Published in final edited form as:

Clin Chest Med. 2015 December ; 36(4): 621–630. doi:10.1016/j.ccm.2015.08.005.

Identifying novel biomarkers in sarcoidosis using genome-based approaches

Nancy Casanova[#], Tong Zhou[#], Kenneth S. Knox, and Joe G.N. Garcia^{*}

University of Arizona Health Sciences Center, Arizona Respiratory Center and Department of Medicine, University of Arizona, Tucson, AZ 85721, USA

[#] These authors contributed equally to this work.

Synopsis

We briefly review conventional biomarkers used clinically to 1) support a diagnosis and 2) monitor disease progression in patients with sarcoidosis. We describe potential new biomarkers identified by genome-wide screening and the approaches to discover these biomarkers.

Keywords

sarcoidosis; biomarkers; molecular signature; GWAS; microarray; sequencing

Sarcoidosis: Overview and need for biomarkers

Sarcoidosis is a systemic heterogeneous inflammatory disease characterized by the presence of non-caseating epithelioid granulomas in one or multiple organs with the lung affected in ~ 90% of cases. Lung involvement is commonly manifested as bilateral hilar lymphadenopathy (BHL) and pulmonary infiltration with more severe cases developing pulmonary fibrosis. Ocular and skin lesions may become sight threatening or disfiguring requiring aggressive treatment. Cardiac and neurologic involvement may cause morbidity and death [1, 2]. Löfgren syndrome, an acute presentation of sarcoidosis, is characterized by erythema nodosum, BHL, and polyarthralgia and is associated with spontaneous regression [3, 4]. Additionally, more than 50% of sarcoidosis patients will experience remission within 3 years after diagnosis, with over 66% of patients experiencing remission within 10 years [5,

^{*} **Corresponding author:** Joe G.N. Garcia, MD, Senior Vice President for Health Sciences, Arizona Health Sciences Center and Department of Medicine, University of Arizona; Drachman Hall, Room B-207, 1295 North Martin Avenue, Tucson, AZ 85721; Tel: 1-520-626-1197; Fax: 1-520-626-1460; skipgarcia@email.arizona.edu.

Nancy Casanova, MD, Arizona Health Sciences Center and Department of Medicine, University of Arizona; Medical Research Building, Room 412, 1656 East Mabel Street, Tucson, AZ 85721; Tel: 1-520- 626-4110; ncasanova@email.arizona.edu

Tong Zhou, PhD, Arizona Health Sciences Center and Department of Medicine, University of Arizona; Medical Research Building, Room 410, 1656 East Mabel Street, Tucson, AZ 85721; Tel: 1-520- 626-8433; tongzhou@email.arizona.edu

Kenneth S. Knox MD, Arizona Health Sciences Center and Department of Medicine, University of Arizona; Pulmonary Division, 1501 N Campbell, Tucson, AZ 85721; Tel: 1-520-626-6114; kknox@deptofmed.arizona.edu

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Disclosures: The authors have nothing to disclose.

6]. It is clear that while certain disease phenotypes and chest radiograph stages portend a good prognosis, a large proportion of patients would benefit from technological advances in biomarker development.

The need to develop useful biomarkers in the diagnosis and prognosis of subjects with sarcoidosis has long been recognized as sarcoidosis is a diagnosis of exclusion and may mimic multiple other rheumatologic illnesses [7, 8]. Furthermore, a significant percentage of patients with sarcoidosis (~ a third) develop complications with granulomatous involvement of vital organs with progressive disease. Thus, monitoring subclinical disease activity and likelihood of progression or remission in a longitudinal fashion remains a challenge. Finally, significant racial and gender differences in disease development and prognosis have been reported in African Americans, [9], Irish, Scandinavian and Hispanic populations. Again, these differences represent a compelling need for racial and ethnic-selective biomarkers to assess disease progression and response to therapy in diverse populations. In this Chapter, we review traditional biomarkers in sarcoidosis as well as biomarkers emerging from technology driven strategies. We discuss the integration of genome-wide gene expression signature strategies and genetic variation data as additional opportunities to generate useful biomarkers in sarcoidosis, particularly in assessment of personal risk for developing complicated sarcoidosis.

Traditional biomarkers in sarcoidosis

Despite limitations, the use of biomarkers to support diagnosis and predict disease activity remains a focal point in routine clinical care. Multiple methodologies have been applied to detect biomarkers in serum, lung tissue, bronchoalveolar lavage fluid (BALF) and exhaled breath condensate (EBC) using enzyme-linked immunosorbent assays (ELISA), proteomic analysis and mass spectrometry [10]. Traditionally, clinical biomarkers measured in sarcoidosis were limited to soluble factors measured in blood, BAL or cerebral spinal fluid. Data remains inconsistent regarding the validity of EBC biomarkers [11]. As technology improves, biomarker “panels” generated from array data will continue to emerge.

Soluble biomarkers associated with monocyte-macrophage activation

Although specific biomarkers have been suggested as useful tools in sarcoidosis, reduced specificity has been a major limitation with **angiotensin-converting enzyme (ACE)**, the most commonly used biomarker in sarcoidosis. ACE is derived from activated macrophages in granulomatous pulmonary remodeling [12], and is useful in supporting a diagnosis and monitoring disease activity in some patients [12, 13]. ACE levels are increased in approximately two-thirds of patients with sarcoidosis with elevated levels reported in neurosarcoidosis [14]. However, elevated ACE levels are not specific for sarcoidosis, and are observed in granulomatous diseases such as tuberculosis, fungal infections, and Gaucher disease [12, 15, 16]. Moreover, there is no evidence that ACE levels significantly differ between active and inactive sarcoidosis [13] and do not reliably correlate with the severity of disease [17]. Serum concentrations at diagnosis were noted to be significantly lower in acute sarcoidosis and in Löfgren's syndrome [18]. In addition, ACE concentration and activity is influenced by genetic polymorphisms with enzymatic activity significantly higher in

individuals with the DD genotype than in individuals with the II genotype [16, 19, 20]. The utility of ACE as a biomarker in diagnosis of sarcoidosis remains limited, although future studies utilizing conformational fingerprinting of ACE may yield better specificity [73]. While ACE is the prototypic sarcoid biomarker, similar conclusions are made regarding other soluble biomarkers outlined below.

Serum amyloid A (SAA) is an acute phase protein produced in the liver and upregulated by monocyte and macrophage-derived cytokines [21]. SAA levels were significantly higher in sarcoidosis patients than in healthy controls [10, 22, 23] and significantly higher in sarcoidosis patients with active disease [13]. However, similar to ACE, this biomarker suffers from low specificity although SAA levels may be more useful during follow-up as SAA levels are less sensitive to immunosuppressive drugs, such as corticosteroids. Cytokines such as **tumor necrosis factor (TNF- α , TNF- β)** play a major role in granuloma formation and are released in greater quantities from alveolar macrophages obtained from sarcoidosis patients [24].

Lysozyme is an enzyme produced by macrophages to degrade bacteria and is elevated in numerous inflammatory conditions, including sarcoidosis. Lysozyme historically has been associated with extrapulmonary sarcoidosis, particularly uveitis [25] and in a recent study was more sensitive than ACE [26]. **Serum chitotriosidase** concentrations were significantly higher in advanced (stage 3) sarcoidosis compared to healthy controls, directly correlated with ACE levels and were highest in those with persistent disease on therapy [26].

Soluble biomarkers associated with lymphocyte origin

Soluble Interleukin-2 receptor (sIL-2R) levels shed from lymphocytes are increased in active sarcoidosis and, similar to ACE levels, may predict response to therapy [16]. Elevated levels have been correlated with parenchymal infiltration and lung function [27, 28]. Similarly, persistently elevated sIL-2R may suggest extrapulmonary manifestations of sarcoidosis [13, 24, 29]. **IL-17**, an interleukin important in mucosal immunity and autoimmunity, is increased in patients with ocular sarcoidosis [30] and in the BAL fluid of patients with Lofgren's syndrome.

Biomarkers associated with fibrosis and the extracellular matrix

Tenascin-C, an extracellular matrix molecule expressed during wound healing in various tissues, is increased in granulomatous sarcoidosis [29] and in BALF in patients with parenchymal infiltration on chest radiographs [30]. **Transforming growth factor TGF- β 1** is associated with tissue healing and recruits fibroblasts and myofibroblasts to the matrix, with over expression of TGF- β 1 fibrosis can occur. Significantly higher levels of TGF- β 1 and ACE were reported in sarcoidosis patients [12].

Measurement of biomarkers to diagnose and predict remitting or progressive disease remains promissory and relevant in the management of sarcoidosis. Despite a rich history and intense study, however, it is clear that no single soluble biomarker has proven to be sufficiently sensitive and specific to be recommended for widespread clinical use.

Emerging strategies for biomarker development

Recent advances in genome-wide expression profiling techniques, such as gene expression microarray and RNA sequencing, provide opportunities to discover novel disease mechanisms. These high-throughput approaches have been applied to diseases of unknown cause to help understand disease pathogenesis [31-33] including identification of disease-associated candidate genes with diagnostic and prognostic applications [34-37].

Sarcoidosis biomarkers identified by expression profiling in lung tissues

In a targeted approach utilizing primers to well established sarcoid inflammatory pathways, Christoph and coworkers isolated RNA from paraffin-embedded samples and compared sarcoid granulomas (mostly lymph nodes) gene expression to granulomas of other causes (fungal granulomas from lung and foreign body granulomas from skin). Interestingly, T-bet mRNA expression was the only marker significantly greater in sarcoid granulomas than both suture granulomas and fungal granulomas. Expression of Cox-2, IFN γ and IRF-1 was significantly higher in sarcoid granulomas than suture granulomas. IL-13 was more highly expressed in fungal granulomas than in suture granulomas, but not significantly differ from sarcoid granulomas [38]. In a more comprehensive study to identify the genes contributing to inflammation and lung remodeling in patients with pulmonary sarcoidosis, Crouser *et al* compared the gene expression pattern between normal lung tissues (n=6) and tissues obtained from untreated patients with pulmonary sarcoidosis (n=6), using Affymetrix Human Genome U133 Plus 2.0 Array (Gene Expression Omnibus [GEO] accession number: GSE16538) [39]. This whole-transcriptome expression analysis of lung tissue identified interacting gene networks that engage in the maintenance of granulomatous inflammation of the lung and associated lung damage [39]. As expected, many of the genes identified in the most over-represented network are associated with Th1-type immune responses, such as *STAT1*, *CCL5*, *IL7*, and *IL15* [39]. Furthermore, the expression of two genes regulating macrophage-derived proteases, matrix metalloproteinase 12 (*MMP12*) and ADAM-like, decysin 1 (*ADAMDEC1*), was dramatically up-regulated in sarcoidosis lung tissues [39]. These findings were validated in an independent series of patients with sarcoidosis, wherein *MMP12* and *ADAMDEC1* gene/protein expression in bronchoalveolar lavage (BAL) fluid correlated with disease severity [39]. Therefore, Crouser *et al* suggested that *MMP12* and *ADAMDEC1* are potential mediators of lung damage and/or remodeling and may serve as biomarkers of pulmonary sarcoidosis [39].

The majority of patients with pulmonary sarcoidosis recover spontaneously, whereas a significant number of patients exhibit progressive disease leading to varying degrees of pulmonary fibrosis [1, 40, 41]. To understand the molecular basis of fibrotic progression in pulmonary sarcoidosis, Lockstone *et al* examined the gene expression pattern in transbronchial biopsies of granulomatous areas in lung of patients with self-limiting sarcoidosis and patients with progressive-fibrotic sarcoidosis, using Affymetrix Human Gene 1.0 ST Array (GEO accession number: GSE19976) [42]. In total, 334 genes were found to be differentially-expressed between self-limiting and progressive-fibrotic sarcoidosis. Gene Set Enrichment Analysis showed that the gene up-regulated in lung samples obtained from patients with progressive-fibrotic sarcoidosis comprised

predominantly the genes involved in host defense and immune responses [42]. In addition, genes overexpressed in patients exhibiting progressive-fibrotic sarcoidosis are also significantly enriched for genes up-regulated in hypersensitivity pneumonitis, another granulomatous lung disease [42]. Lockstone *et al* suggested that the gene expression profiling in transbronchial lung biopsy samples can be used for prognostic purposes in pulmonary sarcoidosis.

Gene signatures in blood serve as potential diagnostic tool for sarcoidosis

The above-cited studies suggest that gene expression profiling in lung tissues displays potential diagnostic and prognostic power in sarcoidosis. However, development of less-invasive biomarkers that predict clinical course or sarcoidosis disease status remains an urgent need [43]. To determine whether gene expression profiling of blood elements reflects inflammatory pathways in the lung of sarcoidosis patients, Koth *et al* analyzed the transcriptomic gene expression data from whole blood of sarcoidosis patients enrolled at University of California, San Francisco (UCSF cohort), using Affymetrix Human Genome U133 Plus 2.0 Array (GEO accession number: GSE19314) [44] and built a machine-learning algorithm-based classifier using the UCSF microarray data, which distinguished sarcoidosis patients from healthy controls in an external validation cohort with 92% sensitivity and 92% specificity [44]. To understand how whole blood gene expression patterns relate to lung granulomatous tissues, differentially-expressed genes of sarcoidosis in blood (USCF dataset) were compared to lung tissue (dataset generated by Crouser *et al* [39]). Gene expression profiles induced in blood were found to significantly overlap with those in lung biopsies [44] with concordantly dysregulated genes identified as critical transcriptional regulators in interferon signaling pathways [44]. These authors conclude that the gene expression signature identified in blood provides a highly useful, non-invasive method to assess inflammatory activities in sarcoidosis.

Recently, we analyzed genome-wide gene expression in peripheral blood mononuclear cells (PBMCs) in sarcoidosis patients enrolled from the Chicago area (Chicago cohort), using Affymetrix Human Exon 1.0 ST Array (GEO accession number: GSE37912) [45]. A 20-gene signature was identified that distinguishes uncomplicated sarcoidosis from complicated sarcoidosis defined as progressive lung sarcoidosis with fibrosis, or extrapulmonary manifestations such as cardiac sarcoidosis and neurosarcoidosis. The expression levels of the genes within 20-gene signature showed a pattern of an additive model between uncomplicated and complicated sarcoidosis, i.e., when the signature gene is up-regulated, patients with complicated sarcoidosis exhibited higher expression levels than patients with uncomplicated sarcoidosis. Conversely, complicated sarcoidosis cases exhibit lower expression levels than patients with uncomplicated sarcoidosis when the signature gene is down-regulated [45]. This gene signature exhibited a substantial predictive accuracy when classifying sarcoidosis patients from healthy controls in two independent external cohorts. Additional validation strategies included significant genetic association of single nucleotide polymorphisms (SNPs) in signature genes with sarcoidosis susceptibility and severity [45].

We have evaluated the performance of our blood gene signature (Zhou *et al*, Chicago signature), the signature developed by Koth *et al* (UCSF signature) and an independent

validation dataset from Oregon Health Sciences University (Oregon cohort; GEO accession number: GSE18781) [46], in which genome-wide gene expression data in PBMC are available for 25 healthy controls and 12 sarcoidosis patients. The predictive power of the top 20 discriminative genes proposed by Koth *et al* was compared with that of the 20-gene signature proposed by Zhou *et al* (**Table 2**). Despite the absence of genes shared by these two gene sets, the principal component analysis (PCA) indicated that both the Chicago and UCSF signatures significantly distinguish patients with sarcoidosis from healthy controls in the Oregon cohort (**Figure 1A and 1B**). In order to systematically compare these two signatures, a classification score was assigned to each subject based on the first principal component of the given signature [47]. We found that the classification score based on both signatures distinguishes sarcoidosis patients from healthy controls with good accuracy: the areas under the receiver operating characteristic (ROC) curve (AUC) were 0.957 and 0.963 for the Chicago and UCSF signature respectively (**Figure 1C**). Significant differences between the AUCs of both signatures were not detected (DeLong's Test: $P = 0.843$).

A controversial computational study [47] suggested that the majority of published gene signatures failed to perform significantly better than gene sets of identical size that were randomly selected from human genome. To address this issue in our study, we conducted a resampling test and obtained 10,000 random gene signatures by randomly selecting 20 genes from human genome (the same size as the Chicago and UCSF signatures) and calculated the AUC for each random gene signature. Our alternative hypothesis was that the AUCs of the Chicago and UCSF signatures should be more positive than expected by chance if the predictive power of both signatures was significantly better than the random gene sets. Our analysis indicates that for both signatures, the null hypothesis that the predictive power is by chance could be rejected. The AUC of both signatures is significantly larger than that of random gene signatures (Right-tailed: $P = 0.0005$ for the Chicago signature; $P = 0.0003$ for the UCSF signature) (**Figure 1D**). However, it should also be noted that the AUC is larger than 0.8 for most random gene signatures (**Figure 1D**). Therefore, the performance of a signature to identify gene signatures for a given disease, cannot be measured by the prediction accuracy (e.g. AUC) as many randomly-generated gene signatures could also classify subjects with a fairly low error rate (indicated in **Figure 1D**). In light of this finding, we suggest that resampling test should be a standard procedure when generating biomarkers or molecular signatures for specific human disease. Prediction accuracy or nominal P -values do not address the appropriate statistical question as to whether a given set of genes is related to disease and more related to disease than random gene sets [47].

Distinguishing pulmonary sarcoidosis from tuberculosis by blood gene expression profiling

While the etiology of sarcoidosis remains unclear, there has long been an implicated linkage to mycobacterial and propionibacter organisms [48-53], although a consensus on the nature of a microbial pathogenesis in sarcoidosis and environmental factors (e.g., mold/mildew exposure) [15] has not yet been reached. Koth *et al* pointed out the significant overlap in gene expression profiles between sarcoidosis and tuberculosis due to the histologic similarities (e.g., interferon signaling-related genes) [44]. Similar observations were reported

by two independent groups [54, 55]. However, blood transcriptional heterogeneity between sarcoidosis and tuberculosis has been explored. For example, Maertzdorf et al analyzed the difference in PBMC gene expression between sarcoidosis and tuberculosis, using Agilent-014850 Whole Human Genome Microarray and found that the gene expression profiles of sarcoidosis showed an enrichment of down-regulated genes involved in mitochondrial oxidative phosphorylation and translational activity compared with tuberculosis patients [54]. Furthermore, sarcoidosis patients displayed significantly lower expression levels of genes related to antimicrobial defense responses [54]. Bloom et al compared the transcriptional profiles in whole blood between sarcoidosis and tuberculosis, using Illumina HumanHT-12 V4.0 expression beadchip [55] and identified a gene signature consisting of 144 transcripts that showed good sensitivity and specificity in all three independent cohorts from their own study (training, test, and validation sets) [55] and an external cohort from the Maertzdorf et al study [54]. However, few overlapping transcripts existed between the Bloom et al and Maertzdorf et al studies [55].

High-throughput genetic studies to identify novel candidate genes of sarcoidosis

Complex diseases such as sarcoidosis are likely influenced by multiple environmental factors and genetic variation. Previous studies based on *a priori* assumptions from clinical observations have suggested that the susceptibility to sarcoidosis is partially affected by the genetic variation in the genes and pathways related to granuloma formation and immune response such as HLA Class I and Class II genes [1, 2, 56, 57], *IL1A* (interleukin 1 α) [58], *IFNG* (interferon γ) [59], *NRAMP1* (natural resistance associated macrophage protein) [60], *IL18* (interleukin 18) [57], *CFTR* (cystic fibrosis transmembrane regulator) [61, 62]. The heterogeneity of sarcoidosis in clinical course and development of organ involvement [1, 2] suggests that an individual patient's genetic make-up may interact with various genetic and non-genetic factors to present certain clinical manifestations. High-throughput techniques such as whole genome genotyping array and next-generation sequencing technologies, allow for interrogation of the entire human genome with a high resolution for common genetic variants associated with sarcoidosis patients, as well as for those genetic markers potentially associated with complicated sarcoidosis by carefully defining the sarcoidosis subphenotypes. These high-throughput techniques significantly expand the coverage and resolution of the human genome, compared with several previously published genetic studies using microsatellite markers for identifying sarcoidosis susceptibility loci [63, 64]. More importantly, whole genome screening for genetic variants associated with sarcoidosis will aid identification of novel candidate genes implicated in sarcoidosis pathology, which potentially serve as biomarkers in this disease [65].

The first genome wide association study (GWAS) in sarcoidosis was conducted by Hofmann *et al* [66], using Affymetrix Genome-Wide Human SNP Array 5.0 in a predominant German population. This study identified a nonsynonymous single nucleotide polymorphism (SNP), rs1049550, residing in the first of four annexin core domains within the gene *ANXA11* (annexin A11), which was associated with sarcoidosis susceptibility [66], a finding confirmed in an independent case-control study in Czech patients [67] and in African and

European Americans [68, 69]. In a separate GWAS conducted by Hofmann *et al*, a nonsynonymous SNP, rs1040461 in gene *RAB23* (RAB23, member RAS oncogene family), was associated with sarcoidosis in German population with this association replicated in African Americans but not in European Americans [69]. More recently, a sarcoidosis GWAS in both African and European Americans, using Illumina HumanOmni1-Quad array [69], identified a novel sarcoidosis-associated SNP, rs715299, within the intro region of the gene *NOTCH4* (notch 4) in African Americans but not in European Americans [69]. The locus of this SNP is close to several MHC Class II genes known to be associated with sarcoidosis [70, 71] and within the region of high linkage disequilibrium [72]. Using stepwise conditional association analyses, it was confirmed that the observed signal within *NOTCH4* is independent of the SNPs within the MHC Class II genes [69].

Conclusions

Sarcoidosis is a challenging disease of undefined etiology and heterogeneous clinical course. Although there has been significant progress over the past years; significant and vexing questions remain unsolved. Relevant and promissory approaches have been developed to identify potential, blood transcripts and genomic signatures to facilitate diagnosis and predict disease course. Better understanding of environmental factors and molecular approaches such as profiling whole genome gene expression could provide an opportunity to explore genome-based biomarkers associated with sarcoidosis and advance precision medicine into clinical practice.

Acknowledgments

Funding source: This work was supported by National Institutes of Health U01 HL112696 (JGNG).

References

1. Iannuzzi MC, Rybicki BA, Teirstein AS. Sarcoidosis. *N Engl J Med*. 2007; 357(21):2153–2165. [PubMed: 18032765]
2. Newman LS, Rose CS, Maier LA. Sarcoidosis. *N Engl J Med*. 1997; 336(17):1224–1234. [PubMed: 9110911]
3. Byun CW, Yang SN, Yoon JS, Kim SH. Lofgren's Syndrome-Acute Onset Sarcoidosis and Polyarthralgia: A Case Report. *Ann Rehabil Med*. 2013; 37(2):295–299. [PubMed: 23705129]
4. Statement on sarcoidosis. Joint Statement of the American Thoracic Society (ATS), the European Respiratory Society (ERS) and the World Association of Sarcoidosis and Other Granulomatous Disorders (WASOG) adopted by the ATS Board of Directors and by the ERS Executive Committee. *Am J Respir Crit Care Med*. Feb; 1999 160(2):736–755.
5. Nunes H, Bouvry D, Soler P, Valeyre D. Sarcoidosis. *Orphanet J Rare Dis*. 2007; 2:46. [PubMed: 18021432]
6. What Is Sarcoidosis?. [<https://www.nhlbi.nih.gov/health/health-topics/topics/sarc/>]
7. Sweiss NJ, Patterson K, Sawaqed R, Jabbar U, Korsten P, Hogarth K, Wollman R, Garcia JG, Niewold TB, Baughman RP. Rheumatologic manifestations of sarcoidosis. *Semin Respir Crit Care Med*. 2010; 31(4):463–473. [PubMed: 20665396]
8. Drent M, Cremers JP, Jansen TL. Pulmonology meets rheumatology in sarcoidosis: a review on the therapeutic approach. *Curr Opin Rheumatol*. 2014; 26(3):276–284. [PubMed: 24614277]
9. Rybicki BA, Major M, Popovich J Jr, Maliarik MJ, Iannuzzi MC. Racial differences in sarcoidosis incidence: a 5-year study in a health maintenance organization. *Am J Epidemiol*. 1997; 145(3):234–241. [PubMed: 9012596]

10. Bons, JA.; Drent, M.; Bouwman, FG.; Mariman, EC.; van Dieijen-Visser, MP.; Wodzig, WK. *Respir Med.* Vol. 101. Vol. 1687-1695. England: 2007. Potential biomarkers for diagnosis of sarcoidosis using proteomics in serum..
11. Rozy A, Czerniawska J, Stepniewska A, Wozbinska B, Goljan A, Puscinska E, Gorecka D, Chorostowska-Wynimko J. Inflammatory markers in the exhaled breath condensate of patients with pulmonary sarcoidosis. *J Physiol Pharmacol.* 2006; 57(Suppl 4):335–340. [PubMed: 17072062]
12. Ahmadzai H, Cameron B, Chui J, Lloyd A, Wakefield D, Thomas PS. Measurement of neopterin, TGF-beta(1) and ACE in the exhaled breath condensate of patients with sarcoidosis. *Journal of Breath Research.* 2013; 7(4)
13. Gungor S, Ozseker F, Yalcinsoy M, Akkaya E, Can G, Eroglu H, Genc NS. Conventional markers in determination of activity of sarcoidosis. *Int Immunopharmacol.* 2015; 25(1):174–179. [PubMed: 25623898]
14. Tahmoush AJ, Amir MS, Connor WW, Farry JK, Didato S, Ulhoa-Cintra A, Vasas JM, Schwartzman RJ, Israel HL, Patrick H. CSF-ACE activity in probable CNS neurosarcoidosis. *Sarcoidosis Vasc Diffuse Lung Dis.* 2002; 19(3):191–197. [PubMed: 12405488]
15. Beneteau-Burnat B, Baudin B. Angiotensin-converting enzyme: clinical applications and laboratory investigations on serum and other biological fluids. *Crit Rev Clin Lab Sci.* 1991; 28(5-6):337–356. [PubMed: 1663362]
16. Baughman, R. *Sarcoidosis.* Vol. 210. Taylor and Francis Group; New York, NY: 2006.
17. Baudin B. [Angiotensin I-converting enzyme (ACE) for sarcoidosis diagnosis]. *Pathol Biol (Paris).* 2005; 53(3):183–188. [PubMed: 15781381]
18. Alia, P.; Mana, J.; Capdevila, O.; Alvarez, A.; Navarro, MA. *Scand J Clin Lab Invest.* Vol. 65. Vol. 691-697. Norway: 2005. Association between ACE gene I/D polymorphism and clinical presentation and prognosis of sarcoidosis..
19. Floe, A.; Hoffmann, HJ.; Nissen, PH.; Moller, HJ.; Hilberg, O. *Dan Med J.* Vol. 61. Vol. A4815. Denmark: 2014. Genotyping increases the yield of angiotensin-converting enzyme in sarcoidosis--a systematic review..
20. Pietinalho A, Furuya K, Yamaguchi E, Kawakami Y, Selroos O. The angiotensin-converting enzyme DD gene is associated with poor prognosis in Finnish sarcoidosis patients. *Eur Respir J.* 1999; 13(4):723–726. [PubMed: 10362030]
21. Salazar A, Pinto X, Mana J. Serum amyloid A and high-density lipoprotein cholesterol: serum markers of inflammation in sarcoidosis and other systemic disorders. *Eur J Clin Invest.* 2001; 31(12):1070–1077. [PubMed: 11903494]
22. Bargagli E, Bennett D, Maggiorelli C, Di Sipio P, Margollicci M, Bianchi N, Rottoli P. Human Chitotriosidase: a Sensitive Biomarker of Sarcoidosis. *Journal of Clinical Immunology.* 2013; 33(1):264–270. [PubMed: 22878841]
23. Zhang, Y.; Chen, X.; Hu, Y.; Du, S.; Shen, L.; He, Y.; Zhang, X.; Li, H.; Yung, RC. *Respir Res.* Vol. 14. Vol. 18. England: 2013. Preliminary characterizations of a serum biomarker for sarcoidosis by comparative proteomic approach with tandem-mass spectrometry in ethnic Han Chinese patients..
24. Bargagli E, Mazzi A, Rottoli P. Markers of Inflammation in Sarcoidosis: Blood, Urine, BAL, Sputum, and Exhaled Gas. *Clinics in Chest Medicine.* 2008; 29(3):445–458. [PubMed: 18539237]
25. Herbort CP, Rao NA, Mochizuki M. International criteria for the diagnosis of ocular sarcoidosis: results of the first International Workshop On Ocular Sarcoidosis (IWOS). *Ocul Immunol Inflamm.* 2009; 17(3):160–169. [PubMed: 19585358]
26. Febvay C, Kodjikian L, Maucourt-Boulch D, Perard L, Iwaz J, Jamilloux Y, Broussolle C, Burillon C, Seve P. Clinical features and diagnostic evaluation of 83 biopsy-proven sarcoid uveitis cases. *Br J Ophthalmol.* 2015
27. Miyoshi S, Hamada H, Kadowaki T, Hamaguchi N, Ito R, Irifune K, Higaki J. COMparative evaluation of serum markers in pulmonary sarcoidosis. *Chest.* 2010; 137(6):1391–1397. [PubMed: 20081103]

28. Vorselaars AD, van Moorsel CH, Zanen P, Ruven HJ, Claessen AM, van Velzen-Blad H, Grutters JC. ACE and sIL-2R correlate with lung function improvement in sarcoidosis during methotrexate therapy. *Respir Med*. 2015; 109(2):279–285. [PubMed: 25496652]
29. Grutters JC, Fellrath JM, Mulder L, Janssen R, van den Bosch JM, van Velzen-Blad H. Serum soluble interleukin-2 receptor measurement in patients with sarcoidosis: a clinical evaluation. *Chest*. 2003; 124(1):186–195. [PubMed: 12853522]
30. Wu, W.; Jin, M.; Wang, Y.; Liu, B.; Shen, D.; Chen, P.; Hannes, S.; Li, Z.; Hirani, S.; Jawad, S., et al. *J Transl Med*. Vol. 12. Vol. 152. England: 2014. Overexpression of IL-17RC associated with ocular sarcoidosis..
31. Calvano SE, Xiao W, Richards DR, Felciano RM, Baker HV, Cho RJ, Chen RO, Brownstein BH, Cobb JP, Tschoeke SK, et al. A network-based analysis of systemic inflammation in humans. *Nature*. 2005; 437(7061):1032–1037. [PubMed: 16136080]
32. Rolph MS, Sisavanh M, Liu SM, Mackay CR. Clues to asthma pathogenesis from microarray expression studies. *Pharmacol Ther*. 2006; 109(1-2):284–294. [PubMed: 16203040]
33. Selman M, Pardo A, Kaminski N. Idiopathic pulmonary fibrosis: aberrant recapitulation of developmental programs? *PLoS Med*. 2008; 5(3):e62. [PubMed: 18318599]
34. Herazo-Maya JD, Noth I, Duncan SR, Kim S, Ma SF, Tseng GC, Feingold E, Juan-Guardela BM, Richards TJ, Lussier Y, et al. Peripheral blood mononuclear cell gene expression profiles predict poor outcome in idiopathic pulmonary fibrosis. *Sci Transl Med*. 2013; 5(205):205ra136.
35. Zhou T, Wang T, Garcia JG. Expression of nicotinamide phosphoribosyltransferase-influenced genes predicts recurrence-free survival in lung and breast cancers. *Sci Rep*. 2014; 4:6107. [PubMed: 25146220]
36. Desai AA, Zhou T, Ahmad H, Zhang W, Mu W, Trevino S, Wade MS, Raghavachari N, Kato GJ, Peters-Lawrence MH, et al. A novel molecular signature for elevated tricuspid regurgitation velocity in sickle cell disease. *Am J Respir Crit Care Med*. 2012; 186(4):359–368. [PubMed: 22679008]
37. Pitroda SP, Zhou T, Sweis RF, Filippo M, Labay E, Beckett MA, Mauceri HJ, Liang H, Darga TE, Perakis S, et al. Tumor endothelial inflammation predicts clinical outcome in diverse human cancers. *PLoS One*. 2012; 7(10):e46104. [PubMed: 23056240]
38. Christophi GP, Caza T, Curtiss C, Gumber D, Massa PT, Landas SK. Gene expression profiles in granuloma tissue reveal novel diagnostic markers in sarcoidosis. *Exp Mol Pathol*. 2014; 96(3):393–399. [PubMed: 24768588]
39. Crouser ED, Culver DA, Knox KS, Julian MW, Shao G, Abraham S, Liyanarachchi S, Macre JE, Wewers MD, Gavrilin MA, et al. Gene expression profiling identifies MMP-12 and ADAMDEC1 as potential pathogenic mediators of pulmonary sarcoidosis. *Am J Respir Crit Care Med*. 2009; 179(10):929–938. [PubMed: 19218196]
40. Neville E, Walker AN, James DG. Prognostic factors predicting the outcome of sarcoidosis: an analysis of 818 patients. *Q J Med*. 1983; 52(208):525–533. [PubMed: 6657915]
41. Hillerdal G, Nou E, Osterman K, Schmekel B. Sarcoidosis: epidemiology and prognosis. A 15-year European study. *Am Rev Respir Dis*. 1984; 130(1):29–32. [PubMed: 6742607]
42. Lockstone HE, Sanderson S, Kulakova N, Baban D, Leonard A, Kok WL, McGowan S, McMichael AJ, Ho LP. Gene set analysis of lung samples provides insight into pathogenesis of progressive, fibrotic pulmonary sarcoidosis. *Am J Respir Crit Care Med*. 2010; 181(12):1367–1375. [PubMed: 20194811]
43. Su R, Li MM, Bhakta NR, Solberg OD, Darnell EP, Ramstein J, Garudadri S, Ho M, Woodruff PG, Koth LL. Longitudinal analysis of sarcoidosis blood transcriptomic signatures and disease outcomes. *Eur Respir J*. 2014; 44(4):985–993. [PubMed: 25142485]
44. 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(10):1153–1163. [PubMed: 21852540]
45. Zhou T, Zhang W, Sweiss NJ, Chen ES, Moller DR, Knox KS, Ma SF, Wade MS, Noth I, Machado RF, et al. Peripheral blood gene expression as a novel genomic biomarker in complicated sarcoidosis. *PLoS One*. 2012; 7(9):e44818. [PubMed: 22984568]

46. Sharma SM, Choi D, Planck SR, Harrington CA, Austin CR, Lewis JA, Diebel TN, Martin TM, Smith JR, Rosenbaum JT. Insights in to the pathogenesis of axial spondyloarthritis based on gene expression profiles. *Arthritis Res Ther*. 2009; 11(6):R168. [PubMed: 19900269]
47. Venet D, Dumont JE, Detours V. Most random gene expression signatures are significantly associated with breast cancer outcome. *PLoS Comput Biol*. 2011; 7(10):e1002240. [PubMed: 22028643]
48. Song Z, Marzilli L, Greenlee BM, Chen ES, Silver RF, Askin FB, Teirstein AS, Zhang Y, Cotter RJ, Moller DR. Mycobacterial catalase-peroxidase is a tissue antigen and target of the adaptive immune response in systemic sarcoidosis. *J Exp Med*. 2005; 201(5):755–767. [PubMed: 15753209]
49. Gupta D, Agarwal R, Aggarwal AN, Jindal SK. Molecular evidence for the role of mycobacteria in sarcoidosis: a meta-analysis. *Eur Respir J*. 2007; 30(3):508–516. [PubMed: 17537780]
50. Drake WP, Dhason MS, Nadaf M, Shepherd BE, Vadivelu S, Hajizadeh R, Newman LS, Kalams SA. Cellular recognition of Mycobacterium tuberculosis ESAT-6 and KatG peptides in systemic sarcoidosis. *Infect Immun*. 2007; 75(1):527–530. [PubMed: 17088357]
51. Oswald-Richter KA, Beachboard DC, Seeley EH, Abraham S, Shepherd BE, Jenkins CA, Culver DA, Caprioli RM, Drake WP. Dual analysis for mycobacteria and propionibacteria in sarcoidosis BAL. *J Clin Immunol*. 2012; 32(5):1129–1140. [PubMed: 22552860]
52. Ishige I, Usui Y, Takemura T, Eishi Y. Quantitative PCR of mycobacterial and propionibacterial DNA in lymph nodes of Japanese patients with sarcoidosis. *Lancet*. 1999; 354(9173):120–123. [PubMed: 10408488]
53. Chen ES, Moller DR. Etiologic role of infectious agents. *Semin Respir Crit Care Med*. 2014; 35(3):285–295. [PubMed: 25007081]
54. Maertzdorf J, Weiner J 3rd, Mollenkopf HJ, Bauer T, Prasse A, Muller-Quernheim J, Kaufmann SH. Common patterns and disease-related signatures in tuberculosis and sarcoidosis. *Proc Natl Acad Sci U S A*. 2012; 109(20):7853–7858. [PubMed: 22547807]
55. Bloom CI, Graham CM, Berry MP, Rozakeas F, Redford PS, Wang Y, Xu Z, Wilkinson KA, Wilkinson RJ, Kendrick Y, et al. Transcriptional blood signatures distinguish pulmonary tuberculosis, pulmonary sarcoidosis, pneumonias and lung cancers. *PLoS One*. 2013; 8(8):e70630. [PubMed: 23940611]
56. Iannuzzi MC, Rybicki BA. Genetics of sarcoidosis: candidate genes and genome scans. *Proc Am Thorac Soc*. 2007; 4(1):108–116. [PubMed: 17202299]
57. Kelly DM, Greene CM, Meachery G, O'Mahony M, Gallagher PM, Taggart CC, O'Neill SJ, McElvaney NG. Endotoxin up-regulates interleukin-18: potential role for gram-negative colonization in sarcoidosis. *Am J Respir Crit Care Med*. 2005; 172(10):1299–1307. [PubMed: 16100009]
58. Hutyrova B, Pantelidis P, Drabek J, Zurkova M, Kolek V, Lenhart K, Welsh KI, Du Bois RM, Petrek M. Interleukin-1 gene cluster polymorphisms in sarcoidosis and idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*. 2002; 165(2):148–151. [PubMed: 11790645]
59. Akahoshi M, Ishihara M, Remus N, Uno K, Miyake K, Hirota T, Nakashima K, Matsuda A, Kanda M, Enomoto T, et al. Association between IFNA genotype and the risk of sarcoidosis. *Hum Genet*. 2004; 114(5):503–509. [PubMed: 15004750]
60. Malirak MJ, Chen KM, Sheffer RG, Rybicki BA, Major ML, Popovich J Jr. Iannuzzi MC. The natural resistance-associated macrophage protein gene in African Americans with sarcoidosis. *American journal of respiratory cell and molecular biology*. 2000; 22(6):672–675. [PubMed: 10837363]
61. Bombieri C, Luisetti M, Belpinati F, Zuliani E, Beretta A, Baccheschi J, Casali L, Pignatti PF. Increased frequency of CFTR gene mutations in sarcoidosis: a case/control association study. *Eur J Hum Genet*. 2000; 8(9):717–720. [PubMed: 10980579]
62. Schurmann M, Albrecht M, Schwinger E, Stuhmann M. CFTR gene mutations in sarcoidosis. *Eur J Hum Genet*. 2002; 10(11):729–732. [PubMed: 12404105]
63. Iannuzzi MC, Iyengar SK, Gray-McGuire C, Elston RC, Baughman RP, Donohue JF, Hirst K, Judson MA, Kavuru MS, Malirak MJ, et al. Genome-wide search for sarcoidosis susceptibility genes in African Americans. *Genes Immun*. 2005; 6(6):509–518. [PubMed: 15951742]

64. Schurmann M, Reichel P, Muller-Myhsok B, Schlaak M, Muller-Quernheim J, Schwinger E. Results from a genome-wide search for predisposing genes in sarcoidosis. *Am J Respir Crit Care Med*. 2001; 164(5):840–846. [PubMed: 11549543]
65. Rybicki BA, Sinha R, Iyengar S, Gray-McGuire C, Elston RC, Iannuzzi MC. Genetic linkage analysis of sarcoidosis phenotypes: the sarcoidosis genetic analysis (SAGA) study. *Genes Immun*. 2007; 8(5):379–386. [PubMed: 17476268]
66. Hofmann S, Franke A, Fischer A, Jacobs G, Nothnagel M, Gaede KI, Schurmann M, Muller-Quernheim J, Krawczak M, Rosenstiel P, et al. Genome-wide association study identifies ANXA11 as a new susceptibility locus for sarcoidosis. *Nat Genet*. 2008; 40(9):1103–1106. [PubMed: 19165924]
67. Mrazek F, Stahelova A, Kriegova E, Fillerova R, Zurkova M, Kolek V, Petrek M. Functional variant ANXA11 R230C: true marker of protection and candidate disease modifier in sarcoidosis. *Genes Immun*. 2011; 12(6):490–494. [PubMed: 21562576]
68. Levin AM, Iannuzzi MC, Montgomery CG, Trudeau S, Datta I, McKeigue P, Fischer A, Nebel A, Rybicki BA. Association of ANXA11 genetic variation with sarcoidosis in African Americans and European Americans. *Genes Immun*. 2013; 14(1):13–18. [PubMed: 23151485]
69. Adrianto I, Lin CP, Hale JJ, Levin AM, Datta I, Parker R, Adler A, Kelly JA, Kaufman KM, Lessard CJ, 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(8):e43907. [PubMed: 22952805]
70. Rossman MD, Thompson B, Frederick M, Maliarik M, Iannuzzi MC, Rybicki BA, Pandey JP, Newman LS, Magira E, Beznik-Cizman B, et al. HLA-DRB1*1101: a significant risk factor for sarcoidosis in blacks and whites. *Am J Hum Genet*. 2003; 73(4):720–735. [PubMed: 14508706]
71. Grunewald J, Eklund A, Olerup O. Human leukocyte antigen class I alleles and the disease course in sarcoidosis patients. *Am J Respir Crit Care Med*. 2004; 169(6):696–702. [PubMed: 14656748]
72. Miretti MM, Walsh EC, Ke X, Delgado M, Griffiths M, Hunt S, Morrison J, Whittaker P, Lander ES, Cardon LR, et al. A high-resolution linkage-disequilibrium map of the human major histocompatibility complex and first generation of tag single-nucleotide polymorphisms. *Am J Hum Genet*. 2005; 76(4):634–646. [PubMed: 15747258]
73. Danilov SM, Balyasnikova IV, Danilova AS, Naperova IA, Arablinskaya NE, Borisov SE, Metzger R, Franke FE, Schwartz DE, Gachok IV, Trakht IN, Kost OA, Garcia JG. Conformational fingerprinting of the angiotensin I-converting enzyme (ACE). Application in sarcoidosis. *J Proteome Res*. 2010; 95782-93(11) PMID: 20873814.

Figure 1. Performance comparison between the Chicago and UCSF gene signatures in the Oregon cohort

(A) Principal component analysis on the expression of the Chicago signature. X-axis: the first principal component; Y-axis: the second principal component. (B) Principal component analysis on the expression of the UCSF signature. X-axis: the first principal component; Y-axis: the second principal component. (C) The ROC curves of the Chicago and UCSF signatures in classifying the subjects in the Oregon cohorts. (D) Superior predictive power of the 17-gene signature compared with random gene set. The gray area shows the distribution of the AUC for the 10,000 resampled gene signatures (with the identical size as the Chicago and UCSF signatures) randomly picked up from human genome. The solid triangle stands for the AUC of the Chicago signature and the empty triangle denotes the AUC of the UCSF signature. Right-tailed P -values of the sampling distribution were calculated.

Table 1**Conventional Sarcoidosis Biomarkers and their clinical association**

Biomarker	Origin and clinical association
ACE	Monocyte-macrophage origin. Acute stage, levels influenced by polymorphisms.
sIL-2R	Lymphocyte associated. Disease severity, extra-pulmonary organ involvement.
SAA	Monocyte-macrophage origin. Higher level in tissue and serum in sarcoidosis.
Alpha 1-antitrypsin (BALF)	Cytokine associated. Down-regulated only in patients without LS. Associated with spontaneous resolution.
Protocadherin-2 precursor (BALF)	Cell adhesion. Up-regulated in sarcoidosis across all studied phenotypes.
Chitotriosidase	Monocyte-macrophage origin. Disease progression.
Tenascin-C (BALF)	Fibrosis and ECM associated. Levels correlated with infiltrates on chest radiographs in sarcoidosis.
IL-17RC	Lymphocyte associated. Elevated expression in retinal tissues.
TGF-P1	Fibrosis and ECM related. Associated with pulmonary fibrosis.

Abbreviations: ACE, angiotensin converting enzyme; sIL-2R, soluble Interleukin-2 receptor; SAA, Serum amyloid; LS, Lofgren's syndrome ; BALF, bronchoalveolar lavage fluid; ECM, extracellular matrix.

Table 2

The Chicago and UCSF signatures

Cohort	Gene symbol	Gene title	D
Chicago	<i>HBEGF</i>	heparin-binding EGF-like growth factor	↑
	<i>SAP30</i>	Sin3A-associated protein, 30kDa	↑
	<i>APOBEC3D</i>	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3D	↓
	<i>CRIP1</i>	cysteine-rich protein 1 (intestinal)	↓
	<i>CX3CR1</i>	chemokine (C-X3-C motif) receptor 1	↓
	<i>FITM2</i>	fat storage-inducing transmembrane protein 2	↓
	<i>FKBP1A</i>	FK506 binding protein 1A, 12kDa	↓
	<i>KIAA1147</i>	KIAA1147	↓
	<i>KLRB1</i>	killer cell lectin-like receptor subfamily B, member 1	↓
	<i>LOC100132356</i>	hypothetical protein LOC100132356	↓
	<i>LOC100287290</i>	cytokine receptor CRL2	↓
	<i>MEI1</i>	meiosis inhibitor 1	↓
	<i>NOG</i>	noggin	↓
	<i>RBM12B</i>	RNA binding motif protein 12B	↓
	<i>SERTAD1</i>	SERTA domain containing 1	↓
	<i>SESN3</i>	sestrin 3	↓
	<i>TSHZ2</i>	teashirt zinc finger homeobox 2	↓
	<i>ZNF540</i>	zinc finger protein 540	↓
	<i>ZNF614</i>	zinc finger protein 614	↓
	<i>ZNF671</i>	zinc finger protein 671	↓
UCSF	<i>ATF3</i>	activating transcription factor 3	↑
	<i>CEACAM1</i>	carcinoembryonic antigen-related cell adhesion molecule 1	↑
	<i>DHRS9</i>	dehydrogenase/reductase (SDR family) member 9	↑
	<i>GBP2</i>	guanylate binding protein 2, interferon-inducible	↑
	<i>IRF1</i>	interferon regulatory factor 1	↑
	<i>STX11</i>	syntaxin 11	↑
	<i>TAP1</i>	transporter 1, ATP-binding cassette, sub-family B (MDR/TAP)	↑
	<i>CD27</i>	CD27 molecule	↓
	<i>CD3G</i>	CD3g molecule, gamma (CD3-TCR complex)	↓
	<i>CHIC1</i>	cysteine-rich hydrophobic domain 1	↓
	<i>GIMAP5</i>	GTPase, IMA family member 5	↓
	<i>IL7R</i>	interleukin 7 receptor	↓
	<i>KCNA3</i>	potassium voltage-gated channel, shaker-related subfamily, member 3	↓
	<i>LRRN3</i>	leucine rich repeat neuronal 3	↓
	<i>PAQR8</i>	progesterone and adipoQ receptor family member VIII	↓
	<i>TTC3</i>	tetratricopeptide repeat domain 3	↓
	<i>XPO4</i>	exportin 4	↓
	<i>ZNF512</i>	zinc finger protein 512	↓
	<i>ZNF662</i>	zinc finger protein 662	↓

Cohort	Gene symbol	Gene title	D
	<i>ZNF709</i>	zinc finger protein 709	↓

Note – “D” stands for the direction of dysregulation. “↑” indicates up-regulation in sarcoidosis while “↓” means down-regulation in sarcoidosis.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript