There is compelling evidence that sarcoidosis is the product of environmental triggers operating in an immunogenetically susceptible host that results in the formation of immune-mediated granulomas at disease sites. There is also support for the concept that immunogenetic predisposition determines the pattern of organ involvement in sarcoidosis.

Two approaches have been taken to determine the genetics of sarcoidosis: familial studies and case-control studies of sporadic disease. In familial studies, linkage or target gene association approaches are taken to the association of genotype with disease. In linkage studies, a series of linkage markers that span the whole genome are used to track marker pattern with disease. The area of possible association can then be mapped more finely with a second set of markers to target the specific area of interest.

### Familial sarcoidosis

The first indication that sarcoidosis may be of genetic predisposition came from the recognition of familial clustering of disease in several populations. The first record of disease in two German sisters appeared in 1923 [1]. Familial disease has also been confirmed in northern Sweden and central Hokkaido in Japan [2]. Pietinahlo et al. [3] identified familial sarcoidosis with a prevalence of 3.6% in Finnish and 4.3% in Japanese patients. An increased prevalence of sarcoidosis has also been described in the Furano district of northern Japan, with some evidence of familial clustering [4]. Table 1 shows the familial clustering of sarcoidosis in several populations worldwide.

In a study by Headings et al. [5] involving African-American patients a sibling disease prevalence of 10% was reported. Rybicki et al. [6] studied 488 parents and siblings of 179 African-American sarcoidosis cases and found a 2.5-fold increased risk of history of sarcoidosis in case relatives over that in the general population. In the UK, McGrath et al. [7] reported a sarcoidosis risk ratio for siblings of 36–73, indicating significant familial clustering of the disease. The population studied was mainly Caucasian (62.5%). The highest frequency of sarcoidosis among the immediate family of patients was found in an Irish population of 114 patients attending a specialist clinic; 9.6% of the patients were found to have one or more siblings with the same disease [8]. A Case-Control Etiologic Study of Sarcoidosis (ACCESS) is the largest prospective study that set out to investigate, among other things, the familial aggregation of sarcoidosis [9]. It is a large case-control study consisting of 10,862 first-degree and 17,047 second-degree relatives identified by 706 sarcoidosis case-control pairs. The familial relative risk (RR) to siblings
was larger than in parents (odds ratio (OR) = 5.8 versus 3.8). For African-Americans, the familial RR for a sarcoidosis case was 2.8, whereas the familial RR for Whites was markedly higher at 16.6. In summary, in the ACCESS study, cases were almost five times more likely than controls to report a sibling or a parent with a history of sarcoidosis.

Familial genetic studies

Sarcoidosis in all its clinical phenotypes is a genetically complex disease, i.e. the disease and, more specifically, subsets of disease are the product(s) of more than one gene interacting with one or more triggers to produce the end phenotype.

Few linkage studies have been performed in sarcoidosis, largely due to the difficulties in gathering large numbers of familial cases. Schürmann et al. [11] have completed two studies using microsatellite markers and multipoint nonparametric linkage (NPL) analysis. In the first study, they genotyped 122 affected siblings from 55 families for seven DNA polymorphisms in the major histocompatibility complex (MHC) region [11]. NPL analysis showed linkage for the entire MHC region, with the peak score identified at a marker locus in the Class III region (D6S1666). They subsequently suggested that the susceptibility region was nearer the class II MHC loci. A follow-up, genome-wide study using 225 microsatellite markers in 63 families with 138 affected siblings also found the most prominent peak at the MHC (p=0.001), including the D6S1666 marker identified in the first study [12]. However, six additional minor peaks (p<0.05) were found on chromosomes 1, 3 (close to the chemokine receptor genes CCR2 and CCR5), 7 (two peaks, one close to the T-cell receptor B gene), 9 (close to the gene for transforming growth factor-β receptor 1), and the X chromosome (close to the gene coding for the interleukin (IL)-2 receptor gamma chain). The authors of that study did warn that, with 225 markers, 11 similar minor peaks would be expected by chance alone. They also noted the lack of association with markers near other candidate genes, such as angiotensin-converting enzyme (ACE) and NRAMP1, but again warned that for genes with minor effects on susceptibility, a greater marker density and a much larger sample of affected sibling pairs would be required. More linkage studies in different ethnic groups, employing larger samples and more microsatellite markers, are required to allow identification of other candidate genes exerting minor effects on susceptibility.

More recently, Valentonyte et al. [13] utilised a three-stage single-nucleotide polymorphism (SNP) analysis approach to fine focus the region of interest on chromosome 6 in families and sporadic case-control samples of individuals with sarcoidosis. They identified a 15 kb disease-associated segment of the butyrophilin-like 2 (BTNL2) gene. BTNL2 is a member of the immunoglobulin superfamily and demonstrates homology to B7-1. It has thus been implicated as a costimulatory
molecule involved in T-cell activation. The G to A variant results in a truncated protein that is presumed to have aberrant function. Valetonty et al. [13] concluded that the BTN2 association is class II MHC independent, although there is almost complete linkage disequilibrium with HLA-DRB1; this requires further resolution.

**Association studies**

**Major histocompatibility complex**

**Human leukocyte antigen class I.** Some of the first studies of the genetics of sarcoidosis explored human leukocyte antigen (HLA) class I typing using serological techniques; a summary of these studies is presented in table 2.

The earliest such study in 1974 failed to detect any association between sarcoidosis in 132 German patients and HLA locus A [22]. Subsequent studies reported associations with HLA-B8 in Caucasian patients [14–16], but not in Black patients [23]. One of these studies in Czech patients also found an association with HLA-B13, and B8 B13 heterozygotes had a RR of 8.5 for sarcoidosis [16]. When these results were compared with 127 Italian patients, the association with HLA-B8 was confirmed, as well as an association with HLA-A1 [17]. HLA-B8 has been associated with disease of acute onset and short duration [18]. One study of 114 Japanese patients failed to find any association with class I antigens that could not be better explained by linkage disequilibrium with class II antigens, although they did find a reduction in HLA-B7 in patients compared with controls [19]. This is of interest because HLA-B7 has been found in increased frequency in African-American patients (a population with high disease prevalence) [20]. On the basis of these studies, it can be concluded that the association with HLA-B8 seems to be most commonly reported across racial boundaries, but it has not been found in all studies and other reported associations are even weaker. In any case, the immunology of sarcoidosis would suggest that HLA class II molecules are more likely to be involved in antigen presentation; any reported disease associations with class I genes may merely represent linkage with more relevant class II genes. However, in this regard a more recent study of Scandinavian patients found that both HLA-B*07 and HLA-B*08 independently increased the risk of sarcoidosis [21]. Other class I associations were secondary to their linkages to class II alleles. The common allele combination A*03,
B*07, DRB1*15 was most strongly associated with persistent disease and was found in 25.3% of such patients against 7.1% of healthy controls. The authors concluded that HLA class I alleles may have more influence on disease susceptibility and prognosis than has been recently thought.

**Human leukocyte antigen class II.** More contemporary studies have used molecular (usually SNP) methodology to determine associations with MHC class II alleles (table 3) [17, 19, 21, 24, 25–30, 31, 32, 33, 34]. The most compelling of these studies show associations that cross ethnic boundaries.

HLA-DR-associated antigens have attracted most attention recently. In Japanese patients, HLA-DR5, -DR6 and -DR8 [24, 35] and -DR8 and -DR9 [25] have been associated with disease. In Germans, HLA-DR5 has been linked with chronic disease [25], but HLA-DR3 with acute disease [26]. This pattern of differential linkage in acute and chronic disease has been confirmed in Scandinavia; certain alleles have been associated with chronic disease (HLA-DR14 and -DR15), while others have been associated with acute self-limiting disease (DR17 (3)) [27]. HLA-DR14 was also associated with chronic disease in a recent study of Asian Indians [28]. HLA-DR9 has been associated with disease risk in Japanese patients, but with disease protection in Scandinavians.

Some attempt has been made to arrive at a consensus from this confused picture. HLA-DR1 and -DR4 have been found to be protective in Scandinavians, Japanese, Italians, and in a study of UK, Polish and Czech patients [29]. This study showed position 11 of the HLA-DRB1 amino acid sequence to be the most variable, with three susceptibility alleles all coding for small hydrophilic amino acids, and two protective alleles coding for amino acids with bulky, hydrophobic, aliphatic side chains. As the position 11 amino acid is located at a crucial interface between the HLA-DR α and β chains, it is easy to imagine how these different alleles could affect HLA-DR heterodimer conformation and antigen binding.

A recent study reported high-resolution typing for the HLA-DPB1, HLA-DQB1, HLA-DRB1 and HLA-DRB3 loci in the first 474 ACCESS patients and case-matched controls [30]. A significant association was found between HLA-DRB1 alleles and disease in both Blacks and Whites. The HLA-DRB1*1101 allele was associated with sarcoidosis in both Blacks and Whites, with a population attributable risk of 16% in Blacks and 9% in Whites. The only class II allele differentially distributed between Blacks and Whites with respect to disease was HLA-DRB1*1501, being associated with controls in Blacks and with sarcoidosis in Whites. This caused Rossman et al. [30] to speculate that broadly similar HLA class II alleles may be associated with sarcoidosis in both populations, and that the increased risk in Blacks could be due to increased prevalence of susceptibility alleles in the Black population.

The HLA-DP locus has attracted considerable interest due to the finding that chronic beryllium disease (a chronic multisystem granulomatous disease with many immunopathological similarities to sarcoidosis) is associated strongly with the presence of a glutamic acid residue at position 69 of the HLA-DP beta chain. However, this association has not been found in two studies in sarcoidosis [36, 37], although the latter study in African Americans did find an increased risk of disease associated with Val36+ and Asp55+ alleles. In this regard, in the ACCESS study, the HLA-DPB1 locus also contributed towards susceptibility, but with the HLA-DPB1*0101 allele conveying most of the risk [30].

In another recent study in African Americans, 225 sarcoidosis patients with family members as controls were genotyped for six microsatellite markers covering the MHC region [38]. An association was observed between disease and the marker closest to
<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Sample size patients/controls</th>
<th>Susceptibility associations</th>
<th>Protection associations</th>
<th>Phenotype associations</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asian Indians</td>
<td>56/275</td>
<td>DRB1*11</td>
<td>DRB1*07</td>
<td>DRB1*14 severe disease</td>
<td>SHARMA [28]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DRB1*14</td>
<td>DQB1*0201</td>
<td></td>
<td></td>
</tr>
<tr>
<td>American Caucasians</td>
<td>268/268</td>
<td>DRB1*1101</td>
<td>DQB1*0602</td>
<td>DRB1*0401</td>
<td>ROSSMAN [30]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Eye hypercalcaemia</td>
<td></td>
</tr>
<tr>
<td>American Blacks</td>
<td>193/193</td>
<td>DRB1*1101</td>
<td></td>
<td>DRB1*0401 parotid/salivary glands</td>
<td>ROSSMAN [30]</td>
</tr>
<tr>
<td></td>
<td>225/479 first-degree relatives</td>
<td>DQB1*0602</td>
<td>DQB1*0201</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scadinavians</td>
<td>122/250</td>
<td>DR17(3)</td>
<td></td>
<td>DRB3 bone marrow</td>
<td>BERLIN [27]</td>
</tr>
<tr>
<td>Japanese</td>
<td>75/150</td>
<td>DRB1*11 (DR5)</td>
<td>DRB1*0101</td>
<td>DR17(3) acute disease</td>
<td>GRUNEWALD [21]</td>
</tr>
<tr>
<td></td>
<td>40/110</td>
<td>DRB1*14 (DR6)</td>
<td>DQB1*0501</td>
<td>DR14, DR15 chronic disease</td>
<td></td>
</tr>
<tr>
<td></td>
<td>26/247</td>
<td>DRB1*08 (DR8)</td>
<td>DPB*0402</td>
<td></td>
<td>ISHHARA [24]</td>
</tr>
<tr>
<td>Japanese</td>
<td>114/478</td>
<td>DRw5</td>
<td></td>
<td>DQB1*0601 cardiac</td>
<td>NARUSE [33]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DRw8</td>
<td></td>
<td>DR5 mild disease</td>
<td>INA [19]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DRw52</td>
<td></td>
<td>DR8 chronic disease</td>
<td></td>
</tr>
<tr>
<td>Ethnicity</td>
<td>Sample size patients/controls</td>
<td>Susceptibility associations</td>
<td>Protection associations</td>
<td>Phenotype associations</td>
<td>Reference</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-------------------------------</td>
<td>----------------------------</td>
<td>-------------------------</td>
<td>------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>UK and Dutch Caucasians</td>
<td>235/568</td>
<td></td>
<td>DQB1<em>0201 mild disease, including Löfgren's DQB1</em>0602 severe disease</td>
<td></td>
<td>SATO [32]</td>
</tr>
<tr>
<td>Dutch Caucasians</td>
<td>37 Löfgren's/88</td>
<td>DQB1<em>0201-DRB1</em>03(01) and Löfgren's</td>
<td></td>
<td>DQB1<em>0201-DRB1</em>03(01) Löfgren's</td>
<td>ROSMAN [30]</td>
</tr>
<tr>
<td>Czech and Italian Caucasians</td>
<td>233/1010</td>
<td>DR3</td>
<td>DR4</td>
<td>DR3 mild disease</td>
<td>MARTINETTI [17]</td>
</tr>
<tr>
<td>German</td>
<td>78/50</td>
<td></td>
<td></td>
<td>DR4 severe disease</td>
<td>SWIDR [26]</td>
</tr>
<tr>
<td>UK Caucasians</td>
<td>189/288</td>
<td>DRB1*12</td>
<td>DRB1*01</td>
<td>DR5 chronic disease</td>
<td>NOWACK [25]</td>
</tr>
<tr>
<td>Polish</td>
<td>87/133</td>
<td>DRB1*14</td>
<td>DRB1*04</td>
<td></td>
<td>FOLEY [29]</td>
</tr>
<tr>
<td>Czech</td>
<td>69/158</td>
<td>DRB1*15</td>
<td>DRB1*01</td>
<td></td>
<td>FOLEY [29]</td>
</tr>
<tr>
<td>Dutch Caucasian</td>
<td>149/418</td>
<td>DRB1<em>150101/DQB1</em>0602</td>
<td></td>
<td>Severe disease</td>
<td>VOORTER [34]</td>
</tr>
</tbody>
</table>
HLA-DQB1; further analysis of other markers at this locus confirmed that in this African-American cohort, it is HLA-DQB1 and not HLA-DRB1 that confers susceptibility to sarcoidosis. A follow-up study found that HLA-DQB1*0201 was transmitted to affected offspring only half as often as expected, whereas DQB1*0602 was transmitted to affected offspring ~20% more than expected, and was associated with radiographic disease progression [31]. This contrasts with the study on ACCESS patients, in which the major association was with HLA-DRB1; however, that paper did also report an association of disease in blacks with HLA-DQB1*0502, albeit a weaker association than with HLA-DRB1 [30]. In a recent study in a population of 149 Dutch Caucasian sarcoidosis patients, an association was found between severe disease and the haplotype DRB1*150101/DQB1*0602; due to the tight linkage disequilibrium of the alleles in the Caucasian population, it is not possible to assign the primary association [39].

These DQB1 associations are consistent with a European study of the HLA-DQB1 locus in 235 Dutch and UK patients. DQB1*0201 was found to be strongly protective against severe sarcoidosis (i.e. it was associated with stage I disease) whereas the DQB1*0602 allele tended to have the opposite effect [32]. Furthermore, the patient group included 15 patients with Löfgren’s syndrome and 23 patients with erythema nodosum (EN); carriage of the DQB1*0201 allele was greatly increased in these subgroups. In a follow-up study based on this finding and the previously reported associations between Löfgren’s and the DRB1*0301 allele and the A allele at position -307 of the tumour necrosis factor (TNF)-α gene (TNF2) [26], 37 Dutch Löfgren’s patients were genotyped for all three alleles [40]. The associations were confirmed, being strongest for DQB1*0201 (OR=8.6 versus 6.7 and 6.8 for DRB1*03 and TNF2, respectively). In addition, there was 100% linkage disequilibrium between DQB1*0201 and DRB1*03, and 70% LD between these alleles and TNF2. This article, therefore, identified the DQB1*0201-DRB1*03(01)-TNF2 haplotype as an extended MHC haplotype that was present in 76% of Löfgren’s patients (versus 24% of controls) and conferred the risk for Löfgren’s syndrome in Dutch Caucasians (OR=9.9).

Löfgren’s syndrome is extremely rare in Japanese patients; this fits well with the above study, as the Japanese do not have the HLA-DR3 allele. However, cardiac sarcoidosis is more common than in Caucasians. TAKASHIGE et al. [41] reported an association between cardiac sarcoidosis and the TNFA2 allele. Further studies by NARUSE et al. [33] concluded that the strongest association was in fact with the HLA-DQB1*0601 allele, with homozygosity for this allele conferring a RR of 29.5 compared with healthy controls. The TNFA2 allele was not in linkage disequilibrium with the class II allele, so the authors of that study suggested that this allele might confer an additional genetic risk for cardiac sarcoidosis. Small fibre neuropathy has recently been identified as a relatively common problem in a cohort of Dutch chronic sarcoidosis patients. Sarcoidosis patients suffering from small fibre neuropathy showed an increased prevalence of DQB1*0602, suggesting a possible genetic relationship [34].

These studies all demonstrate one of the most taxing issues, that of primary association or association through linkage disequilibrium. Tight phenotyping of different ethnic groups with different linkage profiles, in concert with extended haplotyping and sequence data across the whole MHC region, will assist clarification of this issue.

Infection susceptibility genes

The histopathological and immunological similarities between sarcoidosis and infectious granulomatous diseases, such as tuberculosis, have led to several studies attempting to associate sarcoidosis with allelic variations in genes necessary for innate immunity (table 4).
<table>
<thead>
<tr>
<th>Gene</th>
<th>Population</th>
<th>Sample size patients/controls</th>
<th>Susceptibility associations</th>
<th>Protection associations</th>
<th>Phenotype associations</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNF-α</td>
<td>UK and Dutch</td>
<td>101/216</td>
<td>-857T</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>UK and Dutch</td>
<td>196/576</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>228 (46 Lofgren’s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNF-β</td>
<td>Japanese</td>
<td>110/161</td>
<td>-889C/C homozygotes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL-17</td>
<td>Czech</td>
<td>95/199</td>
<td>-826T</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL-18</td>
<td>Japanese</td>
<td>119/130</td>
<td>-607C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIF</td>
<td>Spanish</td>
<td>28 with EN/122</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL-18</td>
<td>UK and Dutch</td>
<td>205/201</td>
<td>-297T</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCR2</td>
<td>Japanese</td>
<td>100/122</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dutch</td>
<td>137 (47 Lofgren’s)/167</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCR5</td>
<td>Czech</td>
<td>66/386</td>
<td>CCR5Δ32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RANTES</td>
<td>Japanese</td>
<td>114/136</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VDR</td>
<td>Japanese</td>
<td>157/111</td>
<td>B allele intron 8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NRAMP</td>
<td>African Americans</td>
<td>157/111</td>
<td>5′(CA)n 120bp allele</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAP2</td>
<td>UK</td>
<td>117/290</td>
<td>565T</td>
<td>665A</td>
<td>379V</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Polish</td>
<td>87/158</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CR1</td>
<td>Italian</td>
<td>91/84</td>
<td>5507G/G homozygotes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACE</td>
<td>African Americans</td>
<td>183/111</td>
<td>Deletion polymorphism</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>intron 16</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TNF: Tumour necrosis factor; IL: interleukin; MIF: macrophage migration inhibitory factor; IL-18: inhibitor protein kappa B; CCR: CC chemokine receptor; RANTES: regulated on activation, normal T-cell expressed and secreted; VDR: vitamin D receptor; NRAMP: natural resistance-assisted macrophage protein; TAP: transporter associated with antigen processing; CR: complement receptor; ACE: angiotensin-converting enzyme; EN: erythema nodosum. #: not replicated in subsequent studies.
**Mannose binding lectin.** Mannose binding lectin (MBL) is a serum protein that binds oligosaccharides on the surface of pathogens, and triggers complement activation. Foley et al. [59] studied the MBL gene promoter and exon 1 variants known to decrease serum levels of MBL and increase susceptibility to infection. No associations were found in 167 British patients and 164 controls between these MBL gene variants and susceptibility to sarcoidosis, age of disease onset, or severity of disease.

**Vitamin D receptor.** 1,25-dihydroxycholecalciferal (1,25-DHCC) binds the nuclear vitamin D receptor (VDR). As well as its role in calcium homeostasis, it has an immunoregulatory effect, and polymorphisms within the VDR gene on chromosome 12 have been linked to susceptibility to infections, such as tuberculosis and leprosy [60]. 1,25-DHCC is known to inhibit Mycobacterium tuberculosis growth in human macrophages and monocytes. Pulmonary macrophages from sarcoidosis patients show an increased capacity to synthesise 1,25-DHCC, and it then acts on other macrophages, inducing intracellular adhesion molecule (ICAM)-1 expression, and reducing TNF-α release [61]. It is known that 1,25-DHCC is produced at sites of granulomatous inflammation in sarcoidosis and that this is responsible for the hypercalcaemia seen in 10% of patients. A biallelic polymorphism in intron 8 accounts for three genotypes, bb, bB, and BB. The more common bb genotype is associated with reduced VDR mRNA expression. One Japanese study has found an association between susceptibility to sarcoidosis, and the presence of the bB genotype and B allele [54]. However, this study found no correlation of genotype with severity or spread of the disease.

**Natural resistance-associated macrophage protein**

The gene encoding for natural resistance-assisted macrophage protein (NRAMP) was first identified in a murine model resistant to infection by the intracellular parasites leishmania, salmonella and mycobacteria. A single nonconservative amino acid substitution of aspartate for glycine at position 169 (within the second predicted transmembrane domain) confers susceptibility to infection, and Nramp1 gene-disrupted mice are susceptible to all three pathogens [62–64]. Resistance is restored in transgenic mice when the Nraomp1 G169 allele is transferred, proving that Nramp1 is vital in mice to protect against mycobacteria and other intracellular pathogens [65].

NRAMP1 (solute carrier family 11 (proton-coupled divalent metal ion transporters), member 1 (SLC11A1)) is the human equivalent of this gene. It has been mapped to human chromosome 2q35 [66, 67]. The function of the NRAMP1 gene product remains uncertain, but the gene is expressed only in reticuloendothelial cells. The NRAMP family is conserved from prokaryotes to eukaryotes. It has been suggested that it has a role in controlling the phagolysosomal environment, possibly by regulating the concentration of divalent cations, such as iron or manganese.

Bellamy et al. [68] typed NRAMP1 polymorphisms in 410 patients with smear-positive pulmonary tuberculosis and 417 controls in West Africa. They found associations between smear-positive tuberculosis and heterozygosity for two NRAMP1 polymorphisms in intron 4 and the 3' untranslated region.

The putative role of NRAMP1 in macrophage activation, and the association of genetic polymorphisms with susceptibility to tuberculosis in humans, makes it an attractive candidate gene in sarcoidosis. Maliarik et al. [55] analysed several NRAMP1 gene polymorphisms in 157 African Americans with sarcoidosis and 111 ethnically matched controls. They identified a statistically significant difference in allele frequency distribution of 5'(CA)n alleles between patients and controls. The commonest 120-bp allele was over-represented in the sarcoidosis patients, suggesting that the polymorphism
at this site is protective against sarcoidosis. Interestingly, the two variants that were associated with susceptibility to tuberculosis did not affect susceptibility to sarcoidosis in this study. Its role in sarcoidosis susceptibility is uncertain.

**Transporter associated with antigen processing.** Transporter associated with antigen processing (TAP)1 and TAP2 genes, located in the Class II MHC region, encode subunits of a heterodimeric complex that function in the endogenous antigen-processing pathway. TAP is inserted into the membrane of the endoplasmic reticulum, and facilitates the transport of proteosome-generated antigenic peptide fragments from the cytosol into the lumen of the endoplasmic reticulum, for subsequent presentation of the peptides in association with MHC class I molecules [69]. One study examined two SNPs in TAP1 and three SNPs in TAP2, in 117 UK Caucasions with 290 UK controls and in 87 Polish-Slavonic patients with 158 Polish controls [56]. They found different associations with different TAP2 variants in the two ethnically diverse populations studied. In the UK population, a threonine variant at position 565 and an alanine variant at position 665 were found to be significantly under-represented in the patient groups. By contrast, they failed to find these differences in the Polish population; in this group the valine variant at position 379 was under-represented. The location of these genes in the MHC class II region suggests that these associations could be due to linkage disequilibrium, but significant linkage disequilibrium was not found either between the TAP1 and TAP2 loci or between TAP variants and HLA-DPB1 variants. However, in Japanese patients no associations were found between TAP1 or TAP2 polymorphisms and disease susceptibility [70].

**Complement receptor 1.** Complement receptor (CR)1 is a membrane protein expressed on erythrocytes, phagocytes, lymphocytes and dendritic cells that plays a number of roles in the complement system. CR1 on erythrocytes mediates the transport of immune complexes through the bloodstream to phagocytes in the liver and spleen. As sarcoidosis is accompanied by a polyclonal hypergammaglobulinaemia, it could be hypothesised that interaction with the sarcoiold antigen could lead to immune complex formation that may be involved in the pathogenesis of the disease and immune complexes have been shown to be present in this disease. The CR1 gene on chromosome 17 is known to contain a number of SNPs that can be correlated with high or low expression of CR1 on erythrocytes. A study of 91 Italian sarcoidosis patients compared with ethnically matched controls demonstrated a significant association between the GG genotype at the C5507G polymorphism and sarcoidosis [57]. This association was even stronger when the female subgroup was analysed separately. This allele is known to be associated with low CR1 expression on erythrocytes, and Zorzetto et al. [57] postulated that reduced clearance of immune complexes could lead to deposition outside the reticuloendothelial system with the consequent granulomatous inflammation characteristic of sarcoidosis. However, these findings have not been found in other studies (the present author’s unpublished observations).

**Angiotensin-converting enzyme**

Concentrations of serum angiotensin-converting enzyme (ACE) are above normal in only ~60% of patients with sarcoidosis. A 287-bp insertion/deletion (I/D) polymorphism in intron 16 of the ACE gene has been reported to account for 47% of the phenotypic variation in serum ACE levels; the DD genotype is associated with the highest serum levels in patients and controls and the II genotype with the lowest [71]. However, no consistent correlation between ACE I/D polymorphisms and disease prevalence or severity have been demonstrated. In the most recent study, McGrath et al. [72] summarise the previous studies in this area. They also genotyped 118 UK and 56 Czech
patients with sarcoidosis, and attempted to correlate ACE I/D polymorphisms with
disease susceptibility and pulmonary disease severity or progression at 2 and 4 yrs from
presentation. They found no such associations and concluded that this ACE gene
polymorphism has no effect on disease susceptibility or progression in European Whites.

By contrast, in an African-American population, an excess of D alleles was found in
183 cases compared with 111 controls, and a gene dosage effect was seen causing an
increased risk of sarcoidosis in DD homozygotes that was higher than the increased risk
in ID heterozygotes [58]. ACE genotype did not appear to influence disease severity or
progression. This same study showed no difference in allele distribution between 60
Caucasian cases and 48 controls. In summary, there is little strong evidence to relate this
particular polymorphism with either disease presence or severity.

Cytokines

Cytokines are appealing candidate genes, alleles that may affect cytokine levels, may
modify disease pathogenesis and, therefore, susceptibility to disease, and/or the clinical
course [73]. Many case-control studies have picked a candidate gene and speculatively
genotyped for allelic variations; table 4 details the positive reports to date.

Tumour necrosis factor. The TNF gene complex is located within the MHC complex,
and includes the TNF-α and TNF-β (lymphotoxin α) genes. TNF is an early cytokine in
the pathogenetic cascade in sarcoidosis, and the many studies linking sarcoidosis to the
MHC complex have aroused much interest in the genes of the TNF gene complex as
further candidate susceptibility genes.

Early studies explored two biallelic markers in the TNF-α gene (G to A at position -308,
known as TNFA) and the TNF-β gene (in the first intron, known as TNFB). Both have
been shown to affect cytokine production and the former has been linked to granulomatous
disease caused by beryllium [74]. The less common TNFA2 allele (-308A), associated with
high TNF-α production, was first linked with Löfgren’s syndrome in a study of 101 patients
with pulmonary sarcoidosis (including 16 Löfgren’s patients) compared with 216 healthy
controls [42]. This same study found no association of either polymorphism with the
sarcoidosis group as a whole. A survey of 110 Japanese patients (in whom Löfgren’s
syndrome is extremely rare) and 161 controls also found no associations between either of
these polymorphisms and susceptibility to sarcoidosis as a whole [45]. However, this study
did find the TNFB1 allele to be correlated with a prolonged clinical course, with carriage of
this allele significantly prolonging the time to spontaneous disease resolution. Prolonged
disease must be distinguished from severe disease; this association was present in patients
who did not require corticosteroid therapy for severe symptoms or organ-threatening
disease despite a prolonged disease course. Yamaguchi et al. [45] noted that TNFB1 has
been associated with increased production of both TNF-β and TNF-α, so this
polymorphism could have a functional role in sarcoidosis.

The association between Löfgren’s syndrome and TNFA2 has been confirmed in a
more recent study of 196 British and Dutch patients and 576 controls [43]. This study
looked at five bi-allelic TNF promoter polymorphisms, including the G to A allele at
position -308 (now rerenumbered as position -307). This less common TNFA2 allele was
not associated with sarcoidosis susceptibility overall, but was significantly associated
with the Löfgren’s subgroup (n=15). However, a highly significant increase in the TNF-α
-857T allele was found in the sarcoidosis patient group overall. Of the sarcoid patients,
25.5% carried the TNF -857T allele compared with 14.1% of controls; allele frequency
was 13.5% in cases versus 7.3% in controls. This allele has been associated with higher
transcriptional promoter activity.
Interestingly, this allele was not associated with Löfgren’s patients; in fact there was a trend towards it being under-represented in this subgroup. In a follow-up study of 228 sarcoidosis patients including 46 Löfgren’s, six haplotypes were constructed across the five polymorphisms studied in the TNF-α promoter [44]. The -307A allele associated with Löfgren’s syndrome occurred exclusively in haplotype 2. The -857T allele associated with sarcoidosis susceptibility, but under-represented in the Löfgren’s group, occurred exclusively in haplotype 4. There was a very strong association between haplotype 2 and stage I disease with EN, including Löfgren’s patients \( (p=6.6 \times 10^{-11}) \) and with favourable radiological evolution over 4 yrs. By contrast, haplotype 4 was strongly associated with stages II–IV disease \( (p_c<0.01) \) and with unfavourable radiological evolution over 4 yrs.

An association between cardiac sarcoidosis and the TNFA2 allele in Japanese was found to be due to linkage disequilibrium with the HLA-DQB1*0601 allele [33, 41]. It remains unclear whether TNF associations are functional polymorphisms, or whether one or both are in linkage disequilibrium with another site elsewhere in the MHC region.

**Interleukin-1.** RYBICKI et al. [75] studied six polymorphic markers that are closely linked to certain candidate cytokine genes in African-Americans and identified two alleles that may be associated with sarcoidosis. If both the IL-1α*137 and F13A*188 alleles were present, there was a six-fold increased risk of sarcoidosis, rising to a 15-fold increase in patients with a family history of sarcoidosis. The F13A marker is close to the gene for IRF-4 (a member of the interferon (IFN) regulatory factor family, a transcription factor). The importance of IL-1 in granulomatous inflammation is well known; less is known about the role of IRF-4.

A Czech case-control study investigated polymorphisms within the IL-1 gene cluster on chromosome 2q13–21 [46]. They studied bi-allelic polymorphisms in the IL-1α and IL-1β genes, and an 86-bp variable number tandem repeat polymorphism in intron 2 of the IL-1 receptor antagonist (IL-1RN). They found an increase in the IL-1α -889 1.1 genotype (C/C homozygotes) in 95 patients with sarcoidosis compared with 199 controls \( (60.0\% \text{ versus } 44.2\%, p=0.012) \). There was a corresponding fall in the number of IL-1α -889 1.2 (C/T) heterozygotes \( (32.6\% \text{ versus } 47.7\%, p=0.015) \). The RR of disease for IL-1α -889 1.1 homozygotes compared with 1.2 heterozygotes or 2.2 homozygotes was 1.9. This finding was consistent with the study by RYBICKI et al. [75] in African-Americans. However, a study of this same polymorphism in 147 UK patients with 101 UK controls and 102 Dutch patients with 166 controls failed to reproduce this association in either population [76]. This led RYBICKI et al. [75] to surmise that the previously reported associations may have been due to linkage disequilibrium within the IL-1 gene cluster between the IL-1α -889C allele and the unidentified locus that actually confers the risk of sarcoidosis. Population-dependent differences in linkage could explain the failure to reproduce this finding in UK and Dutch populations.

**Interleukin-18.** IL-18 is known to act synergistically with IL-12 to induce IFN-γ production from T-helper type 1 cells. It is constitutively expressed in healthy airway epithelium and this expression is increased in sarcoidosis tissue. Serum and bronchoalveolar lavage (BAL) levels of IL-18 are high in sarcoidosis patients. Therefore, it is an attractive candidate.

One Japanese study genotyped 119 patients with sarcoidosis and 130 controls for two polymorphisms in the gene promoter known to affect promoter activity after stimulation [47]. They found a significant increase in the percentage of patients carrying the C allele at position -607. This position may be a binding site for the cAMP-responsive element binding protein (CREB), and it is postulated that the A allele at this position may
interfere with CREB binding, leading to a reduction in responsiveness with subsequent reduced IL-18 expression.

**Macrophage migration inhibitory factor.** Macrophage migration inhibitory factor is known to stimulate T-cells and drive the delayed type hypersensitivity reaction; polymorphisms in this gene have been investigated in a cohort of Spanish patients with EN [48]. A SNP at position -173 (G to C) was genotyped in 28 patients with sarcoidosis-associated EN, 70 patients with EN of other aetiologies, and 122 healthy controls. The C mutant allele was significantly over-represented in the sarcoidosis EN group compared with the nonsarcoid EN group and with controls. Patients with EN carrying this allele were at increased risk of having sarcoidosis compared with EN patients without this allele. This G to C transition in the 5’-flanking region of the gene creates an AP-4 binding site and may, therefore, be of functional significance in controlling transcription.

**Inhibitor kappa B-alpha.** Nuclear factor (NF)-κB transcription is one of several factors that influence cytokine gene expression. Activation of this signalling pathway occurs in several of the components of the sarcoidosis pathogenetic process. In the resting cell, NF-κB is retained in the cell cytoplasm by the bound inhibitor (I)κB. This inhibitor protein degrades upon phosphorylation, unmasking the nuclear localisation sequence of NFκ-B, allowing nuclear localisation and initiation of transcription. 1κB-α, a 1κB protein that is essential for normal termination of the NF-κB response, has three single nucleotide polymorphisms in the promoter region of its gene on chromosome 14. A total of 205 UK and Dutch sarcoidosis patients were genotyped for these polymorphisms and compared with 201 controls [49]. The T allele at position -297 was strongly associated with disease susceptibility (p=0.008). By contrast, this allele was not associated with disease severity as measured by radiographic staging and pulmonary function test indices. However, the frequency of the T allele at position -826 progressively decreased across the chest radiographic stages of sarcoidosis II–IV, suggesting that it could be protective against fibrotic disease. The functional effects of these polymorphisms are not known.

**Chemokines**

A number of chemokines (and their receptors) play a role in the pathogenesis of sarcoidosis. These include regulated on activation, normal T-lymphocyte expressed and secreted (RANTES), monocyte chemoattractant protein (MCP)-1 and macrophage inflammatory protein (MIP)-1α; all are known to be released by alveolar macrophages and increased levels may correlate with disease severity (table 4).

**C-C chemokine receptor 2.** C-C chemokine receptor (CCR) 2 is a receptor for the MCP family. This receptor also acts as a co-receptor for HIV infection. A SNP in the CCR2 gene at position 64 substitutes isoleucine for valine (V64I). This polymorphism has a protective effect against progression of HIV infection to AIDS. It was first studied in sarcoidosis in a Japanese population, and the V64I allele was found to be associated with a lower risk of sarcoidosis in 100 patients compared with 122 controls (p<0.0017) [50]. This same polymorphism was studied in 65 Czech patients and 80 controls [52]. The allelic frequency of the polymorphism was again reduced in patients compared with controls, but did not reach statistical significance (6.9% versus 11.9%; p=0.17).

Spagnolo et al. [51] investigated the CCR2 gene in more detail, genotyping 304 Dutch individuals (90 non-Löfgren’s sarcoid, 47 Löfgren’s syndrome, 167 controls) for eight SNPs (including V64I) across the whole gene. When analysing all sarcoidosis patients together, no statistically significant differences in any of the allele frequencies were
found. However, analysing the Löfgren’s syndrome patients separately, strongly significant associations were found with five of these alleles when compared both with the healthy controls and also when compared with the non-Löfgren’s sarcoidosis patients. The V64I polymorphism was not one of the alleles showing association with Löfgren’s syndrome. They were able to construct nine haplotypes from the eight polymorphisms. In patients with Löfgren’s syndrome, a strongly significant increase in the frequency of CCR2-haplotype 2, which includes four unique alleles (A at nucleotide position -6752, A at 3,000, T at 3,547, and T at 4,385), was observed compared with control subjects (74% versus 38%, respectively; p < 0.0001), whereas no difference was found between non-Löfgren’s sarcoidosis and control subjects (both 38%). The association remained strong after correcting for two other known associations with Löfgren’s syndrome, female sex and carriage of HLA haplotype DRB1*0301-DQB1*0201. It remains unclear how this association should be interpreted; it is not yet known which of the polymorphisms has a functional effect that could explain the association, or whether the functional effect is due to an as yet unidentified polymorphism in the CCR2 gene or a neighbouring gene that is in linkage disequilibrium with this CCR2 gene haplotype. It also remains unclear whether this CCR2 haplotype predisposes to Löfgren’s in all carriers, or only in those positive for HLA DRB1*0301-DQB1*0201. In Löfgren’s patients negative for HLA DRB1*0301-DQB1*0201, the CCR2 haplotype did not appear to be more common but the number of subjects in this subgroup was too small for definitive statistical conclusions to be drawn.

C-C chemokine receptor 5. CCR5 acts as a receptor for the chemokines RANTES and MIP-1α. A 32-bp deletion in the gene (CCR5Δ32) renders the surface receptor molecule nonfunctional. In a study of 66 Czech sarcoidosis patients and 386 controls, this allele was significantly more common in patients [52]. Furthermore, this allele was associated with clinically more apparent disease; it was present in 39.1% of patients requiring corticosteroids but in only 16.7% of patients who needed no treatment. This study implicates the CCR5 gene as affecting both disease susceptibility and severity.

RANTES (C-C chemokine ligand 5). RANTES, a member of the C-C chemokine family, is produced at sites of granulomatous reaction and is a potent chemoattractant for lymphocytes, monocytes and eosinophils. A Japanese study genotyped 114 sarcoidosis patients and 136 healthy controls for an A/G SNP at position -403 in the promoter of the RANTES gene [53]. Although finding no association between alleles and disease, they did find the less common AA genotype was associated with disease in three or more organs. They also found that this genotype was associated with a higher CD4+/CD8+ T-cell ratio in BAL cells than the GA or GG genotypes. They postulate that homozygosity for the A allele may be a risk factor for more extensive disease in Japanese patients.

Conclusion

In summary, the strongest associations between genotype and phenotype in sarcoidosis are found in the MHC complex region of chromosome 6. Other candidate loci are less appealing and most associations have either not been reproducible or validated functionally. The future will demand very precise phenotyping as part of a trans-ethnic approach to this disease that will allow subtle differences in linkage in different ethnic cohorts to be determined, which will help tease out prime associations from those that are secondary. Familial and case-control studies will provide important complementary approaches.
Summary

Although sarcoidosis is described in almost all populations around the world, there are wide variations in the incidence and prevalence of the different clinical phenotypes, suggesting genetic influence on disease susceptibility and phenotype. It has emerged from multiple family and case-control studies that sarcoidosis does indeed have a significant genetic susceptibility. This review sets out to outline the genetic studies to date and to highlight key genetic targets and the importance of precise phenotyping. Despite the study of many logical (and occasionally illogical) loci, the overwhelming consensus is that the Class II major histocompatibility complex (MHC) region of chromosome 6 is the major association hot spot. In this regard, Löfgren’s syndrome has been strongly associated with an extended MHC haplotype encompassing alleles in human leukocyte antigen (HLA)-DQ, HLA-DR and tumour necrosis factor, suggesting that other clinical subtypes and phenotypes might be associated with different susceptibility genes. In more global sarcoidosis, there is a consistency evolving from studies around the world that there is a small set of susceptibility HLA-DR alleles and other alleles that confer protection, possibly in association with an independent effect from the butyrophilin-like 2 gene.

Other case-control studies using the candidate gene approach have reported some, generally less convincing, associations with genes encoding for cytokines, chemokines and infection susceptibility genes. Of these, the chemokine receptor associations are most promising.

The development of tighter definitions of clinical phenotypes and the emergence of larger cohort studies of sarcoidosis, including familial sarcoidosis, should combine with new technological advances to build a clearer picture of the genetics of sarcoidosis susceptibility and phenotypes in the future.

Keywords: Familial, genetics, human leukocyte antigen, polymorphisms, sarcoidosis.

References


36. Foley PJ, Lympany P, Puscasinska E, et al. HLA-DPB1 and TAP1 polymorphisms in sarcoidosis. *Chest* 1997; 111: Suppl. 6, 73S.


