

Epidemiology of Sarcoidosis: Recent Advances and Future Prospects

Benjamin A. Rybicki, Ph.D.¹ and Michael C. Iannuzzi, M.D.²

ABSTRACT

Sarcoidosis is by definition a disease of “unknown causes,” but recent epidemiologic advances suggest that the long-standing definition of sarcoidosis may soon need to be amended. The recently completed ACCESS (A Case-Control Etiologic Study of Sarcoidosis) study was not able to definitively identify the “cause” of sarcoidosis, but yielded important findings regarding familial and environmental risks that have advanced our understanding of this disease. The *HLA-DRB1* associations reported in ACCESS along with the results of two recently completed genome scans of sarcoidosis in German Caucasians and African-Americans, respectively, have further defined the genetics of sarcoidosis. These studies suggest genetic heterogeneity of sarcoidosis risk between Caucasians and African-Americans and multiple susceptibility genes that interact together and with environmental factors in the disease pathogenesis. Genes that influence sarcoidosis clinical phenotypes may also be largely separate from sarcoidosis susceptibility genes. Although genetic studies of sarcoidosis in African-American populations are confounded by Caucasian admixture, this same admixture may be useful in identifying sarcoidosis genes linked with African ancestry. Case-only methods may be useful in identifying recent acute exposures linked to disease, genetic variants of risk, and gene-environment interactions. In summary, the epidemiology of sarcoidosis has a promising future that should eventually provide the answers to the etiologic origins of this complex disease.

KEYWORDS: Sarcoidosis, etiology, epidemiology, genetics, environmental exposure

CHANGING THE DEFINITION OF SARCOIDOSIS

The definition of sarcoidosis in the American Thoracic Society (ATS) statement on sarcoidosis,¹ is “a multi-system disorder of *unknown cause(s)*.” In general, sarcoidosis-like conditions with a well-defined cause are separately classified disease entities characterized according to their known etiologic agent, even though the disease may be clinically indistinguishable from sarcoidosis. For example, interstitial lung disorders charac-

terized by granulomatous inflammation, such as tuberculosis, silicosis, berylliosis, and hypersensitivity pneumonitis, are defined according to their single etiologic agent (i.e., mycobacteria, silica, beryllium, and endotoxins, respectively). Herein is the conundrum in the practice of sarcoidosis epidemiology, searching for an etiologic agent for a disease that by definition is of unknown cause(s). Although this working definition of sarcoidosis has challenged investigators with what some might characterize as an intractable problem, it

¹Department of Biostatistics and Research Epidemiology, Henry Ford Health System, Detroit, Michigan; ²Division of Pulmonary, Critical Care, and Sleep Medicine, Mount Sinai School of Medicine, New York, New York.

Address for correspondence and reprint requests: Benjamin A. Rybicki, Ph.D., Department of Biostatistics and Research Epidemiology, Henry Ford Health System, Detroit, MI 48202. E-mail:

brybick1@hfhs.org.

Sarcoidosis: Evolving Concepts and Controversies; Guest Editors, Marc A. Judson, M.D., Michael C. Iannuzzi, M.D.

Semin Respir Crit Care Med 2007;28:22–35. Copyright © 2007 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New York, NY 10001, USA. Tel: +1(212) 584-4662.

DOI 10.1055/s-2007-970331. ISSN 1069-3424.

has also more recently motivated ambitious epidemiological efforts to make the “unknown cause” definition of sarcoidosis obsolete. What follows is a brief description of these efforts and the most promising avenues for future epidemiological research of sarcoidosis.

DEFINING WHO HAS SARCOIDOSIS

In the ATS statement paper on sarcoidosis, the authors commented that the epidemiology of sarcoidosis was problematic for several reasons, including:

- (1) lack of a precise, consistent case definition,
- (2) variable methods of case ascertainment, (3) variability in disease presentation, (4) lack of sensitive and specific diagnostic tests, resulting in under recognition and misdiagnosis of the disease, and (5) the paucity of systematic epidemiologic investigations of cause.

Interestingly, four of the five problems in sarcoidosis epidemiological research that were cited involved difficulty with case definition and case heterogeneity. It is difficult to assess the extent to which poor case definitions and case heterogeneity have confounded epidemiological studies to date, but as discussed later in this review, methodological approaches focused on case samples and using case heterogeneity to extract clues about disease etiology are essential to moving the epidemiology of sarcoidosis forward.

A CASE-CONTROL ETIOLOGY STUDY OF SARCOIDOSIS (ACCESS)

The case-control study remains the workhorse of epidemiology, primarily because of its ease of implementation and the large amount of data that can be processed to yield associations that might give clues about disease etiology. In terms of its size, breadth, and methodological rigor, the recently completed ACCESS study is the preeminent sarcoidosis case-control study.² Although ACCESS has yielded important results relating to

familial, genetic, and environmental risk factors for sarcoidosis (some of which will be reviewed in more detail later in this article), it has failed to produce a breakthrough in defining the etiology of sarcoidosis.³ ACCESS has helped refine the risk estimates for some inherited factors long known to be associated with sarcoidosis, such as *HLA-DRB1*⁴ and having a first-degree relative with disease,⁵ and confirmed the consistent perplexing overlap of sarcoidosis risk factors with those of other similar interstitial lung diseases, such as microbial bioaerosols.⁶

THE DEMOGRAPHICS OF SARCOIDOSIS RISK AND ETIOLOGIC CLUES

Understanding the ethnic and sex differences associated with sarcoidosis was one of the ACCESS study goals and should continue to be a focus of future sarcoidosis investigations because this striking feature must contain important clues about etiologic agents. The variation in sarcoidosis incidence across ethnic groups is well documented, with African-Americans and northern Europeans having the highest rates of sarcoidosis incidence. Beyond the disease incidence variations across populations, ethnic differences also exist for clinical presentations, disease course, and genetic susceptibilities (Table 1).^{4,7-19} Although the female preponderance of sarcoidosis appears to be consistent across ethnic groups, some variation in the sex ratio of cases across ethnic groups has been observed. Also, although the age-specific incidence of sarcoidosis has prominent peaks, these peaks occur in different decades of life across ethnic groups. In African-Americans, a prominent peak incidence of sarcoidosis is observed in the fourth decade of life for both males and females.²⁰ In Japanese, the peak incidence of sarcoidosis occurs in the third decade of life,²¹ but some investigators have noted a bimodal peak in sarcoidosis occurrence, with the second peak seen in older female patients (> 50 years old).²¹⁻²³ A similar bimodal incidence peak in women has also been observed in Scandinavian populations.^{22,24} The etiologies behind these early and late cases of sarcoidosis

Table 1 Variations in Sarcoidosis Incidence, Presentation, and Genetic Associations across Ethnic Groups

Ethnic Group	Incidence per 100,000	Peak Decade of Incidence	Percent Increased Risk in Females	Typical Clinical Presentation and Course	Recent Reported Genetic Associations
European Americans	3–10	4th–5th	10–20	Stage I, acute course	<i>BTNL2</i> , ¹⁰ <i>HLA-DRB1</i> ⁴
African Americans	35–80	3rd–4th	30	Stage I–II, extrathoracic involvement	<i>HLA-DRB1</i> , ⁴ <i>HLA-DQB1</i> , ⁷ <i>IGKV</i> ¹¹
Northern Europeans	15–20	3rd	30	Stage I, acute course	<i>HSP70-hom</i> , ¹² <i>BTNL2</i> , ⁸ <i>NRAMP1</i> , ¹³ <i>TAP2</i> ¹⁴
Southern Europeans	1–5	4th–5th	33	Löfgren’s syndrome	<i>NOD2</i> , ¹⁵ <i>CR1</i> ¹⁶
Japanese	1–2	3rd	10–20	Ocular involvement, responsive to therapy	<i>IL-18</i> , ¹⁷ <i>IFNA17</i> , ¹⁸ <i>VEGF</i> , ¹⁹ <i>CCR2</i> ⁹

are likely different and more thorough investigations of early- and late-onset sarcoidosis as separate phenotypes may lessen problems with etiologic heterogeneity. Although the ethnic differences in sarcoidosis have long been appreciated and generally considered in studies designed to examine sarcoidosis etiology, age and sex differences in sarcoidosis presentation are less often taken into account or only considered in the analysis phase. When age/sex variations in disease-risk factors associations are sought, they are generally found. For instance, many case-control studies reporting genetic associations with sarcoidosis often find the associations are limited to, or more prominent in, one sex.^{4,25,26} In ACCESS, *HLA-DRB1*1101* had a higher odds ratio (OR) in males than in females (OR 3.6 for males and 1.51 for females).⁴ Although age and sex did not appear to have a significant effect on familial risk in the ACCESS sample,⁵ in a separate study of African-American families we found familial risk was greater in index cases that were younger and male.²⁷

Heterogeneity in clinical presentation by ethnicity is demonstrated in Fig. 1 taken from the ACCESS study.²⁸ For black subjects, there was more frequent involvement of the eye, liver, bone marrow, extrathoracic lymph nodes, and skin other than erythema nodosum. Abnormalities of calcium metabolism were more frequent among white subjects. For a subset of patients followed for 2 years in the ACCESS study,²⁹ blacks appeared to have more radiographic disease progression compared with whites (Fig. 2A). Longitudinal changes in pulmonary function were also more variable in blacks compared with whites (Figs. 2B,C), but there were no apparent racial differences in level of dyspnea (Fig. 2D).

Other clinical differences by age and sex were observed in ACCESS. Women were more likely to have eye and neurological involvement and erythema nodosum. There was some evidence of more frequent abnor-

malities of calcium metabolism in men. Those diagnosed under 40 year of age were more likely to have involvement of extrathoracic lymph nodes, whereas patients 40 year of age and over were more likely to have abnormal calcium metabolism. Overall, the variation in sarcoidosis risk and presentation across ethnicities, age groups, and the sexes remains fertile ground for finding additional clues about disease etiology.

PROMISING DEVELOPMENTS IN THE EPIDEMIOLOGY OF SARCOIDOSIS

Environmental Risk Factors

Because sarcoidosis most commonly involves the lungs, and because the two other organs most commonly affected, the eyes and skin, also have direct contact with the external environment, the search for environmental causes has centered on exposures to airborne antigens. Some of the earliest studies of sarcoidosis found associations between case status and rural-related exposures, such as wood-burning stoves,^{30,31} tree pollen,^{32,33} and soil exposures.³⁴⁻³⁶ More recently, exposure to wood-burning stoves has been implicated again as conferring increased risk for sarcoidosis,³⁷ as well as exposures to inorganic particles,³⁸⁻⁴⁰ insecticides,^{6,37} and moldy environments.^{6,41} Studies of occupations associated with sarcoidosis have found positive associations with ship's servicemen in the navy⁴²; metal work⁴¹; suppliers of building materials, hardware, and gardening materials⁴³; educators^{41,43}; and fire fighters.^{44,45} The rationale behind occupational studies is that they can help identify common environmental agents or type of exposures that might warrant more in-depth investigation. Unfortunately, occupational associations with sarcoidosis do not appear to have a unifying theme.

The most promising environmental finding from the ACCESS study was an association with potential exposure to microbial aerosols.⁶ Although exposure to work environments with mold/mildew exposures only increased risk for sarcoidosis by 60%, the exposure was robust because it encompassed increased odds ratios for several related specific exposures, including musty odors and home air conditioning. In addition, another independent study of African-American sibs found a similar association with wet or moldy environments⁴¹ as well as possible genetic interactions with this exposure.⁷ One hypothesis is that microbial bioaerosols may harbor mycobacteria, which have been associated with the development of other granulomatous diseases, such as hypersensitivity pneumonitis.⁴⁶ Although this hypothesis suffers from the inconsistency of finding evidence for mycobacterial infection in sarcoidosis granulomas,⁴⁷⁻⁴⁹ it still holds considerable attraction given how strongly the pathogenesis of sarcoidosis mimics that of a disease triggered by an infectious agent.⁵⁰ Future investigations

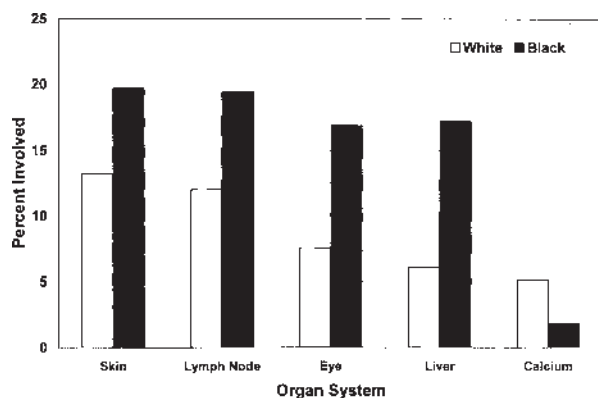


Figure 1 Comparison of proportion of organ involvement in A Case-Control Etiology Study of Sarcoidosis (ACCESS study) in which there was a significant difference between cases and controls on the basis of race. Adapted from Baughman et al.²⁸

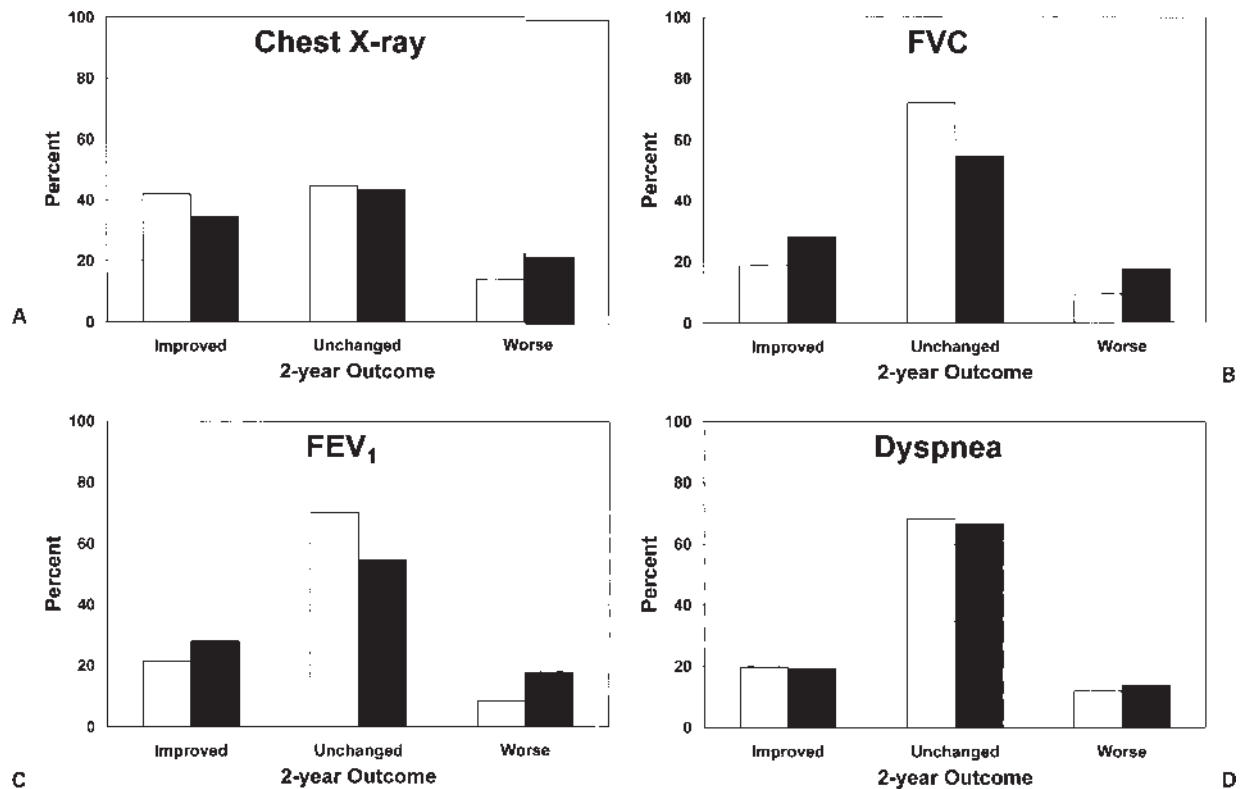


Figure 2 Percentage of African American (darkened bars) and Caucasian (open bars) sarcoidosis cases in A Case-Control Etiology Study of Sarcoidosis (ACCESS study) who, between the baseline and the 2 year follow-up period, improved, worsened, or were unchanged based on (A) chest x-ray, (B) forced vital capacity (FVC), (C) forced expiratory volume in 1 second (FEV₁), and (D) dyspnea. Adapted from Judson et al.²⁹

that can potentially advance this hypothesis include studies of immunologic responses to mycobacterial antigens by sarcoidosis subjects and assessment of genetic correlations with antigen recognition, as well as correlations with particular clinical phenotypes.⁵¹

Familial Aggregation of Sarcoidosis

Until recently, the aggregation of sarcoidosis in families was based primarily on anecdotal reports of familial clustering, with few controlled studies empirically demonstrating that relatives of sarcoidosis are at greater risk. Those studies that included referent populations had either very small samples,⁵² which precluded calculating precise estimates of familial risk, or suspect referent populations,^{53,54} which put into question the validity of their estimates. For African Americans, the first published estimates of familial relative risk for sarcoidosis that were age and sex adjusted to a standardized referent population came from a family-based study performed in a large health system in Detroit, Michigan.²⁷ The familial risk (λ) for sibs and parents was 2.24 and 2.82, respectively. For sibs and parents combined, λ was 2.49. Stratified results indicated that sarcoidosis familial risk was greater for relatives of younger ($\lambda = 2.93$) and male ($\lambda = 3.98$) probands.

This first comprehensive report of sarcoidosis familial relative risk in African Americans was followed shortly by a report from the ACCESS study, in which sarcoidosis familial relative risk was estimated using data on disease occurrence in 10,862 first- and 17,047 second-degree relatives of 706 age, sex, race, and geographically matched cases and controls.⁵ Sibs had the highest relative risk ($\lambda = 5.8$), followed by avuncular relationships ($\lambda = 5.7$), grandparents ($\lambda = 5.1$), and then parents ($\lambda = 3.8$). In a multivariate model fit to the parents and sibs data, the familial relative risk adjusted for age, sex, relative class, and shared environment was 4.7. Caucasian cases had a markedly higher familial relative risk compared with African American cases (18.0 versus 2.8; $p = .098$). These two studies confirmed what had been anecdotally reported for years, that family members of sarcoidosis cases were at increased risk for disease, but more importantly quantified this risk so that appropriate sample size estimates for future genetic linkage studies could be calculated.

Genetic Linkage Studies of Sarcoidosis

Schurmann and colleagues⁵⁵ reported the first genome-wide search for predisposing genes in sarcoidosis. On the basis of 225 microsatellite markers tested in 63 German

families (Caucasians) with affected siblings, they detected seven chromosomal regions of increased ($p < .05$) allele sharing.⁵⁵ The most prominent peak was noted for D6S1661 on chromosome 6p21 ($p = .001$). Minor peaks ($p < .05$) were found on chromosomes 1 (D1S1665), 3 (D3S1766), 7 (D7S821 and D7S3070), 9 (D9S934), and the X chromosome (DXS6789). Using case-control and family-based analyses on data from 69 SNPs across a 16.3 MB interval surrounding D6S1666, Valentonyte et al⁸ identified a 440 kb region of association around the *BTNL2* gene, a costimulatory molecule of the B7 family. By genotyping 48 single nucleotide polymorphisms (SNPs) from this region, they identified the associated variant rs2076530, a SNP in exon 5 of *BTNL2* in which the G to A variant alters the anatomy of a splice site resulting in a truncated protein. The susceptibility A allele of rs2076530 is characterized by odds ratios of 1.6 in heterozygotes and 2.75 in homozygotes. The population attributable risk was 23% for heterozygotes and homozygotes. Potential T cell downregulatory function of *BTNL2* could be impaired by the truncating mutation, which is predicted to cause loss of the immunoglobulin C (IgC) domain and transmembrane helix.⁸

Shortly after the initial genome scan for sarcoidosis susceptibility genes in a German population, the Sarcoidosis Genetic Analysis (SAGA) study conducted a

genome scan for sarcoidosis susceptibility genes in African Americans.⁵⁶ Using the Weber 10 microsatellite marker set (380 markers with an average intermarker spacing of 9.18 cM) in 229 African-American families, results of a Haseman–Elston regression⁵⁷ showed linkage peaks with p values less than .05 on chromosomes 1p22, 2p25, 5p15–13, 5q11, 5q35, 9q34, 11p15, and 20q13.⁵⁸ The most prominent peak was at D5S2500 on chromosome 5q11 ($p = .0005$), with two separate linkage peaks delineated on chromosomes 5p15–13 and 5q35.

Additional analyses of the SAGA data have helped refine the initial linkage results but also reveal the difficulty of doing genetic linkage studies in admixed populations such as African-Americans. Fine mapping excluded most of the original linkage signals, with the exception of the signals on chromosomes 5p15–13 and 5q11.⁵⁹ The results from this fine mapping study also suggested that the effects of the putative locus on 5p15–13 were not independent from another putative locus on chromosome 3p14–11. In another genetic analysis emanating from the original SAGA genome scan, Thompson et al showed evidence that population stratification had an effect on the majority of the original SAGA linkage signals.⁶⁰ In the two genetically distinct populations that were identified within the SAGA sample, only one of the two populations showed evidence for

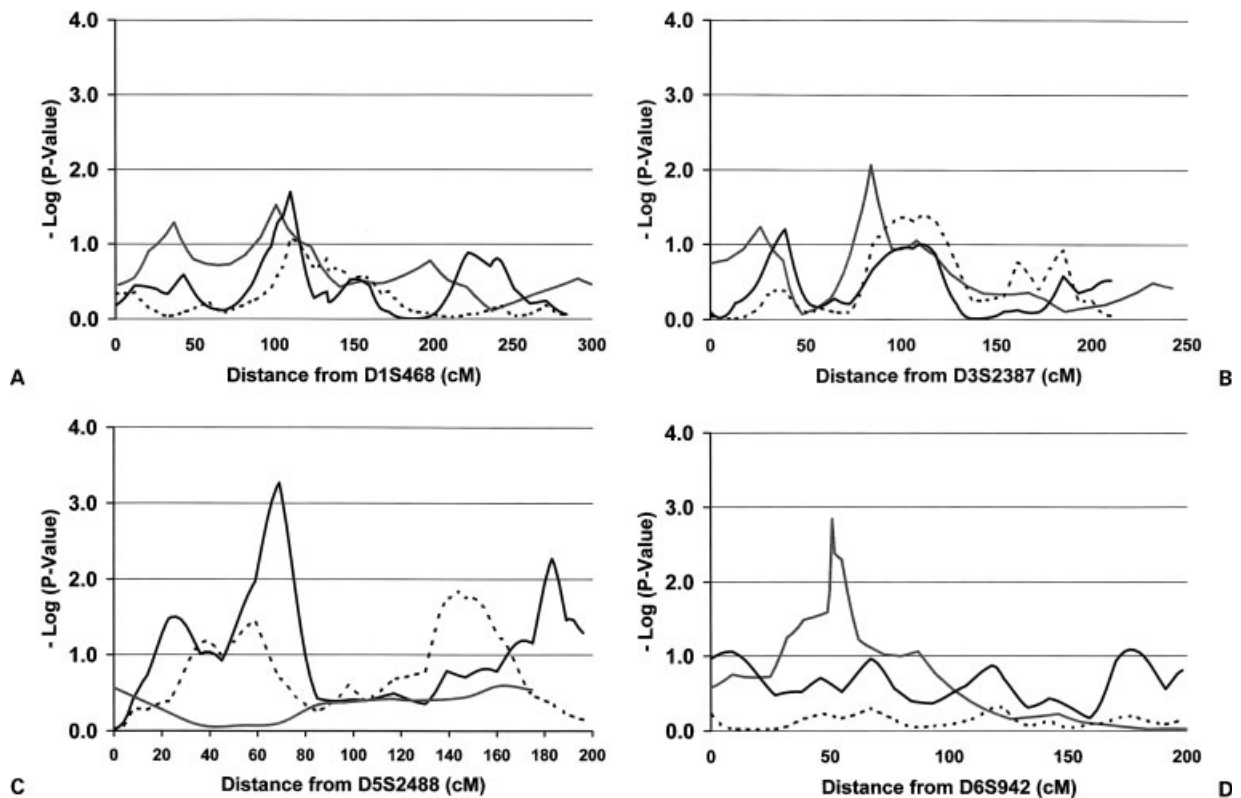


Figure 3 Comparison of results from German (gray line—affected sib pairs) and African American (black line—all sib pairs; dashed line—affected sib pairs) genome scans for sarcoidosis genes. Results from regions with significant signals are shown in (A) chromosome 1, (B) chromosome 3, (C) chromosome 5, (D) chromosome 6,

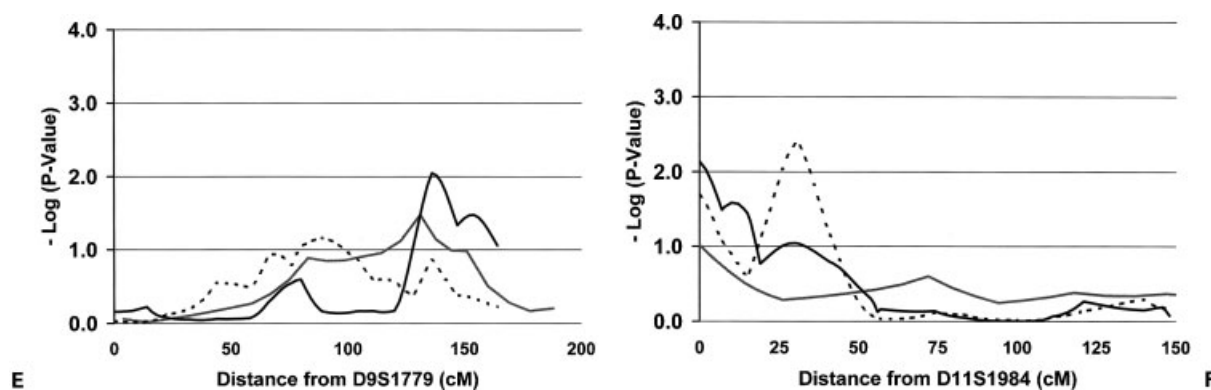


Figure 3 (Con't) (E) chromosome 9, and (F) chromosome 11.

linkage to the 5p15–13 locus, whereas putative susceptibility loci at 1p22, 3p21–14, 11p15, and 17q21 only showed linkage in the other subpopulation. Several of the linkage signals were consistent across both populations, but they were more the exception than the rule.

Comparison of Caucasian and African-American Sarcoidosis Genome Scans

Fig. 3A–F depicts the significant chromosomal linkages from the two sarcoidosis linkage studies done in German Caucasian and African-American populations. The linkage results at chromosomes 3p, 5, and 6p and subsequent additional studies of these regions are illustrative of the complex genetics underlying susceptibility to this disease. In the original SAGA genome scan, no markers on chromosome 3 met the significance level cutoff value of .05. However, in subanalyses of families with three or more sibs affected, the D3S1285 marker on chromosome 3p14 had a p value of .03 and more significantly mapped to the same chromosomal region that had the second strongest signal ($p = .009$) in the original German genome scan.⁵⁵ These findings motivated the fine mapping of this region in the SAGA sample (despite its lack of significance in the overall scan), and as described earlier, it appears that a putative locus in the 3p14 region may act in concert with another putative sarcoidosis locus on chromosome 5p15–13. Furthermore, the fact that the 3p signal appeared in German Caucasians and only in a small subpopulation of the larger SAGA sample⁶⁰ would suggest that this locus may be Caucasian in origin and only discernible in an admixed African American population where analyses of genetically homogeneous subsets that have a high percent of Caucasian genes are possible. Investigation of one of the strongest candidates in the 3p region based on gene function, *CCR2*, has yielded mixed results. Two studies have demonstrated an association with disease susceptibility,^{9,61} and one study found an association with Löfgren's syndrome.⁶² However, the German sample that had a linkage signal at chromosome 3p failed to show an association with

CCR2 in both sarcoidosis patients and the Löfgren's subset within the patient sample.⁶³

Population stratification has also shed new light on the initial chromosome 5 linkages to sarcoidosis in African Americans reported by SAGA.⁵⁸ Thompson et al⁶⁰ showed that the 5p15–13 linkage was unique to one subset of the population, whereas the 5q11 linkage appeared to be consistent across both populations (Fig. 4). Even more perplexing than the putative genetic linkages to sarcoidosis on chromosomes 3p and 5 are the lack of any positive linkages with the major histocompatibility complex (MHC) region in SAGA despite the strong linkage signal to that region observed in German Caucasians. Analyses of the *HLA-DRB1* alleles in ACCESS demonstrated the allelic heterogeneity by race at the locus,⁴ translating in part to *DRB1* alleles that have associations with sarcoidosis in one race, but not the other (e.g., *HLA-DPB1*0101* in blacks, but not whites; *HLA-DRB1*1501* in whites, but not blacks).

Allelic heterogeneity by race at the *BTNL2* locus further illustrates the complexity of MHC associations with sarcoidosis between ethnic groups. In a follow-up to the *BTNL2* exon 5 sarcoidosis allele reported in German Caucasians at the rs2076530 SNP, we genotyped rs2076530 and other nearby *BTNL2* SNPs in independent Caucasian and African-American samples.¹⁰ Four SNPs (G16043A, A16071G, C16113T, and T16165C) captured > 95% of the haplotype variation within this region. The most frequent haplotype, G-A-C-T (haplotype 1), conferred 1.56-fold increased risk for sarcoidosis ($p = .00004$) in whites. Haplotype 1 was not associated with sarcoidosis in either the case-control or the family-based African-American samples. In the sample of whites, the A allele at position 16071 conferred an odds ratio for heterozygotes (AG vs GG) of 1.70 and an odds ratio for homozygotes (AA vs GG) of 2.63. This result is similar to that observed in the German population.⁸ Haplotype 2 (G-G-T-T), the next most frequent haplotype in African Americans, was not associated with sarcoidosis in either the family or the case control sample, but was underrepresented in

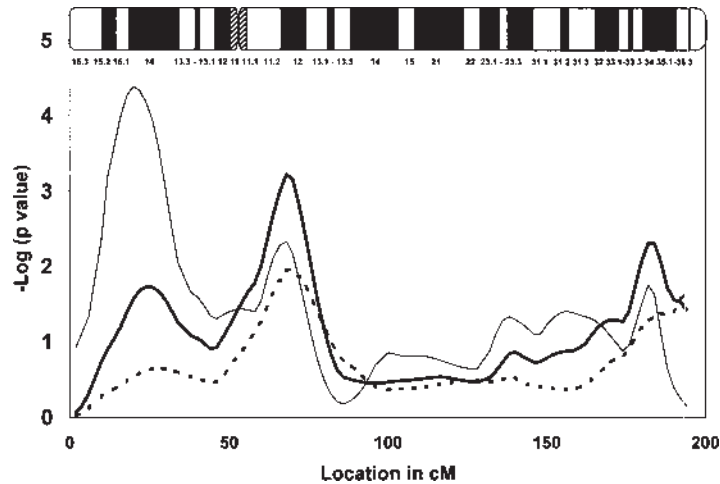


Figure 4 Stratified linkage results on chromosome 5 in the Sarcoidosis Genetic Analysis (SAGA) study population. Linkage results of the original scan (bold line) compared with results from stratified populations 1 (dashed line) and 2 (solid line). Adapted from Thompson et al.⁶⁰

white cases (OR = 0.60; $p = .006$). The second most frequent haplotype in the white population, haplotype 3 (A-G-C-T), was underrepresented in white sarcoidosis cases (OR = 0.74; $p = .02$), in African-American families [transmission distortion (TD) 0.76; $p = .02$], but not in the African-American case-control sample. Haplotype 4 (A-G-C-C), was underrepresented in cases across all three samples, but only in the African-American case-control sample did this haplotype show a significant negative association with sarcoidosis (OR = 0.44, $p = .001$). Overall, the haplotype distribution across *BTNL2* exon/intron 5 was significantly associated with sarcoidosis in all three populations, but the association in African-Americans was more modest than compared with whites. Of note, the association of the *BTNL2* rs2076530 A allele has recently been replicated in a third independent European white population.⁶⁴

FUTURE DIRECTIONS IN SARCOIDOSIS EPIDEMIOLOGY

A Conceptual Model of Sarcoidosis Etiology

If sarcoidosis epidemiology is to move forward, then new methodological approaches are needed. In addition, emerging technological advances that allow detection of hidden molecular signals from antigenic footprints and genomewide analyses of genes more often inherited and/or expressed in sarcoidosis should yield new insights. In all epidemiological studies, investigators need to adequately address the issue of defining who is a case and case heterogeneity. Defining etiologically homogeneous subsets of disease and study designs that are not subject to some of the inherent biases in case-control studies will offer the most in this regard. Sarcoidosis variation by ethnicity and sex should also be used as a

means to tease out etiologic clues. Fig. 5 depicts a general conceptual model of sarcoidosis associations that investigators should bear in mind when designing epidemiological studies. Some examples of methodological approaches that have been or could be fruitful in the study of sarcoidosis epidemiology follow.

Phenotype-Based Analyses

Sarcoidosis researchers have long been aware of the importance of phenotype when searching for genetic risk factors. Most of the initial association studies of HLA class I molecules and sarcoidosis subdivided cases into more homogeneous subsets in an effort to uncover genetic associations that may be masked when examining a clinically heterogeneous group of sarcoidosis cases. For instance, Brewerton et al reported associations between *HLA-B8* and sarcoidosis cases with acute arthritis and *HLA-A1* and sarcoidosis with uveitis.⁶⁵ Six years later, Hedfors and Linstrom extended the *HLA-B8*/sarcoidosis arthritis finding to also include an association with HLA class II *DR3* molecule.⁶⁶

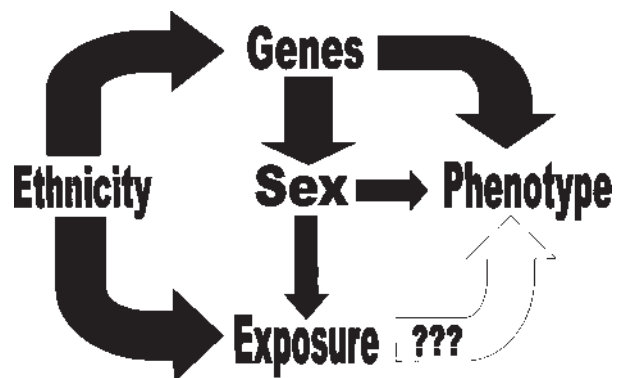


Figure 5 Conceptual model of interrelationship of factors that lead to the sarcoidosis phenotype.

In an effort to simplify the number and types of sarcoidosis phenotypes, sarcoidosis has commonly been divided into “chronic” versus “acute” disease. Although this is likely an oversimplification, genetic factors specific to these two extremes of the sarcoidosis phenotypic spectrum have been identified. For instance, the class II *HLA-DRB1*0301/DQB1*0201* haplotype has been repeatedly found associated with Löfgren’s syndrome,^{67–70} an acute form of sarcoidosis characterized by erythema nodosum, Scadding stage I chest x-ray, arthritis, and a good prognosis.⁷¹ More generally, this haplotype is also associated with a good disease prognosis.^{70,72,73} Genetic associations with sarcoidosis phenotypes reported in the literature are shown in Table 2.^{7,12,26,61,62,67,68,70,74–92} In considering phenotype, racial heterogeneity of disease expression is of utmost importance. In ACCESS, race-specific HLA associations were observed with different organ involvement phenotypes,⁴ and in a multivariate model organic dust and wood burning exposures had race-specific associations with systemic disease.⁹³

Some have posited that sarcoidosis is one clinical manifestation of a constellation of immune-related inflammatory granulomatous disorders that lie on a continuous disease spectrum.⁹⁴ Clearly, inflammatory diseases likely share genetic risk factors,⁹⁵ which suggests that searching for genes linked to specific clinical manifestations of inflammatory disease, rather than the collection of signs and symptoms that constitute a clinically defined inflammatory disorder, might portend a higher level of success. We performed such an analysis with the affected relative pair portion of the SAGA sample to search for loci linked to different pulmonary and organ involvement phenotypes.⁹⁶ In general, higher LOD (log of the odds) scores were attained for covariates that modeled clustered organ system involvement rather than individual organ systems. Interestingly, none of the significant sarcoidosis phenotype linkages overlapped with the regions previously found linked to sarcoidosis susceptibility, which suggested that, in African Americans, genes that influence specific manifestations of disease are likely to be largely separate from those which underlie disease susceptibility.

Gene–Environment and Gene–Gene Interactions

Future genetic epidemiological investigations of sarcoidosis must consider an analytic approach that includes evaluation of gene–environment and gene–gene interactions. A small but growing number of studies are beginning to suggest that such an analytic approach can yield a host of new and important advances in this disease. In our investigation of the *HLA-DQB1* gene and sarcoidosis in African American families, we found evidence for gene–environment interactions between the

*HLA-DQB1*0201* allele and several different exposures, including vegetable dust, high humidity, and water damage.⁷ Interestingly, although the *HLA-DQB1*0201* allele had an overall inverse association with disease susceptibility, in the presence of vegetable dust and water damage an increased risk for disease was observed.

Because it is clear that multiple genes are involved in sarcoidosis pathogenesis, in addition to exploring the front end of the pathogenic cascade, it can also be fruitful to consider events further along the pathogenic pathway that involve interactions between multiple gene products. As mentioned earlier, a fine mapping linkage analysis of the SAGA sample suggested that putative susceptibility loci at chromosomes 3p14–11 and 5p15.2 may be interdependent.⁵⁹ In terms of specific genes that are known to have an important effect on sarcoidosis, Grunewald and coworkers were able to tease apart sarcoidosis risk effects for both HLA class I and II alleles by showing that an increased risk for persistent disease associated with the alleles *HLA-A*03* and *-B*07* required the presence of the HLA class II *DRB1*15* allele.⁷² The novel aspect of this report was that it strongly supported the notion that class I and II HLA molecules may play complementary roles in the promotion of the sarcoidosis disease process. Gene–gene interactions within the HLA region may be even more complex than interactions between class I and II molecules. In the initial report that identified *BTNL2* as a novel sarcoidosis susceptibility gene, there was some suggestion of a possible interaction between the *BTNL2* risk allele at rs2076530 and *HLA-DRB1* risk alleles.⁸ This potential interaction was investigated more in depth by Rybicki et al using the ACCESS Caucasian and African-American case-control samples.¹⁰ Although there did not appear to be any interactions between risk alleles for *BTNL2* and HLA class II in whites, in blacks there was a strong suggestion of a negative interaction between *BTNL2* and HLA class II risk alleles. If such a complex antagonistic effect between antigen recognition and costimulatory molecules coded by HLA class II exists in African-Americans, this could explain the inability to replicate the *BTNL2* rs2076530 association in blacks as well as the lack of a linkage signal at chromosome 6p in the African-American SAGA sample.

Case-Only Study Designs

In terms of environmental exposures, although questionnaire data will continue to be useful, more detail on exposure dose and timing is needed to maximize the use of exposure data. The failure to identify a disease-causing antigen in sarcoidosis strongly suggests that the exposure to the inciting agent(s) of the sarcoidosis immunologic cascade may be very short term. Evidence exists that at least some sarcoidosis cases cluster in time and space, suggesting an acute point source exposure.^{44,45,97–99} In

Table 2 Genetic Associations with Sarcoidosis Clinical Phenotypes

Gene	Allele(s)	Phenotype(s)	Population(s)	Reference(s)	Comment
<i>ACE</i>	D	Advanced radiographic stage, chronic disease	Swedes, Finns	74,75	Equivocal results for this polymorphism in the literature
	I	Advanced radiographic stage, bronchial responsiveness	African-Americans; Japanese	76,77	D allele associated with susceptibility in African Americans
<i>ACE2</i>	haplotype 4	Parenchymal involvement	Dutch	26	Association in males only
<i>CC10</i>	+38A	chronic disease	Japanese	78,79	Association not replicated in other Japanese and Dutch populations
<i>CCR2</i>	haplotype 2	Löfgren's	Dutch	62	Remained significant after adjustment for <i>DRB1/DQB1</i> associated alleles and female sex
<i>CCR5</i>	delta32	Patients requiring corticosteroids	Czech	61	Allele also associated with susceptibility
<i>CTLA4</i>	-318C and +49G	Ocular	Japanese	80	-49G allele also increased in patients with three or more organs affected
<i>DQB1</i>	*0201	Löfgren's	UK and Dutch	70	Allele also associated with erythema nodosum
	*0601	Cardiac	Japanese	81	Allele may be part of a larger MHC haplotype
	*0602	Radiographic progression	African-American, UK, Dutch	7,70	Allele also associated with susceptibility
<i>DRB1</i>	*03	Löfgren's	Czech, Swedes	67,68	Part of a larger MHC haplotype including TNFA and LTA
	*1501	Chronic disease	Dutch, Swedes, Asian Indians	68,82,83	Gene in linkage disequilibrium with <i>DQB1</i> *0602; possible interaction with class I <i>HLA-B</i>
<i>HSP70-hom</i>	+2437C	Löfgren's	Polish	12	Gene in linkage disequilibrium with <i>DRB1</i> *03
<i>IL6</i>	-174C	Stage IV disease	UK, Dutch	84	Allele not associated with lung function
<i>NFKBIA</i>	-826T	Radiographic progression	UK, Dutch	85	Allele not associated with lung function
<i>PTGS2</i>	-765C	Chronic disease	British	86	Gene product is reduced in sarcoidosis lung
<i>TLR</i>	299Gly and 399Ile	Chronic disease	German	87	Allele not associated with susceptibility
<i>TNF-a</i>	-307A	Löfgren's, erythema nodosum	UK, Dutch German	88,89	Other promoter allele, -857T, associated with susceptibility
	A2	Cardiac, Löfgren's	Japanese, German	90,91	Japanese study only included 26 cardiac sarcoidosis patients
<i>TNF-b</i>	*1	Chronic disease	Japanese	92	Patient study population did not require corticosteroids

HLA, human leukocyte antigen; LTA, lymphotoxin alpha; MHC, major histocompatibility complex; TNFA, tumor necrosis factor alpha.

such instances more focus is needed on exposures occurring during the time *directly before* disease onset rather than lifelong exposures. Recent acute exposures are generally easier to recall but are still subject to bias in

comparison with the experience of healthy controls. Study designs that use only case exposure histories, such as the case-crossover design,¹⁰⁰ may be one solution to the recall bias problem inherent in the case-control

design, particularly in situations where the exposure can be quantified on a person time basis.

Other case-only study designs can be used to study genetic variants and combinations of genetic and environmental risk factors. A reported family history of disease implies that shared susceptibility genes among family members will increase disease risk. A study design that requires only case genotype and family history data is the kin-cohort design.¹⁰¹⁻¹⁰³ Essentially, a cohort is formed that is defined by a certain class of case relatives, usually first-degree relatives, in which the familial risk of a case with a certain genetic variant is compared with the familial risk of a case without the genetic variant.^{101,103} Although initially advocated as a tool to confirm case-control results, there is no reason why the kin-cohort approach cannot be used as the primary analytic method. Its main disadvantage is the possible lack of statistical power due to the attenuation of risk from the mixing of carriers and noncarriers in the relatives of the index case. However, in certain situations the kin-cohort approach can be more powerful than a case-control study.¹⁰² In addition, there is no genetic confounding due to population stratification that is an inherent risk in the case-control design. An application of the kin-cohort approach to an African-American sarcoidosis case sample showed *IL1A* may be a promising candidate gene for sarcoidosis susceptibility.¹⁰⁴

Another study design that uses only cases to investigate interactions between risk factors rather than risks associated with individual risk factors is aptly called the "case-only" design.^{105,106} This design makes use of the fact that the multiplicative interaction of two risk factors is based solely on their distribution in cases when the two factors are uncorrelated in controls. Apart from the decreased logistical burden of enrolling and characterizing controls, the case-only design further benefits from the increased statistical efficiency afforded by a lower variance of the interaction odds ratio owing to the reduction of experimental cells from eight to four (i.e., the control cells are no longer relevant). The main limitation of the case-only analytic approach is that it cannot evaluate main effects, and others have noted that lack of independence of risk factors in controls can invalidate this approach.^{107,108} For sarcoidosis, the case-only design may be of limited interest to researchers because definitive risk estimates associated with individual factors are largely unknown and often of interest.

Mapping of Sarcoidosis Genes Associated with Ancestry

A higher incidence of sarcoidosis in African-Americans compared with Caucasians has long been noted,^{109,110} with a three- to fourfold increased incidence of sarcoidosis in African-Americans compared with Caucasians.²⁰ Other studies outside of the United States

comparing sarcoidosis incidence in populations of west African ancestry with Caucasians further support a role of African genes in sarcoidosis susceptibility. In South Africa, Benatar reported a sarcoidosis incidence of 23.2/100 000 in blacks, 11.6/100 000 in coloreds, and 3.7/100 000 in whites.¹¹¹ In Great Britain, Edmondstone and Wilson¹¹² reported an annual sarcoidosis incidence of 19.8 per 10⁵ in blacks and 1.5 per 10⁵ in Caucasians. Overall, these studies across three continents show that sarcoidosis is consistently more common in people of West African ancestry when compared with people of European ancestry. The association between increased sarcoidosis risk and African ancestry and the admixed nature of the African-American population¹¹³ suggests that the newly emerging technique of genetic admixture mapping¹¹⁴ might be an important gene-finding strategy for sarcoidosis.

Mapping by admixture linkage disequilibrium (MALD) involves screening the genome of individuals of mixed ancestry, who have a disease or trait of interest, for chromosomal regions that have a greater percentage of alleles from the parental population with the higher disease risk. The theory behind MALD is that chromosomal segments of affected individuals containing a significantly higher than average composition of alleles from the high-risk parental population are more likely to also harbor a disease-gene allele. MALD is potentially suitable for localizing disease-gene polymorphisms in admixed populations that derive from ancestral populations with large allele-frequency differences that formed in the past few hundred years through admixture. Several criteria exist for MALD to be feasible¹¹⁵: (1) MALD-based identification of disease genes requires a measurable difference in disease frequency (also inferring difference in disease-causing alleles) between the parental populations; (2) admixture ideally needs to be at least two generations old to reduce the initial disequilibrium across chromosomes and between unlinked loci, whereas linkage disequilibrium within chromosomes remains strong¹¹⁶; (3) a set of markers that specifically differentiate chromosomes derived from the parental populations is needed. As described earlier, admixture mapping for sarcoidosis genes in African-Americans satisfies the first of these three criteria. In addition, Caucasian admixture proportions in African-Americans¹¹³ and resulting admixture linkage disequilibrium¹¹⁷ suggests the second criterion is satisfied. Finally, emerging genetic maps of ancestry informative markers^{118,119} provide the necessary genetic tools to successfully undertake sarcoidosis admixture mapping studies.

SUMMARY

Recent findings from large collaborative studies of sarcoidosis and new developments in molecular technologies should move the epidemiology of sarcoidosis

dramatically forward in the coming years. Advances that are on the near horizon in the genetics of sarcoidosis are poised to fill the arsenal of the 21st century physician who will be trained to practice genomic medicine.¹²⁰ The “genetic epidemiology” of sarcoidosis will benefit from continued efforts to find combinations of genetic and environmental factors in the disease pathogenesis as well as efforts to dissect the etiologies of homogeneous disease subsets in which a limited number of risk factors may play a more prominent role. Usage of case-only methods that circumvent problems with identifying and defining proper control groups can provide a fresh look at an old problem. In summary, the majority of epidemiologic advances in sarcoidosis are likely ahead of us, and it behooves those investigators in this field to use all the research tools now available to advance sarcoidosis epidemiology into a promising future in which it will no longer be a disease of *unknown causes*.

REFERENCES

- 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, February 1999. *Am J Respir Crit Care Med* 1999;160:736–755
- ACCESS Group. Design of a case control etiologic study of sarcoidosis (ACCESS). ACCESS Research Group. *J Clin Epidemiol* 1999;52:1173–1186
- Semenzato G. ACCESS: a case control etiologic study of sarcoidosis. *Sarcoidosis Vasc Diffuse Lung Dis* 2005;22:83–86
- Rossmann MD, Thompson B, Frederick M, et al. HLA-DRB1*1101: a significant risk factor for sarcoidosis in blacks and whites. *Am J Hum Genet* 2003;73:720–735
- Rybicki BA, Iannuzzi MC, Frederick MM, et al. Familial aggregation of sarcoidosis: a case-control etiologic study of sarcoidosis (ACCESS). *Am J Respir Crit Care Med* 2001;164:2085–2091
- Newman LS, Rose CS, Bresnitz EA, et al. A case control etiologic study of sarcoidosis: environmental and occupational risk factors. *Am J Respir Crit Care Med* 2004;170:1324–1330
- Iannuzzi MC, Maliarik MJ, Poisson LM, Rybicki BA. Sarcoidosis susceptibility and resistance HLA-DQB1 alleles in African Americans. *Am J Respir Crit Care Med* 2003;167:1225–1231
- Valentonyte R, Hampe J, Huse K, et al. Sarcoidosis is associated with a truncating splice site mutation in BTNL2. *Nat Genet* 2005;37:357–364
- Hizawa N, Yamaguchi E, Furuya K, Jinushi E, Ito A, Kawakami Y. The role of the C–C chemokine receptor 2 gene polymorphism V64I (CCR2–64I) in sarcoidosis in a Japanese population. *Am J Respir Crit Care Med* 1999;159:2021–2023
- Rybicki BA, Walewski JL, Maliarik MJ, Kian H, Iannuzzi MC. The BTNL2 gene and sarcoidosis susceptibility in African Americans and whites. *Am J Hum Genet* 2005;77:491–499
- Pandey JP, Frederick M. TNF-alpha, IL1-beta, and immunoglobulin (GM and KM) gene polymorphisms in sarcoidosis. *Hum Immunol* 2002;63:485–491
- Bogunia-Kubik K, Koscińska K, Suchnicki K, Lange A. HSP70-hom gene single nucleotide (+2763 G/A and +2437 C/T) polymorphisms in sarcoidosis. *Int J Immunogenet* 2006;33:135–140
- Dubaniewicz A, Jamieson SE, Dubaniewicz-Wybieralska M, Fakiola M, Nancy ME, Blackwell JM. Association between SLC11A1 (formerly NRAMP1) and the risk of sarcoidosis in Poland. *Eur J Hum Genet* 2005;13:829–834
- Foley PJ, Lympny PA, Puscinska E, Zielinski J, Welsh KI, du Bois RM. Analysis of MHC encoded antigen-processing genes TAP1 and TAP2 polymorphisms in sarcoidosis. *Am J Respir Crit Care Med* 1999;160:1009–1014
- Gazouli M, Koundourakis A, Ikonomopoulos J, et al. CARD15/NOD2, CD14, and Toll-like receptor 4 gene polymorphisms in Greek patients with sarcoidosis. *Sarcoidosis Vasc Diffuse Lung Dis* 2006;23:23–29
- Zorzetto M, Bombieri C, Ferrarotti I, et al. Complement receptor 1 gene polymorphisms in sarcoidosis. *Am J Respir Cell Mol Biol* 2002;27:17–23
- Takada T, Suzuki E, Morohashi K, Gejyo F. Association of single nucleotide polymorphisms in the IL-18 gene with sarcoidosis in a Japanese population. *Tissue Antigens* 2002;60:36–42
- Akahoshi M, Ishihara M, Remus N, et al. Association between IFNA genotype and the risk of sarcoidosis. *Hum Genet* 2004;114:503–509
- Morohashi K, Takada T, Omori K, Suzuki E, Gejyo F. Vascular endothelial growth factor gene polymorphisms in Japanese patients with sarcoidosis. *Chest* 2003;123:1520–1526
- Rybicki BA, Major M, Popovich JJ, Maliarik MJ, Iannuzzi MC. Racial differences in sarcoidosis incidence: a 5-year study in a health maintenance organization. *Am J Epidemiol* 1997;145:234–241
- Yamaguchi M, Hosoda Y, Sasaki R, Aoki K. Epidemiological study on sarcoidosis in Japan: recent trends in incidence and prevalence rates and changes in epidemiological features. *Sarcoidosis* 1989;6:138–146
- Hillerdal G, Nou E, Osterman K, Schmekel B. Sarcoidosis: epidemiology and prognosis: a 15-year European study. *Am Rev Respir Dis* 1984;130:29–32
- Henke CE, Henke G, Elveback LR, Beard CM, Ballard DJ, Kurland LT. The epidemiology of sarcoidosis in Rochester, Minnesota: a population-based study of incidence and survival. *Am J Epidemiol* 1986;123:840–845
- Byg KE, Milman N, Hansen S. Sarcoidosis in Denmark 1980–1994: a registry-based incidence study comprising 5536 patients. *Sarcoidosis Vasc Diffuse Lung Dis* 2003;20:46–52
- Furuya K, Yamaguchi E, Itoh A, et al. Deletion polymorphism in the angiotensin I converting enzyme (ACE) gene as a genetic risk factor for sarcoidosis. *Thorax* 1996;51:777–780
- Kruit A, Ruven HJ, Grutters JC, Van Den Bosch JM. Angiotensin-converting enzyme 2 (ACE2) haplotypes are associated with pulmonary disease phenotypes in sarcoidosis patients. *Sarcoidosis Vasc Diffuse Lung Dis* 2005;22:195–203
- Rybicki BA, Kirkey KL, Major M, et al. Familial risk ratio of sarcoidosis in African-American sibs and parents. *Am J Epidemiol* 2001;153:188–193

28. Baughman RP, Teirstein AS, Judson MA, et al. Clinical characteristics of patients in a case control study of sarcoidosis. *Am J Respir Crit Care Med* 2001;164:1885–1889
29. Judson MA, Baughman RP, Thompson BW, et al. Two year prognosis of sarcoidosis: the ACCESS experience. *Sarcoidosis Vasc Diffuse Lung Dis* 2003;20:204–211
30. Terris M, Chavis AD. An epidemiologic study of sarcoidosis. *Am Rev Respir Dis* 1966;94:50–55
31. Buck AA, Sartwell PE. Epidemiologic investigations of sarcoidosis, II: Skin sensitivity and environmental factors. *Am J Hyg* 1961;74:152–173
32. Cummings MM, Hudgins PC. Chemical constituents of pine pollen and their possible relationship to sarcoidosis. *Am J Med Sci* 1958;236:311–317
33. Burdick KH. Cutaneous reaction to pine pollen in sarcoidosis: study of forty cases. *J Invest Dermatol* 1964;42:449–451
34. Gentry JT, Nitowsky HM, Michael MJr. Studies on the epidemiology of sarcoidosis in the United States: the relationship to soil areas and to urban-rural residence. *J Clin Invest* 1955;34:1839–1856
35. Comstock GW, Keltz H, Sencer J. Clay eating and sarcoidosis: a controlled study in the state of Georgia. *Am Rev Respir Dis* 1961;84:130–134
36. Hurley HJ, Close HP, English RS. Soil extracts as antigens in sarcoidosis. *Am Rev Respir Dis* 1962;86:100–102
37. Kajdasz DK, Lackland DT, Mohr LC, Judson MA. A current assessment of rurally linked exposures as potential risk factors for sarcoidosis. *Ann Epidemiol* 2001;11:111–117
38. Drent M, Bomans PH, Van Suylen RJ, Lamers RJ, Bast A, Wouters EF. Association of man-made mineral fibre exposure and sarcoidlike granulomas. *Respir Med* 2000;94:815–820
39. Rafnsson V, Ingimarsson O, Hjalmarsen I, Gunnarsdottir H. Association between exposure to crystalline silica and risk of sarcoidosis. *Occup Environ Med* 1998;55:657–660
40. Rybicki BA, Amend KL, Maliarik MJ, Iannuzzi MC. Photocopier exposure and risk of sarcoidosis in African-American sibs. *Sarcoidosis Vasc Diffuse Lung Dis* 2004;21:49–55
41. Kucera GP, Rybicki BA, Kirkey KL, et al. Occupational risk factors for sarcoidosis in African-American siblings. *Chest* 2003;123:1527–1535
42. Gorham ED, Garland CF, Garland FC, Kaiser K, Travis WD, Centeno JA. Trends and occupational associations in incidence of hospitalized pulmonary sarcoidosis and other lung diseases in navy personnel: a 27-year historical prospective study, 1975–2001. *Chest* 2004;126:1431–1438
43. Barnard J, Rose C, Newman L, et al. Job and industry classifications associated with sarcoidosis in a case-control etiologic study of sarcoidosis (ACCESS). *J Occup Environ Med* 2005;47:226–234
44. Prezant DJ, Dhala A, Goldstein A, et al. The incidence, prevalence, and severity of sarcoidosis in New York City firefighters. *Chest* 1999;116:1183–1193
45. Kern DG, Neill MA, Wrenn DS, Varone JC. Investigation of a unique time-space cluster of sarcoidosis in firefighters. *Am Rev Respir Dis* 1993;148:974–980
46. Falkinham JO III. Mycobacterial aerosols and respiratory disease. *Emerg Infect Dis* 2003;9:763–767
47. Fite E, Fernandez-Figueras MT, Prats R, Vaquero M, Morera J. High prevalence of *Mycobacterium tuberculosis* DNA in biopsies from sarcoidosis patients from Catalonia, Spain. *Respiration* 2006;73:20–26
48. Vokurka M, Lecossier D, du Bois RM, et al. Absence of DNA from mycobacteria of the *M. tuberculosis* complex in sarcoidosis. *Am J Respir Crit Care Med* 1997;156:1000–1003
49. Brown ST, Brett I, Almenoff PL, Lesser M, Terrin M, Teirstein AS. Recovery of cell wall-deficient organisms from blood does not distinguish between patients with sarcoidosis and control subjects. *Chest* 2003;123:413–417
50. du Bois RM, Goh N, McGrath D, Cullinan P. Is there a role for microorganisms in the pathogenesis of sarcoidosis? *J Intern Med* 2003;253:4–17
51. Drake WP, Newman LS. Mycobacterial antigens may be important in sarcoidosis pathogenesis. *Curr Opin Pulm Med* 2006;12:359–363
52. Buck AA, McKusick VA. Epidemiologic investigations of sarcoidosis, III: Serum proteins, syphilis, association with tuberculosis: familial aggregation. *Am J Hyg* 1961;74:174–188
53. Headings VE, Weston D, Young RCJ, Hackney RLJ. Familial sarcoidosis with multiple occurrences in eleven families: a possible mechanism of inheritance. *Ann NY Acad Sci* 1976;278:377–385
54. McGrath DS, Daniil Z, Foley P, et al. Epidemiology of familial sarcoidosis in the UK. *Thorax* 2000;55:751–754
55. Schurmann M, Reichel P, Muller-Myhsok B, Schlaak M, Müller-Quernheim J, Schwinger E. Results from a genome-wide search for predisposing genes in sarcoidosis. *Am J Respir Crit Care Med* 2001;164:840–846
56. Rybicki BA, Hirst K, Iyengar SK, et al. A sarcoidosis genetic linkage consortium: the sarcoidosis genetic analysis (SAGA) study. *Sarcoidosis Vasc Diffuse Lung Dis* 2005;22:115–122
57. Haseman JK, Elston RC. The investigation of linkage between a quantitative trait and marker locus. *Behav Genet* 1972;2:3–19
58. Iannuzzi MC, Iyengar SK, Gray-McGuire C, et al. Genome-wide search for sarcoidosis susceptibility genes in African Americans. *Genes Immun* 2005;6:509–518
59. Gray-McGuire C, Sinha R, Iyengar S, et al. Genetic characterization and fine mapping of susceptibility loci for sarcoidosis in African Americans on chromosome 5. *Hum Genet*. In press
60. Thompson CL, Rybicki BA, Iannuzzi MC, et al. Reduction of sample heterogeneity through use of population substructure: an example from a population of African American families with sarcoidosis. *Am J Hum Genet* 2006;79:606–613
61. Petrek M, Drabek J, Kolek V, et al. CC chemokine receptor gene polymorphisms in Czech patients with pulmonary sarcoidosis. *Am J Respir Crit Care Med* 2000;162:1000–1003
62. Spagnolo P, Renzoni EA, Wells AU, et al. C–C chemokine receptor 2 and sarcoidosis: association with Löfgren's syndrome. *Am J Respir Crit Care Med* 2003;168:1162–1166
63. Valentonyte R, Hampe J, Croucher PJ, et al. Study of C–C chemokine receptor 2 alleles in sarcoidosis, with emphasis on family based analysis. *Am J Respir Crit Care Med* 2005;171:1136–1141
64. Li Y, Wollnik B, Pabst S, et al. BTNL2 gene variant and sarcoidosis. *Thorax* 2006;61:273–274
65. Brewerton DA, Cockburn C, James DC, James DG, Neville E. HLA antigens in sarcoidosis. *Clin Exp Immunol* 1977;27:227–229

66. Hedfors E, Lindstrom F. HLA-B8/DR3 in sarcoidosis: correlation to acute onset disease with arthritis. *Tissue Antigens* 1983;22:200–203
67. Mrzacek F, Holla LI, Hutyrova B, et al. Association of tumour necrosis factor-alpha, lymphotoxin-alpha and HLA-DRB1 gene polymorphisms with Löfgren's syndrome in Czech patients with sarcoidosis. *Tissue Antigens* 2005;65:163–171
68. Berlin M, Fogdell-Hahn A, Olerup O, Eklund A, Grunewald J. HLA-DR predicts the prognosis in Scandinavian patients with pulmonary sarcoidosis. *Am J Respir Crit Care Med* 1997;156:1601–1605
69. Swider C, Schnittger L, Bogunia-Kubik K, et al. TNF-alpha and HLA-DR genotyping as potential prognostic markers in pulmonary sarcoidosis. *Eur Cytokine Netw* 1999;10:143–146
70. Sato H, Grutters JC, Pantelidis P, et al. HLA-DQB1*0201: a marker for good prognosis in British and Dutch patients with sarcoidosis. *Am J Respir Cell Mol Biol* 2002;27:406–412
71. Löfgren S. Erythema nodosum: studies on etiology and pathogenesis in 185 adult cases. *Acta Med Scand* 1946;124:1–197
72. 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;169:696–702
73. Martinetti M, Tinelli C, Kolek V, et al. "The sarcoidosis map": a joint survey of clinical and immunogenetic findings in two European countries. *Am J Respir Crit Care Med* 1995;152:557–564
74. Papadopoulos KI, Melander O, Orho-Melander M, Groop LC, Carlsson M, Hallengren B. Angiotensin converting enzyme (ACE) gene polymorphism in sarcoidosis in relation to associated autoimmune diseases. *J Intern Med* 2000;247:71–77
75. 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;13:723–726
76. Maliarik MJ, Rybicki BA, Malvitz E, et al. Angiotensin-converting enzyme gene polymorphism and risk of sarcoidosis. *Am J Respir Crit Care Med* 1998;158:1566–1570
77. Niimi T, Tomita H, Sato S, et al. Bronchial responsiveness and angiotensin-converting enzyme gene polymorphism in sarcoidosis patients. *Chest* 1998;114:495–499
78. Ohchi T, Shijubo N, Kawabata I, et al. Polymorphism of Clara cell 10-kD protein gene of sarcoidosis. *Am J Respir Crit Care Med* 2004;169:180–186
79. Janssen R, Sato H, Grutters JC, et al. The Clara cell10 adenine38guanine polymorphism and sarcoidosis susceptibility in Dutch and Japanese subjects. *Am J Respir Crit Care Med* 2004;170:1185–1187
80. Hattori N, Niimi T, Sato S, et al. Cytotoxic T-lymphocyte antigen 4 gene polymorphisms in sarcoidosis patients. *Sarcoidosis Vasc Diffuse Lung Dis* 2005;22:27–32
81. Naruse TK, Matsuzawa Y, Ota M, et al. HLA-DQB1*0601 is primarily associated with the susceptibility to cardiac sarcoidosis. *Tissue Antigens* 2000;56:52–57
82. Voorter CE, Drent M, van den Berg-Loonen EM. Severe pulmonary sarcoidosis is strongly associated with the haplotype HLA-DQB1*0602-DRB1*150101. *Hum Immunol* 2005;66:826–835
83. Sharma SK, Balamurugan A, Pandey RM, Saha PK, Mehra NK. Human leukocyte antigen-DR alleles influence the clinical course of pulmonary sarcoidosis in Asian Indians. *Am J Respir Cell Mol Biol* 2003;29:225–231
84. Grutters JC, Sato H, Pantelidis P, et al. Analysis of IL6 and IL1A gene polymorphisms in UK and Dutch patients with sarcoidosis. *Sarcoidosis Vasc Diffuse Lung Dis* 2003;20:20–27
85. Abdallah A, Sato H, Grutters JC, et al. Inhibitor kappa B-alpha (IkappaB-alpha) promoter polymorphisms in UK and Dutch sarcoidosis. *Genes Immun* 2003;4:450–454
86. Hill MR, Papafili A, Booth H, et al. Functional prostaglandin-endoperoxide synthase 2 polymorphism predicts poor outcome in sarcoidosis. *Am J Respir Crit Care Med* 2006;174:915–922
87. Pabst S, Baumgarten G, Stremmel A, et al. Toll-like receptor (TLR) 4 polymorphisms are associated with a chronic course of sarcoidosis. *Clin Exp Immunol* 2006;143:420–426
88. Grutters JC, Sato H, Pantelidis P, et al. Increased frequency of the uncommon tumor necrosis factor -857T allele in British and Dutch patients with sarcoidosis. *Am J Respir Crit Care Med* 2002;165:1119–1124
89. Labunski S, Posern G, Ludwig S, Kundt G, Brocker EB, Kunz M. Tumour necrosis factor-alpha promoter polymorphism in erythema nodosum. *Acta Derm Venereol* 2001;81:18–21
90. Takashige N, Naruse TK, Matsumori A, et al. Genetic polymorphisms at the tumour necrosis factor loci (TNFA and TNFB) in cardiac sarcoidosis. *Tissue Antigens* 1999;54:191–193
91. Seitzer U, Swider C, Stuber F, et al. Tumour necrosis factor alpha promoter gene polymorphism in sarcoidosis. *Cytokine* 1997;9:787–790
92. Yamaguchi E, Itoh A, Hizawa N, Kawakami Y. The gene polymorphism of tumor necrosis factor-beta, but not that of tumor necrosis factor-alpha, is associated with the prognosis of sarcoidosis. *Chest* 2001;119:753–761
93. Kreider ME, Christie JD, Thompson B, et al. Relationship of environmental exposures to the clinical phenotype of sarcoidosis. *Chest* 2005;128:207–215
94. Grutters JC, Sato H, Welsh KI, du Bois RM. The importance of sarcoidosis genotype to lung phenotype. *Am J Respir Cell Mol Biol* 2003;29:S59–S62
95. Yamada R, Yamamoto K. Recent findings on genes associated with inflammatory disease. *Mutat Res* 2005;573:136–151
96. Rybicki BA, Sinha R, Iyengar S, Gray-McGuire C, Elston RC, Iannuzzi MCSAGA Study Consortium. Genetic linkage analysis of sarcoidosis phenotypes: the Sarcoidosis Genetic Analysis (SAGA) Study. *Genet Immun.* In press
97. Hills SE, Parkes SA, Baker SB. Epidemiology of sarcoidosis in the Isle of Man, II: Evidence for space-time clustering. *Thorax* 1987;42:427–430
98. Edmondstone WM, Wilson AG. Temporal clustering of familial sarcoidosis in non-consanguineous relatives. *Br J Dis Chest* 1984;78:184–186
99. Visser H, Vos K, Zanelli E, et al. Sarcoid arthritis: clinical characteristics, diagnostic aspects, and risk factors. *Ann Rheum Dis* 2002;61:499–504
100. Maclure M. The case-crossover design: a method for studying transient effects on the risk of acute events. *Am J Epidemiol* 1991;133:144–153
101. Wacholder S, Hartge P, Struwing JP, et al. The kin-cohort study for estimating penetrance. *Am J Epidemiol* 1998;148:623–630
102. Saunders CL, Begg CB. Kin-cohort evaluation of relative risks of genetic variants. *Genet Epidemiol* 2003;24:220–229

103. Gail MH, Pee D, Carroll R. Kin-cohort designs for gene characterization. *J Natl Cancer Inst Monogr* 1999;55–60
104. Rybicki BA, Elston RC. The relationship between the sibling recurrence-risk ratio and genotype relative risk. *Am J Hum Genet* 2000;66:593–604
105. Umbach DM, Weinberg CR. Designing and analysing case-control studies to exploit independence of genotype and exposure. *Stat Med* 1997;16:1731–1743
106. Khoury MJ, Flanders WD. Nontraditional epidemiologic approaches in the analysis of gene–environment interaction: case-control studies with no controls!. *Am J Epidemiol* 1996;144:207–213
107. Albert PS, Ratnasinghe D, Tangrea J, Wacholder S. Limitations of the case-only design for identifying gene–environment interactions. *Am J Epidemiol* 2001;154:687–693
108. Gatto NM, Campbell UB, Rundle AG, Ahsan H. Further development of the case-only design for assessing gene–environment interaction: evaluation of and adjustment for bias. *Int J Epidemiol* 2004;33:1014–1024
109. Gundelfinger BF, Britten SA. Sarcoidosis in the United States Navy. *Am Rev Respir Dis* 1961;84(suppl):109–115
110. Sartwell PE, Edwards LB. Epidemiology of sarcoidosis in the US Navy. *Am J Epidemiol* 1974;99:250–257
111. Benatar SR. Sarcoidosis in South Africa: a comparative study in whites, blacks and coloureds. *S Afr Med J* 1977;52:602–606
112. Edmondstone WM, Wilson AG. Sarcoidosis in Caucasians, blacks and Asians in London. *Br J Dis Chest* 1985;79:27–36
113. Parra EJ, Marcini A, Akey J, et al. Estimating African-American admixture proportions by use of population-specific alleles. *Am J Hum Genet* 1998;63:1839–1851
114. McKeigue PM, Carpenter JR, Parra EJ, Shriver MD. Estimation of admixture and detection of linkage in admixed populations by a Bayesian approach: application to African-American populations. *Ann Hum Genet* 2000;64:171–186
115. Smith MW, O'Brien SJ. Mapping by admixture linkage disequilibrium: advances, limitations and guidelines. *Nat Rev Genet* 2005;6:623–632
116. Stephens JC, Briscoe D, O'Brien SJ. Mapping by admixture linkage disequilibrium in human populations: limits and guidelines. *Am J Hum Genet* 1994;55:809–824
117. Rybicki BA, Iyengar SK, Harris T, et al. The distribution of long range admixture linkage disequilibrium in an African-American population. *Hum Hered* 2002;53:187–196
118. Tian C, Hinds DA, Shigeta R, Kittles R, Ballinger DG, Seldin MFA. Genome-wide SNP panel with high ancestry information for African-American admixture mapping. *Am J Hum Genet* 2006;79:640–649
119. Smith MW, Patterson N, Lautenberger JA, et al. A high-density admixture map for disease gene discovery in African-Americans. *Am J Hum Genet* 2004;74:1001–1013
120. Collins FS, McKusick VA. Implications of the Human Genome Project for medical science. *JAMA* 2001;285:540–544