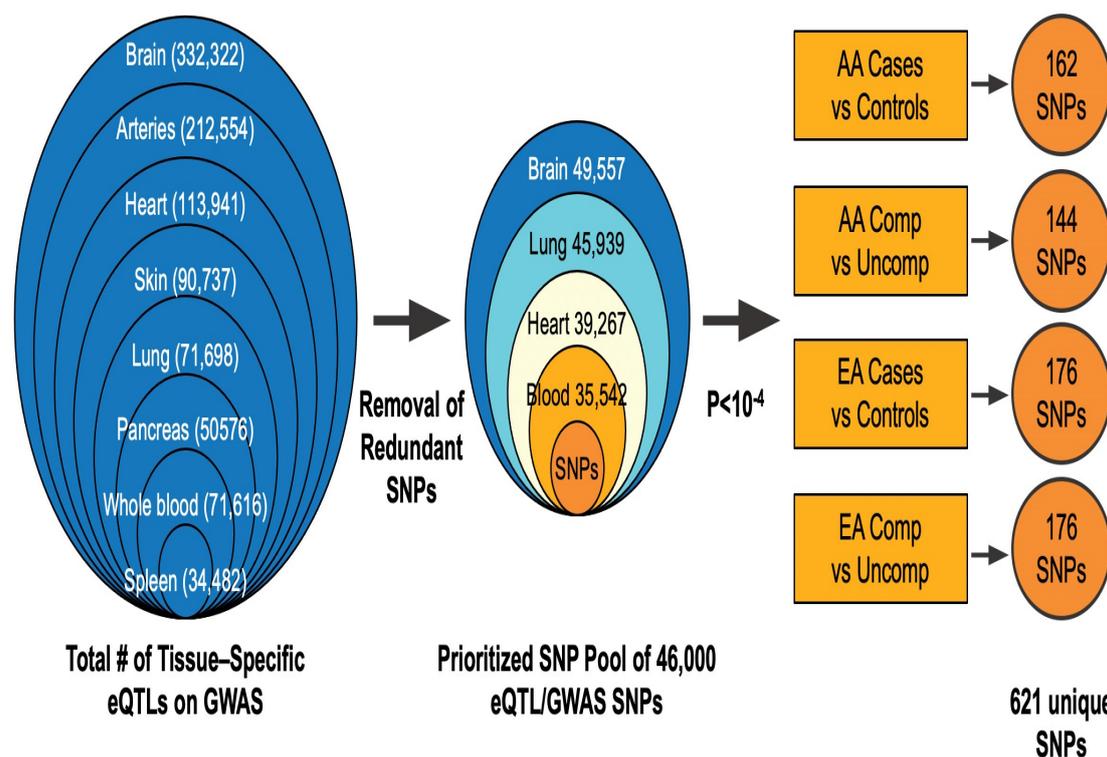
Figure 2. Expression quantitative locus (eQTL) selection and colocalization analysis. Logistic Model of the extracted SNPs from GWAS dataset was performed in each of 4 comparison groups (ED/AD cases vs controls, AD/ED complicated sarcoidosis cases vs uncomplicated) and then intersected with eQTLs from GTEx 7. from different tissues tissues (lung, brain, heart, skin, arteries, spleen, pancreas, and whole blood), excluding duplicates and prioritizing lung, brain, heart and blood tissues and categorized by race and complicated status according to p value (p< -E4).

MassArray genotyping in a validation cohort

Sarcoidosis genomic DNA was extracted from peripheral blood cells using a kit from Thermo Fisher Scientific Inc., according to the manufacturer's instructions. The Agena MassArray platform, iPLEX genotyping multiplex (Agena iPLEX assay, San Diego, CA, USA), was utilized to genotype targeted SNPs. Briefly, DNA was multiplexed end-point PCR followed by base extension, using MALDI-TOF mass spectrometry, the distinct mass of the extended primer identified each allele [46]. The Chi-square test was applied to assess the consistency of MassARRAY. Hardy-Weinberg equilibrium (HWE) was also calculated with a p value <0.05 (two-tailed) considered as statistical significance. The genotyping results of each SNP were calculated.

Statistical analysis

In the GWAS analysis, a Bonferroni correction for multiple testing (n=541,000 SNPs) was used as the genome-wide threshold for significance (alpha =3.50×10^{-8}). For eQTL colocalization analysis, the eCAVIAR program was utilized [47] to identify variants with evidence of disease and eQTL associations. The algorithm estimates the posterior probability that the same variant is causal in both a GWAS and eQTL study while accounting for LD. eCAVIAR requires a correlation matrix among the markers included in the analysis and PLINK v1.07(40) was used to calculate the r2. Logistic regression analysis with adjustment for race, complicated status, age, sex, gender, and population stratification (the first principal component) was used in an additive genetic model. Models were performed separately for EAs and AAs subjects. Enrichment analysis was conducted in ConsensusPathDB

[48]. Massarrayed genotypes were transformed into a VCF files and then analyzed using PLINK 2.0. SNPs having call rate $>95\%$ were included in the statistical analysis Hardy–Weinberg Equilibrium (HWE) was used for assessing the quality of genotypes.

Results

Clinical characteristics of the sarcoidosis cohorts

This study included 402 subjects in our Discovery GWAS Cohort and 911 subjects in the Validation Cohort. The clinical and demographic variables, for the Discovery Cohort are shown in Table 1 and for the Validation Cohort in Table 2. The majority of sarcoidosis subjects were female with ages between 40-60 years old; race was self-reported, the number of EAs were higher in the validation cohort compared to EAs in the Discovery Cohort. Sarcoidosis was categorized as complicated by indicators of parenchymal lung disease (CT scan or radiographic Scadding stages III and IV; Forced vital capacity (FVC) <50%) and/or cardiac or neurological involvement. Non-complicated sarcoidosis was defined as limited to skin or the mediastinal lymphadenopathy (by CT or radiographic Stages I and II).

GWAS analysis in sarcoidosis subjects and controls

Sarcoidosis GWAS case genotype data was merged with GAIN control genotyped data in PLINK format, utilizing the QC and genotyping parameters outlined in Supplemental Figure S2, including MAFs >0.05. For every case subject, three controls were selected (1:3 ratio) matching their gender and age group. Trend test was performed adjusting of 10 PCs with Golden Helix SVS. B. QC was performed for case data first. After merging with control data, strict QC was performed. When ge-

notyping calling for sarcoidosis cases was performed together with the GAIN controls, overall call rate was reduced (88.7%), with Hardy-Weinberg Equilibrium p= 3.92 x 10<-25. Utilization of a genotyping call rate of 0.95 yielded 402,232 SNPs in EAs and 541,309 SNPs in AAs (Supplemental Figure S2). The application of logistic regression and additive genetic model GWAS analyses (PLINK) of sarcoidosis subjects versus GAIN controls, however, failed to identify a single SNP which satisfied the conservative genome-wide threshold of Bonferroni correction (p< 3.50 x 10-8) and likewise failed to identify a single significant SNP in EA complicated vs non-complicated sarcoidosis subjects or AAs complicated vs non-complicated sarcoidosis subjects as depicted in the Manhattan plot shown in Figure 1. Two highest p value SNPs were rs5029975 in gene KNG1 and rs12652595 in gene LINC00992, however, these did not attain statistical significance.

Integration of GWAS and eQTL SNPs for identification of race-specific sarcoidosis risk variants

From the pool of 46,965 annotated eQTL SNPs common to the Affymetrix GWAS 6.0 platform (GTEX/GWAS) and prioritized for representation in sarcoidosis tissues (lung, brain, heart, blood) (Figure 2), we next filtered for eQTL/GWAS SNPs with a p value of <10-4 within each of the 4 phenotypic categories. This revealed 162 eQTL SNPs in the AA sarcoidosis vs GAIN controls analysis, 176 eQTL SNPs in EAs vs GAIN controls, 144 eQTL SNPs in AAs with complicated sarcoidosis and 176 eQTL SNPs in EAs with complicated sarcoidosis yielding a combined total of 657, of those 621 unique eQTL SNPs (p<10-4) (Figure 2). We next generated a list of 730 unique genes known to be linked to the 621 eQTL SNPs which were found to be variably represented in various gene ontology and pathway databases (Figure 3 and 4).

Enrichment analyses of sarcoidosis eQTLs

Enrichment analysis of the 730 gene set revealed that a total of 343 genes (49.1%) were represented in at least 1 Pathway (KEGG, Reactome) and 562 genes (80%) enriched in GO terms (Figure 3). The most significant eQTL/GWAS SNP-identified pathways are depicted in Figure 4 and include: Allograft Rejection, Graft vs Host Disease, Autoimmune, Phagosome, Interferon Signaling, Activation of C3 and C5, Immune System/Cytokine and T-Cell Receptor Signaling pathways.

We next analyzed the biologic pathways derived from eQTL/GWAS SNP-linked genes specific for each of the 4 comparison groups. The eQTL-regulated gene pathways in AA cases vs GAIN controls also yielded allograft rejection, as well as Interferon Gamma Signaling, Staphylococcus aureus Infection, Graft vs Host, Autoimmune, Translocation of ZAP-70 (a surface membrane protein expressed in T cells and natural killer cells) [49] and MHC class II Antigen Presentation (Figure 5A). Interestingly, interrogation of genes involved in these pathways demonstrated very strong overlap in complement-related genes (C2, C4A, C4B) and representation by HLA genes (HLA-DMA, HLA-DOB, HLA-C, HLA-DQA1, HLA-DQA2, HLA-DQB1, HLA-DQB2, HLA-DRB1, HLA-DRB5, HLA-B). Similarly, comparison of eQTL/GWAS-linked gene pathways in EA cases vs GAIN controls also highlighted C3 and C5 complement and coagulation pathway activation, allograft rejection, autoimmune, interferon signaling and antigen presentation and Class II MHC pathways (Figure 5B).

We next compared eQTL/GWAS-regulated gene pathways in EA and AA complicated vs. uncomplicated cases revealing an absence of HLA gene-related enrichment, a significant deviation from the comparisons with controls. In contrast, the most significant EA pathways were in Butyrophylin family interactions, mismatch repair and endosomal autophagy (Figure 5C). The
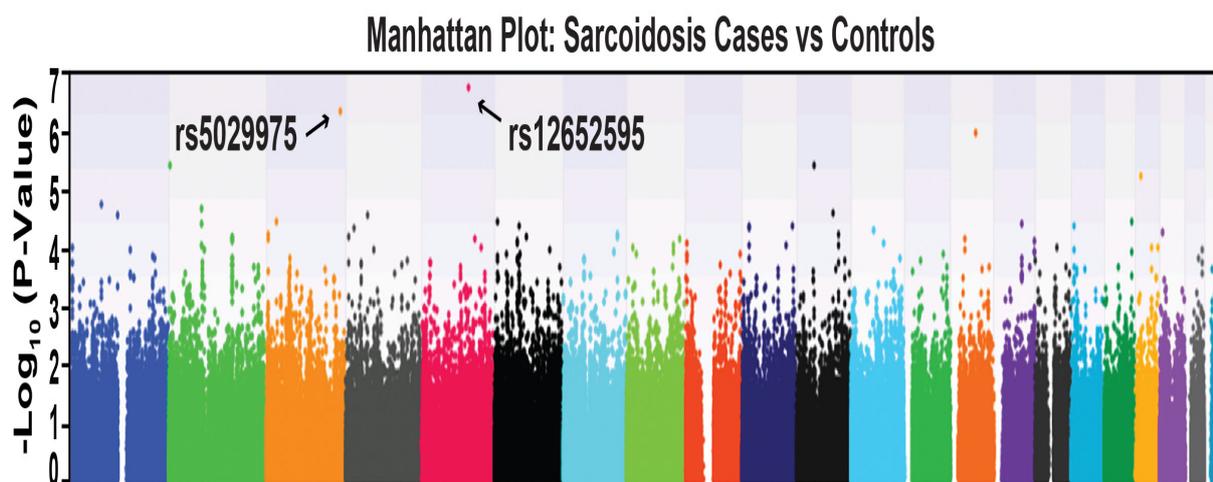

Figure 1. The Manhattan plot shows false positive association in EAs and AAs combines signatures. Application of logistic regression and additive genetic model GWAS analyses in the 404 sarcoidosis subjects (MAF<5%, Call rate 98%, HWE P<0.01) failed to identify any SNP which satisfied the conservative genome-wide threshold. Shown are two highest p value SNPs that did not attain statistical significance, rs5029975 in gene KNG1 and rs12652595 in LINC00992. KNG1 encodes kininogen-1, the precursor protein to high-molecular-weight kininogen (HMWK) and essential for blood coagulation and assembly of the kallikrein-kinin system. LINC00992 is a long non-coding RNAs (lncRNAs) in cellular processes during tumor progression. Neither SNP/gene has been associated with sarcoidosis.
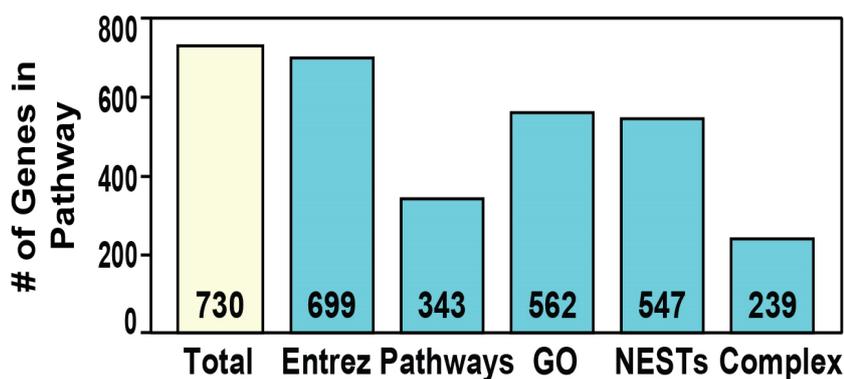
Figure 3. Enrichment analysis of the eQTL-associated genes revealed a total of 343 genes (49.1%) represented in at least 1 pathway (KEGG, Reactome) and 562 genes (80%) enriched in GO terms.
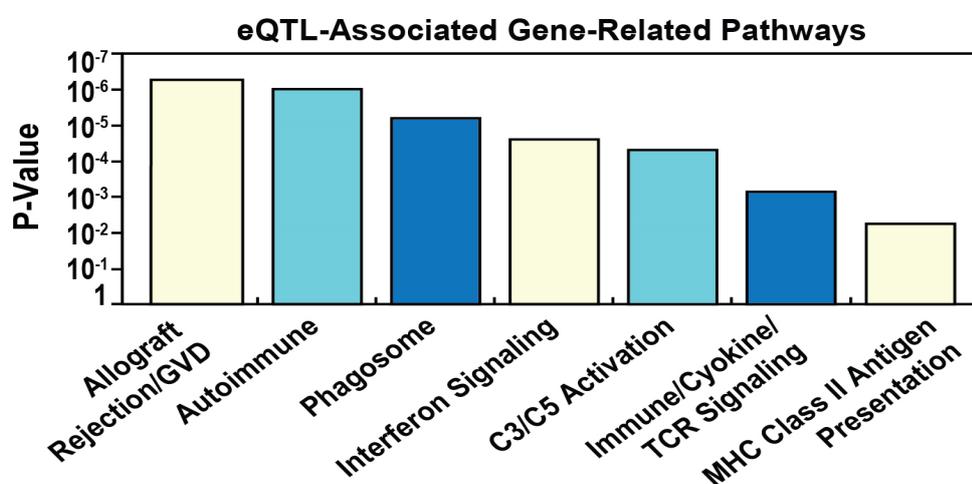

Figure 4. Pathway analysis of sarcoidosis vs controls. A total of 343 of the 730 genes (49.1%) were represented in at least 1 Pathway. The top most significant eQTL/GWAS SNP-related pathways include: Allograft Rejection, Graft vs Host Disease, Autoimmune Phagosome, Interferon signaling, Activation of C3 and C5, Immune system/Cytokine and TCR signaling and MHC related pathways.

BTN3A2 and BTN3A3 genes in the butyrophylin family interactions pathway are known sarcoidosis candidate genes [2, 26, 43]. eQTL/GWAS-regulated gene pathways comparisons in AA complicated vs uncomplicated cases revealed purine nucleotide metabolism (ENTPD6, ADK4/5/6) and directed repair pathways (MRE11, RTEL1, EME2) (Figure 5C).

Confirmation of eQTL/GWAS SNPs in a sarcoidosis validation cohort: high representation of the HLA region

We next sought to validate data derived from our eQTL/GWAS Discovery Cohort and from each of the four EA/AA comparison groups (complicated vs uncomplicated). A total of 113 eQTL/GWAS SNPs were selected, with genotyping data successfully generated for 103 of 113 SNPs included, with 54 SNPs called over 99.5% success rate, 467 of the 1034 (45.2%) samples in the validation cohort, genotyped at a success rate of 100% of valid SNPs, calculated out of 103 SNPs (assays with call rates > 80%). Results analysis revealed 31/103 eQTL/GWAS SNPs (27%) exhibited an eQTL p-value of 10^{-15} or greater, with 13 eQTL/GWAS SNPs identified to be highly significantly associated with risk of sarcoidosis (p <0.005) (Table 3). Supplemen-

tal Figure S3 shows the scatter plots with the mass intensity of each allele of those 13 SNPs. Filtering in each group was based upon p-values on the MassArray assay and the number of genes contained in eQTL pathways (Supplimentary table 1). These results showed that the most significant eQTL/GWAS SNPs resided within HLA region genes on chromosome 6 (rs2596509, rs443532, rs502771), with rs502771 in close proximity to HLA-DRB1 and NOTCH 4, a significant validated SNP reported to be linked to sarcoidosis risk in AAs (p =1.5 x 10^{-4}). In addition to HLA gene-related SNPs, our list of validated SNPs included inflammation-related genes such as IL27RA, RAB23, C100RF67, OS9, CCDC88B, and NOTCH4 genes, previously identified as associated with sarcoidosis risk. Only rs1442533 achieved statistical significance in the complicated EA subgroup (p= 3.28 x 10^{-3}). No SNP was found to be significant in the AA complicated subphenotype analysis (Table 3).

eQTL/GWAS SNPs compared to independent sarcoidosis GWAS datasets

Finally, we compared our 103 validated eQTL/GWAS SNPs to independent sarcoidosis GWAS datasets that included a Dis-
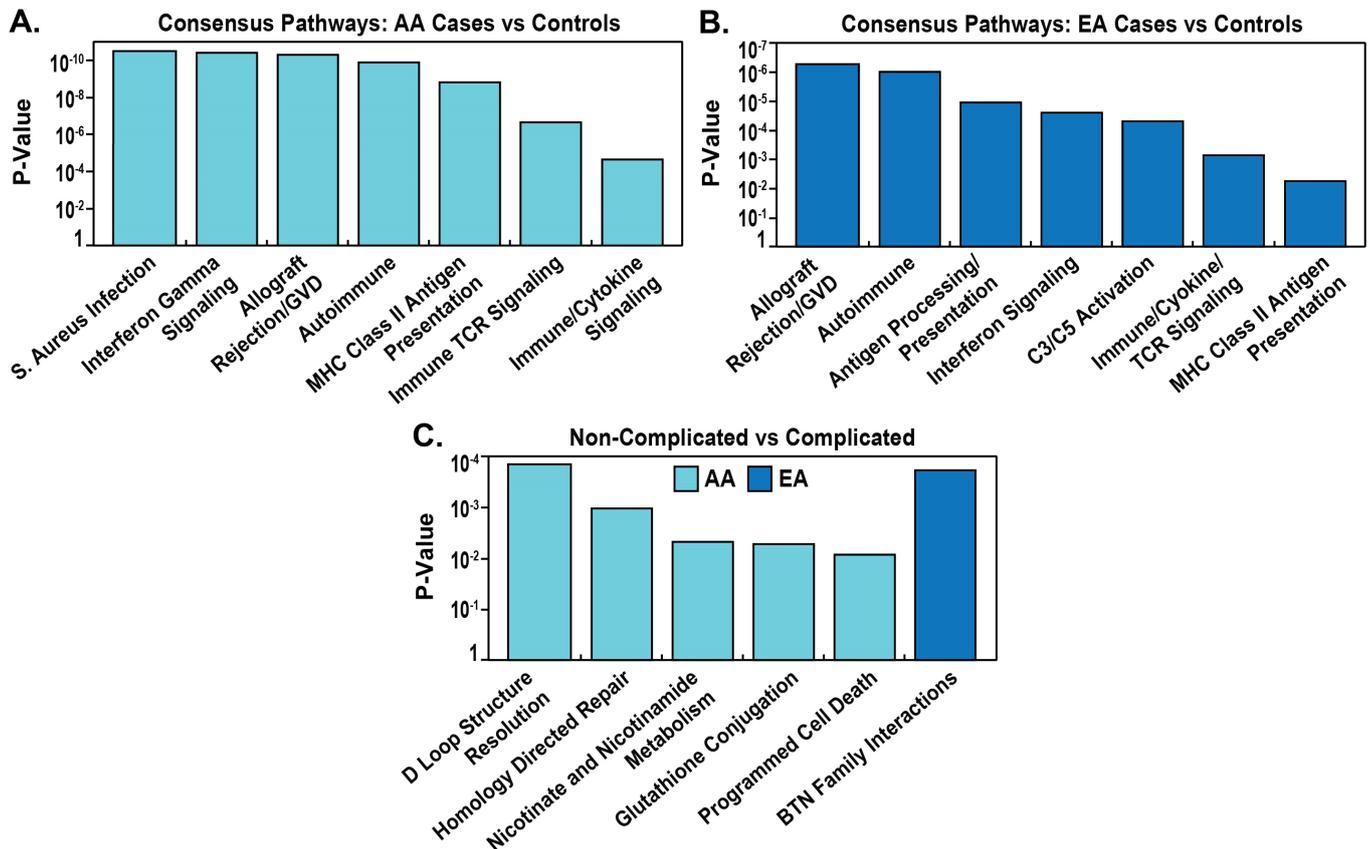
Figure 5. Pathway enrichment analysis of eQTL/GWAS-derived genes in AAs and EAs by complicated-and non-complicated sarcoidosis.

covery cohort of 818 AAs and 1088 controls, an independent set of 442 EAs cases and 455 AAs sarcoidosis cases (Replication Set) [2]. These analyses confirmed five QTL/GWAS SNPs validated by Massarray that were also significantly associated with sarcoidosis in either the EA, AA discovery sets or the AA replication set (Table 4). These eQTL/GWAS SNPs included SNPs present in the AA sets (MICA- rs11753208; POLR1B-rs11677050) and in EA cohorts (LINC00174-rs7787110; HCP5-rs3094228; RPS6KB2- rs1476792) (Table 4). In contrast to MassArray genotyping study results, the rs1476792 SNP (p < 0.0098) in the EA GWAS dataset was significant in AAs RPS6KB2/PTPRCAP -rs1476792 (p<0.007) (Table 4).

Discussion

Prior sarcoidosis GWAS studies have successfully yielded important insights into the genetic basis for sarcoidosis risk and severity [50, 51], a significant feat given the requirement for large sample sizes and sufficient power to survive Bonferroni correction. Somewhat predictably, the analyses of our underpowered GWAS highlighted several SNPs which approached, but did not meet, genome-wide significance. Interestingly, while not significant in our analysis, several SNPs identified (Figure 1) reside within genes which exhibit biologic plausibility for an association with sarcoidosis severity and susceptibility including PTPRR , a MAPK phosphorylation regulator, down-regulated gene in lung fibrosis and IPF [52] and whose expression we have shown to be prominently dysregulated in PBMCs from sarcoidosis subjects [53]. Additional GWAS-identified SNPs included SNPs residing in ZFHX3, which encodes a transcription factor implicated in the development of inflammation and

fibrosis [54], and in FAM19A5, encoding a biologic target in treatment of fibrosis [55]. Similarly, rs6748088 in FAM117B has been associated to sarcoidosis risk [27]. The HURP-generated protein promotes epithelial-to-mesenchymal transition which is recognized as involved in organ fibrosis in lung and neoplasia [56]. Nevertheless, the lack of a significant GWAS signal reduces the reliability of these SNPs as genetic predictors of sarcoidosis risk. As a result, we chose to pivot to an approach which leverages the increasing awareness of eQTL influences on disease risk/ severity through regulatory impact on gene expression [47]. As a result, we integrated the readily available catalogs of eQTL SNPs with our GWAS platform as an approach designed to enhance the capacity for an underpowered GWAS studies to identify genomic regions and affected genes/pathways that associate with sarcoidosis risk and severity.

A total of 621 sarcoidosis-associated SNPs were identified via an eQTL/GWAS filtering approach utilizing our AA and EA GWAS dataset with comparison to GAIN GWAS controls. We found these eQTL/GWAS SNPs to influence 730 genes involved in pathways germane to sarcoidosis pathobiology including allograft rejection, graft vs host disease, and autoimmunity (Type I Diabetes Mellitus, Rheumatoid Arthritis, Inflammatory Bowel Disease), and major histocompatibility complex (MHC) Class II antigen presentation. The GO ontology database also highlighted MHC Class II protein/complex and T-cell receptor activity. Over 50% of the pathway candidates were represented within the 730 gene pool.

The HLA locus on the short arm of chromosome 6 encodes MHC Class II genes/proteins heavily involved in allograft rejec-

Table 3. SNPs Validated via MassArray Genotyping in a Replication Cohort (each category compared to the original GWAS subgroups).

eQTL SNP	Position	GWAS P-value	Validation analysis	MassArray	
P-value	Genes affected				
rs7248735	chr19:14019403	2.70E-03	AD and ED case vs controls	4.90E-03	IL27RA
rs10044736	chr5:138985120	5.46E-04	AD case vs control	8.32E-03	CTNNA1 SIL1
rs11677881	chr2:113760287	5.13E-04	AD case vs control	3.07E-03	DDX11L2 FAM138B PGM5P4 PG-M5P4-AS1 RABL2A RPL23AP7 WASH2P
rs1476792	chr11:67428766	3.78E-03	AD case vs control	6.97E-03	PTPRCAP RPS6KB2
rs17787966	chr11:47048874	8.11E-04	AD case vs control	7.29E-03	ACP2 ARHGAP1 C11orf49 C1QTNF4 DDB2 LRP4 LRP4-AS1 PACSIN3
rs2054517	chr12:75597666	2.09E-04	AD case vs control	2.65E-04	GLIPR1 GLIPR1L2 KRR1 LOC105369844
rs2596509	chr6:31350636	9.82E-04	AD case vs control	6.59E-03	ATP6VIG2 CCHCR1 HLA-B HLA-S MICA PSORS1C1 PSORS1C2 STK19B VARS2 ZBTB12
rs443532	chr17:8884178	4.95E-03	AD case vs control	6.91E-03	PIK3R5
rs502771	chr6:32611193	1.52E-04	AD case vs control	1.50E-04	C4A C4B CLIC1 CYP21A1P CYP21A2 HLA-DMA HLA-DQA1 HLA-DQB1 HLA-DQB1-AS1 HLADQB2 HLADRB1 HLA-DRB5 HLA-DRB9 NOTCH4 PSMB9
rs6779819	chr3:50681914	8.00E-04	ED case vs control	3.28E-03	C3orf18 DOCK3 HEMK1 LINC02019 MAP-KAPK3 TEX264
rs7219	chr17:75319287	2.53E-03	ED case vs control	3.28E-03	NUP85 MRPS7 GRB2 ANXA5 ILKAP
rs1442533	chr9:17379148	3.69E-04	ED complicated vs non-complicated	1.49E-03	CNTLN
rs7549445	chr1:65054080	1.02E-03	AD case vs control	3.21E-05	JAK1 RP11-182I10.3

Table 4. Significant MassArray SNPs Confirmed in Independent GWAS Datasets.

CHR	rs ID	Genes	ORAA_Disc	PAA_Disc	ORAA_Rep	PAA_Rep	PAA_Meta	OREA	PEA
6	rs11753208	MICA	1.205	4.60E-02	8.66E-01	2.98E-01	3.18E-01	1.22E+00	1.17E-01
2	rs11677050	POLR1B	1.258	7.97E-03	7.60E-01	2.14E-02	4.30E-01	1.02E+00	7.71E-01
7	rs7787110	LINC00174	1.025	6.77E-01	8.69E-01	1.42E-01	5.98E-01	1.19E+00	2.50E-02
6	rs3094228	HCP5	0.9457	8.89E-01	1.07E+00	5.53E-01	8.14E-01	1.34E+00	1.24E-03
11	rs1476792	RPSKB2	1.066	8.17E-01	9.76E-01	8.44E-01	9.43E-01	8.16E-01	9.77E-03

ORAA Disc: Odds Ratio for this SNP in the AA discovery set. PAA_Disc: P-value for the SNP in the AA discovery set. ORAA_Rep: Odds Ratio for this SNP in the AA replication set. PAA_Rep: P-value for the SNP in the AA replication set. PAA_Meta: P-value from the meta-analysis of the AA discovery and the AA replication set. OREA: Odds Ratio for this SNP in the EA set. PEA:P-value for the SNP in the EA set

tion, antigen processing and antigen presentation and has been associated with sarcoidosis risk in both black and white populations [4, 11, 33, 57]. Candidate gene studies and GWAS studies in European- and African-descent subjects have indicated the HLA region to harbor genetic risk loci, highlighted by SNPs with significant association with sarcoidosis susceptibility within HLA-C, HLA-B, HLA-DRA, HLA-DRB5, HLA-DRB1, HLA-DPB1, HLA-DQA1, and HLA-DQB1 [37, 43, 57, 58]. However, the exact influence on risk/severity appears to be race-specific. For example, the DRB1 (*1101) allele is associated with increased risk of sarcoidosis in ADs and EDs [35], whereas DRB1 (*0101), DQB1 (*0501), and DQA1 (*0101) alleles were associated with reduced risk of sarcoidosis in diverse ED populations (UK, Netherlands), and in Japanese subjects [11, 43, 59, 60]. Our eQTL/GWAS results implicating MHC Class I and II genes are quite consistent with prior candidate gene/GWAS sarcoidosis studies [60] as we identified MHC Class I genes HLA-B and HLA-C to

be well represented in EAs (compared to controls) and HLA-B primarily in AAs. MHC Class II genes represented in both EAs and AAs included HLA-DRB1 and HLA-DQA2 whereas in our study only AAs exhibited eQTL/GWAS SNPs in HLA-DMA, HLA-DOB, HLA-DQA1, HLA-DRB5, and HLA-DQB2. A recent study utilizing a similar eQTL/GWAS approach (GTEx database) but restricted to ED-only sarcoidosis subjects, identified significant sarcoidosis-associated SNPs near HLA-DQA1, an HLA allele also identified by our eQTL/GWAS approach [51]. This same study identified HLA-DRB9 and HLA-DRB5 as associated with sarcoidosis risk. Although we failed to find HLA-DRB5 as a risk locus in EDs, it was present in AAs in our study. Interestingly, the rs1964995 SNP located between HLA-DRB9 and HLA-DRB5, a risk SNP only in AAs in our study, was protective in a Swedish cohort with non-Lofgren sarcoidosis GWAS [47], but increased the risk of rheumatoid arthritis in ADs, but not in EDs [48]. Another SNP near HLA-

DQA1 (rs2187668) is significantly associated with Lofgren's syndrome [47] and HLA-DRB1*04:01 with ocular sarcoidosis in EDs [61]. The DRB1*1501 allele, in high linkage disequilibrium with DQB1*0602, is associated with increased risk of severe pulmonary sarcoidosis in EDs [20].

As a rigorous validation strategy, we assembled a list of SNPs selected from the top performing eQTLs in each of the 4 comparison groups and performed MassArray genotyping in an ~1000 subject validation sarcoidosis cohort. These results showed that 31/105 eQTL/GWAS SNPs (30%) exhibited an eQTL p -value < 10-15. The list of validated SNPs included rs502771 selected after comparison of ADs with sarcoidosis to GAIN controls residing in the HLA chromosome 6p21.3 region harboring HLA-DQA and HLA-DRB5, as well as rs9268362 and rs3094228 identified in EDs and rs502771 and rs7746922 in ADs with sarcoidosis affecting expression of HLA-DRB6 and NOTCH4, a candidate gene in the NOTCH family of T-cell immune response that regulates the activation of Toll-like receptors in macrophages, known to drive chronicity of lung inflammation [62, 63]. NOTCH4 expression is downregulated in sarcoidosis granulomas [2, 4, 11, 23, 35] and expression of the NOTCH receptor ligand is increased in cutaneous sarcoidosis granulomas [42]. A rare SNP in NOTCH4 (rs715299, MAF <0.00003) is associated with sarcoidosis in ADs and EDs independently from others in the MHC region [2].

Our eQTL/GWAS analysis also identified rs2393915 in EAs with complicated sarcoidosis associated to BTN3A2, BTN3A3 and BTN2A3P, MHC I-associated genes; other BTNL2 polymorphisms (rs2076530, rs206530), have been previously reported associated to sarcoidosis susceptibility [26, 43, 64-69]. Additional SNPS (rs6922431, rs2596509, rs11753208, rs11758964 and rs11753208) were identified as associated to MICA, which encodes the highly polymorphic major histocompatability complex class I chain-related protein A.

In addition to the HLA-related eQTL/GWAS SNPs identified, our MassArray validation studies confirmed 13 of eQTL/GWAS SNPs that were significantly associated with risk of sarcoidosis (p <0.005), including validated SNPs influencing inflammation-related genes such as IL27RA, RAB23, PSORS1C1, CTNNA1, MICA, and NOTH4. The majority of these SNPs were validated in AAs, with rs7549445, the top ranked SNP by p value, related to JAK1 gene, a Janus kinase, key component of the interleukin-6 (IL-6)/JAK1/STAT3 immune and inflammation response, supporting the importance of the JAK-STAT pathway in sarcoidosis [16, 70]. The rs1476792 SNP was also validated in ADs and is associated to RPS6KB2 (ribosomal protein), previously implicated in idiopathic pulmonary fibrosis and lung radiotoxicity via regulation of the mTOR signaling pathway [71, 72] and was dysregulated in World Trade Center responders with associated sarcoidosis [73]. Our MassArray validation study failed to identify eQTL/GWAS SNPs in HLA genes that were involved in sarcoidosis severity in AAs, however we identified, rs1442533, related to CNTLN (Centlein, centriole stabilizer) associated with complicated sarcoidosis in EDs. Finally, our MassArray-targeted SNPs were also cross validated in independent GWAS datasets. This comparison identified five SNPs (rs11753208, rs11677050, rs7787110, rs3094228, rs1476792) of the thirteen eQTL/GWAS SNPs validated in our MassArray genotyping studies that were also present in the independent GWAS sarcoidosis datasets.

There are several limitations of the current study with the most glaring limitation being the small sample size of participants in the Discovery cohort undergoing GWAS interrogation. We also acknowledge the limitation related to using self-reported ancestry for the analysis. However, unlike many other genetic studies that include only individuals of European descent, our study design included sarcoidosis subjects of African descent (52%). Second, our study population was a heterogeneous population of sarcoidosis subjects manifesting multiple phenotypes thereby reducing study power for identify SNPs driving risk and/or severity. Nevertheless, our results indicate success in identifying HLA- associations that have been linked to other sarcoidosis phenotypes in previous studies such as DRB1*1101, *1501 [4, 35]. We believe our integration of eQTL/GWAS SNPs have allowed at least partial mitigation of this limitation. By utilizing the validation cohort of ~1000 sarcoidosis subjects and a mixture of different sarcoidosis phenotypes, we did not restrict enrollment to severe pulmonary sarcoidosis or other specific phenotypes. Future GWAS studies would benefit from focusing on different sarcoidosis phenotypes or those with specific organ involvement to explore genetic drivers of sarcoidosis phenotype manifestations, including neurologic sarcoidosis, cardiac sarcoidosis, or specific types of pulmonary sarcoidosis.

In summary, our findings indicate that combining a robust eQTL pool and SNPs present on our GWAS platform, and implementing multilayered replication studies and adding downstream work such as gene-set enrichment and pathway analysis, provides a useful strategy to address underpowered GWAS issues. Our results provide additional evidence that underscore HLA alleles as important functional contributors to risk of sarcoidosis disease development and progression. These eQTL/GWAS data showed that these loci act as cis-regulators of genes within pathways that have been previously identified as dysregulated in complicated sarcoidosis: allograft rejection, graft vs host, autoimmunity and innate immunity. These findings support the current understanding of disease pathogenesis as a product of the gene environment interaction and suggest that previously unknown factors in the HLA region may be regulating this response. Future studies integrating the eQTL mapping with variants discovered by GWAS may be informative for discovering genes and pathways disrupted in sarcoidosis and influencing severity of other fibrotic pulmonary disorders.

Abbreviations

GWAS: genome wide association studies; eQTL: expression quantitative trait loci; SNPs: small nuclear polymorphisms; AA: African American; EA: European American; AD: Affricant Descent; ED: European Descent; NHWs: non-Hispanic Whites, HWE: Hardy–Weinberg Equilibrium.

Supplementary Material

Supplementary figures.

https://europeanjournalofrespiratorymedicine.com/vol5-1/138Sfigure

Supplementary table.

https://europeanjournalofrespiratorymedicine.com/vol5-1/138Stable

Competing interests

Joe GN Garcia MD is CEO and Founder of Aqualung Therapeutics Corporation. David R. Moller MD is Chairman and CTO of Sarcoidosis Diagnostic Testing, LLC. All other authors report no conflict of interest.

Consent for publication

Not applicable

Availability of data and materials

The datasets generated and/or analysed during the current study are available upon request to the corresponding author.

References

1. Crouser ED, Maier LA, Wilson KC, Bonham CA, Morgenthau AS, Patterson KC, et al. Diagnosis and Detection of Sarcoidosis. An Official American Thoracic Society Clinical Practice Guideline. Am J Respir Crit Care Med. 2020; 201: e26-e51.

2. Adrianto I, Lin CP, Hale JJ, Levin AM, 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.

3. Liu H, Patel D, Welch AM, Wilson C, Mroz MM, Li L, et al. Association Between Occupational Exposures and Sarcoidosis: An Analysis From Death Certificates in the United States, 1988-1999. Chest. 2016; 150: 289-298.

4. Levin AM, Adrianto I, Datta I, Iannuzzi MC, Trudeau S, Li J, et al. Association of HLA-DRB1 with Sarcoidosis Susceptibility and Progression in African Americans. Am J Respir Cell Mol Biol. 2015; 53: 206-216.

5. Lockstone HE, Sanderson S, Kulakova N, Baban D, Leonard A, Kok WL, et al. Gene set analysis of lung samples provides insight into pathogenesis of progressive, fibrotic pulmonary sarcoidosis. Am J Respir Crit Care Med. 2010; 181: 1367-1375.

6. Chen ES, Moller DR. Sarcoidosis--scientific progress and clinical challenges. Nat Rev Rheumatol 2011; 7: 457-467.

7. Sauer WH, Stern BJ, Baughman RP, Culver DA, Royal W. High-Risk Sarcoidosis. Current Concepts and Research Imperatives. Ann Am Thorac Soc 2017; 14: S437-s444.

8. Fingerlin TE, Hamzeh N, Maier LA. Genetics of Sarcoidosis. Clin Chest Med 2015; 36: 569-584.

9. Rybicki BA, Iannuzzi MC, Frederick MM, Thompson BW, Rossman MD, Bresnitz EA, et al. Familial aggregation of sarcoidosis. A case-control etiologic study of sarcoidosis (ACCESS). Am J Respir Crit Care Med. 2001; 164: 2085-2091.

10. Grunewald J. Genetics of sarcoidosis. Curr Opin Pulm Med. 2008; 14: 434-439.

11. Sato H, Woodhead FA, Ahmad T, Grutters JC, Spagnolo P, van den Bosch JMM, et al. Sarcoidosis HLA class II genotyping distinguishes differences of clinical phenotype across ethnic groups. Hum Mol Genet. 2010; 19: 4100-4111.

12. Ascoli C, Huang Y, Schott C, Turturice BA, Metwally A, Perkins DL, et al. A Circulating MicroRNA Signature Serves as a Diagnostic and Prognostic Indicator in Sarcoidosis. Am J Respir Cell Mol Biol. 2018; 58: 40-54.

13. Crouser ED, Hamzeh NY, Maier LA, Julian MW, Gillespie M, Rahman M, et al. Exosomal MicroRNA for Detection of Cardiac Sarcoidosis. Am J Respir Crit Care Med. 2017; 196: 931-934.

14. Crouser ED, Julian MW, Crawford M, Shao G, Yu L, Planck SR, et al. Differential expression of microRNA and predicted targets in pulmonary sarcoidosis. Biochem Biophys Res Commun. 2012; 417: 886-891.

15. Singla S, Zhou T, Javaid K, Abbasi T, Casanova N, Zhang W, et al. Expression profiling elucidates a molecular gene signature for pulmonary hypertension in sarcoidosis. Pulm Circ. 2016; 6: 465-471.

16. Zhou T, Casanova N, Puladi N, Wang T, Lussier Y, Kox SK, et al. Identification of Jak-STAT signaling involvement in sarcoidosis severity via a novel microRNA-regulated peripheral blood mononuclear cell gene signature. Scientific Reports. 2017.

17. Vukmirovic M, Yan X, Gibson KF, Gulati M, Schupp JC, DeIuliis G, et al. Transcriptomics of bronchoalveolar lavage cells identifies new molecular endotypes of sarcoidosis. Eur Respir J. 2021; 58.

18. Nancy GC, Tong Z, Jorge N, Kenneth SK, Joe GNG. Peripheral Blood MicroRNA Signature Differentiates Sarcoidosis. A108 MICRO RNAS, RNA SEQ, LNCRNA: BIOLOGY AND FUNCTION: American Thoracic Society; 2016. p. A2809-A2809.

19. Koth LL, Solberg OD, Peng JC, Bhakta NR, Nguyen CP, Woodruff PG. Sarcoidosis blood transcriptome reflects lung inflammation and overlaps with tuberculosis. Am J Respir Crit Care Med. 2011; 184: 1153-1163.

20. Su R, Li MM, Bhakta NR, Solberg OD, Darnell EP, Ramstein J, et al. Longitudinal analysis of sarcoidosis blood transcriptomic signatures and disease outcomes. Eur Respir J. 2014; 44: 985-993.

21. Zhou T, Zhang W, Sweiss NJ, Chen ES, Moller DR, Knox KS, et al. Peripheral blood gene expression as a novel genomic biomarker in complicated sarcoidosis. PLoS One. 2012; 7: e44818.

22. Garcia AN, Casanova NG, Kempf CL, Bermudez T, Valera DG, Song JH, et al. eNAMPT is a Novel DAMP that Contributes to the Severity of Radiation-Induced Lung Fibrosis. Am J Respir Cell Mol Biol. 2022.

23. Casanova NG, Gregory T, Sun B, Bermudez T, Hufford MK, Oita RC, et al. Biochemical and genomic identification of novel biomarkers in progressive sarcoidosis: HBEGF, eNAMPT, and ANG-2. Frontiers in Med. 2022; 25.

24. Schupp JC, Vukmirovic M, Kaminski N, Prasse A. Transcriptome profiles in sarcoidosis and their potential role in disease prediction. Curr Opin Pulm Med. 2017; 23: 487-492.

25. Casanova N, Zhou T, Knox KS, Garcia JG. Identifying Novel Biomarkers in Sarcoidosis Using Genome-Based Approaches. Clin Chest Med. 2015; 36: 621-630.

26. 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: 357-364.

27. Fischer A, Ellinghaus D, Nutsua M, Hofmann S, Montgomery CG, Iannuzzi MC, et al. Identification of Immune-Relevant Factors Conferring Sarcoidosis Genetic Risk. Am J Respir Crit Care Med. 2015; 192: 727-736.

28. Hofmann S, Franke A, Fischer A, Jacobs G, Nothnagel M, Gaede KI, et al. Genome-wide association study identifies ANXA11 as a new susceptibility locus for sarcoidosis. Nat Genet. 2008; 40: 1103-1106.

29. Mirsaeidi M, Gidfar S, Vu A, Schraufnagel D. Annexins family: insights into their functions and potential role in pathogenesis of sarcoidosis. J Translational Med. 2016; 14: 89.

30. Song GG, Kim JH, Lee YH. Associations between the angiotensin-converting enzyme insertion/deletion polymorphism and susceptibility to sarcoidosis: A meta-analysis. J Renin Angiotensin Aldosterone Syst. 2015; 16: 219-226.

31. Schurmann M, Reichel P, Muller-Myhsok B, Dieringer T, Wurm K, Schlaak M, et al. Angiotensin-converting enzyme (ACE) gene polymorphisms and familial occurrence of sarcoidosis. J Intern Med. 2001; 249: 77-83.

32. Kambe N, Nishikomori R, Kanazawa N. The cytosolic pattern-recognition receptor Nod2 and inflammatory granulomatous disorders. J Dermatol Sci. 2005; 39: 71-80.

33. 33. Spagnolo P, Sato H, Grunewald J, Brynedal B, Hillert J, Mañá J, et al. A common haplotype of the C-C chemokine receptor 2 gene and HLA-DRB1*0301 are independent genetic risk factors for Löfgren's syndrome. J Int Med. 2008; 264: 433-441.

34. Acosta-Herrera M, Kerick M, Lopéz-Isac E, Assassi S, Beretta L, Simeón-Aznar CP, et al. Comprehensive analysis of the major histocompatibility complex in systemic sclerosis identifies differential HLA associations by clinical and serological subtypes. Ann Rheum Dis. 2021; 80: 1040-1047.

35. Rossman MD, Thompson B, Frederick M, Maliarik M, Iannuzzi MC, Rybicki BA, et al. HLA-DRB1*1101: a significant risk factor for sarcoidosis in blacks and whites. Am J Hum Genet. 2003; 73: 720-735.

36. Rybicki BA, Maliarik MJ, Major M, Popovich J, Iannuzzi MC. Genetics of sarcoidosis. Clin Chest Med. 1997; 18: 707-717.

37. Schurmann M, Bein G, Kirsten D, Schlaak M, Muller-Quernheim J, Schwinger E. HLA-DQB1 and HLA-DPB1 genotypes in familial sarcoidosis. Respir Med. 1998; 92: 649-652.

38. Hofmann S, Fischer A, Till A, Muller-Quernheim J, Hasler R, Franke A, et al. A genome-wide association study reveals evidence of association with sarcoidosis at 6p12.1. Eur Respir J. 2011; 38: 1127-1135.

39. Felten M, Ferencik S, Teixeira Alves LG, Letsiou E, Lienau J, Müller-Redetzky HC, et al. Ventilator-induced Lung Injury Is Modulated by the Circadian Clock. Am J Respir Crit Care Med. 2023; 207: 1464-1474.

40. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81: 559-575.

41. Ongen H, Brown AA, Delaneau O, Panousis NI, Nica AC, Dermitzakis ET, et al. Estimating the causal tissues for complex traits and diseases. Nature Genetics. 2017; 49: 1676-+.

42. Rangel JR, Chung Y, Rosenbach M, Ziober AF, Zhang PJ, Schaffer A. Expression of Notch signaling components in cutaneous foreign body and sarcoidal granulomas and fusing macrophages. Am J Dermatopathol. 2014; 36: 409-413.

43. Suzuki H, Ota M, Meguro A, Katsuyama Y, Kawagoe T, Ishihara M, et al. Genetic characterization and susceptibility for sarcoidosis in Japanese patients: risk factors of BTNL2 gene polymorphisms and HLA class II alleles. Investigative ophthalmol visual sci. 2012; 53: 7109-7115.

44. 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: 13-18.

45. Mrazek F, Stahelova A, Kriegova E, Fillerova R, Zurkova M, Kolek V, et al. Functional variant ANXA11 R230C: true marker of protection and candidate disease modifier in sarcoidosis. Genes Immun. 2011; 12: 490-494.

46. Gabriel S, Ziaugra L, Tabbaa D. SNP genotyping using the Sequenom MassARRAY iPLEX platform. Curr Protoc Hum Genet 2009; Chapter 2: Unit 2.12.

47. Hormozdiari F, van de Bunt M, Segre AV, Li X, Joo JWJ, Bilow M, et al. Colocalization of GWAS and eQTL Signals Detects Target Genes. Am J Hum Genet. 2016; 99: 1245-1260.

48. Kamburov A, Wierling C, Lehrach H, Herwig R. ConsensusPathDB--a database for integrating human functional interaction networks. Nucleic Acids Res. 2009; 37: D623-628.

49. Salojin KV, Zhang J, Meagher C, Delovitch TL. ZAP-70 is essential for the T cell antigen receptor-induced plasma membrane targeting of SOS and Vav in T cells. J Biol Chem. 2000; 275: 5966-5975.

50. Calender A, Weichhart T, Valeyre D, Pacheco Y. Current Insights in Genetics of Sarcoidosis: Functional and Clinical Impacts. J Clin Med. 2020; 9: 2633.

51. Rivera NV, Ronninger M, Shchetynsky K, Franke A, Nothen MM, 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: 1008-1022.

52. Bauer Y, Tedrow J, de Bernard S, Birker-Robaczewska M, Gibson KF, Guardela BJ, et al. A novel genomic signature with translational significance for human idiopathic pulmonary fibrosis. Am J Respir Cell Mol Biol. 2015; 52: 217-231.

53. Zhou T, Zhang W, Sweiss NJ, Chen ES, Moller DR, Knox KS, et al. Peripheral blood gene expression as a novel genomic biomarker in complicated sarcoidosis. PloS one. United States; 2012. p. e44818.

54. Zhang L, Zhang T, Shang B, Li Y, Cao Z, Wang H. Prognostic effect of coexisting TP53 and ZFHX3 mutations in non-small cell lung cancer patients treated with immune checkpoint inhibitors. Scand J Immunol. 2021; 94: e13087.

55. Zhu S, Hu X, Bennett S, Mai Y, Xu J. Molecular Structure, Expression and Role of TAFA4 and its Receptor FPR1 in the Spinal Cord. Front Cell Dev Biol. 2022; 10: 911414.

56. Wang Q, Chen Y, Feng H, Zhang B, Wang H. Prognostic and predictive value of HURP in non-small cell lung cancer. Oncol Rep. 2018; 39: 1682-1692.

57. Mortaz E, Sereshki HA, Abedini A, Kiani A, Mirsaeidi M, Soroush D, et al. Association of serum TNF-alpha, IL-8 and free light chain with HLA-DR B alleles expression in pulmonary and extra-pulmonary sarcoidosis. J Inflamm (Lond). 2015; 12: 21.

58. Wolin A, Lahtela EL, Anttila V, Petrek M, Grunewald J, van Moorsel CHM, et al. SNP Variants in Major Histocompatibility Complex Are Associated with Sarcoidosis Susceptibility-A Joint Analysis in Four European Populations. Front Immunol. 2017; 8: 422.

59. Ishihara M, Ishida T, Inoko H, Ando H, Naruse T, Nose Y, et al. HLA serological and class II genotyping in sarcoidosis patients in Japan. Jpn J Ophthalmol. 1996; 40: 86-94.

60. Liao S, Jacobson S, Hamzeh N, Culver D, Barkes B, Mroz P, et al. Genome-wide association study identifies multiple HLA loci for sarcoidosis susceptibility. medRxiv. 2022: 2022.2010.2013.22281070.

61. Garman L, Pezant N, Pastori A, Savoy KA, Li C, Levin AM, et al. Genome-Wide Association Study of Ocular Sarcoidosis Confirms HLA Associations and Implicates Barrier Function and Autoimmunity in African Americans. Ocul Immunol Inflamm. 2021; 29: 244-249.

62. Raymond T, Schaller M, Hogaboam CM, Lukacs NW, Rochford R, Kunkel SL. Toll-like receptors, Notch ligands, and cytokines drive the chronicity of lung inflammation. Proc Am Thorac Soc. 2007; 4: 635-641.

63. Harb H, Benamar M, Lai PS, Contini P, Griffith JW, Crestani E, et al. Notch4 signaling limits regulatory T-cell-mediated tissue repair and promotes severe lung inflammation in viral infections. Immunity. 2021; 54: 1186-1199.e1187.

64. Chaperon M, Pacheco Y, Maucort-Boulch D, Iwaz J, Perard L, Broussolle C, et al. BTNL2 gene polymorphism and sarcoid uveitis. Br J Ophthalmol. 2019; 103: 1690-1694.

65. Javadi A, Zarei M, Pourabdollah M, Sadr M, Kiani A, et al. DNA sequencing validation by PCR-RFLP for evaluating butyrophilin-like 2 rs2076530 polymorphism in Iranian patients with sarcoidosis. Sarcoidosis Vasc Diffuse Lung Dis. 2022; 38: e2021050.

66. Lin Y, Wei J, Fan L, Cheng D. BTNL2 gene polymorphism and sarcoidosis susceptibility: a meta-analysis. PloS one. 2015; 10: e0122639.

67. Morais A, Lima B, Alves H, Melo N, Mota PC, Marques A, et al. Associations between sarcoidosis clinical course and ANXA11 rs1049550 C/T, BTNL2 rs2076530 G/A, and HLA class I and II alleles. Clin Respir J. 2018; 12: 532-537.

68. Pacheco Y, Calender A, Israël-Biet D, Roy P, Lebecque S, Cottin V, et al. Familial vs. sporadic sarcoidosis: BTNL2 polymorphisms, clinical presentations, and outcomes in a French cohort. Orphanet J Rare Dis. 2016; 11: 165.

69. Tong X, Ma Y, Niu X, Yan Z, Liu S, Peng B, et al. The BTNL2 G16071A gene polymorphism increases granulomatous disease susceptibility: A meta-analysis including FPRP test of 8710 participants. Medicine (Baltimore). 2016; 95: e4325.

70. Wang A, Singh K, Ibrahim W, King B, Damsky W. The Promise of JAK Inhibitors for Treatment of Sarcoidosis and Other Inflammatory Disorders with Macrophage Activation: A Review of the Literature. Yale J Biol Med. 2020; 93: 187-195.

71. Hao X, Du Y, Qian L, Li D, Liu X. Upregulation of long noncoding RNA AP003419.16 predicts high risk of aging-associated idiopathic pulmonary fibrosis. Mol Med Rep. 2017; 16: 8085-8091.

72. Li R, Pu X, Chang JY, Ye Y, Komaki R, Minna JD, et al. MiRNA-Related Genetic Variations Associated with Radiotherapy-Induced Toxicities in Patients with Locally Advanced Non-Small Cell Lung Cancer. PloS one. 2016; 11: e0150467.

73. Marchese S, Cancelmo L, Diab O, Cahn L, Aaronson C, Daskalakis NP, et al. Altered gene expression and PTSD symptom dimensions in World Trade Center responders. Mol Psychiatry. 2022; 27: 2225-2246.