



# Multiple Correspondence Analysis and HLA-Associations of Organ Involvement in a Large Cohort of African-American and European-American Patients with Sarcoidosis

Astrid Rasmussen<sup>1</sup> · Bryan A. Dawkins<sup>1</sup> · Chuang Li<sup>1</sup> · Nathan Pezant<sup>1</sup> · Albert M. Levin<sup>2</sup> · Benjamin A. Rybicki<sup>2</sup> · Michael C. Iannuzzi<sup>3</sup> · Courtney G. Montgomery<sup>1</sup>

Received: 2 May 2023 / Accepted: 2 June 2023 / Published online: 15 June 2023  
© The Author(s) 2023

## Abstract

Sarcoidosis is a systemic granulomatous disease with predominant pulmonary involvement and vast heterogeneity of clinical manifestations and disease outcomes. African American (AA) patients suffer greater morbidity and mortality. Using Multiple Correspondence Analysis, we identified seven clusters of organ involvement in European American (EA; n = 385) patients which were similar to those previously described in a Pan-European (GenPhenReSa) and a Spanish cohort (SARCOGEAS). In contrast, AA (n = 987) had six, less well-defined and overlapping clusters with little similarity to the cluster identified in the EA cohort evaluated at the same U.S. institutions. Association of cluster membership with two-digit HLA-DRB1 alleles demonstrated ancestry-specific patterns of association and replicated known HLA effects.

These results further support the notion that genetically influenced immune risk profiles, which differ based on ancestry, play a role in phenotypic heterogeneity. Dissecting such risk profiles will move us closer to personalized medicine for this complex disease.

**Keywords** Sarcoidosis · Precision Medicine · Multiple Correspondence Analysis · African-American ancestry

## Introduction

The pathologic hallmark of sarcoidosis is granulomatous inflammation primarily of the lungs and intrathoracic lymph nodes but which can affect a wide variety of organs. A puzzling aspect of the disease is the heterogeneity that characterizes it. Many patients eventually clinically resolve, sometimes spontaneously and without treatment. However, up to one-third of the cases will have a protracted course with significant morbidity and health care utilization, and ~ 10% will progress to irreversible fibrotic disease [1]. The number

and pattern of organs involved can vary greatly and some subphenotypes impact quality of life more than others (i.e., neurosarcoidosis and cardiac sarcoidosis have greater mortality than cutaneous sarcoidosis) [2, 3]. It is unclear what determines this variability, but there is epidemiologic evidence that a combination of genetic variants and the influence of environmental factors contribute to phenotypic variability. Socioeconomic factors and access to care have also been associated with divergent outcomes [4–7].

Attempts have been made at identifying predictors of worse prognosis, poor quality of life, and increased need for aggressive treatment [6]. Schupp et al. performed multidimensional correspondence analysis (MCA) and clustering of a cohort of > 2000 Caucasian sarcoidosis patients and identified five distinct subgroups based on patterns of organ involvement. The SARCOGEAS study group, analyzing data from 1230 Spanish patients replicated these organ associations [8]. However, the patterns of organ involvement have not been analyzed in populations of non-European ancestry. Our aim was to assess a large cohort of African-Americans (AA) and European-Americans (EA) recruited in the United States to characterize the patterns of organ clustering,

✉ Courtney G. Montgomery  
Courtney-Montgomery@omrf.org

<sup>1</sup> Genes and Human Disease Program, Oklahoma Medical Research Foundation, 825 NE 13th, Research Tower, Suite 2202, Oklahoma City, Ok 73104, USA

<sup>2</sup> Department of Public Health Sciences, Henry Ford Health System, Detroit, MI, USA

<sup>3</sup> Department of Medical Education, City University of New York School of Medicine, New York, NY, USA

compare them with the reported European findings, and assess association with HLA-DRB1, the most confirmed genetic association with sarcoidosis.

## Methods

### Participants

This study was conducted with data from 2072 sarcoidosis cases (1539 AA and 571 EA) originally recruited into three past cohorts, specifically “A Case Control Etiologic Study of Sarcoidosis (ACCESS) Group”[9], “Sarcoidosis Genetic Analysis (SAGA) study”[10], and “Henry Ford Family Study (HFFS)” [11]. (Table 1). Patients in these cohorts were primarily ascertained in outpatient visits to pulmonary clinics and organ involvement was assessed at a single timepoint.

Organ involvement was assessed based on either the ACCESS or the WASOG Organ Assessment Tools [9, 12]; only subjects with confirmed sarcoidosis and data documenting a minimum of two organs affected were considered useful for analysis. Furthermore, adjudication of involvement of at least 50% of the organs of the WASOG Assessment Tool was required for a subject to be included. Thus, the final dataset consisted of 987 AA and 385 EA cases. Two-digit DRB1 HLA genotype data for the complete cohort were available from previous studies [13].

### Clustering

Multiple correspondence analysis (MCA) of the organ involvement, an analog to principal components analysis (PCA) for categorical data, was used for dimension reduction prior to clustering using the FactoMineR package in [14]. Hierarchical clustering with Ward’s linkage method on the first 5 MCA dimensions was used to derive clusters. Cluster stability was assessed using bootstrap resampling ( $n = 1000$  samples), which was implemented with the fpc package in R [15]. The total number of clusters was chosen using the average Jaccard similarity metric across bootstrap samples in order to maximize the total number of stable clusters, defined heuristically as those with bootstrap average Jaccard similarity exceeding 0.7.

### Statistical Comparisons

A hypergeometric test was used for assessing the pairwise association between organs and clusters, where the association was considered statistically significant if its hypergeometric  $p$ -value was less than 0.05. An organ was considered “enriched” within a cluster if its within-cluster frequency exceeded its background frequency; “depletion”

was defined similarly, only the relationship between the organ’s within-cluster frequency and its background frequency was reversed. The standard normal quantile corresponding to the hypergeometric  $p$ -value, scaled positively (enrichment) or negatively (depletion) was used to visualize the association between organs and clusters. Association of MCA cluster membership with HLA-DRB1 alleles (additive model) was determined using logistic regression with inclusion in a given cluster versus not as the outcome variable. The HLA allele was the independent variable, with sex and the first 4 principal components for ancestry as covariates. Statistical significance was defined as  $p < 0.05$ .

## Results

Subjects included in the clustering analyses were a subset of the complete cohorts but are demographically comparable to the subjects excluded due to missing data (Table 1). The clusters identified are named by the primary organ involved (Fig. 1a, b).

### African-American Sarcoidosis Cases

AA patients were predominantly female (~73%) with a mean age at diagnosis of 48.5 years (SD 8.1; median 49). The most frequently affected organs were the lungs (98%) followed by skin (41%), eyes (28%), and extrathoracic lymph nodes (21%).

Six distinct multi-organ clusters were identified in AA (Fig. 1a), excluding a pulmonary-only subset of subjects not included in MCA ( $n = 96$ ). The largest cluster was the calcium cluster ( $n = 396$ ) followed by the skin cluster ( $n = 277$ ). Intermediate in frequency were the neuro/ocular/cardiac/salivary gland and the abdominal organs/extrathoracic lymph nodes clusters ( $n = 97$  and 81, respectively). Two additional small clusters, with a moderate overlap of the organs involved, were the muscle/cardiac/bone marrow ( $n = 21$ ) and renal/cardiac/calcium ( $n = 19$ ) clusters (Fig. 2a). The proportional contribution of each organ to the cluster is shown in Fig. 2a. Age and sex did not differ between the clusters.

Statistically significant association with common HLA-DRB1 alleles was found for all of the clusters except the calcium/vitamin D cluster (2) (Supplemental Table 1). Interestingly, the association of the cardiovascular/muscle/calcium cluster (6) to DRB1\*01:02 is consistent with a previous investigation of the larger cohort from which this sample was derived showing association of this allele with persistent disease in AA [16].

**Table 1** African American and European American Sarcoidosis Cohorts

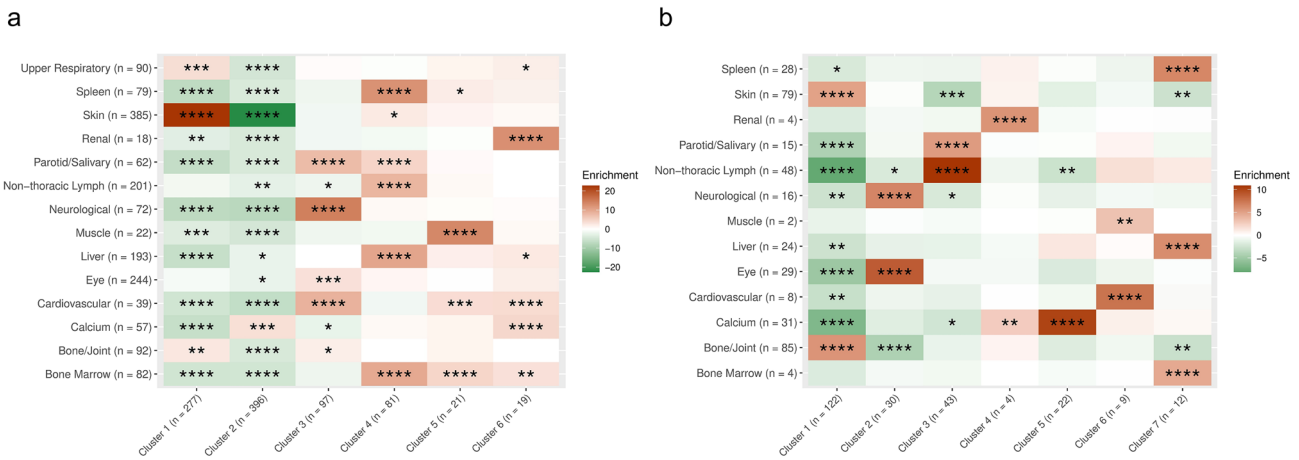
|  | African American            |                          |                      | European American       |                          |                      |
|--|-----------------------------|--------------------------|----------------------|-------------------------|--------------------------|----------------------|
|  | Complete Cohort<br>n = 1593 | Clustered Subset n = 987 | Missing data n (%)   | Complete cohort n = 517 | Clustered Subset n = 385 | Missing data n (%)   |
|  | Organ involved n (%)        | Organ involved n (%)     | Organ involved n (%) | Organ involved n (%)    | Organ involved n (%)     | Organ involved n (%) |
| Sex: female                                      | 1128 (73.3)                 | 716 (72.5)               |                      | 301 (58.2)              | 221 (57.4)               |                      |
| Age at diagnosis (mean ± SD)                     | 48.5 ± 8.1                  | 48.2 ± 7.6               |                      | n/a                     | n/a                      |                      |
| Organ Involvement (ACCESS or WASOG) <sup>1</sup> | Organ involved n (%)        | Organ involved n (%)     | Missing data n (%)   | Organ involved n(%)     | Organ involved n(%)      | Missing data n(%)    |
| Pulmonary  | 1260 (98)                   | 965 (97.8)               | 22 (2.2)             | n/a                     | 369 (95.8)               | 16 (4.2)             |
| Neurological                                     | 104 (10.5)                  | 101 (10.2)               | 12 (1.2)             | n/a                     | 34 (8.8)                 | 0 (0)                |
| Ocular   | 288 (28.9)                  | 274 (27.8)               | 13 (1.3)             | n/a                     | 43 (11.2)                | 0 (0)                |
| Cardiac  | 40 (4.0)                    | 40 (4.1)                 | 4 (0.4)              | n/a                     | 11 (2.9)                 | 0 (0)                |
| Liver  | 207 (20.7)                  | 197 (20.0)               | 5 (0.5)              | n/a                     | 24 (6.2)                 | 0 (0)                |
| Spleen   | 82 (8.2)                    | 79 (8.0)                 | 6 (0.6)              | n/a                     | 28 (7.3)                 | 0 (0)                |
| Calcium/Vitamin D                                | 59 (6.7)                    | 57 (5.8)                 | 123 (12.5)           | n/a                     | 31 (8.1)                 | 0 (0)                |
| Renal  | 21 (2.1)                    | 20 (2.0)                 | 3 (0.3)              | n/a                     | 5 (1.3)                  | 0 (0)                |
| Skin   | 428 (42.1)                  | 401 (40.6)               | 4 (0.4)              | n/a                     | 84 (21.8)                | 0 (0)                |
| Salivary Glands                                  | 78 (8.8)                    | 72 (7.3)                 | 115 (11.7)           | n/a                     | 28 (7.3)                 | 0 (0)                |
| Muscle   | 32 (3.6)                    | 32 (3.2)                 | 113 (11.4)           | n/a                     | 13 (3.4)                 | 0 (0)                |
| Bone Marrow                                      | 94 (9.4)                    | 92 (9.3)                 | 4 (0.4)              | n/a                     | 5 (1.3)                  | 0 (0)                |
| Extra-Thoracic LN                                | 214 (21.3)                  | 202 (20.5)               | 6 (0.6)              | n/a                     | 49 (12.7)                | 0 (0)                |
| Upper Respiratory Tract                          | 97 (16.8)                   | 90 (9.1)                 | 428 (43.4)           | n/a                     | 0 (0.0)                  | 385 (100)            |
| Bone and Joint                                   | 96 (12.1)                   | 92 (9.3)                 | 201 (20.4)           | n/a                     | 85 (22.1)                | 0 (0)                |

Comparison of each complete sub-cohort by race to the subset with enough data to be clustered

<sup>1</sup>Based on the ACCESS or WASOG Organ Involvement Instruments [9, 12]

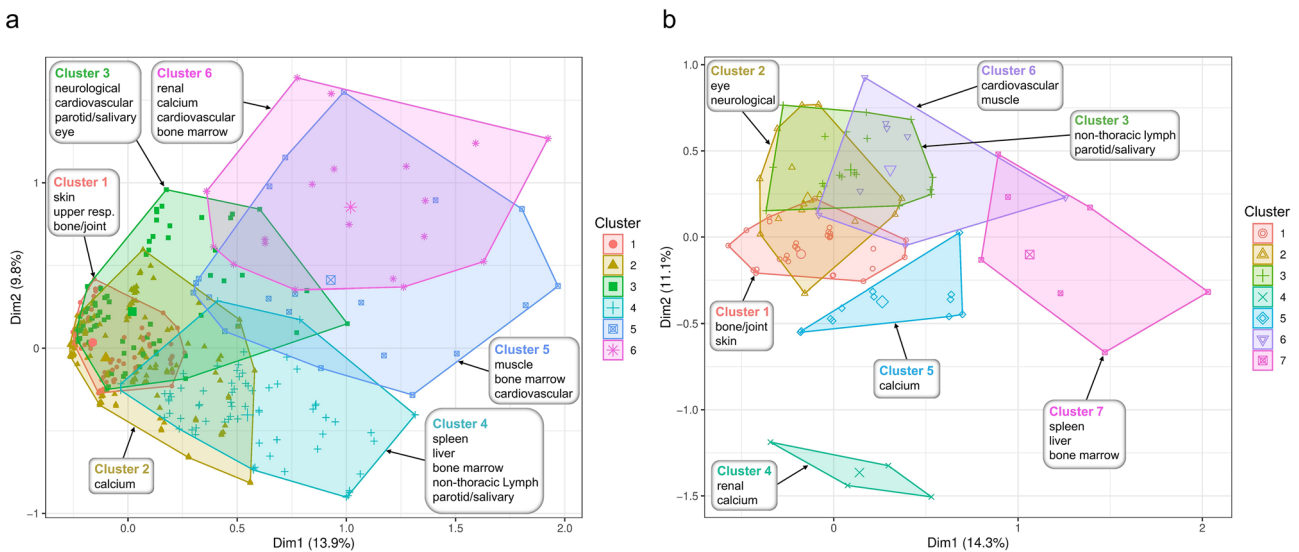
<sup>2</sup>For the subsets of patients included in the clustering analysis, the number and percentage of patients with each organ involved is shown followed by the number of subjects (%) for whom the corresponding data are missing

n/a: Data not available



**Fig. 1** Organ enrichment by cluster in African Americans and European Americans. Panel **a** shows the patterns of organ enrichment in African Americans and Panel **b** shows the patterns of organ enrichment in European Americans. Rows correspond to the involvement of each organ; columns correspond to clusters. The color scale is derived from standard normal quantiles corresponding to hypergeometric p-values for the pairwise association between organ and cluster.

ter. The relationship between an organ’s within-cluster frequency and its background frequency determines the sign or direction of each quantile; positive quantiles denote greater within-cluster frequency (enriched, shown in shades of brown) and negative quantiles denote lesser within-cluster frequency (depleted, shown in shades of green). Asterisks represent hypergeometric p-value magnitude (\*\*\*\*  $P < 0.0001$ , \*\*\*  $P < 0.001$ , \*\*  $P < 0.01$ , \*  $P < 0.05$ )



**Fig. 2** Clustering of systemic sarcoidosis phenotypes in African Americans and European Americans. Each point represents a sarcoidosis-affected individual’s contribution to the first two principal dimensions (Dim1 and Dim2), calculated with Multiple Correspondence Analysis (MCA) on organ involvement data. Enriched (or over-

represented) affected organs are used for annotating each cluster. Panel **a** shows the clustering of organ involvement in African-Americans and Panel **b** shows the clustering of organ involvement in European-Americans

**European-American Sarcoidosis Cases**

Females were the majority of EA cases (~58%) and EA mainly presented with pulmonary (95.8%), cutaneous (21.8%), bone/joint (22.1%), and extrathoracic lymph node (21.8%) involvement. Seven multi-organ clusters, excluding a pulmonary-only subset of subjects not included in MCA

( $n = 143$ ), were identified (Fig. 1b). These are different from those affecting AA subjects and resemble previous descriptions of organ involvement patterns in European populations [8, 17].

The largest cluster identified was bone/joint/skin ( $n = 122$ ), similar to the musculoskeletal-cutaneous [17] and musculoskeletal and skin clusters [8] previously reported.

The extrathoracic lymph nodes/salivary glands cluster ( $n=43$ ), partially replicates the adenopathic anatomic cluster [8], while the neurological/ocular ( $n=30$ ) cluster mimics the ocular-cardiac-cutaneous-central nervous system [17] and the anatomic neuro-ocular clusters [8]. A fourth cluster was calcium ( $n=22$ ) with no direct equivalent in previous studies. Smaller clusters were spleen/liver/bone marrow, similar to Schupp et al.'s abdominal organs cluster [17] and the anatomic hepatosplenic cluster [8]. Very small clusters (heart-muscle and kidney/calcium; Fig. 2b) would fit within the extrapulmonary involvement cluster [17].

All clusters except the calcium cluster had significant association with HLA-DRB1 alleles (Supplemental Table 1). Notably, we observed association of DRB1\*03:01, which has previously been associated with resolving disease [18], with involvement of organs typically associated with disease resolution such as the bone/joint/skin cluster (OR = 2.79; 95% CI 1.73–4.55;  $p=2.92E-05$ ) while being protective for inclusion in the extrathoracic lymph node cluster (OR = 0.27; 95% CI 0.06–0.73;  $p=0.028$ ). Additional relevant associations were with various DRB1\*04 alleles. We observed association of \*04:01 with the neurological/ocular cluster (cluster 2; OR = 6.48; 95% CI 2.72–15.32;  $p=1.97E-05$ ); this allele has previously been associated with ocular involvement [19, 20]. DRB1\*04 has also been associated with hypercalcemia [21] and we show association of the calcium/renal cluster with DRB1\*04:05 (OR = 36.75; 95% CI 1.39–618.85;  $p=0.011$ ).

## Discussion

Sarcoidosis can be a benign disease but for many patients, it results in a debilitating condition that negatively impacts their quantity and quality of life. Understanding the factors that influence the patterns of organ involvement and overall heterogeneity of the disease is important for opportune diagnosis, early management of potentially serious complications, and future targeted therapies. We and others have analyzed the aggregation of disease manifestations and identified patterns occurring in specific populations. We present novel disease clusters for American sarcoidosis patients of African descent. Unfortunately, the AA organ clusters were less defined and overlapped significantly in the first two principal dimensions; this complexity required analyses in 5 dimensions to attain clear separation.

In contrast, the EA clusters were tighter and while they partially overlapped, three studies have replicated similar patterns supporting their value in the appropriate populations [8, 17].

We acknowledge the limitations of our study which include, more prominently for the EA cohort, a large proportion of missing data, and the use of mathematical modeling

for clinical medicine. The goal of methods such as MCA and hierarchical clustering, is to reduce the noise present in data with a large number of variables while retaining the signal of interest and clarifying underlying patterns. As such, they serve to outline steps towards testing hypotheses and validating relationships by other means. [22, 23]

While genome-wide genotyping data and admixture proportions are available for the patients in the cohort described herein, we did not find any correlation between levels of admixture (i.e., the relative contribution of African and non-African ancestry) and specific phenotypic clusters. Each organ cluster is small and thus underpowered to conduct a genome-wide association scan, but additional candidate gene studies, similar to what we present for loci within HLA could give insight into ancestry-specific genetic influences on phenotypic heterogeneity. The overall non-African ancestry in this AA cohort is ~17% which is slightly less than other reports of black Americans [24], and thus, our findings may not be applicable to black populations with different origins or admixture patterns.

Nevertheless, the clear phenotypic differences we observed between the AA and EA subcohorts and our observation of HLA-DRB1 associations to specific patterns of organ involvement support ancestry-specific differences likely resulting from distinct social, cultural, and environmental factors in the setting of unique genetic determinants. Future research should be centered around the multi-omic deconstruction of sarcoidosis in people of all racial backgrounds.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00408-023-00626-6>.

**Author Contributions** All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by CM, AR, BD, NP, CL, AL, BR, and MI. The first draft of the manuscript was written by AR and CM and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

**Funding** This work was supported by the National Institutes of Health [Grant Nos R01-HL113326, U54-GM104938, T32-AI07633].

## Declarations

**Competing interests** The authors have no relevant financial or non-financial interests to disclose.

**Ethical Approval** Institutional Review Board (IRB) approval was obtained prior to sample collection or data acquisition for each of the participating studies, namely “A Case Control Etiologic Study of Sarcoidosis (ACCESS) Group”, the “Sarcoidosis Genetic Analysis (SAGA) study”, and the “Henry Ford Family Study (HFFS)”.

**Informed Consent** Informed consent was obtained and enrollment was HIPAA-compliant and adhered to the tenets of the Declaration of Helsinki. The data was shared with the Oklahoma Medical Research Foundation (OMRF) Sarcoidosis Genetic Studies (SGS) devoid of identifiers



or Private Health Information with the approval and oversight of the OMRF IRB.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Grunewald J, Grutters JC, Arkema EV, Saketkoo LA, Moller DR, Muller-Quernheim J (2019) Sarcoidosis. *Nat Rev Dis Primers* 5:45. <https://doi.org/10.1038/s41572-019-0096-x>
- Kirkil G, Lower EE, Baughman RP (2018) Predictors of mortality in pulmonary sarcoidosis. *Chest* 153:105–113. <https://doi.org/10.1016/j.chest.2017.07.008>
- Drent M, Crouser ED, Grunewald J (2021) Challenges of sarcoidosis and its management. *N Engl J Med* 385:1018–1032. <https://doi.org/10.1056/NEJMra2101555>
- Rice JB, White A, Lopez A, Nelson WW (2017) High-cost sarcoidosis patients in the United States: patient characteristics and patterns of health care resource utilization. *J Manag Care Spec Pharm* 23:1261–1269. <https://doi.org/10.18553/jmcp.2017.17203>
- Ungprasert P, Crowson CS, Matteson EL (2017) Influence of gender on epidemiology and clinical manifestations of sarcoidosis: a population-based retrospective cohort study 1976–2013. *Lung* 195:87–91. <https://doi.org/10.1007/s00408-016-9952-6>
- Brito-Zeron P, Acar-Denizli N, Siso-Almirall A, Bosch X, Hernandez F, Vilanova S, Villalta M, Kostov B, Paradela M, Sanchez M et al (2018) The burden of comorbidity and complexity in sarcoidosis: impact of associated chronic diseases. *Lung* 196:239–248. <https://doi.org/10.1007/s00408-017-0076-4>
- Ronsmans S, De Ridder J, Vandebroek E, Keirsbilck S, Nemery B, Hoet PHM, Vanderschueren S, Wuyts WA, Yserbyt J (2021) Associations between occupational and environmental exposures and organ involvement in sarcoidosis: a retrospective case-case analysis. *Respir Res* 22:224. <https://doi.org/10.1186/s12931-021-01818-5>
- Perez-Alvarez R, Brito-Zeron P, Kostov B, Feijoo-Masso C, Fraile G, Gomez-de-la-Torre R, De-Escalante B, Lopez-Dupla M, Alguacil A, Chara-Cervantes J et al (2019) Systemic phenotype of sarcoidosis associated with radiological stages. Analysis of 1230 patients. *Eur J Intern Med* 69:77–85. <https://doi.org/10.1016/j.ejim.2019.08.025>
- Rybicki BA, Iannuzzi MC, Frederick MM, Thompson BW, Rossman MD, Bresnitz EA, Terrin ML, Moller DR, Barnard J, Baughman RP et al (2001) Familial aggregation of sarcoidosis. A case-control etiologic study of sarcoidosis (ACCESS). *Am J Respir Crit Care Med* 164:2085–2091. <https://doi.org/10.1164/ajrccm.164.11.2106001>
- Rybicki BA, Sinha R, Iyengar S, Gray-McGuire C, Elston RC, Iannuzzi MC, Consortium, S.S. (2007) Genetic linkage analysis of sarcoidosis phenotypes: the sarcoidosis genetic analysis (SAGA) study. *Genes Immun* 8:379–386. <https://doi.org/10.1038/sj.gene.6364396>
- Adrianto I, Lin CP, Hale JJ, Levin AM, Datta I, Parker R, Adler A, Kelly JA, Kaufman KM, Lessard CJ et al (2012) Genome-wide association study of African and European Americans implicates multiple shared and ethnic specific loci in sarcoidosis susceptibility. *PLoS one* 7:e43907. <https://doi.org/10.1371/journal.pone.0043907>
- Judson MA, Costabel U, Drent M, Wells A, Maier L, Koth L, Shigemitsu H, Culver DA, Gelfand J, Valeyre D et al (2014) The WASOG sarcoidosis organ assessment instrument: an update of a previous clinical tool. *Sarcoidosis Vasc Diffuse Lung Dis* 31:19–27
- Levin AM, Adrianto I, Datta I, Iannuzzi MC, Trudeau S, McKie P, Montgomery CG, Rybicki BA (2014) Performance of HLA allele prediction methods in African Americans for class II genes HLA-DRB1, -DQB1, and -DPB1. *BMC Genet* 15:72. <https://doi.org/10.1186/1471-2156-15-72>
- Lê S, Julie J, François Husson (2008) FactoMineR: an R package for multivariate analysis. *J Stat Softw* 25:1–18
- Hennig, C. fpc: Flexible Procedures for Clustering. R package version 2.2–8. <https://CRAN.R-project.org/package=fpc> 2020
- Levin AM, Adrianto I, Datta I, Iannuzzi MC, Trudeau S, Li J, Drake WP, Montgomery CG, Rybicki BA (2015) Association of HLA-DRB1 with sarcoidosis susceptibility and progression in African Americans. *Am J Respir Cell Mol Biol* 53:206–216. <https://doi.org/10.1165/rcmb.2014-0227OC>
- Schupp JC, Freitag-Wolf S, Bargagli E, Mihailovic-Vucinic V, Rottoli P, Grubanovic A, Muller A, Jochens A, Tittmann L, Schnerch J et al (2018) Phenotypes of organ involvement in sarcoidosis. *Eur Respir J*. <https://doi.org/10.1183/13993003.00991-2017>
- Grunewald J, Brynedal B, Darlington P, Nisell M, Cederlund K, Hillert J, Eklund A (2010) Different HLA-DRB1 allele distributions in distinct clinical subgroups of sarcoidosis patients. *Respir Res* 11:25. <https://doi.org/10.1186/1465-9921-11-25>
- Garman L, Pezant N, Pastori A, Savoy KA, Li C, Levin AM, Iannuzzi MC, Rybicki BA, Adrianto I, Montgomery CG (2021) Genome-wide Association study of ocular sarcoidosis confirms HLA associations and implicates barrier function and autoimmunity in African Americans. *Ocul Immunol Inflamm* 29:244–249. <https://doi.org/10.1080/09273948.2019.1705985>
- Sato H, Woodhead FA, Ahmad T, Grutters JC, Spagnolo P, van den Bosch JM, Maier LA, Newman LS, Nagai S, Izumi T et al (2010) Sarcoidosis HLA class II genotyping distinguishes differences of clinical phenotype across ethnic groups. *Hum Mol Genet* 19:4100–4111. <https://doi.org/10.1093/hmg/ddq325>
- Werner J, Rivera N, Grunewald J, Eklund A, Iseda T, Darlington P, Kullberg S (2021) HLA-DRB1 alleles associate with hypercalcemia in sarcoidosis. *Respir Med* 187:106537. <https://doi.org/10.1016/j.rmed.2021.106537>
- Berisha V, Krantsevich C, Hahn PR, Hahn S, Dasarathy G, Turaga P, Liss J (2021) Digital medicine and the curse of dimensionality. *NPJ Digit Med* 4:153. <https://doi.org/10.1038/s41746-021-00521-5>
- Nguyen LH, Holmes S (2019) Ten quick tips for effective dimensionality reduction. *PLoS Comput Biol* 15:e1006907. <https://doi.org/10.1371/journal.pcbi.1006907>
- Bryc K, Durand EY, Macpherson JM, Reich D, Mountain JL (2015) The genetic ancestry of African Americans, Latinos, and European Americans across the United States. *Am J Hum Genet* 96:37–53. <https://doi.org/10.1016/j.ajhg.2014.11.010>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.